Oxfordshire Strategic Model (OSM)

WP 3 - RTM and PTM Specifications

Oxfordshire County Council

October 2013

Plan Design Enable

Notice

Document history

Job number: 5125364			**Document ref: WP3- RTM PTM spec-v1.1.docx**			
Revision	Purpose description	Originated	Checked	Reviewed	Authorised	Date
1	Draft Report	AEA		WW	AEA	18.10.2013
1.1	Draft Report	AEA		WW	AEA	21.10.2013

Table of contents

Tables

Figures

1. Introduction

1.1. Background

The County Council, Local Transport Body (LTB) and Local Enterprise Partnership (LEP) are currently working together to identify, appraise and prioritise potential transport interventions that could form part of Oxfordshire's Strategic Economic Plan (SEP) up to 2021.

A number of schemes have recently been identified as priorities by the LTB. There is a need to produce a fit for purpose transport model that can be used, post-SEP submission, by the LTB and LEP to provide guidance on the detailed scheme design and to produce the value for money elements at the three scheme business case stages.

Within these three stages, there needs to be particular emphasis on using model to identify the impact of transport and development in Oxfordshire. The model needs to be multi-model and WebTAG compliant to underpin specific requirements of the Department for Transport.

The model development involves the delivery of the following Work Packages:

WP 1 Data and Survey requirements
WP 2 Main Demand Model Specifications
WP 3 Road Traffic Model (RTM) and Public Transport Model (PTM) Specifications
WP 4 Study Objectives
WP 5 Calibration and Validation of: RTM
WP 6 Calibration and Validation of: PTM
WP 7 MDM Development
WP 8 Model Forecasting
WP 9 Appraisal Tools

This report covers the third element of the deliverables, WP3 giving the specification for developing the RTM and PTM.

1.2. Scope of Report

Following this introduction, this Model Specification Report consists of the following sections:

- Section 2 maps the use of the model to the key design considerations;
- The standards set for the RTM are discussed in section 3, and the key features of that sub-model are given in section 4;
- Similarly standards set for the PTM are discussed in section 5, and the key features of that sub-model are given in section 6; and
- Finally, the trip matrix development for RTM and PTM are outlined in sections 7 and 8 respectively.

2. Proposed Uses of the Model and Key Model Design Considerations

2.1. Background

The existing suite of strategic models available to OCC include an RTM in SATURN, a PTM in EMME3and a variable demand model also in EMME3. These form the Central Oxford Transport Model (COTM). There are also VISSIM models for Oxford City Centre as well as a number of junction models.

The Local Model Validation Report (LMVR) for the highway and public transport elements of the strategic model were submitted in August and September 2009. The models were calibrated and validated on 2005-06 data for a base year of 2007.

2.2. Key Model Design Considerations

The key considerations for developing the WebTAG compliant Oxfordshire Strategic Model (OSM) are to provide an evidence base for the appraisal of major highway and public transport schemes. The major interventions are principally around Bicester, Oxford, and the Science Vale corridor. The model needs to pay special attention to the A40 corridor between Whitney and J8 of the M40, as well as public transport and P&R.

Other considerations required by Oxfordshire County Council (OCC) for the model are that the run time should not exceed an overnight 16 hour period.

The principal objective of the OSM is to appropriately represent travel conditions on the highway and public transport networks for the appraisal of various schemes. The OSM should provide:

- changes in the travel cost between the base year and forecast years for input to the MDM;
- changes in traffic flows for input to the environmental appraisal of a scheme; and
- changes in travel costs for input to the economic appraisal.

The potential interventions for appraisal will relate to major highway improvements, large traffic management schemes, or large scale complex public transport schemes. The OSM should have the following capabilities:

- reflecting the impact of changes in land use policies, economic conditions and interventions on travel demand;
- testing for scenario development using less detailed modelling; and
- testing of schemes using more detailed modelling to be put forward for inclusion in funding programmes.

3. Model Standards – RTM

In this section we will discuss the standards that the RTM needs to achieve in line with WebTAG.

3.1. Validation Criteria and Acceptability Guidelines

Validation and convergence standards for highway assignment models are specified in TAG Unit 3.19. In general, the advice in TAG Unit 3.19 applies to models created for both general and specific purposes; however, in the case of models created for the assessment of specific interventions, '*it will be natural to pay greater attention to validation quality in the vicinity of the interventions*'.

The unit states that it is important that the fidelity of the underlying trip matrices is not compromised in order to meet the validation standards.

3.1.1. Trip Matrix Validation

For trip matrix validation, the measure which should be used is the percentage difference between modelled flows and counts. Comparisons at screenline level provide information on the quality of the trip matrices. TAG Unit 3.19 describes the validation criterion and acceptability guideline as shown in Table 3.1.

Table 3.1 – Screenline Flow Validation Criterion and Acceptability Guideline

Criterion and Measure	**Acceptability Guideline**
Differences between modelled flows and counts should be less than 5% of the counts	All or nearly all screenlines

Source: TAG Unit 3.19 Table 1

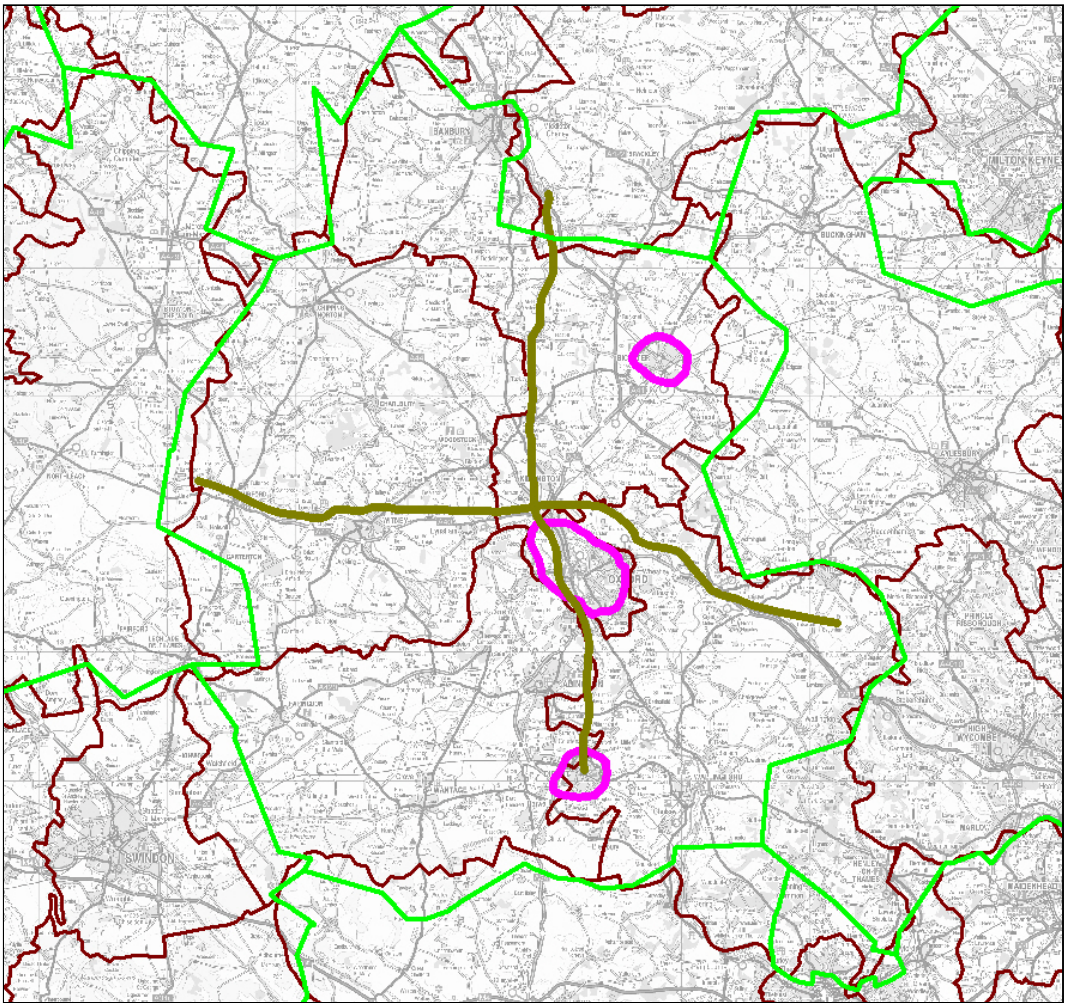
With regard to screenline validation, the following should be noted:

- screenlines should normally be made up of 5 links or more;
- the comparisons for screenlines containing high flow routes such as motorways should be presented both including and excluding such routes;
- the comparisons should be presented separately for
    - roadside interview screenlines where they exist;
    - the other screenlines (made up of ATC[1] for example) used as constraints in matrix estimation (excluding the roadside interview screenlines even though they have been used as constraints in matrix estimation); and
    - screenlines used for independent validation.
- the comparisons should be presented for all vehicles and for cars separately. This is because LGV and HGV do not usually form a large enough sample to make the data reliable; and
- the comparisons should be presented separately for each modelled period.

For this model the comparisons for RSI screenlines will focus on two screenlines, one for the East-West along A40 corridor and one for North-South along railway line between Didcot Parkway, Oxford and Banbury. Meanwhile, there are three calibration/validation cordons around Oxford, Bicester and Didcot. Figure 3-1 shows the locations of proposed screenlines and cordons in Oxfordshire. It should be noted that these screenline and cordons will be finalised at a later stage. In particular, the screenlines or cordons may need to split into a number of mini screenlines or cordons to facilitate model development in line with WebTAG guidance. The screenline or cordon counts will be formed by combining ATC's together, which may be classified. Otherwise MCC[2]'s at representative sites will be used to classify vehicles into cars, LGV and HGV.

[1] Automatic traffic count
[2] Manual classified count

Figure 3-1 Proposed Screenlines and Cordons for OSM



3.1.2. Link Flow and Turning Movement Validation

The two measures which should be used for the individual link (and turning movement) validation are flow and GEH. The flow measure is based on the relative flow difference between modelled flows and observed counts , with three different criteria set depending on the observed flows. The GEH measure uses the GEH statistic as defined below:

$$\textbf{GEH} = \sqrt{\frac{(M-C)^2}{(M+C)/2}}$$

where GEH is the GEH statistic

M is the modelled flow, and

C is the observed flow

TAG Unit 3.19 describes the Link Flow and Turning Movements Validation Criteria and Acceptability Guidelines as shown in Table 3.2.

Table 3.2 – Link Flow and Turning Movements Validation Criteria and Acceptability Guidelines

Criteria and Measures	Acceptability Guideline
Individual flows within 15% for flows from 700 to 2,700 veh/h	> 85% of cases
Individual flows within 100 veh/h for flows less than 700 veh/h	> 85% of cases
Individual flows within 400 veh/h for flows more than 2,700 veh/h	> 85% of cases
GEH <5 for individual flows	> 85% of cases

Source: TAG Unit 3.19 Table 2

With regard to flow validation, the following should be noted:

- the above criteria should be applied to both link flows and turning movements;
- the acceptability guideline should be applied to link flows but may be difficult to achieve for turning movements especially given the strategic nature of OSM covering the whole County;
- the comparisons should be presented for cars and all vehicles but not for light and other goods vehicles unless sufficiently accurate link counts have been obtained;
- the comparisons should be presented separately for each modelled period; and
- it is recommended that comparisons using both measures are reported in the model validation report.

No turning movements were counted for this model. The accuracy of the counts is not sufficient to enable flow and GEH criteria to be examined separately for light and other goods vehicles.

3.1.3. Journey Time Validation Criterion and Acceptability Guidelines

For journey time validation, the measure which should be used is the percentage difference between modelled and observed journey times, subject to an absolute maximum difference. TAG Unit 3.19 describes the Journey Time Validation Criterion and Acceptability Guideline as shown in **Error! Reference source not found.**.3.

Table 3.3 Journey Time Validation Criterion and Acceptability Guideline

Criterion and Measure	Acceptability Guideline
Modelled times along routes should be within 15% (or 1 minute, if higher)	> 85% of routes

Source: TAG Unit 3.19 Table 3

With regard to the journey time validation, the comparisons should be presented separately for each modelled period.

There was no disaggregation of journey time data to enable validation by vehicle type and a single speed/flow relationship was applied to all vehicle types so the validation will be performed for total vehicles only.

Journey times will be obtained from TomTom data that uses SatNav technology.

3.1.4. Impact of Matrix Estimation

The screenline (or cordon) comparison between modelled flows and counts is used to demonstrate the quality of the trip matrices by checking the overall volumes of trips across the modelled area. The changes introduced by the application of matrix estimation should be understood and may be assessed using TAG Unit 3.19 (Table 5), as shown in Table 3.4 below.

Table 3.4 Significance of Matrix Estimation Changes

Measure	Significance Criteria
Matrix zonal cell levels	Slope within 0.98<Slope<1.02 Intercept near zero R^2 in excess of 0.95
Matrix zonal trip ends	Slope within 0.99<Slope<1.01 Intercept near zero R^2 in excess of 0.98
Trip length distributions	Means within 5% Standard deviations within 5%
Sector to sector level matrices	Differences within 5%

Source: TAG Unit 3.19 Table 5

The unit states that it is important that the fidelity of the underlying trip matrices is not compromised in order to meet the validation standards. All exceptions to these criteria should be examined and assessed for their importance for the accuracy of the matrices in the Fully Modelled Area.

The comparisons should be presented by vehicle type (preferably cars, light goods vehicles and other goods vehicles). The comparisons should also be presented separately for each modelled period or hour.

3.2. Convergence Criteria and Standards

The advice on model convergence is set out in TAG Unit 3.19 (Table 4) and is reproduced below in Table 3.5. A more stringent set of standards may be achieved for the RTM with a target of 99% of links satisfying the convergence measure rather than suggested 98% of links.

Table 3.5 Summary of Convergence Criteria

Convergence Measures	Type	Base Model Acceptable Values
Delta & %GAP	Proximity	Less than 0.1% or at least stable with convergence fully documented and all other criteria met
Percentage of links with flow change[2] (P1) < 1%	Stability	Four consecutive iterations greater than 98%
Percentage of links with cost change (P2) < 1%		Four consecutive iterations greater than 98%
Percentage change in total user costs (V)		Four consecutive iterations less than 0.1% (SUE only)

Source: TAG Unit 3.19 Table 5

3.3. Interpretation of the Guidelines

TAG Unit 3.19 states that the achievement of the validation acceptability guidelines specified in Table 1, Table 2 and Table 3 (of TAG Unit 3.19) does not guarantee that a model is ‘fit for purpose’ and likewise a failure to meet the specified validation standards does not mean that a model is not ‘fit for purpose’. Furthermore, in some models, particularly models of large congested areas, it may be difficult to achieve the link flow and journey time validation acceptability guidelines set out in Table 2 and Table 3 (of TAG Unit 3.19) without matrix estimation bringing about changes greater than the limits shown in Table 5 (of TAG Unit 3.19). In these cases, the limits set out in Table 5 (of TAG Unit 3.19) should be respected, the impacts of matrix estimation should be reduced so that they do not become significant, and a lower standard of validation reported. In other words, matrix estimation should not be allowed to make significant changes to the prior matrices in order that the validation standards are met.

4. Key Features of the RTM

4.1. Base Year

The OSM will have a 2013 base year and represents the travel conditions for a typical April weekday. The RTM will be developed by using the latest SATURN version 11.2.05, to take advantages of up-to-date software improvements.

4.2. Modelled Area

TAG Unit 3.19 states that the geographic coverage of highway assignment models generally needs to: allow for the strategic re-routeing impacts of interventions; ensure that areas outside the main area of interest, which are potential alternative destinations, are properly represented; and ensure that the full lengths of trips are represented for the purpose of deriving costs. The modelled area therefore needs to be large enough to include these elements, but within the modelled area the level of detail should vary as follows:

- **Fully Modelled Area**: the area over which proposed interventions have influence, and in which junctions are in SATURN simulation, is further subdivided as:
  - **Area of Detailed Modelling** – the area over which significant impacts of interventions are certain and the modelling detail in this area would be characterised by: representation of all trip movements; small zones; very detailed networks; and junction modelling (including flow metering and blocking back).
  - **Rest of the Fully Modelled Area** – the area over which the impacts of interventions are considered to be quite likely but relatively weak in magnitude and would be characterised by: representation of all trip movements; somewhat larger zones and less network detail than for the Area of Detailed Modelling; and speed/flow modelling (primarily link-based but possibly also including a representation of strategically important junctions).
- **External Area**: the area where impacts of interventions would be so small as to be reasonably assumed to be negligible and would be characterised by: a SATURN buffer network representing a large proportion of the rest of Great Britain, a partial representation of demand (trips to, from and across the Fully Modelled Area); large zones; skeletal networks and simple speed/flow relationships or fixed speed modelling.

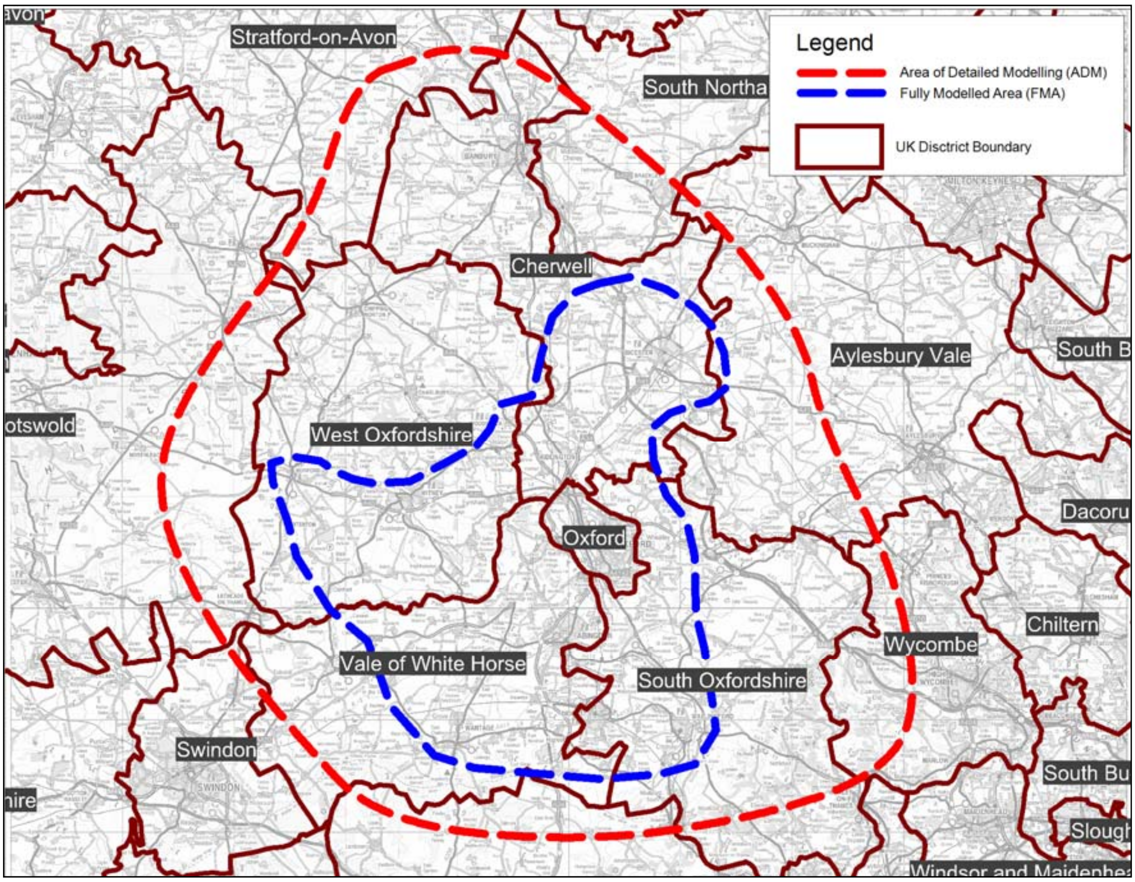
In the OSM the Area of Detailed Modelling (ADM) covers the area bounded by:

- Bicester to the north;
- Wallingford to the east;
- Burford and Witney to the west; and
- Wantage and Didcot to the south.

The Fully Modelled Area (FMA) covers the rest of Oxfordshire plus some hinterland area including Banbury, Swindon, Reading, High Wycombe and Stratford-upon-Avon etc. Figure 4-1 shows the extent of ADM and FMA for OSM.

The External Area covers the rest of Great Britain in a skeletal form and the relationship between the ADM and FMA.

Figure 4-1 Area of Detailed Modelling and Fully Modelled Area for OSM



4.3. Zoning System

The RTM is part of an integrated modelling suite, which links the MDM[3] to both the highway assignment and public transport assignment models. The existing zoning systems was going to build on COTM but after reviewing it, it was found that the zone structure is not fully compatible with TEMPRO or UK Census Output Areas, which contains demographic information such as number of households etc. As a result, a complete zoning system was developed for OSM. A separate working paper will be issued detailing the development of zoning system. In particular, for zones in Oxfordshire, they are aggregated from UK census Output Area zones, attempting to have less than 500 households per zone. This ensures that zones are fine enough in the core study area, with no zone generating a demand of more than 300 pcu[4] per hour, following the latest industry best practice.

The resulting number of zones by area is shown in Table 4.1. In total, there are 716 zones covering the whole of Great Britain, with 567 zones falling in Oxfordshire. In particular, all the five P & R sites and major car parks in Oxford have been given specific zones, together with two separate airport zones for Heathrow and Gatwick. It is also estimated that a total of 50 dummy zones will be added later to allow for future development proposals. The zoning system for OSM will be common for all RTM, PTM and MDM components.

Figures 4-2, 4-3 and 4-4 show zones in the major towns of Oxford, Bicester and Abingdon/Didcot respectively. Banbury has also been split from a single zone in the previous COTM into 24 zones, as shown in Figure 4-5. Figure 4-6 presents the whole zone system in Oxfordshire. The rest of the zones in the hinterland and external area is shown in Figure 4-7.

We will produce a sectoring system which will be of a reasonable size to be viewed easily and be a function of the propensity of travel. The sector will be probably in line with the areas identified in the table above, but

[3] Main demand model
[4] Passenger car unit

with Oxford divided into a number of areas such as CBD, inner and outer, ending up with around 10-11 sector system.

Table 4.1 OSM Zoning System

Area	No. of Zones
Oxford	131
Didcot / Wallingford / Wantage	42
Bicester	25
Abingdon	30
Witney	25
Banbury	24
Rest of Oxfordshire	290
Hinterland	113
Rest of UK	36
Total	716

Figure 4-2 OSM Zones in Oxford

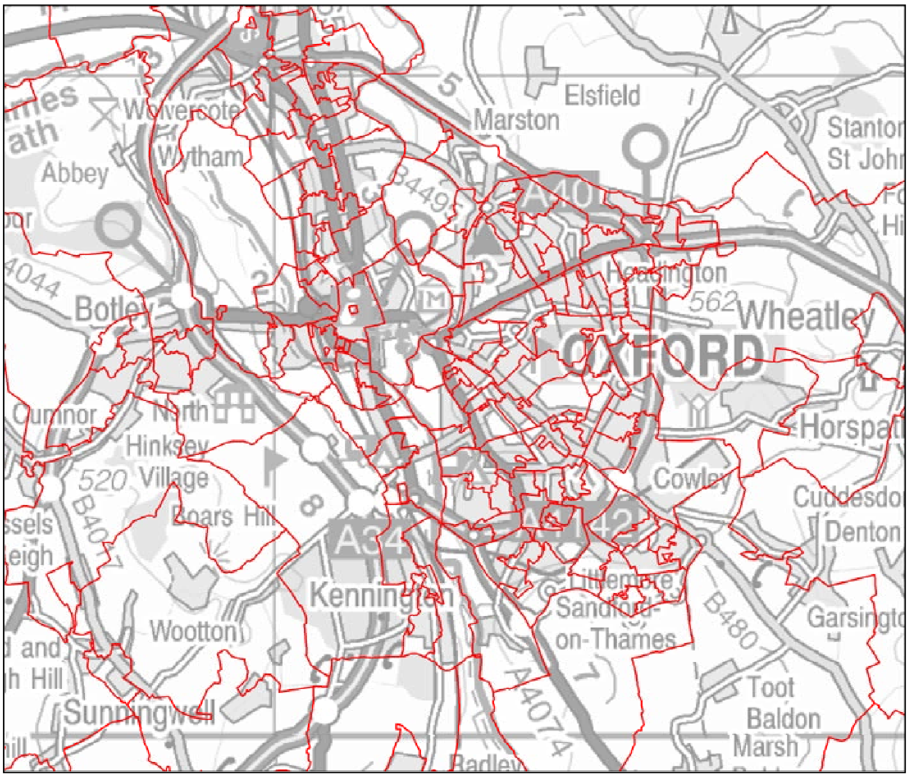


Figure 4-3 OSM Zones in Bicester

Figure 4-4 OSM Zones in Abingdon and Didcot

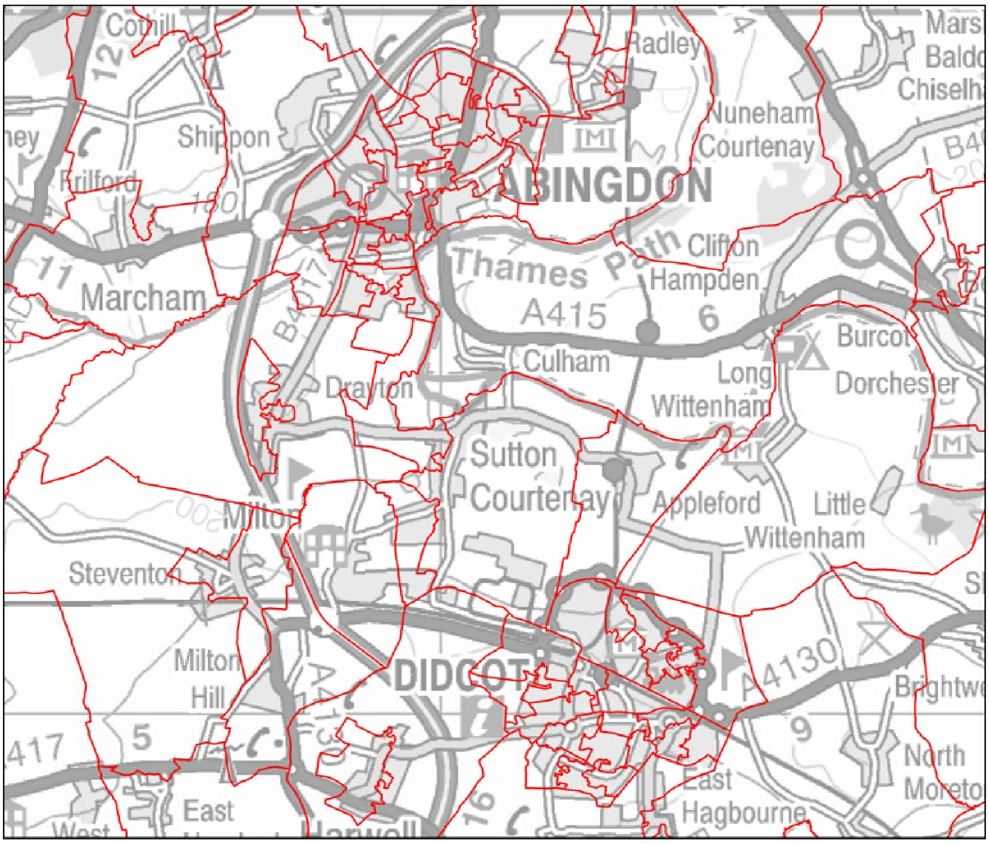


Figure 4-5 OSM Zones in Banbury

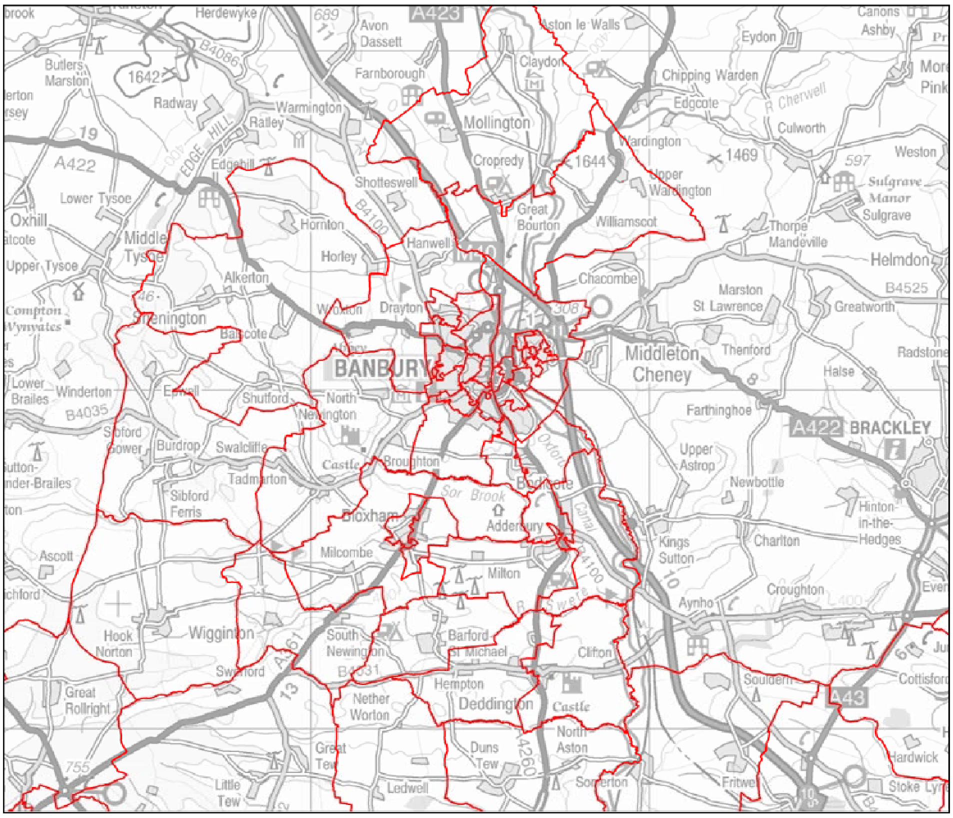


Figure 4-6 OSM Zones in the whole of Oxfordshire

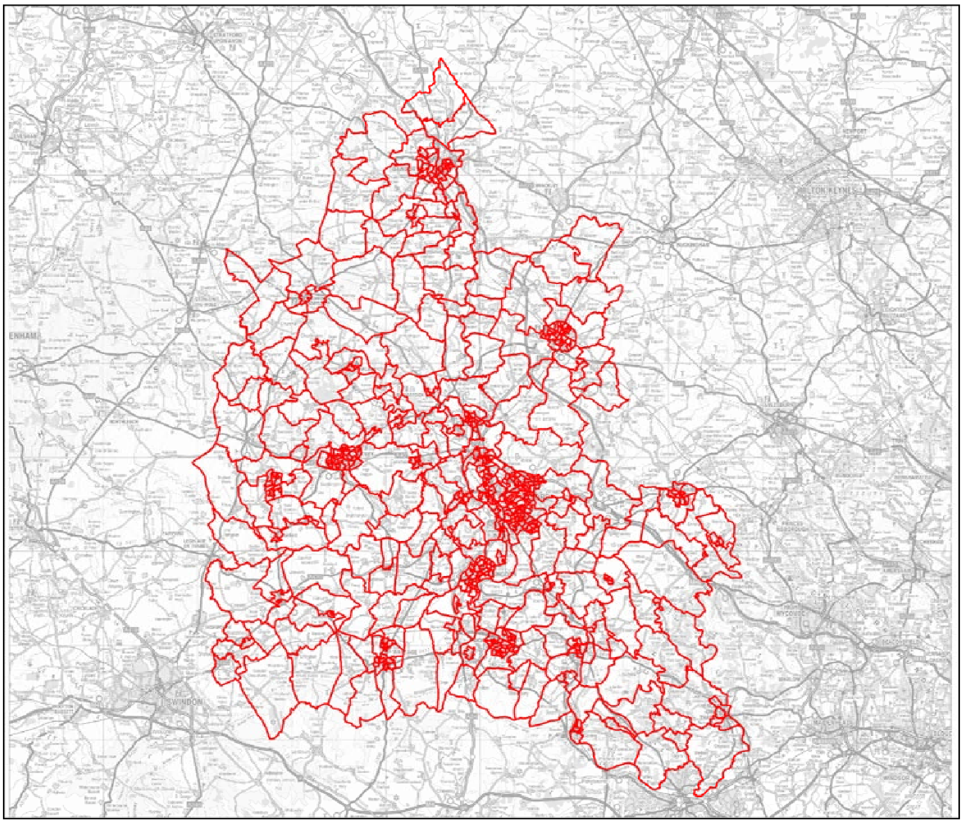
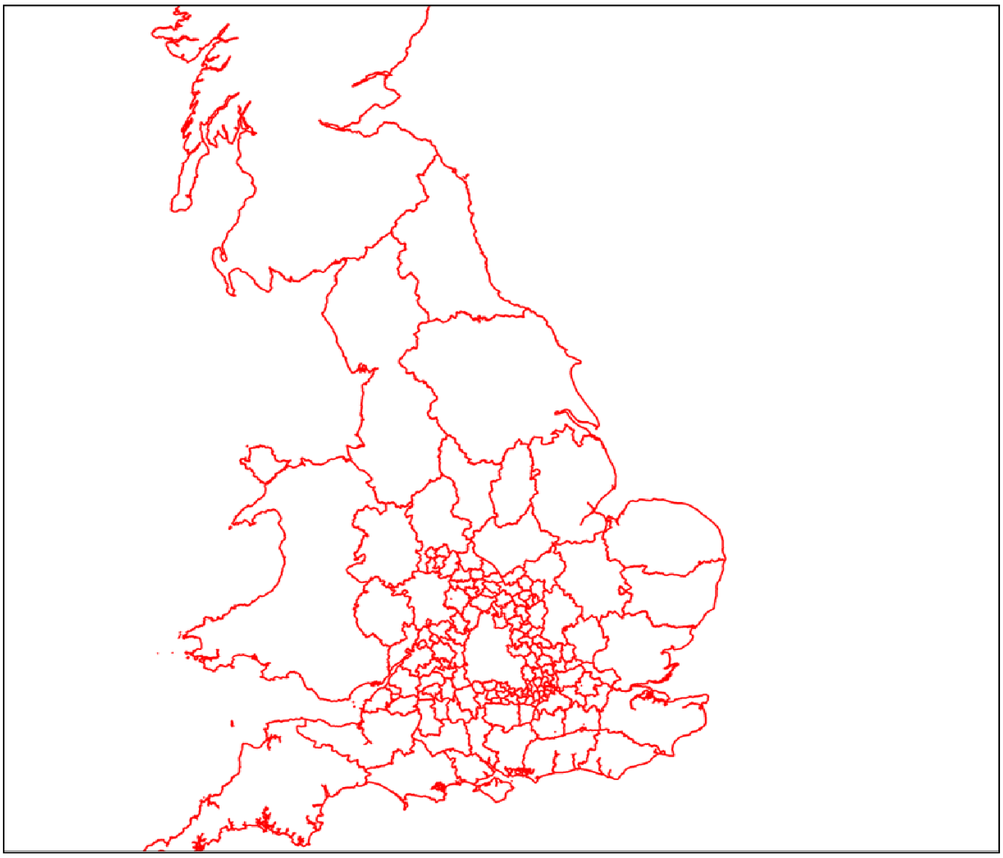


Figure 4-7 OSM Zones in the Hinterland and External Area



4.4. Network Structure

The network structure was developed with COTM as the starting point. The density of the network structure differed between the FMA and External Area as follows:

- within the FMA, all major A-road, B-roads and motorway links were represented along with the main residential roads and access roads to major developments and car parks; whereas
- the External Area only included the major A-roads, B-roads and motorway networks with reducing detail further away from the FMA.

The FMA was coded in the SATURN simulation network (with explicit junction modelling) whilst the External Area was coded in SATURN buffer network. The level of detail and accuracy of the network decreases as progression is made from the ADM to the External Area.

4.4.1. Link Coding

The link coding includes link length and road standard. The link lengths of roads were based initially on COTM and revised as necessary subject to measurements taken from GIS and Google Earth. Within the FMA the links were classified by road type and designated speed limit. For the buffer network the standard Cost Benefit Analysis (COBA) definitions are applied.

4.4.2. Signal Timings

OCC will provide current signal timing data updated for the existing and newly designated junctions. These will be entered into the SATURN junction coding.

4.4.3. Link Speeds

The link speeds in the OSM will be coded using TomTom journey time data disaggregated by road type for the hours 7pm to 7am to reflect the cruise speed as defined in TAG Unit 3.19:

> *"Cruise Speed - the speed of traffic on links between queues at modelled junctions. The cruise speed is dependent on the attributes of the link and activity levels alongside and crossing the link. It is not related to flow to any significant degree and is not necessarily equal to the speed limit".*

The cruise speed is applied to all links within the ADM based on link classification. The cruise speeds were maintained at the same level in all time periods. The fifth percentile speeds across the different time periods were very similarly for all routes indicating little change in conditions affect cruise speed during the modelled peaks.

The centroid connectors enable the zones to be attached on to the link network. The centroid connectors are coded with:

- specific entry / exit junctions from local access roads onto the main road network from self-contained residential areas, business parks, retail areas and car parks for example; or
- selected junctions representing multiple access points (i.e. removing the need to explicitly code every junction on each link).

Judgement is used to determine the number of centroid connectors required from each zone to represent locations where the traffic from the zones was likely to load in reality, using as many or as few zone connectors as was considered appropriate.

4.5. Time Periods

The RTM represents three time periods, namely the morning and evening peak hours and an average inter-peak hour. The three periods explicitly modelled are:

- Morning Peak hour 08:00 – 09:00;
- Average Inter-Peak hour 10:00 – 16:00; and
- Evening Peak hour 17:00 - 18:00.

For the morning peak and evening peak hour, a previous shoulder peak period is also modelled (although this is not separately validated), and queues which build up during this period are carried over to the start of the peak hour using the SATURN PASSQ option.

4.6. User Classes

The RTM represents highway demand with three user classes as detailed below:

- cars;
- light goods vehicles; and
- heavy goods vehicles.

Scheduled local bus services are represented separately and will be discussed the PTM development.

The trip purposes used for car will be made up of the following six purposes in line with WebTAG (unit 3.10.3 par1 1.11.12 for example):

- home-based work – commute HBW;
- home-based employers' business HBEB;
- home-based other – shopping, leisure, personal business HBO;
- non-home-based employers' business NHBEB; and
- non-home-based other NHBO.

4.6.1. PCU Factors

The RTM uses passenger car units (pcus) rather than vehicles as its standard unit for demand and capacities. This allows the effects of longer/slower vehicles that occupy more road space and take longer to clear junctions to be represented. The conversion factors used for the various vehicle types are summarised below in 0.

Table 4.2 Vehicle to PCU Conversion Factors

Vehicle Type	Equivalent PCUs	Comment
Car	1.0	Private cars
LGV	1.0	Goods vehicles using car-based chassis
HGV	2.3[5]	For both OGV1 & OGV2 vehicle types
PSV / Bus	3.0	Scheduled coach and local bus services

Note: All demand matrices used in the assignment represented demand in pcus per hour rather than vehicles.

4.7. Assignment Methodology

The RTM will use SATURN assignment software. SATURN uses the SATALL module to iterate between successive loops of SATASS module (which assigns the user class matrices to the network in accordance with Wardrop's First Principle of Traffic Equilibrium using the Frank-Wolfe algorithm) and SATSIM module (which takes the flows derived by SATASS and calculates the revised flow/delay relationships at each junction within the simulated area) until the resulting travel times and flows do not change significantly (that is, the process has 'converged').

The process starts with SATASS using the free-flow times (without any delays arising from vehicle interactions at the simulated junctions) from the network building program, SATNET. After the first set of path-builds in SATASS, the resulting flows are passed to SATSIM for the turn-based flow/delay curves representing the detailed interactions at each junction to be updated. These revised flow/delay relationships are passed back to SATASS for the travel time and flows to be recalculated. Further details may be found in the SATURN User Manual.

In order to cut down on the assignment run time, it is proposed that the SPIDER network function, a new feature in the latest SATURN release, will be adopted.

4.8. Generalised Cost Formulations and Parameter Values

The route choice within RTM is modelled using the generalised cost of travel time, vehicle operating cost and tolling / congestion charging in accordance with the TAG Unit 3.19, section 2.8. This is so to make it compatible with the MDM which also uses generalised costs. The coefficients for the individual components of generalised costs were calculated using TAG Unit 3.5.6 (October 2012).

The model base year is 2013 with all monetary values calculated at 2010 prices.

4.8.1. Values of Time

Perceived values are used throughout. Note that, in the case of HGVs, and cars and LGVs in work time, the perceived and resource values are the same. The process is summarised below:

- equivalent 2013 values are calculated by applying the specified growth in working and non-working values of time (Table 3 in TAG Unit 3.5.6);
- the relative proportions of Car Non-work for 'Other' and 'Commuting' are calculated from the RSI surveys;
- the equivalent values for vehicles were calculated by applying the occupancies obtained from the RSI surveys;
- HGV travel is assumed to be in work time with the split between OGV1 and OGV2 recorded from the RSI surveys; and
- the values are converted from £ per hour to p/min.

[5] TAG Unit 3.19c provides two pcu values for HGVs: either 2.3 pcus for motorways and all-purpose dual carriageways or 2.0 pcus for all other road types.

4.8.2. Vehicle Operating Costs

Vehicle Operating Costs are calculated using TAG 3.5.6 (October 2012) and defined separately for fuel and non-fuel elements before being combined for the use in the SATURN assignment. Non-fuel costs are only taken into consideration by travellers in work-time.

4.8.2.1. Fuel Costs

The consumption of fuel (in litres per km), adjusted by the fuel efficiency factors, is multiplied by the cost per litre to provide the cost per km in the model base year (2013). Fuel duty is included in the calculations as a perceived cost as businesses are not able to reclaim it. However, VAT is excluded because businesses are able to recover it. For non-work purposes, the perceived cost of the fuel Vehicle Operating Cost was the market price. LGV fuel costs were derived using the same work/non-work proportions used to calculate their average Value of Time.

4.8.2.2. Non-Fuel Costs

The non-fuel cost element is derived using the formulae set out in TAG 3.5.6 Table 15 and is a function of average network speed. No further adjustments are required as the non-fuel costs are assumed to remain constant, in real terms, over time. As noted above, the non-fuel cost element is only included for work trips.

4.8.3. Assignment Parameters

The resulting cost coefficients of PPK and PPM will be calculated later and presented in the RTM LMVR report.

4.9. Capacity Restraint

Capacity restraint is modelled in the FMA (i.e. simulation area) predominantly through junction modelling. All modelled junctions in this area are allocated a junction type, capacities for each turn, lane allocations and traffic signal timings for roundabouts and signalised junctions respectively. The capacity of a link is therefore determined by the junction arm capacities.

4.10. Network Calibration

A number of checks will be undertaken on the network coding including:

- reviewing the warnings produced by SATNET, the SATURN network building software;
- the coded link distances versus crow-fly distances;
- coded link speeds and speed-flow curves; and
- coded junction saturation flows.

Link distances are compared to crow-fly link lengths and those greater than 1.3 times the crow-fly distance will be inspected. During the model calibration changes, the following will be checked:

- **Counts in excess of capacity** – where an observed count is noticeably higher than the coded network capacity the capacities will be checked and amended if necessary;
- **Excessive junction delays** – the largest overall delays, and the largest differences between the link travel times and the moving car observer data are checked and junction coding checked;
- **Low flows** – where the modelled flow is substantially below that counted; this reveals locations where traffic was either restricted at an upstream junction or where a competing route was more attractive; and
- **Poor reproduction of observed travel times** - detailed comparisons of modelled travel times against the observed journey time routes reveal locations where additional modifications to signal settings may be necessary in order to replicate the observed levels of delay.

4.11. Route Choice Calibration and Validation

The accuracy of the assignment depends on the network structure, the trip matrix and the realism of modelled routes. The calibration and validation of the routes chosen by the model will be carried out as follows:

4.11.1. Route Choice Calibration

The ability of model to robustly represent route choice within the network depends on:

- correct zone sizing and definition, network structure and the realism of the zone connections to the modelled network (centroid connectors);
- the accuracy of the network coding and the appropriateness of the simplifications adopted;
- the accuracy with which delays at junctions and link cruise speeds are modelled, which in turn is dependent not only on data and/or coding accuracy but also on the appropriateness of the approximations inherent in the junction flow/delay and link speed/flow relationships; and
- how accurately the trip matrices have been built, which, when assigned, will impact on the route choice process (via the flow/delay and speed/flow relationships).

During the route choice calibration process, any issues such as these, which arise from incorrect or doubtful route choices, will be examined in detail, and where appropriate corrections/changes to the junction coding will be implemented.

4.11.2. Route Choice Validation

No specific criterion exists for validating route choices within a modelled network. However, it is common practice to undertake to review the routing chosen by the model between key locations and TAG Unit 3.19 suggests that the number of routs (OD pairs) should be estimated as:

Number of OD pairs = $(\text{number of zones})^{0.25}$ x the number of user classes

This equates to approximately 16 OD pairs for route validation for each modelled time period.

4.12. Links with MDM and PTM

The RTM is fully integrated within the MDM and PTM modelling system. The RTM provides highway transport costs to the MDM, which in turn provides updated trip matrices for the RTM.

The PTM is closely integrated with the PTM as they share the same network and zoning systems. This common structure enables the automated transfer of link and turn time data from the RTM to the PTM.

The bus services represented in the PTM are automatically transferred to the RTM to ensure that the impact of buses on other highway traffic is taken into account. However, the zone centroid connectors are not shared between the two assignment models, reflecting the different access points to the two networks.

5. Model Standards - PTM

5.1. Validation Criteria and Acceptability Guidelines

As indicated in the public transport calibration guidelines in TAG Unit 3.11.2, the PTAM validation includes:

- validation of the trip matrix;
- network and service validation; and
- assignment validation

5.2. Trip Matrix Validation

WebTAG Unit 3.11.2, para 12.3.2 states that

"Matrix level validation should involve comparisons of assigned and counted passengers across complete screenlines and cordons (as opposed to individual services). At this level of aggregation, the differences between assigned and counted flows should in 95% of the cases be less than 15%."

It should be possible to complete a bus trip matrix validation for the PTM across screenline and cordon counts. No equivalent validation can be carried out for rail as rail passenger patronage surveys along key corridors will not be conducted under the current modelling scope.

5.3. Network and Service Validation

The PTM bus network will be identical in structure to the validated RTM. The routes, stopping pattern and service frequency will be coded with the help of CIF files provided by the bus operators through OCC, together with published bus timetable from individual bus operators.. The speeds for each route will be extracted from the equivalent RTM network after applying a factor to allow for dwell times at bus stops. This factor will be based on judgement from other studies and the implied journey times will be compared to bus timetables.

Checks on the accuracy of the coded network geometry will be carried out to ensure the correct frequency has been coded. The coding of bus services will be verified by checking the modelled flows of buses by route against the roadside bus count data.

5.4. Assignment Validation

TAG Unit 3.11.2, para 10.1.6 states that:

"Across modelled screenlines, modelled flows should, in total, be within 15% of the observed values. On individual links in the network, modelled flows should be within 25% of the counts, except where the observed flows are particularly low (less than 150)."

Some of the observed link counts that will be collected may have flows less than 150. In order to give some measure of the fit of the model to counts less than 150, we will calculate the GEH statistic, as defined in the RTM section 3.1.2. A GEH of less than 5 indicates a good fit of the modelled link flow to the observed count.

Whilst WebTAG does not specify an overall objective for the calibration/validation, we will aim to achieve 85% of links meeting the criteria.

5.5. Bus Assignment Validation

For the bus assignment validation, onboard bus occupancy counts will be collected at proposed key corridors around Oxford, Bicester and Didcot. The counts are disaggregated by time period and bus service. Count comparisons were therefore made at both the overall link and bus service group level.

5.6. Rail Assignment Validation

For the rail assignment validation, (single day) boarding and alighting counts will be collected at Oxford, Didcot Parkway and Bicester North rails stations, supplemented by rail LENNON ticketing data from rail operators.

As with the link flow validation, we have adopted the criterion that modelled boardings and alightings should be within 25% of the counts, except where observed flows are less than 150. We have also reported the GEH statistic as a further guide to the degree of fit of the model to the data.

5.7. Journey Time Validation

The rail assignments are based on timetabled journey times and hence journey time validation is not necessary. With bus, as stated earlier, the implied journey time by applying a factor to link times from the RTM will be compared to bus timetable data.

6. Key Features of the PTM

In this section, mention will only be made where it differs from the key features of the RTM

6.1. Network Structure

The PTM has been developed to represent two public transport modes:

a) bus; and

b) rail.

In addition, the model also includes a bus-based park and ride mode and the performance of the park and ride sub-model is separately reported within the MDM specification.

Separate provision has been reserved for new modes such as LRT and BRT[6], and the assignment procedures allow the flexibility of integrating the new modes into the MDM.

6.2. User Classes

The public transport assignment uses a single user class.

6.3. Assignment Methodology

The Public Transport Assignment Model uses the standard transit assignment implemented in EMME, i.e. a multipath assignment, based on the computation of optimal strategies. Further details of the assignment methodology may be found in the EMME reference manual.

6.4. Generalized Cost Formulations and Parameter Values

The generalised cost function used for the public transport assignment routing, measured in units of time (minutes), is given by:

$$G_{PT} = V_{wk}*A + V_{wt}*W + T + B$$

where:

V_{wk} is the weight applied to time spent walking (walk time weight);

A is the total walking time to and from the services;

V_{wt} is the weight applied to time spent waiting;

W is the total waiting time for all services used on the journey;

T is the total in-vehicle time; and

B is the total boarding penalty applied for each service boarded on the journey

The public transport assignment model uses the parameters based on those provided in WebTAG Unit 3.11.2, which in turn are derived from work undertaken by Institute of Highways and Transportation to establish guidelines for urban transport strategies and further work commissioned by the DfT on the value of travel time savings. Further details, including the various references, may be found in the WebTAG Unit.

The parameter values for assignment are set out below in Table 6.1. In the EMME assignment, the modelled wait time is controlled by the ‘wait time factor’ of 0.5, indicating that the wait time is set at half the service headway.

The assignment parameters can be changed to improve routing calibration.

[6] Light rail transit and bus rapid transit

Table 6.1 – Assignment Parameters

Parameter	Value
Wait time factor	0.5
Wait time weight	2.5
Walk time weight	2.0
Interchange penalty	5 to 20*

** Sourced from WebTAG and adjusted as part of the calibration process.*

6.5. Fares

The public transport sub-mode choice (ie P&R v BRT v Bus v Rail) is undertaken within the MDM based on the standard WebTAG generalised cost formulation (which includes fares). The PTM (assignment) does not consider the impact of fares. The PTM determines the route choice (within each mode) and whilst there will be some influence of fares, it is unlikely to be significant, because:

- Bus Services in Oxfordshire are provided principally by Stagecoach, Oxford Bus Company and Thames Travel. Typically a competitive stage based fare system with a range of day and seasonal ticket types is provided by each bus operator, which limits passenger's choice to choose alternative routes in order to reduce fare costs. Meanwhile rail fares are distance-based and the park and ride mode has a flat fare system;
- The choice of route is sensitive to the difference in the total cost of the journey not the absolute cost and the influence of fare is small compared to the weights attached to In-Vehicle Time, Wait Time and Interchange penalties;
- There are several ticket types such as day returns and season tickets which are purchased independent of route choice; and
- The fare differentials between realistic competing routes for the same O-D pair will be small.

6.6. Bus Journey Times

The bus network is created from the SATURN RTM. This enables a linkage to be established between highway travel times and bus travel times such that, in forecasting mode, the impact of increasing congestion levels on bus travel times is represented.

This linkage also allows the impact on bus journey times of new bus lanes and bus priority measures at junctions to be modelled. At the same time, it models the effects of capacity reduction on general traffic, and the effect this has, in turn, on bus journey times. Further details are provided below.

6.6.1. Mechanisms

The total in vehicle journey time for a bus service is calculated as:

$$\sum(BusLinkTime + BusTurnTime)\ {}_{*}$$

** Bus Turn Time is the time for a bus turning at junctions.*

The link and turn times are calculated using inputs from the RTM. Table 6.2 shows the attributes in the SATURN model that are imported into the EMME model.

Table 6.2 SATURN / EMME Attributes

SATURN Code	Filename	EMME Attribute	Description
2033	*.blk	@bol	Bus Only Lane Marker
4023	*.clk	@clkp	Congested Link time
1633	*.ctu	@tup	Congested Turn Time
1803	*.flk	@flkp	Free flow link time

The congested link time is used when the bus mixes with general traffic. The free flow link time is for bus using a bus-only lane. The bus only lane marker is used to differentiate within EMME which link time is used. The turn time is added to the link time to provide the total journey time.

However, there are some additional complexities that need to be incorporated into the calculation to ensure an accurate representation of the journey time, namely:

- Where there are a large number of other users of the bus lane, such as taxis or high occupancy vehicles, the benefits of the bus lane will be diluted. The magnitude of the effect depends upon who is able to use the bus lanes, and the proportion of traffic this entails;
- The additional priority at junctions resulting in the installation of Selective Vehicle Detection (SVD) will not be recognised within SATURN. Therefore a calculation of the likely effect of additional bus priority is necessary.
- Delays to bus run time occurring through boarding and alighting. Typical boarding times per passenger are as follows[7]:
    - 3 seconds (where majority of tickets are off-vehicle);
    - 6 seconds (where a high proportion involve cash transactions);
    - 9 seconds (where almost all ticketing involves cash transactions and change-giving).
    - alighting times are typically 1 to 1.5 seconds per person7. Therefore alighting times may also have a bearing on journey times, although not as dramatic an impact as boarding.

These impacts are reflected by the model through factoring bus journey times accordingly. Additional attributes within EMME are used to calculate bus journey times as shown in Table 6.3.

Table 6.3 Additional EMME Attributes

EMME Attribute	Description
@svd	Marker for SVD at Signalised Junction
@bsd	Bus Stop Density. Number of bus stops per km

@svd = 1 if there is selective vehicle detection for buses at a given node (signalised junction).
@bsd is calculated from empirical data for all links in Slough that carry a bus route. This is, in effect, the number of bus stops per kilometre.

6.6.2. Link Time Calculation

The following formulae are used to calculate the bus journey time on links:

Bus Link time = (Link time + Link length*BSD*delay)

Where:

- Link time = SATURN congested link time (if no bus lane), or is SATURN free-flow link time (if a bus lane exists)
- BSD = Bus Stop Density per km (derived from actual bus stop intervals)
- Delay = 20 seconds to allow for boarding / alighting

Buses do not always stop at every stop on their route, generally where there are no boarders or alighters for that stop, the combined use of the BSD and Delay function has been used to reflect this. Although some stops may have a longer boarding time than 20 seconds, a number of nearby stops may not, and indeed may not even stop due to lack of demand. The values defined, therefore, have been used to ensure appropriate journey times for travellers and overall generalised cost skim matrices for the demand model. BSD was calculated for each speed profile range (below 30mph; 30 to 59mph and over 60mph) based upon total network length in the speed range divided by the number of NaPTAN[8] bus stops falling on (within 100m buffer) that network.

6.6.3. Turn Time Calculation

The following formula is used to calculate the bus delay at turns:

Bus turn time = Adjustment factor * SATURN turn time

However, there are a number of complications to this formula, depending on the presence of a bus lane that leads up to the stop line and if SVD exists. Little information exists as to the effects on turn times for buses at

[7] The demand for public transport – TRL Report 593, 2004
[8] The **National Public Transport Access Node,** a UK nationwide system for uniquely identifying all the points of access to public transport in the UK

such facilities. The figures in Table 6.4 are considered a best estimate.. Furthermore,, a global factor of 1.1 has also been applied to convert SATURN turning time to bus turning time based on the fact that bus turning manoeuvre is slower than cars at junctions.

Table 6.4 The Assumed Effect of Bus Priority on Turn Times

Bus priority measure		Adjustment factor on turn time
Bus Lane	SVD	
N	N	1.00
Y	Y	0.81
Y	N	0.90
N	Y	0.90

7. Trip Matrix Development - RTM

7.1. Introduction

This section describes the stages associated with the development of the RTM matrix. The matrix development process involved the following steps:

Travel demand data

- collection, editing and expansion of intercept (RSI) survey data at the three sites;
- collection, editing and expansion of INRIX mobile phone survey data. This is a novel feature in this study;
- collection, editing and reconciliation of count data;

Partial matrices

- creation of partial trip matrices consisting of observed data;
- analysis of the accuracy of the partial trip matrices at sector level;

Synthetic matrices

- synthesis of complete car and LGV 'synthetic' trip matrices through gravity models.

External trips

- assembly of matrices of external to external movements.

Merging sources

- assembly of prior matrices of trips by light goods vehicles (LGVs) and heavy goods vehicles (HGVs) by combining partial with synthetic matrices;
- adjustments to the prior trip matrices in the light of the comparisons between modelled flows and counts across screenlines and cordons;

Matrix estimation

- matrix estimation to ensure greater consistency of the trip matrices with the count data;
- adjustments to the prior trip matrices if the magnitudes of the changes brought about by matrix estimation are regarded as significant; and
- adjustments to the prior trip matrices in the light of the journey time validations.

7.2. Checking

WebTAG does not have concise procedures for the way in which matrices need to be built. During our work for the Highway Assignment Models for TfL, a robust process was developed[9] for building trip matrices which has not yet found its way into the guidance. But as this is the best available source of information for building matrices in a consistent and robust fashion, we will be adopting it for this study.

At various key stages of constructing the prior trip matrices checks are required to ensure that the process has derived accurate trip movements. The checks are specified below (Table 7.1). The aims of these tests and the consequent adjustments are:

- to detect errors at each stage of matrix build which otherwise might remain undetected and be compensated for, erroneously, by matrix estimation;
- to ensure that the prior trip matrices are reasonably close to the count data, so as to limit the scale of the changes that matrix estimation will bring about; and
- to maximise transparency by making explicit the factors or adjustments that need to be applied to the various inputs and outputs which are necessary to bring the matrices in line with the counts.

[9] By the Denvil Coombe Practice

Table 7.1 Prior Trip Matrix Development Tests

Stage	Test	Comparison	Measure	Criterion	Acceptability guideline
Partial trip matrices	A	Flows and counts of trips across cordons, for the modelled hours separately.	Flow differences	< 5%	All or nearly all
Synthetic trip matrices	B1	Flows and counts of trips across cordons, for the modelled hours separately, with 3D Furness.	Flow differences	< 5%	All or nearly all
	B2	Flows and counts of trips across cordons, for the modelled hours separately, with 2D Furness.	Flow differences	<7. 5%	All or nearly all
Prior trip matrices	C	Total assigned flows and total counts in both directions across cordons and calibration and validation screenlines, for each modelled hour.	Flow differences	< 7.5%	All or nearly all

Notes: A - Test A should be done without an assignment. B - Test B1 should be conducted following application of the three-dimensional Furness, and Test B2 following a two-dimensional Furness. C - Test C requires assignments

Typically all these tests are carried out at RSI cordons. In this study much of the demand data will be coming from INRIX mobile phone data. The cordons will therefore be formed through ATC rather than RSI data.

Each stage above involves an iterative process of adjustments and refinements to meet the tests described above and to reduce the need for and impact of matrix estimation.

7.3. Mobile Phone Data

7.3.1. Data characteristics

Data will be supplied by INRIX on the basis of O2 mobile phone. The data will correspond with one week's worth of mobile phone movements during November 2013 for any traffic entering or moving within the ADM shown in Figure 4-1 and for three time periods AM, IP and PM separately. The data is treated with complete anonymity and consists of on-call but mostly off-call signals. The following characteristics of the mobile phone data need to be borne in mind:

- We expect the sample to be around 28% on average, from evidence provided by INRIX through previous studies. The data therefore has a significantly higher sample rate than what is normally achieved by RSI and in covering all movements within the FMA is considerably more extensive than RSI, but below we mention some of the shortfalls;
- The signal cannot be traced to the ultimate OD outside the FMA because interrogating the data in this way will be extremely expensive. The data will cover all movements within the FMA, but any external trips will be severed at the FMA boundary;
- The data will include any mobile phone signal and the matrix will thus cover not only cars, but LGV, HGV and even slow modes (pedestrians and cyclists). The slow modes will need to be eliminated through a speed cut-off which will be explained later. Clearly cyclists will certainly require special treatment as they can form a significant proportion in places like Oxford CBD;
- To separate out the HGV, INRIX will provide an HGV trip matrix from their GPS database. Methods of isolating HGV from the full matrix will be discussed later;
- The size of zones to which mobile phone data can be coded is a function of the population density, and is on average the equivalent of the 'postcode sector'. This is the first 4 digits of the postcode (e.g. SE16 4XX) of which there are 9,800 in the UK, not dissimilar to the DfT PASS3 zone granularity they use for their National Transport Model;
- A threshold dwell time must be identified above which the trip is deemed to have ended and another one started. From our experience working with INRIX on the Data Fusion project for the HA, the recommended threshold by INRIX is one hour. This might seem excessive, but given the size of the

zones INRIX data is coded to, and the fact that the more common off-call signal is only received at handover between one mobile phone area and another, it seems to be a threshold we will have to accept;

7.3.2. HGV matrix

The HGV matrix obtained from GPS will have a considerably lower sample rate compared to that from mobile phone data, and is expected to be possibly 0.5-1%. Whilst the prevalence of HGV might be low for the AM and PM peaks, it can be more important in the IP period. There are a number of options for separating the HGV from the total INRIX matrix:

1. Use INRIX (GPS) HGV matrix to subtract from the total INRIX matrix. This must be done after the data is expanded to a common month/day;
2. Ignore INRIX GPS data and use COTM data to subtract from phone matrix as above. This suffers from the fact that COTM is for a different year and assumes COTM was well calibrated, knowing full well that the HGV element of demand is usually the least robust. Alternatively, we could use the implied **proportions** in the COTM matrices at some spatial level to factor out the HGVs, assuming an unbiased INRIX sample.

We would propose to use the second method above working on the proportions of HGV's for each OD.

7.3.3. Cyclists

Eliminating the pedestrians from the INRIX data can be easily done with a 4kph cut-off below which every signal is deemed to be a pedestrian. The treatment of cyclists is more difficult because their speed in congested CBD is not dissimilar to that for cars.

It is pointed out that the estimation of speed is not rigorous within the INRIX data given the zoning system to which their data is coded, and the fact that it deals with 'crow fly' distances. The latter point might not have a significant impact for pedestrians and cyclists, but can make a difference to car speeds due to factors such as one way road systems.

There are several data sources for cyclist speeds. The DfT gives 12 kph for cycling and 4 km/h for walking (WebTAG 3.10.3). These take into account number of roads crossed/level of traffic on roads. www.Cyclestreets.net and states the following speeds for cycling:

- Unhurried 10mph (16 kph);
- Cruising 12mph (20 kph); and
- Quick 15mph (24 kph)

Work we are undertaking in Cambridge on the Data Fusion project showed car speeds[10] in the AM peak hour is similar to cyclist speed in the peak direction within the congested parts of CBD. This means that setting a simple speed threshold of, say 20 kph, to take out cyclist could dispose of many car trips within CBD.

The above issue only affects trips totally within congested parts of the CBD, because car trips with one leg outside CBD will have a higher average speed due to there being less congestion for sections of the route outside CBD. One can safely assume that any record with a speed lower than, say 24 kph is deemed to be a cyclist. There are several methods of overcoming this:

1. Apply a single speed threshold; say 20 kph below which the trip is deemed to be made by a pedestrian or cyclist. This has the danger of not picking up some slow moving car trips made wholly within the congested areas of CBD. The significance of this can be assessed through examining the % of the COTM matrix within this core area of CBD, as a guide;
2. Examine the maximum speed of each record for the full OD. Although average speeds between cycle and car might be similar in the inner areas of CBD, the maximum speed will differentiate between the two modes. However, this might involve considerable data processing for INRIX because the whole route has to traced before the maximum speed can be estimated;
3. Adopt two speed thresholds, one for trips wholly within CBD, and another for all other trips. Inside Oxford CBD we would apply 4kph thereby removing pedestrians, and outside it, a threshold of 24 kph is applied.

[10] Obtained from TomT~~t~~om GPS data

This produces a combined car and cyclists within the CBD sector of the matrix which need to be removed through assigning the resulting matrix and comparing to mode shares of car and cycle;

4. Plot a frequency distribution of car journey times inside the congested sections of Oxford CBD to identify the 85 percentile of the low speed tail. This could then form the threshold below which the trip is deemed to be a cyclist, and apply a higher threshold of, say 24 kph to all other sectors of the matrix; and
5. Apply a single speed threshold of, say 4kph to eliminate pedestrian trips and deal with a combined car and cycle matrix, which is assigned. The mode share of car and cycle on at cordon crossings are used to skim off cyclist trips from the matrix at these screenlines.

The preferred approach would be option 2, assuming INRIX will be able to undertake the processing. The work we carried out on the Data Fusion study in Cambridge showed that after examining modelled car speeds in CBD[11], option 3 was selected with 12 kph for CBD and 24 kph outside it. The 12 kph was sufficient to distinguish between most cars and cyclists within CBD.

7.3.4. Data correction

Before the INRIX sample is expanded, it needs to be corrected for bias. The data will be first checked in terms of the following to assess whether or not the sample rate is uniform:

- Approximate symmetry of the all day matrix;
- Profile of trips by time of day in comparison to NTS[12];
- Compare to NTEM trip ends; and
- Compare sector pair movements to observed volumes from ATC cordons.

The above will provide an idea of the expansion factors and how uniform they are across the matrix. The bias will also be assessed through a comparison of the Trip Length Distribution (TLD) to that from another source such as the NTS. The INRIX trip matrix will be placed in distance bins based on a skim from RTM, or if not available at the time, from COTM. Correction factors by distance bin will be derived as a ratio of the frequency of trips in that bin in NTS compared to the INRIX data. It is hoped that these correction factors are not in excess of +-20%. Once derived, the INRIX matrix is then corrected at the OD level using these factors and depending on the distance for that OD.

This process can of course be undertaken at the daily or period level, and each can give different results. The correction from the daily comparison will be first applied to the period matrices, after which the resulting period TLD will be compared against NTS to ensure there is parity.

7.3.5. Data expansion

There are several steps in expanding the INRIX data. Let us first consider expansion of the external elements, i.e. trips with one leg or more outside FMA:

- The INRIX matrices will be provided by zone of entry into the FMA. We can derive an expansion factor covering an external cordon around the FMA through comparing the sample total to observed traffic counts at the entry point into the FMA;
- This assumes that the expansion factors are the same for all external traffic whether one or both legs are outside the FMA;
- To allocate external trips using to their ultimate OD we can either use COTM or the NTM data. If we assume the profile of external trips has not altered considerably in the past 6 years[13], we can obtain a Select Link Matrix from COTM at each route and use that to allocate INRIX data into OD in proportion to the SLM. However, we believe it will be more robust to use NTM for that purpose, which comes at PASS3 zones (10,000 roughly across the UK) and can be aggregated to our zoning system. However, NTM is fully synthetic and has only been corrected on TLD from NTS and some regional constraints. We can also use the NTS data itself aggregating zones into sectors until the number of (unexpanded) observations becomes of a sufficient size to be reliable. We will compare these approaches and will only decide upon the one to adopt after reviewing the results and discussing with OCC.

[11] as output by the SATURN model and allowing for crow-fly distance
[12] National travel survey
[13] COTM base year was 2007.

As to the internal elements of the matrix, the following will be applied:

- If the zonal expansion factors from the comparisons in the previous section are similar, we will have more confidence that the expansion factor is uniform and can be applied across all the internal elements of the matrix;
- We will compare the INRIX matrix to the calibrated validated COTM matrix and derive expansion factors on a sector to sector level. The problem with this method is that the model base year is 2007 and so at best this process will expand the INRIX matrix to 2007 levels. More importantly the assumption is implicitly made that the 'profile' of trips has not changed in the last 6 years, which is unrealistic;
- The preferred approach will be to compare to current counts. We will collect traffic counts around cordons and screenlines, and this will provide in effect total (vehicular) trips on a sector to sector level. These are divided by sector movements from the INRIX matrix and expansion factors are derived.

After expanding the INRIX matrix, we will assign it onto the network (a fully loaded network with one iteration) in line with step A of Table 7.1.

7.4. RSI Data

Three RSI surveys will be carried out primarily to derive purpose split and occupancy which will be used to segment the demand obtained from the INRIX data. Two sites will be in Oxford (one on B480 Garsington Road West of A4142 Eastern Bypass, and the other on A420 Botley Road West of Seacourt P & R) and one in Abingdon at A415 Marcham Road between A34 and B4017. The sites are located intentionally to be in parity with those surveyed in 2007 and the purpose of that is to compare the matrix profile (in terms of trip end and purpose profile and TLD) to the old matrices. If it is demonstrated that the old and new matrices are similar, then old RSI can be used so long as there has been no major changes in demand (e.g. development) and supply (i.e. network) in the area, and growth in traffic has been moderate.

7.4.1. Data processing

ATC data is available for a continuous one or two week period. The data is classified into Cars/LGVs, OGV1 and OGV2 and processed to obtain the average weekday (Mon-Fri) flows by available vehicle type where the OGV1 and OGV2 data were combined to provide data for HGV. At each of the three new RSI sites, an additional MCC is conducted on the day of the survey.

The OSM will have a 2013 base year and represents the travel conditions for a typical April weekday.. Factors will be derived to account for monthly and yearly variations between the sites if need be. Traffic levels are continuously monitored and information from these sites will enable factors to be determined to normalise the data to the model base of April 2013. The seasonal factors (SF) are used to adjust counts between months.

The data is collected in vehicles and model assignment uses PCU so the factors in Table 4.2 are applied.

7.5. Partially Observed Trip Matrices

7.5.1. Combining Data Sources

One of the advantages of mobile phone data is that it provides a (substantial) sample of trips within the full coverage of the ADM, unlike the RSI surveys that only gives trips intercepted at screenlines and cordons. If the comparison of the old and new RSI proves that old RSI data can be used as discussed in the section above, then the old RSI data is expanded to current counts and needs to be merged with INRIX data.

If merging is required, the ERICA5 manual provides guidance on merging data for the same origin and destination that has been observed at two or more independent screenlines or cordons, as in the following section. Merging will only be required for the (combined old and new) RSI data at the cordons intercepted by the RSI surveys.

7.5.2. Variance Merging

The two sources of demand data (INRIX and RSI) are combined using variance weighting to give an output matrix that makes use of the most reliable estimate of demand for each OD pair.

Thus for cell i,j:

$$M_{ij} = \frac{I_{ij}^{O} W_{ij} + I_{ij}^{W} O_{ij}}{I_{ij}^{O} + I_{ij}^{W}} \qquad (1)$$

where: M_{ij} = Merged matrix

W_{ij} = ETM matrix

O_{ij} = Observed matrix

I_{ij}^{W} = Index of dispersion matrix for INRIX data

I_{ij}^{O} = Index of dispersion matrix for RSI data

and the Index of dispersion I_{ij} is a function of the variance of the trip estimate:

$$I_{ij} = Var(T_{ij}) / T_{ij} \qquad (2)$$

The variance of the trip estimate may be calculated directly:

$$Var(T_{ij}) = \sum_{n} e_{ij}(e_{ij} - 1) \qquad (3)$$

where: e_{ij} is the expansion factor for each recorded journey;

n is the number of recorded journeys from origin i to destination j; and

$T_{ij} = \sum_{n} e_{ij}$ is the total number of trips for cell ij.

For in-filled reverse trips, the value of e(e-1) is doubled and capped at a minimum value of 2 to reflect the added uncertainty in the trip estimate.

7.5.3. Summary of results obtained

Following the merging process described above, the number of movements by data source, time period and vehicle type are obtained, and these are called the partial matrices. Having built the partial matrices these are compared with the count data using Test A (Table 7.2) to ensure that the data had been processed correctly and to ensure the merging / filtering process had resulted in matrices closely reflecting the count data for the cordon crossing movements.

Table 7.2 Matrix Development Test A

Comparison	Measure	Criterion	Acceptability guideline
Flows and counts of trips across cordons, for the modelled hours separately.	Flow differences	< 5%	All or nearly all

7.5.4. Accuracy of Partial Matrices at Sector Level

The partial matrices are not statistically reliable on a cell by cell basis (at zone level) for car trips segmented into purposes. This information is required to produce statistically reliable sector level constraints for gravity modelling. These constraints apply to car and LGV trips; the limited HGV data means that such an assessment should instead be used to determine at what level of detail the partial trip matrices can be reliably used to adjust / constrain the HGV matrices.

The first stage of this process is to determine whether the sector system used within the partial matrix build would allow gravity modelling to be undertaken by trip purpose. Since spatial detail is more important in the RTM (as opposed to the MDM where purpose segmentation would be more important), only the vehicle types car and LGV are considered for gravity modelling. This is also supported by the trip end data being only available for light vehicles combined. By combining purposes and considering only car and LGV trips it is more likely to have statistically reliable data at the sector system adopted.

7.5.5. OD versus PA

The highway matrices are built by model time period in OD format, whereas the demand modelling within the MDM will be at the daily level and in Production Attraction (PA) format. This is completely acceptable because the assignment model will require the trip matrices to be by period and in OD. These can be readily converted into PA matrices (by purpose and time period) using factors derived from existing and new RSI's. The period PA matrices are then combined to form daily matrices that feed into the MDM calibration.

7.6. Trip Synthesis

7.6.1. Introduction

Trip matrices derived from the RSI and INRIX data are partial and samples of the full movements, some of which are not intercepted creating a lumpy matrix. To resolve this problem, synthetic matrices, based on the partially observed data are developed. These matrices have the advantage of including estimates of the movements not intercepted in the surveys and smoothing out the lumpiness in observed data.

The creation of synthetic matrices is a four stage process, and results in matrices that match the partially observed matrix movements at a sector to sector level as follows:

- assembly of synthesised trip ends;
- assembly of generalised cost matrices;
- assembly of trip cost distributions from the partial matrices; and
- trip matrix synthesis using either a gravity model, including constraints to the partial matrices.

7.6.2. Synthesised Trip Ends

TEMPRO provides a source of synthetic trips ends for building the gravity model. However, according to the DfT's website:

> "TEMPRO provides data on trips on foot, by bicycle, motor vehicle (both as a driver and passenger) by rail and by bus. Users should note, however, that TEMPRO trip ends by mode are based on average rates over a wide area and do not necessarily take into account the accessibility of each zone by each transport mode".

This means that whereas it can be a good measure for overall trip making at the daily level, the absolute number of trips cannot be relied upon for a robust estimate of car trips.

Since the matrices from the existing validated COTM highway models are in the dimensions required (assignment hours and OD format), these are identified as the most convenient source for the full set of synthetic trip ends. It is necessary to split the trip ends from light vehicles into car and LGV, and this is obtained from using split factors from the MCC data.

The COTM trip ends are for the model base year of 2007 and have to be expanded to current 2013 levels. TEMPRO 6.2 will be used to produce growth factors between 2007 and 2013 trip ends. Allowance will be made for specific significant developments that have taken place within the FMA during that period if such information is readily available from OCC. This would then give a more realistic profile of growth, to which the existing 2007 highway car demand can be furnessed to. A zonal correspondence table between existing COTM and OSM will be derived to convert the 2013 COTM demand into OSM zonal format.

7.6.3. Generalised cost matrices

In order to obtain the synthesised matrices, a set of generalised cost matrices are required through a skim of the inter-zonal times and distances from SATURN assignment models, which are available by assigning 2013 COTM matrices[14] onto the updated 2013 OSM highway networks.

The assignment will involve calculating the values of time and distance as described in TAG Unit 3.5.6 (DfT October 2012).

Intra-zonal costs are assumed to be 80% of the minimum (non-zero) inter-zonal trip cost for the particular origin zone. The same principle is applied for any other zero cost inter-zonal trips (when both origin and destination are new zones).

7.6.4. Synthetic Matrix Building

- WebTAG (3.10.3) recommends that synthetic infilling should be undertaken to overcome the problem of partial sampling with the observed matrices created.
- A gravity-based trip length distribution model is built using the MVGRAM program from CUBE, replicating the existing trip ends (O_i, D_j). The model is of the form:

$$T_{ij} = A_i B_j O_i D_j f(c_{ij}$$

Where Oi and Dj are zonal end origins and destinations respectively. The balancing factors A_i & B_j are calculated iteratively whilst the deterrence function is based on the Tanner function,

$$f\left(c_{ij}\right) = c_{ij}^{\alpha} e^{-\beta c_{ij}}$$ where α and β are calibration parameters.

- The calibration and forecasting stages within MVGRAM will adopt partial matrix modelling technique, by which the synthetic matrices are only manipulated and generalised at I-I, I-E and E-I level. The synthetic matrix is only created for car by three purposes, i.e., business, commuting and other for each time period (AM, IP and PM).
- The trip length distribution obtained from the partial observed matrices is used in gravity model calibration, along with the skimmed minimum distance matrices from the interim SATURN model. The use of distance skim instead of generalised cost skims is mainly due to the lack of fully loaded network in the initial stages of calibration. Therefore the generalised cost skims are deemed less reliable than distance skims.
- The (synthetic) trips ends for each internal zone in the FMA derived as described in section 7.6.2.

The first stage of gravity modelling is calibration of the Tanner[15] function parameters to best match the partial matrices based on the generalised cost. The second stage is to use the parameters found in the calibration stage with the (synthetic) trip ends to obtain a synthetic matrix. The matrix may then be subjected to 3D Furnessing to origins, destinations, and sector pair totals from the partial matrix. These sectors are defined at a spatial level to provide a sufficiently reliable estimate of trips.

The calibrated synthetic matrix is compared to the partial matrix to check that there were little changes to the trip costs as a result of the synthesis. Having built the synthetic matrices these are compared with the count data using Test B (Table 7.3) to ensure that the data had been processed correctly and to ensure the merging / filtering process had resulted in matrices closely reflecting the count data for the cordon crossing movements. The resulting matrices will be effectively the ‘prior matrices’.

[14] The LGV and HGV demand will also be lifted by global factors derived from National Transport Model (NTM) from 2007 to 2013.
[15] The Tanner function produces a new trip matrix to reflect the change in demand bought about from the trip ends using the generalised cost (distance) of movements between two zones. Essentially the higher the generalised cost the lower the trips.

Table 7.3 Prior Trip Matrix Test B

Comparison	Measure	Criterion	Acceptability guideline
Flows and counts of trips across cordons, for the modelled hours separately.	Flow differences	< 5%	All or nearly all

7.7. External Movements

External movements include movements to and from zones outside the outer cordon, some of which would pass through the cordons and exist in the partial matrix (and hence synthetic matrix).

If information from the new RSI's demonstrates that the old RSI data is still valid, and can be used, then that data will form the bulk of the external movements. Otherwise, the INRIX data has to be the sole source of observed external data, albeit such data (see section 7.3.5) will not have the ultimate OD's. In this case, for external movements, we will use the COTM profile of trips expanded to external cordon counts, supplemented by NTS or NTM[16] data.

7.8. Prior Matrix Creation

The prior trip matrices are assigned and the assigned flows are compared to the count flows for each screenline and cordon using Test C (Table 7.4). If a screenline fails to meet the criterion of having a flow difference less than 7.5% then any sector pairs found to have movements crossing it are altered to match the observed flow, in an iterative.

Table 7.4 Prior Trip Matrix Test C

Comparison	Measure	Criterion	Acceptability guideline
Total assigned flows and total counts in both directions across RSI cordons and screenlines, for each modelled hour.	Flow differences	< 7.5%	All or nearly all

[16] National transport model

8. Trip Matrix Development - PTM

8.1. Bus Trip Matrices

8.1.1. Overview of Methodology

The trip matrix development methodology aims to make the best use of each of the available sources of origin-destination data, namely onboard origin-destination survey data, ETM[17] data and other supporting data. The data will be combined using variance weighting as in section 7.5.2.

8.1.2. Onboard Bus Origin-Destination Surveys

A series of onboard bus occupancy and onboard origin-destination surveys will be undertaken in October and November 2013. The bus services surveyed will cover a sample of the most important services run by the major bus operators in Oxfordshire, namely Stagecoach, Oxford Bus Company, Whites Coaches, and Heyfordian.

8.1.3. ETM Data

ETM ticketing data will be supplied by the bus operators for all of their services in Oxfordshire, for a neutral 4-week period, from Mondays to Fridays between 7am and 7pm in an accounting period between September and October (With no school holidays / half-term).. The data were grouped into three time periods:

- AM peak period (07:00 - 10:00);
- Inter-peak period (10:00 – 16:00); and
- PM peak period (16:00 – 19:00).

The ETM matrices will be converted to OSM zoning. The data will be processed into two matrices, representing those services for which survey data were also available, and services for which no survey data were available.

8.1.4. Education trips

The HBED trips in the bus onboard origin-destination interview matrix will be replaced with the school census data provided by OCC. As there is no time information in the census data, a peak period to peak hour factor will be derived from the bus OD interview data (or as in section 8.1.7) will be applied to obtain the AM peak hour (08:00-09:00) education demand from the total school census records. When considering the education trips from the census data, those pupils that would use a dedicated school bus service will not considered since the OSM does not take account of these services. Note that the school census data will not be used for refining the education trip for modelled Inter Peak and PM Peak bus matrices, mainly due to the difficulty in obtaining accurate return times.

8.1.5. Partial Trip Matrices from Surveys

The new data from the onboard origin-destination surveys will be processed by:

- a series of checks to correct transcription errors and remove any inconsistent records;
- coding trip origins and destinations to OSM zones using the coordinates of the origin and destination postcodes;
- calculating expansion factors; and
- infilling reverse direction trips for certain PM peak period routes.

Expansion factors will be calculated for each surveyed bus service, taking account of:

- the proportion of timetabled bus services actually surveyed (by time period); and
- the proportion of passengers on each surveyed bus who completed the survey questionnaire

[17] Electronic ticketing machines

8.1.6. Multi-stage Bus Trips and Bus-Rail Trips

Trips using rail as their mode of access to the bus stop or onwards mode to their final destination will be separated out and stored in a separate matrix. Due to the hierarchical definition of public transport trips that will be adopted, these journeys are included in the rail matrix and not the bus matrix.

If there are trips using bus as their mode of access to the bus stop or onwards mode to their final destination, double counting will be eliminated through an analysis of origins and destinations.

8.1.7. Converting to Hourly Demand

The period demands, including ETM data, will be converted to hourly values by dividing by the following factors, should survey data is not found reliable enough to support such values:

- 2.5 for AM and PM peak period to peak hour; and
- 6 for Inter peak period to average Inter peak hour.

8.1.8. Merging Data from Surveys and Ticket Records

The observed onboard origin-destination survey and ETM matrices will be merged to produce the matrix. The Each data source has its own particular strengths and weaknesses:

- the survey matrix gives the best indication of true origins and destinations, but relates to a single day, and is derived from a sample of trips such that each recorded trip is assumed to represent a number of actual trips (how many is governed by the expansion factor). This can result in a “lumpy” matrix distribution whereby the demand is concentrated among an arbitrary subset of the true origins and destinations;
- The ETM matrix is based on average trip making over a 4 week period and (in principle) includes all trips rather than just a sample. However, various approximations have been required to convert from fare stage to true origin-destination. In some respects, the ETM matrix can be considered “synthetic” because the trips to/from each stage will have been spread synthetically among appropriate origin and destination zones. This means that the ETM matrix is “smooth”, as opposed to the “lumpy” survey matrix.

The following steps will be carried out to merge the survey and ETM matrices:

- combine ETM data for surveyed routes with the observed data using variance weighting techniques;
- control demand to observed totals at the sector-sector level; and
- add in ETM demand for non-surveyed routes.

The two sources of demand data will be combined using variance weighting to give an output matrix that makes use of the most reliable estimate of demand for each origin-destination pair. The variance weighting will adopt the procedure set out in section 7.5.2.

For any in-filled reverse (un-surveyed) trips, the value of e(e-1) obtained will be doubled and capped at a minimum value of 2.0 to reflect the added uncertainty in the trip estimate.

8.1.9. Variance of Trip Estimate for the ETM Data

For the ETM data, the variance could not be calculated directly in the same way as for the observed data. The observed will be analysed to find a relationship between the demand estimate (T_{ij}) and the variance ($Var(T_{ij})$), an example of the result is shown in

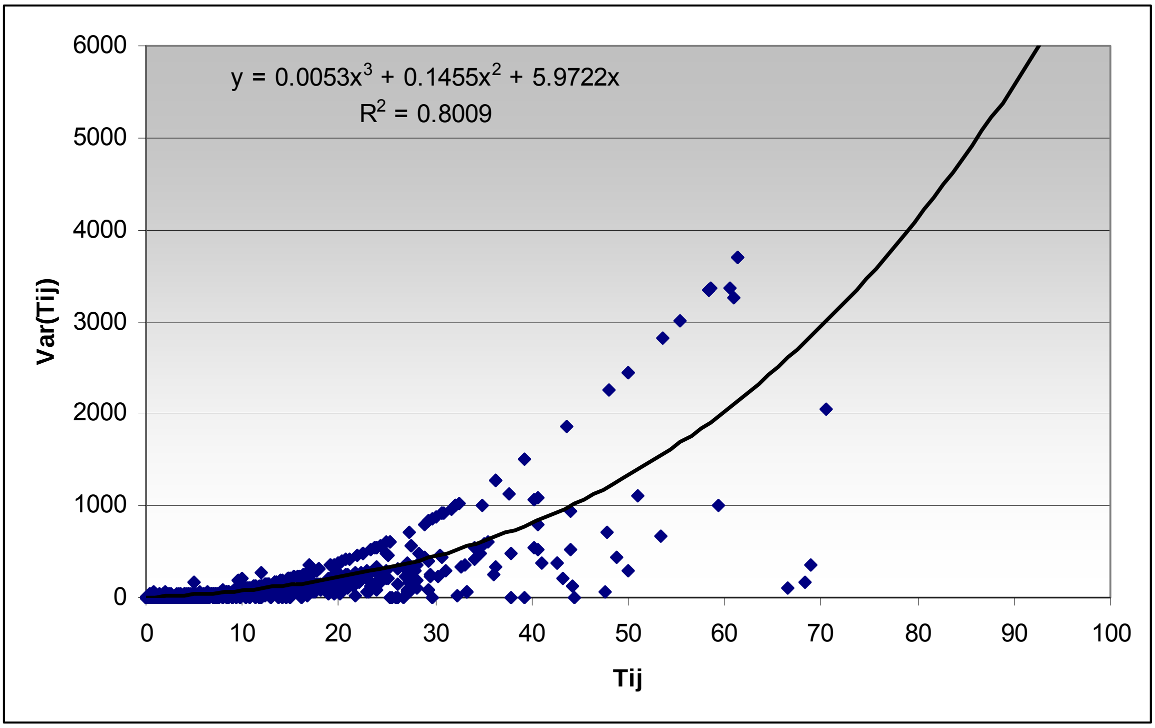
Figure **8.1**.

The function Var(T_{ij}) thus derived will then be used to estimate variances for the ETM data based on the ETM demand for each ij pair. The ETM variance will then be multiplied by 2.0 to reflect added uncertainty in the synthetic process of allocating the ETM demand to ij pairs.

8.1.10. Control Sector-Sector Movements

The merging process causes changes in the number of trips in the matrix. To deal with this, the matrix will be factored to retain the observed demand estimates on a sector-sector basis.

Figure 8.1 – Regression analysis of T_{ij} and $Var(T_{ij})$



8.1.11. Comparison of Observed, ETM and Merged Matrices

Trips to selected city centre zones in the AM peak will be compared to show the impact of the merging process. Desire line diagrams for observed, ETM and merged matrices will be plotted to illustrate how the merging process smoothes the observed demand over a greater range of origins and destinations, while hopefully still retaining the observed pattern of trips.

8.1.12. Adding Demand for Non-Surveyed Routes

ETM data for routes passing through the surveyed area, but for which no observed data are available will be added in to the merged matrix described above to complete the bus matrix.

The matrix totals for each stage in the matrix building process will be reviewed. The trip length distribution of the bus demand matrix at each stage in the process will also be examined.

8.2. Park and Ride Matrix Development

The Park and Ride distribution will be based on onboard origin destination Park and Ride passenger surveys collected in November 2013. We will assume that the volume of inbound trips from the P&R sites to the city centre is insignificant in the evening peak. However, the model will include trips returning to the P&R site in the evening.

The parking charge for each of the three P&R sites will be modelled as bus fares. Site specific constants for each site will need to be estimated during the calibration process, so that modelled and observed patronage are in parity. These constants will influence the choice of park and ride site in the absolute site allocation logit model described in the MDM Report.

8.3. Rail Matrix Development

The observed rail platform OD interviews and NRTS[18] matrices will be merged to produce the initial rail matrix for assignment. It is noted that each data source has its own particular strengths and weaknesses as

[18] National rail travel survey

for the comparison of ETM and OD bus interview data. In addition, the NRTS matrix is based on 2001 (and some 2005 data) expanded by growth rates from ORR[19] station usage data ;

- **OD Interview Data.** OD Interview data will be collected at Oxford, Bicester North and Didcot Parkway stations as well as boarding/alighting counts on platform. The cleaned and processed data will be expanded to the platform counts and then factored down to the gate entry/exit counts.
- **NRTS Data.** The NRTS data will be geo-coded to the OSM zoning system using the postcode sector information provided with the record. Expansion factors for the NRTS data to will be obtained from the entry and exit counts from ORR data. The data will be split by purpose and time period but the factors are the same for each purpose and time period because ORR data is an annual total.

8.3.1. OD and NRTS Matrix Merging Process

Observed data will only be collected at Oxford rail station. Therefore the merging observed OD interview demand with NRTS can only be carried out at this station. The NRTS data for other stations in the FMA will simply be added to the merged Oxford rail station data to create the final rail demand for assignment.

The process of merging NRTS and OD interview data at Oxford rail station will be carried out at sector level. The quality of the data for cells around 3 surveyed rail stations is more accurate from the OD surveys[20]. Therefore the two data sets were merged with 90% taken from the OD matrix and the remaining 10% from NRTS.

After the merge, the matrices will be factored back to the observed boarding/alighting counts at 3 surveyed rail stations. Demand from the bus interview survey for the passengers who use bus to access rail was also added as described in section 8.1.6.

[19] Office of Rail Regulator
[20] Due to data confidentiality, only the first part of the postcode of trip origin and destination is provided in NRTS. Therefore in general the zone allocation in NRTS is less accurate than OD interview data.

Graham Bown

Epsom Gateway
2 Ashley Avenue
Epsom, Surrey KT18 5AL

graham.bown@atkinsglobal.com