Department for Transport

NATIONAL TRANSPORT MODEL VERSION 5 PEER REVIEW

June 2020

Quality Control

Issue/revision	First issue	Revision 1	Revision 2	Revision 3
Remarks		Updated post DfT comments	Updated with final comments from DfT and first draft of ES	Minor revision to ES
Date	11th May 2020	1st June 2020	26th June 2020	29th June 2020
Prepared by	Andrew Stoneman John Bates Ian Williams	Andrew Stoneman John Bates Ian Williams	Andrew Stoneman John Bates Ian Williams	Andrew Stoneman John Bates Ian Williams
Checked by	Chris Sanders	Chris Sanders	Chris Sanders	Chris Sanders
Authorised by	Andrew Stoneman	Andrew Stoneman	Andrew Stoneman	Andrew Stoneman
Project number	70006059	70006059	70006059	70006059
Report number	1	2	3	3.1

File reference	\\uk.wspgroup.com\central data\Projects\70006xxx\70006059 - T-TEAR NTM Futures\E Models and Drawings\05 Documentation

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

Overview

This Peer Review of the recently completed National Transport Model (NTM) was commissioned by the Department for Transport (DfT) to advise on whether the completed model had met its quality and capability requirements. Referred to as NTMv5, the model was commissioned in 2016 and other than retaining a link with the Department's National Trip End Model (NTEM) to provide estimates of exogenous growth over time, in all other respects this is a completely new model.

The Peer Review was conducted collaboratively by the review team of Andrew Stoneman (WSP), John Bates (John Bates Services) and Ian Williams (Ian Williams Services) between January 2020 and June 2020.

The NTMv5 has the general structure of a conventional transport model, in that it contains modules representing (highway) assignment and multi-modal demand responses in terms of mode and destination choice in its Variable Demand Model (VDM). In addition, the assignment "pivots" off a set of base highway matrices which have been specially developed. Relative to a conventional transport model, the challenge is to deal with the size of the study area, which here concentrates on England.

In relation to the key components of the highway assignment model and the demand model (VDM), we have a number of detailed comments, but overall we consider these modules to have been competently and carefully constructed and generally fit for purpose. We are also generally satisfied with the "external" components relating to other vehicles and other mode costs. We have much more concern about the pivoting process and the construction of the base demand as well as the impact of the apparent lack of model convergence on the results of the reported sensitivity tests.

The Use Cases for NTMv5

The model was required to deal with the following "Use Cases" which were defined by the DfT:

- UC1 Strategic Roads Investment and Resilience: To analyse the impacts of packages of roads schemes at a national level;
- UC2 Road User Charging and other potential policy interventions: To adapt to road policies in future Parliaments such as pricing, on strategic roads or urban roads, or parking policy in urban areas, or other behavioural devices.
- UC3 Local Investment and Policy: A variety of analysis including national impacts of congestion relief schemes; policy impacts of introducing public transport improvements (e.g. light rail). This could potentially include travel demand management.
- UC4 General Support for DfT Teams (other than Roads / Local): Environmental analysis of transport policies relating to carbon and/or an approximation of air quality emissions, and aviation surface access.

- UC5 Scenario-based National Traffic Forecasting: Understanding of changes in population or travel trends (e.g. driving rates amongst young people), with scenarios around GDP, car ownership, fuel price, road tax.
- UC6 Exploring the unknown: Testing new policies or technical developments that have not been modelled before (e.g. Connected and Autonomous Vehicles) and others of whose existence we are not currently aware.

The DfT has indicated that use case 5 (UC5, National Traffic Forecasting) should be considered the primary Use Case as it is an existing use and the most fundamental purpose of the National Transport Model.

The Peer Review Structure

The documents which the Peer Review team considered were:

- NTMv5 Developer Guide version 2.0, in 6 volumes:
  - Volume 1 Model Zones and Highway Network (referred to as Volume 1 in this review)
  - Volume 2 Base Year Demand (Volume 2)
  - Volume 3 Highway Assignment Model (Volume 3)
  - Volume 4 Non-car Modes (Volume 4)
  - Volume 5 Variable Demand Model (Volume 5)
  - Volume 6 Model Testing (Volume 6)
- NTMv5 Future Development: Quality Report version 4.0 (the Quality Report)
- NTMv5 User Guidance: Installation and User Guidance version 4.0 (the User Guide)
- National Transport Model for England: Estimation of the mode-destination models version 13.

This Peer Review document is divided into the following chapters:

- Chapter 1 provides the background to the NTMv5 through a general description of the model and introduces the use cases.
- Chapter 2 is where we assess the construction of the model in detail, considering all its constituent components.
- Chapter 3 deals with tests carried out on the model from its base year position. This covers the validation of the highway model, the realism tests, and Sensitivity Tests 2 to 5.
- Chapter 4 moves on to the additional issues associated with future year forecasts where the main discussion relates to Sensitivity Test 1.
- Chapter 5 discusses the other documentation submitted to us for review; the Quality Report and the User Guide. Our assessment of the Quality Report essentially concerns its structure and content, rather than any discussion of the results *per se*.
- Chapter 6 assesses the achievements of the model in respect of the Use Cases, with a view to drawing conclusions about how far it can be used with confidence.
- Chapter 7 considers how we took the opportunity to run the model ourselves, and for this purpose we have carried out a number of tests, both to replicate existing results, and to understand the usefulness of the model and the plausibility of the results for some independently devised strategies.
- Chapter 8 draws out our conclusions and recommendations.

Main Observations of the Peer Review

Given that the predominant focus of NTMv5 is on highway travel, we consider the general structure of a conventional transport model, containing modules representing (highway) assignment and multi-modal demand responses in terms of mode and destination choice, to be sensible, and the decision to align external changes in demand associated with land-use etc with the DfT's National Trip End Model (NTEM) to be correct.

In relation to the highway assignment model our key observations are: the lack of evidence that the model is producing speeds that accord with observation rather than reproducing the Regional Transport Model's (RTM) SATURN speed-flow curves; the unstated rationale for the maximum speeds assumed for different vehicle types; the unverified use (in forecasting) of the relationship for urban areas between speed and total demand by all modes; and the use of national average speeds, rather than link-specific speeds, in the fuel consumption relationships – with special reference to the implications for HGV routeing.

We have also identified that the assignment results fall below the thresholds expected within the TAG guidance and as such the DfT needs to be satisfied that the model is fit for the purposes for which it is being used for each application.

In relation to the demand model our key observations are: the understating of the critical problems stemming from the lack of spatial detail for destinations in the National Travel Survey (NTS) estimation data; the understating of issues associated with the destination constraint; the low value of time for rail; and the absence of evidence about the trip length distribution. In spite of these, we noted that the mean modal costs were well reproduced for all purposes and the first round elasticities were plausible.

In relation to the pivoting process, we have general concerns relating to the use of the period-specific highway assignment matrices as a pivot without any corresponding checks on the 24-hour demand matrices. This has a number of repercussions. Firstly, while the commute and education purposes are constrained at the destination, the number of trips attracted to destinations for other purposes is subject to considerable potential error, and the non-home based trips are dependent on these. This leads to even greater uncertainty when converting the (unconstrained) home-based purposes to origin-destination format and adding in the non-home based trips. This concern is borne out as, with the exception of the doubly constrained purposes, the comparison of the resulting trip length distributions with NTS data is not encouraging. Hence the application of the calculated ratio to the pivot highway matrices is far from robust.

This concern is exacerbated by the issues associated with the construction of the pivot matrices themselves (the description of which is one of the poorest parts of the model documentation). While for the doubly constrained purposes the use of substantial external datasets (Census Journey to Work and the schools census) should produce reasonable estimates, the matrices for the remaining purposes are likely to be much less robust, and the non-home based matrices are consequently likely to be based on different attractions from those predicted by the VDM. Further, little information is available as to the extent of the "matrix estimation" process, following the conversion to time-period format. Finally, the level of mismatch seen in the future year test (Sensitivity Test 1) in terms of NTEM growth does not give confidence about the underlying quality of the base year home based productions.

We are surprised at the lack of validation relating to a) the match between the 24-hour matrices which form part of the construction of the highway pivot matrices and the corresponding VDM-based matrices from the Base Year Run, or b) the match between the pivot matrices (after matrix estimation) and the corresponding synthetic base matrices from the Base Year Run. Only if these could be shown to be reasonably aligned would we be willing to accept the pivoting procedure as being robust.

Finally, on pivoting, the fact that the process is confined to the car mode reduces the value of the model for producing forecasts for other modes. Here again, a greater level of validation for the Base Year Run against external data sources (for example, LENNON) could potentially increase confidence in the model's ability to represent the non-car modes.

During our review a number of systematic unexpected results have emerged from the Sensitivity Tests which strongly suggests that there may have been issues of lack of convergence in these runs that then reduce the range of conclusions that can be drawn from them.

In particular, there are unexpected similar levels of growth in car driver and passenger trip productions in London indicated in each of Tests 2, 3 and 4, even though these policies were not necessarily expected to have such an impact within London. Accordingly, considerably uncertainty about the overall performance of the model still remains because of these difficulties encountered in assessing the results of the various sensitivity tests.

The potential impact on the test results from lack of convergence may have been amplified in some cases by the relatively small scale of some of the actual policy tests.

If the various counterintuitive results from sensitivity tests could largely be resolved simply by running the model for a few more iterations, then this would be a positive development. The extra run time burden that this entails would in the medium term gradually become less of a problem, assuming that computer facilities continue to increase in power and speed through into the future.

If on the other hand the lack of good convergence is more structural within the modelling system and cannot be universally resolved by a few extra iterations then the issues remain serious. Moreover, if it transpires that some of the significant counterintuitive results that arose are not simply due to lack of convergence then more major issues with the model design or implementation would need to be considered.

Based on the information currently published on the model results and on the lack of information presented on the convergence achieved, it is difficult to be certain about which of the above situations holds. Consequently, further detailed experimental model running and analysis would be required in order to provide more informed recommendations on the capabilities of the model, so this uncertainty should be noted when considering the comments below on individual Use Cases.

The NTMv5 and the Use Cases

The NTMv5 has been considered against five Use Cases which were defined by the DfT. We have considered how the NTMv5 has been constructed in terms of its overall demand and supply structure alongside how the highway network and its supporting zone system have been coded.

We have concluded that the NTMv5 can provide some evidence for most of the use cases but for each of them we advise caution with the inputs for any test being carefully specified and the outputs being thoroughly reviewed. This is primarily due to the focus of the NTMv5 being on highway, and more accurately, the more strategic highway network, whereas many of the use cases focus on urban travel policy and public transport interventions. We agree with the model Development Team in saying that the NTMv5 should only be used by modelling practitioners capable of conveying how the limitations of the NTMv5 impact on the evidence which it produces.

For Use Case 5 in particular, the performance of the model in forecasting mode indicates that the road traffic forecasts in the regions other than London are not implausible. It was concluded that, subject to the caveats above, the model should be suitable for use in forecasting the growth of road traffic in most areas other than those adjacent to or within major urban areas.

Recommendations to the Department

Based on this review of the NTMv5 structure and performance and on its potential usage for policy testing, we have assembled a set of recommendations to be considered for potential future enhancements to NTMv5. These recommendations for NTMv5 are grouped by the time scale over which the enhancement tasks could be implemented, distinguishing: short term tasks that could be introduced relatively soon through minor adjustments to the model set-up or usage; medium term tasks that would require more substantial modifications to the model inputs and so might necessitate some limited adjustments to the model calibration; and longer term more fundamental changes in which the model structure, software, calibration or base matrix might undergo significant modifications so that a subsequent full validation and model testing exercise would then be appropriate. As these tasks fall in reality along a continuum of complexity and of resource and data requirements this allocation by time horizon is necessarily fluid.

Our recommendations also extend to enhancements to the NTMv5 inputs from NTEM and from its underlying car ownership forecasting model, so these requirements are also outlined. Finally, some of the recommendations imply that changes to a small number of elements within the current TAG guidance would be beneficial so the underlying reasons for these suggested changes are explained.

Immediate Model Checks and Adjustments

We recommend that the DfT investigate and resolve the source of the systematic pattern of noise in the results identified for the Sensitivity Test runs 2 to 5. They should ensure that the model is always run to an adequate level of convergence and that the degree of convergence achieved is always published for each policy test run, using an appropriate cross-section of indicators. Success with this improvement could increase confidence in the resulting revised outputs from the Sensitivity Tests, which in turn could improve the capability for tackling some of the Use Cases.

We would encourage the DfT to assess the quality of the synthetic base matrix that is the foundation for the VDM and of its match to the Base matrix and to patterns observed in the NTS. The understanding gained on the strengths and weaknesses of this match will aid in understanding and interpreting the results of policy measures being tested.

There are various further adjustments that are likely to be beneficial to the design and implementation of this synthetic base matrix. DfT should reconsider the decision to estimate the population segmentation within MSOAs in the base year using the Any Year Census (AYC) procedure. This disaggregates population segments from the District level to the MSOA/zonal level based just on the property type mix in the MSOA, whereas methods that start instead from the detailed population segmentation available from the 2011 Census at MSOA level should provide more accurate results for most zones through lessening disaggregation error. In particular the underlying car ownership pattern in the base year should be improved, ensuring that it accords realistically with the pattern observed in denser urban areas. The population segmentation 16-64 and 65+ should be realigned to match to the current NTEM segmentation 16-74 and 75+, so that the current inconsistency for the 65-74 year age group between NTMv5 and NTEM is circumvented. Finally, the form and implementation of the Non-Home based trip purpose models should be re-considered.

An analysis of the soon to be published 2020 DfT van survey would help obtain a better understanding of the current spatial patterns and of the trends through time for both freight and non-freight LGV trips and vehicle kilometres. The results from this analysis should help in assessing the ability of the model to forecast LGV growth and to assess LGV responses to policy measures being tested.

Medium Term Improvements to the Model Performance

Switching throughout to use link-based vehicle operating costs (VOCs), rather than VOCs based either on system average speed in the highway assignment model or on zone-pair average speed in the VDM, should significantly improve the performance of the choice of routes in the assignment for HGVs in particular. This link based VOC approach would also have the additional advantage that the estimated fuel consumption within the VOCs should now align consistently with the environmental emissions calculations in the post-processing of the forecast vehicle flows on links.

The pivoting process should be re-considered to see whether a) it could be additionally implemented on a Production / Attraction (24- hour) basis for the home based purposes (as this would stabilise the application of the non-home based purposes) and b) extended to other modes (such as rail) in order to improve the robustness of the model in forecasting non-car modes.

Longer Term Model Enhancements

The segmentation adopted within the spatial distribution model should be re-considered to ensure that it distinguishes realistically the differences in travel patterns between segments. In particular as part of the model estimation procedure, tests for differences in deterrence parameters between a range of segments should be carried out in order to ensure a good match to the observed clear differences in trip length and destination zone patterns: between home based education movements of primary and secondary and other students; as well as between home based work movements of industrial sectors and full-/part-time workers. A sequential estimation approach would facilitate making effective use within the distribution model estimation of other data sources such as the Census Journey to Work matrices and the School Census data.

Recommendations for NTEM and NATCOP

The performance of the car ownership model in dense urban areas should be i**mproved** to take appropriate account of the impact of densification on car ownership rates and trends. It should be ensured that the spatial pattern of its forecast changes in car ownership rates across areas of different densities are broadly consistent with recent trends, except where there are clearly identified reasons for any forecast breaks in trends.

Recommendations for TAG

In the course of the review some aspects have emerged where the NTM developers have followed the current TAG recommendations but where this may have impacted on model performance. A number of current TAG recommendations could be reconsidered by DfT, as now listed.

The guidance in TAG Unit M3.1 should recommend VOC calculations for all road vehicle types to be based on the vehicle speed on individual links, rather than being based on the average speeds for the area.

A related aspect within TAG Unit M3.1 is the recommendation to double the driver's VOT for HGVs to "take account of the influence of owners on the routing of these vehicles". The more appropriate approach would be instead simply to apply the link based VOC in a form that takes full account of the operating cost of the vehicle on that link. Revise the road vehicle fuel consumption formula specified in TAG Unit A1.3 to be based also on link type (speed limit and road type) and not solely on vehicle speed - the current formulation.

Contact name Andrew Stoneman

Contact details 0117 930 6271 | andrew.stoneman@wsp.com

1 Introduction

1.1 Background

1.1.1 A new version of the National Transport Model has been developed during the period 2016-2019 and is referred to as NTMv5. While this retains the link with the Department's National Trip End Model (NTEM) to provide estimates of exogenous growth over time, in all other respects this is a completely new model.

1.1.2 The model has the general structure of a conventional transport model, in that it contains modules representing (highway) assignment and multi-modal demand responses in terms of mode and destination choice. In addition, the assignment "pivots" off a set of base highway matrices which have been specially developed. Relative to a conventional transport model, the challenge, of course, is to deal with the size of the study area, which here concentrates on England but needs to take some account of other parts of Great Britain.

1.1.3 For the purpose of this Peer Review, we have been provided with the following documentation:

- NTMv5 Developer Guide version 2.0, in 6 volumes:
    - Volume 1 Model Zones and Highway Network (referred to as Volume 1 in this review)
    - Volume 2 Base Year Demand (Volume 2)
    - Volume 3 Highway Assignment Model (Volume 3)
    - Volume 4 Non-car Modes (Volume 4)
    - Volume 5 Variable Demand Model (Volume 5)
    - Volume 6 Model Testing (Volume 6)
- NTMv5 Future Development: Quality Report version 4.0 (the Quality Report)
- NTMv5 User Guidance: Installation and User Guidance version 4.0 (the User Guide)

1.1.4 In addition, since we judged that Volume 5 did not provide sufficient information on the Demand Model, we requested and obtained the further document:

- National Transport Model for England: Estimation of the mode-destination models version 13.

1.1.5 We have some general comments on the documentation, which are presented at the end of this review. However, our assessment of the model is entirely based on the documents listed above.

1.1.6 In the following section we provide a high-level description of the model, and then go on to consider its various components in more detail in subsequent chapters. After that, we consider the tests of the model (as set out in Volume 6), and on that basis provide our assessment.

1.2 The Model in Outline

1.2.1 As noted, the model consists of a multi-modal demand model and a highway assignment model, pivoting off a base matrix. The zoning system is generally at the

MSOA level, so that there are 6,772 zones in England, together with 264 additional "bespoke" zones that represent gateways, major non-residential trip attraction areas and major future year growth sites. In addition, there are 68 zones representing Wales and 27 in Scotland, and a further 30 "bespoke" zones representing airports and seaports.

1.2.2 Trip rates extracted from the National Trip End Model (NTEM), owned by the Department for Transport, are applied to the local 2015 population, estimated using the Any Year Census process developed within the study, to generate trip productions by all modes combined for the 6,772 MSOA zones in England, at a considerable level of segmentation by purpose and person/household characteristics: these constitute a basic input to the Model. After a certain amount of aggregation of segments, these are distributed by destination and mode by the Demand Model, being influenced by modal costs and level-of-service [LOS], as well as land-use Attractions (also from NTEM). The modes considered are:

- Car driver
- Car passenger
- Train
- Bus
- Cycle
- Walk

1.2.3 The trips allocated to the Car Driver mode are interpreted as vehicle trips and, in addition to matrices for LGVs and HGVs provided by independent models, assigned to the highway network. As noted, the assignment pivots off a set of base matrices. The assignment then delivers a revised set of highway costs/LOS to the Demand model, and the process iterates until acceptable convergence is reached.

1.2.4 Separate assignments are carried out by three time periods (AM peak, Interpeak and PM peak): for this allocation, a set of fixed time of day factors are used, varying by trip purpose. The outcome costs/LOS are averaged in different ways according to trip purpose within the Demand Model. The trip purposes considered are:

- Home based Work (HbW)
- Home based Employers Business (HbEB)
- Home based Education (HbEd)
- Home based Shopping and Personal Business (HbShopPB)
- Home based Recreation, Social and Visiting Friends and Relatives (HbRecV)
- Home based Holiday and day trip (HbHol)
- Non-Home based Employers Business (NHbEB)
- Non-Home based Other (NHbO)

1.2.5 For future years, NTEM-based growth is applied to the base year trip ends (and corresponding growth in the HGV and LGV matrices).,The model responds to changes in the highway network and to changes in the input costs/LOS for other modes. Note that there is no treatment of public transport crowding.

1.2.6 The model was required to deal with the following "Use Cases" (the Quality Report Section 2.4):

- UC1 Strategic Roads Investment and Resilience:
  - To analyse the impacts of packages of roads schemes at a national level. This could include high-level calculation of value-for-money (VFM), points of expected congestion, and analysis of resilience of the network.
- UC2 Road User Charging and other potential policy:
  - Flexibility to adapt to road policies in future Parliaments. This could include various forms of road pricing, including pricing on strategic roads (tollbooth, distance-based or vignette), urban roads (e.g. congestion charging vignette), or parking policy in urban areas, or other behavioural devices such as High Occupancy Vehicle (HOV) lanes.
- UC3 Local Investment and Policy:
  - Variety of analysis including national impacts of congestion relief schemes; Policy impacts of introducing public transport improvements (e.g. light rail). This could potentially include travel demand management in future parliaments. It is noted that parking is covered in use case 2.
- UC4 General Support for DfT Teams (other than Roads / Local):
  - Environmental analysis of transport policies relating to carbon and/or an approximation of air quality emissions, and aviation surface access.
- UC5 Scenario-based National Traffic Forecasting:
  - Understanding of changes in population or travel trends (e.g. driving rates amongst young people), with scenarios around GDP, car ownership, fuel price, road tax.
- UC6 Exploring the unknown:
  - Testing new policies or technical developments that have not been modelled before (e.g. CAVs). Testing new policies or technical developments of whose existence we are not currently aware.

1.2.7 The Department has indicated that Use Case 5 (UC5, National Traffic Forecasting) should be considered first for the following reasons:

- Production of National Road Traffic Forecasts (RTF) is an existing use and the most fundamental purpose of the National Transport Model;
- UC5 includes scenario-based forecasting, and it is likely that many of the questions arising in UC1 and UC2 will be considered as scenarios for RTF; and
- In order to robustly forecast road traffic for the RTF, it will be necessary for NTMv5 to include the influence of a wide range of factors. Those factors highlighted in other use cases can therefore be considered as implicitly features of UC5.

1.2.8 While the focus on UC5 is appreciated, the validity of the model will be assessed for all the Use Cases in this review.

1.2.9 Six Sensitivity Tests have been reported, and these are reviewed as part of the model assessment.

1.3 Structure of this Report

1.3.1 In preparing our review, we have taken the general line that the main assessment of the model should be able to be carried out on the basis of the 6-Volume Model Development Report (though we have already noted that for the Demand Model, this needed to be supplemented by the Estimation Report). Our overall impression of the quality of this documentation is very favourable, in that we have usually been able to find the necessary detailed information. There are, however, a small number of cases where the information could only be found in the Quality Report: we have drawn attention to these, as we think these omissions from the Development Report are undesirable[1].

1.3.2 In Chapter 2, we assess the construction of the model in detail, considering all its constituent components, along the lines of Figure 2.1 of the Quality Report.

1.3.3 Chapter 3 deals with tests carried out on the model from its base year position. This covers the validation of the highway model, the realism tests, and Sensitivity Tests 2 to 5.

1.3.4 In Chapter 4 we move on to the additional issues associated with future year forecasts where the main discussion relates to Sensitivity Test 1.

1.3.5 Chapter 5 discusses the other documentation submitted to us for review; the Quality Report and the User Guide. In line with the remarks above, our assessment of the Quality Report essentially concerns its structure and content, rather than any discussion of the results *per se*.

1.3.6 In Chapter 6, we assess the achievements of the model in respect of the Use Cases, with a view to drawing conclusions about how far it can be used with confidence.

1.3.7 In addition to the assessment based on the documentation, we have also had the opportunity to run the model ourselves, and for this purpose we have carried out a number of tests, both to replicate existing results, and to understand the usefulness of the model and the plausibility of the results for some devised illustrative strategies. These are reviewed in Chapter 7.

1.3.8 Finally, in Chapter 8, we draw out our conclusions and recommendations.

[1] We have been informed that the documentation is “tiered” with the expectation that User Guide readers should be familiar with the Quality Report, and Developer Guide readers should be familiar with both the Quality Report and User Guide. Nonetheless, we think that all the model detail should be in the Quality Report.

2 The Model

2.1 Highway Assignment Model Structure

Documentation

2.1.1 The development of the Highway Assignment Model (HAM) has been documented in:

- Volume 1 Model Zones and Highway Network
- Volume 3 Highway Assignment Model
- Volume 5 Variable Demand Model

The Zone System

Requirements of the Zone System

2.1.2 The HAM zone system had the following requirements specified:

- maintaining a consistent level of aggregation for all tools and datasets, which includes consistency with the National Trip End Model (NTEM) dataset and the Highways England Regional Traffic Models (RTMs);
- identifying bespoke zones that represent gateways, major non-residential trip attraction areas and major future year growth sites;
- representing a level of spatial detail that can be supported by sample data, including consideration for freight movement data;
- providing consistency with FORGE1 geography (sub region and area type, based on NTS); and
- providing consistency with other models, e.g. DfT models of access travel to Heathrow and Gatwick, and Great Britain Freight Model (GBFM).

2.1.3 To fulfil these requirements within the modelling software would imply substantial computational resources, both in terms of software and hardware which in turn would have meant excessive run times. The model development team refined and re-defined a series of zone systems which balanced the requirements above with the pressure on model run times.

2.1.4 The model development team identified a series of candidate zone systems which were already in use at the national or sub-national level, specifically from National Trip End Model version 7.0, the Highways England Regional Traffic Models and PLANET. Each of these systems had limitations which would have implied significant effort to modify them to fulfil the NTMv5 requirements which meant that the conclusion to build a new zone system was reached.

2.1.5 This conclusion is a reasonable position to adopt and allows the zone system to be tailored to accommodate the NTMv5 requirements more effectively than repurposing an alternative zone system.

Spatial Coverage of the Zone System

2.1.6 The zone system ultimately implemented in NTMv5 can be summarised as:

- Internal zones, predominantly in England but with some areas of North East Wales included where there are very strong commute patterns. These zones have been based on MSOAs;
- External zones, the balance of Wales and all of Scotland which are based on NTEM zones.

2.1.7 Having established the internal and external zones further reviews were undertaken:

- to identify whether there were any candidate zones for aggregation to reduce the total number of zones (and associated connections to the transport networks); and / or
- to decide whether there were specific locations which required isolation from the generic rule-based approaches for trips.

2.1.8 The aggregation review concluded that there was no benefit to such a simplification and as such MSOAs were retained across the internal zones.

2.1.9 The creation of bespoke zones, either as a spatial area within an existing MSOA or a point to represent a specific location did increase the number of zones. Implementation was determined based on:

- The mix within the zone of commercial / industrial and residential uses. MSOAs are defined by resident population which means that commercial / industrial zones tend to be spatially large. The implication of this is that a simple single zone with one connector is likely to lead to incorrect loading and excessive demand at the first point of contact with the coded highway network. This issue also applies to sparsely populated areas;
- Specific uses that were considered as subject to exogenous factors impacting demand included ports, airports, freight hubs, hospitals and enterprise zones. These were isolated from the MSOA zone in which they are located to allow different rules to be applied where appropriate.

2.1.10 Zone attributes have been applied such that zones can be aggregated using different characteristics. This includes:

- Zone type – MSOA or bespoke zone and internal / external;
- Region of England or Wales / Scotland;
- NTEM area type (e.g. 1 = London, 8 = rural);
- Sector which is an indicator based on a combination of region and where the zone lies with respect to screenlines used for calibration and validation;

2.1.11 In addition to the attributes, each zone has other data applied to indicate the MSOA name, the Local Authority District it resides in and the Ordnance Survey Grid Reference of its centroid.

2.1.12 Table 2-1 shows a summary of the zone system:

Table 2-1: Summary of NTMv5 zones (v6.5) by Type and Region

Region	Standard	Enterprise Zone	Port / airport	Major Attractors	Freight hub	Total
North East	339	3	10	5	3	360
North West	924	8	16	4	2	954
Yorks & Humb	691	6	9	5	4	715
East Midlands	573	3	4	5	1	586
West Midlands	735	4	1	1	4	745
East of England	736	9	20	4	1	770
London	983	1	7	5	1	997
South East	1,091	9	23	12	1	1,136
South West	700	6	24	12	1	743
Wales	68	0	7	0	0	75
Scotland	27	0	23	0	0	50
Total	6,867	49	144	53	18	7,131

The Highway Network

Source of Highway Network

2.1.13 The NTMv5 network was initiated from an early version of the Highways England Regional Traffic Models (RTMs[2]). Although these networks were coded in SATURN it was possible to migrate them into VISUM[3] for use in the NTMv5. The RTMs were coded to a common coding manual, but each has its own bespoke modifications.

2.1.14 Volume 1 paragraph 3.2.1 indicates that network quality assurance tests were undertaken to check the following issues:

- Ensure there are no SATURN semi-fatal errors;
- Check of coded link length against crow-fly distance;
- Check of consistency in the reverse direction;
- Check coded speed flow relationships, mid-link capacity and number of lanes; and
- Check junction type of the nodes and banned turn information.

2.1.15 Volume 1 then states that the audit process concluded the networks were of an appropriate level of detail to act as a starting point for NTMv5. There is no further evidence of the checks provided in the reporting nor any discussion of how errors once identified were corrected either before or after migration into VISUM.

2.1.16 Using the basic topography of the RTMs in combination with one another the NTMv5 model development team then coded a range of additional information to each link and added centroid connectors to provide the join between the zones (and matrices) and the network itself.

[2] The RTMS themselves have been independently developed since them by Highways England.
[3] To streamline the functioning of this large model, it has been implemented largely within a single transport modelling package VISUM, provided by PTV.

Centroid Connectors

2.1.17 A series of rules governing the number and length of centroid connectors associated with each zone have been applied and secondary rules about the link types which connectors can join to have also been considered. The rules are all reasonable in terms of maintaining the balance between demand and model stability as although more connectors may lead to more realistic local trip loading, too many connectors would introduce additional route choice.

Link Types, Volume Delay Functions and Speeds

2.1.18 Link type definition allows carriageways of different standard to be identified and grouped together such that general rules can be applied to them. For example, a link type might be a derestricted rural single carriageway or an urban dual carriageway with direct access from developments.

2.1.19 The link type definitions which were coded during the RTM development have been retained in the NTMv5 network albeit with the ability to modify these types during the calibration process. However, slip roads have been singled out for additional attention based on the notion that this type of link contributes to a discrepancy between total link lengths in NTMv4 and FORGE[4]. Slip roads have been allocated one of 13 different types. Evidence has not been provided regarding the distance discrepancy assertion nor is there information about how well the link types from the RTMs match other link length summary data from NTMv4 or FORGE.

2.1.20 The relationship between traffic volume and speeds on links is discussed as SATURN and VISUM have different functions for estimating speed based on volume of traffic – and, by inference, delay, defined as the difference between the free-flow time and the actual time achieved on the link.

2.1.21 A series of tests to identify the most appropriate coefficient values for the VISUM volume delay relationships is reported and evidence provided that a match between the two curves can be generated.

2.1.22 The evidence which is missing from the analysis is whether the speed flow curves used within the RTM SATURN networks are producing speeds which correspond with observations. Effectively, it has been assumed that the RTM speed flow relationships are producing the correct speeds so the VISUM model needs to be able to replicate those relationships successfully. Providing evidence that the NTM replicates observed journey times would mitigate this issue.

2.1.23 The main HAM documentation[5] does not explain that distinct free flow link speeds are used within the model in order to take account of the differences between cars, LGVs and HGVs. Nonetheless this speed differentiation is a potential strength of the model.

2.1.24 In the Quality Report we read that for each link type the different vehicle types are assigned different "maximum speeds." It is not discussed whether these are intended to refer to a) the speed limits applying to different vehicle types or b) the

[4] FORGE is the highway supply module of the NTMv4 and stands for Fitting On of Regional Growth and Elasticities

[5] The only reference appears to be in in the Quality Report Section 8.4.

actual performance of different vehicle types. Since the values implemented for cars and LGVs are 130 and 120 kph respectively, and these are higher than the existing maximum speed limit of 112 kph (70 mph), it must be concluded that they do not represent the speed limit.

2.1.25 Less clearly, for HGVs the maximum speeds implemented are “70 kph for single carriageway links and 96 kph for dual carriageway links”: the speed limit for HGVs on dual carriageways (including motorways) is indeed 96 kph, though for built-up areas it is 48 kph and for single carriageways 80 kph. Further, the implemented list of link-types indicates that for some link types different maximum speeds have been chosen.

2.1.26 The Quality Report goes on to say the… “free flow speed for the mode is then taken as the minimum of the free flow speed for the link type and the maximum speed for the mode (on the link type).” However, our investigations suggest that it is more complicated than that. As well as the maximum speeds for each vehicle type, there is a link “free-flow” speed (corresponding to SATURN’s S_0) which is the base point for the volume-delay function [VDF] which calculates the link time as a function of total flow relative to capacity. In any iteration, if this calculated time is **less than** that implied by the maximum speed for a vehicle type, we believe that the calculated time is replaced, for that vehicle type, by the higher time associated with the maximum speed (note that the form of the VDF ensures that the link travel time cannot be lower than that associated with the link free-flow speed).

2.1.27 Since the maximum speed for both cars and LGVs is currently set higher than the value of S_0 for all possible link types, the adjustment just described will not apply to them, and the maximum speed will have no effect, By contrast, the mechanism does allow a different (lower) link speed for HGVs, and implies that the speed of HGVs is not affected by congestion until the general level of congestion brings down the link speed for all vehicles to the maximum HGV speed. This appears to be a realistic approach. It is unfortunate that this is not fully described in the available documentation.

Urban Area Speeds

2.1.28 These are briefly touched upon in Volume 1 Section 3.6 with more details provided in Volume 5 Section 7 and the Quality Report Section 4.7.

2.1.29 The approach to Urban Area Speeds (UAS) has been adopted from the RTMs but is recognised in Volume 1 to have significant limitations. Essentially, the speed on an urban link is “fixed” from a process outside the HAM itself and remains constant regardless of changes in traffic flow on that link. This simplification is required as observed speeds in urban areas are not simply a function of link type and flow nor are all the links in an urban area coded, so each modelled link may represent several links.

2.1.30 Volume 5 Section 7 addresses the UAS concept in greater detail and explains that five options were explored to estimate highway speeds in urban areas:

- Option 1 Fixed Speeds which are defined once for the project and not altered;
- Option 2 Speeds respond to overall growth in trip ends and not any capacity change;
- Option 3 Mode / Distribution response whereby speed changes are fed back into the demand response;
- Option 4 Aggregate traffic kms across urban areas which are effectively area wide speed relationships that vary in response to trip length and route choice; and
- Option 5 Link speed flow relationships.

2.1.31 Option 2 was favoured on the basis that:

- Option 1 is overly simple and would not be appropriate;
- Option 5 is not appropriate as a large proportion of all urban links are not included in the model network.
- It is not clear how speeds would be estimated in Option 3 as a function of mode choice or trip distribution although more research is suggested.
- Similarly, more research is suggested for Option 4 and so it is not ready to be implemented.

2.1.32 The function which has been implemented essentially scales highway speed in each specific urban area in relation to the ratio of observed base speed within that urban area to observed speed during the off-peak factored by all day home-based trip end growth by all modes. The use of all modes is presumably meant to capture all growth in demand as opposed to a segment of demand such as car driver.

2.1.33 Further observations are now made about the function which has been implemented, which imposes specific limitations on the validity of outcome speeds:

- Forecast speeds are not influenced by any changes in road capacity in the urban area, so measures to increase or decrease road space for cars (e.g. due to the growth in bus lanes, cycle lanes and pedestrianisation) are not taken into consideration;
- Forecast speeds are a ratio of forecast total trip ends to base year total trip ends, which means that when testing scenarios in which the competition between car and competing PT modes is adjusted, which in turn will generate changes in the level of car traffic and congestion within urban areas, the urban car speeds will nevertheless remain unchanged in the model. Also, in policies where urban car ownership or availability is significantly changed this will only have a minor influence on the speed in an urban area via the NTEM differences in <u>total</u> trip rates between car ownership classes.

2.1.34 The calculation of the speed reduction factors identified specific values for each combination of modelled road type and local authority area. These have been summarised in Volume 5 Table 7.2 and show that speeds in London are generally 30% lower in the daytime than off-peak with most of the rest of the country between 25% and 20% lower.

2.1.35 Volume 5 Section 7 demonstrates how the UAS function as defined operates within the model and concludes that the function is operating as anticipated when implemented in a forecast which introduced a notional motorway scheme in the South West. Figures are also provided to indicate where speed reduction factors change over time to demonstrate the sensitivity of the function to both base observed speed reduction factors and trip end growth.

2.1.36 Ultimately, whilst the UAS approach has employed statistical techniques to produce values mapping off-peak speeds to modelled period speeds on a systematic basis and those speeds have been varied in forecasts as a function of trip end growth, the relationship between trip ends and urban speeds has not been proven in this work nor has evidence of its validity been provided. The extent to which it can be relied upon as the basis for forecasting speed in urban areas should be challenged. Beyond the unproven historic link between urban speeds and trip ends the function cannot explain how urban speeds may vary based on factors outside of trip end growth.

Junction modelling

2.1.37 The allowed turns and priority rules for junctions have been imported from the RTMs. Travel time to traverse a junction uses a simple relationship that relies on volume delay functions for each movement.

2.1.38 Essentially, the resultant delays are a function of the turning flow making the movement and not affected by opposing movements. This form of junction modelling is suitable for use in large models as it minimises model run times by simplifying the delay calculations within an assignment which subsequently reduces the assignment (route choice) and simulation (journey time calculation) loops.

Other Restrictions

2.1.39 Where there are HGV restrictions on the A road network these have been applied in the highway coding.

Generalised Cost

Form of the Generalised Cost Function

2.1.40 The generalised cost or impedance function within VISUM builds up the route travel costs between each origin zone and destination zone. The function is presented in Volume 1 Section 3.7 as:

Equation 1

$$GC = \frac{VOC \times distance + toll}{VOT} + Time$$

2.1.41 It is expanded in Volume 3 Section 3.4 to:

Equation 2

$$GenCost_{o,d,auc} = \frac{VOC_{auc} * dist_{o,d,auc} + toll_{o,d,auc}}{VOT_{auc}} + time_{auc}$$

Where:

VOC_{auc} = denotes the vehicle operating cost per unit distance travelled for the assignment user class auc

VOT_{auc} = is the value of time for the auc

$dist_{o,d,auc}$ = is the distance travelled from origin o to destination d by the auc in the HAM

$time_{o,d,auc}$ = is the time taken to travel from origin o to destination d by the auc in the HAM

$toll_{o,d,auc}$ = is the toll incurred by the auc travelling from origin o to destination d in the HAM

Value of Time

2.1.42 Volume 1 Table 3.4 states that values for VOT have been taken directly from the December 2017 TAG databook[6]. This appears to be correct for car commute and car other trips but is not correct for car business trips. These values appear to have been adjusted from market to perceived cost which should be clearly stated in the report text.

Table 2-2: Value of Car Business Time Comparison (pence per minute)

Period	Vol 1 Table 3-4	TAG databook
Morning Peak	32.28	38.41
Interpeak	33.08	39.36
Afternoon Peak	32.74	38.97

2.1.43 The value for LGV in Volume 1 Table 3.4 of 22.81 pence per minute does not appear in Table A1.3.5 of the TAG databook as the weekday average LGV value is reported to be 26.76 pence. Again, clarifying that this conversion has been applied should be stated in the report text.

2.1.44 The value adopted for HGVs in Volume 1 Table 3-4 is 57.90 pence per minute to implement guidance from TAG Unit M3.1 paragraph 2.8.8 which suggests that the value of time may be doubled in the assignment for HGVs. This guidance does appear to have been applied but as the value of time in the TAG databook is 27.56 pence per minute the correct value should be 55.12 pence per minute.

Vehicle Operating Cost

2.1.45 It is noteworthy that the variation in link speeds by vehicle type[7] is not used within the VOC calculations. The method adopted for the HAM is to convert the TAG VOC function in use from being a function of link travel time, distance and speed to instead being just a simple function of distance:

> *"For the HAM v(o,d,auc) is replaced by an assumed speed, applying to all O-Ds, for each user class (auc) to determine the VOCs by segment using*

[6] https://webarchive.nationalarchives.gov.uk/20181113125647/https://www.gov.uk/government/publications/webtag-tag-data-book-december-2017 accessed 05/03/2020. User parameters "Price Year" and "Value Year" both set to 2015.

[7] note that given the high maximum speeds assumed for cars and LGVs, it is only HGVs which may have a different speed (as discussed in 2.1.27)

the TAG databook formula. This simplification of an assumed average speed is typically used in assignment models for route choice since the TAG formulae are intended to be used with average speeds of travel and not considering variations on a link by link basis." Volume 3 Section 3.2 (underlining added)

2.1.46 In contrast to the HAM assignment stage, zone pair specific car speeds are used in VOCs for the VDM stage:

> *"The VDM includes the VOCs for cars as part of the utility calculations and uses speed (distance / time) for each zone pair to determine the VOCs by segment."* *Section 2.2.3 of Vol. 5 (underlining added)*

2.1.47 Two potential issues arise with the approach adopted to calculate VOCs across NTM component stages.

2.1.48 Firstly, it is unhelpful that the distribution of vehicle speeds across individual links does not impact appropriately on the VOCs used in the HAM. Certainly, the TAG formulae were originally derived (Ricardo-AEA, 2014)[8] from analyses of **average** speeds, rather than representing consumption rates just at a **specific** speed. The fuel consumption at each average speed was derived from a specific drive cycle. But the set of individual drive cycles used included: congested urban; free flow urban, congested motorway; free flow motorway, etc. (Barlow et al. 2009)[9], so that the formulae are calibrated to capture the different fuel consumption rates experienced in these different situations.

2.1.49 Accordingly, the adoption of a single national average speed will imply substantial aggregation errors within the calculation of the fuel costs and consumption in the HAM. For example, the VOCs per car mile calculated in congested start-stop conditions in Central London[10] will be deemed the same as those in free-flowing rural Lincolnshire. Likewise, the major increase in HGV fuel consumption per mile that occurs in reality on low grade rural roads (because the vehicle is subject to regular speed changes, stops at junctions and irregular gradients) are not distinguished from the much lower consumption levels experienced when travelling at a constant speed on a motorway. In this way, the assignment will tend to wrongly allocate HGVs away from faster motorway links and onto marginally shorter minor roads or onto shorter routes straight through congested urban areas.[11]

[8] Ricardo-AEA (2014) Production of Updated Emission Curves for Use in the National Transport Model. Report to the Department for Transport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662795/updated-emission-curves-ntm.pdf accessed 06/05/20

[9] Barlow TJ, S Latham, IS McCrae & P G Boulter (2009) A reference book of driving cycles for use in the measurement of road vehicle emissions. TRL Project Report PPR354. https://www.gov.uk/government/publications/road-vehicle-emission-factors-2009 accessed 06/05/20

[10] This shortcoming may partially explain the inappropriate routing noted in Volume 3, Section 5.6, where long distance traffic through London and Manchester opted to use smaller, local roads that pass through these urban areas, as opposed to using more plausible ring roads, such as the M25 or M60.

[11] This shortcoming may partially explain why it has been found necessary to double the VOT being applied to HGVs

2.1.50 A better approach would calculate the VOC on each link, based on the speed of that vehicle type on that link. This approach is described in greater detail below in Section 8.5 It should not significantly increase run times but would produce more realistic costs to govern the VDM and should also improve the assignment, particularly for HGVs, by comparison with the current use of a zone pair specific speed and a single national average speed, respectively, to calculate VOCs in these model stages.

2.1.51 Secondly, for each car auc, the national total VOC cost summed over all trips between O-D pairs which is based on each individual O-D speed in the VDM, is not expected to match the total cost from the HAM, which is based on the national average speed. This difference between model components arises because of aggregation error differences between these procedures. It would be helpful to confirm that the size of this cost gap is small and would remain so across a range of policy tests.

2.1.52 In purely urban models the two issues listed above would be likely to be of diminished significance, since on purely urban roads the range of speeds between road types would be limited because of urban speed limits. However, in national models where some routes will be entirely on high speed interurban roads, others entirely on low-speed congested urban roads, while many are a mix of these two extremes, these issues are of considerable significance and they should be assessed thoroughly in order to help in interpreting the likely impacts of aggregation errors on the model results from specific types of policy tests.

2.1.53 The documentation of the HAM in Volume 3 would be improved if it provided a full and clear explanation of how the model represents both: the basic differentiation of free flow vehicle speeds by vehicle type and road type; and how these different speeds are subsequently adjusted in model iterations in response to link congestion changes.

Tolls

2.1.54 Tolls have been coded onto the appropriate links in the model at their financial values. For most individual tolls listed in Table 7.7 of the Quality Report the charges presented for each vehicle type appear suitable. It is explained there that the representation of the London congestion charge has been adopted from the Transport for London Central London Highway Assignment Model (CLoHAM) but then provides no further information on the underlying logic behind the charge values indicated. For example, the charge coded in Table 7.7 for the Car non work category is £1.01, whereas the real 2015 congestion charge was £11.50 per vehicle per day. Even if the charge was coded both inbound and outbound, so as to capture through trips, plus assuming that a significant number of the car entrants were discount holders and that some vehicles will make multiple entries within charging hours, the resulting charge might still be expected to be larger than that currently coded. Further explanation would be helpful to describe for each vehicle type the precise method by which the real congestion charge was converted into the charge through which it is represented as a toll on the network.

Public Service Vehicle Pre-Loads

2.1.55 The incidence of bus services on the coded network is represented by mapping bus routes from TRACC[12] as it describes timetabled services within an Ordnance Survey layer. Routes are pre-loaded into the coded network such that the vehicles take up some capacity.

2.1.56 There is no link within the model suite between the Public Service Vehicles (PSVs) and public transport costs so there is no attempt to model assignment of bus patronage or crowding by the NTMv5 suite.

Quality Assurance

2.1.57 A comprehensive series of network checks has been defined along with acceptance criteria which all appear to be reasonable. These checks and the evidence provided in Volume 1 Section 3.12 demonstrate that the network which has been coded is consistent with the corresponding links in the real network.

2.1.58 Volume 1 Table 3-10 which reveals the extent of the coded network in comparison to DfT statistics showing the length of highway network within England, Scotland and Wales. The positive results reported are the similarity in the coded network of the values for motorways, A roads and B roads. The substantial difference is found in unclassified and other roads. The results for England are summarised Table 2-3 below:

Table 2-3: Network Statistics (km)

Classification	NTMv5 (network 18)	DfT 2015+	Relative
A road	34,560	32,315	106.95%
B road	19,556	19,966	97.95%
Motorway	3,325	3,056	108.80%
Other / unidentified	259	247,143	0.10%
Total	57,700	302,480	

+ Source: DfT Road Statistics from Table RDL0201, 2015

2.1.59 As Table 2-3 shows, there is a reasonable correspondence between A roads, B roads and motorways. Discrepancies between the coded network and DfT statistics are likely to be related to definitions of slip roads and links which are common to more than one A road or where single carriageway A roads split to become dual carriageway.

2.1.60 The main link type in which the network is clearly lacking is the “Other / unidentified” roads. Whilst Volume 1 Table 3-9 states that “The network includes all motorways, A Roads and the majority of B Roads in England, with additional minor roads where this has been identified as helpful…”, given that this only includes 259 kilometres throughout England the definition of “helpful” is clearly extremely limited.

2.1.61 The fact that 85% of the highway network is missing from the model does impose significant limitations on where the model can provide insights especially for car-based trips. The extent of this limitation can be estimated by considering Road

[12] https://www.basemap.co.uk/tracc/

Traffic Statistics Table TRA0104[13] reveals that 35% of car and LGV travel measured in miles is undertaken on minor rural and minor urban roads.

2.1.62 Further, it is worth noting that within the routing checks review which has been reported in Volume 1 Appendix B and Volume 3 Appendix B a number of issues which were identified are a function of route planning software using links which are not in the NTMv5 network.

2.2 Highway Assignment Model Initial Runs

Documentation and Scope of Testing

2.2.1 Volume 3 Section 2 describes some initial testing of the HAM which aimed to identify how well the modelling software coped with the scale of the zone system and its associated network, as well as considering the most efficient User Equilibrium assignment algorithm available to achieve convergence and acceptable run times.

Assignment Algorithm

2.2.2 The initial HAM runs concluded that the LUCE[14] assignment algorithm had the most favourable combination of run times, ability to restart model runs from later loops (saving CPU time) and ability to achieve good levels of convergence within the highway model.

Data for Calibration and Validation

2.2.3 Volume 3 Section 4 introduces the observed data which is used for the calibration and validation of the HAM including a discussion of the development of screenlines – and mini-screenlines – and confirmation that these were only used for model calibration. Figure 4.2 shows that the screenlines are distributed throughout the model network, but it is not revealed how many counts there are on individual screenlines. Given the sparse network it is likely that even long screenlines will have relatively few count sites.

2.2.4 Volume 3 Section 4 also provides details of the journey time routes which have been defined for the purpose of model calibration and validation. Whilst there is a statement which aims to provide reassurance that the routes have been identified with the Use Cases in mind there is no indication of which routes will be used to provide performance metrics for which Use Case.

2.2.5 The HAM network was finalised with reviews of assigned traffic which enabled a review of the implications of the coding of centroids, specifically where they join to one-way links. The source of the assigned matrix is not stated in Volume 3.

2.2.6 A final comment in Volume 3 Section 5.6 which reviews route choice indicates an issue which may have ramifications for several use cases. The comment suggests that the model is selecting routes through London and Manchester, avoiding their

---

13 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/801188/tra0104.ods accessed 29/04/20

14 Gentile, G and Noekel, K (2009) Linear User Cost Equilibrium: The New Algorithm For Traffic Assignment In Visum, Association for European Transport, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.678.3796&rep=rep1&type=pdf (accessed 07/04/20)

respective motorway ring roads and concludes this “could indicate issues with the level of highway congestion in urban areas which may require monitoring in future model runs.” As identified in paragraph 2.1.47 this finding should have been investigated more thoroughly as it has implications for all the use cases.

Matrix Estimation

2.2.7 Although the provenance of the matrix used for analysis presented in Volume 3 is undeclared there is discussion of how matrix estimation has been employed to refine and reshape it. The process referred to as the Method of Least Squares was adopted as the alternative method, T-flow fuzzy, could not be implemented in the NTMv5.

2.2.8 The description of the estimation process and the constraints implemented reveals that the matrix was reshaped by considering distance bands and provides an example using LGV home based work trips. Volume 3 Section 6.3.1 speculates that there will be a higher proportion of long-distance trips by LGV for home-based work than short distance trips and provides Figure 6-1 as evidence. Although the values shown in Figure 6-1 sum to 71% there is no clear evidence that there is a dominance of longer distance trips as 26 of the 71 percentage points are showing trips less than 15 miles.

Other Adjustments

2.2.9 Volume 3 Table 6.3 introduces a series of scaling factors which appear to have been applied uniformly to all trips within a purpose and distance band cell. The justification for this correction is that the trip length distributions derived from examining the matrix did not replicate the National Travel Survey trip length distributions. This suggests that that gravity model calibration, reported elsewhere, cannot have been successful.

2.2.10 Further manual adjustments have been reported and recorded in Volume 3 Table 6.6. It is noted that the pre-factoring total is less than the factored total, such that the manual adjustment process is adding trips, but the table reports a reduction in the difference column.

2.2.11 Manual adjustments have also been applied to HGV trips to increase the overall number of vehicles in the assignment matrix for each modelled period. The justification for this is that the supplied matrices are in PCUs which are subsequently factored to vehicles by a uniform value. This process is acknowledged not to have taken into account the different use of the highway network by rigid and articulated goods vehicles and that notes the uniform value is likely to have underestimated HGVs.

2.2.12 A final manual adjustment was made to LGV matrices to improve the match between LGV mileage results from the preliminary assignments and national statistics based on the split between personal and freight trips.

2.2.13 The factoring and adjustments of the HGV and LGV matrices appears to be quite arbitrary. Whilst there is evidence that these adjustments are producing a better outcome from the model in terms of the vehicle kilometres it is not clear whether the model is now producing the right answer for the wrong reasons. Furthermore, it is not clear how these factors will be handled during the forecasting stages of the model development which will have implications for forecast year results and using the model outputs to inform any of the use cases.

2.3 Base Matrices – Zonal Population and Car Trips

Documentation

2.3.1 The creation of the base year prior matrices for cars, LGVs and HGVs is described in Volume 2. These vehicle matrices were subsequently further adjusted based on observed traffic patterns using the procedures documented separately within Volume 3, as reviewed in Section 2.2 above.

2.3.2 To increase the robustness of the forecasts of car demand patterns, Volume 2 Section 2.1 explains that the objective underlying the selection of the approach and of the data sources adopted for building the base year highway matrices was to make them consistent with those used subsequently for forecasting in the VDM. For this reason, "the approach adopted was therefore to generate a full set of total personal trip ends (all modes of travel) for the base year for use in the VDM and from these to derive the Base Year car matrices." This is an appropriate approach to adopt.

2.3.3 It is welcomed that despite the general shortage of good LGV data, the approach endeavours to separate out the personal trips that take place in LGVs from those in cars. This aims to contribute to a suitable foundation for the modelling of LGVs, which are by far the fastest growing component of both interurban and urban traffic.

Zonal Segmented Population and Household Estimates for the Base Year

2.3.4 Volume 2 Section 3 discusses the wide range of data sets that were examined for potential use in the creation of the car base matrices and this includes most potentially useful data sources. It describes Highways England's Trip Information System (TIS) in some detail but it is not made clear subsequently in this or in other Volumes[15] whether or how any use has been made in practice of the highway movement matrices captured within this dataset.

2.3.5 Section 4.2 provides detailed information on how the Any Year Census (AYC) tool was developed to project historic Census population data from 2011 up to the base year 2015 through use of residential property data assembled from the land use registry and Ordnance Survey (OS) mapping. Section 4.2.3 then describes in a well-structured and detailed fashion the complex population processing steps used for this projection procedure. However, the general approach adopted for this projection raises some questions of suitability.

2.3.6 Step 3 allocates the population segmentation characteristics to an MSOA/NTEM zone using only property type specific patterns that are constructed from the average segmentation pattern observed at **district-level**. This approach is likely to introduce substantial and unnecessary local aggregation errors in a situation where the working status, gender and age band characteristics had already been known in 2011 for

[15] It was subsequently made clear to the reviewers that TIS was not used in the model development.

each such MSOA, through the Census file DC6107EW, as well as the MSOA car-ownership profile being available from other Census files.

2.3.7 It is not clear why this observed segmented 2011 population by MSOA was not the starting point for the 2015 estimate of the segmented population by MSOA. As most MSOAs will only have had very limited new dwelling construction in the 4 years to 2015, starting from the known 2011 data and applying a simpler bottom-up approach would avoid introducing unnecessary, probably large aggregation errors for such MSOAs.

2.3.8 Through adjusting for any incremental changes observed in the number of dwellings by property type by MSOA up to 2015, the incremental changes in population segments could be estimated in a form that is likely to be more accurate and substantially simpler than the approach used in the AYC, that is based on the aggregated district level segmentation pattern.

2.3.9 The underlying assumption that differences in population structure and car ownership patterns between the MSOAs of a district can be wholly and accurately associated simply with the differences in the property type mix between MSOAs, seems highly debateable and would require an explicit assessment of the evidence to justify it. At present it is not obvious that this AYC approach adopted to estimate the segmented 2015 population by zone was as effective or more importantly was as accurate as a simpler alternative approach.

Zonal Car-Ownership levels for the Base Year

2.3.10 Section 4.2.3, step 1 states that the "household size and car ownership composition of the 2011 households is updated to 2015 using NTEM profiles". However, equation (4.1) implies that the NTEM 7.2 car-ownership composition is aggregated up from the MSOA level to the district level prior to being applied, which again introduces aggregation error.

2.3.11 In step 1, the underlying car-ownership pattern for 2015 within NTEM 7.2 is derived from applying to the 2015 person totals the 2011-based ownership estimates[16] within the National Car Ownership Model (NATCOP)[17] that have been projected forward from 2011 to 2015. It would have been much safer to control these to match directly to observed car ownership changes through time at the LA level, for the reasons explained below.

2.3.12 This newer version of NATCOP endeavours to improve its previous representation of the lower car ownership rates observed in dense urban areas, by adopting car ownership saturation rates that are differentiated by household type and area type (London, Metropolitan, etc.). However, the major feature that requires to be represented in many of the dense urban areas is that car ownership rates have been **declining** for many years in response to an increase in population density, rather than to income changes. It is important that these car ownership reductions should

[16] Atkins (2017) NTEM Planning Data Version 7.2: Guidance Note
[17] Rand Europe (2016) Estimation of the National Car Ownership Model for Great Britain, 2011 Base https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662879/estimation-national-car-ownership-model.pdf accessed 29/4/20

be included both in the base year and in the model forecasts through to the future. It does not appear that the 2011-based NATCOP has this capability.

2.3.13 These observed declines in car ownership rates should not be ignored because they are not trivial. Combining the mid-year population estimates with the DfT statistics Table VEH0105[18] on registered cars by year by LA, indicates that between 2011 and 2015 the cars per 1000 residents:

- **Increased** by 3% for England as a whole and increased in all regions except London;
- In Outer London, they remained **unchanged** on average, with seven individual boroughs exhibiting a decline;
- In Inner London, they **declined** by -6% on average, with all but one borough exhibiting a decline, including major declines for Islington, Camden and Westminster, in order of increasing declines that range from -9% through to -13%;
- **Declines** are also observed in some other major cities and metropolitan areas, including Birmingham and Leeds.

2.3.14 To represent the national mixture of growth and decline in car ownership levels, the 2015 local car ownership rate should instead be based on applying observed Local Authority level changes between 2011 and 2015 to the 2011 observed MSOA car ownership level[19]. This alternative approach would be likely to lessen aggregation and measurement errors, but without increasing the complexity of generating the base year segmented population and household car ownership patterns that are the foundation for the base year trip productions.

2.3.15 In summary, it seems questionable to adopt the projection approach that disaggregates down from the district level, using only the MSOA property type mix, to generate car ownership rates and other household and person type segmentation details at the MSOA level for the base year 2015. It should require careful analysis and testing to quantify whether the aggregation errors that it generates are acceptably small. Simpler and more accurate approaches would appear to be available. Also, it is highly desirable to ensure that the estimated 2015 pattern of car ownership in dense urban areas has taken appropriate account of well-established trends there of major reductions in ownership rates.

Trip Rate Determination

2.3.16 The use of trip rates consistent with NTEM is appropriate and the use of the Department's CTripEnd tool to estimate these all-mode trip ends through a more disaggregate procedure than NTEM should provide more discriminating estimates at the local scale. Note that the segmented home-based all mode productions which are

18 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/794433/veh0105.ods accessed 29/04/20

19 Care is needed when using the DfT statistics Table VEH0105 on registered cars by year by Local Authority, to identify periods in which large observed changes are observed over a short time due to local upheavals in registrations of company owned rather than of privately owned cars.

the result of applying the trip ends to the population estimates described above will be used to produce the base year demand to which the VDM model will be applied.

2.3.17 The NTEM modal split factors for any particular person type only vary by the eight area types. Within any one area type, the output car modal shares of different zones will only reflect variation in the distribution of person types (and, in particular, car ownership levels). There is therefore a potential advantage that the zonal car trip productions that are used as the foundation for the finalised base year car matrix have been based on modal proportions from the VDM, which will reflect variations in LOS, rather than on the modal proportions provided within NTEM.

2.3.18 The presentation in Volume 2 Figure 4.4, illustrates the large scale of the differences that the VDM has produced for the home-based work trip purpose. In particular, the left-hand side chart that presents the NTEM-based estimates illustrates in dark blue the low car trip proportions estimated for commuting by residents of many of the inner-city zones, both in the metropolitan areas and in dense southern cities, such as Bristol, Bath, Brighton, Oxford, Cambridge, etc. This more realistic pattern is entirely missing from the VDM chart, other than perhaps between Inner and Outer London.

2.3.19 This contrast between the patterns is the reverse of what might have been expected a-priori. Instead it suggests that the VDM based spatial pattern of mode split for HbW is less plausible than that originally in NTEM, in the dense urban areas at least.

Car Trip Matrix Creation

2.3.20 Volume 2 Section 5 documents how the base car matrices were constructed for each trip purpose using the zonal segmented population and the trip rate estimates discussed above.

2.3.21 The estimation of the car commuting matrix for the base year 2015 took as its starting point the observed 2011 Census Journey to Work (JTW) matrix. This is a suitable foundation and the inherent differences between it and the required 2015 base matrix are correctly identified. The step by step procedure used to create the base car commuting matrix is presented in detail in Section 5.1 and it adopts an effective approach. The comparison in Figure 5.1 of the trip length distribution produced for the base year car commuting matrix with that from the NTS shows a good match, other than over the shortest distances where the size of zones will inevitably complicate any matching.

2.3.22 Section 5.2 explains how the School Census dataset has been used to produce the base year education matrices. The switch to use this data source, rather than the initial model-based estimation, is appropriate, as are the further steps adopted to create the base education matrix for cars.

2.3.23 In the absence of suitable observed matrices, the base year car matrices for each of the remaining eight trip purposes were constructed using synthetic models, as documented in Section 5.3. The final set of gravity models used trip end inputs for car drivers for each trip purpose. Section 5.3 states:

> *"The mode shares from an interim base year VDM model run (Run A154) were applied to the total trip productions to obtain improved estimates of*

the daily car driver trip productions, as set out in Section 4.4 and shown in Figure 4.4".

2.3.24 However, Figure 4.4 relates only to the HbW purpose, which is not one of the "other" eight purposes under discussion, so it is ambiguous from Volume 2 Sections 4.4 and 5.3 whether the VDM estimated car driver mode shares by production zone that were used were: those specific to the trip purpose in question, which would be most appropriate; or instead were a common all-purpose mode share, which would be less satisfactory.

2.3.25 These synthetic car matrices were calibrated for each trip purpose to match closely to the generalised travel time[20] distribution from the NTS. For all trip purposes, other than HbHol, Figure 5.3 to Figure 5.10 indicate that the calibrated model appears to match the NTS distance band distribution better than did earlier interim model estimates. The subsequent validation of these matrices checked the resulting distance band distribution versus that in the NTS. Figures 5.3 to 5.10 indicate that a reasonable match is often obtained above 10 miles for individual trip purposes but that shorter distance bands are not well matched for most of the eight purposes, perhaps exacerbated by the size of the model zones. Accordingly, the match in Table 5.15 of the observed and modelled average car trip lengths is within +/- 5% for four of these trip purposes but is overestimated by >40% for both HbShop and HbHol and differs by +13% / -14%, respectively, for the purposes HbRecV. These latter four differences are not reassuring as they are much larger than any sampling error within the NTS data.

2.3.26 In summary, both the general approach that has been taken to the development of the car matrices and the data sources that have been utilised, each appears to be appropriate. For four of the trip purposes (HbPB, HbEB, NHbEB and NHbO) the quality of the resulting base matrices looks reasonable based on the limited validation that has been reported. However, four of the trip purposes (HbShop, HbHol and HbRecV), do not appear to validate well.

LGV Non-freight Trip Matrix Creation

2.3.27 Volume 2 Section 5.5 briefly outlines how the LGV vehicle matrices have been split from within the overall "car" matrices. Although in principle this is a necessary step, some serious doubts arise with respect to the practical details of how it has been implemented.

2.3.28 The implemented approach is based on splits that have been calculated using NTS 2010-2015 data, which differentiates between personal trips completed by cars and vans. As vans comprise 15% of road traffic in 2017, and because most van mileage is not for freight purposes (DfT van surveys 2003-05), the van proportions used in Table 5.17 look very low for HbW (4%), HbEB (8%) and NHbEB (7%), presumably due to NTS under-reporting. It is unclear: whether any validation has been carried out of the estimated total van mileage relative to cars that was derived from the

[20] The generalised travel time was constructed from the NTS reported distance, travel time and speed, using the standard TAG formulation (i.e. including non-fuel costs only for business purposes) together with the TAG cost and value of time parameters appropriate to the travel purpose in question.

approach that was implemented; or whether there may be serious bias within the estimated separation between the car and van matrices.

2.3.29 Le Vine et al (2013)[21] discuss in detail the apparent downward trend from 1995 to 2010 in van traffic that is reported in the NTS. In 1995 the NTS was recording about 35% of the total van traffic from the Road Traffic Estimates (RTEs) but by 2010 this had dropped down to just 16% of total van traffic. This downward NTS trend contradicts the strong upward trend in van traffic (+48%) observed in the RTEs[22] in the same period. They examined the evidence regarding the substantial and increasing under-recording within the NTS of van trips. It is unclear whether or how this issue of downward bias in NTS LGV totals has been addressed within the use of NTS 2010-15 van data to create the NTM LGV matrix and also how it has been reconciled with the evidence of a further 14% increase in RTE total van traffic from 2010 to 2015.

External Trips and Airport and Port Access

2.3.30 The creation of the external car trips between England to/from Wales or Scotland is documented in Section 5.4 and has been implemented in an appropriate manner.

2.3.31 Likewise, the approach to generating the car trips to and from airports and ports that is described in Sections 5.6 and 5.7 appear to be appropriate and to make effective use of the available data.

2.4 Base Matrices – Other Vehicles

Documentation

2.4.1 The creation of the base year freight matrices for HGVs and LGVs is summarised in Volume 2 Section 7, which provides a high level overview of the original matrix development work by MDST[23] and describes the additional processing carried out to create the base year prior trip matrices for input to the matrix calibration procedure.

Road Freight

2.4.2 The data sets that have been used in Section 7.2 for the development of the HGV matrix are those that are most appropriate. The main data source used for HGV matrix was the Continuing Survey of Road Goods Transport (CSRGT)[24] but this is missing some types of movements (e.g. foreign registered vehicles) and its limited sample size precludes its use to provide the fine level of detail of the NTM zoning system. Methods were developed to resolve these shortcomings.

2.4.3 The road freight matrices were segmented between domestic, unitised port and bulk port traffic for commodity types NST 01 to 14[25]. Matrices for these segments were generated synthetically at the NTM zone pair level using gravity models within which

[21] Le Vine, Luan and Polak (2013) "Van travel in Great Britain" http://www.theitc.org.uk/docs/111.pdf
[22] Source DfT Traffic Table TRA0101.
[23] This is documented more fully in the report "HGV & Van Origin-Destination Matrix Documentation, for National Transport Model" (MDS Transmodal, September 2019, which lies outside the material within the scope of this NTM review task.
[24] https://www.gov.uk/government/statistics/continuing-survey-of-road-goods-transport-gb-respondents-section
[25] https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Standard_goods_classification_for_transport_statistics_(NST)

the deterrence function is the transport cost between zone pairs, raised to a power. The use here of the power function, rather than of the more conventional exponential deterrence function, would benefit from supporting evidence to indicate why this power formulation was preferred.

2.4.4 Section 7.2.7 explains that for each commodity type in turn, the synthesised domestic + bulk port NTM – zone matrix was scaled to the relevant CSRGT NUTS 4 – NUTS 4 control totals. If the CSRGT sample size was fewer than 5 records a more aggregate geography of NUTS zones was used. This use of 5 records as the cut-off to avoid HGV sampling errors seems much too low in this context. In part, because the CSRGT samples vehicle not trips, all 5 trips might just be from 1 vehicle, implying a sample of 1 and associated very large sampling errors. The O-D controls by commodity type would need to have been applied at a much more aggregate zone pair scale in order to achieve acceptably low sampling errors.

2.4.5 The road tonnes for each commodity type are split by vehicle type based on CSRGT data for each region-region by commodity grouping, and then are translated to PCUs using PCU factors from TAG. The final matrix was then scaled up to account for known underreporting levels within CSRGT.

2.4.6 The conversion of the annual HGV matrices to the time periods of an average weekday is documented in Section 7.2.10. It appears that average HGV factors may have been used throughout these conversions, whereas it would have been more accurate to differentiate the factors between rigid and articulated vehicles, due to their very different time profiles across the hours of the day and the days of the week.

Rail Freight

2.4.7 Section 7.2.6 explains that rail freight matrices were created from Network Rail data that covers all movements. It is noted that rail freight is not modelled and mentioned here for completeness. The matrices segment between:

- movements from terminal to terminal that do not have road feeder legs;
- movements that have a road feeder leg at the origin end; and
- those with a road feeder at the destination end.

Light Goods Vehicles

2.4.8 Section 7.3 explains how the LGV prior matrices for vehicles carrying freight have been developed. The analysis is based on the DfT van surveys of 2003-05, which had relatively small sample sizes. Furthermore, they are now 15 years old within a rapidly growing and evolving traffic sector. However, in the absence of any subsequent detailed van surveys that provide better data being available, they represent the most suitable data source available at present.

2.4.9 The same issue of needing to avoid major sampling error, as discussed above for HGVs, arises equally with LGVs regarding the need to increase the cut-off from the 5 sampled trips that was used, up to a substantially larger number. It is made even more important by the much smaller sample size available for non-personal van trips than that available within the CSRGT, once it is summed across years. Likewise, the use of a power function form for the deterrence function in equation (7.1), rather than an exponential form, would require some evidence for its justification.

2.4.10 The strategic modelling of multi-stop journeys is a known difficult challenge. Accordingly, the approach adopted for such journeys on LGVs that is based on TrafficMaster data appears to be a useful step forward, despite the assumed biases in that dataset that have been discussed in the report.

2.4.11 Although only limited documentation is provided to cover the finer details of the methods used to produce the freight HGV and LGV prior base year matrices, the high-level overview of these methods suggests that the available data have been used reasonably productively, other than the need to switch to larger minimum sample size cut-offs so as to minimise sampling errors.

2.4.12 However, the underlying survey data available from DfT on the movements of LGVs is now very out of date and has small sample sizes. Accordingly, it would be very valuable to carry out a rerun of the 2003-05 van surveys with minimal changes to the questionnaire, but now with an increased sample size[26]. This would greatly improve understanding of current trends and behaviour in the LGV sector as a whole and by improving the evidence base available it would then provide the foundations for much improved LGV base matrices to be produced in the future.

2.5 Other Mode Costs

Documentation

2.5.1 Detailed information is provided in Volume 4 on how the distance, fare and time matrices were developed for public transport and active modes.

Public Transport

2.5.2 The main source used for data on public transport journey times and distances is the application of the TRACC software to timetable information of public transport services for specific time/day combinations. A variety of well thought through experimental runs of TRACC were carried out in order to ensure that the results produced were suitably configured to meet the specific needs of the inputs for the model, which include the need for a separation between the supply characteristics for the main modes bus and rail. A range of relevant checks were made of intermediate run results in order to select the most productive avenue to progress.

2.5.3 Intrazonal characteristics are not available through TRACC so alternative approaches were required. It was eventually assumed that rail intrazonals were "included for model zones with two or more stations and all zones in Inner London" (Section 2.6.4). While acknowledging the inclusion of LU within the "Rail" main mode, the assumption of intrazonal rail being a universal option for all NTM zones of Inner London appears very questionable in practice. It would have been helpful to check directly whether every Inner London zone was guaranteed to include at least two LU/rail stations and refine this assumption.

2.5.4 The production of the fares matrix for rail used the MOIRA2 system to access the LENNON national database for station to station fares. Separate matrices of fares

[26] We understand that a new van survey bas been carried out, and the results should be available at the end of May or early June 2020

were defined by time period for four broad rail ticket types: Full, Reduced, Season and Advance.

2.5.5 Fares for bus were based on the distance travelled, applying costs per kilometre dependent on the area types: London, Metropolitan or Other, whereas longer journeys used coach fares. Although initially it had been envisaged that the interzonal distances would be those output from TRACC, because only 16.25% of the zone pairs output in the TRACC data had bus/coach as their main mode, the NTMV5 HAM network skim distances for car were used instead. It would have been instructive to have included some analysis of the match between the TRACC bus service-based distances and those by car from the HAM, in order to estimate a transformation rule that should provide a more realistic estimate of the actual bus distances and consequent fare levels.

Active Travel

2.5.6 TRACC was also used to generate the matrices of distances and travel times for the active modes walk and cycle.

2.5.7 Overall, excepting a few minor issues raised above, it appears that the matrices of base year supply characteristics for public transport and active modes have been documented appropriately and assembled and then checked in a thorough and effective fashion.

2.6 Demand Model Specification

Documentation

2.6.1 Detailed information is provided in the supplementary report, “National Transport Model for England: Estimation of the mode-destination models version 13,” and Volume 5 Section 2 and Section 3 about how the demand model has been specified.

Overview

2.6.2 The NTMv5 has a completely new demand model, modelling the choice of mode and destination for a number of different journey purposes. The approach is similar to the urban demand models developed by RAND Europe for West Midlands (PRISM) and London (MoTiON) although there is here less commitment to the notion of tours (as opposed to trips).

2.6.3 As set out in Volume 5 Section 2.2, the criteria for success, in addition to the model fit to the estimation data, were a) the implied values of time by mode, b) other ratios of key parameter values, for example ratios of car in-vehicle time to train and bus in-vehicle time, c) the ability of the model to replicate the observed trip length distributions by mode and purpose, and d) the model elasticities.

2.6.4 The model has the general form of a hierarchical logit structure with mode choice above destination choice, though for some purposes this collapses to a simple multinomial logit (MNL). From this point of view, it follows the general TAG guidance. The possibility of further nests for public transport sub-modes and active sub-modes was investigated but no convincing evidence was found to support them.

2.6.5 The demand (choice) data is taken from the NTS: NTS 2010-15 data is used, confined to English households, with weekend trips excluded. The key level-of-service (LOS) variables and fares/costs are provided as described in Sections 2.1 and 2.5 of this Peer Review. The TAG formulae for vehicle operating cost were applied, but it is unclear what speed was used. The RAND Report "National Transport Model for England: Estimation of the mode-destination models version 13" notes (p 15) that "average speed" is used for the calculation but does not say whether this is O-D or network based. We suspect that the implied average speed for each O-D has been used (derived from the distance and time skims for the highway model): if so, there is a potential inconsistency in the VOC between the demand model and the highway assignment, as discussed above in paragraph 2.1.50.

Purposes and Modes

2.6.6 Since trip ends are provided by the Department's NTEM model, there is a general requirement to align the purposes. Separate models are estimated for:

- Home based Work (HbW)
- Home based Employers Business (HbEB)
- Home based Education (HbEd)
- Home based Shopping and Personal Business (HbShopPB)
- Home based Recreation, Social and Visiting Friends and Relatives (HbRecV)
- Home based Holiday and day trip (HbHol)
- Non-Home based Employers Business (NHbEB)
- Non-Home based Other (NHbO)

2.6.7 While slightly less detailed than NTEM, this covers the range effectively. In line with NTEM, escort travel is merged with related purpose trips: the implications are mainly for the education purpose. Note that the EB definition is not quite compatible with NTS (NTS would classify a Business to Work trip as "Business", while here the purpose is based on the destination).

2.6.8 As noted, this is not a pure tours approach, but neither is it a conventional trips-based approach. The Hb modelling is based on the outward movements only, and on this basis, the purpose is defined by the destination and the mode by the mode on the from-home leg: however, costs are defined for the round trip, assuming "simple" tours, using the most commonly used periods for outward and return legs for the purpose. Thus, HbW and HbEB are assumed to go out in the AM peak and return in the PM peak, HbEd to go out in the AM peak and return in the Interpeak, and the other Hb purposes and the two NHb purposes take place in the interpeak.

2.6.9 As in NTEM, motorcycle drivers are classified as car driver: however, both taxi drivers and passengers are classified as car driver, while in NTEM taxi passengers are treated as car passengers. In addition, domestic air travel is re-defined as rail. While these are potential anomalies, they affect a very small proportion of the data.

Treatment of Destinations

2.6.10 Although the estimation approach follows rigorous statistical methods, there is a critical issue relating to the geographical information for destinations available in

NTS. For Hb trips, it was possible to identify the home wards used for the primary sample units, allowing a reasonable mapping to the NTM zones, but for destinations the data cannot be identified below the Local Authority District (LAD). Since on average there are about **20** NTM zones per LAD, this presents a major challenge for a destination choice model. This also means that for NHb trips, both the origin and destination are only available at LAD level.

2.6.11 Generally, our view is that the documentation makes too light of this issue. The solution adopted for the estimation is to introduce an additional level in the hierarchy, so that the choice of NTM zone is modelled conditional on the choice of LAD. Essentially, however, this is no more than an averaging process. If Z is NTM zone and D is LA District, with A_Z the "size variable" (reflecting the attractiveness of the zone) and the U_{iZm} the "utility" (including among other things the cost and LOS of travelling between i and Z by mode m), we have, at the lowest level:

Equation 3

$$p_{Z|D,im} = \frac{A_Z.\exp(U_{iZm})}{\sum_{Z'\in D} A_{Z'}.\exp(U_{iZ'm})}$$

which yields the composite or "logsum" utility for the district D

Equation 4

$$U_{iDm} = \ln \sum_{Z'\in D} A_{Z'}.\exp(U_{iZ'm})$$

2.6.12 In a true hierarchical logit model, we would expect a "nesting parameter" (≤ 1) to be applied to this composite utility for the choice of LAD. However, there is no data which would allow this to be estimated. Hence this part of the model defaults to an MNL structure:

Equation 5

$$p_{D|im} = \frac{\exp(U_{iDm})}{\sum_{D'} \exp(U_{iD'm})} = \frac{\sum_{Z'\in D} A_{Z'}.\exp(U_{iZ'm})}{\sum_{Z'} A_{Z'}.\exp(U_{iZ'm})}$$

2.6.13 If we write $U_{iZm} = \bar{U}_{iDm} + \Delta_{ZD|im}$, then

Equation 6

$$p_{Z|D,im} = \frac{A_Z.\exp(\bar{U}_{iDm} + \Delta_{ZD|im})}{\sum_{Z'\in D} A_{Z'}.\exp(\bar{U}_{iDm} + \Delta_{Z'D|im})} = \frac{A_Z.\exp(\Delta_{ZD|im})}{\sum_{Z'\in D} A_{Z'}.\exp(\Delta_{Z'D|im})}$$

from which it can be seen that the composite district utility:

Equation 7

$$U_{iDm} = \ln \sum_{Z'\in D} A_{Z'}.\exp(\Delta_{Z'D|im}) + \bar{U}_{iDm}$$

is essentially composed of a weighted size variable + the average utility over the district.

2.6.14 Further, for NHb, where the zone i is only known at LA level, the LOS is simply averaged over the origin zones within the LA.

2.6.15 This must raise some doubt about the ability of the model to reproduce short trips, although it should be noted that the mean trip lengths per mode in fact seem to be very well re-produced. It is perhaps surprising that no account was taken of the reported trip distance in the NTS data, which could have provided further information for shorter trips within the LAD or between adjacent LADs.

Utility Formulation

2.6.16 In comparison to the standard practice of a linear specification for "generalised cost", a rather more detailed formulation is used, building on the substantial experience of RAND Europe in developing these kinds of models. The detail relates principally to a) the formulation of the cost and b) the treatment of drivers and passengers. In what follows, we briefly review these two aspects.

2.6.17 As TAG M2 Section 3.3 notes, "There is strong empirical evidence that the sensitivity of demand responses to changes in generalised cost reduces with increasing trip length", and one of the recommended methods is the "log cost plus linear cost" formulation. In principle, separate coefficients, both of which should be negative, can be estimated for both cost and log cost, but in practice, this can often not be achieved. In such cases, RAND Europe have proposed the formulation: $U(\text{cost}) = \beta_c \left[\gamma.\text{cost} + A.(1-\gamma).\ln(\text{cost})\right]$, where the constant *A* is defined as the ratio of the mean cost to the mean log cost in the estimation sample (this ensures that over the sample, the average contribution is $\beta_c.E\left[\text{cost}\right]$). The "mixing factor" γ was obtained by means of a grid search, and the selected values for the different purposes were in the range 0.1 to 0.6.

2.6.18 The formula has the result that the implied Value of Time varies for each mode and ij combination according to the formula $VOT_{ij,m} = \frac{\beta_{t,m}}{\beta_c} / \left[\gamma + A.(1-\gamma)/\text{cost}_{ij,m}\right]$. To comply with the TAG M2 requirement that "values of time need to be reported and acceptable over all appropriate values of" cost, the formula is invoked to provide values at the mean modal cost for each purpose[27]. While this generally gave slightly higher values for car drivers than the TAG recommendations (with the exception of shopping/personal business which was 30% lower), the overall agreement was reasonable, but for public transport (and rail in particular) the values were very much lower. This is discussed further below.

2.6.19 For the NHb Business purpose, it was not possible to obtain a satisfactory result with the mixing factor, and the TAG relationship with distance was imported directly.

2.6.20 Whereas the car driver and car passenger modes are usually modelled together, with an allowance made for average vehicle occupancy, the RAND formulation separates

[27] though note that in the Quality Report this is referred to as "median" modal cost.

the modes, and relies on the data to decide the allocation of costs, according to a formula based on the average occupancy. The Driver is allocated a proportion $\alpha = 1 - \frac{S(O_{cd} - 1)}{O_{cd}}$ of the vehicle costs, while the passenger is allocated a proportion $\varepsilon = \left\lfloor \frac{S}{O_{cp}} \right\rfloor$. The terms O_{cd} and O_{cp} are the average occupancies for car drivers and car passengers respectively, calculated from the data separately for each purpose. The NTS records the vehicle occupancy for both car drivers and passengers: the mean occupancy O_{cd} must be greater than 1 and O_{cp} must be greater than 2 (since the driver is included). While O_{cd} is a reasonable estimate of average vehicle occupancy, the direct interpretation of O_{pd} is more difficult. It is noted that the driver and passenger cost components do not sum to exactly the total car cost.

2.6.21 The "sharing parameter" S is again decided by means of a grid search procedure. If S = 0, the driver bears all the cost. In fact, with the exception of the Education purpose, where a value of S = 0.25 was used (implying that passengers bear a small share of the cost), a value of S = 1 was selected for all other purposes. In this case, the driver is allocated 1/O_{cd} of the cost and the passenger 1/O_{pd}. The car driver mode was considered available to all individuals over 17, regardless of licence or household car availability. Note that a possible consequence of the formulation is a positive fuel cost elasticity for car passengers (presumably as people switch away from car driver).

2.6.22 Aside from the treatment of cost, separate time coefficients are estimated by mode (car driver and passenger share the same coefficient). An attempt to estimate separate coefficients for public transport components was not successful, so fixed weights were used: 1 for IVT, 2 for walk time and wait time, and 5 (minutes) for interchange. Relative to the time coefficient for walk, the car coefficients tend to be of similar magnitude, and the cycle coefficients somewhat larger, which is not entirely intuitive (though there is some variation by purpose): however, public transport coefficients are significantly lower – between 30-60% of the value for bus, and in most cases even lower for rail. There is an additional distance effect for car passenger for some purposes: while this is negative for HbW and HbEd (as might be expected), it is, oddly, positive for Hb Shopping and both NHb purposes (so that people making long distance shopping trips are more likely to choose car passenger mode than the generalised cost would imply).

2.6.23 A large number of segmentation effects are identified, based on car availability, age, gender and work status. These were sensibly limited to those in NTEM, apart from household income. While income effects on the cost coefficient were found for some purposes, they were dropped in order to keep the level of segmentation down.

2.6.24 The structural tests carried out imply significant nesting (destination below mode) for most purposes - the exceptions are HbW, HbEB and HbHol where a multinomial structure is used. However, the structural coefficients in the main text do not agree with those representing the final coefficients in Appendix B, and for HbShopPB and HbRecV the final coefficients are much closer to 1 than those reported in the main text. No explanation is provided.

Double Constraints

2.6.25 While TAG guidance is to use “doubly constrained” destination choice (distribution) models for Hb Work and Education purposes, for Estimation purposes the RAND Europe approach is not to apply such constraints (this would imply estimating a constant for each destination zone, though, given that the estimation data is an unexpanded sample, this would only ensure that the sample destinations were achieved). In the light of the TAG guidance, some testing was carried out for the Hb Work data. Other investigations have suggested that (NB for aggregate trip data) that in the presence of destination constraints, the coefficient on generalised cost may need to be significantly higher.

2.6.26 While the key variables – car time and rail distance – were slightly, and significantly, higher in the doubly constrained run, the general impression was that the differences were small, and hence it was concluded acceptable to carry out the estimation without the zonal constants. The consultants also suggest that since UK demand model are typically implemented using some form of incremental approach, the base matrices will effectively capture the attraction constraints: however, given that in this model it is the O-D assignment base matrices that are used (see Section 2.8 below), the argument is hardly valid.

2.6.27 We note that the test, with zonal constants being added for each destination, was carried out prior to the PT LOS being received, so distance was used as a proxy for generalised cost in the case of bus and rail modes. Other question marks remain over the suitability of this single test and over the interpretation of its results. Firstly, the two significantly different parameters between the singly and doubly constrained models: car driver time and train distance are two parameters that are central to representing the impedance effect within the distribution model. They are relevant to the longer distance movements that would potentially be influenced by the application of destination constraints. Accordingly, the fact that the differences in the other parameter values are not significant is of little relevance to the spatial distribution model formulation.

2.6.28 Secondly, it is unhelpful that the only segment-based differentiation in the spatial distribution model is between full-time and part-time workers. Moreover, this is applied solely through a part-time worker distance parameter, instead of adopting separate parameters between full- and part-time for all distance and/or all time terms as well as a separate destination size term (and constraint) specific to each segment. Section 13.2.2 of the Quality Report discusses the estimation procedure for the model parameters stating that: the time and cost coefficients "do not vary by demand segment (by definition)"; only the mode specific constants (and car distance in some instances) may vary between demand segments. It is difficult to understand from a behavioural perspective why this restriction has been applied as it will inhibit the performance of the destination choice model. Census journey to work multivariate data analysis indicates that trip lengths differ substantially across at least 3 independent dimensions:

- Gender
- Full-time, part-time
- Industry type (SIC)

2.6.29 The only distribution model tested includes just one of these three dimensions and even that in a very partial fashion, so that without testing a more suitably segmented distribution model design it would be dangerous to derive any general conclusions either about the most appropriate choice hierarchy or about the ability of a singly constrained estimation procedure to replace the need for a more time-consuming doubly constrained estimation procedure.

2.6.30 Similar issues arise with respect to the parameter estimation for the education purpose. Due to differences in the size and degree of specialisation of primary versus secondary schools, as well as due to the difficulties of travel of very young children, the average trip length for primary students will be considerably shorter than that of secondary/sixth form students almost universally across the country. The expectation accordingly would be that a larger deterrence parameter should be necessary for the primary students to enable the spatial distribution models to be realistic for each group of students.

2.6.31 A related problem the education purpose is the absence (Volume 5 Section 4.8.2) of any segmentation between primary and secondary students within the demand model, despite them having systematically different destination constraints, trip length distributions and mode choice preferences. This absence appears particularly odd in a situation where Table 4.12 of Volume 5 indicates that segmentation by Gender [2 classes], Employment type [3], Car availability [2] and Age [4[28]] is applied. It is difficult to believe that in reality every one of these segments in use would be of greater behavioural significance for education trips than splitting between primary and secondary students. It appears likely that a simpler segmentation of primary/secondary/ sixth form/tertiary [4] by car availability [2] should lead to a smaller, faster and behaviourally better model. Such a structure did not however appear to have been explored.

2.6.32 In summary, it is not clear that this set of results provides any sound evidence for the decision to estimate doubly constrained distribution models based only on a singly constrained estimation. Secondly, the paucity of segmentation included in the spatial distribution model for commuting is a serious shortcoming that is likely to substantially reduce the realism of the model results and may lead to an inappropriate choice hierarchy being adopted. Likewise, the formulation adopted for the education model appears to be far from appropriate.

2.6.33 Aside from a fully constrained estimation, some more generic destination constants were considered, but in practice they are confined to Inner London to help reproduce the significant differences in the mode share between Inner London and other destinations.

Assessment of Model

2.6.34 As noted in Section 2.6.3, four criteria for judging the model were stated: values of time by mode, other ratios of key parameter values, replication of trip length

[28] The four age groups distinguished are: 0-15, 16-29, 30-64, 65+, so no segmentation between primary and secondary children within the 0-15 age group appears to have been considered from the outset.

distributions by mode and purpose, and model elasticities. We briefly assess these in turn.

2.6.35 The implied values of time (at mean cost) by mode are shown in Table 2-4 compared with the average (all mode) perceived values of time provided in the TAG databook. The values for NHbEB are not presented, as these make direct use of the TAG formula.

Table 2-4: Implied values of time (£/hr) at median trip cost by mode (2015 prices and values)

Trip purpose	TAG A1.3.1	Car driver	Rail	Bus
HbW	11.47	10.50	3.24	4.63
HbEB	18.66	21.83	4.77	16.66
HbEd	5.23	6.58	2.49	1.78
HbShopPB	5.23	3.63	1.68	1.75
HbRecV	5.23	5.27	1.11	2.26
HbHol	5.23	8.26	1.92	3.59
NHbO	5.23	7.39	1.49	2.97

Source: TAG databook A1.3.1 (version 1.9.1 December 2017)

2.6.36 While the car driver values appear acceptable, both in terms of level and variation by purpose, the public transport values are significantly lower, and while bus values are typically found to be low (perhaps reflecting lower average incomes of users), the even lower values for rail are highly questionable[29]. The main reason for this appears to be the low estimated coefficients on rail generalised time compared with car, and it is suggested this reflects the longer distances associated with rail travel. However, the higher mean costs should compensate for this. This requires further investigation. It would be useful also to show the variation in Value of Time with cost, as well as that based on the mean values.

2.6.37 In respect of other ratios of parameter values, there is little to be said, since it was not possible to identify separate coefficients for the various components of public transport: for these modes, “generalised time” was used, adding to in-vehicle time walk and wait times multiplied by 2, and 5 minutes per interchange, in line with ranges in TAG Unit M3.2. Of some interest are the relative values of the modal time coefficients (already alluded to in the previous paragraph). These are presented in the following table, based on Table 13.2 of the Quality Report:

[29] In Table 7.18 of the Phase 2 Report "Provision of Market Research for Value of Travel Time Savings and Reliability", the "All distance" estimates (£/hr) for car and rail are as follows: Commute [Car 11.70, Rail 12.42], Other non-work [Car 4.91, Rail 8.68], EB [Car 16.74, Rail 27.64], though "rail" excludes Underground/metro.

Table 2-5: Time (/generalised time) parameters by purpose

Mode	HbW	HbEd	HbShopPB	HbRecV	HbHol	HbEB	NHbEB	NHbO
Car Driver and Passenger	-0.0379	-0.0821	-0.0506	-0.0438	-0.0220	-0.0204	-3.4510	-0.1171
Cycle	-0.0447	-0.0839	-0.0910	-0.0650	-0.0390	-0.0441	-0.0630	-0.0978
Walk	-0.0349	-0.0441	-0.0466	-0.0417	-0.0169	-0.0339	-0.1127	-0.0952
Bus (gen. time)	-0.0151	-0.0133	-0.0137	-0.0136	-0.0082	-0.0199	-3.4510	-0.0312
Rail (gen. time)	-0.0105	-0.0194	-0.0136	-0.0067	-0.0043	-0.0033	-3.4510	-0.0163
Relative values								
Car Driver and Passenger	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*
Cycle	*1.179*	*1.022*	*1.798*	*1.484*	*1.773*	*2.162*	*0.018*	*0.835*
Walk	*0.921*	*0.537*	*0.921*	*0.952*	*0.768*	*1.662*	*0.033*	*0.813*
Bus (gen. time)	*0.398*	*0.162*	*0.271*	*0.311*	*0.373*	*0.975*	*1.000*	*0.266*
Rail (gen. time)	*0.277*	*0.236*	*0.269*	*0.153*	*0.195*	*0.162*	*1.000*	*0.139*

Source: Quality Report Table 13.2

2.6.38 It appears that the driver and passenger parameters were constrained to be the same, and the very large motorised mode coefficient for NHbEB requires some explanation. The relative values for walk are generally lower than might be expected.

2.6.39 With regard to the trip length distributions, it is notable that nothing is presented, either in the Mode-Destination estimation report, or anywhere else. This seems a significant omission, even though the mean predicted tour distances (using highway network data) show good agreement with the reported distances in the estimation data set.

2.6.40 The elasticities implied by the estimated model (NB these do not take account of supply-side effects, as would be required for TAG realism tests) are generally convincing, in terms of relative magnitudes by purpose and mode. Interestingly, while the Km- and trip-elasticities are very different for car (the main effect being destination choice), they are much more similar for PT which is indeed in line with the empirical evidence.

2.6.41 Overall, the most important criteria are probably the elasticities and the car driver values of time, and on this basis the model appears to be fit for purpose. Nonetheless, questions remain in relation to: the zonal aggregation, the doubly constrained purposes, and the non-car values of time, particularly those for rail.

2.7 Demand Model Implementation

Documentation

2.7.1 The implementation of the demand model is described in in Volume 5 Section 4.

Discussion

2.7.2 Given the estimated model, the next consideration is how to implement it. As noted in Volume 5 Section 4.7.1, the parameters fall into three main categories:

- alternative (mode and destination) specific constants which vary by purpose and in some cases by demand strata (traveller type segments);
- attraction size terms; and
- coefficients to the input cost, time and distance skims (essentially, the "generalised cost" formulation)

2.7.3 While the "generalised cost" formulation is quite complex, it does not present any particular problems for implementation (some minor adjustments are required to meet the VISUM specification for hierarchical logit). The same is generally true of the attraction size terms, where the only issue is to allow the terms to vary in some cases with the person-type characteristics (this is the case with HbEd, where the destination constraint is applied separately for persons under 16 and over 16). For those under 16, the attraction is based on numbers of primary and secondary, while for those over 16, the attraction is based on all education (including tertiary education).

2.7.4 The rationale comes from the Mode/Destination Estimation report (Section 5.2.1) where we find:

> *"The Hb education purpose includes both tours made by school pupils and students, tours made by adults escorting school pupils and, to a much lesser extent, students. Thus, a significant fraction of tours made to primary and secondary education locations are made by adults on school escort tours. Therefore, the primary and secondary enrolments attraction variables are applied to individuals of all ages and not just to the age range within which children go to that type of school. However, the tertiary education enrolment variable is only applied to individuals aged 16 and over on the basis that it is highly unlikely that someone aged under 16 would escort an older person to a tertiary education location."*

2.7.5 The significant problem is how to deal with the segmentation variables. In combination with the set of purposes, the Mode/Destination Estimation report identified 316 "demand strata", and earlier run-time tests had indicated that in order to meet the specified criteria ["to ensure the full model runtime remains within the target of 48 hours for 4 demand-supply loops (3 iterations)"], something of the order of 100 would be desirable. In addition, there were issues of consistency between Hb and NHb purposes, given that the latter are derived from the former. However, after allowing for some aggregation here, this only reduced the number to 312. It was therefore decided to remove the "gender" segmentation, and this was done by obtaining the gender proportions for each zone and using these to calculate a zone-specific an average value of the gender constant (applied to males in the estimation).

2.7.6 The resulting 156 combinations could be further reduced because some of them were not feasible (e.g. children in employment): note that for this purpose it was assumed that there were "no full-time employees or students either under the age of 16 or over the age of 65." It seems questionable that there are no full-time employees over 65. With this reduction, the final number of "demand strata" was 121.

2.7.7 Careful testing was carried out to ensure that the implementation was compatible with the model estimation, in terms of replicating the mode share and trip-length distributions.

2.7.8 For the home-based purposes, the implementation applied to the input productions which are derived by factoring up the base values by NTEM growth (though, as we discuss in relation to Sensitivity Test no. 1, there were some issues relating to incompatibility of the upper age group definition). For non-home-based purposes, however, the "productions" are derived from the number of trips (by various Hb purposes) destinating in each zone. In contrast to most comparable RAND Europe models, there is no set of specifically estimated NHb "frequency" models which would generate the NHb productions.

2.7.9 According to Volume 5 Section 4.8.3,

> *"Parameters were initially taken from NTEM, and then updated using the same methodology for NTS (2010-2015) data but differentiating business trip rates for full time and other employees. This derives trip rates for each Hb purpose into the corresponding NHb purpose, which are then summed to give a NHbEB and NHbO trip rate for each Hb purpose."*

2.7.10 From this it can be seen that the NHb trip rates do not vary by age or car availability. However, these segmentations are available from the Hb models.

2.7.11 The process is described as follows:

> *"This is implemented in Visum by taking the total demand matrix for each demand strata (summed across all modes) and multiplying it by the corresponding trip rates. Age band factors are used to calculate the number of trips that are within each NHb segment for HbShopPB, HbRecV and HbHol, where the age bands in these purposes are more aggregate than the equivalent NHb segment (as shown earlier in Table 4.5). These factors are calculated from the disaggregate Hb trip productions, and sum to 1 for each purpose. For example, HbEd uses the age band 30-64, which is more aggregate than the 30-44 and 45-64 bands used for NHbEB. Hence, the proportion of HbEd 30-64 productions that are in the 30-44 and 45-64 age bands are calculated. These factors are applied to the matrices that require splitting, and the contributions from each Hb purpose are summed to give one matrix per NHb segment. The column sums of these matrices are then taken and are used as the productions for the NHb trips."*

2.7.12 This is not completely clear (and the subsequent equations do not clarify). The NHb trip rates vary only by Hb purpose, apart from the "full time and other employees" for NHbEB, so the "demand strata" for the trip rate (notated as γ in the equations) are considerably more aggregate than the demand matrices. Essentially, it seems that for each Hb purpose, the matrices by demand stratum are multiplied by the (scalar) NHb trip rate (separately for NHbEB and NHbO, and respecting the "full time and other employees" for NHbEB). Where the Hb purposes do not provide sufficient segmentation for the NHb, further factoring is applied to the matrix **rows**, based on the "disaggregate Hb trip productions". Having thus converted to the required demand strata, the column sums are aggregated across the Hb purposes, separately for NHbEB and NHbO, to provide the required NHb productions.

2.7.13 The output from the demand model is confined to internal-internal, internal-external and external-internal trips.

2.7.14 Some indication is given of the ensuing run times. The T1 sensitivity test (2030 demand) run [A246], including trip end growth, 3 iterations (4 passes) of the VDM and HAMs, results exports and final HAMs with tighter convergence criteria, took 59hrs 36mins. While this is "slightly in excess of those specified" [48 hours…], it has apparently been agreed with the DfT that this is "not an issue."

2.8 Preparing Matrices for Assignment

Documentation

2.8.1 The preparation of matrices for input into the HAM is described in Volume 5 Section 5.3.

Discussion

2.8.2 While in demand terms the VDM deals with all modes, with Hb movements on a P/A basis, supply effects (capacity) in NTM only relate to highway. As noted in Section 1.2, the highway assignment operates on a time-period basis, and the segmentation is different from that in the VDM. A more or less mechanical process is required to convert the VDM matrices into the form required for the assignment.

2.8.3 As is well set out in §5.3 of Volume 5, there are a number of steps in this process. Since the volume of vehicles can be aligned with the Car Driver mode, the first step is to aggregate the detailed segments for this mode so that only the purposes are retained. Then, applying factors developed from NTS, the trips are factored to allocate a proportion to the LGV user class: these factors vary by purpose and distance.

2.8.4 The resulting matrices of car trips by Hb purpose on a P/A basis now need to be converted to time period O-D matrices. This is done by means of the NTEM methodology using the rho (ρ) and phi (φ) factors. The ρ factors provide the proportions of outbound trips in each cross classification of mode and time of day, conditional on purpose and area type. The φ factors give the proportion of trips with home outbound purpose and outbound period which return with purpose home in time period. In fact, it has been assumed that the return trip will be the same purpose as the outbound trip, although the NTEM factors show that this is not always the case. The return trip matrices need to be transposed to convert them to an O-D basis.

2.8.5 For NHb trips there is a corresponding set of ρ factors, but as these trips are already on an O-D basis, the φ factors are not required.

2.8.6 The final step is to aggregate over purposes into the required User Classes and convert to hourly demand. All this is in line with best practice.

2.9 Pivoting

Documentation

2.9.1 The pivoting processes employed in the model are described in Volume 5 Section 5.2 and Section 5.4.

Discussion

2.9.2 TAG Guidance M2 Section 4.3 states "The Department's recommendation for scheme appraisal is to use an incremental form of model, whether pivot-point or based on incremental application of absolute estimates, unless there are strong reasons for not doing so." In Volume 5 Section 5.2, these are referred to as IPP (incremental pivot point) and AMAI (absolute models applied incrementally) respectively, and a justification is made for the chosen AMAI method.

2.9.3 However, a crucial distinction is being ignored here: IPP methods relate only to demand models, which must be on a P/A or tour basis – they cannot be applied to O-D matrices. AMAI methods can be applied to both, but there are significant reservations when applying to O-D matrices, as here. A careful reading of TAG M2 indicates that "incremental modelling" is conceived in the context of demand (though the **additional** possibility of "incremental assignment" is also noted). In fact, it takes some time before the reader of Volume 5 becomes aware that for NTMv5 AMAI is applied on an O-D basis. This makes the discussion of the relative merits of IPP and AMAI essentially irrelevant.

2.9.4 While the practice of AMAI allows for special cases relating to small numbers or zeros, the essence of the approach for any matrix cell is as follows:

2.9.5 Given a reliable "base" value B, the "pivoted" forecast is derived as $F = B.\frac{S_f}{S_b}$, where

S_f and S_b are respectively the model-based estimates for the forecast and base scenarios.

2.9.6 In this case, the model-based estimates refer to the period-specific highway O-D matrices, whose derivation was described in the previous section. No discussion is provided as to how close the synthetic base matrices are to the base matrices B, either on a P/A or O-D basis. This is a significant defect.

2.9.7 The available literature relating to AMAI (see, in particular, Daly[30] *et al*, 2012) notes that the aggregate effects of this correction can lead to very different forecasts of growth between the model and the outcome, including the possibility of different signs. It is therefore recommended that normalisation should always be applied when pivoting. The chosen method is as follows:

initial pivot $$F_{ij} = \frac{S_{ij}^f}{S_{ij}^b}.B_{ij}$$

[30] A Daly, J Fox, B Patruni, F Milthorpe (2012), Pivoting in Travel Demand Models, Australasian Transport Research Forum Proceedings, Perth, Australia. Publication website: http://www.patrec.org/atrf.aspx

adjustment factor $a_{ij} = \frac{\sum_{ij \in R} S^f_{ij}}{\sum_{ij \in R} S^b_{ij}} \cdot \frac{\sum_{ij \in R} F_{ij}}{\sum_{ij \in R} B_{ij}} \forall ij \in R$

where R is a "sector to sector" movement

final $F^n_{ij} = a_{ij} F_{ij}$

2.9.8 The sectors are defined as the 133 NUTS3 (January 2015) areas for England, with 5 further areas defined for Wales and Scotland. The normalisation ensures that the sector to sector growth in the pivoted matrix is in agreement with the model-based forecast growth.

2.9.9 Given the potential problems with the pivoting, various checks were carried out, and the results illustrated do not indicate significant problems. Nonetheless, such tests should ideally be done based on significant (rather than quasi-uniform) changes to the matrices, and it is not clear that this has been done. More testing is probably required to provide assurance about the normalisation.

2.9.10 More significantly, the AMAI process on a P/A basis can be normalised in various ways, but primarily to control to total productions (or possibly, productions by mode) which can be considered as being more reliable. However, once we pivot on an O-D basis, we are essentially pivoting using quantities of the type $S^{b,OD}_{i'j'mt} = S^b_{ijm}.\pi^{d=1}_{tmij} + S^b_{jim}.\pi^{d=2}_{tmji}$, where t is time period, and π represent the time period factors (separately by direction d), and there is no longer any particular reliability to these quantities, as they are subject to the considerable potential inaccuracies of the distribution model, and hence no obvious basis for normalisation.

2.9.11 Further, because **multiple** adjustment ratios (by direction and by time period) are being applied to the same underlying P/A demand element, there is no obvious way in which the adjustments can be conveyed to the demand model. This in turn means that the adjustment will not have any effect on model elements that are not subject to pivoting. It is similarly unclear what role normalisation can play in this case.

2.9.12 The importance of this will depend on whether the model is primarily a highway model or is intended to represent all modes. If there is reasonable agreement between B and S^b, the proposed pivoting should deliver sensible results for the highway mode. But the results for other modes are subject to much greater uncertainty.

2.9.13 It is unclear whether any part of the other vehicle matrices is subject to pivoting: in principle, this might be applied to the LGV portion which is factored out of the car driver matrices.[31]

2.9.14 The approach to matrix development which has been adopted introduces a disconnect between the highway demand assignment matrices and the demand model P/A matrix. For highway demand this is acceptable and has been allowed for in the matrix conversion process. However, the absence of a pivoting approach for

[31] We have subsequently been informed that pivoting **is** applied to LGV Personal trips.

other modes limits the NTM's value, and effectively means that it is primarily a highway assignment model.

2.10 Iteration and Convergence

Documentation

2.10.1 Various approaches to damping the supply-demand iterations were considered, and a volume-averaging approach was chosen, applied to the highway assignment matrices (by time period) after pivoting. The step length factor is generally set to 0.5, though an option is allowed for the standard MSA (1/n).

2.10.2 The impact of the resulting cost change is illustrated for the changes in total car driver trips between iterations from the VDM for the 8 NTM purposes. From this it appears that the choice between the two step-length alternatives is more or less even, though sensibly it is decided to fix at 0.5 because of its superior performance in later iterations.

2.10.3 It is noted in Section 9.4 of the Quality Report that the supply-demand convergence "process is stopped after the third iteration (fourth pass)." Figure 5.5 of Volume 5 illustrates for the fixed step length option adopted, that changes of 0.003% approximately are still occurring in car driver trips after the third iteration for some trip purposes. Though it is not explicit that this measure of change is calculated as the net national total difference, assuming that this is the case, it would not be a very discriminating measure so that this low numerical value achieved may not necessarily be reassuring. An alternative measure that identifies the scale of the differences within that subset of zone pairs that exhibit the largest changes between iterations would provide more useful guidance. When using the model to test policies, there needs to be a guarantee that the model is fully converged in all individual areas, rather than simply ensuring that local patterns of increases and decreases cancel out overall at the national scale.

2.10.4 The illustration in Figure 5.5 appears to relate to the realism tests, where small perturbations to costs are typically made. It would be helpful to discuss the level of convergence more thoroughly, and in the presence of larger changes in demand relative to highway capacity.

2.10.5 The level of convergence for the supply-demand loop is not discussed more generally in Volume 5, nor how it might be measured. The only reference to this we found was in the Quality Report.

2.10.6 The Quality Report Section 9 discusses in a single paragraph the approach to convergence which has been adopted within the model development confirming that the model has not been run to full convergence but stopped after the third iteration. Whilst adopting a stopping value limits the run times for the model it is neither clear from the documentation whether the model has stabilised by that point nor the scale of model "noise" which remains in the model. Because of such uncertainty the Quality Report appropriately recommends that scenario-specific checks are conducted on key elements of the model, making comparisons of final and penultimate iterations to check that convergence is satisfactory for specific scenario tests. Unfortunately, there is no evidence provided in Volume 6 (except for the fuel cost realism test, Section 2.2.1) that measures the level of convergence actually achieved for any of the other test scenarios.

2.10.7 The Visum user guide also discusses a "Nested Demand Calculation" as a measure of convergence which appears to be what is referred to in Section 9.4 (though it seems not to be used). This is described as a "cost-weighted relative deviation of current demand from the demand of the last iteration", but unlike the TAG M2 Section 6.3, it does not seem to allow for volume averaging between iterations. The online guide notes that "A prerequisite for gap calculation is that the values of utilities in the Nested Demand procedure are smaller than or equal to zero." Note that Visum generally uses "utility" to denote the negative of generalised cost. It is not clear how this could be reconciled with the NTMv5 model formulation.

2.10.8 We agree that the recommendation in the Quality Report on the need for convergence checks should be followed rigorously and universally. The form of the measurements used to determine the level of convergence achieved should be selected to be comprehensive and incisive for all types of policy test situation. Furthermore, the resulting level of convergence achieved should be published in detail in a standard form for each test run carried out.

2.10.9 A number of queries have emerged related to possible lack of convergence in the sensitivity tests which are discussed in greater depth below in Section 3.5 to Section 3.8 and Section 6.2 below.

3 Model Validation, Realism and Base Year Sensitivity Tests

3.1 Synthetic Base Matrix

3.1.1 The synthetic base year matrix (Sb) output from the VDM is the foundation for the pivoting exercise used in all policy tests. However, its creation appears not to have been described in any detail, other than within a footnote in Section 3.4 of the Quality Report:

> *"The operation of the model differs in that the base runs used for the base year model set-up and highway calibration do not include any iteration of the demand model and HAM: a single pass of the demand model is used to create synthetic base year demand, with a separate and single assignment of the base year vehicle matrices to produce costs."*

3.1.2 The Quality Report Section 12.2 provides a description of tests of model outputs which will be used to confirm whether the demand model validates. There are three outputs which are reported:

- Mode share and trip length distributions;
- Comparison of estimated value of time; and
- Sensitivity of demand to changes in cost and time of travel (elasticities).

3.2 Highway Assignment Model Calibration and Validation

3.2.1 The Quality Report Section 12 sets out standards against which the performance of the HAM can be tested but starts with a discussion about the implications of failing to meet those standards in Section 12.3. This identifies that failure to achieve the threshold considered to be the minimum acceptable standard for a model needs to be further considered against three more tests:

- Whether validation is sufficient to meet the Use Cases;
- The extent to which the conditions have not been met;
- Whether attempts to implement improvements would be worthwhile.

3.2.2 The first test could be defined as the "fit for purpose test." As is commonly understood by modellers it does not necessarily hold that a TAG compliant model is adequate for all purposes. Similarly, a model which has some limitations is often good enough to inform decision making.

3.2.3 The second test aims to explore the scale and direction of the failure to determine how the noted failure might impact upon how the model can be used. As is often the case, models can fail to meet flow criteria in locations which have limited relationship to the locality in which an intervention is planned.

3.2.4 The third test should be clear that worthwhile is being defined in terms whether improvements offer value for money as inevitably improvements cost time and money to explore and implement. Clearly, reviewing the model against test 1 and test 2 will determine what needs to be done to improve the model. The third test is whether implementing changes can be justified in terms of cost.

3.2.5 Section 12 then continues to discuss some relaxations of the TAG criteria which are described as the acceptability criteria for the NTMv5. The challenge this now presents is that if the NTMv5 fails to reach its own specific criteria it becomes more difficult to implement the three tests described above in paragraph 3.2.1.

3.2.6 For example, using values taken from Table 12.2, if the NTMv5 were to show 72% of links meet the TAG flow criteria this would usually mean that further work to improve the model would be required but for NTMv5 this 'score' is deemed acceptable or further analysis of the scale and pattern of failures would need to be undertaken to demonstrate the model is fit for purpose.

3.2.7 Highway model calibration and validation is discussed in Section 14. The flows are shown to pass the relaxed calibration thresholds in Table 14.3 which is positive. However, as no data has been provided either by screenline, mini-screenline or individual site it has not been possible to determine if there are any areas of concern other than London, which is identified in Section 14.7.

3.2.8 Route choice calibration is reported to be satisfactory in Section 14.8 with an important observation about trips using the highway network through London or through Manchester. Essentially, the comment indicates that the model is adopting the shorter distance routes in these urban areas due to "issues with the level of highway congestion." This "issue" is a function of the network coding which results in no congestion in urban areas. As this has an impact on route choice and subsequently summaries of all travel metrics in the model it presents significant challenges interpreting results within urban areas. The aggregation errors discussed above in paragraph 2.1.49, which result from the simplified application of vehicle speeds within VOCs may also be a contributory cause of these routings through congested dense urban centres.

3.2.9 Section 14 discusses convergence and validation within the HAM and the measures which have been reported are %GAP, defined as a "proximity measure" in Section 12.7, which considers how much variation in total travel generalised cost there is between consecutive iterations there is, and flow stability on links[32].

3.2.10 As VISUM does not directly measure the flow stability an examination of the link flows has been undertaken outside of the software itself. This has used two measures, the relative and absolute difference, to demonstrate that links which would have failed the relative measure – the metric in guidance – have low flow differences in absolute terms – generally less than 30 vehicles. The argument is that although the model fails against the relative measure the absolute differences are small so the model should be considered to be converged.

3.2.11 Whilst the absolute differences are small it is not clear why the GEH statistic has not been applied as this is recommended in TAG Unit M3.1. It is unlikely to show that the model has not converged but the test which has been undertaken is not recommended in TAG and no explanation for not following TAG has been provided.

3.2.12 Flow validation results are presented in Section 14.11.2 and indicate that the model has a poor fit against independent observations. As there are only 137 datapoints in the validation dataset it is difficult to draw a conclusion from the evidence presented.

[32] Note that this is different from the measure required to assess the convergence of the demand-supply loop, though in Appendix A2 of the Quality Report it is wrongly described as such..

There would be merit in including more data in the validation dataset and it is unclear how the 137 locations were chosen. Given the dataset is so small the subsequent analysis by region does appear to be redundant.

3.2.13 The final piece of validation data presented is a comparison between observed vehicle kilometres and modelled vehicle kilometres. A description of the calculation of the observed values is provided stating they have been derived from the "DfT's AADF database [https://www.dft.gov.uk/traffic-counts/download.php, downloaded 14th March 2018] to give the three dimensions of region, vehicle type and road type – and the recalculated totals were checked against the published table 'TRA0204_(2015)_Veh_Roadtype'."

3.2.14 Table 14.15 does show that the match between the model and observed vehicle kilometres is not very good and the commentary in the Quality Report suggests a range of contributory factors which mean that finding an accurate observed value is challenging meaning that drawing a conclusion is not possible.

3.2.15 Overall, Section 14 does not present a compelling case in favour of accepting the model as well validated. The implication for the DfT is that the model has not been proven to represent traffic volume observations against which it has been tested. The case is presented that there are in some cases insufficient observations to draw reliable conclusions and in other cases the observations are deemed to be unreliable in themselves.

3.3 Base Year Run

3.3.1 As the highway assignment model makes use of a pivoting process for the matrices to be assigned, it is necessary for the general operation of the model to derive one of the pivoting components – the base year synthetic matrix (Sb). In fact, there is virtually no description of how this was done: the only reference we have found is in footnote 3 in Chapter 3 of the Quality Report.

3.3.2 Our assumption is that the matrix was created by taking the highway LOS from the best validated highway model (as discussed in the previous section), inputting these and the base year costs for other modes to the demand model, and applying the mode and destination shares to the zonal segmented trips for 2015 discussed in Section 2.3 above. It is noted in the Quality Report Section 3.4 that "an exercise was carried out to calculate trip productions and attractions by zone, separated by trip purpose and person type." We have not found any more detailed description.

3.3.3 In principle, we would expect some validation of the resulting synthetic matrix, to see whether the zonal application has retained the general properties of the estimation sample in respect of the trip length distribution and overall mode shares. Furthermore, given the critical role that this element plays in the pivot, a comparison of the implied period-specific O-D assignment matrices with those derived from the highway model calibration would seem essential. We are surprised at the lack of documentation on this issue.

3.3.4 In Section 10.2 of the Quality Report, the criteria for model success are re-stated (as in 2.6.3 above) and can now be assessed in the light of the application of the estimated model to the zonal segmented demand. While the parameter values (and hence the value of time etc.) are not affected, both the model elasticities and the trip length distribution are. We discuss the elasticities in the next section under the heading of Realism tests, but here we review the validation presented in the Quality Report section 13.4 for mode shares and trip lengths.

3.3.5 We have noted that there was no discussion of the trip length distribution in the Mode Destination Report (nor, for that matter, of the outcome mode shares, although given the inclusion of modal constants, one would expect the model results to replicate the estimation sample in this respect). Volume 5 presents the criteria for success, and while it does not provide any results, says that they are set out in Volume 6. However, while other validation information can be found there, the trip length distributions are not in Volume 6, and can only be found in the Quality Report.

3.3.6 The summary validation of modelled versus observed (NTS) mode shares and modal trip lengths in Tables 13.4 to 13.7 indicates a reasonably good match in mode shares for many of the home-based trip purposes but with some larger mismatches for the non-home-based purposes[33]. However, rail trip lengths for most trip purposes exhibit large differences (-33% to +50%) from the observed values, while the overall trip lengths are within +/-10% only for three of the purposes.

3.3.7 The Quality Report notes that "For non-home-based trips the challenges with the level of spatial detail in the estimated model has led to slightly larger differences for some modes, driven in part by differences for the very short (primarily walk) trips." Given that for NHbOther the walk mode share has increased by 11% compared with the estimated model and the associated sample, this seems like an understatement: the increase is at the expense of both car driver and passenger. Indeed, the application of the NHb procedure raises a number of questions, as we shall see.

3.3.8 While, as noted earlier, the estimated model provided an acceptable fit to the mean trip lengths, in application for the base year run car driver and passenger trip lengths are substantially lower than the sample-based values for HbEB, NHbEB and NHbOther purposes. The mean modal trip lengths for the model as applied to the estimation sample correspond with minor discrepancies to those reported in the Mode/Destination Estimation Report, if the home-based purpose values are divided by 2 to convert from tours to trips. However, the "observed" values, based on NTS show significant discrepancies, particularly for the walk mode, even though in both sources it is claimed that the "reported" distance is used. As a further possible source of inconsistency, the mean modal trip lengths for the base run use "mode-specific" distances, while those for the model as applied to the estimation sample use highway distance throughout.

3.3.9 In terms of trip length distributions, results are presented for base year run as compared with the NTS sample (as noted earlier, no corresponding information is available on the fit of the model to the estimation sample). The critical data is in Appendix B of the Quality Report, and reveals considerable discrepancies, with the worst cases relating to the NHb trips. Because the lack of corresponding data from the estimation report, it is not possible to judge whether these discrepancies emanate

[33] Though it is unclear why the mode shares for the model as applied to the estimation sample do not correspond exactly to the observed.

primarily from the model estimation or have been brought about by the application of the model to the base case. More discussion might have been expected.

3.4 Realism Tests

3.4.1 In paragraph 2.6.349, we briefly noted that the elasticities implied by the estimated model appeared reasonable, but were not true realism tests, as they omitted any supply effects. They were estimated using the estimation sample, perturbing the relevant variables by 10% and noting the implied change as predicted by the model.

3.4.2 In this section we assess the true realism tests, with the model being run using the total zonal demand and (except in the case of the car times elasticity) iterated to convergence between demand and supply to take into account the congestion (or removal of congestion) feedback. These are reported in Volume 6 Section 2, where both sets of elasticities can usefully be compared.

Fuel Cost Realism Test

3.4.3 The fuel cost realism test quantifies the change in vehicle kilometres as a result of an increase of 10% in fuel costs. Volume 6 Section 2.2.1 says that the increase in costs within the NTMv5 has been implemented by factoring the a, b, c and d values from TAG Unit by 1.1 in both the demand model and the assignment model.

3.4.4 Section 2.2.1 explains the TAG thresholds within which the model can be taken as performing in a realistic manner and includes an additional criterion regarding the car passenger elasticities. For this mode, the proposal is that the car passenger elasticity should be higher (i.e. less negative) than car driver on the basis that there will be an expected increase in car occupancy as fuel costs increase. Note that this implies that the car passenger elasticity could be positive in some cases.

3.4.5 Results presented in Table 2.1 show that the supply effect results in lower car driver elasticities than were found for the estimation sample, but they are generally plausible, and within TAG guidelines. The small positive elasticity for the HbEB is perhaps surprising, but not impossible given the reduction in congestion and the higher value of time for this purpose. Elasticities for Car passenger are in all cases less negative than for Car driver, and have positive values for three purposes (HbW, HbEB and NHbEB): these purposes have the lowest occupancy. Our judgment is that these are acceptable results.

3.4.6 Table 2.2 shows the change in the number of trips from the converged model. Taking the HbEB result for car drivers there is a change of -0.3% in trips. In combination with the increase in vehicle kilometres the outcome for HbEB from the model is that short distance trips decline more rapidly than long distance trips and those trips are getting longer.

3.4.7 It is also possible to inspect the outcome O-D matrices after pivoting, where the evidence for HbEB trips indicates that the pattern is not uniform across modelled periods, as shown in Table 2.6. There is an increase in business vehicle kilometres in the commuter peaks and a reduction in the interpeak period. This again seems plausible, since the reduction in congestion associated with fewer car trips by other purposes will be greater in peak periods. However, on an O-D basis the period-specific elasticities for the other purposes appear more negative than the overall values from the converged demand model.

3.4.8 Further inspection of the demand and distance matrices for the VDM and the HAM is reported in Section 2.2.3 with analysis of the distribution of trips and vehicle kilometres presented in Table 2.7 to Table 2.9.

3.4.9 The analysis of the evidence concludes that there are discrepancies between the VDM and HAM due to the pivoting process which become exaggerated especially in the zone pairs which sit in the >100 miles distance bands.

3.4.10 The source of this discrepancy lies not in the pivoting procedure itself, since the "normalisation" process ensures that the sector-to-sector growth from the VDM is correctly translated to the pivot matrices, but rather is caused by the major differences in the trip distribution pattern beyond 50 miles between the synthetic base matrix and the Base matrix or the NTS, as illustrated in Tables 2.7 to 2.9. For example, for car driver trips in the PM peak, the Base matrix contains 2.6% trips (23.5% trip kilometres) over 50 miles, whereas the synthetic base matrix contains only 1.0% trips (7.3% trip kilometres) for such trip. Accordingly, more than two thirds of the long-distance car kilometres are missing from the synthetic matrix. Similar issues arise for the PM peak for the Car Other user class within the illustration of the pivoting procedure in Volume 5, Figure 5.1. Because the usage of the NTM is oriented more to interurban rather than to urban traffic, it is very important that it should be able to represent longer distance trips realistically. Pivoting does not provide reasonable results in situations where the Base and synthetic base matrices have fundamentally different patterns.

PT Fare Realism Test

3.4.11 The PT fare realism test seeks to determine how demand on public transport responds to an increase of 10% in fares.

3.4.12 The impact on total trips is presented in Table 2.14 and demonstrates that each purpose except Home based Shopping and Personal Business (HbShopPB) returns a value within the guidance threshold. Compared with the elasticities implied by the model estimation, the values have in most cases become less negative. The HbShopPB purpose does appear to be highly sensitive, with an elasticity value of -1.11.

3.4.13 Complementary results indicate that for the HbShopPB elasticity this high sensitivity applies to both bus and rail. There is no further analysis of these values and it is unclear whether the model development team were content with this result or whether work to consider this outcome further would have been beneficial.

Journey Time Realism Test

3.4.14 Section 2.4 describes the journey time realism test, which aims to identify the response to an increase of 10% on highway journey times. The manner in which this test is implemented, simply factoring the highway costs which are submitted to the demand model by 1.1 means that it is not possible to iterate between demand and supply, so the test is a single loop of the demand model, as stipulated in TAG.

3.4.15 As the test is relatively blunt the target elasticity threshold is relatively wide with guidance saying that car driver trips should have a response of less than -2.0. Again, an additional criterion has been included regarding the car passenger elasticities: they should be negative.

3.4.16 The results presented in Table 2.21 and Table 2.22 show that elasticities at the whole model level are comfortably within the threshold and the overall impact on trips across modes is logical. Interestingly, the elasticities are generally stronger (more negative) than those implied by the estimation sample: it is claimed that this is due to subsequent refinements made in the HAM to the skims provided for VDM estimation, as well as to differences between the sample of trips and the full set of trip ends.

3.4.17 Table 2.23 introduces a noteworthy result for Home based Work (HbW) car passengers which have a positive elasticity value in the segment referred to as "full car." This demand segment includes households where there are as many cars as adults. The result is peculiar in that it means that as journey times increase there are more trip makers in this segment. As there is no other data in the report to isolate this group it is not possible to comment on whether this impact is specific to a region or general across the model.

3.4.18 The result appears to contradict the additional requirement which the model developers decided to impose on the car passenger elasticities. Nonetheless, there is no further discussion of these results nor comment on whether further work is recommended to understand.

3.5 Sensitivity Test 2 – Highway Capacity

Objectives

3.5.1 The objectives of Test 2 are to demonstrate that highway network changes can be coded into the model, that the model operates with those changes within it and to explore how responsive the model is to the defined changes.

Brief and Specification of the Test

3.5.2 An entirely fictitious scheme was coded which provides a 3-lane motorway standard highway route between Bournemouth / Southampton via Bath and Market Drayton to Stoke-on-Trent / Chester.

3.5.3 The route is a combination of new coding and recoding of existing network, which includes disconnecting and reconnecting some secondary roads and providing grade separated junctions at A road and motorway crossings on the route.

3.5.4 The route has sections in four of the regions as shown in Table 3-1 below.

Table 3-1: Route Length for Test 2 Highway Capacity

Region	**Route km**
South East	13.72
South West	220.73
West Midlands	126.22
North West	48.64
Total	409.31

Source: Volume 6 Table 4.1

3.5.5 Volume 6 Section 4.5.1 discusses changes in person trips by region with outputs from the NTMv5 presented in a series of tables showing trips by different modes. The total values for the model are summarised in Table 3-2 below.

Table 3-2: 24hr Productions All Modes

Mode	Trips Base	Trips Test	Difference	Difference %
Car driver	31,212,273	31,209,240	-3,033	-0.01%
Car passenger	13,789,033	13,790,286	1,253	0.01%
Bus	5,097,160	5,097,727	567	0.01%
Rail	2,533,026	2,534,291	1,265	0.05%
Walk	17,218,473	17,218,327	-146	0.00%
Cycle	1,179,733	1,179,825	92	0.01%
Total	71,029,698	71,029,696	-2	0.00%

Source: Volume 6 Table 4.3 to Table 4.8

3.5.6 The first observation to make is that the changes are exceedingly small and would fall well within model noise thresholds. This means that it is not sensible to draw a conclusion about the model performance at a national level from the results shown.

3.5.7 The second observation is that the decrease in car driver trips is counterintuitive as the intervention being modelled increases highway capacity. The increase in rail patronage is similarly a counterintuitive result given the scheme.

3.5.8 An increase in car passengers appears to be a sensible outcome of providing more highway and potentially an increase in bus use could be anticipated if a bus service had been added to the new highway network.

3.5.9 Inspecting Table 4-3 more closely for the car driver mode the decrease is dominated by -5,785 car driver trips in London. There is also a change of -1,189 trips in the South East of England. It is not clear why these decreases would occur in these locations nor is there any explanation offered except a statement to say that the London result is worthy of further investigation.

3.5.10 Table 4-3 reveals increases in car driver trips of +3,419 in the South West and +2,051 in the West Midlands. These results are more logical albeit very small in relative terms.

3.5.11 Results reported for trip lengths, journey times and speeds, and trip length distribution indicate very little change between the base and the test scenario.

3.5.12 Section 4.5.4 discusses the changes in total vehicle kilometres by region due to the introduction of the test highway capacity.

3.5.13 The tables report to two significant places for both million vehicle kilometres and the percentage change from the base. Unfortunately, the tables do not provide the base values, so these have been calculated from other values in the tables.

3.5.14 Doing this exposes some gaps that should be completed to enable a fuller understanding of how traffic is reassigning both in parts of the network close to the scheme and those areas remote from the scheme.

3.5.15 For example, we consider the South West and England results from Table 4.13, which are repeated in Table 3-3 below for ease of reference.

Table 3-3: Difference in vehicle kilometres

Region	Total Absolute difference Million vkm	Absolute difference	Motorway Absolute difference Million vkm	Relative difference	A Road Absolute difference Million vkm	Relative difference	B Road and Other Absolute difference Million vkm	Relative difference
SW	0.47	6.0%	0.71	44.4%	-0.20	-5.00%	-0.03	-2.8%
Eng	0.78	1.2%	1.20	7.9%	-0.35	-1.00%	-0.07	-0.9%

Source: Volume 6 Table 4.13

3.5.16 From this information we can see that 0.47 million vehicle kilometres produces a relative increase of 6.00%. We can infer that the base value is approximately 7.83 million vehicle kilometres in the morning peak hour and the "with scheme" value is approximately 7.39 million vehicle kilometres.

3.5.17 Repeating the same calculation for each highway type and summing the base and "with scheme" values produces approximately 6.67 million vehicle kilometres and approximately 6.42 million vehicle kilometres respectively. The difference between the total value and sum of the road types is approximately 1.16 million vehicle kilometres and 0.97 million vehicle kilometres respectively or approximately 14% of the distance.

3.5.18 A similar calculation for the England and, where possible other regions of England as well as in Wales and Scotland reveal a similar discrepancy between the base and "with scheme" values.

3.5.19 As the links in the networks must all fall into one of the four categories an explanation about the missing vehicle kilometres is required. Without the missing information it is not possible to fully appreciate the implications of the reassignment between the different highway types.

3.5.20 Notwithstanding the missing information, the results for all vehicles show that the majority of change in vehicle kilometres is in the regions with the scheme with very minor changes further from the scheme.

3.5.21 Reviewing the results between periods does again reveal some peculiar results, but as the information provided in the tables is very limited it is not possible to draw conclusions with regard to the performance of the model..

3.5.22 A series of plots of Great Britain are provided to illustrate changes in trip ends due to the introduction of the scheme. As the bulk of zones fall into the "-50 to +50" category where there are larger changes commentary would assist to confirm the changes are where they would be expected.

3.5.23 The changes in vehicle flows in the network which are presented in Section 4.5.6 demonstrate that the scheme and its feeder links experience the largest changes in flow but that there are links remote from the scheme which also experience some changes. This is model noise and common in large models, however, as noted in the text the reductions appear to be happening in the urban areas – which coincide with the urban area speeds treatment in the coding.

3.5.24 The combination of questions about the changes to the volume of trips and the changes in the amount of travel as well as routing changes suggests that the approach to network coding in the urban should have been investigated further during the model development.

3.5.25 One surprising result which is not shown in any of the flow plots is a reduction on the M6 passing through Birmingham. The new route provides an alternative motorway standard connection west of Birmingham which should be more attractive to trips to or from South Wales, the West of England and the South West. As there is no change in flows on the M6 the inference is that these trips must be a relatively small proportion of trips on the M6.

3.5.26 An omission from the analysis of traffic flows is an actual record of the volume of trips using the scheme. The only evidence presented is the relative change from the base which is insufficient information for understanding how busy the new route will be.

3.5.27 Section 4.5.8 provides insights into the route choice between a range of origins and destinations which appear to be candidate trips for the new infrastructure. Below are some observations on the information presented.

- There is clear evidence that HGVs are very sensitive, perhaps overly, to distance. This is most clearly demonstrated in Figure 4.20 and Figure 4.21 where trips between Stoke-on-Trent and the South Coast choose to route through or close to Birmingham and avoid the new route. Southampton to Telford in particular shows light vehicles choosing to route via motorways, whereas heavies, as expected are taking a more direct route" is quite questionable. As can be seen from transport statistics, long distance HGVs will divert to travel on motorways or high quality roads so as to avoid the high rates of fuel consumption experienced on lower quality more direct roads that are not conducive to travelling at a constant speed; and
- There do appear to be some coding issues in the network, as shown in Figure 4.25 where light vehicles are looping once around M4 J13 for no apparent reason. Similarly, Figure 4.26 shows trips looping near Gloucester.

3.5.28 It is not clear from the evidence how many trips are making a journey between the named origins and destinations. It is expected the overall number of trips using the route for the full length will be very low.

3.5.29 The evidence does not consider how the route is impacting on shorter distance trips which is unfortunate as it is likely that these will be the much greater proportion of traffic on any one link on the new route.

3.5.30 Section 4.5.9 shares data on travel time changes between sectors with Table 4.19 indicating flow weighted changes for car trips. There are some large absolute changes, but without information on the number of trips nor the relative change in travel time it is not possible to infer the impact on demand. Given the extremely small changes in demand reported earlier in Section 4 the scheme does not appear to have much impact.

3.5.31 HGV travel times have not been reported. This is unfortunate as in combination with the route choice above this evidence would have shown more clearly how HGVs balance time and distance in the model.

3.5.32 Section 4.5.10 provides a summary and concludes that new infrastructure can successfully be incorporated into the highway network, which appears correct, and that outputs from the model are "largely intuitive." Whilst there are some intuitively correct results there are counterintuitive results which are dismissed as model noise.

3.5.33 The issue which this dismissal causes is that the counterintuitive results are often larger in absolute terms than those which are being accepted as correct, for example the decrease in car drivers in London of 5,785 is dismissed as noise as London is remote from the scheme, whereas the increase in car drivers of 2,051 in the West Midlands is accepted as being a sensible result. So, a reasonable conclusion might be that the scale of the infrastructure introduced to the model is insufficient to create results which are outside of model noise or that the results are being selectively reported as noise or impact.

3.5.34 The conclusion of this review is that any results from modelling of a highway scheme will need to be treated with caution as even large infrastructure interventions cannot escape the thresholds of model noise.

3.6 Sensitivity Test 3 – Public Transport Changes

Objectives and Specification

3.6.1 Here the objective was to test the functionality for modifying the exogenously specified public transport attributes and to explore the responsiveness of the model to changes in public transport supply.

3.6.2 The approach agreed with the DfT was to downgrade rail services between the northernmost regions in the model (NE, NW and Y&H) and London (in both directions). The test halved the frequency of these services and increased their rail travel times by 30 minutes. The frequency reduction was represented through increasing the initial wait times plus the interchange wait times.

3.6.3 The changes apply only for trips between the region pairs in question, as opposed to those that might have to travel into London and out again, e.g. Yorkshire to Essex. There are no changes either to fares or to access and egress times.

3.6.4 The same relative frequency changes were applied across all time periods and hence to all trip purposes.

Discussion of Results

3.6.5 Volume 6 Table 5.11 presents the change in rail travel times between selected sub-region pairs and confirms that the supply change in the test has been implemented correctly. It shows increases of around 60 minutes where envisaged in the test and times remain largely unchanged otherwise.

3.6.6 The impact of the test is summarised for each individual mode in the Report Tables 5.1 to 5.6 that illustrate the regional difference in 24hr trip productions between Test 3 and the base run. These absolute differences are summarised in Table 3-4 for the six modes. It indicates that the total change in trip productions over the whole study area is a fall by 1 trip (in a total of over 71 million) so this is understandable as a (marginal) rounding error. At the regional level, the differences are more noticeable. However, our analysis of the tables disaggregated by trip purpose provided within the Appendix E indicates that regional level differences in total trip productions occur only in the non-home-based purposes and are negligible for the home-based purposes, as should be expected. The reduction of -1,900 trip productions in London is a combined result of -797 fewer NHbEB and -1103 fewer NHbO productions there.

Table 3-4: Regional difference in 24hr trip productions from Test 3 to base run, by mode

Region	Car driver	Car pass.	Bus	Rail	Walk	Cycle	Total
NE	207	35	70	-353	115	28	102
NW	-226	-527	510	-930	1,126	171	124
Y&H	1,512	263	209	-1,950	358	97	489
EM	-70	-9	28	176	29	18	172
WM	-641	-347	229	240	486	92	59
EoE	-276	-106	108	690	284	56	756
Lon	-5,271	-1,942	1,482	1,161	2,077	593	-1,900
SE	-678	-232	18	990	125	42	265
SW	-677	-130	259	181	199	95	-73
IWa	-281	-139	108	59	212	19	-22
EWa	-74	-7	4	98	0	0	21
Sc	0	0	0	6	0	0	6
Total	-6,476	-3,140	3,024	369	5,012	1,210	-1

Source: Volume 6 Table 5.1 to Table 5.6

3.6.7 This summary table raises a number of queries about the results obtained from this rail supply change test.

3.6.8 The reduction in rail trip productions for each of the three northern regions is plausible.

3.6.9 However, this test produces a minimal 0.01% **increase** in overall rail share, rather than a decrease in share which is what should be expected from a reduction in the quality of rail supply, with no changes to any other modes.

3.6.10 London in particular exhibits an increase in rail trips of 1,161, rather than the expected decrease of trips heading north. This result is not explained.

3.6.11 There is similarly no explanation as to why London shows a large reduction in car trips of -7,213 (much larger than any changes in rail trips) which is balanced by gains in bus, walk and cycle trips.

3.6.12 Other unexplained results include car driver and passenger trips **decreasing** by 0.02%, with these reductions being focussed throughout Wales, the West Midlands and the South of England, not just London.

Table 3-5: Rail P/A trips; Test 3-Base (Absolute difference)

Region	HbW	HbEd	HbShopPB	HbRecV	HbHol	HbEB	NHbEB	NHbO	Total
NE	-100	14	0	0	-10	-81	-179	2	-354
NW	-464	-34	16	13	-50	-83	-342	13	-931
Y&H	-1,205	7	0	-2	-75	-231	-455	11	-1,950
EM	152	2	-1	0	2	17	3	2	177
WM	167	4	6	10	7	36	5	4	239
EoE	597	49	3	16	28	-30	20	7	690
Lon	883	-795	47	153	73	769	46	-15	1,161
SE	440	379	5	12	14	113	24	2	989
SW	143	7	0	1	0	28	4	-2	181
IWa	36	1	5	6	1	6	2	2	59
EWa	70	0	1	3	1	21	1	0	97
Sc	0	0	0	0	0	0	5	1	6
Total	720	-365	82	211	-8	566	-866	29	369

Source: Volume 6 Table E.11

3.6.13 Within Table 3-5 the analysis of the absolute change in rail P/A trips due to Test 3 indicates differences between trip purposes in their patterns of response. For a number of purposes, the rail trips do reduce for the three Northern regions, but they change little for the other trip purposes that typically have relatively short trips, all of which is not unreasonable.

3.6.14 For London in contrast, rail trips **increase** rather than decrease for most trip purposes, which is not the expected response from a rail travel time increase. Only HbEd has a reduction in rail trips for London, with little corresponding change in the North but it also has an increase in the South East region that off-sets half of London's reduction, despite the South East experiencing no change in its rail supply characteristics.

3.6.15 Only the HbEd and NHbEB purposes exhibit a non-trivial overall reduction in total rail trips, whereas HbW, HbShopPB, HbRecV and HbEB each exhibit overall increases in total rail trips, despite the decline in rail supply characteristics. These contradictory patterns of response are difficult to understand.

3.6.16 Volume 6 Figure 5.10 and Figure 5.11 that map the change in rail trip productions and attractions, respectively, are not very informative because they lack a neutral band that would remove the convergence related noise pertaining to small changes on either side of zero. Nonetheless, while rail productions have marginally fallen almost throughout the three Northern regions (except around Manchester), they have unexpectedly marginally increased throughout almost all of the rest of the country, including most of London.

3.6.17 Within the model hierarchy because mode choice is above destination choice for a number of segments, one would expect that the first response would be to shift the affected movements to shorter/less expensive distances, and this is might be why London is getting more rail trips (though this would still be counter-intuitive). Even then one would still expect that the reduction in rail utility via the logsum would result in some small shift away from rail overall.

3.6.18 It seems really strange that in a number of regions (London, WM & SW in Table 3-4) the absolute impact on car is greater than that on rail. The effect on highway speeds of the switch from rail should really be second order, so the additional impact on demand ought to be more or less negligible.

3.6.19 Volume 6 Table 5.13 presents the percentage changes in 24-hour rail trips between selected sub-region pairs and these at first appear plausible, being within the range of -50% to -70% for those pairs with rail time increases and being around +1% or less elsewhere. However, the associated Table 5.12 of absolute differences, reproduced here in Table 3-6, tells a less convincing tale. The absolute growth in rail trips travelling just between Inner and Outer London is on a scale similar to the interregional rail reductions elsewhere due to the test!

Table 3-6: Changes in 24hr trips between selected sectors; Test 3 - Base, Rail

	Nbria+Tees	**Cbria+Lancs**	**Manchester**	**Cshire+Mers**	**NEYrk+NLinc**	**SYorks**	**WYorks**	**InnerLdn**	**OuterLdn**	**Grand Total**
Nbria+Tees	68	29	1	1	1	1	3	-460	-8	**-364**
Cbria+Lancs	-11	76	18	5	-1	1	0	-397	-15	**-325**
Manchester	0	13	95	7	0	-0	-3	-192	-8	**-86**
Cshire+Mers	0	9	38	72	0	0	2	-731	-14	**-624**
NEYrk+NLinc	11	4	7	2	30	7	26	-639	-11	**-562**
SYorks	2	2	7	1	23	61	40	-1,114	-10	**-987**
WYorks	2	11	20	2	8	0	65	-570	-9	**-471**
InnerLdn	-5	-1,030	-5	-10	-5	-5	-8	577	101	**-391**
OuterLdn	-3	-162	-7	-8	-3	-3	-5	1,451	141	**1,403**
Grand Total	**65**	**-1,047**	**174**	**72**	**53**	**62**	**121**	**-2,075**	**168**	**-2,407**

Source: Volume 6 Table 5.12

3.6.20 Volume 6 Table 5.16 indicates that the rail trip productions from the NW to London switch in considerable percentages to SE +29% and EoE +21% regions, while those **to** the NW increase by +70% from SE and +58% from EoE. In contrast, there is only a reasonably small shift to SE and EoE for either NE or Y&H, despite their similar reductions in trips to/from London to those for the NW.

3.6.21 Based on these many results that run counter to broad expectations for this specific type of rail test, it is not possible to accept the assertion in Section 5.3.1 that "the significant changes to trips are exclusive to public transport modes".

3.6.22 Instead, the pattern of results appears more similar to that which would be obtained from a pair of model runs where the noise due to lack of model convergence was greater than the signal from the test itself. For this reason, it is particularly unhelpful that information is not provided on the level of convergence actually achieved within the four iterations that were run. Perhaps an alternative rail sensitivity test that was designed to lead to larger scale rail changes might have produced more straightforward and logical responses, particularly if it was certain that a high level of model convergence had been achieved. Other more fundamental explanations for these perverse elasticities could be practical issues with the operation of the choice hierarchy that has been adopted.

3.6.23 In summary, these results provide no confidence in the overall responsiveness of the model within this test of the impacts of rail supply changes. They emphatically do not support the summary statement in Section 5.3.8:

> *"The impact and results of the test confirms that NTMv5 responds in the expected manner to changes in PT characteristics, with changes in rail trips restricted to those regions directly impacted by the input changes, and that these have a proportionate impact on rail trips in the model."*

3.7 Sensitivity Test 4 – Highway Cost Changes

Objectives and Specification

3.7.1 The objective of this test was to explore how distance-based road user charging could be implemented in the model in specific geographical areas and for particular link types and investigate the plausibility of the results from such a test.

3.7.2 A 10 pence per kilometre charge was applied to all vehicles in all modelled periods on all dual carriageway links which are not part of the Strategic Road Network [SRN] in the South East, East of England, South West and Wales.

3.7.3 The subject network is summarised in Volume 6 Table 6.1 which is repeated below Table 3-7 for ease of reference.

Table 3-7: Length of Network with Charges Applied

Region	**Links**	**KM**	**Proportion of model network**
SE	2,592	1,522.62	8.30%
EoE	1,719	1,223.71	8.50%
SW	1,002	670.82	3.40%
Wales	104	113.56	2.50%
Total Charge Regions	5,417	3,530.71	6.20%
Total Model			2.80%

Source: Volume 6 Table 6.1

3.7.4 Checking these statistics against Volume 1 Table 3.10 it should be noted that the per cent values relate to the total network length not just the A road network length.

3.7.5 Section 6 Figure 6.1 indicates the links which have charges applied to them but as it includes the Humber Bridge, Mersey Tunnel, Birmingham Toll Motorway and the Severn Crossings the plot is not “new” tolls but all tolls.

Discussion of Results

3.7.6 The impact of the test on all trips during the day is the sum of Volume 6 Table 6.2 to Volume 6 Table 6.7 and presented in Table 3-8 below.

Table 3-8: Trip Productions Base versus Test 4

Mode	Trips Base	Trips Test	Difference	Difference %
Car driver	31,212,273	31,052,768	-159,505	-0.51%
Car passenger	13,789,033	13,722,145	-66,888	-0.49%
Bus	5,097,160	5,147,557	50,397	0.99%
Rail	2,533,026	2,554,930	21,904	0.86%
Walk	17,218,473	17,360,458	141,985	0.82%
Cycle	1,179,733	1,191,839	12,106	1.03%
Total	71,029,698	71,029,697	-1	0.00%

Source: Volume 6 Table 6.2 to Table 6.7

3.7.7 These results appear to be reasonable, with car trips falling and non-car modes all increasing. It might be argued that some switch from driver to passenger is likely as passengers may share the costs with drivers, so that the proportional decrease in passengers would be smaller than that for drivers: this is only marginally the case..

3.7.8 Table 3-9 considers the impacts by region and mode in more detail.

Table 3-9: Regional proportions of Total mode changes from Test 4 vs Base

Region	**Total**	**Car Driver**	**%**	**Car Pass**	**%**	**Bus**	**%**	**Rail**	**%**	**Walk**	**%**	**Cycle**	**%**
NE	1	-116	0.07%	-18	0.03%	48	0.10%	17	0.08%	53	0.04%	17	0.14%
NW	-542	-1,432	0.90%	-649	0.97%	390	0.77%	308	1.41%	708	0.50%	133	1.10%
Y&H	102	-20	0.01%	23	-0.03%	28	0.06%	46	0.21%	-3	0.00%	28	0.23%
EM	1,132	-110	0.07%	-35	0.05%	275	0.55%	192	0.88%	736	0.52%	74	0.61%
WM	345	-732	0.46%	-365	0.55%	323	0.64%	265	1.21%	733	0.52%	121	1.00%
EoE	-2,649	-53,453	33.51%	-22,075	33.00%	15,802	31.36%	7,773	35.49%	45,699	32.19%	3,605	29.78%
Lon	4,816	-6,386	4.00%	-3,293	4.92%	3,422	6.79%	1,283	5.86%	8,464	5.96%	1,326	10.95%
SE	-3,650	-77,941	48.86%	-33,370	49.89%	23,322	46.28%	9,466	43.22%	69,535	48.97%	5,338	44.09%
SW	49	-17,634	11.06%	-6,780	10.14%	6,208	12.32%	1,751	7.99%	15,109	10.64%	1,395	11.52%
Sc+W	395	-1,681	1.05%	-326	0.49%	579	1.15%	803	3.67%	951	0.67%	69	0.57%
	-1	-159,505	100.00%	-66,888	100.00%	50,397	100.00%	21,904	100.00%	141,985	100.00%	12,106	100.00%

Source: Volume 6 Table 6.2 to Table 6.7

3.7.9 The first result of note is the apparent redistribution of trips shown in the Total column. Taken together, the sum of positive values for London, the West Midlands and East Midlands (+6,293 trips) counters the negative values in South East and East of England (-6,299 trips).

3.7.10 We have not been provided with a breakdown by purpose for this test, but, in line with what was found for the PT test, we assume that these changes in total regional productions are due to NHb purposes. While these changes are generally small, in the case of London they are of the same order of magnitude as the change in car driver and passenger trips, as was also seen in the PT test.

3.7.11 The second set of results to consider more closely is the mode choice response. In Table 3-9 there are 111,311 fewer car trips for the South East (the sum of car driver and car passenger). There is an increase in non-car trips of 107,661 trips with the balance of trips redistributing.

3.7.12 The result to query in the mode choice model is that there is an increase of 69,535 walk trips which represents 64.59% of trips that have switched mode. Similar results are observed for the East of England and South West.

3.7.13 A further query about mode choice is that all other regions also return a result showing that walk is the dominant alternative mode to car-based modes except in Yorkshire & Humberside where rail is favoured.

3.7.14 Reviewing the statistics for average trip lengths, which are presented by purpose at the national level as opposed to by region, there does appear to be a tension between the outcome of the mode choice model and the input trips.

3.7.15 Car driver and car passenger average trips lengths in the base are both greater than 12 kilometres and the change due to the Test being implemented is a minor reduction in all cases. In contrast, the walk mode in the base is averaging 1.49 kilometres with no change due to the Test being implemented.

3.7.16 The immediate tension is that car-based trips appear to be substantially longer than a reasonable walk trip distance suggesting that the outcome in terms of total trips by mode should be questioned.

3.7.17 The trip length distribution does not shed any more light on this issue as it is limited to car driver trips. The results are showing a move towards shorter trips. The combination of redistribution and increase in walk is consistent with the demand model hierarchy but does introduce further questions of whether the results are realistic especially for purposes such as HbW and HbEd.

3.7.18 Other results reported in Section 6 all appear to be consistent with there being fewer car trips in the network with fewer vehicle kilometres in most regions, especially those with the tolls charged, and flows falling in the regions with the tolls imposed. The V/C results are meaningless as there is no comparison to the base. The changes in average speed are similarly difficult to discern as the scale in Figure 6.13 to Figure 6.15 show a per cent value change which is probably the least informative way of presenting speed change.

3.7.19 Section 6 should have also included some analysis of LGVs and HGVs separately, as these vehicles are likely to have responded differently to the car-based trips on the basis that the demand is fixed but routes will have varied. Additional exploration by purpose within the regions affected by charging would have also improved the understanding of how well the model is operating.

3.7.20 Section 6.3.9 summarises the model run as being successful with intuitive results across all metrics.

3.7.21 Whilst this might be a reasonable assertion for car-driver trips and overall car travel in the network there do appear to be responses where alternative modes are too attractive. Similarly, some of the spatial responses appear to pose questions about how the model is redistributing trips.

3.8 Sensitivity Test 5 – Urban Area Strategy

Objectives and Specification

3.8.1 The objective of this test is defined in Section 7.1 as a need to demonstrate that urban policies can be tested in the model for a nominated group of urban areas. The urban areas selected for the test are those in the North East, Yorkshire & Humberside and East Midlands regions which have a resident population greater than 25,000 people.

3.8.2 The tests applied included reducing speeds on the urban network to 20 mph for single carriageway roads, reducing speeds on the urban network to 50 mph for dual carriageway roads (except where the base speed was already below 50 mph), adding 2 minutes to access and egress time for car trips, increasing the proportion of trip makers paying for parking and reducing the perceived travel time of travelling by bicycle.

3.8.3 The text in Section 7.3 indicates that parking charges for NTMv5 are taken from NTMv2R and have been applied to all NTMv5 zones within the zone classes from NTMv2R. There is no indication what the values of the parking charges are nor what proportion of trip makers are paying for parking within the description of the test.

3.8.4 This is a serious omission from the earlier reporting and makes it very difficult to determine whether any changes have been implemented correctly or whether their impacts are reasonable.

Discussion of Results

3.8.5 A summary of mode shares by region is provided in Table 3-10 below based on Volume 6 Table 7.2 to Table 7.7.

Table 3-10: Regional proportions of Total mode changes from Test 4 vs Base

Region	Total	Car Driver	%	Car Pass	%	Bus	%	Rail	%	Walk	%	Cycle	%
NE	-691	-121,226	20.56%	-72,507	21.15%	31,530	21.65%	7,952	15.71%	82,479	21.86%	71,081	19.80%
NW	3,435	-2,124	0.36%	-2,053	0.60%	1,669	1.15%	1,564	3.09%	3,617	0.96%	762	0.21%
Y&H	-7,359	-286,770	48.63%	-166,730	48.64%	69,757	47.91%	25,304	50.00%	170,990	45.32%	180,090	50.16%
EM	-3,395	-173,644	29.44%	-97,666	28.49%	38,874	26.70%	11,555	22.83%	111,980	29.68%	105,506	29.39%
WM	4,027	-283	0.05%	-1,188	0.35%	1,066	0.73%	1,014	2.00%	3,004	0.80%	414	0.12%
EoE	2,867	405	-0.07%	-343	0.10%	533	0.37%	524	1.04%	1,549	0.41%	199	0.06%
Lon	-387	-5,451	0.92%	-1,893	0.55%	1,513	1.04%	2,244	4.43%	2,492	0.66%	708	0.20%
SE	1,305	71	-0.01%	-195	0.06%	307	0.21%	195	0.39%	800	0.21%	127	0.04%
SW	47	-492	0.08%	-131	0.04%	260	0.18%	70	0.14%	242	0.06%	98	0.03%
Sc+W	152	-217	0.04%	-102	0.03%	94	0.06%	191	0.38%	168	0.04%	18	0.01%
	1	-589,731	100.00%	-342,808	100.00%	145,603	100.00%	50,613	100.00%	377,321	100.00%	359,003	100.00%

Source: Volume 6 Table 7.2 to Table 7.7

3.8.6 As expected the car driver and car passenger trips in the North East, Yorkshire & Humberside and East Midlands show the largest decreases. Each of these regions experiences a corresponding increase in bicycle use as well as an increase in walking.

3.8.7 It is noteworthy that London loses approximately 5,000 car driver trips and 1,900 car passenger trips. This reduction is similar to Test 2, 3 and Test 4. This suggests that there is a degree of instability in London that is an artificial result common to each test and needs better understanding before outputs for London can be relied upon in reporting.

3.8.8 Returning to the responses recorded in Table 3-10 there is a clear explanation for the increase in bicycle use – the perceived cost of the mode has reduced by a considerable value. The justification for switch to walk is less obvious as it has not improved as a mode and as such the increase is a function of increased costs to travel by car.

3.8.9 Table 3-11 shows the distribution of displaced car driver and car passenger trips to other modes by region.

Table 3-11: Mode Switch by Region Test 5

	Mode Switch	**Bus**	**Rail**	**Walk**	**Cycle**	**Total**
NE	193,042	16.33%	4.12%	42.73%	36.82%	100.00%
NW	7,612	21.93%	20.55%	47.52%	10.01%	100.00%
Y&H	446,141	15.64%	5.67%	38.33%	40.37%	100.00%
EM	267,915	14.51%	4.31%	41.80%	39.38%	100.00%
WM	5,498	19.39%	18.44%	54.64%	7.53%	100.00%
EoE	2,805	19.00%	18.68%	55.22%	7.09%	100.00%
Lon	6,957	21.75%	32.26%	35.82%	10.18%	100.00%
SE	1,429	21.48%	13.65%	55.98%	8.89%	100.00%
SW	670	38.81%	10.45%	36.12%	14.63%	100.00%
Sc+W	471	19.96%	40.55%	35.67%	3.82%	100.00%
Total	932,540	15.61%	5.43%	40.46%	38.50%	100.00%

Source: Volume 6 Table 7.2 to Table 7.7

3.8.10 Table 3-11 shows that in the regions where the urban policies are being implemented the contribution which bus and especially rail are making as alternative modes is limited. This appears to be counterintuitive as these modes are more attractive to medium to long distance trips. From this result the response to the urban policy appears to rely primarily on redistribution and then walking.

3.8.11 The results presented in Figure 7.5 and Figure 7.7 should have had much more discussion as there are reductions in both trip-ends throughout the Wales, South West, the South East and London. As there are no policy interventions in these areas the results must be a function of model noise. Given most of the changes in these regions are of similar scale to the changes in the areas where policy changes have been implemented the results in those areas should be challenged.

3.8.12 Section 7.4.3 states for light vehicles that:

"The breakdown by road type for the three regions shows that reduction in vehicle kilometres is more on the A Roads compared to motorway, and B Roads (and other more minor classifications). These patterns appear to be consistent with the basis of the test."

3.8.13 However, for the results presented in Tables 7.8 to 7.10 by time period for the three regions, examining the percentage rather than the absolute numbers actually indicates a larger percentage reduction on motorways than on A-roads in almost all case. This percentage pattern appears to be less consistent with the basis of the test.

3.8.14 The results presented in the trip length distributions shown in Figure 7.9 and Figure 7.10 are consistent with the conclusion that the outcome of implementing the policies as described will be an overall shortening of trip lengths.

3.8.15 The results for highway flows, congestion and highway speeds all appear to be consistent with a reduction in highway demand.

4 Forecasting with NTMv5

4.1 Sensitivity Test 1: Demand growth - Objectives and Specification

4.1.1 The sensitivity tests 2 to 5 that have been analysed above all operate as variations on base year conditions. In contrast the "Sensitivity Test 1 – Demand growth" exercises the forecasting features for future years within the NTM. The objectives stated for this test included: demonstrating consistency in demand growth patterns with those from NTEM trip ends; demonstrating the impacts of changes in future GDP and values of time; and confirming adequate model performance and responsiveness under conditions of increased demand levels.

4.1.2 This test was implemented through applying the following set of changes within the 2030 model run.

- Changes in zonal trip productions and in zonal trip attraction weights and constraints were derived using growth factors based on NTEM trip end forecasts, which will include the impact of GDP growth on car ownership patterns.
- Urban fixed speeds were adjusted based on the local changes in trip ends.
- The base year LGV and HGV matrices were each scaled to 2030 using a global growth factor, while bus pre-loads on the network links also were adjusted.
- GDP growth impacted on travellers' values of time. Note that no changes have been made to fuel prices, fares etc.

4.1.3 This test specification provides useful insights into how the demand growth mechanisms in the model function from 2015 to 2030, even though it has not been specifically designed to provide a full reference case forecast run[34].

4.1.4 Firstly below, the trip end forecasting procedure itself is reviewed, then various aspects of the results of the forecast 2030 travel patterns are examined in turn.

Personal Trip End Forecasting Procedure

4.1.5 Volume 5 Section 6 explains that growth forecasts are developed externally to the model and applied using a combination of multiplicative and additive procedures. The additive procedures are to accommodate bespoke adjustments, and are not discussed here, as they are not used in Sensitivity Test 1. The multiplicative growth factors are derived from NTEM. While it is noted that "there will be some variance between the two forecasts due to differences in segmentation and base year values", these are not clearly explained. They turn out to be significant, though the extent of the discrepancy is never made clear.

4.1.6 In Volume 5 Section 6.2.1 it is noted that "the segmentation variables included [in NTM] have been kept consistent with those implemented in NTEM where they were significant and appropriate." With only minor discrepancies, the zoning system and purposes can be considered compatible (although NTM purposes are somewhat aggregated).

[34] For example, no changes in highway capacity have been implemented.

4.1.7 Hence, we can concentrate on the person/household segmentation. According to Volume 5 Section 6.4, this “is set out in Section 4.2 [should be 4.3] and uses the same types of categories as NTEM, namely gender (implicitly), employment, car availability and age. There is not a direct correspondence for some segments and hence careful consideration must be paid to how the segments are matched to each other”.

4.1.8 Volume 5 Table 4.5 gives the final segmentation, which is derived from the mode/destination choice model estimation results. For the Gender, Employment and Car Availability categories, the segmentations can all be derived from aggregations of the 88 NTEM categories (with the minor exception of children, where NTEM does not differentiate by Gender). The conspicuous exception is Age, where NTEM provides only 0-15 (children), 16-74, and 75+. Although the NTM age segmentation varies with purpose, for non-children the categories are generally 16-29, 30-44, 45-64 and 65+. While additional segmentation of NTEM categories in NTM could, with some effort, be handled, the non-alignment of the highest age group is much more problematic.

4.1.9 It is surprising that this has been allowed to happen. In Section 4.6 of the Mode/Destination Report, it is stated: “The models will be applied using the National Trip End Model (NTEM) to forecast growth in trip ends. Therefore, the segmentation variables tested were limited to those predicted by NTEM and household income.” This is followed by a table of “NTEM Segmentation variables” where Age categories are listed as: 0-15, 16-29, 30-44, 45-64, **65-74 and 75+**. It is not clear what the basis for this is. No further mention is made of these last two categories in the Volume 5 – they are always conjoined to 65+. But even if no significant effect could be found for the 75+ segment, it should surely have been kept separate to match the NTEM requirements.

4.1.10 Volume 5 Section 6.4.2 discusses the resulting problem, but it is very unclear what is meant here.

> *“The simplest (and likely most commonly used*[?]*) matching is given in Table 6.3 below [repeated for convenience as Table 4-1]. This assumes the growth in the 75+ age group is not too unrealistic for the age 65+. Making any adjustments within the processing tool would be complex to do robustly (with differential trip rates etc).”*

Table 4-1: Age band correspondence (Volume 5Table 6.3)

NTMv5 Age Band	NTEM Age Band
0-15	0-15
16-29	
30-44	16-74
45-64	
65+	75+

Source: Volume 5 Table 6.3

4.1.11 Almost immediately, however, it is found (as might have been expected) that “Assuming the 75+ growth rates from NTEM apply equally to the NTMv5 trip ends for those aged 65+ was however found to significantly overestimate the growth in travel demand."

4.1.12 Presumably by "growth rates from NTEM" is meant the zonal growth rates in productions for different purposes. These NTEM growth rates will be affected by the population shifts together with further "internal" shifts between NTEM categories (household size, car availability etc.). Moreover, these growth rates will be applied to the NTM base year population which may differ substantially from the synthetic segmented population used for NTEM application.

4.1.13 Unfortunately, we are not given any account of how serious the over-estimate is. In Volume 5 we are merely told that two adjustments were made:

"• to account for differential growth in the population by detailed age band sourced from ONS population projections; and

• a global adjustment for each trip purpose to bring growth in line with that forecast in the NTEMv7.2 dataset."

4.1.14 For NHb trips, the discussion is much more succinct. In Volume 5 Section 6.6 we find:

"In a similar manner to NTEM, the NHb productions in NTMv5 are derived from the Hb attractions. Therefore, it is not necessary to derive growth factors for the NHb productions explicitly."

4.1.15 However, the NTEM process is different in a number of ways: it uses the concept of "balancing areas" and ensures that within each balancing area, the purpose-specific attractions add up to the corresponding productions. These attractions then become the base to which NHb trip rates are applied to produce NHb Productions. Within NTM, however, the base for the NHb productions is the number of Hb Productions ending in particular zones via the destination choice process, with destination constraints applied in the case of HbW and HbEd, but not for other purposes. The resulting growth could be very different.

4.1.16 Because of the multiplicative growth factor process, some rules are required for dealing with possible zeros. Volume 6 Table 3.2 sets these out, but also notes that no cases arise where the NTEM values for either base or future (2030) year are zero: this implies that there are cases, however, where the cell in the NTMv5 Base is zero. In this case, the cell remains at zero: no further information is provided as to how prevalent this is.

4.1.17 In Volume 6, we read that the application of the growth process to NTMv5 Base trip ends:

"was found to give growth in trip productions that was not sufficiently close to the NTEM growth and produced too much growth particularly for shopping and recreation purposes"

4.1.18 Again, no figures are provided, and it is unclear whether only Hb trips are included. It is claimed that the reasons are the segmentation mismatch (generally, but with particular emphasis on the age issue) and "Differences in the profile of trip ends in

NTMv5 compared with those interpolated from NTEMv7.2" – it is unclear what is meant by "profile".

4.1.19 Apart from the age discrepancy, and perhaps the zero values issue, it is hard to see any reason why growth in Hb productions should diverge between NTEM and NTMv5. The "solution" that was adopted is along the lines noted in Volume 5: making use of the underlying ONS population projections shows that the growth in the over 75s between 2015 and 2030 is higher than the growth in the over 65s (though the data is presented in an extremely uninformative way [Volume 6 Figure 3.1]), leading to "population correction factors" for the 65+ group varying by Region between 0.88 and 1.01, with further compensating corrections for other age groups.

4.1.20 However, "the results were still not as close as desired" (though once again we are not informed as to the level of discrepancy) and hence a further set of global factors for each Hb purpose were derived: these are close to 1 for most purposes, but for HbShopPB and HbRecV they are around 0.935. While this does ensure that the overall NTEM growth for productions is reproduced, the method seems essentially *ad hoc* and, presumably, would need to be separately invoked for every forecast year. There could also be significant differences at the zonal level, and no indication is given of this. Table 3.10 suggests that there remain discrepancies at the detailed age band level (though it seems that the table has been produced on the original "matching" assumption (treating the combined 16-29, 30-44, 45-64 age groups as "16-74" and 65+ as 75+), so it is not a reliable comparison. All in all, this seems a significant flaw in the model design.

Freight Matrix Growth

4.1.21 Turning to the freight matrix growth, the Road Traffic Forecasts of 2018 (RTF18) were the source for the vehicle kilometre-based growth factors of 22.3%for LGVs and 1.1% for HGVs that were applied universally to scale these base **freight** vehicle matrices from 2015 to 2030. Section 3.3.3 explains that **non-freight** personal LGV trips are forecast instead by the VDM within which they comprise a proportion of the car plus van personal travel demand. However, it does not clarify whether this proportion of vans is assumed to remain unchanged from the base year assumed split or whether this proportion has been adjusted upwards for 2030.

4.1.22 This is an issue because the 22.3% growth rate for **all** van kilometres from the RTF18 is considerably larger than the forecast growth rate (Table B.3) for car (plus van) driver trips for the important van purposes HbW (5.1%), HbEB (9.6%) and NHbEB (8.5%). Moreover, the car (plus van) average trip lengths for these purposes are not forecast to change significantly (Table 3.23). Accordingly, either a growth rate greater than that for RTF18 should be applied to the freight LGV matrix or else the scaling factor of vans within cars plus vans should be increased substantially between 2015 and 2030, in order to ensure that the RTF18 overall (i.e. freight plus non-freight purposes) growth rate is achieved in the forecast for LGVs.

4.2 Discussion of Results - Responsiveness of the Model

4.2.1 **Given** the NTEM growth in Hb purposes, the remainder of Sensitivity test 1 really has only three further components: additional highway congestion associated with the higher demand (including changes in freight vehicles), the increase in VoT which will induce a move away from walk and cycle modes by reducing the impact of costs, and some change in destination attractiveness inasfar as the growth rates of the attraction variables vary between zones.

4.2.2 The model is iterated to achieve equilibrium with the highway supply, though nothing is said about convergence. The analysis of the output presented within Volume 6 covers

a) changes in total trips by region and mode;
b) zonal trip end changes;
c) changes in mode share by car ownership;
d) highway vehicle trips by region;
e) trip lengths, journey times, speed and cost;
f) trip length distributions;
g) vehicle kilometres by road type and region;
h) changes in highway flows;
i) congestion on the highway network; and
j) changes in highway speeds.

4.2.3 There are some significant issues relating to the presentation which will be carefully discussed.

4.2.4 Section 3.5.1 presents Table 3.12 to Table 3.17 giving "24-hour production trips" by region and mode. While the all modes total for 2030 is not given, it can be calculated as 76,287,030 trips. In the earlier Table 3.8, the total Hb productions is given as 58,557,033. It is assumed that these represent both outward and return movements, and that the difference (17,729,997) is the total NHb trips. Percentage changes in the external zones (Scotland, and parts of Wales) are smaller, for reasons which are not clear. The total 24-hour production trips allocated to car driver is 35,644,213.

4.2.5 Presenting the results in terms of total modal trips by region is not very informative, as it subsumes both the overall growth in productions, the effects brought about by a re-weighting of the segments (and car ownership in particular), and the effects brought about by the further three components (congestion, VoT, attractions) of the mode/destination choice model. Of more interest is the change in overall mode shares, as shown below in Table 4-2.

Table 4-2: Change in Mode Share 2015 to 2030

	Base share	**2030 share**
Car driver	43.9%	46.7%
Car passenger	19.4%	19.1%
Bus	7.2%	9.4%
Rail	3.6%	4.0%
Walk	24.2%	19.5%
Cycle	1.7%	1.4%

Source: Tables 3.12 to 3.17 from Volume 6

4.2.6 Of course, there is some regional variation, but the general pattern is similar, and the changes are not very large. The car modes increase (presumably mainly on account of increased car ownership) but with a bias towards car driver, and bus and rail also show increase, at the expense of the uncosted modes (walk and cycle), as would be expected due to the increase in VoT (and as noted in Section 3.5 of Volume 6).

4.2.7 Section 3.5.2 gives the absolute changes in zonal productions and attractions, for car drivers and for all modes combined. This is presented in map form, but an alternative presentation that corrected for either zone size or population, would convey more useful information.

4.2.8 Section 3.5.3 takes the absolute overall mode share changes and breaks them down by purpose and car ownership level. Commensurate with the table above, the changes are generally very small, and do not show much variation either by purpose or by car availability level, though the changes are usually a little smaller for lower levels of car availability. Of interest is the fact that car mode share (both driver and passenger) for HbW and HbEB falls slightly – perhaps due to increased congestion but no explanation is given.

4.2.9 Section 3.5.4 shows car and all vehicle trips on an O-D basis by time period. Trips are distinguished by region on the basis of "origin" (though it is not clear whether this might actually mean "production" in the case of Hb movements).

4.2.10 The tables are presented as absolute and percentage change from the base, though, given that the base is given in a separate table, it is possible with some effort to deduce the actual forecasts. It may also be noted that the "%diff from base" columns in Tables 3.20-3.22 are all wrong – the figures are generally too low. Adding up over all three periods gives a total of 13,257,153 car trips for 2030 which is completely different from the 35,644,213 given from Table 3.12. This suggests that these values are hourly values but as the interpeak total is similar in volume to the two peaks this is unlikely to be the case. While some of the car trips may have been switched to vans, this cannot possibly account for the discrepancy. And presumably these results are obtained after pivoting off the base highway matrices, but even so such a difference would not be expected. The level of discrepancy is similar for the base year. An explanation is urgently required.

4.2.11 It is claimed that "London shows a marginally greater increase in trips, particularly in the IP and PM, compared with other Regions", but this is not borne out – either in the existing tables or when the correct figures are calculated: in fact, while London growth is slightly above the growth in the overall total for the AM peak, it is **below** it for the other two periods. All in all, this section requires serious reassessment.

4.2.12 Section 3.5.5 shows, for each combination of mode and purpose, the change in average trip length (described, confusingly, as "average daily P/A distance"), duration, speed and money cost. The changes are generally extremely small and must be largely attributable to the combination of increased VoT and highway congestion. Concentrating on the car driver results, the overall results for the four variables are: trip length up by 1.7%, duration up by 4.1%, speed down by 2.3% and cost up by 1.8%. In relative terms these are reasonable, but it might have been more interesting to see how these effects were partitioned between the components referred to above (by first isolating the effect of the change in demand in productions from NTEM, then the impact on destination choice of the change in attractions, then changing the VoT and finally iterating to see the congestion effects).

4.2.13 It is noted that the increases are greater for the bus mode, especially for HbEd, HbShopPB and HbRecV, but part of the proposed explanation ("the increased volumes of trips from these purposes using the bus mode, where trips are generally longer") is difficult to understand. It is also the case that Employers Business trip lengths have **fallen** slightly for the car mode, for reasons which are not discussed (impact of congestion?).

4.2.14 Section 3.5.6 investigates the change in trip length distributions by purpose, for all modes combined. Overall there is little change, and the fact that the data is presented as graphs using different scales between the purposes for the differences in the proportions at various distances is not helpful. The changes are greatest for the HbEd, HbShopPB and HbRecV purposes, where the increased congestion is probably less important, so that the effect of increased VoT is to facilitate longer journeys for motorised modes. However, all purposes demonstrate a shift from shorter distances to longer.

4.2.15 Section 3.5.7 gives the change in Veh-Km by road type and region separately for lights (i.e. cars and LGVs combined) and HGVs, by time period. The base figures are not given, only the absolute and proportional change.

4.2.16 For HGVs the increases in vehicle kilometres are very small (about 1.5%) though there is some tendency for a shift to motorways. For lights in Tables 3.27 to 3.29 the increase is 15% to 24% across the periods: there is not much variation by either region or road type, other than that the growth rate in London is about 3 percentage points lower than the national average.

4.2.17 Section 3.5.8 presents network maps, for the three assignment periods, of the % increase in light vehicle flows. The general impression, confirming that of the previous section, is of uniformity, both spatially and by time period.

4.2.18 Section 3.5.9 again presents network maps, for the three assignment periods, this time for the volume/capacity ratios in both base and test years, as an indication of congestion. Although the changes do not seem unreasonable, it is difficult to compare the different maps, either between base and test, or between time periods. Some tabular form is desirable.

4.2.19 The same is true of Section 3.5.10, where the network maps for the three periods show per centage change in link speed. While the figures purport to illustrate the percentage changes in link speeds, the keys to the coding allow for a wide range of differences described as "link bars coloured by % diff, scaled by abs diff": hardly any of these differences can be seen on the maps, nor is it possible to align the shading with the ranges in the key. Apart from a general impression of decline, it is almost impossible to read these maps.

4.2.20 Overall, there are significant defects in the information provided for Sensitivity Test 1. Some of these are presentational, some of them relate to unexplained discrepancies between different pieces of evidence, and some reflect on the model itself.

4.2.21 Of these, the most serious is the potential divergence – without further correction – between the NTEM-based forecasts and the NTM outcomes. This has not been satisfactorily explained: clearly some of it relates to the mismatch in age segmentation which should never have been permitted, but it seems that this is not the whole story. Further, the divergence between tables relating to car driver productions and tables relating to car trips on the highway needs to be explained/corrected.

4.2.22 The nature of the test itself can be questioned, in terms of what have been here identified as its “components”: the overall growth in productions and the effects brought about by a re-weighting of the segments (and car ownership in particular), the change in destination attractiveness insofar as the growth rates of the attraction variables vary between zones, the increase in VoT which will induce a move away from walk and cycle modes by reducing the impact of costs, and the additional highway congestion associated with the higher demand (including changes in freight vehicles). No clear impression has been gained of the relative contributions of these components.

4.3 Demand growth - Comparisons with Recent Trends

4.3.1 Having examined the responsiveness of various outputs from the demand model to the input assumptions used to generate the forecast, a further assessment is presented below that compares the future trends estimated by the forecast, with the corresponding trends observed over recent years in different areas of England. Clearly, there is no guarantee that any specific current transport trend will continue unchanged into the future. However, if any such trend is shown to change radically within the forecast then some plausible reasons should be provided to explain why such a trend change is expected to arise.

4.3.2 In Table 4-3 to Table 4-5 below, the modelled changes in car (driver plus passenger) mode share forecast for England for the period 2015 to 2030 have been extracted from the Tables 3.12 to 3.17. These regional shares are contrasted with observed changes over the eleven year period from 2004/05 to 2015/16 as measured from the NTS. The NTS is a suitable benchmark as it was the source for many of the calibration targets that were used to implement the demand component of the NTM.

Table 4-3: Mode split for all car (driver + passenger) person trips by region: modelled 2015 & 2030, observed 2004/05 & 2015/16

Region	2015 Mod	2030 Mod	% point diff.	2004/05 Obs	2015/16 Obs	% point diff.	Mod – Obs 2015 % point diff
NE	65%	68%	3.2%	63%	64%	0.8%	1.6%
NW	65%	68%	2.8%	64%	67%	3.1%	-1.3%
Y&H	65%	68%	2.9%	64%	67%	3.3%	-2.0%
EM	67%	69%	2.8%	64%	67%	2.9%	-0.5%
WM	67%	69%	2.7%	68%	70%	2.5%	-3.5%
EoE	66%	68%	2.2%	68%	69%	0.2%	-2.5%
Lon	50%	52%	1.8%	44%	38%	-5.6%	12.5%
SE	65%	68%	2.4%	69%	69%	-0.4%	-3.5%
SW	65%	67%	2.7%	68%	67%	-1.9%	-1.8%
Eng ex Lon	66%	68%	2.7%	66%	68%	1.2%	-2.1%
Total	63%	66%	2.4%	63%	64%	0.2%	-0.3%

Source: Mod (Modelled) Volume 6 Table 3.12 to Table 3.17; Obs (Observed) NTS published table 9903

4.3.3 The columns to the left of Table 4-3 show the model's forecast growth from 2015 to 2030 in mode share for all car person trips (i.e. driver plus passenger) which is a 2.4 percentage points increase in car share for England overall. This growth rate of 2.7 percentage points across all regions excluding London is estimated to not vary greatly by region, ranging from 3.2 points in the North East to 2.2 points in the East of England. London has a lower estimated growth of 1.8 percentage points and is the region with the lowest estimated car mode split of 50% in 2015, compared to 66% on average across the other regions.

4.3.4 The columns to the right of Table 4-3 show similar information but for the observed past growth rate from 2004/05 to 2015/16 in mode share for all car person trips over an 11-year period, which is a 1.2 percentage points growth for England outside London, a rate less than half that forecast for the 15-year period from 2015 to 2030. Because average incomes increased much more slowly over the period 2005-2016 than in earlier years, provided that more rapid income growth re-emerges soon, it is conceivable that the future growth rate in car mode share outside London may accelerate to match that forecast by the model.

4.3.5 For London however the picture is not convincing. The observed all car person trips mode share is 38% from the NTS in 2015/16, whereas that in the model in 2015 is 50% - a major difference of 12.5 percentage points. Moreover, the model projects a future **gain** of 1.8 percentage points in car share, whereas over the period 2005-16 a major **decline** of 5.6 percentage points was observed in London.

4.3.6 Because the NTS is based on a sample of households it is subject to sampling errors that lead to apparent short-term changes in results between pairs of years, particularly for those population segments with small incidences in the population and so with small sample sizes. Accordingly, the comparisons of Table 4-3 have been repeated for an alternative set of NTS sample values also for an 11 year period but shifted to a year later from 2006/07 to 2017/18. Although this latter comparison changes many of the individual regional NTS values somewhat, overall the changes due to sampling errors are not sufficient to modify any of the key conclusions relating to major differences for London between model results and those observed.

Table 4-4: Mode split for car driver trips by region: modelled 2015 & 2030, observed 2004/05 & 2015/16

Region	2015 Mod	2030 Mod	% point diff.	2004/05 Obs	2015/16 Obs	% point diff.	Mod – Obs 2015 % point diff
NE	45%	48%	3.7%	40%	41%	1.3%	3.5%
NW	45%	48%	3.3%	40%	42%	2.1%	2.6%
Y&H	44%	48%	3.4%	41%	43%	2.1%	1.4%
EM	47%	50%	3.0%	40%	43%	2.6%	3.8%
WM	46%	49%	3.1%	43%	44%	1.4%	1.7%
EoE	47%	49%	2.3%	45%	45%	0.3%	1.6%
Lon	33%	35%	2.2%	26%	24%	-2.4%	9.1%
SE	46%	49%	2.7%	45%	46%	0.2%	0.7%
SW	46%	49%	2.9%	45%	43%	-1.6%	3.0%
Eng ex Lon	46%	49%	3.0%	43%	44%	1.0%	2.1%
Total	44%	47%	2.8%	41%	41%	0.4%	2.9%

Source: Mod (Modelled) Volume 6 Table 3.12 to Table 3.17; Obs (Observed) NTS published table 9903

4.3.7 Table 4-4 repeats the same results as those in Table 4-3 but now focusing solely on the car driver component of the car person trips. It indicates a faster rate of growth throughout for car driver trips than for all car person trips, implying that the passenger trip component must be growing more slowly than the car driver component. However, it again shows major differences for London, both in the 2015 match and in the trends through time, between the modelled forecasts and the observed values for car driver mode share.

Table 4-5: Mode split for other mode trips by London and rest of England: modelled 2015 & 2030, observed 2004/05 & 2015/16

Region	**2015 Mod**	**2030 Mod**	**% point diff.**	**2004/05 Obs**	**2015/16 Obs**	**% point diff.**	**Mod – Obs 2015 % point diff**
Bus							
London	9%	12%	2.8%	14%	14%	0.0%	-5.0%
Eng ex Lon	7%	9%	2.1%	6%	6%	-0.1%	1.3%
Total	7%	9%	2.2%	7%	7%	0.0%	0.5%
Rail+LRT							
London	10%	10%	0.7%	11%	14%	3.3%	-4.6%
Eng ex Lon	2%	3%	0.3%	1%	2%	0.4%	0.7%
Total	4%	4%	0.4%	3%	3%	0.9%	0.1%
Walk							
London	28%	23%	-4.9%	29%	31%	2.0%	-3.5%
Eng ex Lon	24%	19%	-4.8%	25%	23%	-1.7%	0.3%
Total	24%	19%	-4.8%	26%	24%	-1.2%	-0.2%
Cycle							
London	2.8%	2.4%	-0.4%	1.8%	2.2%	0.4%	0.6%
Eng ex Lon	1.5%	1.2%	-0.3%	1.5%	1.6%	0.1%	-0.1%
Total	1.7%	1.4%	-0.3%	1.5%	1.7%	0.2%	0.0%

Source: Mod (Modelled) Volume 6 Table 3.12 to Table 3.17; Obs (Observed) NTS published table 9903

4.3.8 Table 4-5 presents comparisons for the modes other than car but focussing only on the contrast between the values for London and those for the rest of England outside London.

4.3.9 Although outside London the base year mode shares in the model are reasonably close to those observed, in contrast within London the bus, rail and walk mode shares are each strongly underestimated in 2015.

4.3.10 For the rest of England outside London, only the forecast growth rate for rail is similar to that observed in the past. In particular, bus is forecast to increase its mode share substantially, though outside London it has not gained share at all in the past 11 years. Furthermore, within London, for most individual modes the forecast future pattern of growth differs strongly from the trend observed in the past:

- Bus share which has previously been **constant** - is forecast to **gain** 2.8 percentage points;
- Rail plus LU share which has previously been **growing rapidly** by 3.3 percentage points - is forecast to **gain only** 0.7 percentage points;
- Walk share which has previously been **growing** by 2.0 percentage points - is forecast to **decline** by a major 4.9 percentage points;
- Cycle share which has previously been **growing** by 0.4 percentage points (a 22% increase in absolute mode share from a small base) - is forecast to **decline** by 0.4 percentage points.

4.3.11 This decline in walk and cycle mode share in the forecast is a standard modelling response to the assumed increase in VoT, as explained in Volume 6 Section 3.5 (and was noted above). However, the fact that in reality cycle shares have been increasing rapidly in many dense urban areas, though not necessarily in all other parts of the country, indicates the need to have included within the model a mechanism that takes proper account of the very different modal trends in dense urban areas compared to trends in low density rural areas.

4.3.12 The trends in forecast highway flows also indicate differences with respect to observed trends in London. Based on Tables 3.27 to 3.29, the forecast growth from 2015 to 2030 in light vehicle (car plus LGVs) kilometres is 19% for England, summed across the three periods, while for London it is lower at 16%, comprising 13%, 21% and 15% growth respectively for the AM, IP and PM periods.

4.3.13 The observed change[35] in car kilometres within London over the 15 year period to 2015 was -14%, while the change in the subsequent period 2015 to 2018 is -1.2%. The corresponding change[36] for all motor vehicle kilometres (i.e. including cars plus LGVs and HGVs) within London over the 15 year period to 2015 was -10%, while the change in the subsequent period 2015 to 2018 is +1.1%. These car traffic declines occurred despite high population growth between 2002 and 2018 of +25% in Inner London and +18% in Outer London.

Conclusions

4.3.14 The forecasts for each mode in London of major changes in direction from its pattern of past growth or decline in mode share; together with poor base year matches in mode share, do not appear convincing. They are out of line with the reasonable mode split matches and plausible trends that are generally indicated for the regions outside London. The rapid growth in London's car traffic that is forecast, in contrast to observed major past declines, likewise is not convincing,

4.3.15 These results suggest that the model could not be safely used to examine policies that relate specifically to London.

4.3.16 Notwithstanding the fact that London is not the main focus of interest for use of the NTM (because policies specific to London are modelled independently by TfL using its own suite of models), it will be important to ascertain whether:

- The modelling issues indicated here are particular only to London itself; or
- They are wider issues that relate more generally to many rapidly growing dense urban areas across England as a whole, including but not confined to Inner London.

4.3.17 In this latter case, then the implications would be more significant for the suitability of the model, particularly if it transpired that overestimates of car share and of future car traffic growth in these dense urban areas had been counterbalanced within the calibration and the forecasts by underestimates of car usage in lower density areas. The significance of this topic would be best examined through a further analysis of car traffic trends by link type over time for aggregates of local authorities segmented into broadly similar levels of population density.

[35] Source DfT Road Traffic Table TRA8902 car traffic 1993-2018 by local authority
[36] Source DfT Road Traffic Table TRA8901 all motor vehicle traffic 1993-2018 by local authority

4.3.18 Although the trends through time in car traffic growth outside London are not implausible, for the other modes only those for rail are in line with trends observed in the rest of England in the years leading to 2015. No convincing explanation is provided for why bus passengers outside London would commence strong growth (39% growth in trip numbers to 2030) after years of minimal growth. The DfT Statistics Table BUS0103 indicates that in the decade to 2014/15 the overall growth in local bus passenger trips outside London was just 2%, whereas such trips have declined by 6% in the subsequent four years. Accordingly, the forecasts of non-car modes do not appear convincing,

5 Other Documentation

5.1 Quality Report

Structure

5.1.1 The NTMv5 Quality Report is presented in four parts:

- Part 1: Model Overview. A description of the model, its purpose, structure and main functions.
- Part 2: Technical information. A detailed technical specification and description of the model, including an explanation of the model design and development.
- Part 3: Model performance. This demonstrates highway and demand model validation, demand model realism testing, and whole model sensitivity tests.
- Part 4: Quality Assurance. This section details the quality assurance framework applied in the model development.

Part 1: Overview of NTMv5

5.1.2 Part 1 explains the justification and purpose of the NTMv5 along with client requirements. A key discussion in Section 2.4 is the explanation of how the Use Cases have been prioritised and confirming that Use Case 5 is paramount. The remainder of Section 2.4 describes how well the model addresses the other Use Cases and generally provides a reasonable set of caveats which caution the DfT about how much reliance can be placed on outputs from the model.

Part 2: Technical Specification

5.1.3 Part 2 provides a summary of the Developer Guides following a similar progression to those reports. A useful addition to the Quality Report would have been a schedule of correspondence between its content and the Developer Guides to demonstrate where content is repeated from a Developer Guide and which content is unique to the Quality Report.

5.1.4 Due to the quantity of material across the Developer Guides and the Quality Report it has not been possible within the Peer Review to ensure that all the content which is unique to the Quality Report has been identified.

Part 3: NTMv5 Performance

5.1.5 Part 3 introduces new material which does not appear in any of the Developer Guides that indicates specifically how the model performs using thresholds published in TAG Unit M2 for the demand model and M3.1 for the assignment model.

5.1.6 These standards are discussed in detail in Section 12 which sets the context in which they are being used, that is, the standards are focussed on models which are developed to consider schemes or specific localities and are therefore not directly applicable for a national model. Our review of the model performance against these standards is presented in Section 2.10.1 above.

5.1.7 The discussion about the process of calibrating and validating each model component receives very little attention in this report. The Quality Report is presented as being the first tier of the documentation with the User Guide and Developer Guides as the second and third tiers respectively. The expectation of the reader is that the Quality Report provides an overall summary of key topics for which the lower order documents provide greater discussion and detail to the more specialist reader. This would suggest that the Quality Report should not have any unique material but should just present summary information and distilled discussion, with the greater details provided elsewhere.

5.1.8 Unfortunately, this is not the case for convergence, calibration and validation with the paucity of information about the processes and performance a significant omission from the documentation with the result that it is difficult for a reader to determine whether the model converges, calibrates or validates to an acceptable level.

Part 4: Quality Assurance

5.1.9 Section 16 explains the approach to Quality Assurance (QA) which has been implemented through the project. Core to the QA process are the "Output Specification and Quality Assurance (OSQA) Plans." An OSQA was prepared for each phase of the project as it was initiated and then used throughout the phase to manage quality processes.

5.1.10 The plans have not been reproduced within the Quality Report due to the scale and duration of the NTMv5 development project. Summary tables of the key QA activities have been provided in Section 16.2.

5.1.11 The summary tables in Section 16.2 are developed from the DfT Quality Assurance for Analytical Models[37] which has been split into five key areas requiring QA checks:

- Inputs: Checks on the quality and reliability of raw data and other inputs, as well as their appropriateness for NTMv5;
- Processing: Checks on the importing, manipulation and formatting of data, robustness of processes and replicability once model is transferred to DfT;
- Validity of model: Checks on the quality of outputs against observed data and general assurance on quality of the model as a whole;
- Checks on functionality: Ensuring that the model, either as a whole or specific components, functions to the required standard, in terms of implementation, modelling processes and run times; and
- Checks against scope: Ensuring that all elements of the model, including inputs, processing stages and final outputs, meet the standard set out and agreed with the DfT.

5.1.12 The Tables in Section 16.2 include a comprehensive schedule of areas where the model development was considered, reporting these against whether they were part of the initial scope for development or a refinement thereof along with a demonstration of checks for the mathematical processes within the model and verification of outputs which the model was producing.

5.1.13 The end column in each table is entitled "Details of the QA undertaken." Generally, each cell in this column is a statement of an action, agreement or piece of technical work which has been recorded as the evidence that QA has been undertaken. The

37 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/350904/qa-modelling-guidance_pdf.pdf

evidence would be substantially enhanced if it referred to a meeting minute or document in which the evidence exists.

5.1.14 For example, the evidence that the treatment of urban speed areas was agreed with the DfT is simply the statement “Details of documentation and agreement with DfT on strategy for modelling urban areas…” with no reference to where the document showing that agreement can be found.

6 Consideration of the Use Cases

6.1 The Use Cases

6.1.1 As presented in the Quality Report Section 2.4, the model has been set the challenge to provide evidence to the DfT in respect of six Use Cases which are defined below:

- UC1 Strategic Roads Investment and Resilience:
  - To analyse the impacts of packages of roads schemes at a national level. This could include high-level calculation of value-for-money (VFM), points of expected congestion, and analysis of resilience of the network.
- UC2 Road User Charging and other potential policy:
  - Flexibility to adapt to road policies in future Parliaments. This could include various forms of road pricing, including pricing on strategic roads (tollbooth, distance-based or vignette), urban roads (e.g. congestion charging vignette), or parking policy in urban areas, or other behavioural devices such as High Occupancy Vehicle (HOV) lanes.
- UC3 Local Investment and Policy:
  - Variety of analysis including national impacts of congestion relief schemes; Policy impacts of introducing public transport improvements (e.g. light rail). This could potentially include travel demand management in future parliaments. It is noted that parking is covered in use case 2.
- UC4 General Support for DfT Teams (other than Roads / Local):
  - Environmental analysis of transport policies relating to carbon and/or an approximation of air quality emissions, and also aviation surface access.
- UC5 Scenario-based National Traffic Forecasting:
  - Understanding of changes in population or travel trends (e.g. driving rates amongst young people), with scenarios around GDP, car ownership, fuel price, road tax.
- UC6 Exploring the unknown:
  - Testing new policies or technical developments that have not been modelled before (e.g. CAVs). Testing new policies or technical developments of whose existence we are not currently aware.

6.1.2 The Department has indicated that use case 5 (UC5, National Traffic Forecasting) should be considered first for the following reasons:

- Production of National Road Traffic Forecasts (RTF) is an existing use and the most fundamental purpose of the National Transport Model;
- UC5 includes scenario-based forecasting, and it is likely that many of the questions arising in UC1 and UC2 will be considered as scenarios for RTF; and
- In order to robustly forecast road traffic for the RTF, it will be necessary for NTMv5 to include the influence of a wide range of factors. Those factors

highlighted in other use cases can therefore be considered as implicitly features of UC5.

6.2 Convergence Achieved in Test Runs of the Model

6.2.1 It had originally been envisaged that the set of sensitivity test runs that were reported in Volume 6 would provide much of the information on the performance in practice of the model that would underpin the guidance below on its suitability for different Use Cases.

6.2.2 However, a number of systematic unexpected results have emerged from these tests that strongly suggest that there may have been issues of lack of convergence in these runs that then reduce the range of conclusions that can be drawn from them.

6.2.3 In particular, there are unexpected similar levels of growth in car driver and passenger trip productions in London indicated in each of Tests 2, 3 and 4, even though these policies were not necessarily expected to have such an impact within London. There are also many other counter-intuitive variations indicated in localities far from where the individual policy test is expected to have its impacts. These results appear likely to have been partially corrupted by noise created by inadequate levels of convergence being achieved in the comparisons of the base with the sensitivity tests runs. Accordingly, considerably uncertainty about the overall performance of the model still remains because of these difficulties encountered in assessing the results of the various sensitivity tests.

6.2.4 The potential impact on the test results from lack of convergence may have been amplified in some cases by the relatively small scale of some of the actual policy tests.

6.2.5 If the various counterintuitive results from sensitivity tests could largely be resolved simply by running the model for a few more iterations, then this would be a positive development. The extra run time burden that this entails would in the medium term gradually become less of a problem, assuming that computer facilities continue to increase in power and speed through into the future.

6.2.6 If on the other hand the lack of good convergence is more structural within the modelling system and cannot be universally resolved by a few extra iterations then the issues remain serious. Moreover, if it transpires that some of the significant counterintuitive results that arose are not simply due to lack of convergence then more major issues with the model design or implementation would need to be considered.

6.2.7 Based on the information currently published on the model results and on the lack of information presented on the convergence achieved, it is difficult to be certain about which of the above situations holds. Consequently, further detailed experimental model running and analysis would be required in order to provide more informed recommendations on the capabilities of the model, so this uncertainty should be noted when considering the comments below on individual Use Cases.

6.3 Comments on Use of the Model with respect to Each Use Case

Use Case 1 Strategic Roads Investment and Resilience

6.3.1 The zone system and network have been designed with access to and use of the Strategic Road Network their primary focus. The links in the network are primarily the higher order highway network. Junction coding has been implemented in a simplified manner which reduces run time of the model but retains the majority of the junction delay characteristics.

6.3.2 The approach to simplify urban areas, from the perspective of considering their relationship with the strategic network, is prima facie acceptable but there is evidence from the route choices being exhibited that this simplification is impacting inappropriately on route selection along the strategic network close to large urban areas.

6.3.3 As this is evident in the base year it is likely to become a stronger feature in forecasts where congestion on the strategic network increases but the congestion response as a function of network coding in urban areas is less pronounced.

6.3.4 The most relevant Sensitivity Test for this Use Case is Test 2 which considers adding a substantial new piece of highway infrastructure into the network. Implementing this type of intervention within the model is a straight-forward and common task and as such does not present any challenges to a user in specifying the input changes or interpreting the outputs.

6.3.5 However, as discussed in Section 3.5 the results which are presented in Volume 6 Section 4 showed very minor changes in all the measures considered, including some which were counterintuitive.

6.3.6 The implication of this is that although the model does appear to be suitable to test investment on Strategic Roads either the scale of the schemes that would need to be implemented has to be greater than that which has been tested and reported or else the model noise issues that appear to be related to lack of convergence would first need to have been resolved.

Use Case 2 Road User Charging and other Potential Policies

6.3.7 There is a series of issues that are identified and briefly discussed in the Quality Report Section 2.4.3 which mean that implementing many of the interventions included within this Use Case is not straightforward and will require considerable work outside the model itself.

6.3.8 In addition to the issues which are identified in the Quality Report the limited representation of existing alternative modes to car trips within urban areas, such as bus and rail, means that the response to charging strategies cannot be adequately represented.

6.3.9 Sensitivity Test 4 is the most relevant test for the road pricing element of this Use Case and Sensitivity Test 5 does include modifications to parking charges and availability.

6.3.10 As discussed above in Section 3.7 and Section 3.8 the results from the Sensitivity Tests have some plausible elements but do raise a number of questions about the alternative modes selected or extent of redistribution.

6.3.11 The implication of these limitations is that the model in its present form cannot provide sufficiently robust outputs to inform Use Case 2.

Use Case 3 Local Investment and Policy

6.3.12 The interventions included within Use Case 3 are focussed on congestion relief in urban areas. The representation of urban areas in NTMv5 is characterised by a combination of simplifying assumptions that have been implemented to manage model run-times and stability. A consequence of the simplifying assumptions is that urban interventions cannot be directly modelled.

6.3.13 The Quality Report Section 2.4.4 does suggest that the representation of local impacts in NTMv5 could be achieved by making some adjustments to standard model assumptions. Guidance on what these might be has not been provided.

6.3.14 The general performance of the model in the base year and in forecasting mode within London has been shown above in Section 4.3 not to be to the standard achieved in other regions. It appears likely that similar types of issues may also arise within other major urban areas that provide a wide choice of competitive modes, though specific information on model performance in such areas is not generally presented in the reports.

6.3.15 Sensitivity Test 5 does attempt to address some of the interventions included within Use Case 2 but as noted in Section 3.8 the results are inconclusive and cannot be disaggregated to isolate the contribution of specific interventions.

Use Case 4 General Support for DfT Teams (other than Roads / Local)

6.3.16 There are two areas of interest within this use case, firstly that of creating outputs for environmental assessments of transport policy and secondly, assessing access to and egress from ports and airports as a function of policy objectives.

6.3.17 The Quality Report Section 2.4.5 rightly identifies that the model is limited in its ability to inform environmental assessments. NTMv5 is consistent with all other strategic transport models in this respect as environmental analysis requires outputs at a level of detail that is not readily available from a strategic model.

6.3.18 In addition to this limitation the structure of the zone system and the simplified coding in urban areas means that small area data, of the type used for environmental assessment, is particularly unreliable. In Section 2.1 above, the discussion on the aggregation errors within their estimation of fuel consumption (costs) by type of road vehicle, coupled with the unhelpful differences between the VDM and HAM in these calculations, imply that great care would be needed in their interpretation.

6.3.19 Turning to the ports and airports whilst separate zones are defined for these locations the trips associated with them are included in the standard user classes and vehicle types.

6.3.20 There has not been any specific validation of trips using these facilities, as explained in the Quality Report Section 2.4.5 and as such any intervention to modify access or increase demand cannot be directly implemented.

Use Case 5 Scenario-based National Traffic Forecasting

6.3.21 The model has been set up to fulfil the requirements of this Use Case to the greatest extent. The network has been coded in a form that allows for disaggregation of results by region, link type and purpose meaning that the form of the National Road Traffic Forecast tables can be replicated. It maintains a clear distinction throughout the HAM between cars, LGVs and HGVs, with distinct speeds coded for each vehicle type.

6.3.22 The procedure used to generate the future overall rate of growth in HGV demand and the freight component of the LGV demand, currently is external to this model so some further work or external source may be needed for this.

6.3.23 The results discussed above in Chapter 4 on the performance of the model in forecasting mode indicates that the road traffic forecasts in the regions other than London are not implausible. In contrast, they indicate an unexplained and unlikely major growth in bus trips outside London and breaks in trends there also for cycle trips. Only rail mode is in line with previous growth trends in trips outside London. The poor base year and growth trend matches identified within London for road and other modes are likely to be relevant also to other major urban areas so that forecasts of future road traffic growth within such urban areas are likely to be substantially overestimated.

6.3.24 Accordingly, the model should be suitable for use in forecasting the growth of road traffic in most areas other than those adjacent to or within major urban areas.

Use Case 6 Exploring the Unknown

6.3.25 The challenge for this use case is extremely difficult. Many of the issues which emerge about yet to be defined policies and unknown technologies are likely to influence decisions about whether to travel and so primarily to impact on trip rates but these lie outside the scope of the NTMv5.

6.3.26 However, it is noted that connected and autonomous vehicles (CAVs) are explicitly mentioned in the Use Case which implies that the NTMv5 should at least be able to recognise how these vehicles will use the network.

6.3.27 The only variable within the model which could be changed to represent this efficiency gain is the PCU value. For CAVs this is assumed to be a lower value as the vehicle itself, although the same size as a car, will be more efficient passing through junctions and travel more closely behind the vehicle in front on high speed roads. This modification can be implemented easily in NTMv5 as a new vehicle type can be added into the assignment procedure.

6.3.28 Returning to the behavioural challenge, it is not clear how CAVs would be included in the demand model so that creating the assignment matrix remains a challenge.

6.4 Summary of Using NTMv5 for its Defined Use Cases

6.4.1 As discussed above the NTMv5 is has been specified to fulfil the requirements of Use Case 5 most completely but due to the manner in which the forecasts for demand are generated there are challenges to completely address this Case.

6.4.2 Its design means that the model can provide some outputs to inform Use Case 1 with the caveat that the tested schemes will need to be of sufficient magnitude to create a response which is greater than model noise.

6.4.3 Of the other four use cases the NTMv5 needs significant exogenous effort to either design appropriate inputs or to distil reliable data from the outputs so that it should not be used other than by very experienced modellers.

7 Independent Model Runs

7.1 Review Summary

7.1.1 In parallel to reviewing the reports associated with the NTMv5 development the model in its entirety was shared with the Peer Review team. The purpose of this exercise was to demonstrate:

- whether the model was portable from one organisation to another;
- that a modeller with sufficient skill and experience could repeat the model runs to replicate base year results;
- that a modeller with sufficient skill and experience could implement tests which have been undertaken to demonstrate the model performance; and
- that a modeller with sufficient skill and experience could design and implement independent tests without support from the model development team.

7.1.2 This section of the Peer Review summarises the Practitioner Review.

7.2 Installing and Running NTMv5

7.2.1 The hardware and software requirements to run the NTMv5 are provided in the User Guide Section 5.

7.2.2 Although the document suggested that as a minimum 256GB of RAM would be required for smooth running of the demand model, it was found that the runs could be successfully replicated using 128GB RAM. Moreover, it is observed that the RAM specification is particularly crucial for successful running of 2 procedures: the doubly constrained nested demand model of HbW and HbEd. When the setup was replicated on another machine which had 64GB of RAM these procedures failed to launch.

7.2.3 NTMv5 is mostly implemented in the VISUM software version 17.01-15, and the same version was used to replicate the runs.

7.2.4 The User Guide Section 5.1 advises to have a minimum of 150GB free space before launching any NTMv5 run. However, it was found that a minimum of 300GB was required as the peak hour highway assignments crashed with an error message about insufficient memory. It is noticed that this issue arises only when intermediate version files of the demand model are to be saved. In cases where intermediate files are not required, a free space of 150 GB was found to be adequate.

7.3 Observations on the User Guide

7.3.1 The User Guide is divided into three parts:

- Model Overview
- Installation Guide
- Running NTMv5

Part 1: Model Overview

7.3.2 Part 1 starts with an explanation of how the majority of the model has been constructed within the VISUM software and identifies which elements rely on exogenous assumptions and inputs and how these feed into the model.

7.3.3 Section 3 introduces the concept of policy tests that the model has the capability to explore. These are aligned to the Use Cases discussed in the Quality Report. This discussion is potentially one of the more useful sections of text in the suite of documents and its location in the User Guide risks hiding it from view.

7.3.4 Table 3-1 which records the policy tests and explains to the model user where to modify inputs to achieve an appropriate output could be expanded to include known limitations that the policy test if implemented in NTMv5 would inherit. For example, "Public transport test" could have the known limitations that the NTMv5 does not have public transport assignment so changes to public transport patronage cannot include any route data or crowding.

Part 2 Installation Guide

7.3.5 The installation guide provides advice on the dimensions for hardware required to run the NTMv5, the version of the software used during its development and how the folders are structured such that files are read from and written to the correct locations during the operation of the model.

7.3.6 As noted above, these instructions were followed during the Practitioner Review and the model was successfully installed onto the WSP hardware. Observations on the model which was provided are set out below.

Part 3 Running NTMv5

7.3.7 Part 3 describes to the User the series of processes which are to be undertaken when specifying new scenarios. The instructions which were relevant to the Peer Review independent tests were implemented successfully.

7.3.8 Section 7 describes the different processes which allow a user to modify the structure of the model by adding or subtracting network and zones. These follow standard VISUM procedures so do not present issues at this point. However, some of the processes within the model as a whole would require reviewing to ensure they are working correctly – specifically pivoting should a zone be added.

7.3.9 Section 9 described the output analysis workbooks which have been prepared in MS Excel and processes to export data to GIS platforms which mean that analytics and results can be interrogated from the model output run files. A number of these tools were reviewed during the Peer Review and were verified as operating correctly.

7.3.10 However, it should be noted that some many of the tools appear to be structured to fit exactly to the dimensions of the NTMv5 and do not offer any flexibility should changes be made to the underlying model. This relates primarily to the zone system and the tools which aggregate and report spatial data.

7.4 Checks and Reviews of the Model Components

Model Version

7.4.1 The model provided for Peer Review by the Department was numbered version 31, which is referred to as the Benchmark Model, whereas the model upon which the reporting was based was version 30, referred to as the Reported Model. The minor difference between these models was in the user defined attributes for LGVs. This difference has led to some minor differences in the outputs, which are noted in this section.

Review of Model Run Times

7.4.2 The User Guide Section 8.3.4 details the procedure to undertake the record of model run time. This procedure was implemented during the Benchmark Model runs. The recorded run times for a single demand/supply loop are compared against those of the Reported Model in Table 7-1 below.

Table 7-1: Model Run Times

Hh:mm:ss	Reported Model	Benchmark Model
Imports & Processing	00:59:35	01:17:28
VDM	02:23:02	02:25:24
PA-OD & Pivot	00:34:12	00:43:00
HAM AM	01:07:37	01:50:17
HAM IP	01:08:11	01:49:10
HAM PM	01:14:28	02:23:00
Total	07:27:05	10:27:55

7.4.3 As can be seen in running the model the time taken to replicate the runs are generally on the higher side which is due to using a machine with less RAM. The main impact has been on highway assignments.

7.5 Ability to Replicate Results from the Core Models

Ability to Replicate Results from VDM

7.5.1 The Reported Model results are available in the Quality Report Section 13. The Base run of the demand model was primarily checked for robustness using the measures mode share and trip length distributions.

7.5.2 The mode shares and average trip length in the Benchmark Model were checked and as shown in Figure 1 and Figure 2 respectively.

Figure 1: Mode Share across the Demand Strata

Source: Quality Report Table 13.4 and Table 13.5 and WSP independent model runs

7.5.3 As Figure 1 shows the Benchmark Model produces the same mode share results as the Reported Model.

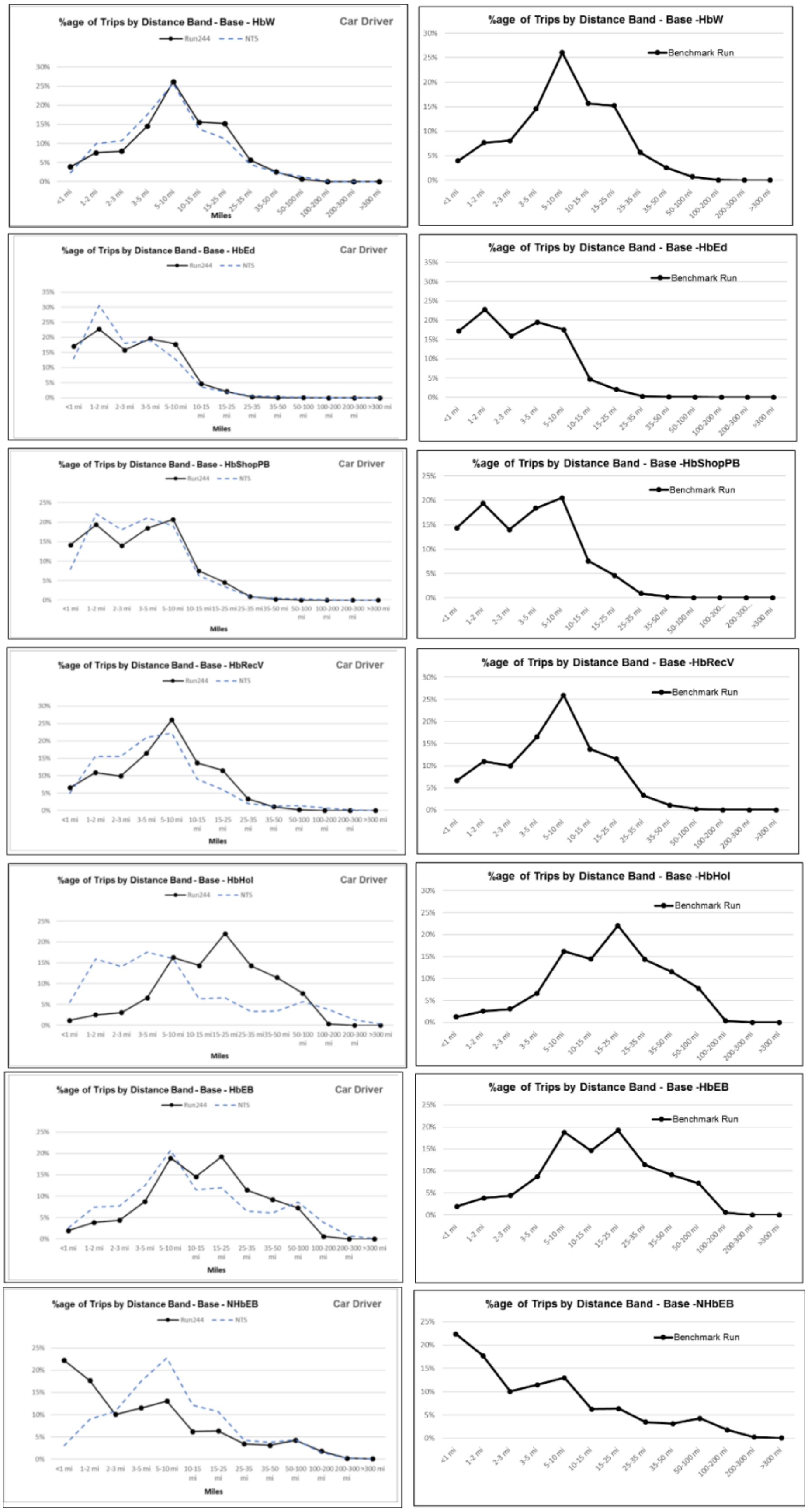
Figure 2: Average Trip Length across the Demand Strata

Source: Quality Report Table 13.6 and Table 13.7 and WSP independent model runs

7.5.4 As Figure 2 shows the Benchmark Model produces a very good match for average trip length when compared to the Reported Model although not an exact match especially for non-home based rail trips.

7.5.5 The dominant mode in the NTMv5 is car driver so further investigation was undertaken for this mode to identify whether the trip length distribution in the Benchmark Model matched the Reported Model. This is shown in Figure 3 below.

Figure 3: Base Year VDM Trip Length Distribution for Car Driver



Source: Quality Report Appendix B and WSP independent model runs

7.5.6 As Figure 3 shows the Benchmark Model is matching the Reported Model exactly for the car driver trip length distribution by purpose.

Ability to Replicate Results from HAM

7.5.7 The Quality Report states 1,901 links were used as screenlines but the Benchmark Model only identifies 1,894 links with observed counts which means that for some reason data 7 for links was removed.

7.5.8 Calibration statistics for each assignment period have been produced from the Benchmark Model and compared with those reported in the Quality Report.

Table 7-2: AM Peak HAM Screenline Result Comparison

		NTM Report (1901 links)			Benchmark (1894 links)		
% Pass		**Lights**	**HGV**	**Total**	**Lights**	**HGV**	**Total**
Links	Flow Difference	85%	99%	83%	87%	100%	86%
	GEH (<5)	77%	96%	76%	74%	95%	73%
	GEH (<7)	87%	99%	86%	85%	98%	84%
	GEH (<5) or Flow Diff	86%	99%	85%	88%	100%	87%
Screenlines (68)	Flow Difference (5%)	96%	96%	96%	97%	96%	97%
	GEH (<4)	88%	100%	88%	96%	100%	96%
	GEH (<7)	99%	100%	99%	100%	100%	100%
Mini-Screenlines (134)	Flow Difference (5%)	91%	89%	91%	95%	93%	96%
	GEH (<4)	90%	100%	90%	97%	100%	97%
	GEH (<7)	98%	100%	98%	100%	100%	100%

Source: Quality Report Appendix E. Table E.1

7.5.9 As Table 7-2 shows the Benchmark Model performs better than the Reported Model on the screenline summaries but worse on individual count sites. This appears to be a contradictory outcome and suggests that further analysis of the balance between sites with too much traffic and sites with too little traffic should be undertaken. The changes do not affect the overall conclusion that the model is almost calibrating to an acceptable standard.

Table 7-3: IP Peak HAM Screenline Result Comparison

		NTM Report (1901 links)			Benchmark (1894 links)		
% Pass		**Lights**	**HGV**	**Total**	**Lights**	**HGV**	**Total**
Links	Flow Difference	91%	99%	90%	91%	100%	90%
	GEH (<5)	84%	95%	82%	80%	95%	78%
	GEH (<7)	92%	99%	91%	89%	98%	88%
	GEH (<5) or Flow Diff	92%	99%	90%	91%	100%	90%
Screenlines (68)	Flow Difference (5%)	97%	93%	97%	97%	94%	97%
	GEH (<4)	93%	100%	91%	97%	100%	97%
	GEH (<7)	99%	100%	99%	100%	100%	100%
Mini-Screenlines (134)	Flow Difference (5%)	89%	86%	90%	93%	88%	94%
	GEH (<4)	94%	100%	91%	99%	100%	98%
	GEH (<7)	100%	100%	100%	100%	100%	100%

Source: Quality Report Appendix E. Table E.2

7.5.10 Table 7-3 again shows the Benchmark Model performs better than the Reported Model in some areas but worse in others. The pattern of poorer performance for indivual sites but improved screenlines is similar to the AM peak and confirms that additional analysis comparing the models would be assist in understanding the reassignment impacts of the differences between the models.

Table 7-4: PM Peak HAM Screenline Result Comparison

		NTM Report (1901 links)			Benchmark (1894 links)		
% Pass		**Lights**	**HGV**	**Total**	**Lights**	**HGV**	**Total**
Links	Flow Difference	84%	97%	83%	87%	100%	87%
	GEH (<5)	77%	96%	75%	74%	96%	72%
	GEH (<7)	87%	99%	87%	85%	99%	84%
	GEH (<5) or Flow Diff	86%	97%	84%	88%	100%	87%
Screenlines (68)	Flow Difference (5%)	96%	90%	96%	97%	93%	97%
	GEH (<4)	87%	100%	85%	97%	100%	96%
	GEH (<7)	99%	100%	99%	100%	100%	100%
Mini-Screenlines (134)	Flow Difference (5%)	87%	83%	88%	91%	86%	93%
	GEH (<4)	87%	100%	87%	94%	100%	94%
	GEH (<7)	97%	100%	97%	99%	100%	99%

Source: Quality Report Appendix E. Table E.3

7.5.11 Table 7-4 is similar to both the other two assignment periods with some improvements and some areas of degradation in performance with a similar pattern with regard to individual sites and screenlines.

7.5.12 Table 7-5 shows the change to the calibration results between the Reported Model and Benchmark Model.

Table 7-5: Difference between Benchmark and Reported values for Regional Stats: All vehicles

		NE	NW	Yorks	EM	WM	EoE	Lon	SE	SW
AM peak - % pass	Flow Diff	-1%	1%	-3%	-2%	-1%	0%	-3%	-1%	1%
	GEH (<5)	-6%	-4%	-2%	-5%	-4%	0%	-6%	-2%	0%
	GEH (<7)	-5%	-1%	-4%	-8%	-4%	1%	0%	0%	-1%
	GEH (<5) or Flow Diff	-1%	0%	-2%	-3%	-2%	0%	-3%	-1%	1%
Inter-peak - % pass	Flow Diff	1%	-1%	-3%	-2%	-4%	-2%	0%	-1%	-1%
	GEH (<5)	-4%	-6%	-3%	-5%	-6%	0%	-3%	-2%	-3%
	GEH (<7)	-5%	-4%	-3%	-8%	-6%	-3%	-6%	-1%	-1%
	GEH (<5) or Flow Diff	1%	-1%	-1%	-2%	-4%	-1%	0%	-1%	0%
PM peak - % pass	Flow Diff	-1%	1%	-1%	-5%	-3%	0%	3%	1%	0%
	GEH (<5)	-4%	-1%	-3%	-6%	-5%	-1%	-3%	0%	0%
	GEH (<7)	-5%	-4%	-3%	-7%	-4%	-1%	-3%	-1%	-1%
	GEH (<5) or Flow Diff	0%	1%	-2%	-6%	-3%	0%	0%	1%	1%

Reference: NTM v5 Quality Report v4.0 Appendix E. Table E.4

7.5.13 Table 7-5 generally shows that the model is worse in terms of passing the flow and GEH criteria in each region. As this is a function of the same data discussed above in Table 7-2 to Table 7-4 this is to be expected. Whilst the screenline data is not provided, it is assumed that the results would be similar. This negative impact caused by the changes between the Reported Model and Benchmark Model should be investigated.

7.5.14 The Quality Report Section 14.11 details the validation statistics of NTMv5 model which are generally observed to be poor. It is also seen that validation is primarily undertaken through a set of 137 Ad-hoc links. Although these links could be identified within the model, there were no observed data to recreate these statistics.

7.6 Ability to Replicate Results of the Realism Tests

7.6.1 Volume 6 Section 2 states that, as part of Realism testing of NTM, all of the standard realism tests defined in TAG unit M2 were carried out by varying fuel costs, public transport fares and highway journey times respectively. Considering the scope of present study and model run times, the ability to replicate the results of Fuel Cost elasticity test was undertaken as an appropriate measure to ascertain the VDM performance.

7.6.2 As specified in the document the fuel cost realism test has been implemented by increasing only the fuel components by 10%. This is implemented by modifying the input values in the VOC calculation as shown in Table 7-6 below.

Table 7-6: VOC parameters increased by 10%

User class	a	b	C	d
HbEB	81.31101	5.40419	-0.03443	0.00033
HbW	97.57319	6.48461	-0.04136	0.00044
HbOther	97.57319	6.48461	-0.04136	0.00044

7.6.3 The revised values for the HAM VOC are presented in Table 7-7.

Table 7-7: Vehicle Operating Costs for Realism Tests

User class	Average network speed (km/hr)	Benchmark HAM VOC parameter (pence per km)	Realism Test VOC parameter (pence per km)
Car: Commute	54	6.66	7.34
Car: Business	65	13.01	13.42
Car: Other	54	6.66	7.34
LGV	54	14.66	15.48
HGV	65	48.13	51.30

7.6.4 Volume 6 Section 2.2.1 indicates that the demand / supply iterations model reaches a reasonable level of stability after three iterations. The Reported Model was run for eight iterations: however, to retain a reasonable run time the Benchmark Model was run for three iterations. Table 7-8 shows the calculated elasticity values from both the Benchmark Model run and the Reported Model run.

Table 7-8: Car driver and passenger trip km elasticities by purpose

Trip purpose	Reported Model	Benchmark Model Iteration 4	Reported Model	Benchmark Model Iteration 4
	Car driver	Car Driver	Car passenger	Car Passenger
HbW	-0.28	-0.29	0.05	0.04
HbEd	-0.41	-0.43	-0.20	-0.22
HbShopPB	-0.43	-0.43	-0.42	-0.44
HbRecV	-0.40	-0.43	-0.22	-0.26
HbHol	-0.39	-0.45	-0.07	-0.11
HbEB	0.02	-0.01	0.10	0.06
NHbEB	-0.17	-0.06	0.17	0.23
NHbO	-0.25	-0.26	-0.15	-0.17
Total	-0.28	-0.30	-0.19	-0.22

7.6.5 The results of the Benchmark sensitivity test show a close match with the Reported Model values with HbEB now showing a negative sign in the Benchmark Model. Where there are larger changes, specifically the NHbEB, it would be prudent for the DfT to explore the underlying cause of these differences.

7.6.6 Table 7-9 shows the relative change in the number of trips when comparing the two model runs against one another.

Table 7-9: Change in trips by mode for each trip purpose due to 10% fuel cost increase (all areas)

	Car Driver		Car Passenger		Bus		Rail		Cycle		Walk	
Purpose	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**
HbW	-0.8%	-0.9%	0.7%	0.7%	1.9%	2.0%	1.2%	1.3%	1.8%	1.9%	1.8%	1.8%
HbEd	-2.1%	-2.4%	-1.9%	-2.2%	1.6%	1.7%	1.6%	1.7%	1.8%	1.9%	1.7%	1.8%
HbShopPB	-1.8%	-2.1%	-2.2%	-2.6%	4.1%	4.4%	4.3%	4.6%	3.9%	4.2%	3.8%	4.1%
HbRecV	-1.1%	-1.2%	-0.8%	-0.9%	2.3%	2.5%	2.5%	2.7%	2.2%	2.4%	2.3%	2.4%
HbHol	-1.2%	-1.3%	0.3%	0.2%	2.6%	2.9%	3.1%	3.5%	2.3%	2.6%	2.3%	2.6%
HbEB	-0.3%	-0.2%	0.1%	0.1%	1.0%	0.8%	0.9%	0.6%	1.0%	0.9%	1.0%	0.8%
NHbEB	-0.5%	-0.6%	0.5%	0.5%	1.3%	1.4%	1.4%	1.5%	1.4%	1.4%	1.4%	1.4%
NHbO	-0.4%	-0.4%	-0.3%	-0.4%	0.5%	0.6%	0.6%	0.7%	0.6%	0.6%	0.5%	0.6%
Total	**-0.8%**	**-0.9%**	**0.7%**	**0.7%**	**1.9%**	**2.0%**	**1.2%**	**1.3%**	**1.8%**	**1.9%**	**1.8%**	**1.8%**

R = Reported Model
B = Benchmark Model

7.6.7 As Table 7-9 shows the Benchmark model closely matches the Reported model with minor differences which are likely to be due to the differences between versions of the model than errors within the model runs.

7.6.8 Table 7-10 compares the Reported model against the Benchmark model for total vehicle kilometres elasticities in each region.

Table 7-10: Car Driver km elasticities by region (internal productions)

	HbW		HbEd		HbShopPB		HbRecV		HbHol		HbEB		NHbEB		NHbO	
Regn	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**	**R**	**B**
NE	-0.26	-0.27	-0.41	-0.42	-0.43	-0.44	-0.36	-0.38	-0.40	-0.43	-0.01	0.00	-0.19	-0.06	-0.24	-0.25
NW	-0.27	-0.29	-0.46	-0.48	-0.44	-0.44	-0.37	-0.41	-0.36	-0.40	0.02	-0.01	-0.19	-0.08	-0.24	-0.26
Y&H	-0.30	-0.31	-0.40	-0.42	-0.43	-0.43	-0.40	-0.42	-0.42	-0.45	0.01	0.00	-0.17	-0.05	-0.26	-0.25
EM	-0.29	-0.30	-0.39	-0.41	-0.42	-0.41	-0.42	-0.43	-0.40	-0.44	0.02	0.00	-0.15	-0.04	-0.26	-0.25
WM	-0.27	-0.28	-0.43	-0.44	-0.42	-0.41	-0.38	-0.40	-0.38	-0.41	0.00	0.02	-0.15	-0.04	-0.24	-0.24
EoE	-0.30	-0.31	-0.40	-0.43	-0.42	-0.43	-0.42	-0.45	-0.41	-0.46	0.01	-0.02	-0.16	-0.03	-0.26	-0.27
Lon	-0.23	-0.26	-0.40	-0.41	-0.47	-0.50	-0.36	-0.44	-0.38	-0.55	-0.01	-0.17	-0.25	-0.33	-0.20	-0.25
SE	-0.27	-0.29	-0.38	-0.41	-0.43	-0.44	-0.42	-0.46	-0.38	-0.44	0.05	0.01	-0.13	-0.03	-0.26	-0.28
SW	-0.31	-0.32	-0.36	-0.39	-0.39	-0.39	-0.41	-0.41	-0.42	-0.45	0.00	0.00	-0.16	0.00	-0.26	-0.23
IWa	-0.20	-0.24	-0.39	-0.42	-0.42	-0.45	-0.42	-0.45	-0.35	-0.40	0.03	-0.13	-0.10	-0.05	-0.28	-0.27
Total	-0.28	-0.29	-0.41	-0.43	-0.43	-0.43	-0.40	-0.43	-0.39	-0.44	0.02	-0.01	-0.17	-0.06	-0.25	-0.26

R = Reported Model
B = Benchmark Model

7.6.9 The Benchmark model is replicating the Reported model very closely with only minor differences shown in Table 7-10 except in London for HbHol and HbEB. As the Peer Review does not have access to the Reported Model the underlying cause of these differences has not been investigated.

7.6.10 Table 7-11 repeats the calculation of O-D base vehicle kilometres for the Reported and Benchmark models.

Table 7-11: O-D trip kilometre elasticities from highway assignment model matrices (all areas)

Assignment User Class	AM R	AM B	IP R	IP B	PM R	PM B
Car driver (vehicle) business	0.004	0.026	-0.049	0.003	0.019	0.027
Car driver (vehicle) commuting	-0.370	-0.404	-0.319	-0.362	-0.345	-0.390
Car driver (vehicle) other	-0.599	-0.639	-0.684	-0.722	-0.679	-0.735
Total	-0.370	-0.396	-0.544	-0.570	-0.452	-0.495

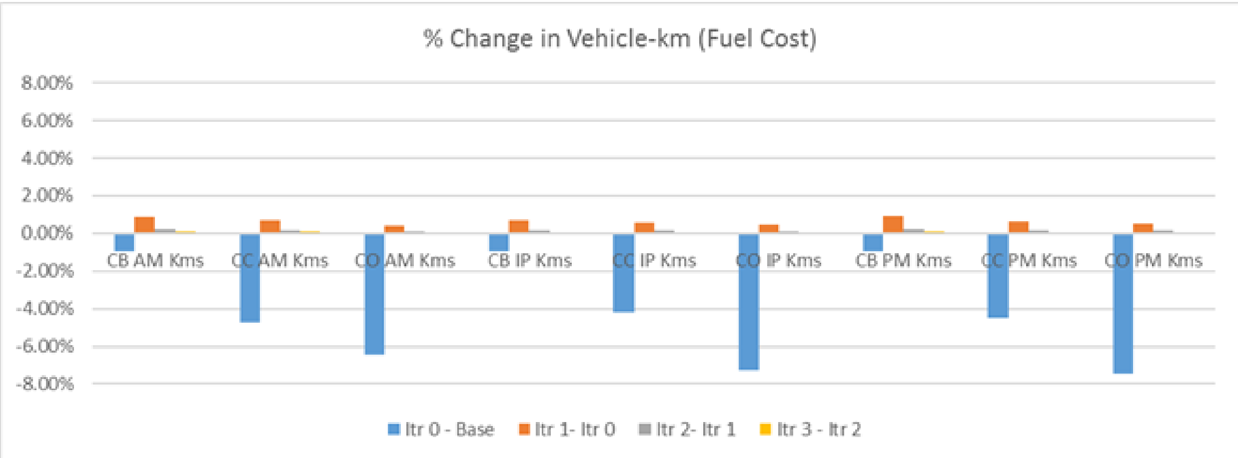
R = Reported Model
B = Benchmark Model

7.6.11 Table 7-11 shows that the calculated elasticities of the Benchmark model are marginally different from those of the Reported model. The direction and scale of the elasticity value is the same in every case except the Car driver business Inter Peak value which becomes marginally positive.

Model Stability

7.6.12 A review of the stability of the assignment models which measured the change between iterations in the total daily vehicle kilometres was undertaken to test the assertion that the model is stabilising after four iterations. This is presented in Figure 4 below.

Figure 4: Changes in Vehicle kilometres by Demand Model Iteration



7.6.13 As Figure 4 shows the difference between successive iterations reduces as more iterations are run which confirms that the model does become more stable.

7.7 Additional Sensitivity Tests

7.7.1 Three independent tests were designed by the Peer Review team to assess whether the model was performing as would be expected. These tests used the Benchmark Model and modified a different supply side component to identify how this would impact through the demand and assignment models.

7.7.2 Each of the tests was applied across the whole model, in contrast to the Sensitivity Tests which were undertaken during the model development phase. The purpose of applying tests nationally was to assess whether the modifications made to the supply side conditions had differential impacts in different regions or whether a similar scale and direction of change would be observed.

7.7.3 The supply side changes which were tested included modifying public transport costs, making highway travel faster and the making longer distance public transport trips less attractive.

7.7.4 As the purpose of these tests was to determine if such changes could be implemented within the model and to identify if the outcome changes in travel appear logical the results have no intrinsic value. Table 7-12 below summarises how the models were setup and assessed.

Table 7-12: Summary of Independent Sensitivity Tests

Test	PT Costs Reduced	Highway Costs Reduced	PT Deterrence Increased
Notes	The reference PT cost skim includes some very large values for trips which are not viable in the existing network which need to be treated with caution.	Coding approach of link types meant universal changes could be easily implemented	Increase in value of perceived travel time.
Expected main outcome	Large reductions across England in car driver / car passenger with large increase in rail and bus use. Largest changes where PT is more available (London and SE)	Reassignment of highway trips to more favourable routes and minor changes to mode choice.	Reduction in long distance public transport use and increase in car driver / car passenger trips. Redistribution away from PT dense areas.
Model outcome	Car driver national mode share falls from 44% to 26%, passengers from 19% to 8%. Rail share increases from 4% to 15%; bus from 8% to 38%. Redistribution of trips to London. Walk / cycle modes also show substantial decrease.	Assignment results car modes reassign to take advantage of reduced costs. HGVs second order response to reassign away from links made busier by cars.	Redistribution away from London and mode share moves away from rail and bus with majority to car-based modes.
Observations	Changes in mode share and redistribution are logical in scale and location.	Mode share changes are extremely small suggesting that PT use is captive to PT.	Logical responses across the model.

7.7.5 Table 7-12 summarises the outcomes of the model runs which were designed to stretch the model. The tests were implemented on the base year model and as such their performance on a forecast model has not been assessed.

7.7.6 The tests which have been undertaken all demonstrate that global changes to the model can be implemented and model runs through the demand / supply loop can be successfully operated.

7.7.7 The nature of these tests differed from those undertaken by the Model Development Team as they were more extreme and therefore even less plausible for implementation in the real world.

7.7.8 The purpose of making the tests more extreme was to confirm that the model did respond outside of the very marginal changes of the tests reported in Volume 6. The exercise confirmed that the model does allow more extreme outcomes to emerge.

8 Conclusions and Recommendations

8.1 General Observations about NTMv5 Documentation

8.1.1 The overall impression of the documentation is that it is accessible to the reader and covers the process of developing the model, its functionality and its implementation. The tiered structure does mean that there is considerable duplication of material and, in some cases as observed above, material appears in one document that might have been more or at least as appropriate to appear in another document.

8.1.2 Where the documentation was lacking tended to be in Developer Guide Volume 6 in which the reporting of results from the realism and sensitivity tests often meant that the results were obscured by rounding and aggregating. This meant that the Peer Reviewers were disadvantaged when trying to determine the impact of sensitivity tests and were only able to identify markers signifying there may be an issue without being able to isolate the cause.

8.2 General Observations about the Structure of the NTMv5

8.2.1 The model structure is set out in Figure 2.1 of the Quality Report, though this omits the pivoting process. It also does not make clear the distinction between home-based and non-home-based trips in the model (in particular, the NTEM growth applies only to Hb productions).

8.2.2 Given the decision not to deal with public transport capacity issues, and with the predominant focus on highway travel, we consider the general structure of a conventional transport model, containing modules representing (highway) assignment and multi-modal demand responses in terms of mode and destination choice, sensible, and the decision to align external changes in demand associated with land-use etc with the DfT's NTEM to be correct.

8.2.3 In relation to the key components of the highway assignment model and the demand model (VDM), we have a number of detailed comments, but overall we consider these modules to have been competently and carefully constructed and generally fit for purpose. We are also generally satisfied with the "external" components relating to other vehicles and other mode costs. We have much more concern about the pivoting process and the construction of the base demand.

8.2.4 In relation to the highway assignment model our key observations are: the lack of evidence that the model is producing speeds that accord with observation rather than reproducing the RTM SATURN speed-flow curves; the unstated rationale for the maximum speeds assumed for different vehicle types; the unverified use (in forecasting) of the relationship for urban areas between speed and total demand by all modes; and the use of average speeds in the fuel consumption relationships – with special reference to the implications for HGV routeing. We have also identified that the assignment results fall below the thresholds expected with TAG and as such the DfT needs to be satisfied that the model is fit for the purposes for which it is being used for each application.

8.2.5 In relation to the demand model our key observations are: the understating of the critical problems stemming from the lack of spatial detail for destinations in the NTS estimation data; issues associated with the destination constraint; the low value of time for rail; and the absence of evidence about the trip length distribution. In spite of these, we noted that the mean modal costs were well reproduced for all purposes and the first round elasticities were plausible.

8.2.6 In relation to the pivoting process, we have general concerns relating to the use of the period-specific highway assignment matrices as a pivot without any corresponding checks on the 24-hour demand matrices. This has a number of repercussions. Firstly, while the commute and education Hb purposes are constrained at the destination, the number of trips attracted to destinations for other purposes is subject to considerable potential error, and the NHb trips are dependent on these. This leads to even greater uncertainty when converting the (unconstrained) Hb purposes to O-D format and adding in the NHb trips. With the exception of the doubly constrained purposes, the comparison of the resulting trip length distributions with NTS data (Quality Report Appendix B) is not encouraging. Hence the application of the Sf/Sb ratio to the pivot highway matrices is far from robust.

8.2.7 This concern is exacerbated by the issues associated with the construction of the pivot matrices themselves (the description is one of the poorest parts of the model documentation). We have noted our concerns about the car modal splits which are applied to the “all modes” productions. While for the doubly constrained purposes the use of substantial external datasets (CJtW and the schools census) should produce reasonable estimates, the matrices for the remaining purposes are likely to be much less robust, and the NHb matrices are consequently likely to be based on different attractions from those predicted by the VDM. Further, little information is available as to the extent of the “matrix estimation” process, following the conversion to time-period O-D format. Finally, the level of mismatch seen in the future year test (Sensitivity Test 1) in terms of NTEM growth does not give confidence about the underlying quality of the base year Hb productions.

8.2.8 Overall, we are surprised at the lack of validation relating to a) the match between the 24-hour matrices which form part of the construction of the highway pivot matrices and the corresponding VDM-based matrices from the Base Year Run, or b) the match between the pivot matrices (after matrix estimation) and the corresponding Sb-matrices from the Base Year Run. Only if these could be shown to be reasonably aligned would we be willing to accept the pivoting procedure as robust.

8.2.9 It should also be noted that the fact that the pivoting process is confined to the car mode reduces the value of the model for producing forecasts for other modes. Here again, a greater level of validation for the Base Year Run against external data sources (for example, LENNON) could potentially increase confidence in the model's ability to represent the non-car modes.

8.3 General Observations about the Operation of the NTMv5

8.3.1 The use of VISUM for the NTMv5 means that it can be used more readily throughout the industry than previous versions. The substantial caveat is that the computing resources required are quite considerable and the licence size required is expensive.

8.3.2 Once those two issues are addressed, experienced users of the software demonstrated that NTMv5 could be successfully installed and operated to replicate the Reported Model, subject to minor changes made between the Reported Model and the Benchmark Model. Sensitivity tests were implemented successfully which provided confidence that NTMv5 can be used independently of the DfT.

8.4 General Observations about the Use Cases and NTMv5

8.4.1 As discussed in Chapter 6 above, the NTMv5 has been considered against five Use Cases which were defined by the DfT. The NTMv5 can provide some evidence for most of the use cases but for all of them it would need to be used with caution with the inputs for any test being carefully specified and the outputs would similarly need to be thoroughly reviewed.

8.4.2 The observations made by the Peer Reviewers are consistent with those documented in the Quality Report Section 2.4.

8.5 Potential Adjustments to NTMv5

8.5.1 Based on this review of the model structure and performance and on its potential usage for policy testing, we have assembled a set of recommendations to be considered for potential future enhancements to NTMv5.

8.5.2 These recommendations for NTMv5 are grouped by the time scale over which the enhancement tasks could be implemented, distinguishing: short term tasks that could be introduced relatively soon through minor adjustments to the model set-up or usage; medium term tasks that would require more substantial modifications to the model inputs and so might necessitate some limited adjustments to the model calibration; and longer term more fundamental changes in which the model structure, software, calibration or base matrix might undergo significant modifications so that a subsequent full validation and model testing exercise would then be appropriate. As these tasks fall in reality along a continuum of complexity and of resource and data requirements this allocation by time horizon is necessarily fluid.

8.5.3 Some of these NTMv5 tasks would also require external enhancements to the inputs from NTEM and from its underlying car ownership forecasting model, so these requirements are also outlined. Finally, some of the recommendations imply that changes to a small number of elements within the current TAG guidance would be beneficial so the underlying reasons for these suggested changes are explained. References back to where the enhancement topic has been introduced earlier in this review are denoted by section or paragraph numbers in square brackets, e.g. [1.1.1]

Immediate Model Checks and Adjustments

8.5.4 Investigate and **resolve** the source of the **systematic pattern of noise** in the results identified [3.6.22, 6.2] for the Sensitivity Test runs 2 to 5. Ensure that the model is always run to an adequate level of convergence and that the degree of convergence achieved is always published for each policy test run, using an appropriate cross-section of indicators. Success with this improvement could increase confidence in the resulting revised outputs from the Sensitivity Tests, which in turn could improve the capability for tackling some of the Use Cases.

8.5.5 **Assess the quality of the synthetic base matrix** that is the foundation for the VDM and of its match to the Base matrix and to patterns observed in the NTS. The understanding gained on the strengths and weaknesses of this match will aid in understanding and interpreting the results of policy measures being tested. This should be done on both a P/A and an O-D basis [8.2.8].

8.5.6 There are various further adjustments that are likely to be beneficial to the design and implementation of this synthetic base matrix. **Reconsider the decision to estimate the population segmentation within MSOAs** in the base year **using the Any Year Census (AYC) procedure.** This disaggregates population segments from the District level to the MSOA/zonal level based just on the property type mix in the MSOA [2.3.6]. Methods that start instead from the detailed population segmentation available from the 2011 Census at MSOA level, and then adjust this population through to 2015 using the AYC procedure, should provide more accurate results for most zones through lessening disaggregation error. **In particular improve the underlying car ownership pattern** in the base year, ensuring that it accords realistically with the pattern observed **in denser urban areas** [2.3.13, 4.3.5]. Realign the **population segmentation 16-64 and 65+ to match** to the current **NTEM** segmentation 16-74 and 75+, so that the current inconsistency for the 65-74 year age group between NTMv5 and NTEM is circumvented [4.1.9]. Finally, re-consider the form and implementation of the NHb models [8.2.9].

8.5.7 **Analyse** the soon to be published 2020 **DfT van survey** to obtain a better understanding of the current spatial patterns and of the trends through time for both freight and non-freight LGV trips and vehicle kilometres. Make adjustments as appropriate [2.3.29, 4.1.22] to the assumed balance of the growth for each of these components. Because no adequate LGV data source has been produced since the original DfT Van Surveys of 2003-05, the current empirical foundations within NTMv5 are necessarily weak: for the creation of the freight and non-freight LGV matrices; and for the relative rates of growth of each into the future. The results from this analysis and adjustment should help in assessing the ability of the model to forecast LGV growth and to assess LGV responses to policy measures being tested.

Medium Term Improvements to the Model Performance

8.5.8 **Switch throughout to use link-based VOCs** [2.1.48], rather than VOCs based either on system average speed in HAM or on O-D average speed in VDM. Because the assignment already updates the link times within each iteration, the additional updating of the link VOCs should not add appreciably to the computational burden. However, it may impact on the performance of the congested route optimisation algorithm or on the complexity of setting up existing packages to operate in this changed iterative fashion. Accordingly, as a fall back the initial link based costs could be left unchanged through these iterations, noting that this this fall-back would still greatly reduce the aggregation errors inherent in the original approach. In particular, if the assignment has a "warm start" based on a previous converged similar run, then these initial link costs would be expected to be very close to those at convergence. The further influence of VOC changes caused by adjustments in overall levels of congestion between successive model iterations could be applied within the VDM by a factor based on the ratio of the O-D specific speeds in the two previous iterations. This fall-back should not significantly increase current run times or generate problems with convergence. This change should significantly improve the performance of the choice of routes in the assignment for HGVs. This link based VOC approach would also have the additional advantage that the estimated fuel consumption within the VOCs should now align consistently with the environmental emissions calculations in the post-processing of the forecast vehicle flows on links.

8.5.9 Re-consider the pivoting process to see whether a) it could be additionally implemented on a P/A (24- hour) basis for the Hb purposes (as this would stabilise the application of the NHb purposes) and b) extended to other modes (such as rail) in order to improve the robustness of the model in forecasting non-car modes [???].

Longer Term Model Enhancements

8.5.10 **Reconsider the segmentation adopted within the spatial distribution model** [2.6.27] and ensure that it distinguishes realistically the differences in travel patterns between segments. In particular as part of the model estimation procedure, test for differences in deterrence parameters between a range of segments in order to ensure a good match to the observed clear differences in trip lengths patterns and destination zones: between HbEd movements of primary and secondary and other students [2.6.31]; as well as between HbW movements of groups of SICs, males/females and full-/part-time workers [2.6.28]. The focus within the model estimation procedure should be on obtaining the most appropriate model structure and parameters. Computing resource limits may imply that sequential, rather than simultaneous estimation of the model structure and parameters becomes necessary for those purposes that are doubly constrained, in order to ensure that this destination constraint is explicitly maintained within the estimation procedure. Provided that this sequential estimation is carried out appropriately, it is unlikely to significantly reduce the quality of the resulting model. It is better to estimate the best model form using sequential estimation while ensuring that the destination constraints are fully included, rather than to discard the destination constraints in order to make a simultaneous estimation methodology become feasible. A sequential estimation approach would also facilitate making effective use within the distribution model estimation of other data sources such as the Census Journey to Work matrices and the School Census data.

Recommendations for NTEM and NATCOP

8.5.11 **Improve the performance of the car ownership model in dense urban areas** to take appropriate account of the impact of densification on car ownership rates and trends [2.3.14]. Ensure that the spatial pattern of its forecast changes in car ownership rates across areas of different densities are broadly consistent with recent trends, except where there are clearly identified reasons for any forecast breaks in trends [4.3.5].

Recommendations for TAG

8.5.12 In the course of the review some aspects have emerged where the NTM developers have followed the current TAG recommendations but where this may have impacted on model performance. A number of current TAG recommendations could be reconsidered by DfT, as now listed.

8.5.13 For the reasons discussed above, the guidance in **TAG Unit M3.1 para 2.8.4** should recommend **VOC calculations** for all road vehicle types to be **based on the vehicle**

speed on individual links, rather than being based on the average speed by O-D or for the whole study area.

8.5.14 A related aspect within **TAG Unit M3.1 para 2.8.8** relates to the recommendation to double the driver's VOT for HGVs to "take account of the influence of owners on the routing of these vehicles". The more appropriate approach would be instead simply to apply the link based VOC in a form that takes full account of the operating cost of the vehicle on that link. The major elements of the HGV VOC should include: the driver's wages per hour of travel; the fuel costs at the estimated speed for that type of link; and the annual fixed vehicle operating costs factored per annual vehicle hour in movement. In this form, the influence of owners on HGV routing would be represented in a realistic fashion that is sensitive to any policy measures that could impact on VOCs. This approach should also be applied for LGVs on freight movements.

8.5.15 The accuracy of routing within assignments, as well as the calculations of fuel consumption, could be further improved by **revising the fuel consumption formula** specified in **Section 5 of TAG Unit A1.3 to be based also on link type** and not solely on vehicle speed - the current formulation. An important reason why long distance HGVs are concentrated onto the motorway network lies in their ability to achieve a constant speed there, which in turn generates relatively low levels of fuel consumption per kilometre. Uncongested motorway travel is more fuel efficient than driving on dual carriageways that generate regular deceleration/acceleration phases at roundabouts or traffic lights. Observed average HGV speeds on high quality roads that are significantly lower than the speed limit are often due to unavoidable variations in speeds caused by road conditions or road congestion and so these lower speeds will increase, rather than decrease fuel consumption rates. Travelling on uncongested high quality roads with a 40 mph speed limit would not create high fuel consumption rates, whereas travelling on a motorway also at 40 mph but now as a result of congestion would lead to very high rates of fuel consumption. Accordingly, the fuel consumption functions for a specific vehicle type on a link should take account of: its speed limit; its road type; and of course the speed achieved by the vehicle. The underlying data on which to construct improved the required fuel consumption formulae may already be available through reuse and analysis of relatively homogeneous sub-cycles within the wide range of drive cycles listed in the original TRL emissions modelling[38]. Individual sub-cycles representing: motorway (at various speeds), rural, suburban, urban, congested and various other traffic conditions are included in the 256 drive cycles analysed in that study. The benefit arising from this enhancement should both improve the realism of route assignments, primarily for HGVs, as well as improving the accuracy of the estimates of environmental costs created by vehicle emissions.

[38] Boulter, PG, TJ Barlow, S Latham and IS MacCrae (2009). “Emission Factors 2009: Report 1 – A Review of Methods for Determining Hot Exhaust Emission Factors for Road Vehicles” TRL Report for DfT, Published Project Report PPR 353, Version 7,June 2009. https://www.gov.uk/government/publications/road-vehicle-emission-factors-2009 accessed 06/05/20

9 Glossary

Terms and abbreviations used in this report:

AADF	Annual Average Daily Flow
AM	Morning modelled period, usually referring to the peak hour 8am to 9am
AMAI	Absolute Model Applied Incrementally
AUC	Assignment User Class
AYC	"Any Year Census" (Software)
BPR	Bureau of Public Roads (US)
CAV	Connected and Autonomous Vehicle
CJtW	Census Journey to Work dataset
CLoHAM	Central London Highway Assignment Model of TfL
CO	Car Other trip purpose
CPU	Computer Processor Unit
CSRGT	Continuing Survey of Road Goods Transport
CTripEnd	Software used for NTEM
DfT	Department for Transport
DO	(NTEM term: Day of week/time period Outward movement)
DR	(NTEM term: Day of week/time period Return movement)
FORGE	(Fitting on Regional Growth and Elasticities) aggregate highway assignment component of NTMv2
GBFM	Great Britain Freight Model
GDP	Gross Domestic Product
GIS	Geographic Information System
HAM	Highway Assignment Model
HbEB	Home based Employers Business
HbEd	Home based Education
HbHol	Home based Holidays
HbPB	Home based Personal Business
HbRecV	Home based Recreation and Visiting Friends and Relatives
HbShopPB	Home based Shopping and Personal Business

HbVFR	Home based Visiting Friends and Relatives
HbW	Home based Work
HGV	Heavy Goods Vehicle, >3.5 tonnes gross vehicle weight
HO	(NTEM term: Home-based purpose Outward movement)
HR	(NTEM term: Home-based purpose Return movement)
IP	Inter-peak modelled period, usually referring to the average of the hours between 10am and 4pm
IPP	Incremental Pivot-Point
IVT	In-vehicle Time
JTW	Journey to Work
LA	Local Authority
LAD	Local Authority District
LENNON	(Latest Earnings Networked Nationally Over Night) rail industry's ticketing and revenue system
LGV	Light Goods Vehicle, <= 3.5 tonne gross vehicle weight
LOS	Level of Service of the transport network
LU	London Underground
LUCE	Linear User Cost Equilibrium
MDST	MDST-Transmodal, developers of the GBFM model
MOIRA2	Updated rail industry Model Of timetable Impacts and Revenue Allocation
MoTiON	Model of Transport in LondON
MSA	Method of Successive Averages
MSOA	Middle layer Super Output Area
NATCOP	NATional Car OwnershP model
NHb	Non-Home Based trips
NHbEB	Non-home based Employers Business
NHbO	Non-home based Other
NST	(Eurostat) Nomenclature uniforme des marchandises pour les Statistiques de Transport
NTEM	National Trip End Model
NTMv2R	National Transport Model version 2 Revised

NTMv4	National Transport Model version 4
NTMv5	National Transport Model version 5
NTS	National Travel Survey
NUTS	Nomenclature des Unités Territoriales Statistiques (EU-wide Common classification of territorial units for statistics)
O-D	Origin to Destination matrix
OGV	Other Goods Vehicle- equivalent to HGV
OP	Off peak modelled period, usually referring to the average of hours between 7pm and 7am
OS	Ordinance Survey
OSQA	Output Specification and Quality Assurance plan
P/A	Production / Attraction
PCU	Passenger Car Unit
PLANET	Rail demand forecasting model
PM	Evening modelled period, usually referring to peak hour 5pm to 6pm
PRISM	Policy Responsive Integrated Strategy Model for the West Midlands
PSV	Public Service Vehicles
PT	Public Transport
QA	Quality Assurance
RTE	Road Traffic Estimates
RTF	Road Traffic Forecast, published by the Department for Transport https://www.gov.uk/government/publications/road-traffic-forecasts-2018
RTF18	Road Traffic Forecasts for 2018 by DfT
RTM	Regional Traffic Model, developed by Highways England
SATURN	Simulation and Assignment of Traffic in Urban Road Networks
Sb	Synthetic base year matrix of trips
Sf	Synthetic future year matrix of trips
SFC	Speed/Flow curve
SFR	Speed to Flow relationship
SIC	Standard Industrial Classification
SRN	Strategic Road Network

TAG	Transport Appraisal Guidance published by the Department for Transport
TfL	Transport for London
TIS	Highways England Trip Information System
TRACC	Multi-modal journey time calculation software from Basemap
TrafficMaster	TraffcMaster is a database of vehicle journey times and routes
UAS	Urban Area Speed - aggregate road speed adjustment procedure
UC	Use Case
V/C	Volume/Capacity Ratio
VDF	Volume/Delay Function
VDM	Variable Demand Model
VISUM	Transport modelling software of PTV
VOC	Vehicle Operating Cost
VoT	Value of Time

Kings Orchard
1 Queen Street
Bristol
BS2 0HQ

wsp.com