Department for Transport

TAG unit M3.2

Public Transport 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 - Assignment 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 4
   1.1 Scope of the Unit 4
   1.2 Relationship of this Unit to Other Advice 5
2. Designing a Public Transport Assignment Model 6
   2.1 Introduction 6
   2.2 Modelled Area 9
   2.3 Design of the Zoning System 11
   2.4 Network Representation 13
   2.5 Time Periods 13
   2.6 User Classes 14
   2.7 Assignment Methods 15
   2.8 Representing Responses of Public Transport Operators 17
3. Generalised Cost Definition 18
   3.1 Introduction 18
   3.2 Access/Egress Time and Waiting Time 20
   3.3 Fare Structures 23
   3.4 Boarding Times/Interchange Times 24
   3.5 Quality Factors for Stops and In-Vehicles Times 25
   3.6 Capacity Constraints and Crowding 27
   3.7 Path Building 30
   3.8 Cost Skimming 33
4. Validation and Convergence Standards 36
   4.1 Introduction 36
   4.2 Fitness for Purpose 36
   4.3 Validation Criteria and Guidelines 37
   4.4 Convergence Measures and Acceptable Values 41
5. Calibration and Validation Data 42
   5.1 Introduction 42
   5.2 Data for Model Calibration and Validation 43
   5.3 Journey Time Data 44

6. Coding and Calibration of Network and Services 44
6.1 Introduction 44
6.2 Network Data Sources and Coding 44
6.3 Public Transport Services Data Sources and Coding 50
6.4 Pre-Calibration Checks 57
7. Route Choice Calibration 59
7.1 Introduction 59
7.2 Calibration 60
8. Trip Matrix Calibration 61
8.1 Introduction 61
8.2 Refinement of Prior Trip Matrices by Matrix Estimation 62
9. Assignment Calibration 63
10. Validation 64
10.1 Introduction 64
10.2 Network and Service Validation 64
10.3 Route Choice Validation 65
10.4 Trip Matrix Validation 66
10.5 Assignment Validation 66
10.6 Testing Responsiveness 67
11. Reporting 67
12. References 69
13. Document Provenance 69
Appendix A: Glossary of Terms 70
Appendix B: Assignment Methods 72
B.1 Deciding on an Assignment Approach 72
B.2 Path Choice Methodology 77
Appendix C: Model Convergence 80
C.1 Base Year and Future Year forecasts 80
C.2 Assessing accuracy of final results 80
C.3 Presentation of convergence results 81
Appendix D: Reporting Requirements 82
D.2 Public Transport Assignment Model Specification Report 82
D.3 Public Transport Assignment Model Validation Report 84

1. Introduction

1.1 Scope of the Unit

1.1.1 This TAG unit provides detailed advice on developing, calibrating and validating public transport assignment models and covers the following topics:

- designing a public transport assignment model including a discussion on assignment methods in section 2
- generalised cost, path building and cost skimming in section 3
- validation and convergence standards in section 4
- calibration and validation data in section 5
- coding and calibration of network and services in section 6
- route choice calibration in section 7
- trip matrix calibration in section 8
- assignment calibration in section 9
- validation in section 10
- reporting in section 11

1.1.2 The unit has been structured to recommend a sequential, logical approach to calibration. Effort should be made to ensure that each element in the sequence (such as zones, network structure, centroid connectors, network and service coding, uncalibrated generalised cost parameters, capacity restraint procedures and trip matrices) is developed as accurately as reasonably possible before moving on to the next element (see Figure 1). Some assumptions can only be tested as the process develops and so practitioners should review and refine assumptions as the model development progresses. It is also possible to read this guidance as a reference to individual elements of public transport models during model development or update.

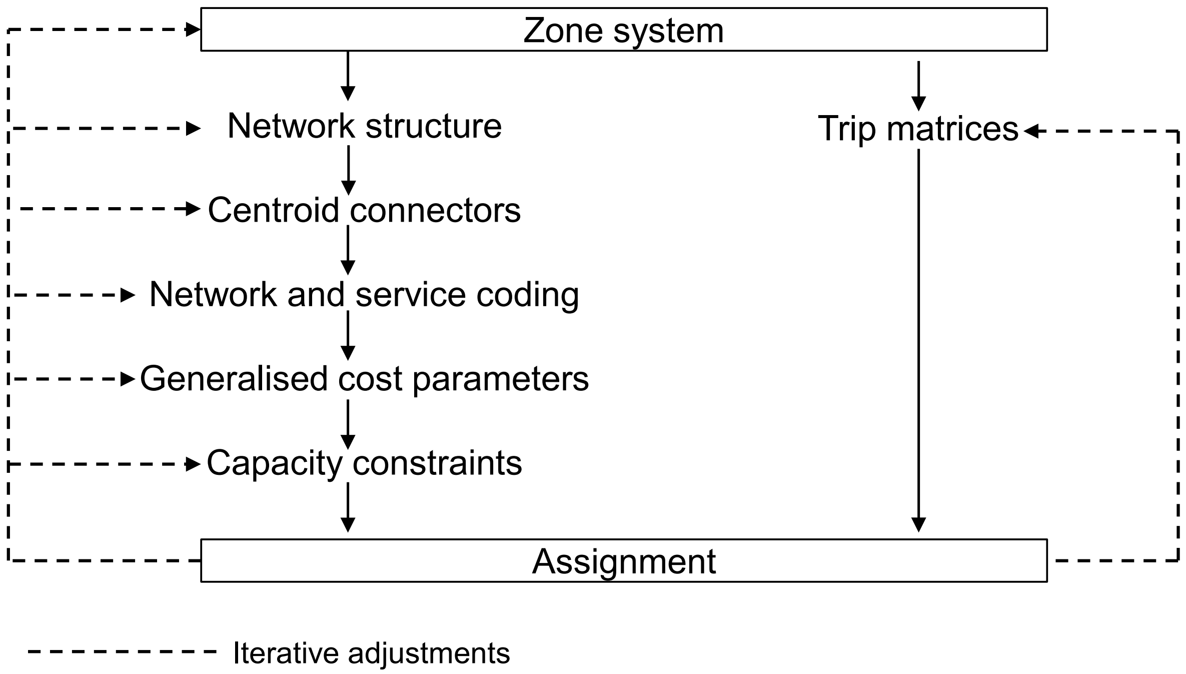

Figure 1 Calibration Process

1.1.3 This unit is relevant to all public transport sub-modes. However, its focus is on public transport assignment in a multi-modal context and can be applied for rail, bus, light rail schemes or other public transport schemes.

1.1.4 Uni-modal (rail-only) models have been used for rail-only studies where interactions with other modes are of secondary importance. Heavy rail schemes are also modelled in uni-modal contexts using pre-existing industry models such as MOIRA, focusing on demand elasticities to represent the net impact of other demand drivers, based on the Passenger Demand Forecasting Handbook (PDFH) guidance. Whether a uni-modal or a multi-modal approach is most appropriate is also covered in TAG unit M1.1 Principles of Modelling and Forecasting.

1.2 Relationship of this Unit to Other Advice

1.2.1 This unit **excludes** advice on the following topics which are related to public transport assignment modelling and are covered in other TAG units:

- travel demand data collection, including public transport intercept surveys, in-vehicle passenger counts and boarding / alighting counts (TAG unit M1.2 Data Sources and Surveys)
- the development of prior trip matrices for public transport assignment models (TAG unit M2.2 Base Year Demand Matrix Development)
- developing a variable demand model with a public transport component (TAG unit M2.1 Variable Demand Modelling)

- preparing forecasts using public transport assignment models (TAG unit M4 Forecasting and Uncertainty)

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

- TAG unit M1.1 Principles of Modelling and Forecasting
- TAG unit M1.2 Data Sources and Surveys
- TAG unit M3.1 Highway Assignment Modelling
- TAG unit A5.3 Rail Appraisal
- TAG unit M4 Forecasting and Uncertainty (Modelling a Scenario – Rail Schemes)
- TAG unit M5.1 Modelling Parking and Park and Ride

2. Designing a Public Transport Assignment Model

2.1 Introduction

2.1.1 The key role of public transport assignment models in demand forecasting is the provision of levels of service, the travel times, distances and costs associated with public transport trips between origin-destination pairs. This involves the estimation of passenger demand, per mode, per route and route section. Level of service measures distinguish components such as transfer and wait times, and, where relevant, for different transport modes for example bus, tram and heavy rail. These measures are used in scheme appraisal and demand modelling.

2.1.2 This section provides advice on the following topics:

- the role of public transport assignment in the overall model structure
- the specification of the Internal and External areas of the model
- the design of the zoning system
- the representation of the network
- the time periods which should be modelled
- the specification of the classes of user

- the assignment method
- the representation of the response of public transport operators

2.1.3 Whilst the focus of this unit is on modelling for scheme appraisal, a public transport assignment model may also support other applications such as operational support, fleet sizing, design of stations and other infrastructure, logistics, timetable changes, vehicle upgrades etc. This guidance will not cover these applications, but many of the principles for developing these use cases are the same. Therefore, it is important to consider how the model will be used, and the applications it will support, as this impacts the model design.

Role of public transport assignment in overall model structure

2.1.4 The role of public transport assignment modelling within the modelling and forecasting process is discussed in TAG unit M1.1 Principles of Modelling and Forecasting. A public transport assignment model allows the estimation of passengers on services and associated travel behaviours. This unit focuses on the development of a multi-modal public transport network to estimate travel costs, and how passengers use the public transport services modelled. The estimation of base year public transport demand is the focus of TAG unit M2.2 Base Year Demand Matrix Development, whilst the development of demand forecasts is covered in other units referenced above.

2.1.5 A multi-modal public transport network is often an important component of a variable demand model (TAG unit M2.1 Variable Demand Modelling). Such a model considers that changes to transport conditions will also cause changes in transport demand, including changes to mode of travel. Consideration should be given to whether certain travel choices should be represented in the demand modelling or assignment modelling (this unit). Such choices are:

- Station choice
- Access mode
- Park and ride access to public transport
- Sub-mode choice

2.1.6 **Station choice** involves the allocation of access trips via any mode which is not explicitly represented, from a zone to individual stations. This can be undertaken within a public transport assignment model however it may not be proportionate to model competition between access modes or between competing stations in this way. For this reason, it may be more appropriate to implement station choice within a variable demand model (VDM).

2.1.7 Public transport assignment software typically offers limited scope to distinguish **access modes**. Where the model purpose particularly focuses on access provision, such as station parking or city centre cycle hire, it may be proportionate to represent such choices in a variable demand model that explicitly represents the generalised costs of each of the access/egress alternatives.

2.1.8 The decision to model **Park and Ride** is made as part of the wider mode choice consideration of whether to use car or public transport as the main mode for a trip. For this reason, it is most appropriate to include Park and Ride choice within a VDM (see TAG unit M2.1 Variable Demand Modelling). A station choice layer may be implemented below the park and ride layer if there is more than one park and ride site available. Detailed advice on park and ride modelling is contained in TAG unit M5.1 Modelling Parking and Park and Ride.

2.1.9 Demand should be considered as complete door to door journeys. Where data sources describe individual legs[1] of a journey it is important to avoid double-counting trips, for example where mixed-mode trips will have been observed from surveys for more than one mode, for example, a bus count that includes passengers travelling to a rail station, and a rail survey that also includes those passengers when they board a train.

2.1.10 Mixed-mode trips should be loaded onto the appropriate networks for all legs of their journey. For example, a trip that uses bus to access rail should appear in the bus flows between its origin and the railway station and in the rail flows from the railway station to its destination. This may be achieved either with a single network model with certain modes and routes banned to certain segments, or with separate network models with zones at mode interchange points. The former is simpler to implement and easier to ensure consistency.

2.1.11 The form of the public transport assignment model is determined by whether trips are allocated between public transport **sub-modes** at the mode choice stage or at the assignment stage. The key factors which govern the choice of method are:

- the proportion of passengers who use more than one public transport mode
- the numbers of different mode combinations used to a significant extent
- the level of competition between different public transport modes
- The public transport scheme under consideration and its dependency on or competition with other modes.

Although allocation as part of the assignment process could be generally less reliable than an explicit mode choice model where competition between modes exists and is important, mixed modes are generally easier to handle at the assignment stage, especially if there are a number of mode combinations to consider. A number of multi-modal public transport models for major cities use the assignment process to allocate to mixed modes.

[1] A particular stage or portion of a trip, such as accessing the public transport services from the trip origin.

2.2 Modelled Area

2.2.1 The modelled area should be designed to represent all transport impacts, with greatest detail local to the scheme and reducing detail away from the scheme towards a simplified external area.

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

2.2.3 The geographic coverage of public transport 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

2.2.4 The second and third requirements are particularly important where a public transport assignment model will be linked to a demand modelling system.

2.2.5 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 model analysis more complex than necessary. 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. In some cases, it may be more appropriate to add more detail to the local area of interest rather than removing areas of the network. Practitioners should consider the appropriate and proportionate approach depending on the purpose of the modelling.

2.2.6 Within the overall modelled area (in many models this encompasses 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.

- **Internal Area**: This is the area over which proposed interventions have the largest influence. Modelling detail in this area would be characterised by small zones, representation of all trip movements, detailed representation of most or all public transport services and detailed representation of access/egress legs[2]. This area of the model should have consistent and detailed demand and supply representation.
- **External Area**: This is the area where the impacts of interventions are expected to be small and assumed to be reasonably negligible. It would be characterised by the representation of a large proportion of the rest of Great Britain with a simple network and large zones. It would have a partial

[2] A particular stage or portion of a trip

representation of demand where the main consideration is to ensure passengers approach the Internal Area in the right corridor/services.

2.2.7 When crowding is modelled, there may be impacts on routeing distant from the proposed interventions. Checks should be made for how and where existing and future/potential crowding may affect demand (for example suppressing it) and how the interventions/schemes will interact with this. If corridors or areas are identified where crowding is considered important, these should be included in the internal area.

2.2.8 The highway model (part of a multi-modal modelling system) needs to include the areas where material changes to highway flows might result from the public transport scheme so that highway costs and congestion relief benefits can be calculated.

2.2.9 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. For some studies, it may be appropriate to define a smaller Internal Area, typically representing the immediate geographical areas around the scheme being assessed. This is where transport effects are most pronounced. In this case the External Area, where transport effects are smaller, may be extensive.

2.2.10 In other cases, the model may be conceived as a general-purpose tool where the Internal Area should be determined by an increasing geographical level of details near the area or corridor of interest. There is a need for clarity about the uses to be made of the model. Therefore, the key question which needs to be considered is **the purposes for which the model can be used?**

2.2.11 A further consideration when defining the area boundaries is the potential need to undertake analysis and reporting, in line with, for example administrative boundaries. The need to report relevant and effective outputs to stakeholders should be considered at this stage of model design. The drawback of relying on administrative boundaries is they are not based on current or anticipated future travel catchments. This can be mitigated by exploring other definitions such as travel-to-work areas.

2.2.12 Physical features, such as rivers or railway lines can often form suitable boundaries where they allow for limited crossing opportunities and therefore enable large travel demand flows to be intercepted.

2.2.13 If there is an existing public transport assignment model, initial tests could be modelled of the scheme (or a rough approximation of it) to determine the approximate area over which transport effects occur, thus informing the model boundaries (where it is appropriate to modify the model area).

2.3 Design of the Zoning System

Design principles

2.3.1 This section discusses the design principles for the zoning system, explaining considerations relevant to public transport assignment models.

2.3.2 It is important to avoid the temptation to allow the zoning system for the public transport model to be distorted from the ideal by partially adapting a zoning system designed for some other model. The zoning systems appropriate for highway assignment modelling and public transport modelling may be quite different (for example highway zones may follow road boundaries whereas public transport zones may follow rail tracks and natural boundaries) and the former should not unduly influence the latter in the case of models for the appraisal of major public transport schemes.

2.3.3 Whilst the considerations in the design of the public transport model zone system are the granularity of the zone system of other models and enabling logical access/egress trip routeing (discussed below), an overarching consideration should be compatibility of the model data with other data sets. There should be alignment with Office for National Statistics (ONS) boundaries otherwise it is challenging to translate planning data forecasts. Compatibility of datasets will allow more efficient exchange and analysis between the data and the model, also reducing the chance of errors. These data sets could include planning data sourced at different levels of granularity (for example, LAD, MSOA, LSOA, OA and Workplace zones), as well as fare zones and Electronic Ticket Machine (ETM) data.

2.3.4 Often the most critical datasets to link with will be the associated demand and highway assignment models. As the design of the zone system should consider the requirement to represent access/egress legs of trips in a logical manner, it should adopt a level of zone granularity which is suitable given those levels of modelling accuracy required. Unless the models are especially large, there are rarely model run time and data constraints that prevent adapting of a single zone system for a multi-modal model system.

2.3.5 The above conditions illustrate the need for careful design of the model zone system. If zones, or data boundaries cannot be combined, it is recommended to employ a nested zoning system between models that ensure data and spatial alignments.

Detailed considerations

2.3.6 Access and egress of public transport journeys in urban areas is usually represented in public transport assignment models by access modes, such as walking or cycling. These legs of the journey can strongly influence public transport routeing decisions and sub-mode choice.

2.3.7 Therefore, the design of the zoning system should facilitate logical access choices between zones and stations/public transport stops. In densely developed urban areas, this can be achieved through a disaggregate/finer representation of zones and the walk network. This is so that access routes are determined based on the relative attractiveness of actual walk routes. For more rural areas, the competition between alternative stations and stops within zones should still be represented. This could be via access links with direct centroid connectors, simplified walk network representation, or a combination of both. When employing centroid connectors, the impact of natural and man-made barriers on access severance should be taken account of.

2.3.8 Care should be taken if there is a large variation in access across a zone and the stations or stops that serve the zone. Unless the model purpose is for broad area -wide interventions, variations should be limited to a few minutes for zones in the internal area. Often this may mean that zones contain a single rail station or a few bus stops. If the model has a broad strategy / corridor purpose, a more aggregate treatment of zoning and associated access may be considered.

2.3.9 In the External Area, the zone structure will typically be coarse therefore an accurate representation of access choices will be more difficult to achieve. A common approach is to model access to the stations/stops serving a zone with an identical access time (which could be based on a mix of access modes). This is so that access choice is based on onward costs of travel to the Internal Area which will be more accurately represented. The emphasis should be on ensuring that trips are split realistically between corridors and services on the approach to the Internal Area.

2.3.10 A more refined zone structure should enable more accurate calculation of access/egress to stations/stops and therefore a better validation. However, it should be considered what level of accuracy is required and whether the need for higher resolution demand data, the greater development costs and longer model run times are worthwhile costs to incur. A more refined zone structure will also enable demand at individual stations (and therefore on links between stations) to be modelled more accurately. This will be crucial if crowding is an important consideration.

2.3.11 The need to represent specific locations as zones should also be considered to separately account for their transport demand impacts. Examples are airports or individual terminals within airports, park and ride sites and land-use developments. This method can form a useful basis for studies to assess the impacts of land-use developments on public transport demand and options to mitigate their impacts.

2.3.12 Practitioners should avoid changing the base year zoning for the assessment of future year schemes. Zone requirements for future year developments need be considered at an early stage of model development and applied consistently to both the Base Year and Forecast Year networks. This can be achieved at the model specification stage and associated reporting, preventing too many iterations of zone structures.

2.4 Network Representation

2.4.1 The public transport network usually consists of the following:

- Links and nodes – representing the public transport infrastructure (roads, tracks, pavements, stations and bus stops) and accessibility (pedestrian, cycling)
- Centroids – consolidated points, that are generally at the geographic centre of a zone that represents homes, jobs etc
- Centroid connectors – links that load passengers onto the network from a zone centroid
- Public transport services – representing the trains, buses, trams. The coding of public transport service routes will make reference to the links and nodes of the infrastructure network

2.4.2 The ideal structure is for the highway and public transport networks to both have the same node and link structures (with the addition of rail, tram and walk links to the public transport network), although this is not always possible. Representation, coding and calibration of the network and services is covered in section 6.

2.5 Time Periods

2.5.1 Trip purpose and travel patterns vary by time of day, as do the service levels provided by operators. Where there are variations in service levels, public transport assignment models should represent the morning and evening peaks and the inter-peak period separately. There may also be a need to model further time periods, such as off-peak times, 24 hours and weekends, if service levels vary and the impacts cannot adequately be represented by the appraisal 'annualisation factors' (see TAG unit A1.3 User and Provider Impacts).

2.5.2 Travel patterns can also vary significantly between different months and days of the week. Public transport assignment models normally represent a weekday in a 'neutral' period. Refer to TAG unit M1.2 Data Sources and Surveys for information of 'neutral' periods. This may not be applicable in all instances and the days represented should be selected with careful consideration of the purpose of the transport model (refer to TAG unit M1.1 Principles of Modelling and Forecasting for more information on the purpose of the transport model).

2.5.3 The choice of which time periods to represent should be determined by the availability, quality and analysis of survey data considering how demand and supply varies. Practitioners should consider hourly variations to determine modelled periods and should use available data ideally spanning a full year to consider the need for modelling other periods and establish annualisation factors (see TAG unit M1.2 Data Sources and Surveys).

2.5.4 The aim should be to identify those periods where there are variations in the level and quality of service provided. Highway congestion affects bus travel times and operators often run more frequent services in weekday peak periods. If crowding impacts are of importance, these are also typically greater in peak periods. The times during which crowding is of a sufficient level to impact the quality of travel must also be considered. The duration of the peak periods represented in the model (peak hour or peak period) should be based on the pattern of variation in level of service, and where modelled, passenger crowding.

2.5.5 Models which cover multiple cities may require longer time periods to account for different peak times and longer trips. This will help avoid problems in calibrating to counts taken across a variety of locations with different peaking times. Models with longer trips tend to have longer time periods for the same reason.

2.5.6 When deciding whether a particular trip and a particular public transport service should be included in a particular time period, a consistent approach is required. Some illustrative approaches could be:

1. Start (or mid-point, or end) of the service being in the time period
2. The proportion (probability) of the service being in the time period
3. The time the service crosses the most important screenline / cordon

The approach will depend on the nature of the peak period and it is important that the practitioner explains which approach is taken and why.

2.5.7 Consideration should be given to whether the time period in the public transport model should align with the highway model. The different modes may peak at different times and priority needs to be determined based on the scheme being tested and the impacts being assessed. However, for compatibility with the associated demand model, avoidance of errors and ease of calculation, consistency in time periods is highly desirable. However, it is possible to have skims covering a consistent time period, for example, from three 1-hour highway models (weighted by demand) and one 3-hour public transport model.

2.6 User Classes

2.6.1 User classes refer to different segments of demand and include categories of people and trips such as purpose (for example, commuting, education, business or leisure), or ticketing (availability of concessions or ticket products). Within a demand model the most significant segmentation is usually car availability.

2.6.2 Segments of demand respond differently to changes in the transport system in relation to fares, value of time and other travel behaviours. These user classes are likely to be estimated separately in a demand model leading to user classes being assigned separately.

2.6.3 If differences in responses between user classes are less significant, a single user class can be modelled. In this case it is adequate to disaggregate demand and skims subsequently as part of the appraisal (see 3.8 for further information).

2.6.4 As part of its standard calculations, the associated demand model should already have split public transport demand by a variety of segments and therefore it is logical that these form a suitable initial consideration for how user classes should be defined as part of the public transport assignment model. It is likely that the demand segments derived as part of the demand model will be more disaggregate than the needs of the public transport model and that many of these segments can be aggregated, thus reducing run times. The consideration for which segments, if any, should usually be retained is whether those segments respond differently to changes in generalised cost. For example:

- Business trips typically give more weight to travel time and leisure trips to fare.
- School pupil trips may have dedicated school bus services available.
- Where there are sufficient differentials between available fare products such that choice of fare influences route choice. For example, single/return trips may have a preference for longer, less expensive routes whereas season ticket trips may have a preference for comparatively shorter trips if such trips are eligible under the terms of the season ticket. Concessionary and multi-use tickets are also examples of differential fare products.

2.6.5 There may be the need to derive monetised time costs segmented by user class for appraisal purposes. This requires the application of different values of time by user class, usually business, commuting and leisure. The need for this is not in itself a reason to segment assignment. These calculations can be performed after completion of the public transport model assignment using generalised cost skims generated from the model run.

2.6.6 Depending on the model structure, segmentation by sub-mode may also be required. For example, if public transport sub-mode choice is performed as part of the demand model rather than assignment, or if Park and Ride (or car access to PT) trips are explicitly modelled.

2.7 Assignment Methods

Introduction

2.7.1 Public transport assignment calculates how passenger demand is distributed among different services and links in the public transport network. The two different assignment methods that are in common use in public transport assignment software packages are **timetable-based** (or schedule-based) and **frequency-based** approaches, whilst a further distinction can be made between

stochastic and **deterministic** assignment. The approaches are described briefly below, with advice on deciding on which approach to adopt given in Appendix B.

Timetable and frequency-based approaches

2.7.2 Timetable-based approaches reflect the actual clock face vehicle arrival/departure times at the time when users make their choices. The result of this assignment is that every service vehicle in the timetable for the modelled period is represented and has a passenger load after assignment This approach allows practitioners to take into account the dynamics of supply and demand and calculate the dynamics and variation of level of service attributes. It gives a better representation of the perceived inconvenience of waiting for a service or of changing services.

2.7.3 In the case of frequency-based assignment, departure times of individual services are not considered explicitly, but practitioners refer to the service headways, or to their inverse (the service frequencies). Therefore, it is not possible to calculate explicitly attributes that users consider in relation to individual service options, but only the average values that relate to that line. Frequency-based modelling constitutes the classical approach as it is usually simpler, requiring less input data and less computational power than timetable-based approaches. It corresponds to the steady state approach (See TAG unit M3.1 Highway Assignment Modelling) to user equilibrium highway assignment and therefore allows the use of some of the same techniques. It is easier to use for long term forecasting, as no future year timetable is required, and is generally sufficient for urban bus and LRT elements of the public transport network.

2.7.4 A timetable-based approach is more suited for the representation of short-term operational decisions or where services run at a low frequency, for example, in rural areas, or long-distance services. Such an approach is necessary to inform how relatively modest changes to the existing public transport network, such as improved frequency on a particular service, will impact demand and crowding. The UK rail industry has a standard approach to such modelling (contact DfT Rail Modelling team on railmodelling@dft.gov.uk for more information). The Rail Delivery Group's (RDG) Passenger Demand Forecasting Handbook (PDFH) (membership required) also has useful advice on modelling infrequent or low frequency services (see PDFH's Rooftop Model).

2.7.5 A frequency-based approach will be appropriate for most studies requiring a strategic public transport model and often where most services in the area run frequently. Typically, such studies require an assessment of impacts at a strategic level, be it along a transport corridor, within a region, city or suburb. The grouping of individual trains into services which represent entire time periods is an appropriate level of aggregation for such assessment. However, where frequencies are low, and interchanges are planned carefully, a frequency-based approach may overstate the wait time and a timetable-based method may be preferable. In this case, the additional model development and run costs would need to be justified by the value gained, and whilst more refined analysis may be of value, it needs to be considered what level of accuracy can be achieved at such a level of refinement, particularly as definitive assumptions for future year timetables are unlikely to be available, and assuming that these remain unchanged from the base year is also likely to be incorrect.

2.7.6 The choice of assignment approach also influences the modelling of waiting and interchange times (see 3.2 for further details)

Deterministic and stochastic assignment

2.7.7 There are parameters and options in the major public transport assignment packages that allow some deterministic and some stochastic features.

2.7.8 The basic assumption underlying many assignment tools is the user equilibrium principle: 'All used paths are minimum (generalised) cost paths and all paths that are not minimum cost are not used'. In a deterministic user equilibrium (DUE) assignment it is assumed that passengers behave as if they share a perfect and equal perception of the generalised travel costs to their destination and all choose the cheapest option. In deterministic frequency-based assignment, multiple-routeing across the acceptable paths between OD pairs is achieved via the service frequencies. In the absence of crowding, deterministic frequency-based assignment allocates passengers to paths in proportion to the relevant line frequencies which can produce similar results to an all-or-nothing assignment.

2.7.9 All passengers are still choosing their perceived (rather than objectively calculated) cheapest option, but this may not be the same for all since they do not necessarily agree on what the cheapest option is. Stochastic methods are designed to capture this but are out of scope of this unit. Stochastic methods can be considered for modelling users' choice heterogeneity if the software allows for this. This approach will generally add complexity and increase run times.

2.7.10 The practitioner is advised to review the relevant documentation of the chosen software package in order to understand the software specific assignment options. There is a further discussion on this topic in Appendix B.

2.8 Representing Responses of Public Transport Operators

2.8.1 An issue for forecasting is that, in principle, public transport operators can respond to changes in demand by either changing one or a combination of the frequency, capacity or fares they charge. Such responses are most likely with bus services than services with high vehicle and planning costs such as rail. This requires due consideration in order to appropriately represent public transport services in future years. The practitioner will need to assume how the operator might respond in terms of future services, frequencies or fares and in determining the future public transport network. Further information on the inclusion of responses of public transport operators in forecasting, particularly in the context of rail, is in TAG unit M4 Forecasting and Uncertainty.

3. Generalised Cost Definition

3.1 Introduction

3.1.1 Generalised costs are used in the calculation of the utility of paths as perceived by travellers and therefore in determining the assignment or loading of passenger flows to the paths. It is a combination of a number of different attributes of a path with each attribute being given its own weight or coefficient. The coefficients convert components to common units (usually time in minutes, but sometimes costs in pence) and are chosen to ensure that the relative importance of each component for passengers is reflected. Each component combines to give a generalised cost, of which journey times (weighted by the value of time) are only one component.

3.1.2 The generalised cost is defined as follows. Indicative ranges for the various weights are provided in Table 1 for a typical urban multimodal context. These ranges reflect plausible values[3] and practitioners should justify when they adopt values outside these ranges. Such justification may be based on statistical estimation from specific revealed preference (RP) surveys of local behaviour or may be context-related as noted in the table.

Generalised Cost_{PT} = AccessTime_{o} * access time factor +

BoardingPenalty +

WaitTime_{o} * wait time factor +

$\text{TransferTime}_{interchange}$ * interchange time factor +

$\text{WaitTime}_{interchange}$ * wait time factor +

TransferPenalty +

In-VehicleTime * in-vehicle time factor +

Fare / Value of Time +

Crowding (may be represented as additional wait time or in-vehicle penalty) +

EgressTime_{d} *egress time factor

Where

[3] IHT's Guidelines on Developing Urban Transport Strategies (May 1996) and ITS and John Bates's review of value of time savings in the UK in 2003

- Section 3.2 explains
  - $AccessTime_o$; the time from trip origin to public transport stop
  - $EgressTime_d$; the time from public transport stop to trip destination
  - $TransferTime_{interchange}$; the time between public transport stops
  - $WaitTime_o$; the time spent waiting for first service
  - $WaitTime_{interchange}$; the time spent waiting for subsequent services
- Section 3.3 discusses fares
- Section 3.4 explains
  - BoardingPenalty
  - TransferPenalty
- Section 3.5 discusses In-VehicleTime and service quality
- Section 3.6 discusses crowding

3.1.3 Practitioners should refer to software user documentation because commercial software differs in the extent and way in which these are implemented.

3.1.4 Passengers typically seek to minimise the number of times they interchange, as this incurs additional boarding time, interchange time and wait time, all of which are considered to be less desirable than in-vehicle time. The weights (or multipliers) represent the balance or trade-off passengers make between interchanges, walking, waiting and riding on a vehicle. The calibration process described in section 9 involves adjustment of these generalised cost weights to reflect local routeing behaviour.

3.1.5 Indicative ranges for the various weights provided in Table 1 are for a typical urban multi-modal context. These ranges reflect plausible values (referenced in paragraph 3.1.2) and practitioners should justify when they adopt values outside these ranges. Such justification may be based on statistical estimation from specific revealed preference (RP) surveys of local behaviour or may be context-related as noted in the table.

Table 1 Recommended Generalised Cost Weights and Penalties

Weights / multiplier	Sub-mode	Indicative Ranges	Exceptions, notes
$AccessTime_o$ $EgressTime_d$ $TransferTime_{interchange}$	Walk mode	1.5 – 2.0	car mode access or external zones
$WaitingTime_o$ or $WaitingTime_{interchange}$		1.5 – 2.5	
InVehicleTime	Rail	1	
InVehicleTime	Bus	1 – 1.4	
InVehicleTime	Interurban bus	1 – 1.4	
InVehicleTime	Tram/LRT	0.7 – 1.1	
Penalties	**Sub mode, units**	**Indicative Ranges**	**Exceptions, notes**
BoardingPenalty or TransferPenalty	combining parameters used, minutes per transfer	2 – 10	Larger penalties observed for long distance (rail) journeys

3.2 Access/Egress Time and Waiting Time

Access/egress/transfer time

3.2.1 Access time represents the time spent from the trip origin to the public transport stop and the egress time represent the time spent from the public transport stop to the trip destination. The transfer time represents time spent moving between different services, across different modes or different platforms. It is intrinsically related to the detail with which individual stations and stops are represented. In some cases, to ease network coding, access and egress can be simplified to ‘access’ if no differential treatments are being applied.

3.2.2 In studies where the infrastructural component (such as new or refurbished stations or stops) is not the primary scope, the interchange movements are usually simplified in a single stop and the transfer time incorporated in interchange penalties (a time penalty for moving from one mode to the other or within the same mode). When infrastructure improvements need to be assessed at terminals, or for detailed operational modelling, pedestrian movements need to be assessed within a station. It may be necessary to represent the detailed layout of the terminus or station (explicitly coding individual platforms, entrances, etc) to represent transfer within the terminus.

Waiting time

3.2.3 The simplest assumption for the calculation of the mean wait time is to assume that it is half the headway. This assumes that passengers arrive randomly at the stop and that the service is reliable. This may be a reasonable assumption for services with short headways but for long headways it is more realistic to assume that passengers will try to time their arrival at the stop to minimise waiting time. For this reason, some packages allow the definition of ‘wait curves’. These define the waiting time as a function of headway. Examples of wait curves are shown in Figure 2 below.

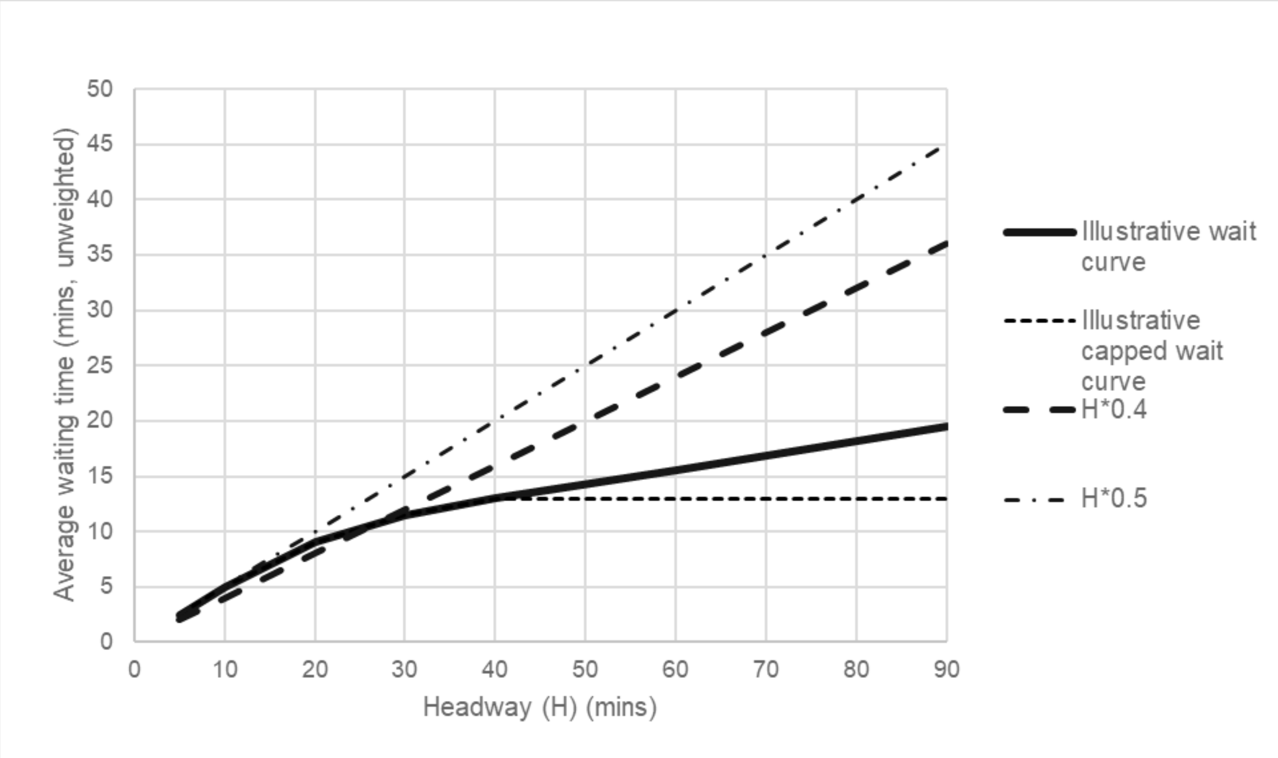

Figure 2 Examples of Wait Curves

3.2.4 The Headway (H) * 0.4 and H*0.5 curves are examples of simplistic representations of the relationship, but this will often be sufficient for the level of assessment required by strategic public transport modelling. This is particularly if a frequent ‘turn up and go’ metro style system is being modelled. For other studies, the adoption of a curve similar to the solid line labelled ‘illustrative wait curve’ may be more suitable whereby wait time increases at a lower rate than headway to reflect that passengers are inclined to time their arrival for infrequent scheduled services. Additional wait curves can be found in the PDFH chapter B4 (membership required). If these are not appropriate for the model, the curves could be informed by local surveys of the relationship between service headway and wait time.

3.2.5 Often it is only appropriate to use wait curves for the first service that is boarded, i.e. the one where the passenger has the most control over the time they arrive at the stop. For subsequent services on the path (i.e. interchanging onto subsequent services) it may still be appropriate to use half the headway, however long that may be – in effect, this assumes random arrivals by services. An exception to this approach might be where a subsequent service being boarded is an infrequent long-distance service for which the departure is known as opposed to a frequent metro style service for which the departure time is not relevant.

3.2.6 When services are irregular (either planned or a result of poor punctuality), half the mean headway may actually be an underestimate of the mean waiting time. In this situation it is worth considering using wait curves where the waiting time is greater than half the headway. Reduced waiting times for a given headway can be used to model the effect of improved reliability and better passenger information, although both effects can be hard to quantify.

3.2.7 When using non-linear wait curves the wait time calculation will usually be a function of the combined frequency of attractive services, not individual services, therefore the waiting time estimate will vary depending on how services are combined when they are coded. This should be carefully considered when defining wait curves.

3.2.8 Different forms of wait curves can be used, including capped wait curves. Capping the wait time represents the likelihood that passengers will arrive closer to the service departure time for services with low headways. This can be reflected through capping the wait curve at a certain point (see Figure 2). The interactions between the wait curve, how the lines are represented and the resulting forecast loadings on competing services that have different headways need to be critically examined before employing a cap. This is important to ensure services with different headways do not have the same utility (all other things being equal).

3.2.9 In timetable-based assignment systems it is possible to calculate intermediate waiting times exactly, due to knowledge of the timetable. For the first wait time, an arrival time must be estimated, or a generic time or boarding penalty may be assumed. In principle this will relate to the characteristics of individuals or journey purpose (e.g. genuine knowledge of the timetable, attitudes towards risk of missing services, etc.), the reliability of services and other factors that are required to be estimated as part of the model calibration process. When modelling low frequency services, and when the skims will feed into a demand model, a timetable-based-based approach or PDFH's rooftop approach may be more appropriate.

3.2.10 The choice of a suitable wait time methodology depends on the project or policy under investigation and the modelling of wait time will be influenced by the modelling approach used:

- When services are frequent (less than 10-15 minute headway) use simple half headway
- When services are less frequent and interchange is important, apply refinements which can approximate a passenger's planned interchange by using wait time curves or node specific boarding penalties
- For the inclusion of optimised, complex interchanges which are designed to minimise wait times, consider using a timetable-based approach.

However, a key indication of the suitability of the wait time methodology applied will be the level of validation achieved for service boardings. Where a suitable fit with observed data has not been achieved, alternative methods can be tested as part of model calibration.

3.3 Fare Structures

3.3.1 Fares only need to be included in the assignment if they influence route choice. Whilst fares are needed for appraisal, they are rarely included in urban public transport assignment models due to prevalence of ‘travelcards’ meaning fares do not affect the service or route chosen. If fares do not significantly influence route choice, then fares can be added to the generalised cost after the assignment and before passing generalised cost matrices to a demand model or appraisal package.

3.3.2 Where fares can influence route choice then it is suggested to include them in the assignment and segment them by value of time (VOT), which in turn implies different user classes by purpose or an average value of time for all purposes. It is accepted that the complexity of some fare systems may prevent them from being represented exactly in the assignment model, but the model representation needs to be ‘acceptable’ its acceptability recorded in associated reporting. Acceptability can be gauged from whether the assignment model validates or not (see section 10).

3.3.3 There are a number of fare schemes that are used in public transport modelling. Fares are converted to time using a value of time parameter in the assignment process. The units of generalised cost and VOT need to be consistent with the units of fare for example if generalised cost is in minutes and fare is in pence then the VOT needs to be in units of pence per minute. Various types of fare structures are listed below.

- Distance-based. This usually consists of a fixed amount (boarding penalty/cost) plus a cost per km:

$$fare = a + b * distance$$

  where a and b are user-defined and need to be calibrated from public transport operators’ fares. A variation of this definition allows b to vary by distance band.

- Stop-to-stop fare. The fare is defined explicitly for each stop-to-stop combination.

- Stage-based**.** For stage based fares, the fare depends on the number of fare stages passed by a trip.

- Zone-based. This is similar to stage-based, but depends on the number of fare zone boundaries crossed (note that in that case model zones should be defined to correspond to fare zones)

Discounted fares

3.3.4 A travel card typically permits unlimited travel within the area and for the time that the card is valid. Typically travel cards offer significant discounts to off-peak travellers or travellers who need to make many trips during a day. For assignment purposes travel card holders can be modelled as not having to pay a fare, or paying a reduced fare. However, a separate model may be required to forecast how many users will choose to purchase the card in the future.

3.3.5 The fare schemes may also be mixed together in some models. In some cases different structures and/or parameters can be used for different modes or sub-modes in the model. Note here that, if traveller types pay different fares (such as concessionaries), or that fare products differentially affect routeing, there may be a need to represent these with different user classes. Also, if fares differ between periods (e.g. off-peak fares) different models need to be constructed for each period. The most appropriate software package will depend on its ability to reflect the fares structure for the system in question.

3.3.6 Whatever levels of segmentation are used, it is important to ensure that the average fare paid across all segments in the model broadly correspond with what is paid in reality. For example, having one segment only that pays the full adult fare will not give an accurate representation of the true fares paid and hence potential operator income.

3.3.7 Total operator revenue needs to be distinguished from fares. It will include not only fares but income from advertising and other commercial activities. Whilst only fare revenues are used in appraisal process, all revenue income is a major consideration for operators when determining the viability of participation in a project. Whilst a public transport model can assist with estimates of total revenue (i.e. number of people that may view advertisements) total revenue is generally calculated externally of a model.

3.4 Boarding Times/Interchange Times

3.4.1 Boarding times, expressed in minutes, (sometimes called boarding penalties), can represent time taken boarding a service at the doors of the vehicle, taking into account demand or crowding. In some software, separate penalties are available for initial and subsequent (transfer) boardings to reflect different perceptions of these boarding components. Overall, it is a parameter broadly representing passenger perceptions of the generalised cost of transfer between services that is calibrated to:

(a) control the numbers of boardings or transfers made per trip (the details are software dependent) and

(b) capture differential passenger perception of the relative attractiveness of the various public transport sub-modes.

3.4.2 Time spent transferring between services is usually perceived to be more onerous than in-vehicle time. If the model does not explicitly represent walk times between platforms, stations and stops, a larger boarding or fixed transfer penalty can be applied.

3.4.3 Boarding penalties may be adjusted to calibrate public transport sub-mode shares. For example, if mode A is preferred to mode B (all other components of generalised cost being equal), the practitioner may reflect this with higher boarding and/or transfer penalties for mode B than for mode A. This should be considered in combination with the in-vehicle time factors for these sub-modes.

3.5 Quality Factors for Stops and In-Vehicles Times

3.5.1 Quality factors, often referred to as “soft measures”, incorporate values for comfort, security, information provision, ease of interchange, etc., into the costs of using public transport[4]. For example, where a bus service is particularly comfortable, with good passenger information systems and facilities which prioritise passenger safety (e.g. CCTV, well-lit interchange facilities, staff presence), passengers may elect to use this service as opposed to a potentially quicker service that is of a lower quality.

3.5.2 This behaviour can be represented in a transport model in the assignment stage or in the mode choice model (see TAG unit M2.1 Variable Demand Modelling). Within the assignment there are a number of ways to incorporate these into generalised cost, for example as variations in boarding penalties, wait times or mode specific values of in-vehicle times that reflect the perceived differences in quality.

3.5.3 The method for deriving IVT factors or mode-specific values should be proportionate to their importance and potential impact on the modelled results. In cases where they are deemed significant (for example where a scheme introduces high quality public information systems), these may be derived from appropriately specified stated preference surveys. Alternatively, manual calibration methods through trial and error may be adopted in cases where there are likely to be more minor impacts on the results.

3.5.4 A potential source for valuations for quality factors associated with rolling stock and station improvements is the Rail Delivery Group’s Passenger Demand Forecasting Handbook (PDFH) v6.0 (membership required). Valuations for bus/tram quality factors are presented in Transport for London’s Business Case Development Manual (a version available on this link).

3.5.5 The TAG Data Book contains values for bus quality factors (expressed in generalised minutes) which can be used as both a starting point and a reference point when reporting.

3.5.6 The following points should be noted when considering these quality values:

- Only overall figures have been presented for audio announcements, climate control and new bus shelters because the segmented bus and car user figures obtained from the models were not significantly different. Overall

[4] See TAG unit A4.1 Social Impact Appraisal for a more complete description of quality factors.

valuations obtained from an unsegmented model are also presented (they are an average of the bus and car user figures).

- The valuations are in generalised minutes. Introduction of a quality measure does not represent a time saving as such, but will increase attractiveness and can therefore be modelled as a reduction in generalised time.

- Valuations may not remain constant over time. The figures have been calculated as additive factors on the basis of respondents‘ current journeys, but as different aspects of these journeys change with time, the quality measure valuations may also change, requiring updated research.

- There is no direct correspondence between the car and bus user categories and segmentation on the basis of car availability. However, it is suggested that the car user figures could be used as a proxy for the car available segment, with the bus user figures used as a proxy for the non-car available segment.

3.5.7 The valuations relate to commuting trips. Valuations for other trip purposes are likely to be different and valuations may be obtained by using the ratio of values of time (VOTs) for other trip purposes and the VOT for commuting. It is recommended that local VOTs are used where these are available, or VOTs from the TAG Data Book otherwise.

3.5.8 When evaluating the impact of introducing a number of measures, quality values can simply be summed. This is because research found little evidence of a “package effect”, i.e. where the introduction of a package of soft measures was valued by travellers differently from the sum of the individual measures. There is some evidence of a “package effect” specifically for rail station facilities (see TAG unit M4 Forecasting and Uncertainty and refer to PDFH – membership required). If this is important to the study, it should be included in the stated preference survey.

3.5.9 It is recommended that where buses are less frequent than every ten minutes, or where there are known problems with punctuality or reliability, the quality values should be reduced to account for this. This is because qualitative surveys did indicate the existence of a “service provision” effect, linking the perceived quality of features to the harder measures of service frequency and reliability.

3.5.10 Where a quality factor is already partially in place but is to be enhanced, the quality values should also be reduced. Additionally, if the scope of the proposed measures is narrower relative to those evaluated during the research, then again, the values should be reduced. These reductions should be in proportion to the relative quality that exists before and after the enhancements are made (relative to the ‘maximum’ values presented) and explained in associated reporting.

3.5.11 Relative valuations attributed to different levels of information provision are also available. These are presented in Table 2 as percentages relative to the bus stop Real Time Passenger Information (RTPI) values presented in the TAG Data Book.

Table 2 Valuation of Information Provision

Attribute	Valuation
Real Time Information in City Centre	83%
Real Time Information at Bus Station	85%
Real Time Information at Bus Stops	100%
SMS Real Time Information (Free)	64%
SMS Real Time Information (10p cost)	31%
SMS Timetable Information (Free)	13%
Web Based Information	29%

3.6 Capacity Constraints and Crowding

3.6.1 Crowding of public transport services as a result of capacity constraint affects passengers' perception of service quality and thus, can influence routeing decisions. Where there is expected to be crowding and it is less likely that operators will respond to demand (such as the case for rail commuter services to major cities) consideration should be given to representing crowding. However, if the dominant response to a change in demand would be the operator adjusting capacity, adding a representation of crowding as a function of demand, without consequent changes to service supply, would worsen the model performance.

3.6.2 There are two different crowding effects:

- The perception of discomfort of crowded conditions on public services. This is usually measured as a crowding factor applied to the in-vehicle time which increases the perceived time and makes the overcrowded services less attractive. For instance, the actual journey time might be 20 minutes but with an overcrowding factor of 1.5 the perceived journey time would become 30 minutes. This would affect all passengers and not just passengers boarding the service. This effect can be included with all standard software.

- The extra time spent at the stop due to the inability to board an overcrowded service. When the first service arriving at the stop does not provide available capacity, users have to wait for another service to arrive, and this will increase the total waiting time. Where software can represent this measure, queueing theory is applied to estimate 'unscheduled' wait time by combining the service headway with the probability that vehicles are too full to board. This is both due to vehicles bunching and due to passengers' inability to board. Only some software includes this, or the practitioner may wish to introduce this through a 'scripted' solution.

3.6.3 The introduction of crowding has significant practical implications for public transport assignment and is generally applied on each modelled service per time period. During assignment, generalised costs will be changed to account for crowding, and a subsequent assignment will be needed to take these new generalised costs into account. Introducing an iterative assignment procedure to achieve convergence impacts on run times and there will be a need to calibrate crowding curves and calculate capacities of service vehicles. For these reasons crowding should only be modelled where it is likely to have a significant effect on traveller behaviour or where one of the objectives of the scheme is associated with crowding. Where crowding is not modelled it is still important to monitor volume to capacity ratios when forecasting to determine whether crowding will become a problem in the future. When crowding is not explicitly modelled, it may be beneficial to manually adjust service in-vehicle weights to reflect future travel conditions using crowding relationships described below.

3.6.4 Modelling crowding adds a requirement to define and code the capacity of services. Perceptions of in-vehicle crowding are affected by whether passengers are seated or standing and by the density of passengers. Accordingly, the following need to be distinguished for different types of vehicles or rolling stock operated:

- the seated capacity of a vehicle, representing number of seats on different vehicle types
- the load distribution factor of a vehicle (the proportion of seats that are occupied before passengers start to stand)
- the ‘crush capacity’ of a vehicle (the total number of standing and seated passengers)

3.6.5 Crowding curves are used to define a factor – or penalty – that reflects the passengers’ perception of in vehicle time in crowded conditions. As illustrated in Figure 3, crowding curves are defined for values of vehicle utilisation. Factors may be capped at an upper limit of 100% (when total number of standing and seated passengers is at ‘crush’ capacity) for stability reasons. The crowding effect starts before a vehicle is full. For most public transport assignment models, it is sufficient to apply separate curves for a limited number of categories, whether by mode or groupings of service types.

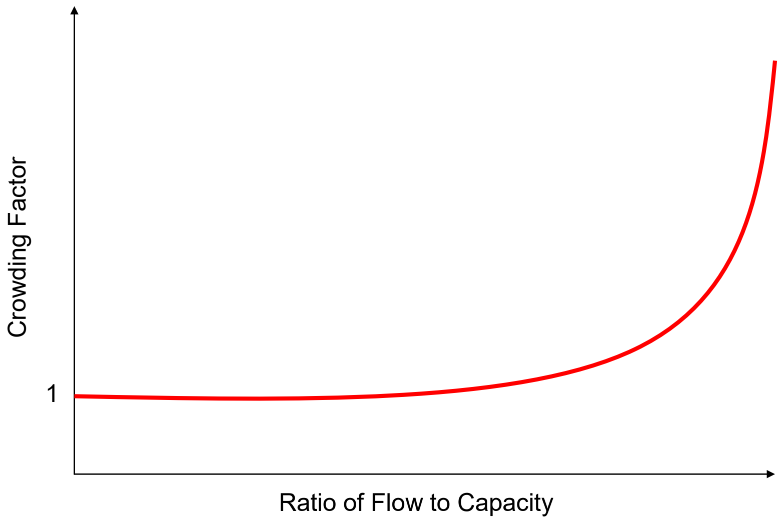

Figure 3 Typical Crowding Curve

3.6.6 Practitioners may consider whether to reduce the input standing area (or capacity) where there is inefficient or uneven use of the standing area. This could be due to passengers' incomplete knowledge about the associated decisions passengers' make about which carriage they choose to use. This reduction in maximum capacity can be referred to as effective capacity and is an input to the modelled services.

3.6.7 The perception of crowding can be defined in different ways depending on the software package or industry research. The underlying assumption is that crowding starts before all seats are taken. If passenger loadings vary materially in different parts of the train (some carriages having empty seats while other carriages are full of many people standing), the practitioner should consider applying an adjustment to represent this.

3.6.8 One approach to adjusting for crowding in the generalised cost equation is to use in-vehicle weights. Table 3 below contains in-vehicle weights for London and the Southeast (LSE) developed from a DfT commissioned study[5] that can be used when train loading exceeds the number of seats. The in-vehicle weights for regional and interurban journeys, and the full study itself, can be requested from the DfT's Rail Modelling Team (railmodelling@dft.gov.uk). Another potential source of information is PDFH (membership required).

3.6.9 The in-vehicle weights should be applied to both seated and standing passengers but depending on study requirements, a weighted average could be applied to all passengers. The weights are applied for each service between each pair of station calls (or line segment) for both seated passengers and standing passengers. The passenger per square metre ('number of standing passengers per metre squared of available standing area') is derived from the assignment step (load divided by standing area) and then looked up in the table to identify the subsequent weights to use in the next assignment iteration until the convergence criterion is satisfied (see section 4.4).

[5] Valuation of Overcrowding on Rail Services, June 2008, MVA Consultancy

Table 3 Suggested In-Vehicle Weights for London and the Southeast (LSE)

Passenger/m²	LSE	
	Sit	**Stand**
0.0	1.00	1.43
0.5	1.05	1.50
1.0	1.09	1.56
1.5	1.14	1.63
2.0	1.18	1.69
2.5	1.23	1.76
3.0	1.27	1.82
3.5	1.32	1.89
4.0	1.36	1.95
4.5	1.41	2.02
5.0	1.45	2.08
6.0	1.54	2.21

3.6.10 ‘Crush capacity’ is viewed differently across the industry. The PDFH refers to around 6 standing passengers per square metre but a service would never realistically be loaded at (or near to) this in the UK. Furthermore, strategic models are modelling "recurrent” crowding (not for example what happens when there is a cancelled service). Therefore, very high standing densities in public transport models are probably indicative that there in an issue with the model that requires investigation and resolution.

3.6.11 The levels of crowding passengers experience varies both by time within the modelled period, between the individual services operated and between the carriages on those (train) services. It may be helpful to apply demand and supply profiles (sourced from surveys) to the crowding calculation. By splitting peak period demand and supply into more refined time intervals and calculating capacity utilisation as a (passenger weighted) average of each time interval, variations in crowding within the time period can better be accounted for.

3.7 Path Building

Introduction

3.7.1 The objective of path identification is to find all potentially attractive paths and calculate their cost. This consists of three stages:

- Identification of least cost paths between specific OD-pairs
- Establishing connectivity between public transport paths
- Selection of acceptable public transport paths

Identifying acceptable paths

3.7.2 There are two common methods for identifying acceptable paths.

Method 1 (frequency-based methods)

3.7.3 In this method the first step is to identify the shortest path excluding the waiting time for the first service from the generalised cost function. This represents the best option if the passenger can time their arrival at the stop/station to avoid any waiting. In practice however the passenger may well have to wait for this 'best path' and it may be possible to get to the destination sooner by taking an earlier service, albeit one which has a slightly longer journey time. In other words, the passenger may choose to take a path with a longer journey time, provided that in return they benefit from a short waiting time, hence all paths with a generalised cost less than the sum of the best path's generalised costs plus headway are acceptable. Any path whose generalised cost excluding wait time is higher than the generalised cost of the best path plus the headway of the best path will not be used, i.e. the passenger will reach the destination sooner by waiting for the service on the best path[6].

3.7.4 The process is illustrated in Figure 4. The best path (excluding the origin wait time) is Path 1. On the other hand, the path with the lowest maximum generalised cost (generalised cost excluding waiting time plus headway (headway=maximum waiting time)) is Path 2. Paths 2-4 all have a generalised cost (excluding the initial wait) that is less than the generalised cost of best path plus the headway of the best path; they are therefore considered acceptable paths, i.e. in some circumstances it would be better to take one of these paths rather than wait for the service on Path 1. Path 5 is not acceptable because the generalised cost is too high – it would always be preferable to wait for the service on Path 2.

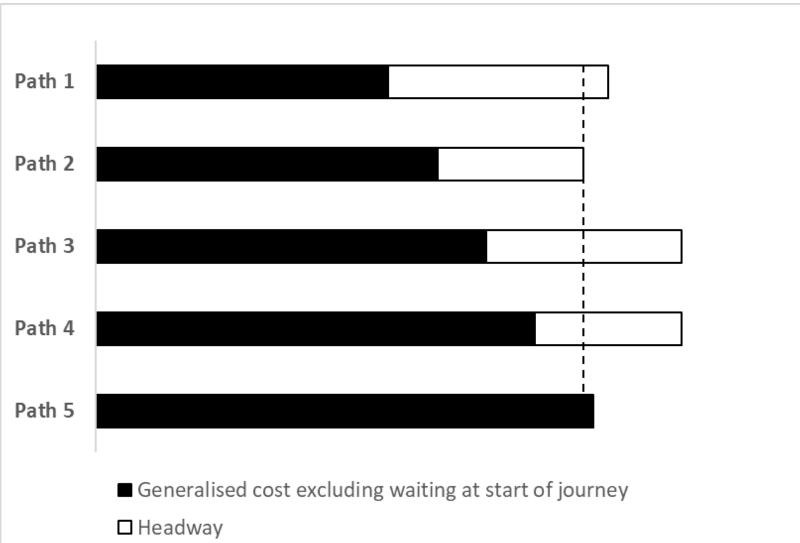

Figure 4 Identification of Acceptable Paths (Method 1)

[6] There is an assumption here that passengers are fully aware of the timetable.

Method 2

3.7.5 Minimum generalised cost paths are identified between OD pairs to establish base path costs. A set of paths between each OD pair is then generated using simple network connectivity rules. These represent sets of links a traveller could use to get from an origin to a destination. These may be constrained to some computationally practical maximum number.

3.7.6 These are then constrained further to a set within the same cost range, i.e. all paths with a cost a certain amount higher (expressed either as an absolute or percentage difference) than the minimum cost path will be discarded. This may also involve limiting number of transfers and interchanges for example (although these may already be built into the generalised cost function) and the total length of walk segments. This path set may be further reduced at the path choice stage, for instance if it is calculated that less than X% of flow uses a path it may be discarded for computational reasons.

Path choice methodology

3.7.7 Ultimately the best test of the adequacy of a particular algorithm is its ability to reproduce observed routeing behaviour. In an all-or-nothing assignment all flow is loaded onto the single minimum cost route for each OD pair. This may be an adequate reflection of reality in some cases. In others, e.g. complex urban networks, there is likely to be observed multi-routing which would require a more complex assignment method to model adequately.

3.7.8 Where multiple paths are identified, some mechanism for allocating flow to each path is required. This is usually as a function of the generalised cost on each path, including all aspects of time (including access and necessary further interchanges), fares and comfort (journey quality). In these cases, path choice is governed by calculations of ‘probability of use’ of each of the acceptable paths between OD pairs. A more detailed explanation of the principles governing path choice methodology is provided in Appendix B.2.

3.7.9 The application of the path choice methodology will differ between software packages. There will also be differences in the treatment of generalised cost elements within the methodology including access/egress costs, differential weightings between first and subsequent service boardings, fares and crowding. Where relevant, the impact of alternative methodologies on run time and ease of undertaking select link or flow bundle analysis should be taken into consideration.

3.7.10 The practitioner is advised to review the relevant documentation of the chosen software package in order to understand the method used to generate the paths.

3.8 Cost Skimming

Introduction

3.8.1 The skimming of costs from assignment is important as skimmed costs are used in demand modelling and in economic appraisal (for instance as input into TUBA). Calculating costs along a particular path is straightforward. However, packages differ in the way that the costs on individual paths are combined to provide a single skimmed cost for each origin-destination pair. It is also likely that not every origin-destination pair is connected by public transport. In these cases, the entire journey may be 'walk only', or deemed impossible. In some software packages the resulting costs will be very high, in others they may be zero. In their particular software package, practitioners need to check the journeys that are not possible by public transport.

Methods

3.8.2 A number of different methods of skimming costs between OD pairs are available, with some packages offering more than one option. It is common practice to use flow-weighted average costs of paths used (as used in TUBA), available as either total generalised cost or for individual components of cost. Other methods are available and may be more appropriate given certain circumstances and the requirements for their subsequent use:

- costs on minimum cost path; usually available as either total generalised cost or for individual components of cost
- straight average over all used paths; usually available as either total generalised cost or for individual components of cost
- frequency-weighted average costs; usually available as either total generalised cost or for individual components of cost
- composite cost (logsum where path choice is based on a logit model); available only for total generalised cost, not individual components

Relationship to variable demand models (VDMs)

3.8.3 Where a variable demand model is used, consistency and convergence between the assignment and the demand model(s) is very important. To optimise processing time and ensure true converged solutions, the travel cost formulations used in both should contain the same ratio of weights of in-vehicle, walking and waiting times, and the same ratio of weights of time and fare.

3.8.4 For input into TUBA, and general appraisal purposes, it is important to be able to skim the individual components of generalised costs separately, particularly travel time and fares. For example, times for business trips will need to be unweighted total OD travel times; for consumer trips they need to be weighted

OD travel times, using the weights for waiting and walking time. Perceptions of journey quality and of crowding should be distinguished as explained in TAG unit A4.1 Social Impact Appraisal.

3.8.5 For input into TUBA and the demand model costs need to be skimmed for each demand segment (e.g. combination of purpose, car availability and/or income). It is worth noting that it is possible to have fewer segments in assignment than in the demand model, provided the demand segments map unambiguously onto the assignment segments. The fundamental principle which should be adhered to is that, whilst cost skims can be extracted at different levels of demand aggregation when used in TUBA and demand modelling, the skims should be generated consistently.

3.8.6 The results of different skimming procedures may lead to rather different results. The method used for skimming costs should be consistent with that used to split flow between routes (see TUBA general guidance and advice). For instance, if a logit model of route choice is used then the skimmed cost should ideally be the logsum measure, although this is not possible in many existing packages and appraisal analysis requires different components of costs rather than a total generalised cost only. The current TUBA recommendation is to use flow-weighted average route costs. For input to TUBA these need to be separated between fares and the different components of time-related costs.

3.8.7 The skims also need to feed the demand model which itself may require skims of individual cost components and apply coefficients that vary by purpose and/or person type. This can be a problem if in the skimming procedure the model aggregates over routes. Any inconsistency here can lead to counter-intuitive results.

3.8.8 Overall, it is recommended that the assignment package's skimming capabilities are assessed before committing to its role in the modelling structure, to ensure that a robust interface between assignment and demand model can be achieved. If different modes are modelled separately, practitioners should be aware of the difficulty of transferring demand, cost skims and highway link speeds) between modelling packages.

3.8.9 A final issue in skimming is the implications of using biased networks (also identified in TAG unit M2.1 Variable Demand Modelling). To enable the skimming of levels of service for each of the public transport sub-modes as input to the choice model, the assignment networks may need to be manipulated or biased to favour one or more of the modes. This may also be done to obtain consistency between the sub-modal split in the demand and assignment models. The preferred approach, where software has the capability, is to avoid the need to do this, with network models producing sufficiently accurate costs to allow a suitably calibrated demand model to reproduce observed sub-mode shares.

3.8.10 If biased networks cannot be avoided, the analyst should ensure that the amount of bias introduced is as little as possible and recorded. A quantitative assessment should be made of the level of bias introduced, and its acceptability. Care should be taken that the biased networks used in application are also used for model estimation, and in future forecasting, the biases should be retained.

3.8.11 TAG unit M2.1 Variable Demand Modelling contains further discussion on the interface between VDMs and public transport assignment models.

Relationship to highway assignment models

3.8.12 Public transport schemes may reduce highway congestion by attracting travellers from car or increase localised road traffic congestion through measures to secure priority for public transport over general traffic. The impacts of public transport schemes on road traffic congestion should therefore be estimated, unless schemes are aimed at serving low car ownership areas (and where modal transfers are therefore likely to be small) and which would not impact adversely on the capacity of the road system for general traffic.

3.8.13 Models for the appraisal of major public transport schemes which are expected to have an impact on road traffic congestion will require a validated highway (road traffic) assignment model. See TAG unit M3.1 Highway Assignment Modelling for details on how to do this. Indeed, there are some cases where highway-based public transport schemes may be appraised satisfactorily using a highway assignment model alone, such as a bus priority scheme.

3.8.14 For sub-modes that run on-street (mainly bus, but some LRT schemes) it is important that journey times in the public transport (PT) assignment model are consistent with the level of traffic congestion. This will require some linkage of on-street public transport mode link times in the public transport network to assigned journey times from a highway assignment model. Especially when modelling forecast scenarios, a link to a highway assignment model will be necessary to estimate public transport on-street journey times for forecast years. The presence of bus or no-car lanes needs to be taken into account in reducing highway capacity.

3.8.15 Similarly, the journey times in the highway assignment should take into account the effect of on-street public transport vehicles on congestion. The presence of segregation, whether for bus or light rail, should be taken into account, as discussed in section 6.

3.8.16 In all cases such as these, the highway assignment model should be sufficiently detailed to model both the road capacity changes required by the public transport scheme and the effects of those changes on road traffic congestion.

3.8.17 The practitioner should investigate how the software package chosen can handle the effects and requirements described in this section. If the highway model is in different software (a common occurrence) then appreciable effort may be needed to exchange data between models.

4. Validation and Convergence Standards

4.1 Introduction

4.1.1 This section provides advice on validation guidelines. It also covers convergence measures and acceptable values in crowding.

4.1.2 Any adjustments to correct errors or improve the accuracy of the model should be regarded as calibration. Validation involves comparing modelled and observed data that is independent from that used in calibration. Using validation metrics to expose and diagnose improvements is a method used to focus calibration.

4.1.3 The criteria presented in this section are intended to be used to assess the suitability of the model for scheme appraisal.

4.1.4 Where models are developed for the assessment of specific interventions, it is natural to pay more attention to validation quality in the vicinity of the interventions versus when a model is developed for a general purpose.

4.1.5 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 4.2. The issues of calibration and validation should be addressed up front in model development. It should be part of the Appraisal Specification Report (see the Guidance for the Technical Project Manager), where the scope of the model and the purpose for which it will be used are agreed.

4.2 Fitness for Purpose

4.2.1 The test of fitness for purpose of a model is to explain the confidence that can be placed on model outputs. The achievement of the validation guidelines specified in this section 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'.

4.2.2 The practitioner should carry out base year tests to give assurance that data have been assembled and processes are functioning. Evidence may also be presented to demonstrate that demand allocation between competing routes, services and/or sub-modes is functioning. For high impact projects as defined in the Uncertainty Toolkit, further evidence of how the model responds is recommended to judge whether the model behaves acceptably and if the outcome is plausible. This can be carried either through the consideration of "backcasting" (seek advice from DfT when this is required) or forecast tests and/or sensitivity testing (see section 10.6).The practitioner should apply their experience to judge whether the nature and scale of the model response is reasonable, and these tests, together with the conclusion, should be reported upon.

4.2.3 For a general-purpose model, it may be useful to carry out demonstration testing so that potential users of the model can gauge the usefulness of the model for particular applications. This may include coding a particular scheme or enhancement into the model and reviewing the appropriateness of the model response.

4.3 Validation Criteria and Guidelines

4.3.1 The differences between modelled and observed data should be quantified and then assessed using the criteria presented in the subsequent paragraphs. The purpose of this assessment is to explain the confidence that can be placed on the model outputs. This should not be interpreted as a target that the model should be constrained to achieve.

4.3.2 In some models, particularly models with high levels of crowding or route choice, it may be difficult to achieve the occupancy or boarding/alighting validation criteria as set out later in Table 5. As explained in TAG unit M2.2 Base Year Demand Matrix Development, 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 practitioner 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 practitioner should explain why the model does not reproduce passenger volumes to these tolerances and should indicate the scale and nature of potential forecasting uncertainty and suitability of the model for its intended purpose.

Trip matrix validation

4.3.3 Validation of the trip matrix should involve comparisons of assigned and counted passengers across complete screenlines, cordons, group of stops or stations (as opposed to individual services). The validation criteria and guidelines for screenline flows are defined in Table 4. Where models do not achieve the guidelines, the practitioner should review the assumptions and quality of data used to develop the trip matrices, as well as the network coding.

Table 4 Trip Matrix Validation Criterion and Guidelines

Criterion	Guideline
Differences between assigned and observed flows should be less than 15% across screenline/boarding/alighting/service groups	All or nearly all screenlines (i.e. 95%)

4.3.4 Section 8.2 explains when matrix estimation techniques may be used. If the decision is taken to apply a process to adjust demand matrices based on differences between counts and comparable assigned prior matrix flows, the advice is set out in TAG unit M3.1 Highway Assignment Modelling. The advice in this unit applies equally to matrix estimation of public transport demand matrices.

4.3.5 Matrix estimation should not be allowed to make significant changes to the prior matrices in order that the validation standards are met and guidance on acceptable scale of changes is given in TAG unit M3.1. Outliers should also be examined, even when the criteria in TAG unit M3.1 are met. The practitioner should include explanations about the relevance of the outliers to the intended uses of the model in the Model Validation Report.

4.3.6 With regards to screenline validation, the following should be noted:

- Screenlines should normally be made up of occupancy or boarding/alighting counts
- assigned flows and counts totalled for each screenline or cordon, as a check of the quality of the trip matrices
- Screenline/boarding/alighting/service groups should be defined at total level (all modes), and at sub-mode level where there are mode-specific matrices. More than one screenline/boarding/alighting/service group should be used.
- The approach taken and data used for matrix validation may draw on or be related to count data. Accordingly, the practitioner should present evidence to distinguish the extent of independence, in particular whether the count data were used in expansion of source data or to constrain trip matrices.
- The comparisons should be presented separately for each modelled period.

Assignment validation

4.3.7 The validation of a public transport assignment model should include comparisons of the assigned flows and counts on individual links or at stops/stations (or groups of them) as a check of the quality of the assignment:

- grouped passenger counts across screenlines and cordons, which should be defined to distinguish allocation between public transport modes or corridors, and sometimes at the level of individual bus or train services
- single passenger counts on stops, links, or services in urban centres

4.3.8 The validation measures, criteria and acceptability guidelines, for these comparisons are detailed in Table 5 below. The phrase ‘passenger counts’ refers to either boarding/alighting or onboard counts. The data sources (such as count data, ETM data, Lennon data) available for validation is discussed in section 5 and in TAG unit M1.2 Data Sources and Surveys.

4.3.9 The validation metrics need to be sufficiently granular for the intended purpose. In some cases, for specific schemes, comparisons by Train Operating Company (TOC) or groups of similar services might be needed, and in other cases they should at least be defined by mode. Another example is if the intent is to relocate bus stops, then the metrics should be reviewed at an individual stop level. Generally, more than one screenline or group of passenger counts should be used. The comparisons should be presented separately for each modelled period.

Table 5 Assignment Validation Criteria and Guidelines

Data aggregation	Observation used to validate	Criteria (difference between modelled and observed flows)	Guideline
Grouped services by corridors	• Onboard counts • Entry & exit counts	< 15%	All or nearly all (95%) services
Grouped stops/stations by fare stages or corridors	• Boarding & alighting • Stop/station entry & exit counts	< 15%	All or nearly all (95%) stops/stations
Individual service/link (total passengers)	• Boarding & alighting • Onboard counts	< 25% (except where observed flows < 150 passengers per hour).	All or nearly all (95%) services
Individual stops/stations (total passengers)	• Boarding & alighting • Stop/station entry & exit counts	< 25% (except where observed flows < 150 passengers per hour).	All or nearly all (95%) stops/stations

4.3.10 If counts are used for matrix development or calibration, then an explanation of how they were used previously in processing demand should be provided. The practitioner should present evidence to distinguish the extent of independence, in particular whether the count data were used in expansion of source data or to constrain trip matrices.

4.3.11 When counts are of less than 150 passengers per hour (i.e. low levels of overall demand), comparisons should still be reported but the criteria above should not be applied. Instead, reporting should explain differences and include histograms, scatterplots, or tables for the relevant counts. Histograms can be plotted on a map to show the geographical variation.

4.3.12 In all cases, together with the trip matrix validation, additional checks should be performed to assess the trip length distribution by mode.

Service frequency and journey time validation

4.3.13 It is common practice to process online timetables into a format required by the modelling software to code services. Verification tests of such processes should be clearly reported in the Model Validation Report (see section 11), demonstrating the accuracy of this processing by comparing, separately for each modelled period, source and processed representation of:

- service frequencies (or, for timetable assignment, reproduction of individual services)
- journey times for individual services
- (where services have multiple stops) service segments or journey time profiles and
- Passenger seated and standing capacities of services, if crowding is being modelled

4.3.14 Processing source data may involve a degree of interpretation and simplification which should be explained to justify differences. Where methods are automated for large network models with many services it is acceptable for reports to set out comparisons for a sample of routes in verification tables.

4.3.15 Where manual coding methods are adopted the same high standard of coding accuracy is expected. Timetable differences of a minute of more may have noticeable impacts on routing and would need justification. Small differences in frequencies or headways may be expected to be less critical. Practitioners should not, for these verification tests, compare different data sources, such as timetable versus vehicle tracking data. Such comparisons will unearth the accuracy of the operator's timetable versus observed journey times but provide no insight on accuracy of the coding of the services in the model.

4.3.16 Where highway congestion feedback is incorporated, it should be applied incrementally as a forecasting adjustment, rather than applied in absolute form in base year network coding. The relationship will normally be calibrated comparing data from the validated base year highway network with the public transport service journey times. Table 6 sets out guidelines practitioners may use for validating this relationship.

Table 6 Journey Time Validation Criteria and Guidelines (linked to highway congestion)

Criterion	Guideline
Modelled journey times along routes should be within 25% of timetable times	>85% of routes

Annual Patronage

4.3.17 Wherever possible, a check should be made between the annual patronage derived from the model and annual patronage obtained from the operators (TAG unit M1.2 Data Sources and Surveys provides advice on annualisation). Differences in total patronage should be as small as possible (for instance less than 5% for services with significant patronage). Greater differences need to be explained, including how model results might be used to predict future patronage and revenue. Practitioners should note and make adjustment for definitional inconsistencies, for example, operator patronage may be boardings and not trips. Boarding information will contain the start location but not the end of the journey, whereas trip data contains start to end information for a journey.

4.4 Convergence Measures and Acceptable Values

4.4.1 Convergence relates to the stability of a model that is solved iteratively, such as when modelling crowding which can cause route choice to change. Before the results of a public transport assignment model are used to influence decisions, the stability (or degree of convergence) of the assignment must be confirmed at the appropriate level. Instability may occur where either stochastic assignment methods are adopted or when crowding is represented. This section only applies in these circumstances; it is then important to demonstrate that the whole public transport assignment model converges to a satisfactory degree.

4.4.2 The required level of convergence necessary for appraisal purposes reflects the scale of the transport user benefits of the scheme being appraised, relative to the network size. For instance, the calculation of benefits from small schemes in large networks will be much more sensitive to convergence than large schemes in small networks. However, the level of convergence required for stable integration with a demand model is not related to the size of scheme but the interactions between the assignment and demand models.

4.4.3 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 unique solution.

4.4.4 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 where % gap is:

$$\frac{\sum_a D_a |C_a^n - C'^{n-1}_a|}{\sum_a D_a C'^{n-1}_a}$$

Where:

- C'^{n-1}_a is cell a of the generalised cost matrix, averaged across previous iterations
- D_a is cell a of the assignment input demand matrix

- C_a^n is cell a of the generalised cost matrix from assigning D_a
- a represents every combination of origin, destination, demand segment/user class, time period and mode

4.4.5 While some software might not have % gap measure included in the assignment process, its value can be calculated using the costs calculated in each iteration and later weighted by demand as per the formulation above (see more on model convergence in Appendix C).

4.4.6 The level of convergence should also be monitored and reported in relation to four other measures aimed at quantifying the degree of switching between competing routes and services:

- Passenger kms per service: less than 1% for 95% of the services

Percentage differences between previous and current iteration should be minimal for the following measures:

- Passenger boardings per service
- Passenger perceived hours per service
- Passenger perceived wait time per service (when crowding methods are applied at the stop)

5. Calibration and Validation Data

5.1 Introduction

5.1.1 This section contains guidance regarding the use of specific public transport data for the calibration and validation of models.

5.1.2 For guidance on the availability of public transport data and the conduct of surveys, see TAG unit M1.2 Data Sources and Surveys. For guidance on their use in matrix development refer to TAG unit M2.2 Base Year Demand Matrix Development.

5.1.3 Calibration and validation data are generally boarding/alighting counts, on-board counts, reported patronage and revenue by operator or mode and ticketing data. There may be issues around commercial sensitivity on some of these data sources (for example revenue data), and some of them may not be available.

5.2 Data for Model Calibration and Validation

5.2.1 Bias, measurement error, and day to day variability will affect the quality of passenger counts. The practitioner should consider the data quality and take steps to ensure that the data can be relied on for its intended purpose. Further guidance is provided in TAG unit M1.2 Data Sources and Surveys.

5.2.2 Counts are best undertaken as passengers board (or alight from) services or enter (or exit) stations. Such counts can be used directly and can be used to estimate the number of passengers on the service as it progresses, known as the demand profile. Direct counts of passengers on a service at a given point can be undertaken but tend to become less reliable on busy and crowded services. Some trains are equipped with technology to count passengers, such as through axle loading weight. The use of cameras to establish public transport passenger counts requires calibration of the video technologies. TAG unit M1.2 Data Sources and Surveys explains the merits of different survey methods.

5.2.3 A programme of public transport passenger counts will require a substantial investment, particularly as counts will be required across several days. Where ticketing data are available, it may be more appropriate to process ticketing data directly. TAG units M1.2 and M2.2 in the Guidance for the Modelling Practitioner explain the potential biases and limitations of ticketing data (for example fare evasion, season tickets, travel cards, overbuying and split ticketing), and how these may be addressed.

5.2.4 Where the practitioner chooses to use ticketing data to derive counts, especially where ticketing data are also used to develop demand matrices, it is important to ensure that the processing methods to estimate counts is independent from other uses. Practitioners should also arrange a limited (small) sample of counts independently to verify the data and processing assumptions.

5.2.5 The nature of count data required will be dictated by the purpose and scale of the model. Large towns and cities are often served by a predominantly radial pattern of services from the centre. In such a model it is likely that the following counts would be required:

- boarding / alighting within the centre, by mode and possibly operator
- at major termini
- for a cordon around the centre, by mode
- at one or more additional cordons, depending on city size and model purpose

5.2.6 Where the model is focused on a corridor (such as a particular rail service or new Light Rail service); the profile of demand along the corridor will be of importance. A number of screenlines intercepting the corridor should be defined, proportionate with the length of operation and nature of the intervention being considered.

5.3 Journey Time Data

5.3.1 As explained in section 4.3, verification of journey time is to ensure coding accuracy. It is not a test of model performance and no 'validation' is therefore expected.

5.3.2 Nevertheless, it may be helpful to present a comparison between timetable and (GPS) vehicle tracking data. This comparison would help practitioners interpret service reliability, how this in turn affects wait times and hence inform adjustments that may be required to assignment parameters better to reflect routeing behaviour. This is particularly relevant when the scheme under consideration may be expected to affect service reliability.

6. Coding and Calibration of Network and Services

6.1 Introduction

6.1.1 This section provides guidance on the coding and calibration of the public transport assignment model network, building on the guidance provided in section 2. It covers the following details:

- Network data sources and coding
- Public transport service data sources and coding
- Pre-calibration checks

6.2 Network Data Sources and Coding

6.2.1 This section covers the base network data required for a public transport model. It comprises of nodes and links that form the network structure, zones (centroids), centroid connectors (connecting centroids to the network) and walk and interchange links.

6.2.2 A node and link numbering protocol is useful for coding road type, local authority, and node type (e.g. road junction, bus stop etc). This will pay dividends in simplifying network calculations, assignment processes and filtering data in reporting.

6.2.3 Public transport networks typically consist of road and non-road networks. The road network would typically be used for buses and the non-road network typically used for rail, light rail, tram and segregated bus lanes. The coverage and detail will depend on the purpose of the model as discussed in section 2.

Where rail station-to-station modelling is undertaken, there may be no need for a road-based network. The following sections provide context to how each type of network should be coded and key considerations for the practitioner.

Road-based Networks

6.2.4 In most practical applications of public transport assignment modelling a road network of some form is needed. This is to overlay bus services on, so that the system may be viewed and edited as it exists (running along actual streets). It is also needed as the core of a walk network for public transport access and interchange. Where bus speed forecasts are based on car speeds from a linked highway model it is usual to have a common network with the highway model.

6.2.5 A common node and link structure will enable easy transfer of bus vehicle volumes from the public transport model to highway model, and congested car speeds from a highway model to public transport model. When developing the road network, consider if the public transport model will use intersection nodes to approximate bus stopping locations, or if buses will stop at specific bus stop nodes. If the latter, decide whether to include bus stops in the highway model network (so that the network can be exported to the public transport model with bus stops already included). There is often a decision point for the practitioner around the level of detail required. For example, whether or not to include explicit detail of bus routes running through local estates that are not represented in the highway model. This level of detail can either be added to better represent local access or be included but simplified.

6.2.6 It is common practice to import all links from the highway model, but to reduce detail in the public transport model. Irrelevant nodes and links could be removed (bearing in mind the walk access needs).

6.2.7 If there is no linked highway model, the road network can be built from digital GIS sources, simplified to relevant links to make it easier to manage and to avoid excessive detail (see TAG unit M1.2 Data Sources and Surveys for data sources).

Non-road-based networks

6.2.8 In developing a network, the nodes and links needed to carry segregated public transport services (for example rail or bus-only links) must be coded to allow access to public transport services and any available interchange between them. The structure of the network can be derived from track layouts in GIS, large-scale OS maps, or OpenStreetmap, depending on level of detail required.

6.2.9 Road network data sourced from a highway network or GIS data will need additions and edits:

- **Bus stops.** Bus stopping locations could be approximated by using intersection nodes (discussed in paragraph 6.2.5). However, if bus services come from a digital source, every individual bus stop will be identified. For strategic modelling this can be too much detail, making it hard to maintain the coding and impacting potential routeing and interchange choices. To address this, stops that together form a logical group can be combined. Bus stops could be combined where multiple bus stops exist along a specific stretch of road representing different service groups because in practice they will operate as a single stop with the potential to interchange. Clustering analysis may help to reveal other potential groupings of bus stops.

- Rail, metro, light rail, major bus or tram **stops and stations** (see section 6.3).

- **Links** for rail, metro, light rail, ferry, and bus-only links.

- **Walking links** within, major bus and tram interchange stations as well as other walk access/egress links will also need including. The level of detail for stops and stations needs to be considered. In some cases, it is appropriate to represent a whole station with one node. In others, the station can be disaggregated with multiple links and nodes representing station entry/exits points and different platforms for different services or directions. The approach taken should be governed by:

    - the purpose of the model,

    - the outputs needed,

    - the computation requirements,

    - observed data available to calibrate.

6.2.10 If there is detailed representation of the station, the vertical and horizontal distances and speed (travelators, lifts, etc.) should be coded appropriately. Signage and wayfinding could also be considered in the coding of perceived walk times. It is important to ensure that interchange volumes between different areas of the station do not mix (for easy reporting of transfers). 'Triangle effects' should be avoided whereby it is quicker to go from A - B via C so that the demand representation and interchange time are correct within the station.

6.2.11 Figure 5 below illustrates a disaggregated node/station where each platform is represented by a node with links connecting to other platform nodes and connected to the walk network at the main station entrance (shown in blue). Figure 6 shows a further example where two lines connect at a station and interchange connections are included (shown in green dashes).

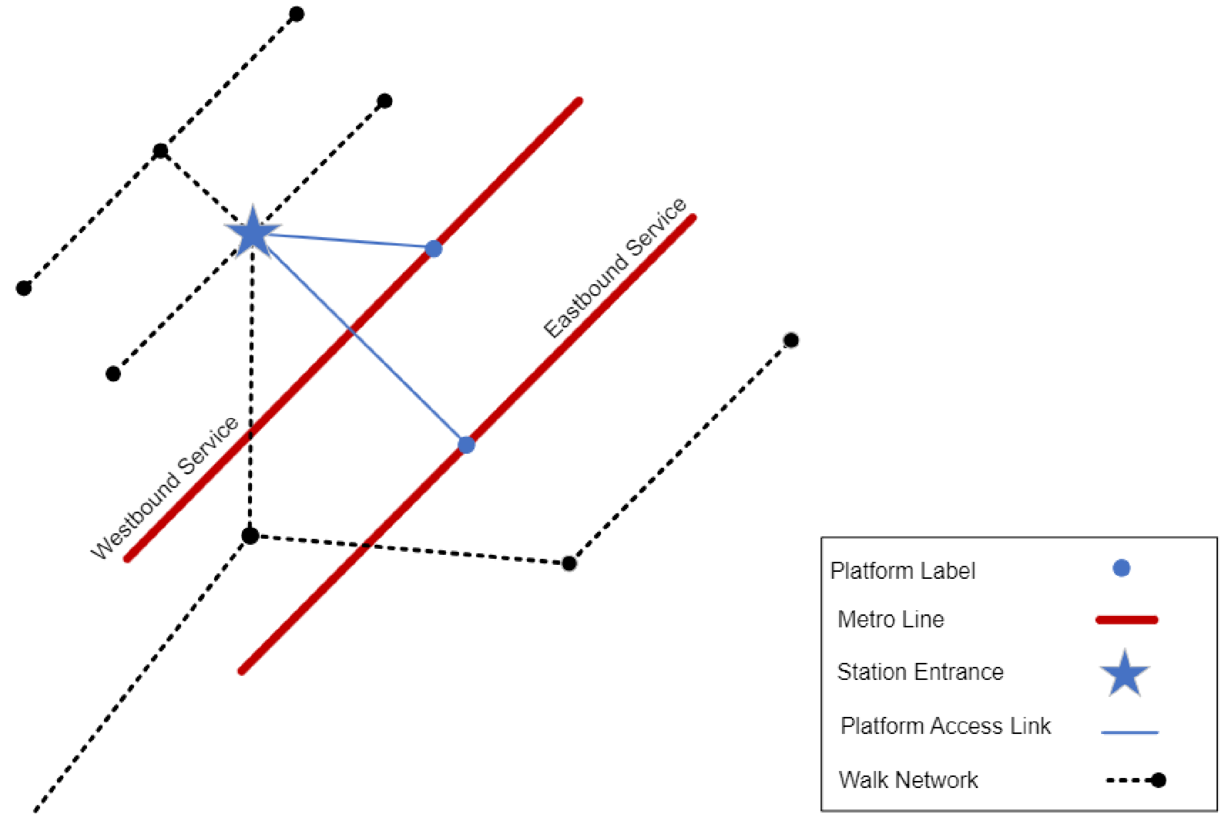

Figure 5 Example of Simple 'Disaggregated' Station Coding

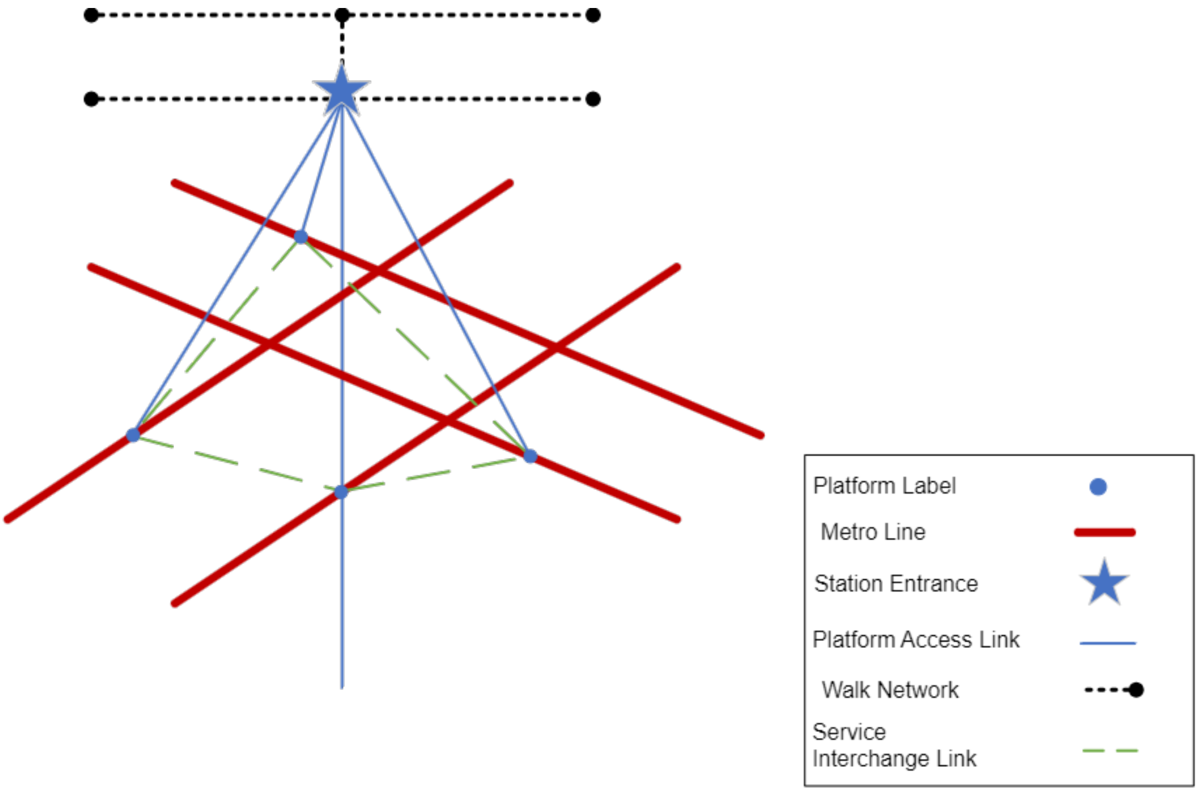

Figure 6 Example of Complex 'Disaggregated' Station Coding

6.2.12 The level of detail in the network should be considered at the outset and a decision taken whether detailed or simplified coding is most appropriate regarding the study purpose and computational limitations. It is important to develop a coding manual or guidance note to assist practitioners in applying consistent principles across the base and future networks. Coding manuals and user guides have been developed by other organisations, such as Transport for London (TfL) and Ireland's National Transport Authority in their model development. These represent examples of best practice and contain examples of simple and complex station coding. Whilst these are not publicly available, they may be made available on request to the organisations.

6.2.13 Extra walk links may need to be added to access public transport. These include:

- the reverse direction of one-way streets
- pedestrianised links
- footpaths and footbridges
- key through-routes via malls, parks, open space
- interchange links for example between bus stops, rail stations or platforms (if the coded street network does not already perform this role)
- external area network reflective of a skeleton / schematic external public transport representation

6.2.14 Walk should be disallowed on motorways and other roads without a footway, unless these are used in practice.

Centroid Connectors

6.2.15 Public transport journeys comprise a sequence of 'legs'. The simplest case comprises three legs: first is a walk from the journey origin (model zone) to a bus stop or station, secondly a ride on a vehicle, and after alighting, a walk to the journey destination (model zone). Centroid connectors provide access to the modelled network and their representation will depend on the level of detail represented. Section 2.3 contains guidance on what to consider when designing the zoning system.

6.2.16 In most public transport models, there will be an explicit road/walk network. The practitioner needs to create connections between the zones and the road/walk network. In this case the access / egress leg of the journey will be a combination of the centroid connector together with the network links between the centroid connector and the stop or station. The form of connection will vary where the model is more or less detailed; reflecting the part of the journey that is not explicitly represented by the network.

6.2.17 In dense urban networks, connections with zones should be made using a walk network. Connections from zones should be undertaken systematically identifying the population or employment centre and loading to an appropriate part of the walk network, allowing access to nearby bus or tram/LRT stops. Adding centroid connectors to specific stations/interchanges directly should be avoided as this could lead to a bias towards a particular mode or service. In the external network it is appropriate to connect directly to stations.

6.2.18 If centroid connectors from a highway model are used, they will need to be reviewed and updated for applicability in public transport models. Alternatively, public transport centroid connectors can be coded from scratch.

6.2.19 Regardless of the approach, centroid connectors must retain consistency between base and future models (or between Do Minimum and Do Something). This is important if the practitioner relies on the software to control the centroid connector set, as this could cause issues with future accessibility and route choice. There are some limited situations (e.g. the quality of access such as cycle paths) where zone connector characteristics may change between base and future scenarios. If there are any zones for future year developments, they should also be included in the base model.

6.2.20 The access mode in built-up areas is generally assumed to be walking. This is coded typically at 4-5 km/h (based on a variety of studies[7] of average walk speeds) but can be adjusted to reflect local data. In less built-up areas where the network coding may be simplified or where zones are large, the access mode may represent a mix of car, taxi, cycling, and other modes including uncoded public transport services. An average speed representing a notional mix of access modes can be applied.

6.2.21 This concept is shown in Figure 7 which shows an illustration of time to access/egress public transport stops increasing non-linearly with distance. This reflects how some public transport stops (for example in rural areas) are accessed from longer distances and that access time is a function of motorised or cycling speeds.

6.2.22 It is recommended to avoid connecting zones directly to stops or stations. The time/cost on centroid connectors should reflect the true distances and times so that the correct costs are captured in skims for demand modelling.

6.2.23 The number of access/egress/transfer options provided for each OD pair is important in the calibration process and representation of user behaviours. This can be governed by the software if there is a well-represented walk network linking various stops and stations. Checks should be made that where there are multiple connectors for a zone, there is a true 'spread' of demand across the connectors. Usually, each zone has several potential public transport options

[7] Various sources quote 1.2m/s usually referenced to 2009 DfT guidance on walking speeds for signal design:
https://content.tfl.gov.uk/traffic-modelling-guidelines.pdf
https://www.livingstreets.org.uk/media/1796/review-of-pedestrian-walking-speeds-report-v4b-280814-docx-2.pdf

which should be reflected in the coding (for example more than one accessible stop per zone).

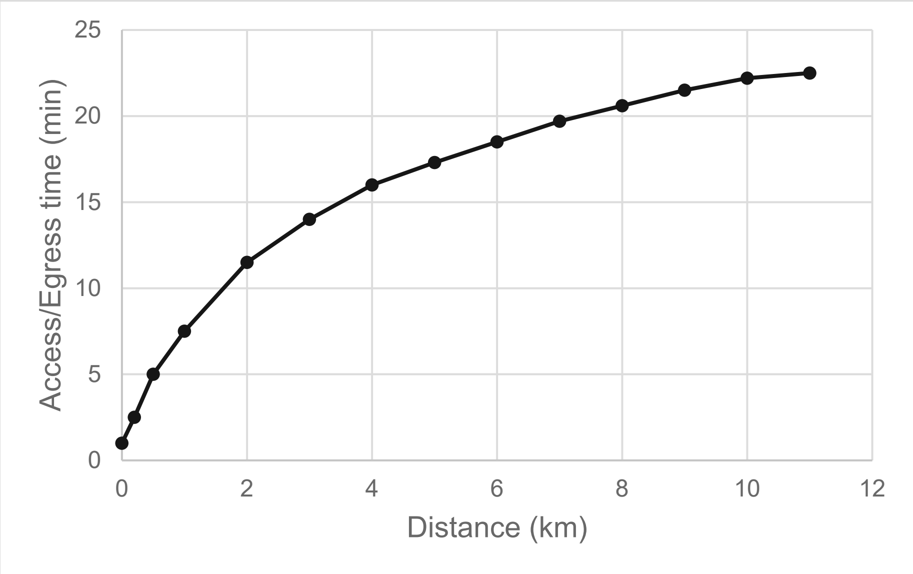

Figure 7 Illustration of Access/Egress Time as a Function of Distance

6.3 Public Transport Services Data Sources and Coding

6.3.1 TAG unit M1.2 Data Sources and Surveys summarises available data sources for public transport services:

- the National Public Transport Data Repository ATCO-CIF
- BusOpenDataService (BODS)
- Traveline National DataSet (TNDS)
- Rail timetable data can be obtained from the Rail Delivery Group (RDG) website
- Automatic Vehicle Location (AVL) data and GPS data available from operators.
- Some authorities and Passenger Transport Executives (PTEs) maintain their own bus service databases.

6.3.2 For proposed light rail services, the network structure can be obtained by scaling from engineering layout plans. Details of the services, including their frequency, stops served and fares, can be postulated by the system design team. Of considerable importance, though, will be the operating speed in order to estimate segment journey times, and rolling stock assumptions indicating capacities. This should be available from operational modelling undertaken by the system design team. For street-running sections of light rail systems, it is important to reflect the interactions between road traffic and the light rail vehicles in the coding.

6.3.3 The subsequent sections cover details on how to approach coding public transport services, how to define services, considerations when mapping nodes and coding travel time between stops.

Approaches to Coding

6.3.4 Data obtained from open data sources need to be processed to create the input in the right form for a public transport model. Each software has specific requirements and data input formats. GTFS (General Transit Feed Specification), is an open data format for public transportation schedules and associated geographic information. Most software package can import GTFS data, although not all of the data sources referenced in paragraph 6.3.1 are in GTFS. Practitioners should check the allowable formats in their chosen software, and scripting may be required for successful import of some data sources.

6.3.5 The following information is a typical minimum requirement:

- mode and operator
- route/description including origin, destination and stopping patterns
- headway for frequency modelling or departure and arrival time for timetable modelling
- nodes forming the route and travel times

6.3.6 Other data items may be added, particularly when modelling fares and crowding. For example, mode and operator information might be required for modelling fares and vehicle type information required for crowding.

6.3.7 Travel times for non-road services can be hard coded as segment times (based on the timetable) in the input. Road based services (i.e. buses) should either be hard coded or if incorporating highway congestion, applied as an incremental change in forecasting (not an absolute change in the base network).

6.3.8 Depending on the software being used, it can be advantageous to use different look up codes for operators or service groups to aid filtering during model use. Practitioners should consult the relevant software guidance to ensure coding styles provide adequate flexibility during model application.

Service Definition

6.3.9 Once the modelled time periods have been determined (covered in section 2.5), a criterion needs to be applied for inclusion of services in each time period. A logical approach is to include services which are timetabled to serve a defined geographical location within the time period concerned. This could be a central urban station or cordon within the Internal Area. The adopted approach should depend on the requirements of each study and in particular the characteristics of peak demand and supply provision. **A consistent approach should be adopted** to determine the services to include, with consideration given to services that either only start or end within the time period but also ~~or~~ services that arrive or depart just before the hour (for example a service departing at 0659 for a 0700 to 0959 model).

6.3.10 The standard approach for a frequency-based model is to identify recurrent patterns in services such as common origins, destinations and stopping patterns and to group the recurrent services as a single coded entity, with an average headway. All the (individual) services represented by this entity are similar enough to be grouped for strategic modelling purposes and are referred to as a “transit line”. Some judgement is required in deciding the best approach to this. Grouped services should generally share the same operator, similar stopping patterns and have the same destinations. The practitioner should explain the process and criteria used for grouping services, explaining which approach is taken and why and carry out checks to ensure the capture of services is representative.

6.3.11 Where sub-mode choice is determined within the assignment model, typically all public transport services are modelled within the Internal Area as they provide either an alternative (competing) or complementary (integrating) route to the main scheme being assessed and as such would be material to the assessment of the main scheme.

6.3.12 In **timetable-based** assignment modelling a definition of every individual service is required. This aligns with the open data sources where individual services are encoded. However, such models are relatively unusual, and many PT assignment models are typically **frequency-based**.

6.3.13 A model can be created that includes all route variants or even all individual services as separate transit lines, but this should generally be avoided. For example, if there is a bus route with several variants, it is recommended to code just one transit line for the bus route based on the most common route variant. This is because coding all variants is inefficient and will result in a slow assignment. In assignments where demand apportionment depends on service frequency and journey time, the demand apportioned to service variants with slightly different in-vehicle times may not be as expected. If every individual service, or variant, is coded in the base model, then future models should be coded in this way too. This can be onerous, requires a level of detail about the future not often available, and therefore requires more assumptions.

6.3.14 For low frequency routes, there may be time periods where there is no service, resulting in unassignable trips or unreasonably high skim costs being passed to a demand model. The practitioner needs to decide how to deal with this (see paragraph 3.8.2). This may involve copying a service from another time period, adding a 'dummy' service, but preferably the journey should be allowed by access via the 'non-transit' (e.g. walk) network. This is also another reason centroid connectors should not be connected to bus stops and stations directly, but to the walk network instead (see paragraph 6.2.17).

6.3.15 Representing train services that split and join is not straightforward to code. Most modelling software platforms have limited functionality to code this and the practitioner usually needs to develop a method to deal with this. Two such approaches are:

- coding in 'dummy nodes' where services split/join which do not represent real platforms, with a near zero boarding penalty, with one of the service branches represented in full and an additional (partial) service coded to represent the missing part with an interchange where the services split. For consistency, these nodes will need to be used by all services not just split/join ones.

- separating the split/join services into two full length services. The 'split' sections often have lower demand, and in this approach it may be most appropriate to represent each with half of the frequency of the combined service, more accurately to represent the level of service on the busier section.

The limitations of the approaches need to be acknowledged (for example when using 'dummy nodes' the boardings skim matrix will show the incorrect number of boardings).

Mapping of stops and nodes

6.3.16 When coding railway station nodes, the nodes in the RDG data are stations which should correspond with the coded public transport model base network.

6.3.17 When coding bus stops, the bus data in TNDS is coded to National Public Transport Access Nodes (NaPTAN) bus stops. Practitioners should ensure that the location and completeness of stops coded from data sources is checked in key areas of the model. This will help to ensure that any process to allocate stops to nodes will work appropriately. Depending on the approach to coding the network (discussed in paragraph 456.2.5) NaPTAN may not align with the node system used in the public transport base network. A method is needed to map NaPTAN nodes to the modelled public transport network nodes and fill any gaps. Common approaches are:

- Using spatial joins in GIS to 'snap' the NaPTAN stop sequence onto the public transport network

- Using a look-up / dictionary to recode NaPTAN nodes to public transport network nodes

- Using public transport assignment modelling software to interpolate for missing nodes (if any) for example on a bus route

6.3.18 Whichever process is adopted, considerable time may be taken up checking and correcting routeings in the model. A common issue that arises when using spatial joins to snap bus stops on to a road network is incorrect snapping. When the road network nodes are dense, the spatial join can snap the bus stop to the incorrect location. This can result in the modelled bus routes making long detours and/or U-turns on the network. The practitioner should undertake thorough checks of the networks generated or infilled with GIS processes, including visual checks. These checks can include assuming standard vehicle speeds over distance versus timetable time to pick out unduly fast or slow sections.

6.3.19 An example of the type of issue that can arise is illustrated in Figure 8. Here the north/eastbound bus stop on the side road is incorrectly snapped to a node on the westbound road (marked by a black square) resulting in the bus turning left and towards the main road, leading to an unrealistic loop for the bus route. The matching of node to bus stop needs to be corrected to fix the bus route, as shown on the right-hand image where the route turns right and proceeds eastbound as expected.

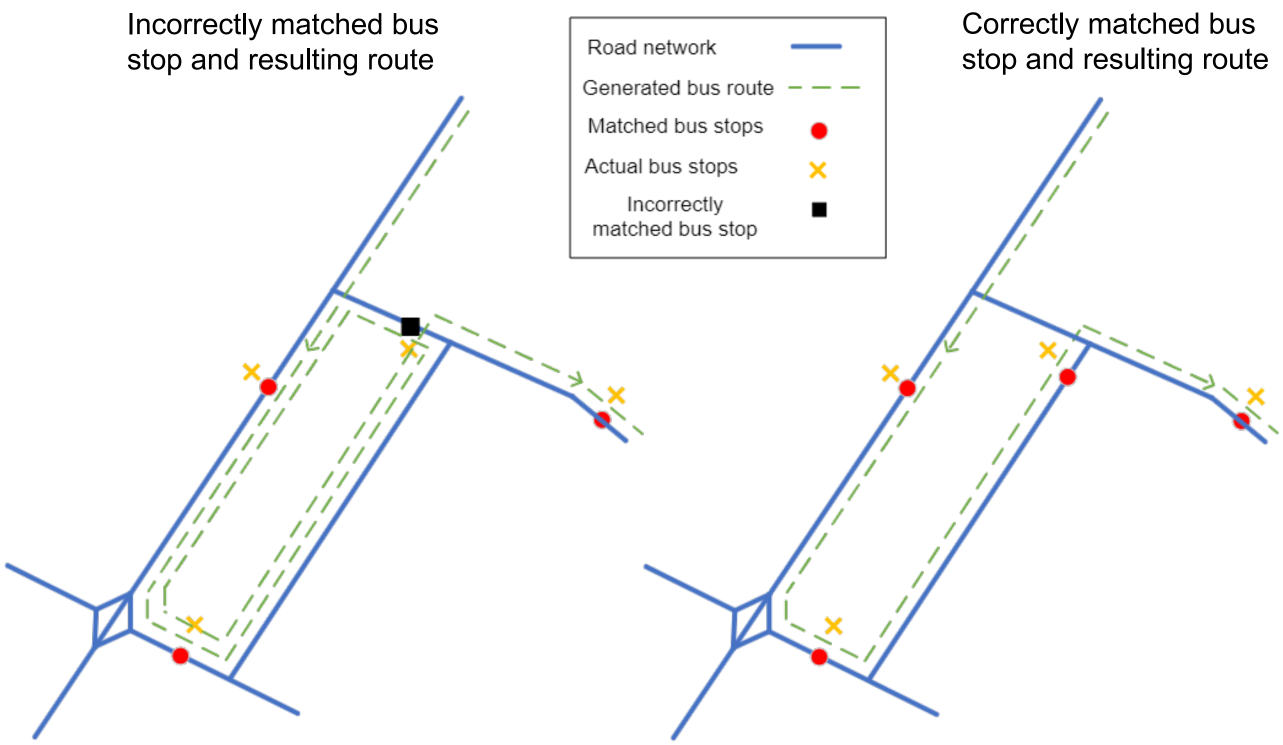

Figure 8 Examples of Matched Bus Stop and the Resulting Bus Route

6.3.20 Another common issue arises when algorithms use quickest path to connect two successive stops and generate the associated bus route. Buses do not always take the quickest path because of banned streets/turns or requirements for U-turns on narrow streets. Figure 9 illustrates an example of the quickest path connection not being correct for a bus route.

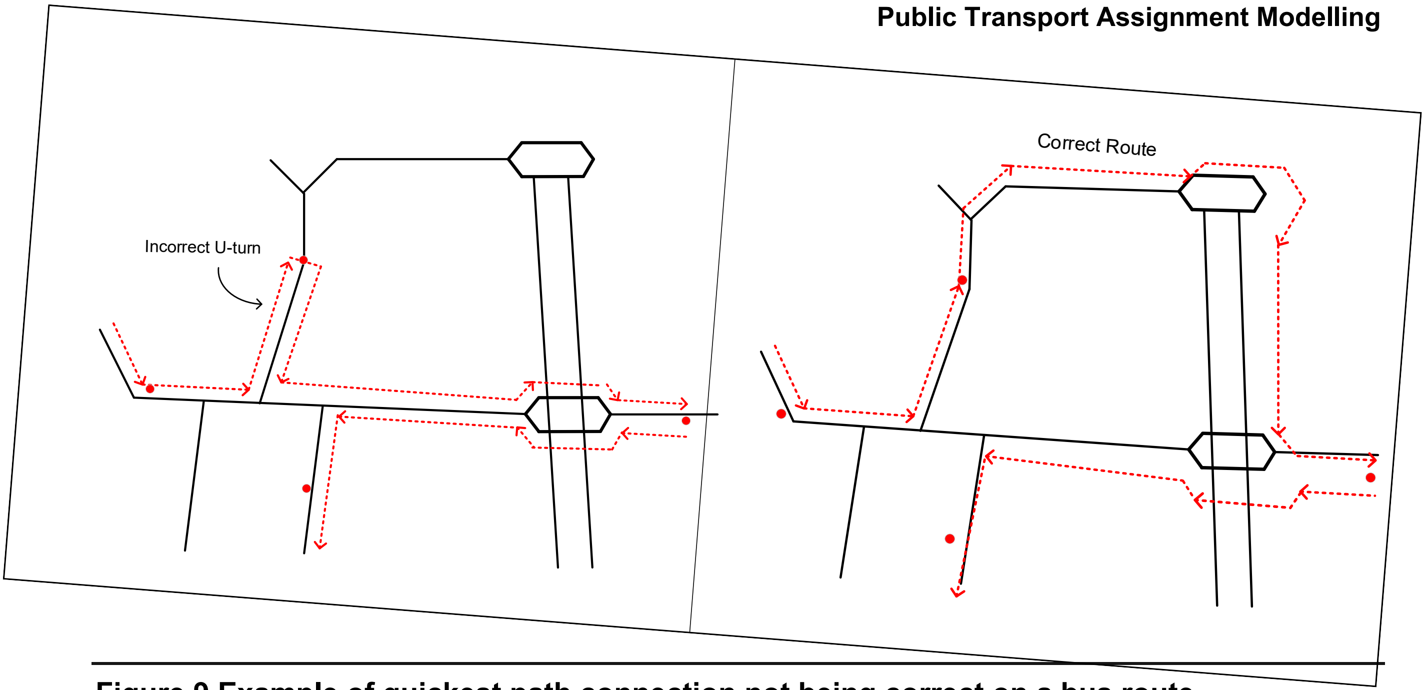

Figure 9 Example of quickest path connection not being correct on a bus route

6.3.21 A similar problem could also occur if the bus network includes links that are omitted from the highway network for example a diversion into a housing estate. This can be overcome either by adding the extra links or simplifying the route to miss out the diversion if unimportant (in a strategic modelling sense). Where speeds linked to highway models are used to determine bus journey times adjustments should be considered where such a diversion may add a few minutes of travel time to the route.

6.3.22 In the External Area, more grouping and simplification of bus stops is expected. For example, six bus stops in a small village could be represented by one node (especially where zone sizes are too large to differentiate between multiple stops). The level of detail should be reduced, grouping stops and services to aggregations that make sense given the zonal detail. An alternative approach to simplifying the network is to code proxy public transport services by coding times/speeds on the non-transit links that approximate the public transport generalised costs.

6.3.23 Open data may contain errors, and additional errors may be introduced in the data processing step. It is common to find that some services are missing and some are duplicated. Long distance bus routes that operate as a single service may be split into sections that need to be combined. Careful checks are needed and this is covered in section 6.4.

Travel Times

6.3.24 The travel times or speeds for buses should either be hard coded or have highway congestion applied incrementally. Hard coding is normal for rail and other modes that are segregated from general traffic.

6.3.25 For modes that share road space with cars, particularly bus, it is important for times to be linked to highway times and come either from a linked highway model or GPS data if available. However, there is a risk that the highway model is not reliable at this level of detail. For example, if there are implausible delays at certain nodes, these delays would be passed on to the bus and in turn to the demand model. To overcome this, an incremental approach that aggregates speeds across the route or a group of links or area of city should be used. This will help to dilute individual outlier delays. For bus priority measures there should be a separate link type representing the bus priority speeds or travel times.

6.3.26 Of particular importance will be the accuracy and consistency with which in-vehicle times are represented. Timetables may not be adhered to, of course, and for bus services in direct competition with a proposed new public transport service, in-vehicle times should be defined consistently (i.e. either as the vehicle arrives at stop or leaves the stop), with an allowance for dwell time at stops. Some link times have the dwell time included, and in this case how the dwell time is included in the link time should be consistent (for example it is always added to the departing stop).

6.3.27 The relationship between highway and bus times can be expressed using functions. At its simplest these are some multiple of the highway speed that might vary by area/sector and/or link type. These functions need to be calibrated using base modelled car time as an independent variable and timetable (or GPS/Automatic Vehicle Location (AVL)) time as the dependent variable.

6.3.28 More complicated functions may include other independent variables for example stops per km or boardings/alightings. The aim here is to capture stopping time (dwell time and delay re-entering the traffic stream) separately from the run time. These functions could be quite complex but in general a simple function of highway speed is likely to be adequate unless there is high and variable delay at stops.

6.3.29 Where bus times are hard coded in the base (not linked to highway speeds) then the future year speeds should be adjusted down if road congestion is expected to get worse. This can be done by using a simple travel time index from a highway model (aggregated by areas and/or link type) or from, for example, National Road Traffic Projections (NRTP). This simplified approach may be suitable if full highway model linkage is not proportionate.

6.3.30 The Department's default expectation is for the highway and public transport networks to both have the same node and link structures (with the addition of rail, tram and walk links to the public transport network), although this is not always possible and whilst this may be difficult to achieve, effort to ensure network compatibility will pay dividends when dealing with mode interactions during model running and analysis. When the bus journey times are linked to highway speeds (and where needed, allowance made for bus lanes/bus priority measures) the same structure simplifies the transfer of congested times. This requires a strong coordination when coding new bus schemes. In practice it is challenging to keep both the highway and public transport networks up to date and consistent and the use of separate software for highway and public transport networks presents a particular challenge that needs careful planning to overcome.

6.4 Pre-Calibration Checks

6.4.1 This section covers the pre-calibration checks around network connectivity and service coding accuracy that should be undertaken before assigning a trip matrix. This is a minimum set of checks, but further checks may be necessary for complex or high value schemes. The efforts required for this important stage of model development will ease calibration and validation of the model in later stages.

Network Links and Nodes

6.4.2 Overlaying network links and nodes on a GIS map base is helpful to check all expected links are included, and GIS analysis can also verify the coding of link distances. In undertaking such comparisons it is useful to give particular focus to:

- the walking route connections near stops and stations
- that interchange is possible where known to be material, especially at major interchanges
- avoid the network containing zero length links
- compare crow fly lengths with coded link lengths and review any links where modelled length is significantly more than or shorter than the crow fly (relative to link length)

6.4.3 The number of access/egress/transfer options provided for each connection is usually crucial in the calibration process and representation of ~~the~~ user behaviour. Useful data to calibrate or validate the access/egress component is provided by the National Travel Survey (NTS). A review of interchanges can be carried out using Electronic Ticket Machine (ETM) or smart card data where recorded fares can be traced back to the different operators. Refer to TAG unit M1.2 Data Sources and Surveys for more information on data sources and surveys.

Representation of Services

6.4.4 Using the interactive tools available with most public transport assignment modelling software, the routes and stop locations of each transit line can individually be reviewed where time permits. An alternative approach is to carry out detailed service-checking at a project/study level and ensure any corrections are fed back to the overall model owners. Typically, the review covers three steps as outlined below:

- **Plausibility of routes** – check on reasonableness of derived transit lines
    - Are there duplicated routes?

  - Are there very short routes?
  - Are there unusual routings (looping, detours, changes in direction)?
- **Plausibility of attributes** – check on the reasonableness of transit line characteristics
  - Are there large variations in speed along the transit line?
  - Are there very high or very low speeds?
  - Are the intermediate stopping patterns correct?
- **Plausibility of assignment** – check that the transit line and network connectivity can generate demand on the coded services
  - Stations/stops with no boarders/alighters (through assigning a dummy matrix – see paragraph 6.4.8)
  - Directional symmetry of service coding (headways, assigned volumes) as this may alert to missing services, noting that there is likely to be peak directionality in supply.
  - View the report files from the tool/assignment software looking for reported errors, such as zero or overflow values

6.4.5 The bus time function calibration should be documented. This should include model fit statistics and scattergrams for each calibrated model segment for example A roads, rural roads etc, or for different sectors of the model.

6.4.6 Check modelled service frequencies at key locations against summarised timetable data or survey data (vehicle counts) if available.

6.4.7 Check journey times for some OD pairs against predications from online journey planners or accessibility modelling software. A wide selection of journey times should be checked, including journeys within the area of interest and study area, as well as some externals.

Paths and Skims

6.4.8 Create a dummy matrix (all cells are 1) and then skim inter-zonal distances and times from the network:

- Check that a path can be built between all zone pairs
- Check symmetry of skims between time periods
- Plot isochrone and iso-distance vectors for individual zones and check for unusual patterns and outliers

- Check for OD pairs with zero in-vehicle time or 'walk all the way'
- Check for OD pairs with high numbers of transfers

6.4.9 View best paths between specific start and end points and 'trees' (meaning best paths from a single start node to every end node, or vice versa) for selected nodes. Selected nodes should be identified based on the model use, for example along a specific corridor, or key interchange or terminating locations. To review, practitioners should:

- View the trees for locations across different parts of the model and for a variety of mode subsets for example walk only, walk + bus to identify any unexpected patterns
- Building single OD best paths can provide more detail for investigation of odd results seen in the 'tree' plot
- Look for holes in connectivity and unusual paths and review the plausibility of any gaps to determine if there are any specific missing services

7. Route Choice Calibration

7.1 Introduction

7.1.1 The quality of 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 (credible access/egress times to/from all reasonable stops/stations)
- the accuracy of the network and service coding and the appropriateness of the simplifications adopted.
- the accuracy with which interchanges and generalised costs are modelled, which are dependent not only on data and/or coding accuracy and appropriateness but also on the parameters representing user behaviour

7.1.2 At various stages in the model development process, modelled routes should be examined and their plausibility checked to help identify inaccuracies in the network coding elements listed above. For example, as suggested in section 6.4, early plotting of chosen path routes is a useful way of identifying network and centroid coding errors. Modelled routes may also be usefully considered in the later stages of calibrating the model.

7.1.3 A public transport route involves a walk leg to board the first service, one or more legs on vehicles and, after alighting, a walk to a final destination. Some software tools and assignment methods can represent passengers walking for their entire journey. Such routeing is usually inconsistent with demand matrices, which are derived from public transport passenger data. There is therefore a specific need to monitor and ensure that the extent of such routeing behaviour is negligible.

7.1.4 At this stage in the model development process, before considering travel demand, the focus of the calibration task is around the aggregate corridor rather than detailed consideration of individual services or stops.

7.2 Calibration

7.2.1 As the calibration of the assignment proceeds, checks should be carried out by inspecting the routes through the network. Checks should be undertaken using both:

- select link analysis, identifying the origins and destinations of routes traversing individual network links, and
- a selection of origins and destinations

7.2.2 The number of comparisons should be proportionate to the size and complexity of the model. Where crowding is represented and if there are substantive level of service differences between model periods, separate analysis should be undertaken for every model period. The following factors should be multiplied to estimate a suitable number:

- number of main centres + major trip generators
- distance categories
- number of sub modes or distinct user classes

For example, for a modest size town with a single centre, served by bus and rail networks, three distance categories may be adequate, and it may be adequate to review 1*2*3=6 routes. For a long distance rail corridor, there may just be a single major terminus, but 4 or 5 distance categories may be more appropriate to represent the range of travel distances.

7.2.3 Particular focus should be given to the specific corridor or study area that warrants the need for the model.

7.2.4 Information on ‘observed’ routes can be derived from route finding tools or GPS tracking and may also be informed by local knowledge. The analysis should consider which routes are chosen in practice, and if possible, how demand distributes across these. If the comparison is poor in terms of route choice, the network representation should be reviewed to identify misrepresentation and errors. A review of the network may call for a refinement of the:

- zone centroid connector times, costs and loading points

- representation of interchanges (for example walk link distances and times between nodes)
- fidelity of the service representation, reconsidering simplifications in relation to the network and to the main interchanges

7.2.5 Further steps may involve reconsidering the following parameters in the generalised cost formulation:

- quality factors representing perception of specific modes
- walking and waiting time coefficients
- the boarding or interchange penalties that represent local conditions (specific interchange).

7.2.6 Where systematic variance in flows (e.g. too high or too low) along corridors or more broadly exist, this may be indicative of specification issues relating to segmentation and the parameter values adopted in the mode choice model. Review of any parameter estimation should be undertaken, particularly where ticketing data permits more disaggregate analysis to determine if, for example, greater heterogeneity of time valuations exists between passenger segments that may help to explain any variances.

8. Trip Matrix Calibration

8.1 Introduction

8.1.1 Prior trip matrices should be created following the advice referred to in TAG unit M2.2 Base Year Demand Matrix Development.

8.1.2 Assigned (modelled) passenger flows should be compared with passenger counts across screenlines, cordons, individual or group of stops or stations, individual or groups of services and for major movements. Should this reveal discrepancies, initial courses of action are:

- reconsider the assumptions made interpreting data developing the prior trip matrices with a view to producing new versions which, when assigned, yield modelled flows which accord more closely with the counts (further explained in TAG unit M2.2 Base Year Demand Matrix Development)
- check how the comparison count data was collected and processed. Ensure there are no definitional differences between this data, and that used to create the trip matrices.
- review further the adequacy of the network coding

8.1.3 There may be similar or greater confidence in the trip matrices than the count data. Prior trip matrices may for example have been derived from nearly complete ticketing data and there may be only limited independent count data. In these cases the use of automated matrix estimation methods that adjust demand using count data is **not recommended**. Other calibration actions discussed in this section should also not be undertaken in these cases.

8.1.4 The remainder of this section relates to circumstances where demand matrices are developed from a limited sample of journeys.

8.2 Refinement of Prior Trip Matrices by Matrix Estimation

8.2.1 If the decision is taken to adjust demand matrices based on differences between counts and comparable assigned prior matrix flows, advice is set out in TAG unit M3.1 Highway Assignment Modelling. The advice in this unit also applies to matrix estimation of public transport demand matrices.

8.2.2 Where there are significant systematic variations in the modelled and observed demand, rescaling matrices across sectors or operator groups may be considered. This is recommended prior to any matrix estimation from counts as it will help to preserve the underlying distributions in the data and focus on more systematic adjustments to the demand.

8.2.3 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. Assignment methods are sensitive to particular behavioural parameters and methods and accordingly counts should be grouped across services in a corridor in applying matrix estimation methods. Where distinct user classes are assigned, these should be aggregated to the segmentation available in the count data in applying matrix adjustments.

8.2.4 Where the model represents crowding, it may initially be helpful to test the matrix estimation method without representing crowding. Nevertheless, convergence of the iteration between assignment and matrix estimation should include crowding in the final use of the matrix estimation tool.

8.2.5 Matrix estimation should not be used to only achieve the flow validation criteria at the expense of significantly changing the demand matrix patterns. The tests on the significance of matrix estimation changes are set out in TAG unit M3.1 Highway Assignment Modelling. Where sector-sector changes are compared, the sectors should be aggregated to represent flows which should be of (at least) a few hundred passengers.

8.2.6 The practitioner should report comparisons of count data with assigned flows of both the prior matrix and of the matrix post matrix estimation.

9. Assignment Calibration

9.1.1 Calibration of the model should progress sequentially (see Figure 1). Calibration action should follow the guidance in sections 7 and 8, and focus at a corridor level. This assignment stage focuses at a more detailed level. This may include refining service boarding patterns and stop/station catchment areas along a corridor, or the assignment of demand between services operating in a corridor. Effort put into pre-calibration checks of the model (see section 6.4) will help improve this detailed calibration.

9.1.2 This stage will include use of the same count data used in earlier stages, but the comparisons made will be more disaggregate, comparing flows on individual services, stations of small groups of stops.

9.1.3 Calibration actions are similarly likely to be detailed, consideration may include:

- reviewing the distribution of population within zones and available walking routes to better to reflect average walk times and refine centroid coding assumptions
- at stations or park and ride locations, consider how the use of car access is modelled, including possible 'fast walk' representation of zone connectors (where not explicitly represented in demand models)
- service reliability and effects on perceived wait times in refining service quality parameters and wait time parameters
- the quality and perception of interchange facilities and consideration of timed interchanges (where frequency based assignment methods are used) to refine representation of station interchange facilities

9.1.4 Depending on the extent and nature of local refinement, it may also be appropriate to review the corridor comparisons (as described in section 7) and refine previous calibration assumptions.

9.1.5 When crowding is represented it can be helpful to compare validation measures with and without crowding. Where service levels are otherwise similar between time periods it may be sufficient to compare peak and off-peak periods. These comparisons are undertaken to indicate whether the detailed calibration adjustments required are consistent. If there are differences between modelled periods it is more likely that the representation of crowding should be reviewed further.

9.1.6 It is also helpful at this stage to 'stress test' the model, particularly when crowding is represented. This may be by increasing demand by a fixed proportion (such as 20%) or by factoring up in vehicle time of the dominantly used services. This may reveal faults in the route choice or interchange points which previous checks have not detected. For example, some passengers may be modelled using direct services which are severely congested rather than diverting to alternatives on less direct routes (slower services or more interchanges). All tests should be recorded and reported.

10. Validation

10.1 Introduction

10.1.1 Once the practitioner has created a calibrated model, the model results should, ideally, be compared against data that is not used in calibration. This is referred to as **validation**. Validation is a test against independent evidence and involves comparing modelled and observed data that is independent from that used in calibration. Practically, it is expensive to obtain sufficient data and may be proportionate to process and aggregate data independently to test model performance, however this should be explained in reporting. Section 4 provides guidelines to assess the differences between the model results and observed data.

10.1.2 A model is calibrated by modifying all its elements (input data, data processing, assumptions, parameters and factors on overall behaviour). In assessing the impact of each change, a metric is needed to understand how and in what way the model is impacted.

10.1.3 Using validation metrics to expose and diagnose improvements is a method used to focus calibration. Clear reporting on the validation criteria provides the components to judge the level of calibration achieved. This helps to interpret the outcomes of the model, providing reassurance on the level of trust in the model. Adjustments to the model intended to reduce the differences between the modelled results and observed data should be documented. Complementary data used during the process should be explained.

10.1.4 The Model Validation Report (see section 11) should state if any data in the validation was used for elements of the calibration and in what form. Whilst the extent of the data available for model development and calibration is often limited, the data should not be used in the same form in assessing validation. For example, if ticketing data are used to derive trip matrices, further processing of the data to derive counts would not provide any additional verification of the trip matrices themselves (these comparisons would be expected as part of the matrix development). However, their independence from the model routeing or assignment methods means that such derived counts are suitable for assignment validation.

10.2 Network and Service Validation

10.2.1 This section covers verifying the network validity through comparisons with observed data and the reporting that should accompany the comparisons. The network and service items to verify are discussed in section 4.3. These comparisons first provide evidence of the network coding accuracy.

- The differences in service frequencies should be minimal and comparisons should be made in each period and in each direction. The differences in service end-to-end run times should be minimal. Practitioners can provide a scattergram or table of observed vs predicted end-to-end run times for each route (the observed data can be the same data used for the calibration because this is not a validation check but a processing verification check). In this case observed times could be either planned run times or actual run times. Note that any differences between planned and actual run times should be reconciled in data processing. If not, this stage will simply expose definitional differences.

- The differences in journey times should be minimal and comparisons should be made of point-to-point run times from the model against timetables for a sample of services as well as all stop-to-stop journey times. This is to demonstrate that the data have been correctly processed.

10.2.2 Additional validation evidence including isochrone plot(s) centred on one or more key locations (for example a bus or rail station in the city centre) will add confidence in the model's network. This could be annotated with journey times from an online journey planner or accessibility software for a sample of locations.

10.2.3 If suitable comparator data can be sourced, high level checks on total operations can be carried out (for example bus kilometres by route in each period or daily or annually).

10.2.4 The differences between modelled journey times along routes compared to timetable times should be within the criteria set out in Table 6. Any exceedances of these acceptability guidelines should be explained.

10.3 Route Choice Validation

10.3.1 To demonstrate the route choice is valid, the following evidence may be shown, alongside the checks listed in section 7.2. These plots should be included in the Model Validation Report and any differences should be reported and justified.

10.3.2 The following plots should be included:

- A selection of origin to destination (OD) routes that represent both the route corridor under study and any alternatives.

- Journey times for a sample of OD pairs for a variety of movement types. This can then be compared with online journey planners or accessibility software. Movements could include radial, orbital, inner, outer, external etc.

10.3.3 The number of routes should be proportionate with respect to the size of the model and the complexity of the scheme, 10.3.4 When the model includes crowding, route choice should be reviewed at the validation stage. Plots should be provided with and without crowding for different time periods and with stress tests. These are likely to expose whether there is credible diversion to non-crowded routes. Observations of routes are not usually available, so these checks should be based on local experience and judgement.

10.4 Trip Matrix Validation

10.4.1 The quality of the trip matrix should be explained as set out in TAG unit M2.2 Base Year Demand Matrix Development. This will include comparison with screenline count evidence. 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.

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

10.4.3 While it is possible to validate trip matrices using cordons or screenlines, care needs to be exercised to ensure what is used is appropriate for the model. Sometimes discrepancies between modelled flows and counts along a screenline are the result of erroneous routeing in the model. In this case long screenlines will show the quality of the matrix more clearly. In other cases where the project involves primarily radial network and there is little option to route outside a cordon, cordons should be used.

10.4.4 The validation of the trip matrix should involve a comparison of modelled and observed passengers across complete screenlines and cordons, as stated in section 4.3. Trip Matrix Validation (see Table 4). If the criteria are not met for all or nearly all screenlines and cordons, remedial action should be considered. Any remaining differences outside the validation criteria should be justified.

10.4.5 The comparison should be included in the Model Validation Report as a table indicating which screenlines are only used for validation, and which ones were also used in the matrix calibration. Consider differentiating which were used in the various steps of matrix development.

10.4.6 Where matrix estimation methods are adopted the extent of change must be reported as set out in TAG unit M3.1 Highway Assignment Modelling, together with screenline flow comparisons of both the prior and post estimation matrices.

10.5 Assignment Validation

10.5.1 In addition to evidence of network, services, route and trip matrix validation, the Model Validation Report should include evidence of the assignment validation with the below comparisons.

10.5.2 Validation of the assignment should involve comparing modelled and observed data such as grouped boarding/alighting/onboard flows, usually by public transport mode. In simpler models the comparison can be undertaken at the level of individual bus or train services. In urban centres this also includes comparing single boarding/alighting/onboard flows on stop/links/services. The differences should be in line with the criteria and guidelines given in Table 5 (assignment validation in section 4.3).

10.5.3 The annual patronage from operators/TOC should be compared to modelled results. The comparison should show the assumptions of the annualisation factors for the model results. The differences should be in line with the criteria and guidelines given in section 4.3.

10.6 Testing Responsiveness

10.6.1 In addition to providing evidence of the validation of the assignment in the base year, it is useful to include information on how the model responds to changed inputs to help judge the potential suitability of the model for forecasting.

10.6.2 The main functionalities required for the specific scheme/study should be considered and the response assessed to determine if it is in line with expectations. The Public Transport Model Validation Report should include this information. The responses that may be considered will reflect the model purpose and may include:

- changes to the frequency of specific services, which should result in increased/decreased attractiveness of them
- changes to the capacity of specific vehicle types, which should be reflected in the change of magnitude and location of the congested sections, where crowding is represented
- journey time changes, which should impact on the passenger routeing and loading of the affected services
- where fares are modelled, changes to fares for specific modes, operators or services should impact on passenger routeing and service loading.

11. Reporting

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

- Public Transport Assignment (Local) Model Specification Report (or as part of the Appraisal Specification Report); and
- Public Transport Assignment (Local) Model Validation Report.

11.1.2 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.

11.1.3 The recommended structures of these reports are set out in Appendix D. A network coding manual should be appended to one of these reports.

11.1.4 The Model or 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
- the methodologies for network development, trip matrix development, and for calibrating and validating the network, route choices, trip matrices, and assignment

11.1.5 The Model Validation Report will be the last task in the development of the base year model (the development of forecasting model may proceed after and should also be reported). The report should include:

- updated sections on proposed uses of the model and key model design considerations
- model standards (including convergence)
- key features of the model such as a description of the calibration and validation data used and descriptions of the network and trip matrix development
- descriptions of the calibration and validation of the network, route choices, trip matrices, and assignment

11.1.6 It is important to report on the level of confidence that can be placed on model outputs. The achievement of the validation guidelines 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'. Model Validation Reports should therefore **not** include statements such as 'because the validation standards have been (largely) achieved, the model is necessarily fit for forecasting purposes'. The findings from the responsiveness testing discussed in section 10.6 should be reported and commented on in the Model Validation Report.

12. References

12.1.1 Institution of Highways and Transportation (1996) Guidelines on Developing Urban Transport Strategies.

12.1.2 Institute of Transport Studies, University of Leeds, in association with John Bates Services (2003), Values of travel time savings in the UK.

12.1.3 MVA Consultancy (2008) Valuation of Overcrowding on Rail Services

12.1.4 Australian Transport Assessment and Planning T1 Travel Demand Modelling: https://www.atap.gov.au/tools-techniques/travel-demand-modelling/appendix-d

13. Document Provenance

13.1.1 This Transport Analysis Guidance (TAG) unit is based on the previous TAG unit M3.2 (May 2020) but was significantly restructured and expanded in spring 2024 with the inclusion of more detailed best practice guidance.

Appendix A: Glossary of Terms

Calibration:	Adjustments to the model intended to reduce the differences between the modelled and observed data.
Centroid connections:	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.
Cost skimming:	Calculating costs along a particular path, for each origin-destination pair.
Crowding:	A measure of demand in relation to capacity.
Demand model:	A model which forecasts changes in trip frequency, mode of travel, time of travel, and trip destination.
External Area:	The area outside the **Internal Area**, where the impacts of interventions would be so small as to be reasonably assumed to be negligible.
Frequency-based Approach:	For this assignment method, departure times of individual services are not considered explicitly, but practitioners refer to the service headways, or to their inverse (the service frequencies).
Generalised Cost:	A linear combination of time and money costs, expressed in time or monetary units.
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.
Internal Area:	This is the area over which proposed interventions have influence. Modelling detail in this area is at its most detailed. This area of the model should have full consistent demand and supply representation.

Matrix estimation (from counts):	The adjustment of prior trip matrices so that, when assigned, the resulting flows accord more closely with counts used as constraints in the process.
LENNON	Latest Earnings Networked Nationally Over Night (rail ticket database)
Path building:	The identification of all potentially attractive paths and the calculation of their cost.
Prior trip matrix:	The trip matrix to be subjected 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.
Route choice:	The generation of alternative routes through a network on the basis of generalised cost or time.
Timetable-based Approach:	Timetable-based approaches reflect the actual clock face vehicle arrival/departure times at the time when users make their choices.
User class:	Combinations of vehicle types and trip purposes which are assigned separately in a multi-user class assignment.
Validation:	The independent 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 a set of prescribed measures) and then assessed using these criteria.

Appendix B: Assignment Methods

B.1 Deciding on an Assignment Approach

B.1.1 The choice between frequency-based and timetable-based approaches, and between deterministic and stochastic models is driven by practical questions. The impact of these considerations on the choice of the model to be used is summarised in Table 4.

- Is the public transport system operating with high or low frequency?
- How punctual is the system?
- How regular is the system?
- What kind of passenger information is available?
- Does the demand vary significantly over the modelled period?
- How detailed is the demand information (by day, by hour, or even more specific)?
- Does the system experience capacity problems?
- How big is the network to be modelled?
- Is the network complex, so that regular users behave differently compared to occasional users?
- How homogenous is the likely user group? For example, is there a large difference in perception or valuation of travel time?
- What are the levels of interchange between services?
- How many different sub modes are there?
- What fare structures are used and do they differ between services/modes/operating companies?
- Is the necessary data available for timetable-based modelling?

Practical considerations

B.1.2 Compared with frequency-based models, an advantage of timetable-based approaches is that vehicle loadings can be predicted for specific services at specific points in time. This means timetable-based models are better suited for operational considerations. However, if passenger arrivals and/or vehicle departures are highly variable, frequency-based approaches may give more realistic results, whilst the extra data and calculation efforts of timetable-based approaches may be unnecessary with high-frequency systems.

B.1.3 Despite their theoretical advantages timetable-based methods suffer from a number of practical disadvantages:

- there are greater data and associated resource requirements
- results can be very sensitive to the actual timetable specified
- it can be difficult to predict the timetable accurately a) for a scheme which does not yet exist or b) for several years into the future
- the way in which unreliable services should be handled is not clear
- run times are much higher than for frequency-based approaches

B.1.4 A timetable-based approach can often better capture the complexities within a public transport network. However, frequency-based approaches may be proposed where practical considerations mean that timetable-based approaches may be disproportionate or impractical, or in cases where it is judged that the model will be fit for purpose even in the absence of a full representation of the service and network complexities. Where this is the case, for instance with a model of a complex urban area such as a major city, a thorough appreciation for the approximations being made in the model and the impact these have on the outcomes (e.g. the appraisal) is of clear importance and should form part of the model specification agreed in the Appraisal Specification Report.

High and low frequency services – wait times

B.1.5 Passengers typically dislike waiting as part of their journey, hence it is weighted more highly than IVT and passengers will typically seek to time their arrival at a station or stop to minimise the amount of time that they wait. Where services operate at a high frequency, a commonly accepted threshold of which is services of less than 10-15 minutes headway, passengers are more likely to arrive without consulting timetables (i.e. ‘turn up and go’), and the differences between desired and actual departure time can be treated as a constant, for example half the headway. However, if the service operates with a lower frequency than this suggested 10-15 min threshold[8], travellers may be expected to time their arrival at the station for specific services. From this it follows that frequency-based models are less suitable for services that operate with headways larger than this threshold. On low frequency services, initial passenger wait times should either be capped to a maximum value or derived from a wait curve (see section 3.2), and this applies to both frequency, and timetable based assignment models.

[8] Combined headway of all used services for the OD movement

Passenger information and service punctuality

B.1.6 The more information a traveller has and the more reliable this information is, the more the choice will be service-based rather than route-based and hence a timetable-based approach will be more valid. Frequency-based models will be more suitable if services operate with low punctuality and/or a low level of user information. Delays and irregularity have to be treated implicitly or explicitly in timetable-based models. An implicit treatment is possible by adding error terms to the path choice model. A Monte Carlo technique allows the explicit treatment of delays, but availability of this may be limited in commercial packages.

B.1.7 Only in the case of a high frequency service that is unreliable or has poor passenger information is a frequency-based model sufficient. If the service is not yet operating, estimates of the services will have to be made and this reduces the advantages a timetable-based approach might have. In the case of low frequency services or high frequency services that are reliable and have good provision of passenger information, timetable-based approaches are advised.

Service regularity

B.1.8 Service regularity is a separate issue from punctuality. In this case it is the regularity in the headway or scheduled intervals between the arrivals of the vehicles rather than unplanned delays. Frequency-based models assume an equal headway per service and thus an equal share of passengers between the runs of this service. If a service is not scheduled to arrive with regular headways, (for example 00, 15, 30, 45 after the hour), but say 10, 15, 40, 45 after the hour, this might lead to line loading errors in frequency-based models. Further, a timetable-based approach might be required if there is a major influx of passengers during a certain period (like an underground station connected to a train station that brings a large number of passengers to the underground network once every hour) in order to show overloading of certain services. An additional consideration in such a case is that a timetable-based approach is better equipped to estimate the correct average wait times, particularly if transferring passengers have no choice.

Crowding and capacity constraints

B.1.9 In principle, if in-vehicle crowding is, or is expected to be, so severe that demand for the mode concerned is, or would be, constrained, some means of representing the costs of the crowding for use in the demand model would be required. In practice, crowding is more likely to be of importance in the allocation of trips between alternative routes through a combined network model than in models of separate networks. Section 3.6 offers advice on how to represent in-vehicle crowding costs.

B.1.10 Congestion in highway assignment and capacity restraint in public transport assignment are not the same. This is for two reasons. Firstly, the cost function in the case of public transport is not increasing continuously, but the finite capacity of public transport vehicles will lead to a step function; either a traveller can board the arriving vehicle or not, in which case the waiting time will increase by one headway. It is worth noting that in practice this can have implications for model convergence. Secondly, this capacity restraint will only be experienced by boarders. Passengers on-board have priority and do not perceive the same increase in cost, although they may experience some increase in discomfort due to crowding. In frequency based-models it is possible to handle capacity restraint implicitly through a concept referred to as **effective frequency**. The idea is to increase the perceived costs of boarders through a local reduction in service frequency, reflecting the fact that the passenger may not be able to board a vehicle at a particular point because of overcrowding. This approach can be criticised for two reasons: a) a cost increase based on the number of passengers wanting to board and spaces available is still a continuous cost function; b) an increase in cost does not prevent line capacities being exceeded, leading to inaccuracies elsewhere in the network. Additionally, it is not clear how the correct wait time can be extracted for demand response modelling and appraisal.

Scale of network

B.1.11 Because of the more detailed network description and because of the dynamic representation of supply and demand, timetable-based approaches are computational more demanding and data hungry, particularly in larger networks.

Variation in user behaviour

B.1.12 If the variation in user behaviour is an important issue, models using Stochastic User Equilibrium (SUE) assignment are needed. A ‘dispersion factor’ can be used to model the different cost perception of different travellers, and these dispersion factors can differ between demand segments. SUE assignment can be applied to timetable-based as well as frequency-based models. SUE assignment should also be applied if one wants to reflect the behaviour of occasional users in complex networks. Occasional users might not know about all available routing options and therefore the route choice might not be restricted to the least generalised cost path only. For low frequency services it is of less importance to distinguish frequent and occasional users, firstly because the route choice is in most cases not as complex and secondly because in low frequency services passengers will not often change their path en-route. The SUE models differ in their assignment assumptions. Logit, nested logit or probit models are most common. Where paths overlap significantly and hence path utilities are positively correlated (in practice usually the case), it is advisable to use the nested logit, C-logit or probit model. Logit models tend to be more tractable than probit models.

Summary

B.1.13 Table 7 summarises which assignment models are advised, depending on network characteristics, availability of data and (to a lesser extent) passenger behaviour and the options to be modelled.

Table 7 Summary of recommendations for public transport assignment model applicability

		Timetable-based (TB) or frequency-based (FB)	Stochastic User Equilibrium (SUE) or Deterministic User Equilibrium (DUE)
Service frequency	High		SUE
	Low	TB	DUE
Passenger information & service punctuality	High	TB	
	Low	FB	SUE
Transfer choice-making by travellers	Pre-trip	TB	
	En-route	FB	
Regular schedule	Yes		
	No	TB	
Crowding/ Congestion	Yes	TB	
	No		
Capacity problems	Yes	TB	
	No		
Scale of network	Large	FB	
	Small		
Day-by-day variations	Yes	TB	
	No		
Significant dispersion of behaviour	Yes		SUE
	No		DUE

Note: blanks indicate that either option is appropriate

B.2 Path Choice Methodology

B.2.1 Ultimately the best test of the adequacy of a particular algorithm is its ability to reproduce observed routing behaviour. Several methods are described in this section that should be considered in order to achieve this. Section 3.7 also discusses path building.

B.2.2 Where multiple paths are identified, some mechanism for allocating flow to each path is required, usually as a function of the generalised cost on each path, including all aspects of time (including access and necessary further interchanges), fares and comfort (journey quality). Ideally, explicit consideration would be given to common/overlapping and parallel paths (i.e. where the 'common line' dilemma occurs) and some way of including the representation of individual preferences may be necessary through probabilistic or stochastic methods. Path choice is governed by calculations of 'probability of use' of each of the acceptable paths between OD pairs. As noted earlier a useful distinction can be made between deterministic and stochastic methods.

All-or-nothing (deterministic) assignment

B.2.3 In an all-or-nothing assignment all flow is loaded onto the single minimum cost route for each OD pair. With frequency-based methods there is therefore no multi-routing. This may be an adequate reflection of reality in some cases, particularly in timetable-based models or simple networks. In others, e.g. complex urban networks, there is likely to be observed multi-routing which would require a more complex assignment method to model accurately.

B.2.4 The all-or-nothing assignment is a deterministic method. Methods below are probabilistic or stochastic.

Simple discrete choice

B.2.5 In these stochastic methods no consideration is made of whether paths are overlapping or in parallel. Only the generalised cost on each path is considered. The following discrete choice functions are used:

- Logit: the most commonly used discrete choice model where passengers are distributed over a set of paths according to the absolute difference in cost between the paths
- Power function (Kirchoff): passengers are distributed over paths according to the power of the ratio of the costs of alternative paths
- Box-Cox: a flexible model form that includes power and logit as special cases
- Lohse: uses the ratio of path costs relative to the minimum cost
- Probit: similar to logit, although error terms are normally distributed rather than a logistic distribution

B.2.6 In each case the 'spread' of the path choice can be controlled by a user-defined parameter. This determines how strong the preference is for the minimum cost path. This will depend on the level of taste variation among passengers and how complete their knowledge of services is ('errors in perception').

B.2.7 With the exception of probit (which is not actually used in commercial packages) all of the above have theoretical shortcomings regarding their ability to deal with a choice between correlated alternatives. Path utilities will be correlated if, for example, they share a common segment.

Models with 'independence'

B.2.8 The choice models given above in their basic form do not cater adequately for timetable-based stochastic assignment. Temporal factors are therefore incorporated into the models in order to make them more suited to timetable-based public transport routeing. In order to do this, interactions between different connections are defined:

- the temporal proximity of the connections with regard to departure and arrival
- perceived journey cost differences between connections
- fare differences between connections

B.2.9 These factors are combined to derive an **independence of connection** factor which defines the attractiveness of a particular connection relative to all others. They ensure that identical alternatives are assigned same volumes of passengers if no other connections with temporal proximity have an effect.

Service frequency model

B.2.10 Passengers are assigned to a path according to the frequency (or a function of frequency) of services along available paths, i.e. the probability of using a path is proportional to its frequency or a function thereof. This is a simple approach where travellers are assumed to possess no knowledge of timetables or journey times and take the first reasonable service from the stop.

Service frequency and cost model

B.2.11 In this extension of the service frequency model the path choice probability is modified to reflect the difference in costs between the paths. Passengers are assumed to have some knowledge of the frequencies and journey times of alternative services and will decide whether to take the first feasible service from the stop or wait for a faster one.

B.2.12 In all but the simplest public transport networks, travellers between certain OD pairs are likely to be split between different paths and services. Therefore a multi-routing algorithm must be used to reproduce this behaviour. Most path identification methods are acceptable; the crucial part of the algorithm is how the flow is allocated to the used paths. Methods that take into account generalised costs, rather than just frequencies are likely to produce better-validated results. Where there are overlapping routes methods that consider the degree of independence between competing routes should, ideally, be used.

Appendix C: Model Convergence

C.1 Base Year and Future Year forecasts

C.1.1 Not all assignment packages produce all the measures mentioned in section 4.4 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.

C.1.2 As convergence is greatly affected by the level of crowding 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.

C.1.3 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.

C.2 Assessing accuracy of final results

C.2.1 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

C.2.2 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.

C.2.3 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 public transport 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.

C.2.4 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.

C.2.5 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.

C.3 Presentation of convergence results

C.3.1 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 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 D: Reporting Requirements

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

- Public Transport Assignment (Local) Model Specification Report
- Public Transport Assignment (Local) Model Validation Report

D.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.

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

D.2 Public Transport Assignment Model Specification Report

D.2.1 This report should form part of an overall Appraisal Specification Report (ASR), 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

Internal and External Area

Zoning System

Network Structure

Centroid Connectors

Time Periods

User Classes

Assignment Methodology

Generalized Cost Formulations and Parameter Values

Crowding and capacity constraints

Relationships with Demand Models (TUBA) and Highway Assignment Models

5 Calibration and Validation Data Specification

6 Network Development Methodology

7 Trip Matrix Development Methodology

Travel Demand Data

Partial Trip Matrices from Surveys

Trip Synthesis

Merging Data from Surveys and Trip Synthesis

8 Network and services 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

D.3 Public Transport Assignment Model Validation Report

D.3.1 This should refer to the public transport model specification report where needed. 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

Internal and External Area

Zoning System

Network Structure

Centroid Connectors

Time Periods and justification of the choice

User Classes

Assignment Methodology

Generalized Cost Formulations and Parameter Values

Crowding and capacity constraints

Relationships with Demand Models (TUBA) and Highway Assignment Models

5 Calibration and Validation Data

6 Network Development

7 Trip Matrix Development

The recommended content of this section is detailed in TAG unit M2.2 Base Year Demand Matrix Development (Appendix F).

8 Network and Service Calibration and Validation

Network and service calibration

Network and service validation

9 Route Choice Calibration and Validation

Route Choice Parameters and source

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