Department for Transport

TAG Unit M3.1

Highway Assignment Modelling

May 2024

Department for Transport

Transport Analysis Guidance (TAG)

https://www.gov.uk/transport-analysis-guidance-tag

This TAG Unit is guidance for the **Modelling Practitioner**

This TAG Unit is part of the family **M3 - Supply-side Modelling**

Technical queries and comments on this TAG Unit should be referred to:

Transport Appraisal and Strategic Modelling (TASM) Division

Department for Transport
Zone 2/25 Great Minster House
33 Horseferry Road
London
SW1P 4DR
tasm@dft.gov.uk

Contents

1. Introduction

1.1 Scope of the Unit

1.1.1 This unit provides guidance on developing, calibrating and validating a highway assignment model. It provides advice on aggregate highway assignment models, for both general and specific purposes, which represent average conditions over the modelled period, often one hour. The focus of this guidance is hence on steady-state equilibrium models.

1.1.2 The features and methods used in dynamic assignment and microsimulation are not covered explicitly. However, it should be recognised that many of the principles of this guidance are equally applicable, irrespective of model type. Further context about these methods, particularly relating to microsimulation model development, is provided in appendices at the end of this unit.

1.1.3 The unit provides advice on the topics shown in Figure 1. The numbers in brackets relate to the sections in this unit, with the aim of providing a means by which topics can be located relatively easily.

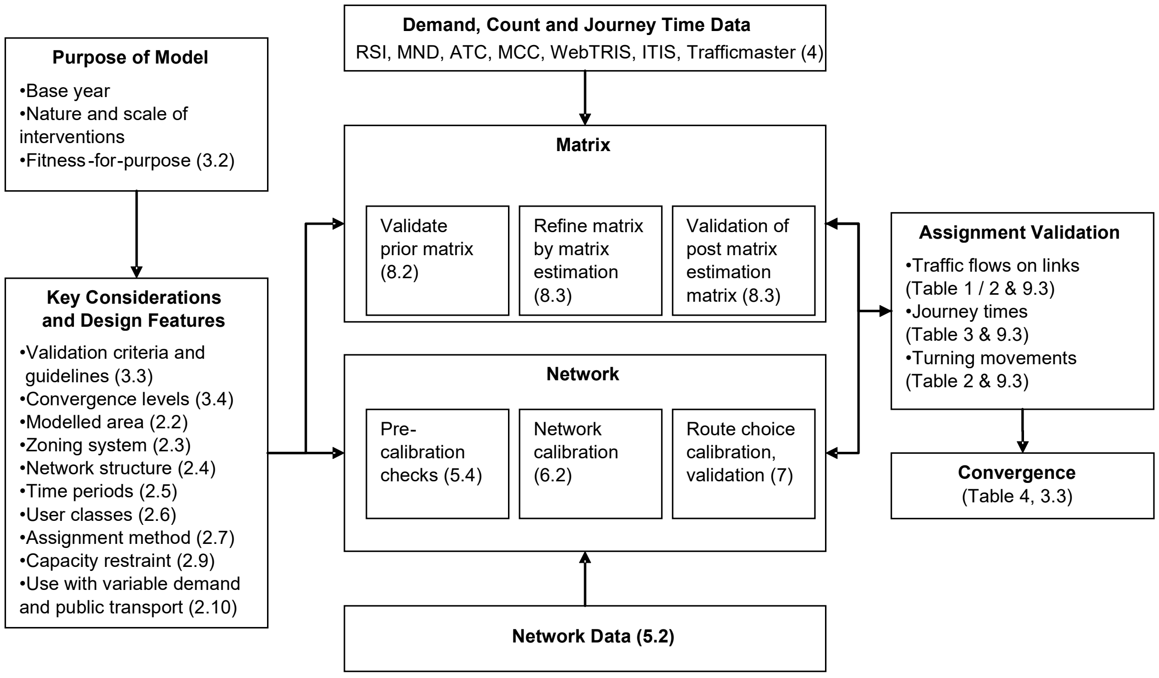


Figure 1 Structure of this Unit

1.2 Relationship of this Unit to Other Advice

1.2.1 This unit excludes advice on the following topics which are related to highway assignment modelling:

- travel demand data collection, including roadside interview surveys, traffic counts and journey time surveys
- the development of base or prior trip matrices for highway assignment models
- preparing forecasts using highway assignment models

1.2.2 This unit is a companion to the following TAG units:

- M1.2 - Data Sources and Surveys
- M2.1 - Variable Demand Modelling
- M2.2 - Base Year Demand Matrix Development
- M3.2 - Public Transport Assignment Modelling

1.2.3 A glossary of the terms used in this unit can be found in Appendix A.

2. Designing a Highway Assignment Model

2.1 Introduction

2.1.1 This section provides advice on the following topics:

- overall consideration of fitness for purpose of a highway assignment model
- the specification of the Fully Modelled and External Areas of the model
- the design of the zoning system
- the structure of the network representation, including centroid connector design
- the time periods which should be modelled
- the specification of the classes of user which should be assigned separately

- the assignment method
- the specification of generalised cost and the sources of the operating costs and values of time which should be used
- capacity restraint mechanisms, including the use of junction modelling and speed/flow relationships
- the relationships of the highway assignment model with variable demand models and public transport assignment models

2.2 Modelled Area

Fully Modelled and External Areas

2.2.1 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
- ensure that the full lengths of trips are represented for the purpose of deriving costs

2.2.2 It is important to establish at the outset the nature, scale and location of the interventions which are to be tested using the model.

2.2.3 The second and third requirements are particularly important where a highway assignment model will be linked to a demand modelling system. See section 2.10 for further details.

2.2.4 The modelled area should ideally be no larger than is necessary to meet these requirements. A larger than necessary modelled area will add to model run times and make acceptable convergence harder to achieve. This means that, where a model is developed from an existing model, consideration should be given to removing or simplifying redundant areas of network and zoning so that the model is no larger or more detailed than is necessary to meet the requirements set out above.

2.2.5 Within the overall modelled area (in many models encompassing the whole country), the level of modelling detail will vary. It is useful to consider this variation in terms of a classification of modelled area type as set out below.

- **Fully Modelled Area**: the area over which proposed interventions have influence. This is further subdivided as set out below:
  - **Area of Detailed Modelling.** This is the area over which significant impacts of interventions are certain. 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.** This is the area over which the impacts of interventions are considered to be quite likely but relatively weak in magnitude. It 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**: In this area impacts of interventions would be so small as to be reasonably assumed to be negligible. It would be characterised by: a 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.

2.2.6 In traffic routeing terms, a primary objective for the External Area is to ensure that traffic enters the Fully Modelled Area at the right locations and that opportunities to avoid travelling through the Fully Modelled Area are properly represented. The same principle applies to the relationship between the Rest of the Fully Modelled Area and the Area of Detailed Modelling. This will usually involve an appreciation of the catchment areas of the main roads crossing the boundaries of the Fully Modelled Area and Area of Detailed Modelling.

2.2.7 For microsimulation models the areas covered are typically characterised as the Area of Detailed Modelling. Typically, the external area consists of a series of zones representing the entry and exit points equivalent to a cordon model described further below. Routeing of external trips from/to or through this area would typically be represented in a higher tier model, this is described further in section 2.10.

2.2.8 The key to determining the boundaries of these areas is to understand the nature and scale of the interventions to be tested using the model. In some cases, models will be built for the appraisal of a specific scheme. In other cases, though, the model may be conceived with only some provisionally specified interventions in mind or as a general purpose tool. There is a need for clarity about the uses to be made of the model, that is, **the purposes for which the model can be used.**

2.2.9 Once an understanding about the uses to be made of the model has been gained, there are a number of ways to define the boundaries of the various model areas. One method is to make use of an existing model with geographic coverage as wide as or wider than that for the proposed new model. Testing of a range of potential interventions using such a model will, through analysis of link flow changes from a base position, allow an indication of where impacts are strong, weak or negligible. These categories broadly correspond to the three model areas as set out above. Where interventions could have a particularly strong impact on travel demands and where only highway assignment modelling is available rather than a fully specified variable demand model, the use of elastic assignment techniques could be made in this analysis.

2.2.10 Without an appropriate existing model, other less directly informative methods will need to be employed. Commercially available digital networks (with fixed speeds by road type) could be incorporated in a provisional highway assignment network and interrogated as above. Local knowledge and professional judgement should also be applied in the definition of the areas, although in this case the need to err on the side of wider geographic coverage may arise.

Cordon Models

2.2.11 Cordon models are assignment model variants where the modelled area is curtailed at a specified boundary. Trips with one or both end points outside of this area are not represented for their full length, being curtailed to one of the zones that represent the cordon boundary.

2.2.12 Cordon models are generally created from larger models. Most assignment software has the facility to create a cordon model automatically from a larger model. The primary element of the process is the aggregation of demand on highway links passing through the cordon entry/exit points to form the demand for the zones on the cordon boundary. Such a cordoning process has been demonstrated, in general, to maintain the flow and journey time validation of the initial model in respect of the internal area.

2.2.13 Cordon models as described above are of value when there is a need to examine significant numbers of relatively minor variants of proposed interventions. They will run in shorter times and achieve better convergence than the larger donor model. It would be usual to cordon (from the main model) a number of versions of the cordon model which relate to a core scenario and a central case definition of an intervention or set of interventions. Care needs to be taken in the use of a cordon model that the variants tested will not significantly change the flows at the entry and exit points to the cordoned area, as this would suggest the cordon was too tightly drawn or the full version of the model should be used. Once a preferred option has been identified, testing using the full model will be required. An assumption implicit in the above is that the intervention variants to be tested would have negligible impact outside of local route choice (ie no significant demand effects).

2.2.14 Creation of free-standing cordon models is not recommended practice. As discussed above, such a process would involve making assumptions about where the boundary of the impact on traffic routeing lies and, without a wider area network, this is not normally practical. More importantly, cordon models cannot be used in conjunction with demand models, as full origin to destination costs are not known for many movements, and alternative destinations for trips currently with an end point in the cordoned area are excluded from destination choice modelling.

2.2.15 Microsimulation models will tend to follow the principles of cordon models as described above and are often developed as free-standing tools. When developing, it is important to acknowledge the potential for wider traffic re-routeing and develop appropriate methods to account for such effects. Further details around developing such models and their connection to higher tier models are provided in sections 2.7 and 2.10.

2.3 Zoning System

2.3.1 The design of the zoning system should be closely related to the classifications of the modelled area defined in section 2.2. Zones should be smallest in the Area of Detailed Modelling, becoming larger for the Rest of the Fully Modelled Area and progressively much larger for the External Area. At the boundary between the classifications of area type, it is important to avoid sudden changes in average zone size and a graduated approach is desirable.

2.3.2 The primary building block for the zone system should be Census and administrative boundaries, and boundaries relating to national forecasts, in particular:

- Census Output Areas
- Wards
- Districts
- Counties
- National Trip End Model Zones (which are based on Census Output Areas)

Digital boundaries are available for all of the above.

2.3.3 It is very important that boundaries of zones do not straddle any of the primary building block boundaries as set out above, as this will make assembly of base year planning data and use of exogenously available land-use forecasts quite problematic. It has to be accepted that these administrative area definitions have not been designed with highway or transport modelling specifically in mind and the boundaries are often less than ideal. However, the adverse consequences of not following this approach outweigh any detailed highway modelling advantage in the context of forecasting using transport models.

2.3.4 In addition to the above, boundaries defined for more strategic models that may be linked to, or need to be consistent with, the new model should also be respected (provided their zone definition principles have followed compatible rules). This advice also applies in respect of the zoning system of any other data sources to be used in constructing the model.

2.3.5 Within the above constraints, further factors that should be taken account of and exploited in the definition of zones include:

- natural barriers (rivers, railways, motorways or other major roads)
- areas of similar land use that have clearly identifiable and unambiguous points of access onto the road network to be included in the model

2.3.6 Zone boundaries should take account of the need for internal screenlines (interviews and counts) for trip matrix building, calibration and validation. It is desirable that all counts should be located on zone boundaries or sufficiently close to these, such that the positioning of zone connectors makes the modelled traffic on the relevant links consistent with the observed flows.

2.3.7 Once zone boundaries and survey locations are defined, it is generally quite straightforward to fit a study-specific programme of counts, for calibration and validation, to the zone boundaries. However, where development of the model is dependent upon use of existing counts for such purposes, the process will be more problematic, and the location of these counts would also need to be taken account of in the zone definition process.

2.3.8 In some models, especially those covering urban centres in considerable detail, it will be necessary to represent explicitly the location of the main car parks or groups of car parks as unique zones. Trip matrices for highway assignment models should represent vehicle rather than person trip ends, and in urban centres these can be significantly different[1].

2.3.9 There will always be a degree of interdependence between the definitions of the zoning system and the network, such that one should not be defined without reference to the other. If zones are significantly larger than implied by the detail of the network, it will often be impossible to locate zone connectors realistically. This may lead to distorted traffic flows on nearby links and turning movements at nearby junctions, which may themselves distort traffic patterns elsewhere in the network. In urban area models in particular, zones should be small enough to avoid this type of problem.

2.3.10 Generally, zones should be designed so that the number of zone connectors for each zone is minimised. Use of multiple connectors leads to loadings at the periphery of zones, underestimating travel and therefore traffic within the zone itself and should, if at all possible, be avoided[2].

2.3.11 An important feature of the zoning system for the Area of Detailed Modelling in a highway assignment model is that the resultant numbers of trips to and from individual zones should be approximately the same for most zones and that the numbers of trips to and from each zone should be some relatively small number, such as 200 or 300 per hour, to avoid unrealistically high loads appearing at some points in the network. Some will need to be smaller than this, in order to model their loading points sufficiently accurately. Beyond that, any

[1] By contrast demand models require person trip ends, so the zone system must also be designed to represent these in an adequate manner.

[2] Some software have the ability to allocate traffic to zone connectors in a way that reflects the user assumed distribution of trip ends within a zone, and this would allow for the appropriate use of multiple connectors and somewhat larger zones. The practice is not common in highway assignment modelling, although it is widely used for public transport modelling.

major increase in the number of zones, with fewer trips per zone, is likely to increase the complexity of the model without adding significantly to real model accuracy. Small zones should be concentrated in the key parts of the modelled area. In the Rest of the Fully Modelled Area, where zones will be significantly larger than for the Area of Detailed Modelling, it will be often necessary to relax the constraint on numbers of trips to and from each zone.

2.3.12 The level of zonal detail adopted will be a matter of judgement. However, analysts should bear in mind that, although a finer zoning system and network will give a better loading of traffic onto the local road system, this better definition is achieved at the expense of less data certainty, model building costs and computer run times.

2.4 Network Structure

Network Structure: Area of Detailed Modelling

2.4.1 Within the Area of Detailed Modelling, a relatively high level of detail will generally be appropriate. Guidelines for Developing Urban Transport Strategies (Institution of Highways and Transportation 1996) suggests that "all roads that carry significant volumes of traffic" should be included and more generally that networks "should be of sufficient extent to include all realistic choices of route available to drivers".

2.4.2 For a model created for a specific scheme, the network should include all main roads, as well as those secondary routes, and roads in residential areas (especially 'rat-runs'), that are likely to carry traffic movements which could use the scheme being assessed, either in the base year or in future years, and that are significant in relation to the capacity of the scheme. Modelling this 'rat run' traffic may present some technical difficulties, but it is desirable that the effectiveness of the scheme in attracting this traffic back to the main road network, is accurately assessed. Local highway authorities will normally be aware of the common 'rat-runs', but some independent assessment may also be required. In the absence of count data, accident plots may also give an indication of alternative routes that vehicles are using to avoid local congestion points.

Network Structure: Rest of the Fully Modelled Area

2.4.3 Impacts of interventions could occur over the Rest of the Fully Modelled Area, albeit at a relatively weak magnitude, as ensured by well-considered model design. This part of the model would be characterised by somewhat larger zones and less network detail than in the Area of Detailed Modelling and capacity restraint modelling by means of speed/flow relationships (primarily link-based but possibly also including a representation of strategically important junctions).

2.4.4 The appropriate balance between demand and supply for this area should be carefully considered. An appropriate balance is ensured in the Area of Detailed

Modelling through use of quite small zones and detailed networks. In the External Area, travel costs should not be responsive to levels of demand and therefore the need to ensure an appropriate balance between demand and supply does not arise.

2.4.5 In the Rest of the Fully Modelled Area, while demand would be fully represented, zones would be relatively large and networks would not be comprehensive. The possibility exists that the demand loaded to the network links could be either excessive, where the network is too sparse in relation to the level of zoning, or too low, where zones are large and many trips are intra-zonal and therefore not loaded onto the network. Either possibility would distort the representation of journey times and could introduce undesirable levels of sensitivity of journey times to demand changes. An appropriate balance therefore needs to be found. The ideal would be a situation where the loaded demand is appropriate for the capacity of the roads that are included in the network.

2.4.6 Some key points to bear in mind when defining networks and zones for the rest of the Fully Modelled Area are as follows:

- intra-zonal trips are not loaded onto the network
- it would be appropriate to consider including in the network all significantly trafficked links that cross zone boundaries
- some increase in the capacity of the modelled roads to reflect that of roads that have not been included could be considered
- the larger size of zones in the Rest of the Fully Modelled Area may mean that it is necessary to relax the constraint of 200 to 300 trips to and from each zone in each modelled time period referred to in section 2.3
- the level of detail of the network in this area may be adjusted during the model calibration process
- tests should be undertaken of the sensitivity of journey times to demand changes, with any necessary adjustments made to ensure realism
- some count and journey time information should be obtained for this network

Network Structure: External Area

2.4.7 The level of detail in the External Area should be less than within the Fully Modelled Area. Movements between the Fully Modelled Area and the External Area need to be represented at some level of detail for two reasons:

- if only internal movements are defined, then the representation of traffic on the network at the periphery of the Fully Modelled Area will not be representative as trips from outside the Fully Modelled Area will not be present

- the ability of the model to reflect changes affecting choice of travel to these peripheral areas will be inconsistent and dependent on whether the changes relate to movements within the Fully Modelled Area or between the Fully Modelled Area and the External Area

2.4.8 Movements between the Fully Modelled Area and External Area will have an impact on the network in the Fully Modelled Area. In addition, through traffic, that is, external to external movements that pass through the Fully Modelled Area, will also have an impact on the network of the Fully Modelled Area.

2.4.9 External networks are often characterised by a more skeletal network and zoning system. The fixed speeds used in this area should be based on cruise speeds (which are time period specific) rather than speed limits and information on suitable cruise speeds to use in the External Area could be obtained from the Fully Modelled Area and/or from external data sources such as tracked vehicle data and WebTRIS (National Highways' Traffic Information System).

2.4.10 It is important to avoid the occurrence of unrealistic re-assignment to routes which avoid the Fully Modelled Area or the 'collar' part of the External Area which may result from the differing levels of detail at which delays and congestion are modelled within these areas. This is likely to be more prevalent where fixed speeds are used and no capacity restraint is applied.

Centroid Connectors

2.4.11 Trips are loaded onto the network using special links usually referred to as **centroid connectors**. Each zone can be viewed as having a centre of gravity and the centroid connectors are the means by which the demand from or to zones loads onto or leaves the network. The position of these connectors is often a critical factor in achieving realistic results from the assignment model, especially within the Fully Modelled Area.

2.4.12 The general principle is that centroid connectors should, wherever possible, represent actual means of access to and egress from the modelled network. They should not cross barriers to vehicular movement.

2.4.13 It is important to bear in mind the design of the centroid connectors when designing the zoning system. For example, in a zone bounded by main roads, the centroids can connect to these main roads via more local roads in a generally realistic fashion. If, on the other hand, two or more main roads pass **through** a zone, it becomes much harder to design centroid connectors which load traffic appropriately onto the network; in these instances, whether traffic loads onto one or other of the roads will depend on the analyst's judgment about the location of the centre of gravity of the zone in relation to the road network.

2.4.14 In general, centroid connectors should not be connected directly into modelled junctions unless a specific arm exists to accommodate that movement. In practice, certain packages allow a range of different types of connection. It is important that the analyst understands how flows are output by the package being used and ensures that this information is used and reported in a consistent manner.

2.4.15 It is generally preferable to minimise the number of centroid connectors from a single zone to a network. Multiple connections can lead to instability during assignment and model convergence problems. There are also associated difficulties introduced where multiple centroid connectors can straddle traffic count locations and this should be avoided. Centroid connectors from adjacent zones should not be loaded onto the same point as this will lead, at worst, to movements between the zones not appearing on the network. Indeed, this can be a common source of errors and warnings from TUBA.

2.4.16 The times, distances and money costs coded onto centroid connectors should represent the average costs of accessing the network from the development in the zone as realistically as possible.

2.5 Time Periods

2.5.1 Traffic patterns, trip purpose and vehicle type proportions, traffic flows and congestion vary by time of day. Highway assignment models should therefore normally represent the morning and evening peaks and the inter-peak period separately as a minimum. There may also be a need to model further time periods, such as off-peak times and weekends, if this materially affects the analysis of the scheme impacts, including economic and environmental, and cannot be accommodated adequately by the appraisal 'annualisation factors' (see TAG unit A1.3 - User and Provider Impacts).

2.5.2 The peak periods should be identified by analysis of Automatic Traffic Counts (ATCs) at as many points as are available throughout the Fully Modelled Area. The aim should be to designate those periods where traffic flows are markedly higher as the peak periods[3] with the inter-peak period being a period between the two peaks during which flows are approximately constant. It is conventional to define peak periods, and inter-peak periods, as multiples of hours rather than as hours plus fractions of hours.

2.5.3 In the inter-peak period, it is usually appropriate to model an average hour. In the peaks, however, models could represent one of the following:

- each individual hour within each peak period
- the actual peak hour within each peak period
- the average hour within the peak periods

2.5.4 In theory, a separate model run should be prepared for periods where demand differs significantly from others. During the peaks, it may therefore be necessary to model each hour separately where the profile of demand across the peak is substantially different, rendering the use of an average hour an unrealistic representation. In practice, constraints on resource to collect the adequate data

[3] It should be noted that flows are often tidal, so the flows in the peak period in one direction may actually be lower.

and calibrate and validate the model for each time slice can be prohibitive. Where this is the case, some compromise will be necessary and where possible a peak hour model should be preferred to an average hour where such a distinct peak exists.

2.5.5 Actual peak hour models are therefore to be preferred in most circumstances. However, highway assignment models are commonly used to provide generalised costs for demand models and demand models normally represent peak periods (as well as inter-peak and off-peak periods). The use of actual peak hour assignment models, and possibly models of the peak shoulder hours, to drive peak period demand changes is considered further in section 2.10.

2.5.6 Peak hour models have the following advantages:

- traffic flows and congestion at peak times will be more accurately modelled, which will not be the case if average conditions are lesser congested

- a peak hour is more representative of a situation in reality. While traffic counts and journey times can, in principle, be averaged over the peak hours, it is hard to judge the plausibility of the routes modelled for a period which does not exist in reality

2.5.7 Average peak hour models may be suitable to use in the following circumstances, noting that it will be rare for these conditions to be met for models of regions and for urban areas outside Inner and Central London:

- capacity on the network is more than adequate to cater for the forecast demand in the base year and forecast years and/or

- traffic levels are approximately constant throughout the period and/or

- a substantial proportion of the trips in the Fully Modelled Area are longer than one hour (although this may be more appropriately handled through modelling using longer time periods or through dynamic methods)

2.6 User Classes

2.6.1 Operating costs vary by vehicle type and values of time vary by the purpose of the trip being made. Values of time may also vary by income group. This means that different combinations of vehicle and user may have different distance coefficients (defined as the vehicle operating cost / value of time) and therefore be modelled as choosing different routes through the network.

2.6.2 The total trip matrix should therefore be split into a number of **user classes**, each of which should have distinctly different distance coefficients in their generalised cost formulation. Unless there are special circumstances, cars on business, other cars, LGVs and HGVs should be treated as individual user classes and assigned separately in a multi-user class assignment. Non-work car demand should also be split by income band where tolling and charging schemes are to be assessed.

2.6.3 It is also possible that separate user classes may need to be distinguished for other reasons. Examples of such reasons include the following.

- In some models, some vehicle types may be exempt from or prone to certain restrictions. For example, goods vehicles are sometimes permitted to use bus priorities, or restricted from using certain roads. They should therefore be treated as a separate user class, assuming the network coding can accommodate these exemptions or restrictions.

- Certain vehicle types are important for particular appraisals. For example, HGV noise and air pollutant emission levels are much higher than those for cars and therefore should be treated as a separate user class. Also, employers' business trips have relatively high values of time and the impacts of these trips will play a significant role in the Transport Economic Efficiency appraisal. Being able to identify the demands and costs for these trips separately and realistically is important.

2.6.4 Having decided on the categories of vehicle and user combination that should be treated separately, consideration should be given to the following two further factors.

- The proportion of the total matrix formed by any particular user class should be considered. Doubling the number of user classes will roughly double model run times, so the value of adding each additional user class should be considered carefully.

- The robustness with which the trip matrices for each user class can be assembled should also be considered. It is unlikely to be worth assigning a user class separately if the accuracy with which the relevant trip matrices can be created is poor.

2.7 Assignment Methods

Introduction

2.7.1 This unit is concerned with aggregate highway assignment models which represent average conditions over the modelled period, often one hour (as explained in section 2.5). These models are also known as **Steady State Assignment** models. Some description of dynamic assignment methods, including microsimulation, can be found in Appendix B. Further details on how the guidance in this unit should be considered in relation to microsimulation models is provided in Appendix C.

Steady State Assignment Models

2.7.2 The assignment model predicts the routes that drivers will choose and the way that traffic demand interacts with the available road capacity. The importance of this aspect of modelling in urban areas has already been emphasised. By definition, a Steady State Assignment model is one in which a single (average) set of cost conditions is assumed to apply across a fixed time period, typically one hour, and a fixed volume of trips (with uniform demand) is 'assigned' to satisfy these conditions.

2.7.3 Using these models in scheme assessment, the aim of the assignment model is to reach an equilibrium such that costs and traffic flows are in balance, under the assumption that individual users will seek to minimise their own costs of travel through the network. The underlying principle is expressed as Wardrop's First Principle of Traffic Equilibrium which may be stated as:

Traffic arranges itself on networks such that the cost of travel on all routes used between each OD[4] pair is equal to the minimum cost of travel and all unused routes have equal or greater cost.

2.7.4 Further details of the various assignment methods can be found in Appendix B.

Choice of Assignment Method

2.7.5 The choice of the assignment method will depend on the degree of congestion in the network, the availability of alternative routes, and the extent to which delays at junctions influence route choice. The general assumption should be that, in congested networks, junction delays are critical and therefore junctions should be modelled explicitly and congested assignment techniques applied. A converged Equilibrium Assignment is required for a Steady State Assignment.

Other Features of Assignment Models

2.7.6 Some assignment models have been extended to take into account other special features of urban road networks. These include:

- the use of time slices, or demand profiling
- bottleneck (or **flow metering**) effects
- queue interaction (or **blocking back**) effects

2.7.7 As described in section 2.5, it may be deemed necessary to model more than one hour to represent a period within which the demand profile varies significantly. This is particularly important where there is significant queuing delay necessitating explicit junction modelling.

2.7.8 The use of **time slices** to model the effects of varying traffic levels within peak periods may be considered to address the above issues. However, the use of time slices can introduce additional complexity into the assignment process. This needs to be balanced against the improved modelling of the growth and decay of queues and delays. The method chosen for a particular scheme appraisal will depend on the circumstances surrounding the scheme. If

[4] Origin-Destination

assignments are undertaken for time-slices of less than one hour, the traffic measures should be presented in terms of hourly flow rates.

2.7.9 One specialised use of a time slice approach that may be considered is the use of a pre-peak assignment to represent queued traffic present at the start of the main modelled time period. Some assignment suites allow for existing queues (and the volumes of traffic represented) to be used as initial conditions in the network. In such cases, it is often convenient to introduce a 'feedback' loop, whereby estimates of demand in the main time period are factored appropriately and used as demand for the pre-peak period. This is particularly important for microsimulation models where the performance of the network at the start of the time period can be essential to achieve realistic network conditions through the time period.

2.7.10 Some congested assignment models also incorporate procedures for estimating the effects of capacity restrictions on downstream traffic flows (sometimes referred to as **flow metering**). This is an important feature of many congested road networks, and failure to take it into account can lead to serious over-estimation of queues and delays at downstream junctions and poor estimation of overall network delays. It also leads to the distinction between 'demand' flow (ie the flow that would exist if there was no upstream capacity restriction) and 'actual' flow (after taking such restrictions into account). If this modelling feature is in operation, care must be taken to use the correct flow definition in the subsequent traffic, environmental and economic appraisal calculations (including the validation of traffic flows). Where aggregate flows over a complete model time period (or a series of time periods) are being used, the difference will often be negligible but, if flows for shorter periods are being considered, as for example in certain environmental and design calculations, it is the 'actual' flow values that are most appropriate.

2.7.11 The operation of junctions in urban areas is sometimes influenced by queues **blocking back** from an adjacent junction and limiting flow through upstream junctions. Some modelling packages have procedures for dealing with this situation, and particularly in microsimulation, this is a critical part of the model. The use of blocking back procedures can generally be justified in terms of producing more realistic assignments in congested networks, although users should be aware of potential problems where very short links are coded. Where blocking back effects are significant and the focus of interventions is to relieve congestion, microsimulation models can be useful tools to enable a better understanding of changes in congestion.

2.8 Generalised Cost

2.8.1 In principle, the basis for route choice in a highway assignment model should be generalised cost, defined as follows:

Generalised cost = (time) +

(vehicle operating cost per km x distance / value of time) +

(road user charges / value of time)

It should be noted that where user classes are defined by income group (for example where road user charges are important), the values of time used in the generalised cost formulation should vary by income group.

2.8.2 **Generalised cost is expressed in units of time.** This removes the difficulty of changes in costs over time, due to inflation and other changes, which may produce inconsistencies from year to year. In this regard, time is the more stable measure to use and does not require further adjustment, beyond the change in values of time over time.

2.8.3 The vehicle operating cost per km should be calculated using the formulae in TAG unit A1.3 - User and Provider Impacts. The values of time used to convert the vehicle operating costs to time units should also be taken from TAG unit A1.4. In a highway assignment model, the vehicle operating costs and road user charges should relate to vehicles (as distinct from persons as required for a demand model). Average vehicle occupancies should be derived from local data, such as roadside interview surveys, in order to convert the values of time per person given in unit A1.4 to the values of time per vehicle.

2.8.4 The operating cost formulae given in TAG unit A1.3 relate vehicle operating costs to speed of travel, by vehicle type. It is not considered necessary to use the speeds of traffic on each link individually, not least because modelled speeds will vary as the assignment iterations proceed, and varying vehicle operating costs (and therefore the basis for route choice) within the assignment process is likely to make convergence harder to achieve efficiently. Some kind of average speed should therefore be used.

2.8.5 Where separate speed/flow relationships are used for light and heavy vehicles, the average speeds used in these vehicle operating cost calculations should vary accordingly. See the next section for further advice on this aspect.

2.8.6 Values of vehicle operating cost and value of time should be derived from TAG unit A1.3 for the base year and for forecasting assignments without amendment. This will maximise consistency between the assignment and the demand model in line with the advice in TAG unit M4 - Forecasting and Uncertainty.

2.8.7 It is a requirement for the assessment of road tolling and charging that demand in the highway assignment model is segmented to reflect the variation in values of time. When introducing segmentation by income (that is, values of time which vary by income group), that variation in the value of time should usually only be allowed to affect the tolls and charges and not the vehicle operating costs; the same distance coefficient should be applied to all income groups in each car purpose[5]. In cases where there are no significant tolls and charges in the base year, this approach means that the assignment model can be calibrated without

[5] Usually, only the consumer purposes.

income segmentation, with merely a final check made that segmenting the matrices by income in the base year does not materially affect the validation.

2.8.8 The value of time given in TAG unit A1.3 for HGVs relates to the driver's time and does not take account of the influence of owners on the routeing of these vehicles. On these grounds, it may be considered to be more appropriate to use a value of time around twice the TAG unit A1.3 values[6]. If the higher value of time is used, a sensitivity test should be run to show the impacts of using the values of time from TAG unit A1.3.

2.9 Capacity Restraint Mechanisms

Introduction

2.9.1 Capacity restraint is the process by which speeds, and therefore travel times and generalised costs, are adjusted so that they are consistent with the assigned traffic flows in macro/strategic models. As the assigned traffic flows are dependent on the generalised costs of each route, and as the generalised costs are dependent on the assigned flows, an iterative procedure is required to find the equilibrium when the assigned flows are consistent with the generalised costs. The methods by which this equilibrium can be determined are outlined in section 2.7 and Appendix B.

2.9.2 In models of congested areas, capacity restraint should be applied throughout the Fully Modelled Area. Capacity restraint may be applied by the use of either:

- link-based speed/flow or flow/delay relationships or
- flow/delay modelling of junctions or
- a combination of both

2.9.3 Junction modelling will be required where junction capacities have a significant impact on drivers' route choice, and where delays are not adequately represented by speed/flow relationships applied to network links. Care must be taken to specify realistic capacities throughout the Fully Modelled Area and in the choice of turning movements for which it is necessary to specify individual turn capacities.

Junction Modelling and Speeds Between Junctions

2.9.4 The Fully Modelled Area will contain the highest level of detail within the model and hence this is the area within which all significant junctions should be modelled in detail (often referred to as 'simulated').

2.9.5 It is usually assumed that, in **urban** areas, flow does not greatly influence link speeds between junctions (cruise speeds), although there are some exceptions,

[6] For further information on values of time, see Hague Consulting Group (1994 (revised 1996)).

such as motorway links and long dual carriageway links. Speeds between junction queues are more related to the type of road and activity levels alongside links and pedestrian flows crossing links than flow levels.

2.9.6 There may be instances where link capacities are lower than junction arm capacities. In these cases, speed/flow relationships should be used to represent the effects of the link capacities.

2.9.7 Appendix E specifies the speed/flow relationships used in COBA[7] (the DfT's link-based COst Benefit Analysis software) and which may also be used in highway assignment models. However, the **urban** speed/flow relationships apply to networks rather than individual links and also include an allowance for junction delays. These **urban** relationships are therefore only suitable for the approximate modelling of capacity restraint effects in areas peripheral to the area over which the main impacts of the interventions being tested would be felt; they should **not** be used in conjunction with junction modelling. The **suburban** and **small town** speed/flow relationships apply to whole routes rather than individual links but exclude delays at major junctions which need to be modelled explicitly.

2.9.8 Generally, in **urban** areas within the Fully Modelled Area, the use of fixed cruise speeds is advised in conjunction with junction modelling, rather than using link-based speed/flow relationships. Cruise speeds should **not** be based on speed limits but should reflect mean speeds on a link. This is particularly relevant for traffic-calmed links, for links carrying heavy bus flows that impede other traffic, and for links with high parking and/or pedestrian activity.

2.9.9 Experience suggests that, in an **urban** network, the times taken to travel along links between queues can, for the network as a whole, be as much as twice the total delay at junctions; that is, as much as two-thirds of total travel times will come from the cruise speeds rather than junction delays. Thus, cruise speeds need to be established accurately. See section 5.2 for further advice.

2.9.10 There may be a need to represent mid-link delays, such as pedestrian crossings, where pedestrian crossing flows are significant, principally near to public transport interchanges, rail stations and shopping streets. Time penalties based on signal timings for pedestrian crossings (see Local Transport Note 2/95 (Department for Transport 1995)) can be allocated to affected links to represent such mid-link delays. Alternatively, pedestrian `crossing signals can be explicitly modelled where outputs from signal control systems are available. In some circumstances, it may be appropriate to use the suburban and small town speed/flow relationships to represent the effects of minor junctions which are not modelled explicitly.

2.9.11 Junction modelling may also be appropriate in **inter-urban** networks where significant delays at major junctions on trunk roads and intersections on motorways occur. In these instances, it will generally be appropriate to use speed/flow relationships on the links between the modelled junctions rather than

[7] The COBA manual has now been withdrawn but is still available from the National Archives: https://webarchive.nationalarchives.gov.uk/20090902134923/http://www.dft.gov.uk/pgr/economics/software/coba11usermanual/

fixed cruise speeds. The **rural** speed/flow relationships given in Appendix E exclude delays at major junctions so the use of these relationships in conjunction with junction delay modelling would not lead to delays being included twice.

2.9.12 Advice on the modelling of delays at motorway merges is provided in Appendix E.

Speed/Flow Relationships

2.9.13 The speed/flow relationships recommended for use in highway assignment models are specified in Appendix E. These relationships are specified in terms of vehicles and the Appendix includes advice on their conversion to Passenger Car Units (PCUs). It also specifies the forms of speed/flow relationship which should be used in highway assignment models to cater for circumstances when assigned demand exceeds capacity, leading to standing queues.

2.9.14 The use of separate speed/flow relationships for light and heavy vehicles is expected in most circumstances. This is because this approach provides more accurate estimates of changes in vehicle operating costs. Having separate speed/flow relationships for goods vehicles may assist modellers in the challenging area of HGV route assignment. This is of crucial importance in environmental assessments, for example of noise and air quality, where accurate flows of goods vehicles will be required at a relatively detailed level.

2.9.15 There are a number of points to note, as follows:

- The **urban** and **small town** speed/flow relationships in COBA apply to all vehicles and do not provide separate estimates for light and heavy vehicle speeds. If the use of these relationships is confined to areas away from the area of influence of the interventions to be tested, the assumption that light and heavy vehicle speeds are the same should be acceptable.

- The **rural** and **suburban** speed/flow relationships provide separate estimates for light and heavy vehicles. Where possible, the modelled journey times for the two vehicle types should be validated separately, which would require journey times to be surveyed in such a way that reliable data for the two separate vehicle types is obtained.

2.9.16 Where a single relationship is used for all vehicle types, it will be necessary to assume a percentage of heavy vehicles so that a flow-weighted average of the relationships for light and heavy vehicles can be derived. The proportion of heavy vehicles will vary by link in the network and will also vary as the assignment calibration proceeds. The obvious source for these proportions is traffic counts by vehicle type. However, the 95% confidence interval for volumes of heavy vehicles from a single day Manual Classified Count (MCC) is ±28%. Moreover, counts will be available for only a sample of the modelled roads. It is therefore recommended that **average** heavy vehicle proportions should be calculated from counts by, at least, road type (motorway, all-purpose dual carriageway, single carriageway) and by type of area (rural, urban central, urban non-central, small towns and suburban). Other categories, such as roads leading to freight generators, should be considered depending on the availability of sufficient count data to support further categorisation.

2.9.17 As the calibration of the assignment model proceeds, the proportions derived from counts should be compared with the proportions derived from the assigned flows. Where discrepancies are significant, the conversions of the capacities and breakpoints in the speed/flow curves should be revised. While it may be impractical to make this comparison at each stage in the calibration of the assignment model, periodic checks should be made and adjustments made as necessary, especially near the end of the calibration process.

2.9.18 The relationships given in Appendix E allow for the effects of road geometry and other attributes to be taken into account. The analyst should consider whether the hilliness and bendiness of the links being modelled are sufficient to have a material effect on the speed/flow relationship. In practice, in most cases only a very small proportion of links within the highway network are likely to require bespoke speed-flow curves to capture their hilliness and bendiness. Highway assignment models often cover a significant study area and measurements of specific geometric attributes should therefore typically only be reserved as a possible calibration technique if modelled journey times provide a poor match against observations.

Driver behaviour

2.9.19 Assignment modelling software typically include parameters that reflect the behaviour of drivers and the capacity restraint behaviour at junctions and along links. Software suppliers typically provide recommendations for how these parameters should be initially set and practitioners should ensure that an appropriate level of consistency across the model is adopted. For example, this may include setting default gap acceptance parameters for particular junction types.

2.9.20 In microsimulation models, driver behaviour is a key component influencing the interaction of vehicles and these assumptions are important in ensuring the model can replicate observed conditions. These behavioural attributes extend to vehicle speed distributions and car following behaviour.

2.10 Relationships with Variable Demand and Public Transport Assignment Models

Introduction

2.10.1 This section provides advice on a number of aspects of highway assignment models which have a bearing on aspects of variable demand models and public transport assignment models. The implications of these matters for demand models and public transport assignment models should be borne in mind when designing a highway assignment model.

2.10.2 It is possible for highway interventions to be modelled using a highway assignment model and demand model on the basis that public transport times are assumed to be constant, both over time and as a result of the highway intervention. In these instances, the implications for public transport models will not be relevant. In multi-modal transport models, however, which are intended for the assessment of both highway and public transport interventions, the implications for public transport assignment models will be important to bear in mind.

Convergence

2.10.3 A high level of convergence for the highway assignment is particularly necessary where the assignment model is linked to a variable demand model because inadequate convergence is likely to result in unstable and unreliable forecasts. TAG unit M2.1 discusses the process of supply/demand convergence.

Compatibility of Highway Assignment, Public Transport Assignment and Demand Model Zones

2.10.4 An identical zoning system for highway and public transport assignment model components of a comprehensive transport modelling system is **not** required or necessarily appropriate. It is often the case that zones for a public transport assignment model need to be sufficiently small around stations and stops for access by walking to be realistically represented by the centroid connector times[8]. It is therefore quite conceivable that the ideal zoning system for a public transport assignment model will be different from the ideal zoning system for a highway assignment model. Forcing one to match the other, or a compromise between zoning systems could lead to unsatisfactory results. A fine zoning system that was a combination of both systems could lead to unacceptably long model run times.

2.10.5 However, within a demand model system there is a need for consistency of highway and public transport zoning to allow the travel choices other than that of route (eg mode or destination) to be appropriately dealt with. To this end, it is desirable that the two zone systems can be nested together, e.g. where public transport zones are smaller than highway zones, the former should be able to be aggregated to the latter. On this basis, and given that public transport zones need never be larger than highway zones, use of the highway model zones as the basis for demand modelling has become standard practice. A further argument supporting these principles is the need for both highway and public transport zone systems to respect the administrative boundaries defined in section 2.3.

2.10.6 An area where the zoning systems for highway and public transport models need to be considered simultaneously is that of the representation of mixed mode trips, commonly referred to as park-and-ride. A common practice is to

[8] This is not always the case as some public transport assignment model software has the capability to spread loading along centroid connectors in a manner that reflects a distribution of trip ends within a large zone.

define an interchange zone representing the park-and-ride car park and public transport stop location. These zones need to be referenced in the zone system for both models.

Highway Networks and Public Transport Assignment Modelling

2.10.7 It is recommended practice that road-based public transport model networks are derived from highway model networks, in order to facilitate adjustment of bus speeds to reflect forecast changes in highway conditions. Alternative practices, where totally separate networks are created (eg a stopping point representation for public transport) are much more problematic in this respect. Where a multi-modal model is being developed, it is therefore necessary to take account of the requirements for public transport modelling in the development of the highway network. The key elements to be considered relate to:

- inclusion of most roads with bus operation
- provision of nodes to represent bus stopping points
- representation of bus priority measures

2.10.8 Development and use of a multi-modal model is most efficient if the representation of bus services is common to both the highway and public transport assignment models. With this arrangement, bus service definitions in the base year and future year scenarios need only be considered once, rather than separately for highway and public transport assignment models.

2.10.9 If roads with bus operations are excluded from the highway network, a unified approach becomes problematic, and so, in general, inclusion of all such roads is desirable. Exceptions could be end points of routes in suburban areas (perhaps where operations are wholly within a single relatively large zone) and minor and infrequent route deviations to serve a location adjacent to a main radial route such as a housing estate.

2.10.10 In many cases, highway network nodes will provide suitable representation of bus stopping points, particularly in denser urban areas, bearing in mind that the aggregation of public transport trip ends to zones means that a precise location for stops can be spuriously detailed[9]. However, this will not always be the case and so some additional highway nodes to represent some bus stops may be needed.

2.10.11 Many urban area highway links contain some form of bus priority such as bus lanes or bus gates. These need to be specifically allowed for in the highway network.

[9] This also applies to large rural zones. There is little point in including large numbers of bus stop nodes within a highway network if they are all included within a single zone.

Consideration of the Interface Between Demand and Supply in the Design of the External Area Network

2.10.12 In the External Area, the use of fixed speeds means that the travel costs are not responsive to levels of demand. However, the fixed speeds should vary by time period to reflect varying levels of congestion.

Time Periods

2.10.13 As discussed in section 2.5, the ideal approach is to ensure consistency between the peak periods covered by the highway assignment and demand models by creating a series of hourly highway assignment models to cover the full peak periods employed in the demand model. For more practical reasons, actual peak hour highway assignment models are often preferred. However, the use of cost changes derived from actual peak **hour** assignment models in the demand model may well lead to peak **period** demand changes being over-estimated.

2.10.14 Sensitivity tests should be carried out using damped cost changes from the actual peak hour assignment model in order to gauge the extent to which the demand changes might be over-estimated. This will assist in determining an appropriate conversion from costs in the peak hour to the peak period.

User Classes

2.10.15 The demand segments usually employed in demand models will be based on trip purpose and car availability (although the car user classes employed in highway assignment models will, by definition, all be in the car available category). The number of trip purposes treated separately in the highway model will normally be fewer than those treated separately in the demand model, so some aggregation will normally be required.

Generalised Costs

2.10.16 In the case of a highway assignment model which operates with a demand model, it is better if the values of vehicle operating cost and value of time derived from TAG unit A1.3 can be retained in the base year assignment without amendment and that they are changed in forecasting in line with the advice in that unit. This approach will maximize consistency between the assignment and the demand model.

2.10.17 In a highway assignment model, the car vehicle operating costs and road tolls and charges should relate to vehicles whereas, in a demand model, they should relate to persons.

Capacity Restraint Modelling in Tiered Model Systems

2.10.18 In some cases, the size and nature of the Fully Modelled Area, and the nature of the preferred assignment method, may mean that convergence between demand changes and cost changes cannot be achieved within practical model run times. Such areas will be characterised by large numbers of small zones and detailed and congested networks. Normally, for areas of this kind, the preferred assignment method would be user equilibrium assignment, with capacity restraint modelled by detailed junction modelling, including flow metering.

2.10.19 One way of reducing run times to more practical levels, while retaining the desired level of detail in the highway assignment model, is to create a **tiered model system**. In such a system, a simplified version of the detailed highway assignment model is created from the detailed version and cost changes from this simplified version are used to derive the demand changes. Because the simplified highway assignment model is able to run to convergence in much shorter times than the detailed model, convergence between demand changes and cost changes can be achieved in shorter times. See Appendix F for further details on tiered model systems.

3. Validation and Convergence Standards

3.1 Introduction

3.1.1 This section provides advice on the following topics:

- validation criteria and guidelines
- convergence measures and acceptable values
- the importance of fitness for purpose of the highway assignment model

3.1.2 In general, the advice in this section 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.

3.1.3 It is important to bear in mind that the role of calibration is to develop a model that is fit for purpose and does not produce unduly misleading or biased results that are material in the context of the schemes or policies being tested. This is discussed further in section 3.2. The issues of calibration and validation should be addressed up front in model development and be part of the Appraisal Specification Report (see Guidance for the Technical Project Manager), agreeing the scope of the model and the purpose for which it will be used.

3.2 Fitness for Purpose

3.2.1 **The test of fitness for purpose of a model is: can robust conclusions be drawn from the model outputs?**

3.2.2 The achievement of the validation guidelines specified in Table 1, Table 2 and Table 3 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'. A model that meets the specified validation standards may not be fit for particular purposes and, conversely, a model that fails to meet to some degree the validation standards may be usable for certain applications. Local Model Validation Reports should therefore **not** include statements to the effect that, **because** the validation standards have been (largely) achieved, the model is **necessarily** fit for purpose.

3.2.3 For a model developed to assess a specific intervention, tests of the sensitivity of the appraisal of that intervention to variations in aspects of the model thought to be weaker and of relevance will show the robustness of the appraisal to uncertainties in the model. Standard output from the model such as assignment flows and journey times are important to check. In addition, it may be useful to conduct Transport Economic Efficiency appraisals, since these are relatively sensitive to uncertainties in the modelling, to ascertain the sensitivity of the appraisal results. This may also be important for assessments of air quality, for example, which are also very sensitive to modelling uncertainties.

3.2.4 For further discussion of this point, particularly in the context of matrix estimation, please also refer to paragraph 3.3.3.

3.2.5 For a general purpose model, it may be useful to carry out a series of demonstration tests so that potential users of the model can gauge the usefulness of the model for particular applications. The range of tests should cover the range of interventions for which the model is intended to be used.

3.3 Validation Criteria and Guidelines

3.3.1 Any adjustments to the model intended to reduce the differences between the modelled and observed data should be regarded as **calibration**. **Validation** simply involves comparing modelled and observed data that is independent from that used in calibration. The extent of data available for model development is often limited and it may be appropriate to use data first for validation through independent testing of other data and model relationships, and then to undertake additional calibration to refine the model. In this case the extent of change from introducing complementary data should be explained.

3.3.2 The differences between modelled and observed data should be quantified and then assessed using some criteria. The recommended proportion of instances where the criteria are met should be assessed. The purpose of this assessment is to explain the confidence that can be placed on the model outputs; it should not be interpreted as a target that the model should be constrained to achieve. Paragraph 3.2.2 gives more detail about the fitness for purpose of models.

3.3.3 In some models, particularly models of large congested areas, it may be difficult to achieve the link flow and journey time validation guidelines as set out later in this section in Table 2 and Table 3 without matrix estimation bringing about changes greater than the limits shown in Table 5. In these cases, the limits set out in Table 5 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.

3.3.4 Outliers should also be examined, even when the criteria in Table 5 are met. Explanations about the relevance of the outliers to the intended uses of the model should be included in the Local Model Validation Report.

3.3.5 The validation of a highway assignment model should include comparisons of the following:

- assigned flows and counts totalled for each screenline or cordon, as a check on the quality of the trip matrices;
- assigned flows and counts on individual links and turning movements at junctions as a check on the quality of the assignment; and
- modelled and observed journey times along routes, as a check on the quality of the network and the assignment.

3.3.6 These are the **main** comparisons that are recommended. Other checks are discussed in sections 6, 7, 8 and 9. The validation **measures**, **criteria** and **guidelines** for each of these comparisons are as follows.

Trip Matrix Validation

3.3.7 For trip matrix validation within traffic assignments, the measure which should be used is the percentage differences between modelled flows and counts. Comparisons at screenline level provide information on the quality of the trip matrices. The validation **criterion** and **guideline** for screenline flows are defined in Table 1. In the first instance the prior matrix should be assigned and tested against this criterion before any impact of matrix estimation takes effect. If matrix estimation is required to improve the model validation, then its impact on screenline flows should be also be monitored.

Table 1 Screenline Flow Validation Criterion and Guideline

Criteria	Guideline
Differences between modelled flows and counts should be less than 5% of the counts	All or nearly all screenlines (ie 95%)

3.3.8 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 (a) where data were used to inform matrix development, (b) for screenlines used as constraints in matrix estimation; and (c) screenlines used for independent validation (as noted in para 3.3.1 there may also be a need to report both validation tests and then the extent of change when data are used to refine the model)
- the comparisons should be presented by vehicle type (preferably cars, light goods vehicles and other goods vehicles)
- the comparisons should be presented separately for each modelled period

Advice on the specification of the screenlines for matrix estimation and validation is provided later in section 4. Further advice on trip matrix verification is provided in TAG unit M2.2.

3.3.9 As explained in TAG unit M2.2, the integrity of the demand matrices with the source data and consistency with the forecasting methods is of particular importance. Where models do not achieve the guidelines the analyst should review the assumptions and quality of data used to develop the trip matrices, but should not impose constraints just to improve the base year flow validation. In reporting the analyst should explain why the model does not reproduce traffic volumes to these tolerances and should indicate the scale and nature of potential forecasting uncertainty and suitability of the model for its intended purpose.

Link Flow and Turning Movement Validation

3.3.10 For flow validation, the measures which should be used are:

- the absolute (vehicles per hour) and percentage differences between modelled flows and counts
- the GEH statistic, which is a form of the Chi-squared statistic that incorporates both relative and absolute errors, and is defined as follows:

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

where: GEH is the GEH statistic

M is the modelled hourly flow

C is the observed hourly flow

These two measures are broadly consistent and flows that meet either criterion should be regarded as satisfactory.

3.3.11 The validation criteria and guidelines for link flows and turning movements are defined in Table 2.

Table 2 Link Flow and Turning Movement Validation Criteria and Guidelines

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

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

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

Consideration of count accuracy in calibration and validation is provided later in section 4. Further advice on assignment validation is provided in section 9. Where models do not achieve the guidelines the analyst should review the network coding quality and the local quality of the trip matrices, but should not impose constraints just to improve the base year accuracy of the model. The focus here should be to ensure that the model is suitable for its intended purpose within the critical area for the interventions that are to be tested. In reporting the analyst should explain why the model does not reproduce traffic volumes to these tolerances and should indicate the scale and nature of potential forecasting uncertainty and suitability of the model for its intended purpose. Paragraph 3.2.2 gives more detail about the fitness for purpose of models.

3.3.13 Experience has shown that the level of model validation outlined in this section results in a robust standard of traffic model used for major scheme appraisal.

The greater the difference in modelled flows from observed flows, noting that there will also be uncertainty and variation in observed flow data, the wider the uncertainty around the performance of the model and hence the resulting appraisal results. Practitioners should examine the extent to which this affects the robustness of their models on a case-by-case basis. This will depend on factors such as the proximity of poorly validating links to the scheme (or schemes) to be tested and the degree to which modelled flows differ from observed.

3.3.14 An important consideration is to add wider context and interpretation to the model performance by including narrative about its fitness for purpose in addition to presenting validation statistics for links or cordons in tabular form. For example, a way to present outputs geographically would be through the inclusion of network diagrams that illustrate how the model performs in the area around the scheme and in the wider modelled area. If there are particular areas where the model validation performs poorly, practitioners should highlight these to ensure that potential weaknesses in the model are understood.

Journey Time Validation

3.3.15 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. The validation criterion and guideline for journey times are defined in Table 3.

Table 3 Journey Time Validation Criterion and Guideline

Criteria	Guideline
Modelled times along routes should be within 15% of surveyed times (or 1 minute, if higher than 15%)	> 85% of routes

3.3.16 With regard to the journey time validation, the following should be noted:

- it is expected that separate speed/flow relationships and/or link speeds are used for light and other vehicles; hence comparisons should be presented for light and other vehicles separately; otherwise, the comparisons should be presented for all vehicle types together;
- for validation of journey times by vehicle type, it will be necessary to obtain journey times by vehicle type to a level of accuracy which will allow a meaningful validation; if journey times by vehicle type are not available but separate speed/flow relationships for light and heavy vehicles have been used, a weighted average of the modelled light and heavy vehicle speeds should be compared with the surveyed all-vehicle speed; and
- the comparisons should be presented separately for each modelled period.

3.3.17 Advice on the specification of the journey time routes and sources of observed journey times is provided later in section 4.3. Further advice on assignment validation is provided in section 9.

3.4 Convergence Measures and Acceptable Values

Introduction

3.4.1 Before the results of any traffic assignment are used to influence decisions, the stability (or degree of convergence) of the assignment must be confirmed at the appropriate level. The importance of achieving convergence, at an appropriate level, is related to the need to provide stable, consistent and robust model results. When the model outputs are being used to compare the with-scheme and without-scheme cases, and especially when estimating the Transport Economic Efficiency (TEE) impacts of a scheme, it is important to be able to distinguish differences due to the scheme from those associated with different degrees of convergence. Similar considerations apply when the benefits and disbenefits of different interventions are being compared. Model convergence is therefore key to robust TEE appraisal.

3.4.2 The following discussion, with convergence criteria summarised in Table 4 at the end of the section, introduces the different aspects of model stability and proximity as indicators of convergence. The background to the topic is set out in greater detail in Appendix D. Achieving these criteria may require a different number of iterations for the without-scheme and with-scheme cases, but the target measures of convergence should remain consistent. Note that the number of iterations is not a measure of convergence and that, in general, models should be run until the convergence criteria are met rather than for a pre-determined number of iterations; the latter is simply an indication of the effort expended in search of equilibrium.

3.4.3 In general, the iterative methods for reaching equilibrium required in most assignment algorithms will not converge absolutely, and user-defined 'stopping criteria' are required to describe the point at which satisfactory convergence is considered to have been achieved. The convergence indicators provided by different software packages vary, as does the availability of a facility for the user to control the assignment process to ensure a given level of convergence. Care needs to be exercised to distinguish between convergence and stability. Stability can often be achieved, indeed some methods such as the Method of Successive Averages (MSA) force it to happen, without there necessarily being convergence to a solution.

3.4.4 The assignment methods to which the convergence and stability measures apply are described in section 2.7.

User Equilibrium Assignment

3.4.5 Most assignment packages that are suitable for congested networks containing alternative routes between zones provide algorithms which seek to achieve

Wardrop's First Principle of Traffic Equilibrium or **User Equilibrium** (see section 2.7). These packages provide one or more of the following convergence indicators:

- the percentage of links on which flows or costs change by less than a fixed percentage (recommended as 1%, see Table 4) between successive iterations, sometimes known as 'P' or 'P2' respectively
- the difference between the costs along the chosen routes and those along the minimum cost routes, summed across the whole network, and expressed as a percentage of the minimum costs, usually known as 'Delta' or the '%GAP'
- the degree to which the total area under the cost/flow relationships is minimised (also the uncertainty in the objective function), sometimes known as 'Epsilon'
- the percentage change in total user costs or time spent in the network between successive iterations, sometimes known as 'V'

3.4.6 The percentages of links with small flow or cost changes both provide pragmatic views of the stability of the assignment, rather than the degree of convergence. The measures are necessary but not sufficient indicators of convergence. It is recommended that, in addition to satisfying the true convergence measures described below, assignment model iterations should continue until at least four successive values of 'P' or 'P2' in excess of 98% have been obtained. If this cannot be achieved, especially in a future year assignment, this may be an indication of instability caused by the level of traffic demand being higher than can be absorbed by the network capacity.

3.4.7 The Delta statistic or %GAP (see Appendix D for definitions) is a truer measure of convergence but may not be provided by all packages. Delta values generally decrease towards a minimum value as the number of iterations increases but will not do so monotonically. Delta has traditionally been preferred over Epsilon and should be used as the first choice measure of assignment convergence.

3.4.8 In all cases, supporting information on stability, including acceptable values when measured against the other criteria, should be provided. The attainment of high degrees of convergence is particularly necessary where the assignment modelling is linked to a variable demand model because inadequate convergence is likely to result in unstable and unreliable forecasts. See TAG unit M2.1.

3.4.9 **Origin-Based Assignment** (OBA) is a more recently developed algorithm for computing a Wardrop equilibrium which, in principle, guarantees convergence to an 'exact' solution by eliminating non-optimum paths commonly found in solutions generated by more conventional equilibrium algorithms. Convergence may be monitored by both Delta and %GAP, as for conventional approaches. For further details see Appendix B:.

Stochastic User Equilibrium Assignment

3.4.10 Stochastic User Equilibrium (SUE) assignment algorithms typically converge more slowly than Wardrop User Equilibrium algorithms, although the speed of convergence depends on the algorithm used. It is also far more difficult to monitor convergence with SUE assignment. The process has to converge with respect to both the effects of congestion (as with Wardrop) plus the effects of random perceptions and, unlike Wardrop equilibrium, there are no parameters which definitely reduce to zero at perfect convergence - even if two successive iterations give identical results this could simply be the result of a singular set of random numbers.

3.4.11 It is recommended that the percentage change in total user costs between successive iterations ('V') should be used to monitor convergence. At convergence, the total costs could be expected to fluctuate randomly about the 'true' value; therefore a consistent increasing or decreasing trend in total costs is indicative of a lack of convergence. In this case, iterations should continue until four successive absolute values of 'V' less than 0.05% have been obtained, recognising that the measure should approach zero at convergence but, for the reasons noted above, will never exactly equal zero.

Acceptable Levels of Convergence

3.4.12 During the development of the base year model, to ensure that reasonable levels of convergence are achieved, sufficient iterations should be carried out to achieve an acceptably low value for %GAP. A guideline target for this is 0.1% or less. In previous guidance this guideline was more relaxed. Improving computing power and software design has enabled many more iterations to make an improved target achievable.

3.4.13 A level of convergence which is sufficient to ensure that scheme benefits can be estimated robustly above model 'noise' is essential and a lower value of %GAP than the 0.1% guideline may need to be achieved. More iterations may be required in the forecast year, when congestion levels are forecast to be higher, simply to achieve the base year value of 0.1%, and even more iterations will be required to achieve the lower %GAP values required for robust economic appraisal.

3.4.14 However, it is clearly difficult to be precise about the appropriate level of convergence at the outset of model development. As soon as is practically possible, the analyst should, assess the model 'noise' within each model run as well as between model runs, that is, between a without-scheme model run and a with-scheme model run. This should assess the overall change in vehicle (or PCU) hours between assignment loops (the model 'noise') and the difference in vehicle (or PCU) hours between corresponding loops in the without-scheme and with-scheme model runs.

3.4.15 As noted in TAG unit M2.1, it is also good practice to check the size of the %GAP in relation to the scale of the user benefits of the tested intervention to the total network costs. Ideally, the user benefits as a percentage of network costs should be at least ten times the % GAP achieved in the without-scheme and with-scheme scenarios, although smaller values are not necessarily indicative of a problem.

3.4.16 Experience has shown that %GAP values of less than 0.05% have often been achieved in order to provide a more robust basis for economic appraisal of highway schemes. The larger the model, the more difficult it will be to attain the requirements set out in TAG unit M2.1. In fact, this target will often be impossible to achieve in a large model and hence a balance needs to be attained between practical model run times and convergence levels. Whilst computers are becoming faster, benefits in processing speed could be countered by the development of larger and/or more sophisticated models. For smaller networks, smaller values of %GAP may be achievable without an excessive number of iterations. The potential need for more stringent convergence standards for scheme appraisal applies to the other measures of convergence as well as to %GAP.

3.4.17 Table 4 summarises the most appropriate convergence measures (of proximity and stability) and the values generally considered acceptable for use in establishing a base model. **Tighter levels of convergence may be required for scheme appraisal**.

Table 4 Summary of Convergence Measures and Base Model Acceptable Values

Measure of Convergence	Base Model Acceptable Values
Delta and %GAP	Less than 0.1% or at least stable with convergence fully documented and all other criteria met
Percentage of links with flow change (P)<1%	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)

4. Calibration and Validation Data

4.1 Introduction

4.1.1 This section contains guidance regarding the use of specific highway survey data for the calibration and validation of models.

4.1.2 For guidance on the availability of highway survey data and the conduct of surveys, see TAG unit M1.2 - Data Sources and Surveys. This unit also contains guidance on the assumed confidence intervals for surveys, the accuracy of survey data and guidance on the factoring of data.

4.1.3 Calibration and validation data are generally of two kinds: traffic counts, and journey times.

4.1.4 Traffic counts may be obtained by automatic means (Automatic Traffic Counts, ATCs) or manually (Manual Classified Counts, MCCs). Journey times may be obtained from commercial sources such as tracked vehicle data or camera observations from Automatic Number-Plate Recognition systems (ANPR).

4.2 Traffic Counts for Matrix Estimation and Model Validation

4.2.1 The main purpose of matrix estimation is to refine estimates of those movements that do not provide an accurate match against observed volumes, often due to the high-level nature of either the data source or expansion method used to derive them. TAG unit 2.2 provides advice on interpreting the relative confidence of synthesised demand and different data sources. For some data sources, such as RSI, counts are directly used for data expansion. A hierarchy may be considered distinguishing screenlines:

- directly used in expanding source data (and not therefore independent) where relevant to the data used to develop prior trip matrices
- for calibration, ie used in matrix estimation or to help inform prior matrix development
- for validation

As noted in paragraph 3.3.1 it may be judged that there are insufficient data to retain screenlines purely for independent validation. Nevertheless, the hierarchy should be considered for staged or sequential use of the data, with the hierarchy designed to give progressively finer spatial granularity (ie 'validation' screenlines would generally intercept movements within sectors defined by calibration screenlines).

4.2.2 The density of screenlines should be designed so that the majority of intra-sector movements of relevance to the model purpose are subject to comparison with counts. In designing the screenlines, account should therefore be taken of the trip length distribution. The best estimate of the mean length of these trips is likely to be the synthetic matrices.

4.2.3 Matrix estimation should be applied to individual vehicle type matrices because the routes used in the matrix estimation will vary by user class. This means that MCCs are required at the sites where constraints are to be applied. The use of average vehicle proportions to obtain vehicle splits by type in the absence of MCCs should be avoided where possible.

4.2.4 To enable matrix estimation to adjust the prior matrices to approximately the correct overall levels, ATCs are also required at the constraint sites. Thus, the ATCs should be used to give the total vehicles, and the MCCs to provide the split by vehicle type.

4.2.5 Turning movement counts should only be used as constraints in matrix estimation if they have been derived from both MCCs **and** ATCs. In the absence of ATCs, turning movement counts should be used mainly as a diagnostic during model calibration. Alternatively, ANPR and video counts should provide sufficient accuracy.

4.2.6 Neither ATCs nor MCCs will yield counts of light and heavy goods vehicles which are sufficiently accurate for the validation of the assigned flows of these vehicle types on individual links. Validation of these vehicle types will therefore generally need to be reported for short screenlines using grouped counts which have sufficiently small confidence intervals.

4.2.7 Validation of assigned flows on individual links will generally have to be restricted to cars and all vehicles. The counts of all vehicles used for this purpose should be ATCs with MCCs being used to determine the proportion of cars.

4.3 Journey Times for Calibration and Validation

4.3.1 The importance of setting accurate cruise speeds has been explained in section 2.9. Cruise speeds must distinguish delays from junctions. Comparisons will show how well total link times are modelled and therefore, given that the cruise speeds may have been largely derived from observations, how well junction delays are represented.

4.3.2 For general purpose models, the routes for the validation of journey times should cover as wide a range of route types as possible and cover the Fully Modelled Area as evenly as possible. For models developed for the appraisal of specific interventions, routes should include those from which it is expected traffic will be affected by the scheme, as well as covering the scheme itself as appropriate.

4.3.3 The validation routes should be neither excessively long (greater than 15 km) nor excessively short (less than 3 km). Routes should not take longer to travel than the modelled time periods (although, a few minutes longer is unlikely to be problematic). Where Moving Car Observer (MCO) surveys are undertaken start times should be staggered, particularly if runs are undertaken on the same day. For models of actual peak hours, journey time routes ought to be no longer than about 40 minutes to allow some staggering of start times.

4.3.4 As described, it is standard practice to use journey time validation at the route level. However, increasingly there is a need to take a more detailed approach and check journey time validation at the link level or for segments of the route as well. This can be very important to assess noise and air quality impacts in the detail that they are required. Where these impacts may be material, the analyst should produce some assessment of the accuracy of speeds at a finer level.

5. Network Data, Coding and Checking

5.1 Introduction

5.1.1 Advice is provided in section 2.4 on the design of the structure of the networks in the Fully Modelled and External Areas. The data required to describe the networks were also identified there. In this section, sources of network data are discussed and advice is provided on the coding and checking of network data.

5.1.2 For additional guidance on the availability of network data sources see TAG unit M1.2 - Data Sources and Surveys.

5.2 Network Data and Coding

5.2.1 Network descriptions in the Area of Detailed Modelling will generally need to include both link and junction details.

Link Representation

5.2.2 **Links** are usually described in terms of:

- the reference numbers at the ends of the link (ie 'nodes')
- the link length
- the cruise speed in the base year, defined as the mid-link speed, separate from any junction delay, during the time period modelled
- the speed/flow relationship (if any) appropriate for the link (but see section 2.9 for further advice)
- whether the link operates in both directions or in one direction only
- any restrictions to particular vehicle types using the link

5.2.3 Highway assignment models should represent **mean** traffic speeds[10]. Therefore, cruise speeds should represent the mean speed of traffic between junction queues, given the activity alongside and crossing the link in the time

[10] In this context, the mean speed referred to here is the mean of the speeds of all vehicles, or all vehicles in each vehicle type for which separate speed/flow relationships have been defined.

period concerned. Cruise speeds are therefore **not** free-flow speeds and should **not** necessarily be set at the speed limits either.

5.2.4 In urban areas, it may sometimes be necessary to consider the impact of traffic management measures such as bus lanes, traffic calming, parking controls and cycleways, on the capacity and cruise speed of individual network links. The network coding should vary between time periods to reflect variations in the application of bus lanes and parking controls by time of day. The way in which the above information is input to a model may vary slightly according to the software package that is being used.

Junction Representation

5.2.5 The way in which **junctions** are described varies between different modelling packages. The usual requirements are:

- the junction type (traffic signals, roundabouts, priority)
- the number of approach arms and their order (in terms of entry link references)
- the number and width of traffic lanes on each junction approach, the flare length, and the lane discipline adopted (including prohibited turns)
- any additional data required to describe the operational characteristics of the junction (eg saturation flows, signal timings, phasing and method of control, turning radii, visibility, yellow boxes and gap acceptance characteristics)

5.2.6 In practice, the way that junctions are used may not accord with the layout painted on the road surface. For instance, where no lane is reserved for right-turning vehicles, there may be sufficient room to accommodate one or two right-turning vehicles without impeding straight-ahead traffic and, in these instances, the coding should reflect actual use rather than the painted layout.

5.2.7 Modern data sources provide the analyst with a large array of information to assist in network coding. The following framework and data sources are suggested for coding a network:

- determine the number of approach arms, and establish their order (in terms of entry link references). A site visit will add value over using desk-based sources alone and can obtain greater clarification
- derive a sensible structure for numbering of nodes (and zones) - preparing a structured numbering system will save time and effort when making modifications later in the process
- identify 'complex' junctions, ie gyratories and signalised roundabouts, at the outset as these will need to be 'expanded' rather than treated as single nodes

- confirm the network structure and a-node/b-node connectivity - time spent in preparing the structure of the network before coding commences will avoid problems of mis-specification of connections later in the process

- derive geographical co-ordinates for all nodes

- ensure one way links/banned turns are incorporated

- for large models, on-site network inventories may not be affordable, in which case, obtain aerial photographs and maps of the area to be coded; a number of products are available, such Google Maps/Earth/Streetview, and Bing Maps, which can be extremely helpful in verifying junction layouts, types and geometry[11]

- obtain link lengths from a reliable source (e.g. Ordnance Survey MasterMap Highways Network), measuring manually where necessary (take care to ensure a->b and b->a link lengths are 'compatible')

- agree and/or set out default standard values for key input parameters (saturation flow rates, minimum gap values, cruise speeds, speed/flow relationships) at the outset, which can be reviewed subsequently, if necessary

5.2.8 Traffic signal timings can be divided into two types: UTC (Urban Traffic Control) and non-UTC. UTC allows the co-ordination and control of signal timings at junctions over an area from a central location. UTC can be further divided into SCOOT (Split Cycle Offset Optimisation Technique) and fixed time plans although the latter may or may not contain demand dependent stages.

5.2.9 Non-UTC junctions operate under vehicle actuation or MOVA (Microprocessor Optimised Vehicle Actuated). Many modelling packages are unable to represent the dynamic nature of such signal timing regimes and compromises need to be made when coding signal timings. The initial source for signal timings should be the relevant Highway Authority(ies) in the area being modelled.

5.2.10 The key requirement is to process the observed signal timing data such that the representative timings for the periods being modelled are coded. Where fixed time plans dominate in the peak periods, these should be used as the basis for the coded signal timings. Analysis of SCOOT/vehicle actuated signal timings will be necessary in order to develop representative cycle times, stage lengths, offsets and timings for groups of junction within a region or group.

5.2.11 The location of traffic signals in close proximity and the presence of upstream traffic signals altering the arrival pattern of traffic at downstream traffic signals often means that improved traffic flow can result by co-ordinating or linking traffic signals. The following terms are used when discussing traffic signal co-ordination:

---

[11] It should be noted that the age of such sources are often indeterminate, so care is required where more recent changes may have occurred to the network.

- phase - a predetermined set of traffic signal movements that operate concurrently
- stage - a group of phases running together
- cycle time - the total time taken to run once through all phases
- offset - this refers to the given time relative to that at a reference junction of the beginning or end of the green period of all other junctions that are co-ordinated

5.2.12 Co-ordination is achieved by:

- traffic signals running on a common cycle time (or in special cases, one half of the common cycle time, known as double-cycling)
- the offset is determined by the distance between the signals, the speed of progression along the road, and the queues of vehicles waiting at signals with a red aspect
- the optimisation of offsets and phase times

5.2.13 Saturation flow rates represent the maximum number of vehicles (or more commonly PCUs) per hour which can cross a stop-line unopposed by other vehicles or red lights. Many packages require saturation flow rates to be specified by individual turns.

5.2.14 In view of the need to maintain consistency, it is recommended that standard procedures/tables of values are derived at the outset. Road Note 34 (Road Research Laboratory 1963) describes the standard (manual) method of measuring saturation flow rates[12]. It is recommended that local measurements of saturation flow rates be undertaken and that these surveys should include a cross-section of junction types and layouts.

5.3 Network Checking

5.3.1 Most network modelling packages contain procedures for checking the integrity of the network description. However, with large models, the multiple levels of warnings and errors and the quantity of diagnostic output that many such packages generate often means that such diagnostics can be over-looked. Resources should be set aside to examine all the warnings and other diagnostics.

5.3.2 The following is a basic checklist of items designed to minimise problems once the network has been coded:

[12] This has been codified into TRL's SATFLOW program, part of their Traffic Engineering Software BUNDLE 3.0. This program enables saturation flow rate measurements to be undertaken on a palmtop computer. TRL note RR67 has further information on the calculation of saturation flows.

- check for appropriate junction types
- check that the appropriate number of entry lanes have been coded and that flaring of approaches, where appropriate, are accounted for
- check that turn restrictions have been correctly identified (these may vary by time period)
- check that one-way roads and no entries, if applicable, have been correctly specified
- check that saturation flows are appropriate (particularly if turn rates appear excessively high or low compared to straight ahead)
- check that link lengths, link types and cruise speeds for each direction of a link are consistent, and that the second and third do not vary unjustifiably along series of links
- compare crow-fly link lengths against actual lengths, check that the coded link length is between 1.1 and 1.3 times the crow-fly distance and inspect links which fall outside this range. It is often the case that analysts may re-locate nodes to enhance the representation of a junction on-screen. Whilst this does not affect the coded link length it will obviously affect the crow-fly distance. Given the eventual need to interface with other models for environmental assessment of noise and air quality, experience shows that it is preferable to ensure that traffic model nodes are positioned according to their actual geographical location.

5.3.3 In addressing errors, the network structure and connectivity should be checked first, followed by the link attributes and then, if necessary, the more detailed (junction-related) data such as junction type, lane usage, saturation flows, signal timings and so on, should be checked.

5.3.4 The use of standardised templates and error checking procedures is recommended. Coding assumptions and default values should be set out as early as possible and before any coding is attempted. The development of a standardised approach to coding is important as this task is often carried out by a number of people and consistency of approach is necessary.

5.3.5 The Ordnance Survey OpenData datasets[13] can be useful in verifying network structure and, in conjunction with GIS software packages, can be used as a backdrop to modelled network and zoning systems. The OS MasterMap Highways Network dataset is also a useful source for this.

5.4 Pre-Calibration Checks

5.4.1 Calibration should not start unless the network has passed a series of basic checks as described here. The network calibration should not be viewed as a means of debugging a network containing errors, although in practice some

[13] Available at www.ordnancesurvey.co.uk/opendata

errors not captured in the initial network coding checks may manifest themselves during that process. The intention should be to identify and eradicate, as far as possible, common sources of error such that the calibration (and validation) process is largely confined to monitoring network performance and fine-tuning it to match observed conditions. This latter process is described in section 6.2.

Network Documentation and Completeness

5.4.2 It is useful to document headline statistics, such as the number of links, nodes and the version of the software used to develop the network. It is possible that new versions of software may be introduced during the course of the development of a model which may update or upgrade certain aspects of network coding.

5.4.3 Data sources should be documented and a map or plan showing which parts of the network have been coded using newly surveyed or derived data and those parts which have been coded or derived from existing models should be prepared. The location of any junction inventory surveys, signal timing and saturation flow measurement surveys should also be documented and mapped. The intention is to provide an overview of the extent of the coded network and data sources.

Network Compilation

5.4.4 Invariably, each transport modelling software package has its own error trapping procedures so it is not possible to be prescriptive on what checks should be applied. Nevertheless, certain errors will cause the network build process to fail and these, clearly, need to be addressed. If the software package provides graded error messages, the analyst should seek to ensure that a satisfactory explanation for each type of warning (not necessarily each warning) is documented if steps to remove the warning(s) are judged to be unnecessary.

Consistency of Coding

5.4.5 It is commonplace, particularly in the development of large models, for network coding to be undertaken by a team of analysts. Consistency of approach is essential to ensure that the network is coded in a uniform fashion using a consistent set of rules. Documentation of network coding assumptions is essential in this respect. Whilst it is recognised that assumptions can be changed (during calibration and validation), the need to record a consistent and sensible starting set of assumptions is important. This should minimise the potential for excessive adjustments later in the model development process.

5.4.6 The selection of appropriate link types and cruise speeds within the Fully Modelled Area is important. Where a link type and/or cruise speed changes on a specific link, checks should be undertaken and mapped to demonstrate that this accords with the source data. A check should also be undertaken comparing the coded link lengths against suitable source data (e.g. the OS

MasterMap Highways Network data). Justification for using coded links of length significantly greater than the crow-fly distance should be documented.

Check of Key Junctions

5.4.7 It is important to demonstrate that key junctions and intersections which have the greatest influence in model calibration and validation, are coded appropriately. All key junctions and intersections should be formally reviewed. The review should seek to ensure that junction types are correctly defined and that the representation of key characteristics is appropriate and accords with available data sources.

5.4.8 Again, different software packages represent junctions in different ways but the following should be viewed as a reasonably common source of information for review. **For all junction types**: number of lanes at the stop line; number of lanes on the main (mid-link) approach; link type classification/cruise speed attributed; representation of flaring; lane usage/definition of turns; representation of bus lanes; and representation of banned turns. **For traffic signals**: representation of filters; definition of stages; cycle times and offsets; green times; and inter-green times.

Network Connectivity

5.4.9 It is important to ensure that the network is correctly connected and that all destinations can be reached. Whilst some software can undertake such tests automatically, the assignment of a matrix of small values in each cell to ensure that all trips assign is a straightforward task. This also allows the preliminary investigation of paths through the network which may highlight unused links (cruise speeds too slow or link lengths too long, perhaps) or heavily used links. Plotting of minimum path trees, even at this early stage, can be a useful way of trapping coding errors.

6. Network Calibration and Validation

6.1 Introduction

6.1.1 Initial checks that should be undertaken during the development of a network were outlined in section 5. It is possible that, despite passing through all such checks, some aspects of the coding may still be 'wrong', in that the model may be unable to replicate observed conditions to a satisfactory degree. Whilst this could reflect on issues outside the quality of the network coding (that is, it could be related to the level of demand), there are certain aspects of the network coding that may benefit from further checking before the need to make adjustment to demand is contemplated. It is recommended that this process of network calibration is undertaken before adjustments to demand (matrix estimation) are carried out.

6.1.2 The process of adjusting an initial network which has been subjected to a standard set of data range and logic checks is that of **calibration**. Network calibration requires examination of preliminary assignments prior to the calibration of trip matrices. This is an important step. Networks should be debugged before being used to adjust the trip matrices, otherwise matrix adjustments may be made which compensate for and hide problems within the network.

6.2 Network Calibration

6.2.1 Section 5.4 provides guidance on the prerequisite checks required on the network coding before an appropriate calibration of the network can proceed.

6.2.2 Network calibration, with an initial version of the trip matrix, should include checks to ensure that speeds and flows on network links and delays at junctions are as expected. For links, these checks should include both speed and flow comparisons at locations where suitable observations are available. It may also be useful to check flow/capacity ratios.

6.2.3 At junctions, remedial action should be considered for any turning movement where:

- the capacity calculated by the model is less than the count or
- calculated delays are significantly greater than observed delays

This will usually involve reviewing the parameters which control the capacity of the movements affected, or possibly reviewing the position of nearby centroid connectors.

6.2.4 Having made such adjustments, a second set of adjustments should be considered where:

- modelled flows are significantly below observed flows for a particular turning movement or
- modelled delays are unacceptably lower than observed delays

6.2.5 Adjustments should only be made to network descriptions if they can be justified **and these should be documented**. Arbitrary adjustments to measurable quantities (eg link length or junction geometry) should **not** be made. Artificial and/or excessive adjustment of cruise speeds, or link or junction capacities to give a closer fit to observed conditions is **not** recommended.

6.2.6 Further checks should be carried out by inspecting the routes through the network taken by selected traffic movements. This is complicated by the fact that capacity restraint procedures calculate several sets of routes. Plotting of routes used between key origins and destinations should be undertaken to verify that the routes appear plausible.

6.2.7 When inspecting routes, the likely differences between the routes taken by HGVs and other vehicles should be recognised. Actual HGV routes may differ from the routes taken by other traffic for a variety of reasons, including different acceleration and speeds, associated fuel costs, possibly different geometric delays at junctions, and general propensity for HGVs to use the main, higher capacity, roads and to avoid use of local, narrower, roads, especially roads with vehicle height and width restrictions. Separate speed/flow curves can partially deal with these issues.

6.2.8 It is not envisaged that large changes to saturation flow rates, signal timings or cruise speeds would be necessary at the calibration stage. Adjustments should be made for valid traffic reasons. For example, closer inspection may reveal that short-term peak hour parking close to a junction may have a detrimental effect on the saturation flow rates assumed which may not have been taken into account, or the poor visibility from a side road onto a main road has been under-estimated or not taken into account and, as a result, the saturation flow rates are over-estimated. While changes in the order of ±20% may be justifiable, larger adjustments would benefit from further explanation and documentation. In some cases, site visits may be necessary to gain more insight into how the network and junctions are used.

6.2.9 In addressing problems, it is recommended that structure and connectivity should be checked first, followed by link attributes and then, if necessary, more detailed (junction-related) data such as junction type, lane usage, saturation flows and signal timings.

6.3 Network Validation

6.3.1 It is not possible to validate a network in isolation, since the output traffic flows and travel times will reflect not only errors in the network, but also those inherited from the input trip matrix. This is a particularly important consideration in congested urban areas, where relatively small discrepancies in a trip matrix can have a disproportionate impact on junction delays and hence on the routes taken by vehicles through the network.

6.3.2 It would not be possible to undertake a meaningful comparison, or validation, of journey times based on an assignment of a unit matrix as junction delays will not be modelled correctly. Once an initial estimate of a prior matrix is available, more meaningful comparisons of observed and modelled journey times can be made.

6.3.3 Areas of the network which show differences at a route level of greater than 25% should be investigated. This may, nonetheless, hide problems along the route so time/distance graphs are an essential means of highlighting problems en-route.

6.3.4 Once the trip matrix has been finalised, model validation will include the validation of modelled journey times against observed data (see section 9).

7. Route Choice Calibration and Validation

7.1 Introduction

7.1.1 The calibration of an assignment model should be, as far as possible, a sequential process. Thus, effort should be made to ensure that each element in the sequence - zones, network structure, centroid connectors, network coding, capacity restraint procedures and trip matrices - is developed as accurately as reasonably possible before moving on to the next. Nevertheless, some iteration will be inevitable, with adjustments potentially being made to each of the constituent elements as part of the process of refining the accuracy of the assignment.

7.1.2 The accuracy of the assignment will be dependent not only on the accuracy of the constituent elements listed above but also on the realism of the modelled routes. It is very rare for surveys of actual routes to be undertaken and the best that can usually be done is that the analyst assesses the plausibility of the modelled routes.

7.1.3 The modelled routes will depend on:

- the appropriateness of the zone sizes and modelled network structure and the realism of the 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 times along links are modelled, which are dependent not only on data and/or coding accuracy and appropriateness but also on the appropriateness of the approximations inherent in the junction flow/delay and link speed/flow relationships
- the accuracy of the trip matrices which, when assigned, will lead to the times used in the route choice process (via the flow/delay and speed/flow relationships)

7.1.4 At various stages in the model development process, modelled routes should be examined and their plausibility checked. For example, as suggested in paragraph 6.2.6, early plotting of minimum path routes is a useful way of identifying network coding errors. Modelled routes may also be usefully considered in the later stages of calibrating the model.

7.1.5 Whether or not route choice is 'calibrated', it is essential that the plausibility of the modelled routes is displayed and assessed. This may be viewed as 'validating' the modelled routes and the way that this should be done is described in the final part of this section.

7.2 Route Choice Calibration for HGVs

7.2.1 The generalised cost formulation for highway assignment modelling and the source for the coefficients are set out in section 2.8. As explained there and in section 2.6, the distance coefficient (given by vehicle operating cost / value of time) will vary by user class, that is, by vehicle type and trip purpose combination. However, as also explained in section 2.8, changes to the distance coefficients should no longer be used as a means of calibrating route choice.

7.2.2 It is often the case that the routes based on generalised costs given in TAG for heavy goods vehicles do not appear to take full account of the attractiveness of motorways and trunk roads and the unattractiveness of local roads for these vehicles. While the route choice calibration process described above applies, in principle, equally to all user classes, heavy goods vehicle routes may require further special attention.

7.2.3 Distance coefficients given in TAG for HGVs can have the effect of deterring these vehicles from being assigned to longer, faster routes, such as motorways and trunk roads in favour of shorter and slower routes. If HGV routeing is considered implausible, special modifications to the distance coefficients should be considered. One approach is to introduce a link-based calculation of generalised costs so that, for HGVs, longer, faster routes such as motorways and trunk roads appear more attractive. This would enable an alternative distance coefficient to be used on such routes and ensure that HGVs were not deterred from using such routes in the assignment. This would need experimentation to determine appropriate values to ensure plausible routeing but a good starting point would be to adopt the car distance coefficients.

7.2.4 Further adjustments that may also be considered are to use separate HGV speed/flow curves (in rural areas) and/or to set HGV-specific link cruise speeds and/or to include HGV-specific penalties to represent geometric delay at junctions.

7.2.5 Whichever approach is adopted, details should be provided in the Local Model Validation Report.

7.3 Route Choice Validation for Private Travel

7.3.1 As the calibration of the assignment proceeds, checks should be carried out by inspecting the routes through the network taken by selected traffic movements. Plotting of trees, that is routes from an origin to all destinations, is easy to do but not particularly informative. A better approach is to examine the modelled routes between selected origins and destinations. These **selected** origins and destinations should focus on important centres of population and employment or key intersections. These should be chosen so that the routes:

- relate to significant numbers of trips
- are of significant length or cost (eg 20+ minutes)
- pass through areas of interest (eg scheme impacted areas)
- include both directions of travel (to sense check differences)
- link different compass areas (eg north to south, east to west, etc.)
- coincide with journey time routes as appropriate

The routes modelled for each user class should be examined separately.

7.3.2 The number of pairs of zones which should be examined and displayed will be dependent on the size of the model. The following rule of thumb should be used:

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

7.3.3 Observations of routes are not usually available, so these checks must be based on local knowledge and judgement and hence cannot be regarded as true validation. Nonetheless, the results from the zone-to-zone route plots should be documented in the Local Model Validation Report and any remedial action taken explained.

8. Trip Matrix Calibration and Validation

8.1 Introduction

8.1.1 Trip matrices should be created following the advice referred to in TAG unit M2.2. The process of producing trip matrices involves the following steps as outlined in TAG unit M2.2:

- planning
- data assembly
- matrix development
- matrix refinements

8.1.2 It is common for the trip matrices produced by the methods referred to in TAG unit M2.2 to be termed 'prior' trip matrices. The adjective 'prior' usually refers to matrices which have not (yet) been subjected to matrix estimation. That convention is adopted here.

8.1.3 Having designed the zoning system, network structure and centroid connectors, and calibrated and validated the network and routes, it will be essential to validate the trip matrices by comparing assigned flows with traffic counts. Depending on the proximity of the modelled flows and counts, three courses of action are possible:

- reconsider the development of the prior trip matrices with a view to producing new versions which, when assigned, yield modelled flows which accord more closely with the counts
- refine the trip matrices using **matrix estimation**, as discussed below in this section
- explain the model performance and potential limitations for its use

8.2 Validation of the Prior Trip Matrices

8.2.1 It is important that the need for either matrix estimation or other adjustments to the prior trip matrices is clearly established before any such processes are carried out. The prior trip matrices should therefore be validated by comparing total screenline and cordon modelled flows and counts by vehicle type and time period.

8.2.2 As explained in TAG unit M2.2 screenline count evidence should be considered in developing prior matrices. Documentation reporting the matrix development should explain consistency of the data sources and derived matrices with count evidence. This should explain how evidence from the count data has informed the matrix development. For some data sources, such as the use of RSI data, screenlines used to expand the source survey data should be distinguished, from other screenlines.

8.2.3 The flow difference measures specified in section 3.3 should be used. If the criteria given in Table 1 are not met for all or nearly all screenlines and cordons, remedial action should be considered.

8.2.4 All screenlines and cordons used for this and similar purposes should be 'watertight'. They should include all the roads in the actual network that intersect them.

8.2.5 While it is possible to validate trip matrices using cordons, care needs to be exercised to ensure that discrepancies between modelled flows and counts are not the result of traffic erroneously routeing to avoid crossing the cordons. Long screenlines, on the other hand, will show the quality of the matrix more clearly and should be used in preference to cordons where possible.

8.3 Refinement of Prior Trip Matrices By Matrix Estimation

The Purpose of Matrix Estimation

8.3.1 The primary purpose of matrix estimation is to **refine** prior matrices and the refinements should be sufficiently small that they are not regarded as significant. Matrix estimation can be a useful tool to initially establish any inconsistency in the pattern of demand between the matrices and the count data that should be investigated.

8.3.2 Matrix estimation only either increases or decreases non-zero cell values in the prior trip matrix. The technique cannot be used, therefore, to provide estimates of trips not observed in surveys or not contained in the synthesised trips.

8.3.3 Matrix estimation should not be used to factor matrices from one year to another or to factor period or average hour matrices to actual hour matrices or vice versa. While matrix estimation could be used mechanically for this purpose, experience has shown that the resulting changes are often significant - indeed, the longer the period over which matrices are being grown, the larger the changes are likely to be.

8.3.4 Matrix estimation should not be used solely to achieve the flow validation criteria at the expense of significantly changing the demand matrix patterns (refer to paragraph 8.3.17).

Applying Matrix Estimation

8.3.5 It is important that the effects of matrix estimation are minimised. If the prior trip matrices have been developed from matrices in existing models, and the existing model matrices were not developed in accord with best practice, do not meet acceptable validation standards or were not adequately documented, the matrices taken from the existing models should exclude the effects of any matrix estimation carried out during the development of the existing models. Applying matrix estimation to a matrix which has been subjected previously to matrix estimation may not achieve the aim of minimising the effects of matrix estimation. Thus, the trip matrices taken from existing models should be 'prior' trip matrices.

8.3.6 Count constraints should generally be grouped and applied at the short screenline level. The use of counts at individual sites as constraints should be avoided. The reason for this advice is that the mismatch between modelled flows and counts at any one location may be due to a number of reasons and not solely due to deficiencies in the trip matrices. In adjusting the prior matrices, matrix estimation may well compensate (undesirably) for other errors arising from the design of the zoning system, network structure, centroid connectors, network coding and route choice coefficients, which is why all these aspects should be checked before applying matrix estimation. Applying constraints at individual sites is likely to exacerbate the tendency of the matrix estimation procedure to compensate for deficiencies in other aspects of the model.

8.3.7 As explained in section 4.2, the counts used as constraints in matrix estimation should be derived from two-week ATCs at a minimum, with the vehicle type proportions being obtained from MCCs. Turning movement counts should only be used as constraints in matrix estimation if they have been derived from MCCs **and** ATCs.

8.3.8 As also explained in section 4.2, unless unusually accurate counts of LGVs and HGVs have been obtained, it will not be appropriate to apply constraints on the matrices of these vehicle types on individual links, even when there is a case, albeit rare, for applying constraints on the car trip matrices at individual sites.

8.3.9 Constraints should not be applied to individual user classes unless counts of these user classes have been made individually. Thus, in cases where MCCs are not available, constraints can only be applied to total trips, that is, all user classes combined, which is not advisable because routes vary between user classes. In practice, therefore, total counts across all purposes will have to be used.

8.3.10 It will be desirable that the highway assignment matrices are as consistent as possible with the trip matrices used in a related demand model. In addition to applying count constraints, therefore, trip end constraints should also be applied whenever acceptable trip end estimates can be obtained. The trip end constraints should be applied in the same way as the count constraints are applied, in preference to controlling the estimated matrices to exogenously derived trip ends in a Furness procedure at the end of the matrix estimation process.

8.3.11 Some software packages allow weights to be attached to the inputs to matrix estimation which reflect their relative accuracy. In principle, this approach should be adopted. Thus, weights should be attached to prior matrix cells, trip ends, and counts to reflect their relative accuracy. Any weighting calculations and associated assumptions should be reported.

8.3.12 Ultimately, the assignment convergence standards used during matrix estimation should be consistent with those used in the assignment model. However, there is merit in considering a staged approach as initial matrix estimation runs may be based on initial demand estimates and the network may still require further calibration. Hence, adopting tight convergence criteria may be inefficient and a more relaxed value may be used, although not exceeding the guideline %GAP of 0.1%.

8.3.13 The convergence of the iteration between assignment and matrix estimation should be continued until the changes in the matrices, at zone to zone level, between iterations becomes negligible.

Monitoring the Changes Brought About By Matrix Estimation

8.3.14 The changes brought about by matrix estimation should be carefully monitored by the following means:

- scatter plots of matrix zonal cell values, prior to and post matrix estimation, with regression statistics (slopes, intercepts and R^2 values)
- scatter plots of zonal trip ends, prior to and post matrix estimation, with regression statistics (slopes, intercepts and R^2 values)
- trip length distributions, prior to and post matrix estimation, with means and standard deviations
- sector to sector level matrices, prior to and post matrix estimation, with absolute and percentage changes

8.3.15 The changes brought about by matrix estimation should **not** be significant. The criteria by which the significance of the changes brought about by matrix estimation may be judged are given in Table 5.

Table 5 Significance of Matrix Estimation Changes

Measure	Significance Criteria
Matrix zonal cell values	Slope within 0.98 and 1.02 Intercept near zero R^2 in excess of 0.95
Matrix zonal trip ends	Slope within 0.99 and 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%

8.3.16 All exceedances of these criteria should be examined and assessed for their importance for the accuracy of the matrices in the Fully Modelled Area or the area of influence of the scheme to be assessed. Where the exceedances are important and statistically significant, the development of the prior matrix should be reconsidered. Where they are not considered to be important, the reasons should be documented in the Local Model Validation Report.

8.3.17 As outlined in paragraph 3.3.3, **matrix estimation should not be allowed to make significant changes to the prior matrices in order that the validation standards are met. In these cases, the limits set out in Table 5 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**. If issues in the process of creating the prior matrices are identified then these should be rectified before running through the model calibration and validation process again.

Validation of the Post Matrix Estimation Trip Matrices

8.3.18 The trip matrices output from the matrix estimation should be validated by comparing total screenline and cordon modelled flows and counts by vehicle type and time period. These comparisons should be made and presented separately for each of the following types of screenline and cordon:

- the screenlines and cordons used in applying count constraints in the matrix estimation
- the screenlines earmarked for independent validation

8.3.19 Where screenlines are also used to expand source data for the trip matrices, comparisons should also be provided separately at these locations.

8.3.20 The flow difference measures specified in section 3.3 should be used. If the criteria given in Table 1 are not met for all or nearly all screenlines and cordons, remedial action should be considered. This may include re-examining the data sources used for matrix development to identify potential issues in their processing or using insights from a new data source not used previously in the matrix development process.

8.3.21 Where the source data provides high confidence in the pattern of movements (eg at inter-sector level), comparisons of the **post** matrix estimation matrices with source data used to develop the trip matrices may be useful. Where the comparisons of the changes brought about by matrix estimation listed in paragraph 8.3.14 show material changes arising in matrix estimation, or similar verification of the prior matrices indicates differences, post estimation matrices should be compared with source data and the process undertaken in developing the prior trip matrices, as outlined in TAG unit M2.2, should be reviewed and updated if necessary. The confidence that can be placed in the source data should be appropriately reflected in the process.

9. Assignment Calibration and Validation

9.1 Introduction

9.1.1 Advice has been provided on calibration of:

- networks in section 6
- routes in section 7
- trip matrices in section 8

9.1.2 Section 9.2 provides advice on some possible further means of calibrating an assignment.

9.1.3 The final step in setting up a model is a formal validation of the assignment. The requirements are specified in section 9.3.

9.2 Assignment Calibration

9.2.1 If the steps set out in sections 6, 7 and 8 to calibrate the network, routes and trip matrices do not yield an acceptable validation of link flows, turning movements and journey times, the following further steps to calibrate the assignment model may be considered:

- the number of zone centroid connectors, their coded times and the points at which they connect to the network should be reconsidered and adjusted if necessary
- 'forests' should be analysed to understand routeing and adjustments may be made to competing routes
- modelled and surveyed journey times should be compared and analysed in order to:
  - identify queue locations
  - check outturn capacities on congested (queued approaches) and hence adjust signal timings, saturation flows, lane use, etc., accordingly

9.2.2 It may also be useful to 'stress test' the model by increasing the numbers of trips in the matrices by 10% or 20% and reassigning. This may reveal faults in the network which previous checks have not detected. For instance, against expectations, some junctions may become over-loaded while others show no queues despite the increased demands. This is particularly important within microsimulation models to understand the scale of change and whether the model network is of sufficient scale to be appropriate for use.

9.3 Assignment Validation

9.3.1 In addition to evidence of network, route and trip matrix validation, the Local Model Validation Report should include evidence of the validation of the assignment, in the following primary terms.

- **Traffic flows on links**. Modelled flows and counts should be compared by vehicle type and time period for screenlines. For cars, flows on individual links should be compared. For goods vehicles, flows on short screenlines should be compared unless very accurate counts have been obtained. The measures, criteria and guidelines given in Table 2 should be used.
- **Journey times**. Modelled and surveyed journey times should be compared along routes, by vehicle type if separate speed/flow relationships have been used for light and heavy vehicles, and by time period. End to end route times

should be analysed, with the means and 95% confidence intervals of observed times being presented alongside the modelled times. In addition, time/distance graphs should be produced for individual section on each route. For microsimulation models, plots of the profile of travel times throughout the modelled time period should be compared to observed for each route. The measures, criteria and guidelines given in Table 3 should be used.

9.3.2 **Turning movements** at key junctions should also be validated by time period. However, it is rare that turning movements will have been counted using automatic methods over a number of days; most likely, the available or affordable counts will be single day MCCs. For this reason alone, turning movements may not validate to the standards achieved for link flows. Given the 95% confidence intervals usually associated with LGV and HGV counts, it is unlikely that it will be sensible to validate turning movements by vehicle type. Nevertheless, modelled turning flows and counts should be compared by time period and assessed using the link flow criteria and guidelines given in Table 2. When microsimulation models are used to assess traffic performance (particularly at junctions), the turning counts are an essential part of model performance. The importance to model calibration and validation is therefore much greater and a balance will often need to be considered against the reliability of the data. Even where only single day counts are available presentation of these against the model should be included.

9.3.3 In addition to the primary tests above, the following further checks may also be valuable:

- If available, comparisons of the trip patterns assigned to key links with the patterns derived from source data for those links - note that this analysis should be conducted at a sector rather than zonal level, as a close match at the latter level is highly unlikely due to source data statistical accuracy

- comparison of average modelled network speeds with average speeds derived from the more comprehensive journey time data sources, such as tracked vehicle data, by area. For microsimulation, comparison of observed and modelled speed distributions should be considered in key locations to verify the driver and vehicle behaviour parameters within the model.

9.3.4 All significant discrepancies between modelled and observed data should be noted and a commentary provided in each instance. These commentaries should state whether or not the discrepancies might affect the model's usefulness for certain applications.

9.3.5 All text summarising validation results should be carefully considered to ensure that the correct impression about the quality of the model is conveyed.

10. Reporting

10.1.1 The following two reports are required which relate to the advice in this unit:

- Highway Assignment Model Specification Report (or as part of the Appraisal Specification Report)
- Highway Assignment Model Validation Report

10.1.2 The recommended structures of these reports are set out in Appendix G.

10.1.3 The **Appraisal Specification Report** should be prepared as the first task in the process developing a model. The report should include: proposed uses of the model and key model design considerations; model standards; key features of the model; **specification** of the required calibration and validation data; and the **methodologies** for network development, trip matrix development, and for calibrating and validating the network, route choices, trip matrices, and assignment.

10.1.4 The **Model Validation Report** will be the last task in the model development process. The report should include: updated sections on proposed uses of the model and key model design considerations, model standards (including convergence), and key features of the model; a description of the calibration and validation data used; descriptions of the network and trip matrix development; and descriptions of the calibration and validation of the network, route choices, trip matrices, and assignment.

11. References

Akçelik R (2003), Speed-Flow Models for Uninterrupted Traffic Facilities, http://www.sidrasolutions.com/documents/AA_SpeedFlowModels_UninterruptedFacilities.pdf

Bar Gera (2002), Origin-based Algorithm for the Traffic Assignment Problem, Transportation Science 36(4), pp398-417.

Department of the Environment (1971), Advice Note 1A.

Department for Transport (2002), Advice on Modelling of Congestion Charging and Tolling Options for Multi-Modal Studies.

Department for Transport (1995), Local Transport Note 2/95, The Design of Pedestrian Crossings.

Institution of Highways and Transportation (1996), Guidelines for Developing Urban Transport Strategies.

Hague Consulting Group (1994 (revised 1996)), The value of travel time on UK roads.

Parsons Brinckerhoff and The Denvil Coombe Practice (2009), Speed/Flow Relationships: Comparisons of Alternatives to Advice Note 1A, Technical Note A3.11.

Road Research Laboratory (1963), A method of measuring saturation flow at traffic signals, Department of Scientific and Industrial Research Road Note 34, HMSO.

Transport and Road Research Laboratory (1992), Contractor Report 279, Speed/Flow/Geometry Relationships for Rural Dual-Carriageways and Motorways

12. Document Provenance

12.1.1 This Transport Analysis Guidance (TAG) unit is based on the previous TAG unit M3.1 (May 2020) with the addition of a new appendix (Appendix C) on microsimulation modelling.

Appendix A: Glossary of Terms

All-or-nothing assignment: An assignment in which all the trips between each origin and destination are loaded onto a single path, usually the minimum generalised cost path.

Area of Detailed Modelling: The area over which significant impacts of interventions are certain. 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).

Calibration: Adjustments to the model intended to reduce the differences between the modelled and observed data.

Capacity restraint: The restraining effects on demand of capacity limitations or the process by which speeds, and therefore travel times are adjusted so that they are consistent with the assigned traffic flows. Capacity restraint mechanisms include junction **flow/delay** and link **speed/flow** relationships.

Centroid connectors: The means by which the demand from or to zones is loaded onto or leaves the network.

Convergence: An equilibrium or balanced position between two inter-related model outputs. A converged assignment is one where the assigned flows and the resulting travel costs are consistent. A converged demand/supply loop is one where the demands are consistent with the travel costs in the supply model.

Convergence criteria: The values of measures of convergence by which it is accepted that an acceptable level of convergence or equilibrium has been reached.

Cordon model: Assignment model variants where the modelled area is curtailed at a specified boundary. Trips with one or both end points outside of this area are not represented for their full length, being curtailed to one of the zones that represent the cordon boundary.

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.

Demand model: A model which forecasts changes in trip frequency, mode of travel, time of travel, and trip destination.

Deterministic User Equilibrium Assignment:

Assignment procedures designed to achieve **Wardrop's First Principle of Traffic Equilibrium**. Deterministic algorithms take no account of drivers' differing perceptions of costs. See also **User Equilibrium Assignment** and **Stochastic User Equilibrium Assignment.**

Distance coefficient: The coefficient of distance-related money costs which are combined with time and other money costs to form **generalised costs.**

Dynamic assignment model: Highway assignment models which permit the trip matrix to vary in terms of both level and pattern of flow during the modelled period.

Elastic assignment: An extension of a normal assignment procedure which uses an elasticity function to approximate some demand responses in addition to the change of route response modelled by an assignment.

External area: The area outside the **Fully Modelled Area**.

Flow/delay relationship: The relationship between traffic flow and travel time along a link sometimes including delays at junctions and also the relationship between traffic flow and delay for turning movements at junctions.

Frank-Wolfe Algorithm: A method of deriving a **User Equilibrium Assignment** by means of combining successive All-or-Nothing assignments such that an objective function is maximised, with the result that the contributions of the All-or-Nothing assignments to the final assigned flows are optimised (and not determined by the number of All-or-Nothing assignments as in the case of the **Method of Successive Averages**).

Free-flow speed: The speed at zero flow, as defined by a **speed/flow relationship.**

Fully Modelled Area: The area where trip matrices are complete (as opposed to partial in the **External Area**) and the network and zoning are at their most detailed (as opposed to coarser as in the External Area).

Generalised cost: A linear combination of time and money costs, expressed in time or monetary units. See also **distance coefficient**.

Heavy goods vehicles: Heavy goods vehicles (HGVs) are defined as other goods vehicles (OGV1 and OGV2).

Heavy vehicles: Heavy vehicles are defined as HGVs, buses and coaches (PSV).

Highway assignment model: A model which allocates car and goods vehicle trips to routes through a highway network. It includes path building and loading of trips to routes between zones. It excludes all demand responses other than route choice.

Light vehicles: Light vehicles are defined as cars and light goods vehicles (LGV).

Matrix estimation: The adjustment of **prior trip matrices** so that, when assigned, the resulting flows accord more closely with counts used as constraints in the process.

Method of Successive Averages:

A method of deriving a **User Equilibrium Assignment** by means of combining successive All-or-Nothing assignments such that, on the nth iteration, the combination factor used is fixed at 1/n, each route generated therefore contributing 1/n of total flows to the final assignment. See also **Frank-Wolfe Algorithm**.

Microsimulation model: A **dynamic assignment model** in which individual user-decisions are represented at a disaggregate level.

Operational model: A model in which detailed driver behaviour or vehicle operation is analysed to understand the network performance. These can typically cover a standalone junction or a network of junctions using deterministic or microsimulation models.

Prior trip matrix: The trip matrix that is input to **matrix estimation**.

Public transport assignment model:

A model which allocates public transport passenger trips to routes through a public transport network. It includes path building and loading of trips to routes between zones. It excludes all demand responses other than change of route and service.

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. It 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).

Route choice: The generation of alternative routes through a network on the basis of generalised cost or time.

Speed/flow relationship: The relationship between traffic speed and assigned flow on a link. Speed/flow relationships may either exclude or include delays at significant junctions. See also **free-flow speed**.

Steady State Assignment model:

An aggregate highway assignment model which represents average conditions over the modelled period, often one hour.

Stochastic User Equilibrium Assignment (SUE):

Assignment procedures designed to achieve **Wardrop's First Principle of Traffic Equilibrium**, taking account of drivers' differing perceptions of costs. See also **User Equilibrium Assignment and Deterministic User Equilibrium Assignment**.

Stopping criteria: Criteria to determine when an iterative procedure should be stopped. Not to be confused with **convergence criteria**.

Tiered model system: A model in which a simplified highway assignment model (upper tier) is created for the whole of the Fully Modelled and External Areas from a detailed highway assignment model (lower tier) for the same area, and in which demand/supply equilibrium is sought by iterating between the demand model and the upper tier assignment model, with the resultant demands being fed down to the lower tier assignment model.

Tracked Vehicle Data Data derived from Automatic Vehicle Location Systems usually through GPS systems, deriving detailed geographical locations of vehicles with this equipment at specific times.

Trip matrix synthesis: The creation of a matrix of trips by use of non-trip data, usually via a gravity model.

User class: Combinations of vehicle types and trip purposes which are assigned separately in a multi-user class assignment.

User Equilibrium Assignment: An assignment procedure in which a number of assignments are combined such that the resulting network flows satisfy **Wardrop's First Principle of Traffic Equilibrium**. See **Frank-Wolfe Algorithm** and **Method of Successive Averages**.

Validation: The comparison of modelled and observed data. Any adjustments to the model intended to reduce the differences between the modelled and observed data should be regarded as **calibration**.

Validation guidelines: The recommended proportion of instances where the **validation criteria** are met.

Validation criteria: The differences between modelled and observed data should be quantified (using some **measures**) and then assessed using some **criteria**.

Wardrop's First Principle of Traffic Equilibrium:

Traffic arranges itself on networks such that the cost of travel on all routes used between each OD pair is equal to the minimum cost of travel and all unused routes have equal or greater cost.

WebTRIS: National Highways' Traffic Information System.

Appendix B: Assignment Methods

B.1 Introduction

B.1.1 This appendix provides some details on **Steady State Assignment** models. Models of this nature fall into the following categories:

- **Deterministic User Equilibrium Assignment** models, of which the following types exist:
  - **Link-Based Assignment**
  - **Path-Based Assignment**
  - **Origin-Based Assignment** and
- **Stochastic User Equilibrium Assignment** models

These four assignment methods are described below. There then follows a note on an aspect of capacity restraint which has a bearing on the assignment method.

B.1.2 Some brief notes are then provided on **Dynamic Assignment** models and **Microsimulation** models.

B.2 Steady State Assignment Models

Link-Based Assignment

B.2.1 The most common forms of Link-Based Assignment are the so-called Frank-Wolfe Algorithm and the alternative Method of Successive Averages (MSA).

B.2.2 Equilibrium assignment conventionally employs the relatively efficient Frank-Wolfe Algorithm for combining simple All-or-Nothing (AON) assignments (where all trips from a particular origin-destination pair are assigned to a single 'best' route based on current costs) such that the resulting linear combination of network flows satisfies Wardrop's First Principle of Traffic Equilibrium. It works by recognising that the set of flows V_a satisfying Wardrop's First Principle could also be obtained by finding the set of flows which minimised a certain 'objective function', Z, given by:

$$Z = \sum_{a} \int_{O}^{V_a} C\,a(v)dv$$

where: the summation is over all network links, a

$C_a(v)$ is the cost of travel on link a with volume v

the upper integration limit is the volume of trips on each link V_a

B.2.3 This equivalence enables algorithms to be constructed which, by minimising Z, guarantee finding an equilibrium solution. In practice, the Frank-Wolfe algorithm is most commonly used to derive a solution which meets Wardrop's Principle. The algorithm works by iterating over a large number of simple AON assignments and combining these in an 'optimum' way, by seeking on each iteration a 'descent direction' in which the current solution is improved and moving an optimum distance, λ, in that direction which minimises Z.

B.2.4 In terms of link flows the optimum solution may be written:

$$V_a^{(n+1)} = (1-\lambda)V_a^{(n)} + \lambda F_a^{(n)}$$

where:

$V_a^{(n)}$ is the link flow on iteration n

$F_a^{(n)}$ is the all-or-nothing flow

B.2.5 The **Method of Successive Averages** is similar in its use of successive AON assignments, but the combination of these assignments is such that on the nth iteration, the (lambda) combination factor used is fixed at 1/n, each route generated therefore contributing 1/n of total flows to the final assignment. This approach is generally significantly less efficient than Equilibrium Assignment, requiring a far larger number of iterations, but has been shown to converge to identical solutions.

Path-Based Assignment

B.2.6 **Path-Based** Assignments have generally seen less use to date than Link-Based Assignments, but it is possible to base algorithms explicitly on origin-destination path flows. Several assignment suites offer such implementations, with solution processes often involving variants of the Frank-Wolfe algorithm.

B.2.7 There are indications that Path-Based algorithms can be considerably more efficient than Link-Based algorithms, both in terms of obtaining solutions and in terms of analysing the solution once obtained, but the pool of experience in their use is considerably smaller than for Link-Based methods.

B.2.8 Mathematically, in terms of path flows:

$$T_{pij} > 0 \Rightarrow C_{pij} = C^*_{ij}$$

$$T_{pij} = 0 \Rightarrow C_{pij} \geq C^*_{ij}$$

where:

T_{pij} is the flow on path p from i to j, C_{pij} is the path cost on pij

C^*_{ij} is the minimum path cost from i to j

Flows on paths are combined rather than flows on links, as in the case of Link-Based algorithms.

Aggregating path flows yields the equivalent link flows V_a, the flow on link a. This may be equivalenced to a constrained minimisation problem with an objective function 'identical' to that for the link based approach, subject to the assignment being 'feasible', where 'feasibility' is most easily defined in terms of path flows:

$$\sum_p T_{pij} = T_{ij}$$

for $T_{pij} \geq 0$

where:

T_{ij} is the total of trips from origin i to destination j

T_{pij} is the number of trips from i to j using path p

B.2.9 As for Link-Based Assignment, the Frank-Wolfe algorithm can be used to generate a 'descent direction' and an optimum distance, λ, which minimises Z in that direction. In terms of path flows:

$$T_{pij}^{(n+1)} \begin{cases} = (1-\lambda)T_{pij}^{n} & & pij \neq pij* \\ = (1-\lambda)T_{pij}^{n} & +\lambda T_{ij} & pij = pij* \end{cases}$$

where:

T_{pij}^{n} is the number of trips from i to j using path p on iteration n.

B.2.10 Some suites provide solution algorithms that are variants of the Frank-Wolfe algorithm, but which make more explicit use of path flows.

Origin-Based Assignment

B.2.11 **Origin-Based Assignment** (OBA) (Bar Gera 2002) is a relatively recent development that is now offered by a number of mainstream assignment suites. It is effectively intermediate between Link-Based and Path-based Assignment in that it stores the link flows as generated by each individual origin:

$$V_a^i = \sum_i T_{ij} P_{ija}$$

where:

P_{ija} is the proportion of the trips T_{ij} from origin i to destination j which use link a.

B.2.12 The main theoretical advantage of OBA is that it provides virtually exact solutions to the Wardrop Equilibrium without requiring either excessive memory or excessive processing time, eliminating 'non-optimal' paths (and associated 'noise') which are routine by-products of Frank-Wolfe based solutions. In practice, exact solutions may still be difficult to obtain, particularly for complex assignments where an iterative process with a junction simulation model is involved.

Stochastic Equilibrium Assignment Methods

B.2.13 An element of user-perception modelling can be introduced within the equilibrium frameworks described above through the use of **Stochastic User Equilibrium (SUE) Assignments**. SUE Assignment methods try to account for variability in travel costs (or drivers' perception of those costs) by assuming that the perceived cost of travel on each network link varies randomly, within predefined limits. These methods are generally only required where the network is not congested.

B.2.14 The most common method, a variant of 'Burrell multi-routeing', is a form of Monte Carlo simulation. Prior to each AON assignment within the equilibrium (effectively MSA) process, a new set of costs is generated for each link, or each origin to destination pair. Costs are chosen 'randomly' from a distribution whose mean is the current link (or OD) cost before the nth iteration AON paths and flows are generated. A normal distribution of 'perceived' costs with a user defined 'spread' is conventionally assumed. No optimum combination of flows can be calculated for stochastic methods, so the combination of AON iterations reverts to the MSA approach. As a consequence large numbers of iterations are required to achieve convergence.

B.2.15 With this type of (SUE) assignment, the spread in route choice due to the stochastic effects tends to reduce as congestion increases, so the potential advantages of increased realism may be small. In general, where congestion levels are significant, conventional equilibrium assignment should be sufficient. If stochastic assignment methods are used, the randomness factor used in the calculation of link costs should be clearly stated in the model reporting.

B.3 Capacity Restraint Mechanisms

B.3.1 Some form of capacity restraint mechanism is a prerequisite for any form of congested equilibrium assignment. A characteristic of this mechanism will be a relationship between the volume of trips and the cost of travel (on a network link or for a turning movement). In general, a monotonic relationship is required, so that the greater the flow volume, the greater the cost. This provides the mechanism for the transfer of trips from more costly to less costly routes.

B.3.2 At its simplest, this mechanism takes the form of link-based speed/flow curves, whereby speed reduces with increasing flow. Speed/flow curves specific to link categories can be pre-defined to form the basis for an 'equilibrium' solution. Where costs on a link are dependent upon flows on that link only, costs are said to be '**separable**' - a pre-condition for algorithms such as Frank-Wolfe to guarantee convergence.

B.3.3 The quantification and achievement of convergence is an important subject in this context and is dealt with in section 3.4 and Appendix D.

B.4 Dynamic Assignment Methods

B.4.1 The assignment methods described above usually assume steady state conditions, in which the demand (trip matrix) is assumed to remain constant (in terms of level and pattern) throughout the modelled period. A few assignment model packages - sometimes referred to as **Dynamic Assignment Models** - permit the trip matrix to vary in terms of both level and pattern of flow during the modelled period. These packages usually require the trip matrix to be specified in 'time slices'. Each time slice is assigned to the network separately, ensuring that network travel costs during each time slice are consistent with the level of demand being assigned. This ensures that routeing fully reflects the level of demand and the corresponding journey time for each time slice.

B.4.2 The potential advantages of dynamic models in reflecting time-varying conditions, including the build-up and decay of congestion, are often unrealised because of the extra complexity involved relative to steady state models and the more detailed (demand) data required. Model convergence to an acceptable level has also proved challenging, as has model calibration and validation.

B.5 Microsimulation Models

B.5.1 A **microsimulation** model is a dynamic model in which individual user-decisions are represented at a disaggregate level, and the combined effects of these decisions contribute to the overall system state. In the context of traffic assignment, decisions - including randomly generated 'micro' gap acceptance and car following behaviour - dictate traffic flows and delays whose effects may be translated into route choice.

B.5.2 Microsimulation will not always involve route choice, however, some microsimulation packages allow the use of routes extracted from 'equilibrium' type assignment models, e.g. for more detailed operational assessments.

B.5.3 If route choice (assignment) is required, the user will typically choose how often 'new' routes are calculated based on 'current' costs. At a more detailed level, routes may be decided for individual 'vehicles' and updated as progress is made through the network. The concepts of equilibrium and convergence are difficult under such conditions and stability becomes a more crucial concern for microsimulation based assignments, particularly for models of large areas.

Appendix C: Microsimulation Models

C.1 Introduction

C.1.1 Microsimulation models can be a crucial part of the scheme assessment and appraisal process. They offer enhanced operational analysis of highway networks, a more precise representation of network capacity and can provide compelling visualisations to aid with stakeholder engagement. Whilst their design, development, calibration, and validation share many similarities with highway assignment models, there are several key differences.

C.1.2 In contrast to macro models, which use empirically derived relationships, microsimulation models explicitly depict the interactions between users (typically individual vehicles , but also pedestrians and cyclists), thereby providing a more accurate reflection of specific local highway network conditions. They typically cover a road junction or a network of neighbouring junctions and are particularly valuable where the interactions are complex and interdependent. This can extend into public transport interchanges.

C.1.3 Microsimulation models are useful in scenarios where: the interactions between different vehicle types are complex (e.g. the merging and weaving of traffic or the effect of signal control in multiple junctions); the interventions being designed include prioritisation measures; more complex technology such as dynamic signal operation or optimised signal control is being deployed; and there are interactions with pedestrians and cyclists, such as at crossing points. Microsimulation models can help to better understand the impact on traffic flows, queue lengths, delays, and journey times. These models can also give variability of these outputs (through multiple model runs), which provides information on journey time reliability.

C.1.4 In practice, microsimulation models are rarely used for scheme appraisal on their own. For schemes involving complex junctions and interactions with pedestrians and cyclists, a microsimulation model may be used in conjunction with a wider strategic model in a tiered model. Whereas the microsimulation model provides detail on junction operations and delays, the strategic model can provide input traffic flows to the microsimulation model, and is better suited to estimating wider re-assignment effects, and providing the generalised costs to variable demand modelling (see TAG unit M2.1 Variable Demand Modelling).

C.1.5 Practitioners should bear in mind that when developing a tiered model, information can flow both ways: the strategic model may provide initial flow inputs and subsequent re-assignment and variable demand effects, yet the microsimulation model can provide more detailed junction delay information that may be used by the strategic model and affects re-assignment of traffic around problematic junctions or bottlenecks. Practitioners should ensure that cost changes in the higher tier model are consistent with those in the lower tier model (see Appendix F).

C.1.6 In the economic appraisal, the use of the strategic model outputs rather than those of the microsimulation model is likely to create a more robust case, as the effects of the scheme can be thoroughly captured (i.e. the variable demand impact and overall changes in network conditions over a wider area).

C.1.7 Where models are developed for scheme appraisal purposes, a recommended way to do this is through the appraisal specification process at the beginning of the project, described in The Transport Appraisal Process. The production of an Appraisal Specification Report is a useful record of the agreements on the scope of the appraisal and model, including those reached with relevant stakeholders, in particular, between scheme promoters and sponsors.

C.1.8 When using microsimulation models for the appraisal of transport schemes, practitioners should ensure that:

- The survey specification is developed to ensure sufficient detailed data availability for all required time periods, including an allowance for relevant future year conditions (e.g., peak spreading).

- The design and specification of all elements of the modelling methodology reflect and are proportionate to the overall purpose of the assessment.

- The model extent is designed to consider any induced traffic effects that may influence the appraisal, including consideration of all future years and associated predicted traffic growth, as necessary.

- Adequate consideration is given at project inception to assignment methodology. This decision should not be underestimated and should be made impartially, on a project-by-project basis, involving consultation with experienced professional modellers where necessary.

- If microsimulation is chosen as the preferred platform for scheme appraisal after consideration, the method for estimating benefits should be robust. This may dictate the modelling of more time periods, a larger modelled extent, or further integration and interworking with a higher tier model than would have been undertaken for purely design related applications, to capture induced traffic effects arising from strategic re-routeing, mode and destination choice.

C.1.9 This appendix provides detail on the application of the highway assignment modelling guidance to the development of microsimulation models. It is not intended as a detailed step-by-step guidance note on how to develop or use microsimulation models. Practitioners can consult a number of existing guidance sources commonly used in the development of microsimulation models, including:

- Transport for London modelling guidance

- U.S. Department of Transportation guidance

- Software vendor advice notes

C.1.10 This appendix follows the main structure of TAG Unit M3.1 and should be considered in tandem by practitioners.

C.2 Design

C.2.1 Microsimulation models typically differ from strategic highway assignment models in their spatial extent. They usually focus on a smaller network encompassing the most relevant selection of junctions and streets. In these instances, demand is either directly derived from traffic count surveys (see TAG unit M1.2 Data Sources and Surveys) of the relevant network or derived by combining data from a higher tier model to reflect current demand, future growth and/or behavioural responses such as strategic re-routeing.

C.2.2 Microsimulation models can extend to cover larger areas and examples do exist in England of models covering towns and cities. In these instances, the guidance outlined in the main body of this unit becomes more relevant.

C.2.3 Smaller scale microsimulation models can be viewed as a cordon model, as the external area is typically not explicitly modelled. If strategic re-routeing may result from a particular intervention due to reduced or improved capacity, consideration should be given as to whether the full effects can be represented robustly in a microsimulation model. In such cases, the use of higher tier models to support appraisal is recommended.

C.2.4 The design of zoning systems in microsimulation models should consider that of any linked higher tier model. Similarly, if a strategic highway assignment model is intended to interact with a microsimulation model, corresponding consideration should be given. For instance, this may involve considering specific accesses points at a junction or parking location. It may not be practical to create a one-to-one mapping of higher tier model zoning to microsimulation zoning. It is recommended that the process for generating higher-tier demand is deliberated and designed as early as possible, as adjusting zoning at a later stage is often more time consuming. A nested approach is recommended where lower tier zoning is entirely within that of the higher tier model. The process should consider the mapping of demand by zone, by time period, and by vehicle class.

C.2.5 In most cases, the zoning of the microsimulation model may be much more detailed than the associated higher tier model. In all cases, the practitioner must carefully consider the consistency of applying traffic forecasts, and should consider the use of local planning data in the distribution of traffic growth forecast from the higher tier to the microsimulation model. Further detail of how the demand assumptions can be calibrated are outlined under section C.8.

C.2.6 The level of detail of the network structure in microsimulation models is typically consistent throughout and advice in section 5 on network data and coding of the ‘Area of Detailed Modelling’ applies. The intent is to ensure that junction operation and pedestrian or cyclist interaction (where modelled) is appropriate, and a high level of detail in both network and demand representation is essential to achieve this. Pedestrian interaction is often included where traffic behaviour is influenced by pedestrian activity, for example, at uncontrolled or demand-dependent pedestrian crossings. Similarly, cyclists tend to be included where cyclist volumes are high, or are otherwise of particular influence or interest. Practitioners should review the network and include pedestrian and/or cyclist interaction where it influences traffic performance and the scheme being assessed.

C.2.7 Some software allows the inclusion of external network coding to provide access routes into a detailed area and approximate strategic effects on re-routeing. The design of this network should align with the advice in the main body of this unit on network structure (see section 2.4).

C.2.8 The capacity restraint mechanisms in strategic highway assignment models typically differ from those used in microsimulation models. Some software packages offer hybrid approaches where sub-areas of strategic models are micro-simulated. Although these are not currently widely adopted, they are becoming increasingly more so. Advice on highway assignment and microsimulation applies to the respective areas.

C.2.9 In microsimulation models, the physical environment dictating the road capacity is typically modelled explicitly, reflecting cross-sections, design speeds, turning radii, junction control methods, and give way priorities. The resultant traffic behaviour shown by the software when appropriate traffic flows are incorporated, should reflect observed conditions and can be demonstrated through the standard calibration and validation criteria, e.g., flows and journey times, (see section 3.3).

C.2.10 In microsimulation models, the assignment methods used can either be static or dynamic, or can be set up using a combination of both methods. The choice of which assignment methodology is most appropriate is largely dependent upon the intended use of the model being developed and should therefore be carefully considered. In the static assignment, practitioners manually define the routes in the road network (typically taken from a higher tier model), whereby the modelled vehicles follow these routes. The traffic routeing in dynamic assignment takes account of fluctuations in traffic demand and road characteristics (e.g., different signal programs) throughout different times in the model. There is further discussion about routeing choice considerations in section C.7.

C.2.11 There are multiple factors to consider when deciding the assignment methodology when specifying the modelling approach, which include:

- Model size, complexity and purpose, including the level of route choice in real-life.
- Whether the microsimulation model is linked to a higher tier model, which would enable the representation of wider re-routeing (and whilst this doesn’t inherently require either static or dynamic assignment in microsimulation, it is an important consideration).

- The needs of the future year(s) to be assessed. Even without route choice, it can be easier to work with origin-destination matrices in dynamic assignment if the assignment is effectively static but the demand matrices that are received from higher tier models change per scenario.

- Time and budget considerations – either of the assignment (static and dynamic) strategies can add significantly to project time and financial costs, if they were poorly chosen initially, requiring change part-way through the project.

C.2.12 Regardless of assignment method, all microsimulation models include a random stochastic effect that aims to reflect real-world day-to-day variations through the use of probability distributions which depend on multiple runs using random seeds. For example, the profiling of vehicle arrivals into the network, or the exact per vehicle assignment of their desired speed from a given profile, within a sub-time period. This randomised component results in reported variability between model runs when repeatedly tested.

C.2.13 It is important to note that different software handle the concept of convergence differently, and different assignment techniques allow for less or more detailed assessment of stability and proximity, as explained in Appendix D. The practitioner should demonstrate that the resulting model calibration and validation, as well as the model's suitability for application are appropriate. This is discussed further in section C.3.

C.2.14 Unlike most strategic highway assignment model applications, microsimulation models necessitate shoulder time periods on either side of the modelled peak period. The warm-up period prior to the peak period allows for the progressive loading of traffic into the network, ensuring that the entire model accurately represents network conditions at the start of the evaluated peak period. The cool-down period (just after the peak period) allows a suitable period for traffic to complete its journey or for queues to dissipate in the base year. In future year models, these shoulder time periods can be crucial for understanding 'peak-spreading' conditions. As a result, demand should be accurately derived from on-site surveys for all modelled periods, including the warm-up and cool-down. Substantial changes in the demand and operation of the warm-up and cool-down periods might indicate the need to consider peak spreading and time period choice mechanisms in the higher tier model.

C.2.15 Microsimulation models are particularly useful for assessing the performance of a linked system of junctions. Where blocking back or platooning is a key issue to be influenced by an intervention, and the interactions with adjacent junctions are important, microsimulation models can be used to feed higher tier models with improved estimates of network characteristics, such as network capacity or speed/flow relationship. This is particularly important where lane choice at upstream junctions is heavily dependent on the desired direction at downstream junctions, especially with closely spaced junctions, roundabouts or gyratory systems. This feedback can in turn influence the overall demand in the microsimulation model estimated from higher tier models. Practitioners must check the validity of any feedback to ensure that it is giving reasonable and consistent results.

C.2.16 Driver and vehicle behaviour parameters are critical in understanding the specific capacity available from the physical environment coded. These parameters will define, for instance, how aggressively drivers are likely to behave in different situations, such as the size of gaps needed to exit a priority-controlled arm or the scale of lane changing behaviour. The adjustment of these parameters is typically undertaken from a set of starting values and is discussed further under model calibration in section C.9.

C.2.17 Some authorities may have their own recommendations for software set up relevant to local experience. It is recommended that when designing microsimulation models, early engagement with the relevant highway authorities is undertaken to determine any relevant local experience or requirements influencing model design and particularly behavioural parameters. The model parameters chosen should be evidence-based and documented in a Model Validation Report (see Appendix G).

C.3 Calibration, Validation and Convergence Standards

C.3.1 The standards set out in section 3 in the main body of this unit are mainly for strategic highway models. They also have some relevance to microsimulation models. Typically, though, it is expected that further comprehensive assessment is undertaken by the practitioner given the detail present in these models.

C.3.2 Model calibration and validation should be carried out as two separate exercises, involving comparisons against two independent datasets. Although not prescriptive, a sensible approach is as follows:

- **Calibration** – adjustments to the model intended to reduce differences between the model outputs and the input data used in the model development stage (e.g., entry flows, junction counts, signal timings)
- **Validation** – comparing the model outputs (e.g., journey times, exit flows, junction counts, and saturation flows) against an independent dataset (i.e. data not used during the model development stage or as model inputs)

Calibration

C.3.3 It is crucial to ensure that microsimulation models contain a representative traffic volume for detailed assessment. Therefore, all modelled entry flows should be demonstrated to be within 5% of observed flows.

C.3.4 During model calibration, it is essential to compare turning flows, given the significance of accurate junction performance to overall model performance. The modelled and observed turning flows should also be compared as part of model validation (see paragraph C.3.8).

C.3.5 When modelling variable or demand-dependent control, it is important that it is shown to be representative. As adaptive systems like MOVA and SCOOT (see paragraphs 5.2.8 and 5.2.9) can provide detailed logs of actual timings and stage calls, it is possible to demonstrate that similar levels of green time and demand occurrence are being achieved in the microsimulation model.

C.3.6 Although analysis of modelled queue lengths should form part of the model calibration process, this cannot be considered as a strict and quantified criterion. As a visual aid, it is important to provide evidence that network performance is appropriate in terms of the locations of congestion, including the correct buildup and dispersion profiles of queues. This should be demonstrated through the use of comparative charts showing observed vs modelled queues at regular polling intervals (e.g., every 5 minutes) throughout the modelled period.

Validation

C.3.7 Network exit flows are a key metric to demonstrate a suitable volume of vehicles that are completing their journeys. Exit flows should be demonstrated to achieve a GEH value under 5 for 85% or more of such links.

C.3.8 For the validation of junction turning movements, both criteria in Table 2 of this unit should be applied to demonstrate a suitable level of flow validation. This recommends 85% or more movements to meet flow criteria 1 in the table, and 85% or more movements to have a GEH value of less than 5. Practitioners also tend to ensure that all movements achieve a GEH of less than 10. If these standards are not achieved, the practitioner should provide commentary and justification in the associated report.

C.3.9 Saturation flow validation may also be considered to demonstrate accurate throughput at signalised junctions. Practitioners may refer to Transport for London modelling guidance for further advice on saturation flow validation.

C.3.10 Journey times should be considered as the key metric in the demonstration of model validation. With microsimulation modelling, it is recommended to achieve validation within 15% and within 60 seconds of observed values, instead of either/or as detailed in Table 3 of this unit for strategic highway assignment models. For smaller models, the 60 second threshold can be too tolerant, whereas for larger models, the 15% tolerance can be too generous. Meeting both allows a high standard of validation to be achieved either way.

C.3.11 Journey time validation can be required for either all vehicles combined, or can be required for individual vehicle types, dependent on the exact nature of the scheme and its assessment. This nuance will need consideration at project inception, as it will inform the level of data collection required. For instance, where buses are included, and of importance to the scheme/stakeholders, these should be validated against journey times.

C.3.12 Equally, although cyclist behaviour can be difficult to be accurately recreated in microsimulation, if they have any sort of significant impact, are in high volumes, or are of key importance to stakeholders, realistic inclusion would be essential. It is often simpler to model cyclists using fixed demand assignment, based on turning proportions at junctions, as it is particularly difficult to collect reliable origin-destination data. If cycle journey times are important, then specific cycling journey time data (see section C.4) should be collected to validate the model.

C.3.13 Multiple model runs will be required to generate statistically significant calibration and validation output, in a similar manner to determining robust observed data for comparison. The specific number of model runs (see further discussion in paragraph C.9.1) should depend on the type of scheme, the extent of the modelled area, the volume of traffic, the level of available re-routeing, and the severity of congestion in the model. While the production of model statistics from multiple runs is relatively straight forward, presenting visual representations from multiple runs can be difficult and time consuming. Where such visual representations are used from a small sample of runs (or even only one) it is important that the selected run(s) are representative of the overall trends and are not the subject of “cherry picking”.

C.3.14 As a result, practitioners are advised to analyse the variability of the statistics generated from multiple runs (through varying random seeds), as measured during the calibration and validation processes, and report the corresponding confidence intervals. Practitioners should demonstrate that sufficient runs have been conducted to establish a confidence interval where the interventions being assessed can be meaningfully differentiated. Further advice relating to this can be found in the U.S. Department of Transportation guidance.

Convergence

C.3.15 The principles of model convergence are described in Appendix D, and outlined here briefly:

- They relate to stability in model results, as well as proximity to the assignment objective
- In highway assignment modelling the objective is often Wardrop equilibrium, in which the costs of all routes used between an OD pair are equal, and less than that of any unused route
- Convergence considerations come into play when an iterative solution mechanism is used, which enables measuring whether consecutive iterations get closer to the assignment objectives
- Stability indicators apply to both static and dynamic assignment, but proximity to an equilibrium is only meaningful in dynamic assignment applications

C.3.16 In microsimulation, stability is more easily measured than proximity, and can be monitored for both static and dynamic assignment methods, for example, global stability indicators such as network-wide total journey times, total queue lengths or average speed. Disaggregate stability measures, consisting of absolute changes in values of individual link flows, costs or times, can be measured for both static and dynamic assignment, but are more informative for dynamic assignment as in static assignment route proportions are user-input.

C.3.17 Stability measures can be derived through post-processing of consecutive iterations in the application. If the software used does not automatically produce proximity measures during dynamic assignment, these can also be user-derived, by skimming route flows and generalised costs.

C.3.18 Convergence should be measured separately for the peak period, the warm-up and cool-down periods. The latter two should in general not cause problems. Convergence throughout the iterative solution process should be reported, as a table or in graphical format. In the absence of microsimulation-specific criteria, those developed for and used in traditional highway assignment should be used (see section D.2.8). Practical experience suggests that achieving the associated targets may not be practically feasible, and until updated guidance is produced, reporting of the values achieved with an explanation of their implications in the model results should suffice.

C.3.19 If stability and/or proximity indicators show that the solution process is still moving towards improved convergence, the process should be continued until these improvements flatten out. Oscillations in convergence values may be caused by the use of demand-dependent control, but they can also indicate coding issues that lead to alternative solutions being optimal in consecutive iterations, and a review of high delay junctions may be required.

C.3.20 Convergence of the iterative solution method to a stable solution should not be confused with the differences that occur between consecutive model runs with different random seeds.

C.4 Calibration and Validation Data

General

C.4.1 Microsimulation models generally require a much more comprehensive and detailed dataset than typical strategic highway assignment models, due to the increased spatial and behavioural model granularity.

C.4.2 Typically, traffic survey data should be collected via video surveys, with multiple cameras attached to street assets at a suitably high position to allow both data enumeration and a good understanding of site conditions and vehicle behaviour. Where relevant (i.e., for journey times), the use of tracking data (as described in TAG unit M1.2 Data Sources and Surveys) provides the best chance of high sample rates throughout all modelled periods. The aim should be to ensure that the dataset is as broad and robust as possible, although exact requirements should be dependent on, and proportionate to, project specifics.

C.4.3 The survey specification should comprise a suitable selection of the following, although this list is not intended to be exhaustive:

- Automatic Traffic Counts (ATC) – often 1–2-week surveys, providing classified counts and local speed profile data. This data is useful for peak identification, typical condition checks, speed profile creation and screenline or cordon counts.

- Junction Turning Counts (JTC) – usually (although not exclusively) collected for a single day, providing detailed classified counts in 15-minute intervals allowing accurate profiling across the hour for each vehicle class.

- Automatic Number Plate Recognition (ANPR) – usually (although not exclusively) collected for a single day, providing classified origin-destination data in 15-minute intervals allowing accurate profiling across the hour. ANPR can also be useful for journey time validation, although care must be taken if cameras are placed at junction stop-lines and clock settings are correct.

- Queue length data – although there are often inherent difficulties with fully collecting queue length data (e.g., differences between human/software queue definitions and difficulties understanding the extents of long queues), and quantitative comparisons, it is still important to evidence performance visually wherever possible. Queue lengths should be recorded at regular intervals (i.e., every 5 minutes) through all modelled periods. Where possible, video surveys of queues should be conducted to allow the modeller to visually compare the model with observed conditions and to better understand any survey records.

- Journey time data – can be collated from a commercial GPS tracking source, usually accessible through an online portal. This allows journey time data to be collected for the entire modelled area, for single or multiple days, which can then be used for both localised speed profile calibration and journey time validation. Journey time data may be collected using a moving car observer approach (which can also be used to collect journey time data for cyclists). See TAG unit M1.2 Data Sources and Surveys for further advice.

- Saturation flow data – essential for validating stop-line capacity at signalised junctions. This can be particularly valuable where Advanced Stop-lines (ASL) are present along with high cyclist volumes, as this creates non-standard conditions requiring a detailed assessment of on-site conditions.

- Signal control data – usually sourced directly from the local highway authority. This can include signal specification documents, timing sheets, demand dependent call frequency data, variable green time data etc.

- Public transport data – can either be sourced online, or directly from the relevant authority. As a minimum, this should include routes and timetables, but can also include detailed actual journey times and stop dwell time.

C.4.4 There are further model elements which do not have any associated formal calibration or validation criteria but can be essential to correctly calibrate and validate the operation of a microsimulation model. The level of inclusion of these elements is nuanced, proportionate to the size and exact goals of the assessment being carried out.

C.4.5 Examples of such elements are pedestrian behaviour at (and away from, in some cases) formal crossing points (e.g., green time compliance), behaviour at or near on-street parking (particularly if used frequently and for short-stays – i.e., outside shops – which therefore might warrant full parking modelling if feasible in the applied software), and behaviour of buses at and around bus stops.

C.4.6 The following paragraphs provide additional and more detailed advice for calibration and validation data to support the development of microsimulation models.

Queue length data for matrix development, calibration and validation

C.4.7 When developing matrices from count data, and where any level of significant queuing occurs on entries within the microsimulation model, then upstream demand data should be captured during traffic surveys, beyond the back of any queuing rather than using stop-line counts for input flows. This is likely to be part of the calibration process, although it is advised that practitioners should consider this at project inception where possible, so that traffic surveys can be specified accordingly. Queue lengths can be estimated using site visits, local knowledge, from videos, or historical data on typical conditions. This information helps plan surveys effectively to minimise discrepancies when applied in calibration and validation.

C.4.8 The specific definition of a queue is a key issue that the practitioner must consider both in terms of the collection of data and the extraction of data from the software being used. It is important that practitioners recognise how modelling software measures queues (or how the queue lengths are measured visually against visual measurements from observed queues) and ensure that appropriate data is collected and that comparisons are suitable.

C.4.9 Where long queues are observed it can be difficult to appropriately capture this in models, and sometimes surveys are limited in their extent. This can lead to challenges appropriately representing these in simple quantifications. These are the reasons that queue lengths form a visual calibration aid, rather than any sort of strict validation criteria.

Journey Times for Calibration and Validation

C.4.10 The advice in section 4.3 of this unit on the length of journey time routes in general is not directly applicable to microsimulation models. For example, route lengths are often shorter than the 3km recommendation as the network coverage tends to be reduced. Travel times can vary significantly between traffic lanes, so that it is important that the location of journey time routes are accurately reflected in the model against the observed data collected.

C.4.11 This is particularly relevant for journey time routes which navigate turns at junctions. Practitioners should ensure they fully understand how observed data sources account for lane choice, queues and delays and turning behaviour to ensure the comparisons with microsimulation model outputs are consistent. Practitioners should review the processing assumptions used in any floating vehicle data as some aggregate data products available for purchase are link based only, with no differentiation by lane, turn or route.

C.5 Network Data, Coding and Checking

C.5.1 Network coding in microsimulation models typically involves a detailed process of reflecting real-world conditions, including on-street road designs and junction layouts. They will typically include the number, width and intended usage of lanes, any vehicle restrictions (e.g. bans for high polluting vehicles), junction method of control, visibility and the presence of any specific lane markings impacting behaviour (e.g. yellow box markings).

C.5.2 Each software package typically includes specific tooling and data structures to configure the inputs for the network coding, and the practitioner should familiarise themselves with the specific assumptions used.

C.5.3 In microsimulation models, the link speed parameter should reflect the mean traffic speed under free-flow conditions where this is known to be different from the posted speed limits. A model's vehicle speed parameters (including maximum speeds and distributions) will determine the distribution of speeds around this link speed, including (the percentage and type of) vehicles travelling faster than the coded speed.

C.5.4 Where no local information is available it is recommended that practitioners draw on Vehicle Speed Compliance Data[14] produced by the Department for Transport to provide appropriate starting values. This is typically an area for adjustment through calibration and is further referenced below.

C.5.5 The ability to model different traffic signal control types is more advanced in microsimulation models than in strategic highway assignment models. It is usually possible for the specific configuration of controllers to be included so that vehicle actuation and timing optimisation behaviour is properly reflected. Although such a detailed approach is generally preferred, there are exceptions where modelling simplified fixed-time signals, based on averaged data, is proportionate and appropriate. As a general rule, practitioners should model traffic signals as accurately as possible, including elements such as demand dependency, bus priority, and green time optimisation.

C.6 Network Calibration and Validation

C.6.1 Each software will include specific processes for debugging, and practitioners should familiarise themselves with the relevant warnings and messages to determine which require remedial action.

C.6.2 It is often helpful to start the network calibration process in microsimulation models with a reduced level of demand, to check that the basic principles of the network coding work as intended. It is recommended that appropriate time is allocated to checking that the basic network structure and any specific behaviours (e.g. parking or routeing restrictions) are operating as intended prior to adding full demand. In practice, a demand reduction to 75% has been found to be a reasonable starting point. As more demand is added, the focus of

[14] Data available in table SPE0111 from https://www.gov.uk/government/statistical-data-sets/vehicle-speed-compliance-statistics-data-tables-spe typically provides a breakdown by road and vehicle type.

calibration then moves to calibrating each element to the required standard (where applicable), against the relevant dataset. This can draw considerable resource and time.

C.6.3 In order to implement the advice in section 6 of this unit for microsimulation models, the practitioner should spend time observing model operation at 'real-time' speed. This is often essential to ensure all areas are operating realistically and as intended. In addition, it is also relevant to 'follow' or 'trace' specific vehicle movements through the network to understand the exact routes being used by vehicles and to verify these appear plausible in free flow conditions.

C.7 Route Choice Calibration and Validation

C.7.1 When working with a dynamic assignment model, where the route choice varies through the model time period, the assumptions governing this should be reviewed. Each software package uses different configurations of parameters to control this behaviour and the practitioner should ensure they understand which parameters are reasonable to adjust. The sensitivity of software assumptions relating to how dynamic routeing behaviour is established and the basis for new routes chosen can significantly influence the extent of alternative routes that may be used and their plausibility. Evidence should be presented to demonstrate that the route choice dynamics are plausible. For example, this can be undertaken through comparisons to online journey planners.

C.7.2 Although having fewer degrees of freedom compared to dynamic assignment, static assignment still requires calibration and validation of route choice. As with dynamic assignment, origin-destination matrices, static assignment inputs, and routeing should be applied separately for each modelled vehicle class and should be assigned to sub-time periods (i.e., 15-minute intervals) to emulate the profiling across each hour.

C.7.3 The construction of microsimulation networks requires more attention to detail than in the case of highway assignment models - lane choice behaviour, lane restrictions and banned movements play a critical role in route choice and its calibration and validation. This applies to both static and dynamic assignment.

C.8 Trip Matrix Calibration and Validation

C.8.1 Although static assignment models do not use matrices, the data sources and process for generating trip matrices for dynamic assignment microsimulation models are similar to those used to generate end-to-end trip volumes for static assignment.

C.8.2 In either instance, prior data is required to inform initial proportions and can be derived from surveys (e.g., ANPR), calculation (e.g., directly from junction turning counts), or taken from a cordon of a higher tier model.

C.8.3 Where demand is derived from higher tier models, the practitioner will need to consider the quality of data taken from them and make suitable adjustments to reflect that. This may include time profiling or zonal disaggregation assumptions. Care should also be taken when considering flows extracted from upper tier models. If this model has limited representation of junction delays, without blocking back, the flows reaching the entry points of the microsimulation model may be overestimated, particularly in forecast years. Even in cases where junctions are represented in the upper tier model, a distinction between demand flows and actual flows needs to be made during any cordoning process.

C.8.4 If there is some overestimation of model entry flows, this may provide a case where observed queuing information results in the need to scale up or down specific zonal demand to ensure the scale of queuing is appropriate. However, this will usually be incorporated into the prior matrix build directly.

C.8.5 In the case of static assignment, volume data is entered directly into each entry point and route, per vehicle class, per time period. Any necessary adjustments in, for example, entry volumes and route proportions, can then be carried out as an iterative process by the practitioner. With dynamic assignment models, some microsimulation platforms are capable of matrix estimation/refinement directly, although this will generally still require prior matrices for successful use.

C.8.6 Where matrix estimation processes are adopted, the advice contained within the main unit section 8.3 should be followed. In all cases, where adjustments are made to input demand, these should be documented accordingly.

C.9 Assignment Calibration and Validation

C.9.1 All performance metrics should be considered against the average over a number of runs to ensure any stochastic effects are understood and can be included in the comparisons of modelled data with observed. Although not intended as an endorsement of a particular number of runs, a minimum of 10 should be considered as a starting point with each model. Practitioners must monitor the variability of model results across runs, which should be documented in a Model Validation Report (see Appendix G). Extreme or unexpected conditions may arise where specific seed values generate a significant outlier in results. Practitioners should review these instances to determine how they relate to specific coding issues, and these should then be addressed, allowing a subsequent full run of consecutive seed values. ‘Cherry-picking’ or selective use of model seed values is not an acceptable approach and is strongly discouraged.

C.9.2 Microsimulation software typically provides flexibility for the practitioner to modify a range of driver and vehicle behaviour parameters that can influence routeing and queuing behaviour. Any adjustments to the initial coded values should be clearly documented and evidence provided as to why they are acceptable. Where recommendations from software user guidance is adopted, this should be clearly documented in the Model Validation Report (see Appendix G).

C.10 Base Model Reporting

C.10.1 The general requirements for reporting of microsimulation models should follow the requirements set out for strategic highway assignment models in Appendix G. Any changes to a set of model parameters should be based on evidence and documented. Additionally, it is important to document the number of simulation runs and the random seeds used to generate reported results, allowing the reproduction of any results.

C.10.2 In addition to the Model Validation Report, it is also common practice to produce videos showing the operation of key areas of the network in the model. This can help to provide a visual comparison against actual video footage of the same location, or to provide reviewers an ability to compare with local knowledge of how the network performs.

C.11 Forecasting and Future Year Modelling

C.11.1 Whilst the purpose of the microsimulation modelling task can take a range of forms, there are still some common elements of guidance to inform forecasting and future year modelling.

Flow Forecasts

C.11.2 The derivation of flows used in forecasting needs to be appropriate for the study being undertaken. Whilst the initial scoping stage is likely to have informed this, care and consideration should be taken regarding how these flows are derived.

C.11.3 Where manual calculations are proposed, these should be suitable and calculated using industry recognised tools, with all relevant growth and/or local committed developments accounted for. See further details in TAG unit M4 Forecasting and Uncertainty.

C.11.4 Where the input flows are derived from higher tier model outputs, the translation of flows for the microsimulation model needs to be agreed and be suitable for the scheme (for example, deciding whether strategic model growth percentages, proportions or absolute values (or a combination) are to be used in the local model).

C.11.5 It is expected that any assumptions are reflected in the Forecasting Report.

Model Changes

C.11.6 The microsimulation model should also be updated to reflect the future year with scheme and without-scheme design layouts, with changes based on suitable background layouts and all elements accounted for. These include, but are not limited to, network structure, speed and priority modifications, signal control amendments (informed through collaboration with a signal engineer and should be applied to the with scheme and without-scheme forecasts), public transport route and stop line.

Reporting

C.11.7 The future year assumptions and proposed modelling approach should be documented in a report that details the following (as a minimum):

- the changes made to the validated and approved base year model
- a description of modelled scenarios and details of the flow-sets used (i.e. the flows used for each scenario)
- an assessment of the key outputs to provide an understanding of the impacts of the schemes being tested.

C.12 Model Quality and Fitness for Purpose

C.12.1 Choice of software also plays a part in what can, or cannot, be modelled. It is therefore important that modelling practitioners construct suitable models appropriately that are tailored to the situation at hand and to transparently present and communicate the potential risks around the accuracy or use of the model for its intended purpose. TAG unit M1.1 Principles of Modelling and Forecasting provides further discussion on assuring model quality.

C.12.2 To ensure that the modelling undertaken is suitable for its purpose, and to provide an opportunity for industry knowledge sharing and continuous improvement, it is encouraged that the modelling work is independently reviewed by experts with sufficient transport modelling expertise (particularly in microsimulation, as well as strategic modelling if tiered model is used) from a different team within the same organisation or ideally from a separate organisation. This process should be treated with pragmatism, with the aim to arrive at a model(s) that is the best fit for purpose for its intended use, as well as expanding the industry's knowledge and best practice techniques.

Appendix D: Model Convergence

D.1 Introduction

D.1.1 Before the results of any traffic assignment are used to influence decisions, the stability (and degree of convergence) of the assignment must be confirmed at the appropriate level. This appendix is concerned with the measurement of assignment stability and convergence, and the establishment of guidelines against which convergence may be assessed.

D.1.2 For all iterative assignment processes convergence is a significant issue. In some cases the iterative assignment procedure is mathematically guaranteed to converge, so that after a finite (albeit large) number of iterations a perfectly stable flow and delay pattern which meets the assignment objective will be reached. However, with most advanced assignment models involving explicit modelling of junctions or other options, convergence is not guaranteed.

D.1.3 For practical purposes convergence in assignment should be considered as a reasonable point in the iterative process, where the flows and costs are sufficiently stable (and therefore self-consistent) and within an acceptable proximity to the assignment objective. Convergence in practice is best measured in terms of two desirable properties of the flows and costs calculated by the program:

- stability of the model outcomes between consecutive iterations
- proximity to the assignment objective (eg Wardrop equilibrium)

D.1.4 The number of iterations required to achieve convergence in an assignment will generally depend on the network and matrix sizes, and the levels of congestion, as discussed in paragraph 3.4.12.

D.1.5 The specification of the number of iterations in direct terms may also not be straightforward for some more complex but widely used models. Practical testing will always be required to assess the impact on model convergence. Network size or computing requirements should not be a limiting factor in acceptability of convergence standards.

D.1.6 Convergence monitoring is an integral part of congested assignment modelling. It is of particular importance:

- at the initial stages of base year model calibration
- when moving to future year forecasts
- in assessing the accuracy of the final results

D.2 Use of Convergence Criteria

D.2.1 Convergence of congested assignment models can be monitored using a variety of indicators. These can be classified as follows:

- **global stability indicators**, based on comparisons between successive iterations of network-wide values of total journey time, total journey distance, total or average travel costs or average speed
- **disaggregate stability indicators**, based on absolute changes in values of individual link flows, costs or times, origin-destination costs or a combination of these
- **proximity indicators**, reflecting how close the current flow and cost pattern is to the assignment objective

D.2.2 **Stability at global level** (eg total travel time, costs or distance) **is necessary but not sufficient for ensuring model convergence**. Such measures may hide substantial uncertainty at a lower level, such as in individual link flows or OD-costs.

D.2.3 Of a large number of **disaggregate stability indicators**, the following three have been identified as being straightforward to compute, easy to interpret and explain, and robust in their explanation of assignment stability.

- Average Absolute Difference (AAD) in link flows between successive iterations:

$AAD = \frac{1}{N}\sum_{a=1}^{N} |V_a^n - V_a^{n-1}|$ where:

N is the number of links

V_a^n is the flow on link a in iteration n

- Relative Average Absolute Difference (RAAD) in link flows between

$AAD = \frac{1}{N}\sum_{a=1}^{N} |V_a^n - V_a^{n-1}|$ successive iterations:

- %FLOW, the proportion of links in the overall network with flows changing less than 1% from the previous iteration,

D.2.4 **Proximity measures** can only be calculated when an assignment objective has been formulated. This is usually the case with equilibrium assignment, and deterministic extensions (multiple user classes, dynamics, elastic assignment).

The most appropriate proximity indicator is the duality gap delta (δ). The duality gap expresses the flow-weighted difference between current total cost estimates on the network, as determined by the present flow pattern and the speed/flow curves, and the costs if all traffic would use minimum cost routes (as calculated by the next all-or-nothing assignment). The duality gap is a natural convergence indicator for equilibrium process, measuring how far the current flow pattern is removed from the desired equilibrium, and should approach 0 at that equilibrium. In link form, this is given by:

$$\delta = \frac{\sum_a C_a(V_a^n)(V_a^n - F_a^{n+1})}{\sum_a F_a^{n+1} C_a(V_a^n)}$$ where:

$C_a(V_a^n)$ is the cost for link a based on current flow estimate V_a^n

F_a^{n+1} is the all or nothing flow based on $C_a(V_a^n)$

The summation is made over all network links and implicitly all i and j pairs.

Note that an equivalent formulation can be made on a path basis:

$$\delta = \frac{\sum T_{pij}(c_{pij} - c_{ij}^*)}{\sum T_{ij} c_{ij}^*}$$ where:

T_{pij} is the flow on route p from origin i to destination j

T_{ij} is the total travel from i to j

c_{pij} is the (congested) cost of travel from i to j on p

c_{ij}^* is the minimum cost of travel from i to j

D.2.5 Of a large number of **disaggregate stability indicators**, the following three have been identified as being straightforward to compute, easy to interpret and explain, and robust in their explanation of assignment stability.

D.2.6 **'Gap' is the single most valuable indicator of overall model convergence**. It has a definite theoretical interpretation, does differentially weight 'good' and 'bad' fits and is easy to compare between networks of very different sizes, complexity and degrees of congestion. Regardless of which actual stopping criterion is chosen, Gap should always be monitored, its final values reported and convergence judged against the values obtained.

D.2.7 In complex assignment models, two separate 'Gaps' are sometimes reported, delta as above, and the '%GAP', which is a generalisation of the delta function to include the interaction effects within the simulation. It is, firstly, the difference between the current total vehicle costs on the assigned routes and the total vehicle costs if all drivers were to use minimum cost routes with the costs fixed. This measure is then normalised by dividing by the total vehicle costs and expressing it as a %. It is therefore the same as delta except that the costs are calculated after the simulation rather than after the assignment.

D.2.8 Although proximity and stability usually accompany each other, they both should be assessed separately, as each relates to different aspects of the iterative process. The following criteria have been found to lead to stable and robust assignment results, whilst in practice being achievable in most cases and with most assignment packages and for very large models. The following should be satisfied at convergence:

- Proximity measures:
  - Delta (δ) < 0.1%
  - %GAP (if relevant) < 0.1%

AND the following:

Stability measures:

- RAAD in flows < 0.1%
- %Links with Flows changing by less than 1% > 98% ("P1")
- %Links with Costs changing by less than 1% > 98% ("P2")

The RAAD is considered to be a more robust indicator of stability than AAD, hence its inclusion above.

D.2.9 Both stability and proximity criteria should be satisfied for four consecutive iterations before convergence can be judged to be acceptable. At least one of the stability criteria should be satisfied, the values of the other two measures should also be reported. If examination of the statistics in more detail shows that (rather than oscillating about a constant value) all these indicators still move in the same direction, it is necessary to continue the iterative process further, or to investigate further the reasons behind the unsatisfactory convergence.

D.3 Convergence Monitoring for Different Assignment Methods

D.3.1 Not all of the above criteria are applicable to all assignment options:

- in **user equilibrium assignment** both proximity and stability should be satisfied
- in **multiple user class assignment** stability should be monitored for each class separately and proximity should be assessed for total flow
- in **stochastic assignment** stability needs to be addressed within the iterative process and between different seed values for the stochastic process (see below)

- with **stochastic user equilibrium** (SUE) both proximity and stability must be assessed
- with **Origin Based Assignment** both proximity and stability should be satisfied
- **dynamic assignment** convergence monitoring for the whole period should suffice

D.3.2 Pure **stochastic** assignment is based on an iterative application of the randomisation of link times or costs, followed by all-or-nothing assignment steps. Randomisation of link costs results in more than single paths being found in the iterative assignment process, with a greater probability for objectively cheaper routes. The combination of consecutive assignments takes place using the Method of Successive Averages (MSA).

D.3.3 Stability of this process itself can be monitored using stability indicators as above. However, because of the use of MSA, enforced stability will occur after a large enough number of iterations. Therefore, the impact of the randomisation process on stability should also be checked to ensure that this apparent convergence is genuine.

D.3.4 In the initial iterations of a stochastic assignment the effect of different random number 'seed' values on the calculated flow pattern may be substantial, but in later iteration these differences should reduce to acceptable levels. The stability of stochastic assignment with respect to the seed-value should be monitored using the AAD and RAAD in flow estimates after an equal number of iterations with different seeds, with the minimum acceptable values being the same as for stability within the iterative sequence. This process should be carried out for three randomly chosen different seed values. Having generated three stable, similar assignments, the one of these with minimum total network cost should be used.

D.3.5 **Stochastic user equilibrium** (SUE) assignment combines a stochastic element and a capacity restraint element in the iterative procedure. Convergence with respect to both elements should be monitored. However, due to the stochastic element, proximity cannot be measured directly; therefore an equivalent deterministic user equilibrium assignment must be employed to monitor proximity, and comprehensive convergence monitoring should take place as follows:

- carry out a deterministic user equilibrium assignment using the method of successive averages; determine the minimum number of iterations required for proximity: $\delta < 0.1\%$
- carry out stochastic user equilibrium assignment with at least the required number of iterations for sufficient proximity, and monitor for convergence of the stochastic element as above, checking for stability within the iterative procedure and stability between different seed-values for the random number generator

D.4 Base Year Modelling and Future Year Forecasts

D.4.1 Not all assignment packages produce all the above statistics or allow the user to define stop criteria for the iterative process. In such cases users should allow the iterative assignment procedure to run for a fixed, large number of iterations during the initial stages of base year model calibration. Then check to see at which iteration the above requirements are met, and use this as a guide as to the number of iterations required during model development. Similarly, in forecasting it will generally be sufficient to determine the minimum number of required iterations for each scenario and each demand level once. Other runs can then be undertaken using perhaps 110% of the minimum number.

D.4.2 As convergence is greatly affected by the level of congestion in the network, it may lead to greater computational demands in forecast years. Thus in general longer run times and more iterations will be required to achieve a similar level of convergence in forecast years.

D.4.3 If convergence proves difficult, a spatially segregated assessment of convergence in different parts of the network should be carried out, by calculating the convergence statistics over subsets of the network. If this indicates that the problem is remote from the scheme, it may be possible to take results from the converged part only. If not, it is important to examine the coding of the part of the network where convergence problems arise.

D.5 Assessing the Accuracy of the Final Results

D.5.1 A key element of successful and robust scheme evaluation is the relationship between:

- the size of the model (in terms of total network times/costs)
- the time/cost savings of the scheme under consideration
- the uncertainty due to possible lack of convergence

D.5.2 If a large model is used to evaluate a scheme with relatively small network impacts, then convergence requirements need to be very tight. Otherwise the noise in poorly converged models can swamp the difference in total costs between without-scheme and with-scheme cases.

D.5.3 When using assignment models in scheme appraisal, the remaining uncertainty in model results may still be substantial, even after the model has achieved the desired level of convergence. This may arise where very large assignment models are used for relatively minor highways schemes, so that a small relative convergence error in the overall model may be quite large in comparison with the estimated scheme benefits. This can also happen when very high demand forecasts in future years lead to instabilities in the iterative sequence, particularly in the without-scheme scenario.

D.5.4 In some cases the remaining uncertainty in the model cannot be eradicated, as the model oscillates around the optimum flow pattern. It is necessary to assess this uncertainty in comparison with the scheme benefit estimates, to ensure that results are robust.

D.5.5 If the level of uncertainty is considered acceptable (in the context of scheme costs, etc) then the assignment may be taken to be robust. Out of the converged iterations for the without-scheme and with-scheme assignments those should be selected which have minimum total network travel time in each case.

D.6 Presentation of Convergence Results

D.6.1 Final results should always be accompanied by supporting documentation on convergence quality. Convergence monitoring of the assignment models used should form an explicit element of both the Local Model Validation Report and the presentation of forecasts. One suggested form of presentation is a 'convergence monitor' showing iteration number and the values of both proximity and stability indicators over the final four model iterations.

Appendix E: Speed/Flow Relationships

E.1 Introduction

E.1.1 The speed/flow relationships previously used in COBA were derived from national research and are considered to be the most appropriate for use in traffic models. These relationships predict the traffic speeds associated with a given traffic flow. They should be fitted locally by use of local parameters, and then validated by journey time runs.

E.1.2 Speed/flow relationships may be used in various kinds of highway assignment models, as follows.

- In models of urban areas, junction delays should normally be modelled explicitly. In these models, the use of speed/flow relationships should generally be restricted to motorways and dual carriageway links.
- In models of inter-urban areas, junction modelling may be limited to the key junctions and greater use may be made of speed/flow relationships, for single carriageway roads as well as motorways and dual carriageways.
- In some models, in urban, suburban and small town networks away from the main area of interest, junction modelling may be regarded as being unnecessarily detailed. In these cases, network speed/flow relationships for urban and route speed/flow relationships suburban areas and small towns may be used as an approximation.

This Appendix gives details of the speed/flow relationships for all categories of road.

E.1.3 The basic form of the speed/flow relationships varies between road classes. For rural, suburban and small town roads, the speed of vehicles reduces as flow increases until a critical flow level 'break point' is reached, at which the rate of speed reduction becomes greater until capacity is reached. The relationships for urban roads are simpler and they have a uniform speed/flow slope for all flow levels above their nominal capacity.

E.1.4 The standard road classes are as follows.

Table E.1 Road Classes

Road Class	Description	Section in Appendix E
1	Rural single carriageway	2
2	Rural all-purpose dual 2-lane carriageway	3
3	Rural all-purpose dual 3 or more lane carriageway	3
4	Motorway, dual 2-lanes	3
5	Motorway, dual 3-lanes	3
6	Motorway, dual 4 or more lanes	3
7	Urban, non-central	4
8	Urban, central	4
9	Small town	5
10	Suburban single carriageway	6
11	Suburban dual carriageway	6

E.1.5 Class 1 to 6 roads are all-purpose roads and motorways that are generally not subject to a local speed limit. Class 7 and 8 roads are roads in large towns or conurbations subject to 30 mile/h (48 km/h) speed limits only. Class 9 roads are roads in small towns or villages for routes subject to a 30 mile/h (48 km/h) or 40 mile/h (64 km/h) speed limit. Class 10 and 11 roads are major suburban routes in towns and cities that are generally subject to a 40 mile/h (64 km/h) speed limit.

E.1.6 The use of separate speed/flow relationships for light and heavy vehicles is to be preferred because this could provide more accurate estimates of changes in vehicle operating costs. Light vehicles are defined as cars and light goods vehicles (LGV); heavy vehicles are defined as other goods vehicles (OGV1 and OGV2), buses and coaches (PSV). However, there are a number of points to note, particularly that urban and small town class roads are for all vehicles and rural and suburban roads have separate speed/flow relationships for light and heavy vehicles, as described in paragraph 2.9.14.

E.1.7 The rural single carriageway, all-purpose dual carriageway and motorway relationships apply to individual links and exclude delays at major junctions. Major junctions on rural roads should therefore be modelled explicitly. The urban road relationships for both central and non-central areas are network relationships rather than individual link relationships. Therefore, if these relationships are applied to individual links, the approximate nature of this approach should be recognised. These urban relationships include delays at major and minor junctions. The small town relationships apply to routes and exclude delays at major junctions. Major junctions in small towns should therefore be modelled explicitly. The suburban road relationships apply to routes outside urban areas and include delays at both major and minor junctions. However, delays at congested junctions should be modelled separately.

E.1.8 Table E.2 summarises the characteristics and potential applications of the speed/flow relationships.

Table E.2 Summary of Speed/Flow Relationships

Area type	Separate light and heavy vehicle speeds	Link or route or network relationships?	Major junction delays included or excluded?
Rural	Yes	Link	Should be modelled separately
Urban	No	Network	Included
Small Town	No	Route	Should be modelled separately
Suburban	Yes	Route	Should be modelled separately

E.1.9 Highway assignment models in which speed/flow relationships are used in part or all of the modelled network without junction modelling with flow metering may assign flows which exceed capacity. Section E.8 specifies the speed/flow relationships which should be used to represent over-capacity conditions. The relationships specified in sections E.2 to E.6 apply up to capacity.

E.1.10 The speed/flow relationships are specified in terms of vehicles per hour. Traffic models generally require relationships specified in terms of PCUs per hour. The conversion from vehicles to PCUs is dealt with in section E.7.

E.1.11 The final section in this appendix offers advice on the modelling of high speed merges.

E.2 Rural Single Carriageway Roads (Road Class 1)

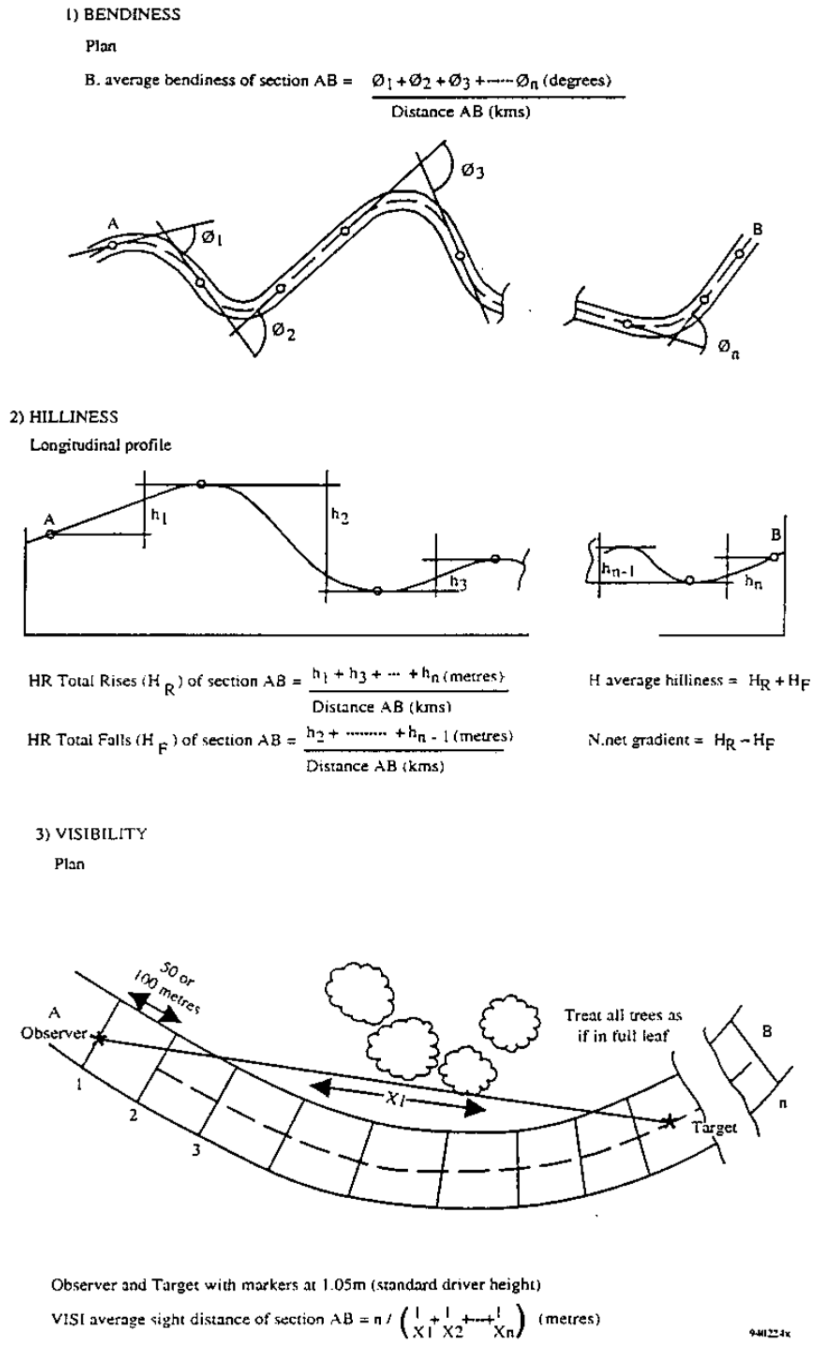
E.2.1 The rural single carriageway speed/flow relationships apply to single carriageways which do not lose priority and are not subject to a local speed limit. Table E.3 defines the geometric parameters and variables used in the relationships and gives the ranges of typical values over which the relationships should apply. The relationships cannot necessarily be taken to apply outside the given ranges of the variables.

E.2.2 Vehicle speeds for a given flow level are dependent on the geometric variables (CWID, BEND, HILL and HR). The value of those variables should be calculated, and the relationships set out below applied, for **at least** each individual road link longer than two kilometres, **on which flows change as a result of the scheme being appraised**. For other links, similar roads may be allocated to one of a number of typical link types (eg 6.7m bendy) each representing averaged characteristics. The use of a single road type for roads with markedly different characteristics (particularly free flow speeds) should be avoided.

E.2.3 The actual width of surfaced road is defined by two parameters. The first (CWID) being the width of carriageway between any continuous white lines which may or may not be delineating a hard strip. The second (SWID) is the total width of any continuous edge line and hard strip, which increases the effective carriageway width (as set out below) by at least 0.8 metres and thus increases free-flow speeds as well.

E.2.4 Figure E.1 explains how measurements of bendiness, hilliness and visibility should be taken. Hilliness is (H_R + H_F) and net gradient NG is (H_R - H_F). On two-way links net gradient is always zero because the two directions of flow are not disaggregated. On one-way links, rises and falls are defined with respect to the direction of traffic flow and, in general, they will not cancel out.

Figure E.1 Measurement of Road Geometry on Rural Roads



E.2.5 The average sight distance, VISI, is the harmonic mean of individual observations. For proposed new roads, VISI should be calculated from engineering drawings. For existing roads, an empirical relationship has been derived which provides estimates of VISI given bendiness and edge details:

Log VISI = 2.46 + (VWID + SWID)/25 - BEND/400

This relationship should normally be used for all existing roads for which bendiness and verge width have been measured. On long straight roads or where sight distance is available outside the highway boundary, VISI should be set to 700 metres for roads with high visibility; otherwise estimates should be made from plans or site measurements.

Table E.3 Definition of Variables Used in Speed/Flow Relationships for Rural Single Carriageways

Symbol	Variable Description	Typical Values	Max
DES	Is road designed to TD9/93 standards?	Yes	No
BEND	Bendiness; total change of direction (deg/km)	0	150
HILLS	Hilliness; total rise (H_R) and fall (H_F) (m/km)	0	45
NG	Net gradient one-way links only (m/km)	-45	45
JUNC	Side roads intersection, both direction (no/km)	0	5
CWID	Average carriageway width between white line edge markings, excluding any painted out portion (m)	6	11
SWID	Average width of hard strip on both sides, including width of white line (m)	0	1
VWID	Average verge width, both sides (m)	0	7
VISI	Average sight distance (m)	100	550
PHV	Percentage of heavy vehicles (OGV1 + OGV2 + PSV)	2	30
V_L, V_H	Speed of light and heavy vehicles (km/h)	n/a	n/a
S_L, S_H	Speed/flow slope of light and heavy vehicles (km/h reduction per 1000 increase in Q)	5	50
Q	Flow all vehicles (veh/hour/direction)	n/a	n/a
Q_B	Breakpoint: the value of Q at which the speed/flow slope of light vehicles changes (veh/hour/direction)	0.8 Qc	
Q_C	Capacity: defined as the maximum realistic value of Q (veh/hour/direction)	900	1600

E.2.6 The capacity of a single carriageway **per direction** is:

Q_C = 2400(CWID - 3.65)/CWID x (92 - PHV)/80 veh/hour

E.2.7 This value of Q_C identifies links which are likely to be overloaded. When flows reach this level the user must decide whether the flows are realistic and what course of action to take. The point of change of slope (Q_B) is given by the relationship:

$Q_B = 0.8\ Q_C$

E.2.8 For flow levels less than the breakpoint QB the speed prediction formulae for light vehicles in km/h is:

$V_L = 72.1$ - 0.09 x BEND or -0.015 x BEND for roads designed to TD9/93

- 0.0007 x BEND x HILLS

- 0.11 x NG (one-way links only)

- 1.9 x JUNC

+ 2.0 x CWID

+ SWID (1.6/SWID + 1.1)

+ 0.3 x VWID

+ 0.005 x VISI

- (0.015 + (0.00027 x PHV)) x Q

Note that, for two-way links, HILLS is the average of $H_R + H_F$; for one-way links HILLS is H_R alone.

E.2.9 For flow values greater than QB the speed prediction formula for light vehicles is:

$V_L = V_B - 0.05 \times (Q - Q_B)$

where V_B = speed at $Q = Q_B$.

E.2.10 For all flow levels the speed prediction formula for heavy vehicles in km/h is:

$V_H = 78.2$ - 0.1 x BEND or ZERO for roads designed to TD9/93

- 0.07 x HILLS

- 0.13 x NG (one-way links only)

- 1.1 x JUNC

+ 0.007 x VISI

+ 0.3 x VWID

- 0.0052 x Q

subject to the constraint that, if the calculated value of V_H is greater than V_L, then V_H is set equal to V_L, which is often the case in non-hilly areas.

E.3 Rural All-Purpose Dual Carriageways and Motorways (Road Classes 2 to 6)

E.3.1 The rural all-purpose dual carriageway and motorway speed/flow relationships apply to dual carriageways and motorways which do not lose priority and are not subject to a local speed limit. Table E.4 below defines the variables used in the relationships and gives the ranges of values over which the relationships should apply. The relationships cannot necessarily be taken to apply outside the given ranges of the variables.

Table E.4 Definition of variables used in Speed/Flow relationships for rural all-

Symbol	Variable Description	Typical Values		
		Min		Max
BEND	Bendiness; total change of direction (deg/km)	0		60
HILL	Sum of rises and falls per unit distance (m/km)	0		45
H_R	Sum of rises per unit distance one-way links only (m/km)	0		45
H_F	Sum of rises per unit distance one-way links only (m/km)	0		45
PHV	Percentage of heavy vehicles (OGV1 + OGV2 + PSV)	2		30
V_L, V_H	Speed of light and heavy vehicles (kph	n/a		n/a
S_L, S_H	Speed/flow slope of light and heavy vehicles (km/h reduction per 1000 increase in Q)	0		55
Q	Flow, all vehicles, two-way or one-way (veh/hour/lane)	0		2300
Q_B	Breakpoint: the value of Q at which the speed/flow slope of light vehicles changes (veh/hour/lane)	1080	or	1200
V_B	Speed of vehicles at flow Q_B (km/h)	80		105
Q_C	Capacity: defined as the maximum realistic value of Q (veh/hour/lane)	1400		2250

E.3.2 Vehicle speeds for a given flow level are dependent on the geometric variables (BEND, HILL and H_R). The value of those variables should be calculated, and the relationships set out below applied, for **at least** each individual road link longer than two kilometres, **on which flows change as a result of the scheme being appraised**. For other links, similar roads may be allocated to one of a number of typical link types (eg D3L uphill) each representing averaged characteristics. The use of a single road type for roads with markedly different characteristics (particularly free-flow speeds) should be avoided.

E.3.3 Q_C, the maximum realistic value of Q, per lane is:

2330 / (1 + 0.015 x PHV) for motorways, and

2100 / (1 + 0.015 x PHV) for all-purpose dual carriageways.

E.3.4 Q_B, the value of Q at which the speed/flow slope changes, is 1200 and 1080 veh/hour/lane for motorways and all-purpose dual carriageways, respectively.

E.3.5 For flow levels less than the breakpoint (Q_B), the speed prediction formula for light vehicles, in km/h is:

$V_L = K_L$ 0.1 x BEND

0.14 x HILL (two-way links only)

- 0.28 x H_R (one-way links only)

S_L x Q

where K_L is 108 for dual 2 lane all purpose (Class 2),

115 for dual 3 lane all purpose (Class 3),

111 for dual 2 lane motorways (Class 4),

118 for dual 3 lane motorways (Class 5),

118 for dual 4 lane motorways (Class 6),

and S_L, the speed/flow slope for light vehicles, is 6 km/h per 1000 vehicles.

E.3.6 At flow levels **greater** than the breakpoint (Q_B) the speed prediction formula for light vehicles, in km/h is:

$V_L = V_B$ 33(Q Q_B)/1000

E.3.7 There is no allowance for merge delays at grade-separated junctions in the above formulae. Any such delays should be separately modelled as set out in detail in section E.9.

E.3.8 The speed prediction formula for heavy vehicles, which applies at all flow levels, in km/h is:

$V_H = K_H$ - 0.1 x BEND

- 0.25 x HILLS (two-way links only)

- 0.5 x HR (one-way links only)

where K_H is 86 for allpurpose (Road - Classes 2 and 3) and

93 for motorways (Road Classes 4, 5 and 6)

subject to the constraint (which is unlikely to apply before the breakpoint) that, if the calculated value of V_H is greater than V_L, then V_H is set equal to V_L

E.3.9 The dual carriageway speed/flow relationships are expressed in flow per lane and not flow per direction as is the case with single carriageways. The research to develop the speed/flow relationships was undertaken on links with close to the standard 3.65 metre width lanes and was not able to detect a significant width parameter for use in the speed prediction formulae. Also it found that, unlike single carriageway links, the average speed of light vehicles on all-purpose dual carriageways is not influenced by the presence of a hard strip. However, if the average lane width of the proposed scheme is significantly less than the standard 3.65 metres, it may be necessary to survey free-flow speeds and use a local speed/flow relationship.

E.4 Urban Roads (Road Classes 7 and 8)

E.4.1 In built up areas, generally subject to a local speed limit, the road network becomes more dense and intersections play a more significant role in determining speeds. The speed/flow relationships that have been developed for urban areas apply to the main road network in towns (population greater than 70,000) and cities where there is a 30 mile/h (48 km/h) speed limit. The relationships are designed to represent average speeds in towns on the roads that function as traffic links. A distinction is made between central and non central areas, and they include an allowance for an average number of junctions. They are linear relationships of fixed negative slope.

E.4.2 Central areas are defined as those including the main shops, offices and central railway stations, with a high density of land use and frequent multi storey developments, as in the widely used classification 'central business district' or CBD. Conurbations will have several CBDs whilst most freestanding towns will normally have only one. Streets which have commercial or industrial development but are not of a high density CBD nature should not be included in the central area category. Non central areas comprise the remainder of the urban area. With this classification, the central areas constituted between 4 per cent and 22 per cent (average 11 per cent) of the total street length in the networks of the 13 towns where the data used to develop the relationships were collected.

Table E.5 Definition of Variables Used in Speed/Flow Relationships for Urban Roads

Symbol	Variable Description	Typical Values	
		Min	Max
INT	Frequency of major intersections averaged over the main road network (no/km)	2	9
DEVEL	Percentage of road network with frontage development (%)	50	90
V	Average vehicle speed (km/h)	n/a	n/a
Vo	Speed at zero flow (km/h)	28	48
Q	Total flow, all vehicles, per standard lane (veh/h/3.65m lane)	0	1200
Qc	Capacity: defined as the maximum realistic value of Q (veh/h/3.65m lane)	800	

E.4.3 Away from the area of immediate interest (ie **where flows are not expected to change markedly as a result of the scheme**), the use of network class 7 or 8 speed/flow relationships which include an allowance for junction delays may be satisfactory. In such cases, all links in a particular area should have the same value for INT or DEVEL, and therefore the same speed/flow relationship per standard lane. Even then, changes in speeds would be predicted on a link by link basis, but the free-flow speed would be set by reference to the network average conditions.

E.4.4 The average vehicle speed V km/h at flow Q veh/h/3.65m lane is given by the relationship:

$$V = V_0 - 30 \times Q / 1000,$$

where V_0, the speed at zero flow, is defined below for central and non central areas.

E.4.5 The maximum realistic flow (Q_C) should be taken as 800 veh/h/3.65m lane. For urban links this value is not affected by the proportion of goods vehicles.

E.4.6 Where the introduction of the scheme being appraised is predicted to produce major flow changes in urban areas, junction modelling is recommended rather than use of these urban speed/flow relationships.

E.4.7 Non central areas (Class 7) are defined as all those areas not included in the central area definition. The average network speed V_0 in km/h at zero flow is given by the relationship:

$$V_0 = 64.5 \quad DEVEL / 5 \text{ km/h},$$

where DEVEL is defined as the percentage of the non central road network of the town that has frontage development, counting business and residential development as 100% and open space as 0%. DEVEL is normally in the range 50 - 90% with average values about 80%.

E.4.8 Central areas (Class 8) are defined as those including the main shops, offices and railway stations, with a high density of land use as in the widely used classification 'central business district ' (CBD). Streets which have commercial or industrial development but are not of a high density central business district nature should not be included in the central area. The average network speed V0 in km/h at zero flow is given by the relationship:

$$V_0 = 39.5 \quad 5 \times INT /4$$

where INT is a measure of the frequency of major intersections averaged over the main road network. INT is calculated by dividing the total number of lengths of road between major intersections in the central area by the total length of main road in the central area. Major intersections will generally be roundabouts or traffic signals, but they may also be uncontrolled junctions where a significant traffic movement loses priority.

E.4.9 This **network-based** value is **not** directly comparable with a **route-based** value used for suburban links. INT should generally be in the range 2 to 9 per km, with an average of about 4.5. A value less than 2 is not appropriate for central areas; if this occurs, part or all of the area classified as 'central' should be re classified as 'non central'.

E.5 Small Town Roads (Road Class 9)

E.5.1 The small town speed/flow relationships have been developed for towns with a population of less than 70,000 (where the main urban relationships do not apply) and for villages or short stretches of development. Table E.6 defines the variables used in the relationships and ranges of typical values.

Table E.6 Definition of Variables Used in Speed/Flow Relationships for Small Towns

Symbol	Variable Description	Typical Values	
		Min	Max
DEVEL	Percentage of route with frontage (%)	35	90
P30	Percentage of route subject to a 30 mile/h speed limit	0	100
V	Average vehicle speed (km/h)	n/a	n/a
V_B	Average vehicle speed at Q_B	38	57
Q	Total flow, all vehicles, per standard lane (veh/h/3.65m lane)	0	1200
Q_B	Breakpoint: the value of Q at which the speed/flow slope changes (veh/h/3.65m lane)	70	

E.5.2 Like the suburban speed/flow relationships (Classes 10 and 11), they do **not** apply to individual links but model traffic speeds **over the whole of a route that is subject to a speed limit of 30 or 40 mile/h**. Unlike the suburban relationships, however, they do **not** distinguish between light and heavy vehicles, and they **specifically exclude junction delays**; hence junctions where the route loses priority must be modelled separately.

E.5.3 The breakpoint flow Q_B is taken as 700 veh/h/3.65 metre lane. The maximum realistic flow (Q_C) should be taken as 1200 veh/h/3.65 metre lane.

E.5.4 The average speed in km/h of all vehicles for flows below the breakpoint (Q_B) is given by:

V = 70 - DEVEL/8 - P30/8 -12Q/1000

where DEVEL is the percentage of the length of route that has frontage development, counting business and residential development as 100% and open space as 0%: the value will normally lie in the range 35% - 90%.

E.5.5 For flows greater than QB, the average speed in km/h of vehicles is given by:

V - VB - 45 (Q - QB)/1000

E.5.6 These relationships should not be used for routes with P30 < 10% (that is for routes with an almost continuous 40 mile/h limit), DEVEL < 65 (that is with less than 65% development), and access friction < 3. Access friction is defined as the total number, both sides, of laybys, side roads and accesses per km (excluding house and field entrances) divided by the carriageway width in metres. In such cases, the route should be split into links, as appropriate, and the standard rural relationships should be used instead.

E.6 Suburban Roads (Road Class 10 and 11)

E.6.1 The suburban speed relationships apply to the major suburban routes in towns and cities where the speed limit is generally 40 mile/h (64 km/h).

E.6.2 The suburban relationships provide estimates of the average journey speed of light and heavy vehicles separately, including delays at junctions. Table E.7 below defines the variables used in the relationships and gives the ranges of values over which the relationships apply. The relationships cannot necessarily be taken to apply outside the given ranges of the variables. The geometric variables INT and AXS should be averaged over a reasonable length of link, generally not less than two kilometres. **Congested junctions should be modelled separately** and not included in the calculation of the value of INT.

Table E.7 Definition of Variables Used in Speed/Flow Relationships for Suburban Roads

Symbol	Variable Description	Typical Values	
		Min	Max
INT	Frequency of major intersections (no/km)	0	2
AXS	Number of minor intersections and private drives (no/km)	5	75
PHV	Percentage of heavy vehicles (%)	2	20
V_L,V_H	Speed of light and heavy Vehicles (km/h)	n/a	n/a
S_L,S_H	Speed/flow slope of light and heavy vehicles (km/h) reduction per 100 increase in Q	0	45

V_0	Speed at zero flow (km/h)	48	64
Q	Total flow, all vehicles, per standard lane (veh/h/3.65m lane)	0	1500
Q_B	Breakpoint: the value of Q at which the speed/flow slope changes (veh/h/3.65m lane)	1050	1050
Q_C	Capacity: defined as the maximum realistic value of Q (veh/h/3.65m lane)	1350	1700

E.6.3 Generally, the use of **area-wide** class 10 or 11 speed/flow relationships which include an allowance for junction delays will be satisfactory **only** away from the area of immediate interest.

E.6.4 There are important differences between the definition of the variable INT for suburban roads and urban roads. For suburban roads, INT is specific to each section of route and major intersections are either roundabouts or traffic signals. Junctions between consecutive links should not be double counted, and classified junctions, whose delays are separately modelled, should be excluded from INT. The number of minor intersections and private drives, AXS, should be the total for both sides of the road.

E.6.5 The maximum realistic flow (Q_C) is the same for both single and dual carriageways and is calculated by the relationship:

$Q_C = 1500\ (92 \quad PHV)/80$ veh/h/3.65m lane

E.6.6 A standard value of 12% heavy vehicles (a typical value for main roads) is used to calculate the point of change of slope (Q_B) of light vehicles by the relationship:

$Q_B = 0.7 \times Q_C = 1050$ veh/h/3.65m lane

E.6.7 The speed for vehicles (V_0) at zero flow ($Q = 0$) in km/h is given by:

$V_0 = C \quad 5 \times INT \quad 3 \times AXS/20$

where, for single carriageways (Road Class 10),

C = 70 for light vehicles, and

C = 64 for heavy vehicles,

and, for dual carriageways (Road Class 11)

C = 80 for light vehicles, and

C = 74 for heavy vehicles.

E.6.8 The rate of decrease in speed (S) with increasing flow is the same for light and heavy vehicles and for single and dual carriageways. For values of flow (Q) less than the breakpoint (Q_B) for light vehicles and for all flow ranges for heavy vehicles:

$S_L = S_H = 12 + 50 \times INT/3$ km/h per 1000 vehicles

E.6.9 For values of flow (Q) greater than the breakpoint (Q_B), the speed/flow slope for light vehicles increases to:

$S_L = 45$ km/h per 1000 vehicles

E.6.10 The speed/flow slope for heavy vehicles does not increase when flow levels exceed the breakpoint. Therefore, the calculated speed of heavy vehicles can exceed the speed of light vehicles, when this occurs the speed of heavy vehicles (V_H) must be set to the speed of light vehicles (V_L).

E.7 Conversion to Passenger Car Units

E.7.1 The capacities and breakpoints in the relationships set out in sections E.2 to E.6 are specified in terms of vehicles per hour. It is common for all trip matrices to be converted to Passenger Car Unit (PCU) equivalents prior to assignment. The capacities and breakpoints in the speed/flow relationships therefore also need to be converted to passenger car unit equivalents. To achieve this conversion, two pieces of information are required: (a) PCU equivalent values, and (b) the proportions of the various vehicle types.

E.7.2 The following PCU equivalent values should be used:

LGVs on all road types:	1.0
HGVs on motorways and all-purpose dual carriageways:	2.5
HGVs on other road types:	2.0

E.7.3 For this purpose, HGVs consist of OGV1, OGV2 and PSV vehicle types.

E.7.4 Given the adoption of a PCU equivalent value for LGVs of 1.0, only the proportion of HGVs is relevant here. The proportion of HGVs will vary by link in the network and will also vary as the assignment calibration proceeds. The obvious source for these proportions is traffic counts by vehicle type. However, the 95% confidence interval for volumes of HGVs from a single day MCC is 28%. Moreover, counts will be available for only a sample of the modelled roads. It is therefore recommended that **average** HGV proportions should be calculated from counts by, at least, road type (motorway, all-purpose dual carriageway, single carriageway) and by type of area (rural, urban central, urban non-central, small towns and suburban). Other categories, such as roads leading to freight generators, should be considered depending on the availability of sufficient count data to support further categorisation.

E.7.5 As the calibration of the assignment model proceeds, the proportions derived from counts should be compared with the proportions derived from the assigned flows. Where discrepancies are significant, the conversions of the capacities and breakpoints should be revised. While it may be impractical to make this comparison at each stage in the calibration of the assignment model, periodic checks should be made and adjustments made as necessary, especially near the end of the calibration process.

E.8 Speed/Flow Relationships for Over-Capacity Conditions

E.8.1 In cases where junction flow metering is not modelled, assigned flows may exceed the capacities defined by the 'COBA' speed/flow relationships given in sections E.2 to E.6. In models where this can occur, a speed/flow relationship is required which applies to assigned flows in excess of capacity.

E.8.2 The Department has identified two relationships which may be used in conjunction with the 'COBA' relationships described in this appendix: (a) the so-called 'Advice Note 1A' (AN1A) relationship, and (b) the Akçelik relationship. These relationships are based on theory rather than empirical data[15].

The Advice Note 1A Relationship

E.8.3 The over-capacity speed/flow relationship specified in DoE Advice Note 1A[16] takes the form:

$$V = \frac{V_c}{1 + \frac{V_c}{8L}(E-1)}$$

$$E = \frac{Q}{Q_c}$$

where

V = speed (in km/h) at the assigned volume, Q (in PCU/h/lane), which is greater than the capacity, Q_c (in PCU/h/lane)

V_c = speed at capacity, Q_c

L = link length (in km)

E.8.4 In terms of time, this function is rather more simply represented as:

$$t = t_c + B(E-1)$$

where

t = link travel time at demand flow

t_c = link travel time at capacity

B = a constant equal to half the model period

---

15 Speed/Flow Relationships: Comparisons of Alternatives to Advice Note 1A, Technical Note A3.11, Parsons Brinckerhoff and The Denvil Coombe Practice, June 2009.

16 Advice Note 1A, Department of the Environment, 1971.

E.8.5 This function is based on deterministic queuing theory and assumes that an increase in delay as E exceeds 1 is equal to the time it would take half of the queued flow to dissipate. Advice Note 1A assumed a 15-minute model period, which gives rise to the factor of 8 in the formula above; this should be changed to 2 for models of one-hour periods.

E.8.6 The delay calculated by this function is constant irrespective of the link length. This means that the delay time on a longer link is a much smaller proportion of the total time than on a shorter link and so, at a given over-capacity assigned flow, the shorter link will have a lower average speed than the longer link.

Choice of Relationship

E.8.7 The Akçelik function is more unwieldy than the AN1A relationship. It needs the delay parameter, k_d, to be defined for different link types. Also, it would only be usable in software that allows custom functions to be defined for link delays.

E.8.8 Both relationships essentially assume that there is one queue that exists at the end (or start) of the link. This means that, when a link is arbitrarily split by a ‘dummy’ node (for connection to a zone centroid, for example), twice as much over-capacity delay would be modelled despite there being no real change to the network.

E.8.9 The achievement of convergence can be affected by the shape of the over-capacity relationship. If the slope of the relationship is too steep, there is a risk of an oscillating assignment; if the slope is too gentle, links may appear too attractive with the result that convergence is achieved with assigned flows in excess of capacity despite under-capacity alternatives being available. If convergence problems are experienced, analyses should be undertaken to assess the extent to which assigned flows are in excess of capacity. If only a few or a small proportion of the modelled links have assigned flows above capacity, the shape of the over-capacity speed/flow relationship may not be a significant cause and other aspects of the model should be considered. If, however, a significant number or proportion of the modelled links have assigned flows above capacity, the shape of the over-capacity speed/flow relationship should be reviewed.

E.8.10 The effect of the Advice Note 1A relationship may be adjusted by altering the length of the affected links, either by removing or adding dummy nodes. Using two links instead of one will effectively double the over-capacity delay and using one link instead of two will effectively halve the delay. The result may be thought of as effectively doubling or halving the overall slope of the relationship.

E.8.11 The shape of the Akçelik relationship may be adjusted by changing the value of k_d. At $k_d = 0$, the Akçelik relationship is equivalent to the AN1A relationship. Increasing k_d increases the amount below the AN1A relationship which the Akçelik relationship tends towards, although the slope of the Akçelik relationship remains essentially the same. However, changing k_d will impact on the slope of the relationship for volume to capacity ratios in the region of 1.0 to 1.2.

E.9 Merge Modelling on High Speed Roads

E.9.1 Research by TRL into traffic behaviour on high speed roads (published in CR 279) has allowed the speed/flow relationships to be updated. Further research then developed a modified speed/flow relationship on the links between junctions and a relationship for the delay caused by merges on high speed roads. The resultant advice on merge delays applies to all merge types (including lane gains) which involve weaving and merging manoeuvres.

E.9.2 The merge delay formula is:

Delay (seconds per vehicle) = 227 x (downstream flow/average capacity - 0.75)

E.9.3 This is applicable when flows exceed 0.75 times average capacity and increase linearly with flow. When flows reach average capacity delay per vehicle is 57 seconds. It applies equally to all traffic when acceleration lanes are provided to Departmental design standards for high speed merges (80 km/h or higher), as then traffic can merge with minimal delay at that point. The extra traffic joining at such merges reduces gap lengths in the downstream traffic flow until the traffic is able to spread out. However, when flows are heavy, average capacity may be exceeded and "flow breakdown" may follow starting 0.5 to 2 kilometres downstream from the merge point, rather than from the merge point itself. After flow breakdown, capacity is reduced by 5-10% and speeds are significantly reduced. Average capacity reflects situations both with and without flow breakdown.

E.9.4 Behaviour at such merges is therefore different from the ‘Give Way’ mechanism, usually based on a gap acceptance, which forms the underlying basis of most congested assignment packages. The mechanism of demand exceeding capacity can be represented in most congested assignment packages by means of their "blocking back" procedures. These use a mechanistic process which at some point usually reduces effective capacity on upstream links and can thus propagate blocked backed queues on either a) the upstream link designated as giving way, or b) where priorities are designated as equal, predominantly on the more congested upstream link (which can be either the mainline or joining feeder).

E.9.5 There is no evidence to support either a priori rule at High Speed Merges. Hence, unless evidence has been collected over 10 or more days which shows that a more appropriate factor can be calibrated locally, blocking back in traffic models of high speed merges should be controlled by the analyst to delay joining and mainline traffic to an equal extent. This can be achieved by keeping the modelled blocking back delay downstream of the merge. This may require careful choice of the model parameter for the average length of a vehicle in a queue, or other measures of stacking capacity. For example at least 15% overcapacity can be handled on a 2km link length by using 5.75m per vehicle (which is in practice less than that observed on High Speed Roads).

E.9.6 Thus, the recommended procedure is to model high speed merges on a ‘No Priority’ basis with a downstream link length of 2 km (unless merges are more closely spaced than that) and check where blocking back is forecast to occur. If any blocking back is forecast upstream of the merge consider reducing the parameter representing the average length of a vehicle in a queue for that link. The number of modelled blocked back vehicles should be monitored and the potential for their interaction with other traffic movements which do not pass through the merge considered. Alternative procedures could then be justified and may be required where such interactions already occur.

E.9.7 Such modelling needs to be complemented by the use of a special node type at the downstream end of the link to simulate the additional merge delay. Alternatively the link may be given a special speed/flow relationship, which mimics standard speed/flow effects up to the breakpoint (1,200 veh/h/lane or 1,080 on all-purpose roads) and introduces the additional merge delays as per paragraph E.9.2. The slope of such a relationship will be dependent on the length of the link and will vary significantly for different percentages of HGV.

E.9.8 It is typically assumed that motorway merge delays apply equally to all streams of traffic, irrespective of whether they are entering at a junction or are already on the main carriageway. Whilst this may be a suitable assumption when flows are typical, it may not be appropriate when there is an exceptionally high level of merging traffic or when there is an atypical mixing heavy vehicles on the main carriageway. In such cases, it can be helpful to extend the calculation of merge delays to include their separate effects on merging and mainline traffic. See the reference in the footnote[17] for further information on this topic.

[17] Liu R (2008). Towards a Generic Guidance for Modelling Motorway Merge. Final Report for Department for Transport, ITEA Division.

Appendix F: Tiered Model Systems

F.1.1 If a tiered modelling approach is being considered, it is a requirement that the demand estimated by the higher tier model and supply estimated by the detailed model meets the guidelines for acceptable demand-supply convergence in TAG unit M2.1. There must always be a check that the demand response to a particular price predicted by the upper level is consistent with that predicted at the lower level, and especially that the aggregate speed-flow relationship of the higher level is fully compatible with that governing the network model.

F.1.2 The ultimate proof that the upper tier demand changes are consistent with the lower tier cost changes is to run the demand model and detailed highway assignment model to convergence. However, this will rarely be practical for more than one or two tests (which is why a tiered system would be considered necessary). One way in which consistency can be achieved is to design the upper tier or simplified highway assignment model so that it replicates as closely as possible the costs and cost changes produced by the lower tier highway assignment model.

F.1.3 The recommended procedure for creating a simplified model is one that is fully automated, using inputs from the detailed version. Ideally, there should be no manual coding, or manual adjustments to the coding, of the simplified model. This applies to both base year and forecast versions of the model. The simplified and detailed models should have the same network link and node structure and the same zone system[18] as the detailed model (i.e. there is no spatial simplification within the recommended process).

F.1.4 For the rest of the Fully Modelled area and the External Area, no simplification is required and the network coding from the detailed model can be used directly. For the Area of Detailed Modelling, where junction delays are modelled explicitly, there are two alternative approaches to simplification:

- summarising of junction turn flow/delay effects to form link flow/delay curves
- direct use of turn flow/delay curves

The former is a simplified version of the latter. Direct use of turn flow/delay curves has proved to be greatly superior in the replication of detailed model flows and journey times. The remainder of this discussion focuses on this method.

F.1.5 Turn flow/delay curves derived for a simplified highway assignment model need to be applicable for the whole of the duration of the modelled period (generally one hour). However, the flow/delay curve for a particular turn is not a static property. It is, to a degree, dependent upon the level of flow for any opposing

[18] Examples of where the simplified model uses an aggregated version of the zone system do exist, but results have not generally been satisfactory.

movements and upon other factors such as traffic signal timings. Within each iteration of the detailed assignment process, these factors are varied to some degree. As a result, turn flow/delay curves have to be developed from the results of a converged assignment and not just from the network coding for the detailed model.

F.1.6 In the development of the base year model, the turn flow/delay curves that are the outturn of the validated detailed model should be used for the simplified model. The process of extracting turn flow/delay curves will vary depending upon the software used for the detailed model. Some software explicitly develops and stores such curves as part of its internal processes, e.g. because a junction simulation process is employed. For other software, points on a curve may have to be developed through a process of assigning varying levels of demand to the base year network.

F.1.7 Following completion of the base year simplified model, the following comparisons with the detailed model should be undertaken (as a minimum):

- assigned flows using base year demand
- inter-zonal times using base year demand
- inter-zonal times and time changes using demand factored by 0.9, 1.1, and 1.3

F.1.8 A formal link flow validation (that is, a comparison of modelled flows and counts) should be carried out as specified in section 9. It would be surprising if the assigned flows in the simplified model validated as well as those in the detailed model. However, whether the assigned flows in the simplified model are sufficiently accurate may be judged from the accuracy with which inter-zonal time changes are replicated. It is these inter-zonal time changes which are the most important output from the simplified model as it is these time changes which will drive the demand changes.

F.1.9 Regression analyses of detailed and simplified times and time changes should be carried out. It is to be expected that the intercepts will be zero or very near to zero. If the slopes lie outside the range 0.95 to 1.05 and/or if the R^2 value is less than 0.90, ways of improving the performance of the simplified model should be investigated.

F.1.10 Preparation of a simplified model for future years is more complicated than for the base year, as the future demand is not known (and indeed is an output of the supply/demand process). The recommended procedure is as follows:

- prepare and check the future year detailed model networks
- assign the future matrices that have been prepared assuming no cost change from the base year (ie the reference case demand)
- from the above prepare the simplified network as per the base year

F.1.11 In iterating the demand and simplified supply models to convergence, it is desirable to re-build the simplified network at least once, ie inserting a run of the detailed highway assignment model.

Appendix G: Reports

G.1 Introduction

G.1.1 The following two reports are required which relate to the advice in this unit:

- Highway Assignment (Local) Model Specification Report
- Highway Assignment (Local) Model Validation Report

G.1.2 Reports maybe also be termed ‘Local’ where they refer to a specific geography or application. The term helps to provide the reader with an understanding of whether the reporting relates to a general use model (e.g. for a number of applications) or a specific application (e.g. for a scheme appraisal). Where general use models are adopted for a specific scheme appraisal, evidence should be provided to justify the model is suitable for such application.

G.1.3 The recommended structures of these reports are set out below.

G.2 Highway Assignment Model Specification Report

G.2.1 This report should form part of an overall Appraisal Specification Report, as required particularly at the end of stage 1 of the appraisal process in order to specify and agree the modelling at an early stage. The following structure should be used.

1 Introduction

2 Proposed Uses of the Model and Key Model Design Considerations

Proposed Uses of the Model: Scenarios to be Forecast and Interventions to be Tested

Key Model Design Considerations

3 Model Standards

Validation Criteria and Guidelines

Convergence Criteria and Standards

4 Key Features of the Model

Fully Modelled Area and External Area

Zoning System

Network Structure

Centroid Connectors

Time Periods

User Classes

Assignment Methodology

Generalized Cost Formulations and Parameter Values

Capacity Restraint Mechanisms: Junction Modelling and Speed/Flow Relationships

Relationships with Demand Models and Public Transport Assignment Models

5 Calibration and Validation Data Specification

Traffic Counts at Roadside Interview Sites

Traffic Counts for Matrix Estimation

Traffic Counts for Validation

Journey Time Surveys for Calibration and Validation

6 Network Development Methodology

Network Data, Coding and Checking

Junctions:	Flow/Delay Relationships
	Signal Timings
	Saturation Flows
Links:	Speed/Flow Relationships
	Fixed Speeds

7 Trip Matrix Development Methodology

Travel Demand Data

Partial Trip Matrices from Surveys

Trip Synthesis

Merging Data from Surveys and Trip Synthesis

8 Network Calibration and Validation Methodology

Network Calibration

Network Validation

9 Route Choice Calibration and Validation Methodology

Route Choice Calibration

Route Choice Validation

10 Trip Matrix Calibration and Validation Methodology

Trip Matrix Estimation

Trip Matrix Validation

11 Assignment Calibration and Validation Methodology

Assignment Calibration

Assignment Validation

12 Summary of Model Development, Standards Proposed, and Fitness for Purpose

Summary of Model Development

Summary of Standards Proposed

Proposed Assessment of Fitness for Purpose

G.3 Highway Assignment Model Validation Report

G.3.1 The following structure should be used.

1 Introduction

2 Proposed Uses of the Model and Key Model Design Considerations

Proposed Uses of the Model: Scenarios to be Forecast and Interventions to be Tested

Key Model Design Considerations

3 Model Standards

Validation Criteria and Guidelines

Convergence Criteria and Standards

4 Key Features of the Model

Fully Modelled Area and External Area

Zoning System

Network Structure

Centroid Connectors

Time Periods

User Classes

Assignment Methodology

Generalized Cost Formulations and Parameter Values

Capacity Restraint Mechanisms: Junction Modelling and Speed/Flow Relationships

Relationships with Demand Models and Public Transport Assignment Models

5 Calibration and Validation Data

Traffic Counts at Roadside Interview Sites

Traffic Counts for Matrix Estimation

Traffic Counts for Validation

Journey Time Surveys for Calibration and Validation

6 Network Development

Network Data, Coding and Checking

Junctions:	Flow/Delay Relationships
	Signal Timings
	Saturation Flows
Links:	Speed/Flow Relationships
	Fixed Speeds

7 Trip Matrix Development

The recommended content of this chapter is detailed in Appendix F of TAG unit 2.2.

8 Network Calibration and Validation

Network Calibration

Network Validation

9 Route Choice Calibration and Validation

Route Choice Calibration

Route Choice Validation

10 Trip Matrix Calibration and Validation

Trip Matrix Estimation

Trip Matrix Validation

11 Assignment Calibration and Validation

Assignment Calibration

Assignment Validation

12 Summary of Model Development, Standards Achieved, and Fitness for Purpose

Summary of Model Development

Summary of Standards Achieved

Assessment of Fitness for Purpose