High Speed Two (HS2) Phase Two

PLANET Framework Model (PFM v7.1)
Model description

High Speed Two (HS2) Phase Two

PLANET Framework Model (PFM v7.1) Model description

Contents

Glossary

AADT	Average Annual Daily Traffic
AAWT	Average Annual Weekday Traffic
ASC	Alternative Specific Constant
AVE	Alta Velocidad Española (Spanish High Speed Rail)
CA	Car Available
CAA	Civil Aviation Authority
CAF	Car Available From
CAT	Car Available To
CBA	Cost Benefit Analysis
CIF	Common Interface Format (for network data)
COA	Census Output Area
COBA	Cost Benefit Analysis (DfT program/methodology)
CP	Crowding Penalty
CSV	Comma-Separated Variables (file format)
CV	Coefficient of variation (standard deviation/mean)
DCF	Discounted Cash Flow
DfT	Department for Transport
FJTS	Frequency and Journey Time Strategy [PT assignment]
FORGE	Fitting On of Regional Growth and Elasticities (NTM)
FY	Financial Year
GJT	Generalised Journey Time
GJTC	Generalised Journey Time including crowding
GJTCAE	Generalised Journey Time including crowding + Access/Egress
GOR	Government Office Region
HAM	Heathrow Access Model
HBEB	Home-Based Employers' Business
HBEd	Home-Based Education

HBO	Home-Based Other
HBW	Home-Based Work (Commuting)
HSR	High Speed Rail
IEP	Intercity Express Programme (DfT initiative)
ITN	Integrated Transport Network
IVT	In-Vehicle Time
LASAM	London Airports Surface Access Model
LATS	London Area Travel Survey
LD	Long Distance (> 50 miles one-way)
LDM	Long Distance Model (DfT)
LENNON	Latest Earnings Networked Nationally Over Night (Rail ticket database)
LF	Load Factor (ratio of passengers to seats)
LOS	Level of Service
LSE	London and South-East
MSA	Method of Successive Averages (algorithm)
MSOA	Middle layer Super Output Area
NAM	National Accessibility Model
NAPAM	National Air Passenger Allocation Model (DfT)
NAPDM	National Air Passenger Demand Model (DfT)
NCA	Non-car available
NHB	Non-Home-Based
NHBEB	Non-Home-Based Employers' Business
NHBO	Non-Home-Based Other
NHBW	Non-Home-Based Work
NPS	National Passenger Survey (rail)
NPTDR	National Public Transport Data Repository
NRTS	National Rail Travel Survey
NSE	Network SouthEast
NTEM	National Trip End Model
NTM	National Transport Model (DfT)

NTS	National Travel Survey
OD	Origin-Destination
P/A	Production/Attraction
PDFH	Passenger Demand Forecasting Handbook
PFM	PLANET Framework Model
PLD	PLANET Long Distance
PM	PLANET Midlands
PN	PLANET North
PS	PLANET South
PT	Public Transport
RoH	Rule of a Half (Benefit calculation)
RPI	Retail Prices Index
SCM	Station Choice Model
SNCF	Société Nationale des Chemins de fer Français (French Railways)
SP	Stated Preference
TEMPRO	Trip End Model presentation PROgram
TGV	Train à Grande Vitesse (French High Speed Train)
TOC	Train Operating Company
TRADS	TRAffic flow Data System (Highways Agency database)
TTW	Travel To Work area
TUBA	Transport User Benefit Appraisal (DfT software)
UA	Unitary Authority
VDF	Volume Delay Function (highway assignment)
VoT	Value of Time
WebTAG	DfT's web-based documentation for Transport Appraisal Guidance

1 Introduction

1.1 Background

1.1.1 The PLANET Framework Model or PFM has been developed by HS2 Ltd as a tool to forecast the demand and benefits of HS2. PFM has been subject to an on-going programme that uses industry best practice and guidance to establish a "fit for purpose" model. The current version of PFM is known as version 7.1 or PFMv7.1.

1.1.2 The aim of this document is to provide a clear account of the PFMv7.1 model. It is a minor update of the corresponding document relating to v6.1, which itself had updated the earlier v4.3 description. This had brought together a large amount of information which had previously been separately reported, often only in working notes, and therefore constituted an authoritative description of the model at a particular stage in its development.

1.1.3 The document does not describe the evolution of the model, but it does discuss in detail those aspects of the model which are unique to PFM. It also attempts to establish a consistent mathematical notation across the disparate elements of the model and to provide a description of all the key calculations.

1.1.4 PFMv7.1 has been used to generate what is known as the HS2 reference case as presented in the Economic Case released in Summer 2017. The current document does not present the forecasting assumptions used for this purpose, nor the outcomes, though the methodology both for forecasting and appraisal is described.

1.1.5 Further explanation of the forecasting assumptions and how the PFM model has been developed can be found in the following documents;

- PFMv7.1 Assumptions Report;

- PFMv7.1 Forecasting report; and

- PFMv7.1 Release Notes.

Part 1: An Overview of PFM

This part contains a single chapter that provides a high level overview and introduction to the model.

2 An overview of PFM

2.1 Introduction

2.1.1 This chapter provides a general overview of the PLANET Framework Model. The general dimensions of the model are presented first, and then the three main model constituents – PLANET Long Distance [PLD], the Regional PLANETs, and the ancillary Heathrow access model – in outline. Each of these will be discussed in detail in subsequent chapters. The chapter ends with a description of the generalised cost specification used in the various model components.

2.2 Model overview

2.2.1 The PLANET Framework Model [PFM] has evolved over a number of years, and builds on existing model components. Its aim is to provide forecasts of demand and (generalised) costs to drive the appraisal of HS2. As is standard in transport modelling, 'generalised cost' is a combination of monetary costs and travel time components (see box below –Demand and Supply in Transport Models – for more detail).

2.2.2 The crucial elements of the forecasting process may be described as follows:

- derivation of base year (FY[1] 2014/2015) demand patterns for rail, road and highway;

- growth of base year demand to future year(s); and

- demand response to changes in the provision of future rail services, including both new High Speed Rail (HSR) services and/or changes to rail services on the existing or 'classic network'.

2.2.3 The main emphasis of this report is on the third element, though the critical aspects of the other two elements will also be discussed. The derivation of base year demand is described in Chapter 9, while the growth in future year demand is described in Chapter 11.

2.2.4 It may be expected that demand for HS2 will be a mixture of demand transferring ('abstracted') from other modes, and partly additional ('generated') demand. This means that PFM needs to represent the supply and demand for those other modes which may transfer to HS2.

2.2.5 PFM is conventional in the sense that it contains a multi-modal supply representation based on networks, together with a multi-modal demand model which is segmented by different types of travel, and responds to changes in generalised cost.

---

[1] FY is Financial Year, ending 31 March

2.2.6   Due to the nature of HS2, the emphasis of PFM is mainly on representing longer distance travel. The main area where shorter distance travel needs to be represented is on the rail side. This is in order to reflect the fact that short distance passengers can cause crowding on long distance trains, and also to be able to model the benefits that arise from running additional local services. . This is done by making use of existing regional network models with the generic title of 'PLANET'. Note that Northern Ireland is excluded and as such PFM only models domestic trips within mainland Great Britain.

Demand and Supply in Transport Models

The notions of demand and supply are fundamental to economic theory, but although the terms are indeed widely used within the field of transport economics, there are certain aspects of the transport problem which require that they, and the related concept of an equilibrium system, be defined with some care. In classical economics it is conventional to treat both supply and demand as functions of (monetary) cost. Since, in addition to costing money, travelling between different locations inevitably involves an expenditure of time, it has become standard in transport economics to deal with so-called 'generalised cost', which explicitly recognizes both kinds of expenditure. In its simplest form, generalised cost is a linear combination of cost and (different kinds of) time, the latter being converted to money units by means of the so-called 'value of travel time savings'.

The notion that demand for travel T is a function of (generalised) cost C presents no difficulties. However, if the predicted travel demand were actually realized, the generalised cost might not stay constant. This is where the 'supply' model comes in. The classical approach defines the supply curve as giving the quantity T which would be produced, given a market price C. However, while certain aspects of the supply function do relate to the cost of providing services (whether it be the cost of highway infrastructure or a public transport service with a specified schedule), the focus of supply relationships in transport has very often been on the non-monetary items, and on time in particular. This is because many of the issues of demand with which transport analysts are concerned impinge on the performance of the transport system rather than on the monetary costs.

It is therefore more straightforward to use the inverse relationship, whereby C is the unit (generalised) cost associated with meeting a demand T. In this sense, the supply function encapsulates both the response of supplying 'agencies' and the performance of the system. Note therefore the different 'directionality' of the two functions: for demand, the direction is from cost to quantity, whereas for supply the direction is from quantity to cost.

The supply model thus reflects the response of the transport system to a given level of demand. In particular, what would the generalised cost be if the estimated demand were "loaded" onto the system? The most well-known 'supply' effect is the deterioration in highway speeds, as traffic volumes rise. However, there are a number of other important effects, such as the effects of congestion on bus operation, overcrowding on rail modes and increased parking problems as demand approaches capacity. Since both demand and supply curves relate volume of travel with generalised cost, the actual volume of travel must be where the two curves cross – this is known as the 'equilibrium point'. A model with the property that the demand for travel be consistent with the network performance and other supply effects in servicing that level of demand is often referred to as an 'equilibrium model'.

2.2.7   The main structure of PFM can be envisaged as:

- a multi-modal model of supply and demand for long-distance movements ('PLANET Long Distance' [PLD]), and

- a set of local rail models ('regional PLANETs') which deal with the supply implications of the changes brought about by, and associated with, HS2.

In addition, there is an ancillary model which deals with the possible use of HS2 for accessing Heathrow as the first part of an international air trip.

2.2.8   Outside of the PFM model - but included within the appraisal of HS2 - there are separate models for valuing the benefits of running international trains on HS2 to Europe, the noise impacts of trains running HS2 trains on the new high speed line, and the carbon impacts of reduced changes in car and diesel train use. In addition there is a separate calculation of wider economic benefits. None of these modules are described in this document, but are instead separately reported.

2.2.9   PFM provides an overall 'framework' linking the components in a consistent way, as well as managing the interaction between supply and demand. The main components are distinguished in figure 1-1.

Figure 1-1: Planet Framework Models (PFM)

Planet Framework Models (PFM)

Planet Long Distance (PLD)

Multi – Modal Demand and Supply (Includes SCM)

Regional Planets

Heathrow Access Model
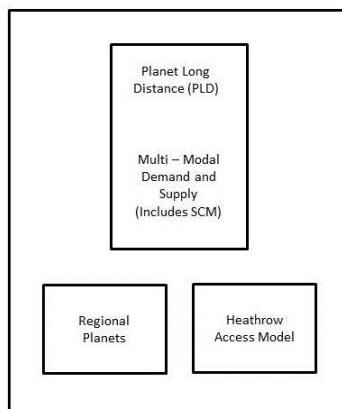
2.2.10   The PLD demand model deals with three modes of travel: rail, car (either driver or passenger) and air. The coach mode is not explicitly recognised, and is generally considered of low relevance in terms of possible abstraction by HS2, since in most cases coach travellers have explicitly opted for a slower (and cheaper) service. Insofar as such abstraction might occur, it is implicitly treated as part of generation.

2.2.11   Apart from the service characteristics, no special recognition is given to HS2 services. The choice between HS2 and 'classic rail' services is made as part of the general route choice in the PLD assignment model. The introduction of HS2 leads to improvements for rail in general and increasing overall demand for rail: that portion of rail demand using HS2 services is obtained from the assignment.

2.2.12   In order to carry out an appraisal of a particular HS2 proposal, PFM needs to be run four times: with HS2 and without HS2, for two modelled (forecast) years. This produces the required output matrices of demand and generalised cost which feed into the appraisal process (discussed in chapter 13). Both the model and the appraisal are, with minor variations, 'WebTAG[2]-compliant'.

2.2.13   For the 'without HS2' (Do-Minimum) runs, the procedure is essentially one of generating costs on the Do-Minimum networks for a fixed level of Do-Minimum demand relating to the year of operation: this is largely the domain of the supply models. The full demand-supply equilibrium PFM process is used for the 'with HS2' (Do Something) runs, together with associated changes in the pattern of classic rail services. This leads to changes in the rail costs, which in turn operate through the demand model to bring about modal shift and generation. This new pattern of demand then leads to further ('second round') changes in cost – largely on the rail network, but to a small extent influencing highway congestion as well. The system is iterated a number of times.

2.2.14   Separate networks are maintained for the rail, car and air modes, and these are input into the corresponding assignment procedures. The assignment model takes a modal matrix of travel (as movements from origins to destinations) and assigns (or 'loads') it onto an appropriate network. While the underlying principles are not mode specific, the different characteristics of highway and public transport networks lead in practice to a rather different set of practical problems.

2.2.15   Although assignment is treated as a single 'stage' in the conventional transport model, it in fact relates to a number of separate processes which may be described as:

- choice of route (or path) for each Origin [O]–Destination [D] combination;
- aggregating O–D flows on the links of the chosen paths;
- dealing with supply-side effects (capacity restraint) as a result of the volume of link flows relative to capacity; and
- obtaining the resulting cost for each O–D combination.

2.2.16   The route-choice process could be considered as a component of demand, but because of its treatment within the assignment it is typically regarded as part of the 'supply' procedures. It should be noted that it is strictly only the implementation of capacity restraint that represents the true supply function. The last process is often referred to as 'skimming' the network.

2.2.17   On the highway side, as only the PLD (long distance) car matrices are assigned, short distance car traffic and freight is represented by 'pre-loads' (i.e. traffic which is not included in the matrices to be assigned, but is independently estimated for each link in the network, and assumed to be fixed). Standard equilibrium assignment techniques are employed, and separate user classes are recognised.

---

[2] WebTAG is the Department for Transport's web-based documentation for Transport Appraisal Guidance - see http://www.dft.gov.uk/webtag/documents/index.php

2.2.18   By its nature, public transport assignment is considerably more complicated. For the regional PLANETs, the assignment needs to allow for different access and egress options to and from stations, in addition to the variations between rail services in terms of frequency, in-vehicle time and interchange requirements.

2.2.19   For the long-distance movements in PLD, station access is separately modelled by means of a 'station choice model' [SCM]. In this case, therefore, the assignment receives a matrix of station-to-station movements (as opposed to zone-to-zone movements). This results in a more complex version of supply-demand iteration.

2.2.20   By contrast, the air network does not in fact represent any supply constraints: aircraft are not assumed to have any capacity constraint (on the basis that airlines could just run bigger aircraft) and thus, fare and frequency of service do not vary with demand level. The network's main function is to provide matrices of cost.

2.2.21   PFM generally runs on the specialist transport modelling EMME/4[3] software platform, making use of standard procedures and macros. However, the SCM has been developed as a separate module in the C++ language.

2.3   Zoning system, demand segments and scope

2.3.1   The basis of the PLD zoning system is the Local Authority District level, of which there are 406 in mainland Great Britain. To focus on the corridor of interest, aggregation has taken place in the more remote areas, so that the number of zones is reduced to 235 (including a separate zone for Heathrow airport). The zoning system is shown in figure 1-2.

2.3.2   The size of the PLD zones means that much of the total travel made within Great Britain is intra-zonal, and therefore not captured by the model. The emphasis of the PLD model is on long-distance movement.

2.3.3   The PLD demand matrix and network assumptions represent an average weekday. As described in Chapter 12, for appraisal purposes, this average weekday demand is converted into an annual demand using 'annualisation factors', to allow for weekend and holiday travel. The regional PLANETs are AM peak models only, relating to the three-hour period [0700-1000] only. Hence an adjustment is required when interfacing demand between the regional peak and PLD all day models.

---

[3] http://www.inro.ca/en/products/emme/

Figure 1-2: PLD Zoning System
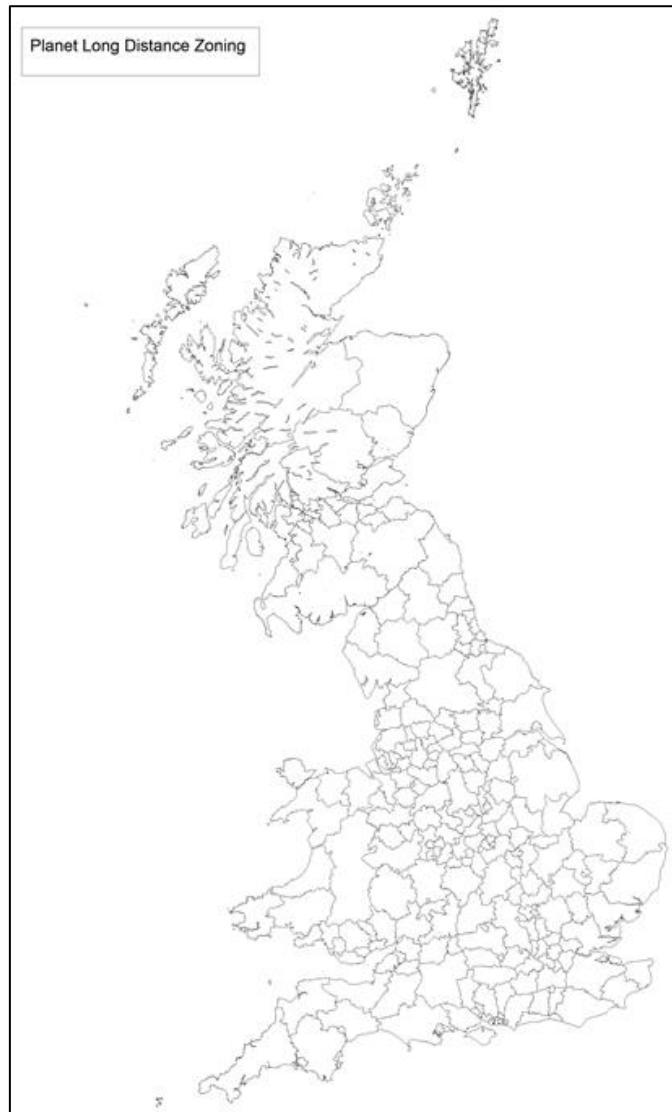
2.3.4   The impact of local rail demand is modelled in the regional PLANETs, while that of long distance demand is modelled in PLD. As described in Section 2.5, the regional PLANETs have a more detailed zoning system in their core areas of interest, generally based on aggregations of Census Output Areas [COAs]. To ensure that no elements of demand appear in two models, PFM makes use of a 'control matrix', which defines which origin-destination movements are held in which of the models. This in turn requires appropriate interfaces to ensure that crowding levels are represented correctly in each of the models. Without further adjustment this could lead to crowding being understated on long distance services in PLD south of Milton Keynes.

2.3.5   Three purposes are recognised in the modelling: Business, 'Other' (essentially Leisure) and Commuting. The definitions, which are in line with NTS definitions for trip purpose[4], are as follows:

---

[4] NTS codes the purposes at both ends of the trip: from this, a "trip purpose" can be deduced. The NTS definitions of Commuting and Business are identical to those given above, but NTS makes further distinctions within the "Other" purpose category.

- Commuting: Trips between Home and Usual Workplace only;

- Business: Trips between Home and 'In course of work', and non-home-based trips with destination in course of work; and

- Other: all other trips, including education.

2.3.6   These trips are further segmented between car available [CA] and non-car available [NCA], though the distinction does not apply in all cases.

2.3.7   In most transport models, the matrices and the model operate on a 'Production/ Attraction' [P/A] basis, distinguishing between home-based trips and non-home-based trips, where the latter are treated on an 'Origin-Destination' [O-D basis][5]. One reason for this is so that the forecasting process can distinguish between factors affecting growth at the home end of the trip (including car availability) and at the 'destination' end of the trip. In addition, this distinction is essential when destination choice is one of the possible responses.

2.3.8   PFM does not explicitly allow for destination choice. This is largely due to the structure of the PFM model and the associated constraints this imposes. Instead, destination choice is reflected indirectly within the 'generation' component of the demand model.

2.3.9   As this removes the most critical reason for moving to a P/A structure, PFM therefore operates on an O-D basis which makes the implementation of the model somewhat simpler. As discussed in Chapter 11, the forecasting procedure is however adjusted to take account of the P/A-based growth. The differential treatment of car availability according to the location of the home is also taken into account at various stages in the model, particularly in connection with the SCM. This is another way in which the advantages of a P/A structure are built into the current model, by distinguishing the direction (from home, to home) in which the journey is being made.

2.3.10   To summarise, the PLD model component operates on a 235 zone level, representing district level spatial resolution in the key areas served by HS2. It models an average weekday, distinguishing three purposes, further segmented by car availability, and deals with three modes. The model operates on an OD basis, with some adjustments to reflect P/A format, and deals with assignment (including station choice), and mode choice and generation as demand responses.

2.3.11   By contrast, the Regional PLANETs (separate models for South, Midlands and North) operate for rail only, and only allow for demand responses by means of elasticities to changes in rail services (though changes in PLD rail demand are conveyed by interface). They are also weekday models, but relate to the AM peak only.

2.3.12   Finally, the 'international' component operates on the same zoning system as PLD, and relates to accessing international air trips to and from Heathrow (Heathrow Access Model [HAM]).

---

[5] The differentiation between PA and OD development is discussed further in the box below.

Alternative formats for transport demand matrices

There are two alternative formats for transport demand matrices, and the distinction is of major importance.

The Origin-Destination format relates to trips starting in zone i and ending in zone j, while the Production/Attraction format relates to trips "produced" in zone i and "attracted" to zone j.

An alternative formulation is the "tour", which is a chain of linked trips beginning and ending at the zone of production. In most cases the zone of production is taken as the zone of residence, though some work-related journeys can be "produced" from the zone of workplace.

To see the difference, consider the following simple 2 zone example. Zone 1 has 10,000 residents who all work in zone 2, and zone 2 has 2000 residents who all work in zone 1. Each person travels once to work and back in a day. The total daily volume of travel can thus be represented as:

	Attraction zone					Destination zone		
		1	2	Total			1	2	Total
Production zone	1		20,000		Origin zone	1		12,000	
	2	4,000				2	12,000		
	Total			24,000		Total			24,000

O-D matrices when taken over a whole day tend to be symmetric. This is not true of P/A matrices. The totals are the same, but the distribution over cells is quite different.

The full range of demand responses (in particular, the modelling of destination choice) cannot be sensibly modelled on an O-D basis. The matrix structure must be on a P/A basis. A minor exception may be made when modelling the morning peak only, when most journeys will be from home. In addition, when forecasting growth over time, to take account of changes in residences, employment etc., the growth factors need to be applied to P/A matrices.

However, when considering the impact on the network, they are transformed to an O-D basis. This involves recognising the separate directions (outward and return) in the P/A format. The "return" portion of the matrix is transposed and added to the "outward" portion.

2.4 PLANET long distance model components [PLD]

2.4.1 The PLANET long distance [PLD] model is the core of PFM and is, in effect, a supply-demand equilibrium model for long-distance travel focussed on the HS2 market. An outline of the PLD model structure is shown in figure 1.3. The figure simplifies by ignoring the different demand 'segments' (purpose and car availability). In addition there is an interaction between PLD and the HAM – not shown in the figure.

Figure 1-3: PLD Model
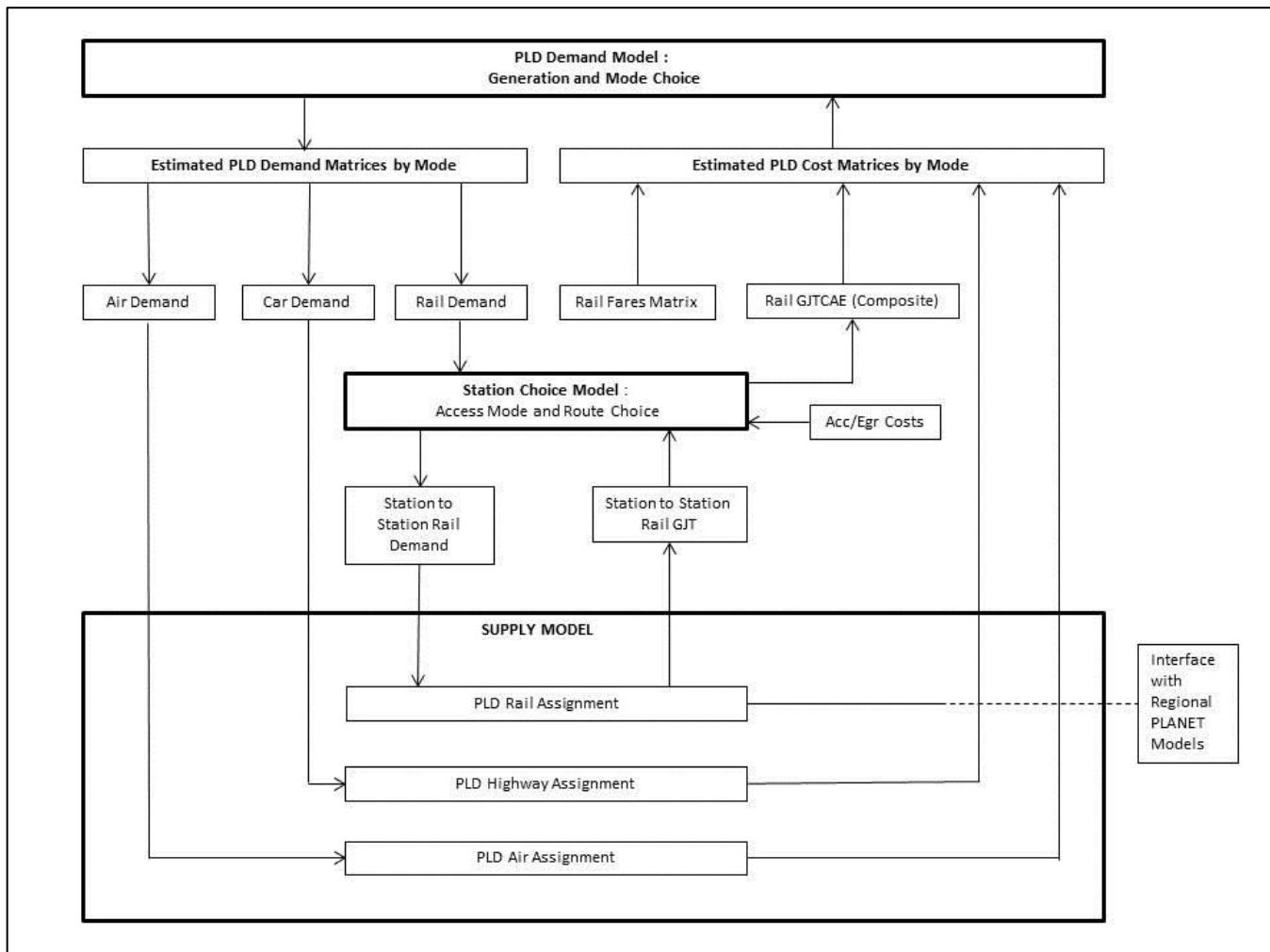
2.4.2 At the top of the structure is the Demand model, which deals with mode choice and generated demand, in response to changes in (generalised) cost (and specifically those brought about by the introduction of HS2). At the bottom of the structure are the modal Supply models, based on networks: in principle these calculate the changes in cost, as a result of changed demand. For rail there is an intermediate Station Choice Model [SCM] which converts the rail demand matrices from a zone-to-zone basis to a station-to-station basis. The rail assignment then calculates the routes through the network for each pair of stations, allocates the demand and derives the costs (there is also an interface with the Regional PLANETs). The costs feed back into the demand model and the process iterates until a stable result is obtained.

2.4.3   As is general practice within public transport assignment, the choice of routes within the rail assignment does not make allowance for differential fares. PFM has not been designed to take into account the different fare options generally available to passengers for travelling between an origin and destination using different routes, Train Operating Companies [TOCs] or types of service. It therefore assumes that the cost of travelling by HS2 is the same as travelling by any other 'classic' rail service.[6]

2.4.4   On this basis, the choice of route reflects only the service characteristics of the route (in-vehicle time – including an allowance for crowding, service frequency/waiting, and the need for interchange). The cost output of the PLD rail assignment is in terms of 'generalised journey time' [GJT][7], but includes an allowance for crowding, so it is denoted as GJTC.

2.4.5   The SCM then introduces the access and egress generalised costs to and from each station, and allows also for choice of access mode (between car and public transport). Because a probabilistic allocation to stations is made in the SCM, the output is in so-called 'composite' GJTCAE [GJTC plus Access and Egress] terms, reflecting the cost associated with all possible station pairs. To convert to generalised cost, the fares – divided by the Value of Time – are added prior to returning to the demand model. More discussion is provided in Chapter 5.

2.4.6   The (revised) estimates of zone-to-zone rail demand (from the Demand Model) then pick up the station choices obtained from the SCM, thus converting the rail demand to a station-to-station basis.

2.4.7   As noted, allowance is also made for a two-way interface in the rail assignment between the PLD assignment and the Regional PLANETs. This ensures that, for the key 'common' services between PLD and the regional models, the total demand from both long-distance and regional passengers is accurately reflected, ensuring that demand and crowding is correctly assessed in both models.

2.4.8   To achieve this, the demand from relevant short distance rail travellers is transferred as 'pre-loads' from the appropriate regional PLANETs to the PLD assignment. In the opposite direction, for PLANET North [PN] and PLANET Midlands [PM], demand arising from long-distance travellers is again transferred in the form of pre-loads, from PLD to the assignment. For PLANET South [PS], PLD demand is transferred by a device known as 'wormholes', both to represent crowding appropriately in PS and to deal with more detailed access/egress options within the local rail system.

---

[6] The possibility of differential fares for HS2 is modelled outside PFM, in the Commercial Modelling workstream
[7] GJT is a term widely used in the rail industry, and in PDFH in particular (see Section B4.1 of PDFH5). Unlike the term "generalised cost" which can contain any number of variables depending on the context under investigation, GJT has a very specific meaning in the context of PDFH and contains only the three timetable related service quality attributes: in-vehicle time (strictly speaking, not including allowance for crowding), an allowance for service frequency (effectively, waiting time), and interchange (both as a "penalty" expressed in minutes, and any time taken to make the change). Note that it does not include access and egress costs.

2.5   Regional PLANETs

2.5.1   Separate network models are used for local movements in different parts of the country, referred to as PLANET South [PS], Midlands [PM], and North [PN]. These are versions of free-standing rail models maintained by the Department for Transport [DfT]. PLANET South [PS] is the oldest and has been used to assess a wide variety of schemes including the Thameslink Upgrade, while the other two were developed more recently. The three models operate along generally comparable principles, though there are some minor differences in the assumptions they make. Their coverage is indicated in figures 1-4, 1-5 and 1-5.

2.5.2   The primary aim of the Regional PLANETs is to estimate the savings in GJTC associated with new schemes taking advantage of capacity released by HS2 and to ensure that local movements are correctly reflected in the PLD assignment. In addition, they can reflect local crowding benefits associated with classic rail passengers switching to HS2.

Figure 1-4: PLANET North Zoning

Planet North Zoning
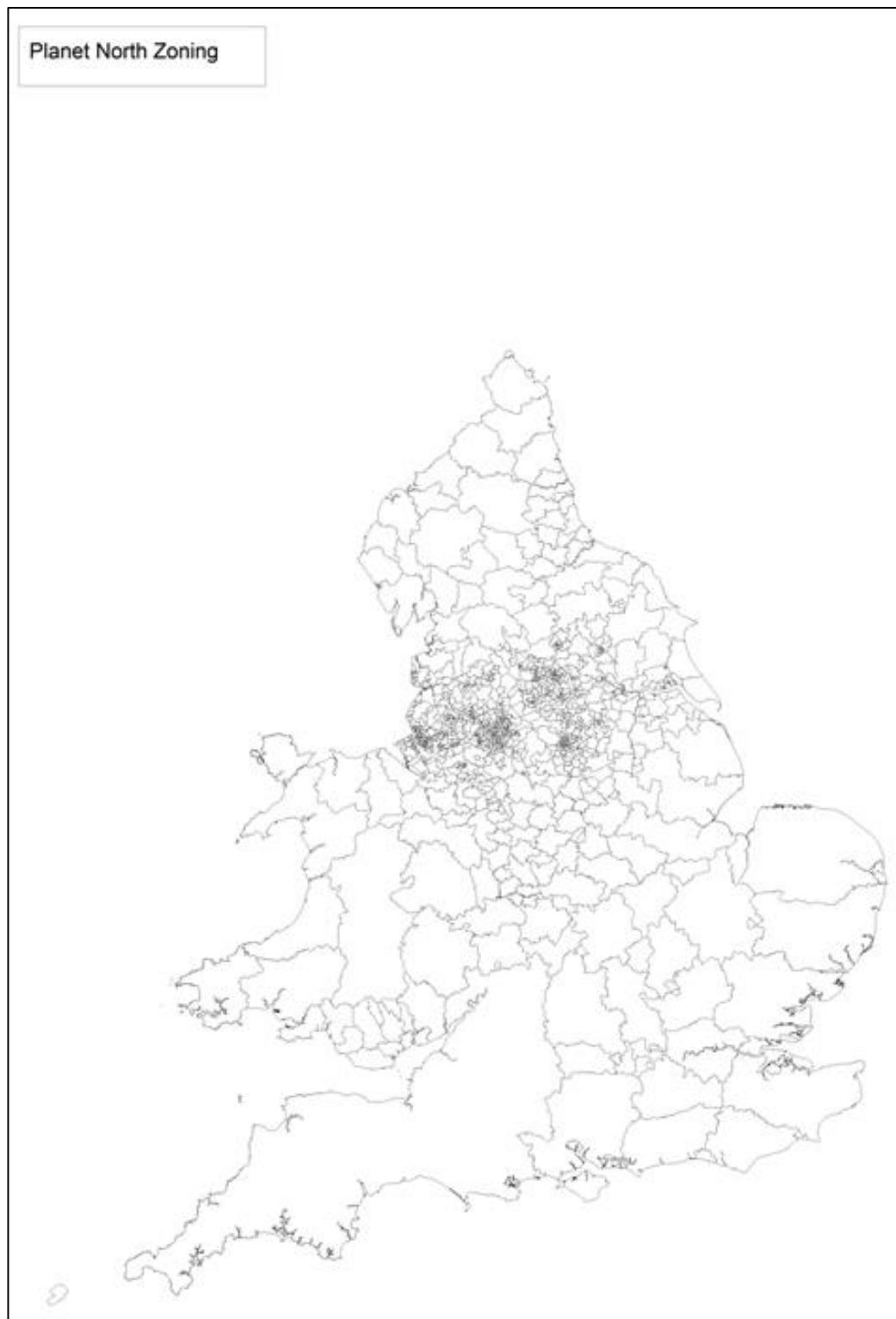
Figure 1-5: PLANET Midlands Zoning

Planet Midlands Zoning
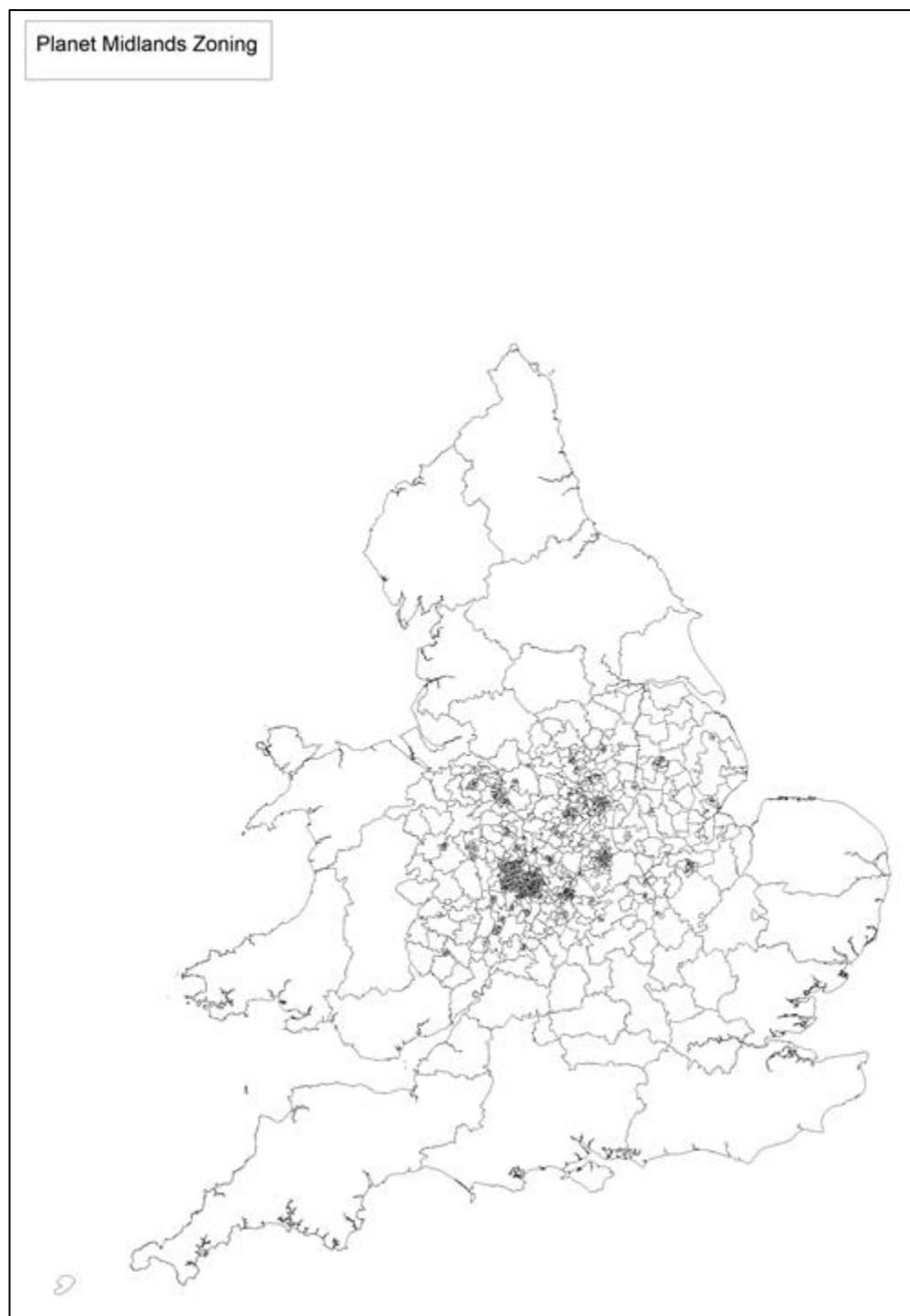
Figure 1-6: PLANET South Zoning

2.6 Other model components

2.6.1 Both PLD and the Regional PLANETs are only concerned with 'domestic' movements between zones within Great Britain. An ancillary model – the Heathrow Access Model [HAM] – has been developed to cater for the mode choice of travellers to and from Heathrow (PLD Zone 90).

2.6.2   The HAM is a bespoke spreadsheet model which derives from the London Airports Surface Access Model [LASAM] but additionally includes the following access mode: Domestic air interlining – mainly between Heathrow and the airports at Manchester, Newcastle, Edinburgh and Glasgow.

2.6.3   Journeys are segmented not only by purpose (Commuting and Other are combined), but also by UK versus foreign resident, in recognition of the interaction with car availability for mode choice. The model only includes flows that could realistically be abstracted by HS2, based on the London – Birmingham – North West – Scotland corridor.

2.6.4   Feeding into the appraisal of HS2 are additional models assessing wider economic benefits, the benefits of running international trains on the HS1-HS2 link and certain environmental impacts. None of these models form part of the PFM framework and are therefore not described in this report.

2.7   Consistency of generalised cost specification

2.7.1   As was noted in the box at the beginning of this chapter, the Demand model makes use of the concept of generalised cost, and in most cases the individual elements are produced from network models. In addition, on the rail side, the SCM allocates demand to stations based on generalised cost[8]. It is important that, as far as possible, the definitions are consistent between all modules in PFM. Note this is not referring to the generalised costs used in appraisal, though these will be discussed in chapter 13.

2.7.2   All references to 'generalised cost' assume that it is measured in units of in-vehicle time minutes. Thus generalised cost is defined as: In-vehicle time + $element_a$_* $weight_a$ + $element_b$ * $weight_b$ + ... etc. For example, $element_a$ might be (expected) waiting time, in which case $weight_a$ would represent the equivalent of one minute of waiting time in in-vehicle time units. The actual values are discussed below.

2.7.3   WebTAG [UnitM2, paragraph 6.5.8] strongly advises that the generalised costs used in both assignment and demand models should be compatible, and the development of PFM has aimed to reduce any inconsistencies to a minimum.

2.7.4   On the highway side, generalised cost relates essentially to in-vehicle time and to monetary costs of travel – these are fuel costs, other operating costs, and 'user charges' such as tolls (and parking). Monetary costs are converted to time units by dividing by the 'value of time' [VoT], which may vary with the demand segment. Generalised cost then forms the basis of route choice (within the Highway Assignment) and mode choice (within the demand model). The assignment can impact on zone-to-zone in-vehicle times (via congestion effects) which in turn influence route choice, but the generalised cost weights are consistent between these two models. Over time, generalised costs can change, both to reflect growth in VoT and to reflect other price changes (eg fuel costs).

---

[8] though, as will be seen, rail fares are not included in this instance.

2.7.5   On the public transport side (i.e. both rail and air), the specification is more complex. In addition to in-vehicle time (and fares, though these are only used in the demand model), the following need to be taken account of:

- access and egress time;

- crowding penalties

- walking time (e.g. between connecting services);

- waiting time; and

- interchange.

2.7.6   There is some potential confusion in the way these terms are used in different component models. The main point is that each element of generalised cost needs to be represented in units of in-vehicle time, and the multipliers applied to the elements are referred to as 'element weights'. For example, one minute of waiting time may be considered equivalent to 2 minutes of in-vehicle time, leading to a waiting time weight of 2.0.

2.7.7   In the case of interchange, each interchange is represented by a time penalty[9]. (Expected) Waiting time is derived as a factor applied to the service interval, and is then further weighted to represent the disutility of having to wait.

2.7.8   For air, the assignment is not sensitive to levels of demand, and its main purpose is to provide the appropriate network costs [in-vehicle time, frequency, wait time, access and egress time and fare 'skims' separately for business and other purposes]. In fact, as shown in table 1-1 below, some high element weights have been used in the assignment to ensure that sensible routes are chosen, in terms of the airport pairs used for any zone-to-zone movement. Element weights are also shown for the HAM, which considers access trips to international flights from Heathrow. There is no reason to expect or require consistency between the element weights here, as the HAM is addressing a different market from the main PFM model.

2.7.9   It should also be noted that element weights for the air mode will only affect the HS2 demand and benefits in situations where the air service level changes between the Do Minimum and the Do Something Scenario. For the assessment of HS2 only rail services are assumed to change between Do Minimum and Do Something Scenario and hence, the discrepancy in generalised cost definition between the assignment and demand model for the air mode is of no practical significance.

---

[9] The convention in the PLANET models is to apply a 'boarding penalty' each time a train is boarded: this means that one more penalty than the actual number of interchanges is included in the GJT. However, this has no impact on the route chosen, nor on the demand model or appraisal. The same convention is maintained for the air assignment.

Table 2-1: Generalised cost element weights and other parameters for air in the different models

Air	Assignment model	Demand model	Heathrow Airport model (UK business)	Heathrow Airport model (UK leisure)
IVT	1.00	1.00	1.00	1.00
Wait time	2.00	1.00	2.72	2.75
Access/Egress time	10.00	2.00	3.06	4.80
Board time penalty (minutes)	163.00	n/a	n/a	n/a
Board time factor	2.00	n/a	n/a	n/a
Wait time factor applied to service interval	0.10	0.40	0.10	0.10
Total factor applied to service interval	0.20	0.40	0.27	0.28

2.7.10　For rail, consistency between the various models is of much greater significance and is discussed in more detail in Chapters 4-6. For PLD, this has been ensured by:

- selecting the assignment model element weights appropriately;

- maintaining the same individual element weights in the SCM; and

- conveying the 'composite' cost over all the station choice alternatives to the demand model.

2.7.11　There remain some minor inconsistencies between PLD and the Regional PLANETs, as can be seen in table 1.2. The regional models have a different wait time factor and a different boarding penalty, with Planet South having a boarding penalty of 3.5 minutes and the other regional models 20 minutes. The PLD assignment model has a boarding penalty of 30 minutes which is, as should be expected, greater for the less frequent travellers on strategic services. It was decided not to alter the values in the regional models, as these reflect different markets to the PLD demand model, and such large differences are also seen in PDFH[10].

2.7.12　The general pattern of recommendations in PDFH is to allow many of the element weights (interchange, walking and waiting time) to vary with overall journey distance: in addition, the treatment of interchange allows for explicit coding of the waiting time for the connecting service. This approach is not really suitable for PLANET network modelling where only the frequency of connecting services is available, and where varying element weights by distance would be a major computational problem[11]. For this reason, the values in the free-standing Regional PLANETs have been accepted as having independent validity.

---

[10] There would also be knock-on consequences for the demand elasticities in the Regional PLANETs.
[11] PLANET assignment does not store path information, so there is no reasonable means of ascertaining the ultimate trip origin when considering passengers boarding a service after interchanging. This applies both to PLD and the Regional PLANETs

2.7.13   The wait time factor converts the service interval to the average (expected) waiting time. The factor that is applied in the regional models is 0.5, suggesting random arrivals[12]. This is reduced to 0.4 in the PLD assignment, to acknowledge some of the large service intervals in the model and the fact that passengers will therefore plan their long distance journeys more, rather than arriving at random; again this is consistent with PDFH. Note that having calculated the expected waiting time, it is further weighted to reflect its greater disutility relative to in-vehicle time.

2.7.14   Other than these two differences the regional models and the assignment models have the same set of parameters. It should be noted that these values do not differ by journey purpose: this again is consistent with PDFH.

Table 2-2: Generalised Cost Element Weights and Other Parameters for Rail in the Different Assignment Models

Rail	Planet long distance	Planet South	Planet North	Planet Midlands
IVT	1.0	1.0	1.0	1.0
Wait time	2.0	2.0	2.0	2.0
Walk time (for connections)	2.0	2.0	2.0	2.0
Access/Egress time*	1.0	2.0	1.0	1.0
Board time penalty (minutes)	30.0	3.5	20.0	20.0
Wait time factor applied to service interval	0.4	0.5	0.5	0.5
Total factor applied to service interval	0.8	1.0	1.0	1.0

*As will be seen, Access/Egress time is dealt with outside the assignment model except in the case of PS.

2.7.15   For the rail modes, fares are only introduced at the Demand model level: implicitly it is assumed that the choice of route between stations, as well as the station choice itself, is not influenced by considerations of fare. As noted, PFM has not been designed to deal with the possibility of 'premium fares' for HS2 or any other services.

2.7.16   With the exception of these minor issues noted, a consistent specification of generalised cost has been used throughout. This applies also to values of time [VoT], used to convert money costs to IVT units. As far as possible, the individual elements are maintained separately, and conveyed to the appraisal procedure.

---

[12] If passengers arrive uniformly in the interval (say, 30 minutes) between two services, then some will just miss the earlier train and wait nearly 30 minutes, while some will arrive just in time for the later train and wait almost 0 minutes, but on average the passengers will wait half the service interval (ie 15 minutes). Thus the required factor is one-half.

2.7.17   As was noted in Section 2.4, the term 'generalised journey time' [GJT] is reserved for the rail mode for the particular combination of IVT, wait time (based on headway) and interchange. It does not include the additional disutility associated with crowding, though the assignment makes use of this as well as GJT. The SCM then introduces further elements relating to access and egress. It is only when the fares are introduced at the demand model stage that a true 'generalised cost' is achieved. All these generalised cost elements - both separately and in combination - are measured in minutes.

Part 2: Long Distance Model (PLD)

This part of the document describes the core component of PFM, known as PLANET Long Distance [PLD]. There are a number of different components to PLD which are described in separate chapters.

Chapter 3 introduces PLD and discusses key theoretical aspects and notation. The three main components of PLD – the rail assignment, the Station Choice Model [SCM] and the Demand model are then discussed in chapters 4 to 6. Although PLD is a multi-modal model, most of the emphasis of this report is on the rail mode as this is key to the assessment of High Speed Rail.

The model inputs, in terms of demand matrices and networks, are not described here, but are described separately in Part 3 alongside those used for the regional PLANET models.

3 Introduction to PLD

3.1 Overview

3.1.1 The overall structure of PLD was indicated in figure 1-3. It can be seen that there are three main components: the Supply Model, the Demand Model, and the Station Choice Model which acts as an intermediary step for the rail mode. The purpose of the key PLD modules is to model behavioural responses to changes in rail services in terms of route choice (assignment), station choice (SCM) and mode and frequency (Demand model).

3.1.2 The various modules (the PLD rail assignment, the SCM, the Demand Model and the Highway and Air[13] Assignment) form part of an iterative system, which also includes the Regional PLANETs and the Heathrow Access Module (HAM), which are the subjects of chapters 7 and 8. For this reason, a full account of the iterative algorithm is postponed till chapters 10 and 11.

3.1.3 The supply model consists of a network and an assignment routine. These operate independently for the three modes: rail, car and air. The highway network is conventional, and the details will be described later, in chapter 7. The rail and air networks are also conventional EMME public transport networks, and their details will likewise be described in chapter 9.

3.1.4 As far as assignment is concerned, the key discussion relates to the rail mode, and it is described in some detail in chapter 4. For the other two modes, it is convenient to describe the assignment more briefly, in connection with the networks in chapter 9.

3.1.5 The Station Choice Model [SCM], described in chapter 5, is the link (on the rail side) between supply and demand. This converts the matrices of zone-to-zone rail demand, output from the Demand Model, to matrices of station-to-station demand, taking account of access and egress costs. It also deals with the effect of car availability on the choice of access mode.

3.1.6 Finally, the multi-modal Demand model, which deals with modal shift and generation, is described in chapter 6. As will be discussed there, the PLD Demand model is an 'incremental' model, so that it requires a set of (base) Matrices from which to 'pivot'. In chapter 7, it was described how the separate modal (rail, car, air) Base Year Matrices, segmented by purpose and car availability, have been obtained. When running the model for future years, the pivot matrices need to take account of demand growth over time, as discussed in chapter 10. For both rail and car, these future pivot matrices are obtained by applying growth to the Base Year Matrices: the air pivot matrices are obtained independently from the DfT Aviation model (see Section 9.4). The resulting pivot matrices represent the Do-Minimum forecasts.

---

[13] as noted earlier, it is not strictly necessary that the Air Assignment is repeated.

3.1.7　The SCM and the Demand model have been specially estimated for use in PFM, and both the estimation and application of the models is described in detail. While the rail assignment makes use of available routines in the EMME/4 public transport software suite, some of the aspects are complex. Hence, much of the description in chapters 4 to 6 is of a technical nature.

3.1.8　For this reason, the final sections of this chapter set out a notation which will be consistently used for the technical description, together with a discussion of the logit (discrete choice) model, which is the model form which underlies both the SCM and the demand model.

3.1.9　Finally, an important feature of PLD is that it uses a consistent definition of generalised cost throughout.

3.2　Consistent notation

3.2.1　In devising an appropriate notation for the mathematical description of the PLD model, there are three types of item that need to be distinguished:

- quantities, such as travel demand, cost, population;
- arguments/categories, such as zones, stations, purpose, car availability; and
- parameters, such as generalised cost element weights, values of time.

3.2.2　While it is not possible to be completely consistent, as a general guide quantities are represented by upper case Roman letters, arguments by lower case Roman letters, and parameters by Greek letters. At the end of this section, a table will be set out for reference.

3.2.3　Arguments can be either superscripts or subscripts. Superscripts will be used for fixed categories such as purpose, while subscripts will be used for categories of 'choice', such as zone or mode. An asterisk may be used to denote summation over the relevant argument. To reduce the notational burden, arguments may be omitted when the context is obvious.

3.2.4　The most important symbols are introduced below. Less used symbols are introduced specifically in the text to which they relate.

Quantities

3.2.5　Travel Demand will be denoted by T, and will typically be superscripted by purpose and car availability, and subscripted by zones (usually production and attraction, but sometimes referring to origin and destination) and mode. Pivot or base travel demand is denoted by B.

3.2.6　(Generalised) Cost will be denoted by C, and is measured in time units (minutes). Where necessary, the monetary component is separately denoted by M. The arguments generally reflect those of travel demand. As will be seen in the next section, some of the models are specified in 'utility' rather than cost terms: for this purpose U is used.

3.2.7　Other quantities are Distance (D), Population (P), Workplaces (W) and Income (Y).

Arguments

3.2.8   For fixed categories, p denotes journey purpose and c car availability: in combination these may be represented as 'user classes' and denoted by u. Categories of choice relate to zones (I, J: note that upper case letters are used to distinguish from the lower level 'mzones' used in the SCM, for which i,j are used), modes (m, as well as x for access mode), and stations (R,S).

Parameters

3.2.9   The most important parameters are the 'scale factors' λ on generalised cost (discussed in the next section), which will depend on the 'level' at which the choice is taking place, the structural parameters θ representing the ratio of two λ factors at different levels of choice, other estimated or assumed parameters generally denoted by β (with appropriate description), and α, used to denote the appropriate 'weight' for a generalised cost element. In some cases, specific abbreviations (e.g. VoT for value of time) are used for improve legibility.

3.2.10   The following table 3.1 sets out the symbols used with, where appropriate, an indication of their range (in the case of arguments), and the sections of the report to which they have most relevance. In most cases, upper case letters denote quantities (e.g. demand) and lower case letters denote arguments/subscripts/superscripts (e.g. mode, purpose). Exceptions are the use of I, J and R,S as arguments, and p as a proportion.

Table 3-1: Table of Notation

Symbol	Type *A* = argument *P* = parameter *Q* = quantity	Interpretation	Use	Section
A	*P*	used for crowding calculation	rail assignment	4.2
B	*Q (P)*	base or pivot travel demand (also, parameter used for crowding calculation)	Demand model rail assignment	6.4 4.2
C	*Q (P)*	generalised cost (time units) (also, parameter used for crowding calculation)	Demand model, Appraisal rail assignment	6.2-6.4, 11.9, 13.2 4.2
CP	*P*	crowding penalty	rail assignment	4.2
D	*Q (P)*	distance (also, parameter used for crowding calculation)	Demand model rail assignment	6.2-6.4 4.2
GJT	*Q*	"generalised journey time" (rail)	rail assignment, SCM	2.4, 2.7, general
GJTC	*Q*	GJT including crowding (rail)	rail assignment, SCM	2.4, 2.7, general
GJTCAE	*Q*	GJTC with access/egress costs (rail)	SCM, Demand model	Chapter 5
G°	*Q*	base or pivot generalised cost	Demand model	6.4, 11.9

Symbol	Type **A** = argument **P** = parameter **Q** = quantity	Interpretation	Use	Section
I	A	zone (production end)	general	
J	A	zone (attraction end)	general	
L	A	public transport 'transit line' (or combination of lines)	rail assignment	Chapter 4
LF	Q	loading factor (ratio of passengers to seats)	rail assignment	4.2
M	Q	monetary cost	Demand model	Chapter 6
P	Q	population	SCM, growth in demand	5.4, 11.2
R	A	station (production end)	SCM	Chapter 5
S	A	station (attraction end)	SCM	Chapter 5
T	Q	travel demand	demand model, appraisal	6.4, 13.2
U	Q	utility	SCM, demand model (estimation)	Chapters 5 and 6
V	Q	estimate of current demand in MSA procedure	assignment, supply demand loop	4.4 11.4
VoT	P	Value of time	demand model (SCM)	Chapter 6 (and 5)
W	Q	High level Managerial jobs (also used for estimate of 'auxiliary' demand in MSA procedure)	SCM assignment, supply-demand loop	5.4 4.4, 11.4
Y	Q	income	SCM, demand model	Chapters 5 and 6
Z	Q	Choice set for station pairs	SCM (and Regional PLANETs)	Chapter 5, 7.4
a	A	link	assignment	Chapter 4
c	A	car availability segment (CA, NCA)	SCM, demand model	Chapters 5 and 6, 8.2
d	A	direction (outward, return)	SCM, demand model	Chapters 5 and 6, 8.2
i	A	mzone (production end) within I	SCM	Chapter 5
j	A	mzone (attraction end) within J	SCM	Chapter 5
k	A	generalised cost element	appraisal	13.2
m	A	main mode (rail, car, air)	demand model	Chapter 6
n	A	iteration no.	assignment, supply demand loop	4.4 11.4

Symbol	Type *A* = argument *P* = parameter *Q* = quantity	Interpretation	Use	Section
p	*A (Q)*	purpose (superscript) (commute, business, other); proportion (quantity)	general SCM, demand model	general Chapters 5 and 6
υ	*A*	user class	assignment	Chapter 4
x	*A*	access mode (highway, PT)	SCM	Chapter 5
y	*A*	year	demand model	6.2-6.3, 13.3
α	*P*	lower bound for crowding penalty	rail assignment	4.2
β	*P*	estimated parameter	SCM, demand model	Chapters 5 and 6
δ	*Q*	dummy (0,1) variable	SCM, demand model	Chapters 5 and 6
ε	*Q*	random term in logit model	SCM, demand model	3.3, Chapters 5 and 6
η	*P*	distance elasticity for VoT (also inverse VoT for LASAM)	demand model HAM	Annex B 8.9
θ	*P*	structural parameter for logit hierarchy	SCM, demand model	3.3, Chapters 5 and 6
λ	*P*	coefficient on generalised cost	SCM, demand model	3.3, Chapters 5 and 6
μ	*Q*	average loading factor for rail link	rail assignment	4.2
π	*P*	proportionate allocation of zone between mzones	SCM	5.4-5.5, 13.2
σ	*Q*	standard deviation of loading factor for rail link	rail assignment	4.2
ω	*P*	proportion of auxiliary demand used in MSA averaging	assignment, supply demand loop	4.4 11.4

3.3 Some notes on logit models

3.3.1 There are many texts which establish these properties, but chapter 3 of Kenneth Train's 'Discrete Choice Methods with Simulation' (2002) which is downloadable (free)[14] is particularly accessible.

---

[14] http://elsa.berkeley.edu/books/choice2.html

3.3.2 The logit model is a 'discrete choice' model widely used in transport demand models to allocate demand among different ('discrete') alternatives, such as modes or destinations. In the classic description, the relative appeal of each alternative i, taking into account all its salient characteristics including cost, is denoted as 'utility'. Individuals are assumed to choose the option which has maximum utility. Alternative i is assumed to have a 'representative' utility U_i which is subject to an additive 'random term' ε_i, which allows for those elements in individual choice which are unknown to the modeller. When these random terms for the alternatives are independently and identically distributed ('IID') with a type I extreme value (or Gumbel) distribution[15], the logit model is obtained, whereby the proportion (probability) choosing alternative i is given by:

Equation 3-1

$$p_i = \frac{\exp(U_i)}{\sum_{i'} \exp(U_{i'})}$$

where the summation in the denominator ranges over all the possible alternatives.

3.3.3 In transport models it is conventional to re-specify the model in terms of 'generalised cost' C, rather than utility, yielding:

Equation 3-2

$$p_i = \frac{\exp(-\lambda.C_i)}{\sum_{j \in Z} \exp(-\lambda.C_j)}$$

where λ is a positive 'scale factor', and the negative sign accounts represents costs or 'disutility'.

3.3.4 Given that each individual is choosing the option with maximum utility, it is of interest to see what the expected maximum utility derived from the choice process is, allowing for the random term. This can be shown to be given by the natural logarithm of the denominator in Eq (3-1), in other words:

Equation 3-3

$$U^* = E[\max(U_i + \varepsilon_i)]_i = \ln\left(\sum_{i'} \exp(U_{i'})\right)$$

Because the formula is the logarithm of a summation, it is colloquially referred to as the 'logsum'.

3.3.5 The formula can also be cast in generalised cost terms, where its interpretation is the expected minimum generalised cost derived from the choice process. This is normally referred to as the 'composite cost', given by the equivalent 'logsum' formulation but converted to cost units:

---

[15] see the cited work by Train for an explanation

Equation 3-4

$$C* = \frac{-1}{\lambda}.\ln \sum_{i'} \exp(-\lambda.C_{i'})$$

3.3.6 The composite cost plays an important role in the 'nested (or hierarchical) logit' model, where the choice set is partitioned between different levels. These levels relate to different degrees of sensitivity to generalised cost. For example, a difference of 10 minutes may have a greater effect on station choice than (main) mode choice. In such a case, the more sensitive choice ('station') is carried out at a lower level, and 'nested' within the mode choice. The model for station choice would then be 'conditional' on the mode used, making use of the formulation in Eq 3-2. However, the choice between modes would make use of the composite cost Eq 3-4 calculated separately for each mode. In such models the scale factor (λ) is different at each level, and it is a requirement that at any given level it may not be greater than the value for the next level down. For more discussion in the context of transport models, see WebTAG Unit M2, §3.2.

3.3.7 When estimating logit models, it is conventional to use the method of Maximum Likelihood[16], and to report the value of the maximum log-likelihood, as well as the value when all parameters are set to zero (the 'null' value). In assessing two alternative models where one of them can be viewed as a restricted version of the other (for example, by constraining certain parameters to have the same value), twice the difference in log-likelihood is distributed as a chi-squared statistic with degrees of freedom equal to the number of restrictions. Another statistic commonly reported is the "rho squared" value [ρ^2], which is an approximate measure of goodness of fit with a range between 0 and 1, though it should not be considered analogous to the R^2 statistic used in regression analysis. For further discussion, see the cited reference by Train, chapter 3.

3.3.8 As noted, for nested (hierarchical) models, the scale factors λ must not decrease as the 'tree' is descended. Thus if station choice is conditional on access mode, the λ relating to station choice cannot be smaller than the λ relating to access mode, otherwise the model is mis-specified. The estimation conventions are to some extent software-dependent, but can all be put on a consistent basis.

---

[16] This is a statistical method of fitting the parameters of a model to data which attempts to maximise the probability that – assuming the model form is correct – the data would be that which is actually observed.

4 Rail (PLD) Assignment

4.1 Introduction

4.1.1 The function of the PLD Assignment is to take matrices of station-to-station demand and assign them to appropriate routes through the PLD Network. This then allows the components of GJTC (in-vehicle time, interchange and an allowance for frequency – see section 2.4, as well as crowding) plus other generalised cost elements to be 'skimmed' from the network and passed to the SCM. As discussed later, the station-to-station demand is provided by the SCM.

4.1.2 In addition, the impact of relevant demand from the Regional PLANETs on crowding in PLD is included, by means of 'pre-loads' (see 2.2.16). PLD loadings on specified services relevant to the Regional PLANETs are also provided as output. The interfaces between PLD and the Regional PLANETS are described in Chapter 7.

4.2 Discussion of assignment

4.2.1 Since public transport assignment is a relatively specialised topic, it is worth providing some description. The problem is that between any two stations there are a number of different routes or services that passengers can choose to travel on, each of which will have different, stopping patterns, journey times, frequencies and crowding levels. Groups of services that have similar characteristics are for the purpose of this discussion referred to as 'transit lines'[17]. Each 'transit line that is coded will have an associated in-vehicle time [IVT] for all pairs of stations 'served', together with a specified frequency or headway. In practice, there may be some judgment in classifying the services actually available according to these criteria (eg there may be marginal differences during the course of the day in IVT, or in the stations actually served, which are considered insufficient to justify defining a separate 'transit line').

4.2.2 Based on appropriate assumptions to convert between frequency and (expected) waiting time (see 2.6.14), it is possible to calculate, for each station-to-station pair, the transit line with lowest GJT[18] (including an allowance for interchanges). But it would be inappropriate to allocate all the demand to this transit line. In particular, passengers arriving at random may find that another transit line will get them to their destination station faster, so that in practice they would not choose the minimum GJT service. For these and other reasons, public transport assignment methods use the notion of an 'attractive set' of transit lines, with the aim of defining those transit lines which in practice are likely to receive some of the demand.

---

[17] This is the terminology used within the EMME software.
[18] As we shall see in section 4.3, in practice the assignment also takes account of crowding, so GJT should be written GJTC. For the moment we ignore this complication.

4.2.3   The attractive set is normally defined by first identifying the transit line with the fastest GJT excluding the allowance for the service interval (i.e. 'waiting time'), and then considering what other transit lines offer a connection which under some circumstances could be worth taking. The rules in respect of how the 'attractive set' is defined may vary, and in particular according to software implementations. Having decided the attractive set, the standard ('frequency-based') approach is to allocate the demand across all attractive transit lines in proportion to their frequency.

4.2.4   The EMME software offers two options for undertaking assignment known as conventional and alternative methodologies, both of which are used in the PFM framework and discussed below.

4.2.5   In conventional EMME methodology, as used by the regional PLANET model, this standard approach is referred to as 'Optimum Strategy Algorithm' and may be described[19] along the following lines:

A: define 'attractive set' of transit lines:

- order transit lines in terms of increasing (generalised) journey time (excluding waiting); and

- proceed through transit lines adding to attractive set as long as the (generalised) journey time excluding waiting for the next transit line is less than or equal to the expected total time (including waiting) from the current attractive set.

B: allocate demand to transit lines:

- at any stage in the definition of the attractive set, the allocation of demand to a transit line is directly proportional to the frequency of that transit line; and

- thus the calculation of expected total time is based on this allocation

4.2.6   Waiting time is calculated as a user-supplied proportion of headway (the 'wait time factor'). In PLD, the wait time factor is 0.4, so that waiting time is calculated as 0.4 * headway, and further weighted by a factor of 2 (Table.2)[20].

4.2.7   Figure 4-1 below gives an indication of how this process operates, for an illustrative example[21]. In this example, the best transit line takes 91 minutes and has a frequency of 12 trains over the 16-hour day, with an implied waiting time of 62.3 minutes: these values correspond to the case where the attractive set contains only the best transit line. It is now shown what happens as further transit lines are considered for addition to the attractive set.

---

[19] A more detailed description, based on EMME documentation, is provided in Annex A

[20] According to WebTAG Unit M32, paragraph 3.2.1" 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." See also the discussion in paragraph 2.7.13 of this report.

[21] For the sake of clarity, options requiring interchange have been excluded from this example. Note that the values of 'wait' in Figure 4.1 reflect the PLD assumptions noted above.

4.2.8   Concentrating firstly on the solid lines, the top (blue) line [av_t] shows how the average in-vehicle time (including possible allowance for interchange) increases (because 'later candidates' have higher IVT) as further transit lines are added to the attractive set. This is compensated for by an increase in frequency which reduces expected waiting time. The mauve line [f], to which the right hand axis applies, shows how the total frequency (in terms of services per 16 hour period) of attractive transit lines increases, from 12 (in the case where the attractive set contains only the transit line with lowest IVT) to 36 when the process ends (because no more transit lines are worth adding): the brown line [wait] shows the corresponding reductions in expected waiting time. In this example, 15 attractive transit lines have been found: thereafter the reduction in effective waiting time is outweighed by the increase in mean IVT.

Figure 4-1: Definition of attractive transit lines [the dashed lines refer to the FJTS
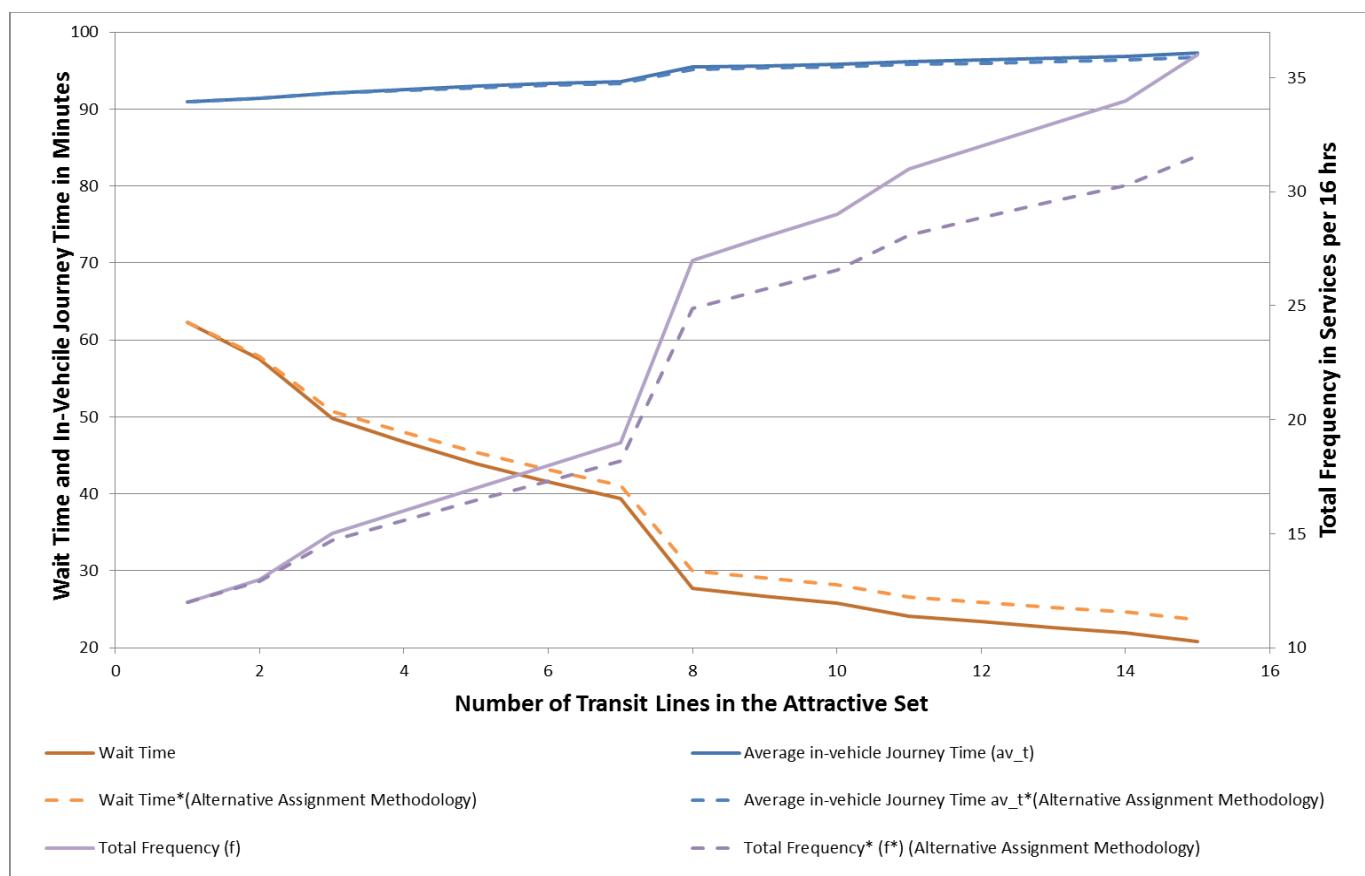
4.2.9   This standard version of public transport assignment, using the Optimum Strategy Algorithm, is widely used, and has been adopted for the Regional PLANETs described in chapter 7. However, an acknowledged concern with the frequency-based approach is that, once they are in the attractive set, two transit lines with the same frequency will get the same demand allocation, even though there may be a non-trivial distinction in their IVTs (in the example given, there was a difference of 14 minutes between the first and last transit lines in the attractive set). For this reason, it is of interest to consider how the frequency-based assumption might be relaxed, particularly in the context of long distance travel.

4.2.10   In the current version of the software (EMME 4.2.9) being used for PFM, an alternative algorithm is offered and this is used in the PLD model. In place of adding in the complete frequency for each new transit line accepted into the attractive set, the frequency is factored down by an amount proportional to the difference in IVT between the new transit line (t) and the mean IVT $\bar{t}$ for the current attractive set. The factor (p′) is given as:

Equation 4-1

$$p' = \left(1 - \frac{t - \bar{t}}{w}\right) < 1,$$

where w is the expected wait time for the current attractive set.

4.2.11   In general, the difference between p′ and 1 can be expected to be quite small, since longer distance services tend to have wider headways, so that the average waiting time will be relatively large (as the graph suggests). It can also be noted that because of its impact on the mean IVT, this approach can lead to a different (larger) attractive set. This revised 'strategy' is referred to as the Frequency and Journey Time Strategy [FJTS]: further details are provided in Annex A.

4.2.12   The dashed lines in figure 4-1 show the corresponding effects when this method is used. It can be seen that, in this example, there is a very slight reduction in average IVT [av_t*], and correspondingly less of a reduction in expected wait time [wait*]. Thus, as expected, the allocation of demand is more towards those transit lines in the attractive set with lower IVT. Overall, the changes are not very large. In this example, the attractive set is the same for both methods.

4.2.13   Nonetheless, the impact of the FJTS can be expected to be greater when high-speed services are introduced. It will be clear that the impacts of HS2 will be most keenly felt when there is a significant IVT improvement, and it is this which will test the assignment model to the greatest extent. Remaining with the previous example, but introduce an illustrative high speed service as an additional ' transit line' ('H'). In the example, the H service has 16 trains (over a 16 hour day) with IVT of 53.2 minutes: with the fastest service had 12 trains with IVT of 91.3 minutes. As H is the fastest transit line, it now becomes the first transit line in the attractive set, and corresponds to the case in Figure 4-2 where there is only one transit line in the attractive set.

4.2.14   With the traditional 'frequency' allocation, only two other transit lines enter the attractive set in this example[22]. For the faster of these, there is still a 38 minute difference in IVT with respect to H, and for the next fastest, the difference is 43 minutes. Thereafter it turns out that the improvement in waiting time from including further services in the attractive set is outweighed by the increase in mean IVT. The results are shown in the solid lines in figure 4-1. According to the frequency-based allocation, the new service H gets only 55% of the overall demand.

---

[22] These are of course the first two transit lines In Figure 4-1, corresponding to the Attractive set of 2 transit lines.

4.2.15   It is worth studying the inclusion of the second transit line in more detail. As noted, the H service in the example has 16 trains with IVT of 53.2 minutes and the next fastest service has 12 trains with IVT of 91.3 minutes. If for convenience it is assumed there are in fact 16 of the next fastest service and that it operates exactly between the H schedule, in other words 30 minutes later, then it will be clear that those catching the next fastest service will actually arrive later than they would if they waited another 30 minutes and took service H. On this reasoning, it seems unlikely that the next fastest service should take 41% of the demand as the frequency-based algorithm predicts (if there were indeed 16 rather than 12 trains it would take 50%).

4.2.16   Now considering the alternative FJTS assignment method, shown by the dashed lines in figure 4-2, it can be seen that the attractive set has increased relative to the frequency-only allocation: however, the additional 5 transit lines are all low frequency transit lines, with only 1 or 2 services per day, so that their impact is small. The essential difference between the two assignment methods occurs because of the 'down-grading' of the next fastest service, owing to its 38 minutes of additional IVT. As a result its share of total demand falls from 41% to 13%, while that of H rises from 55% to 85%. At least with this example, this appears a more reasonable outcome.

Figure 4-2: Definition of attractive transit lines with HS2 included [the dashed lines refer to the FJTS]

41
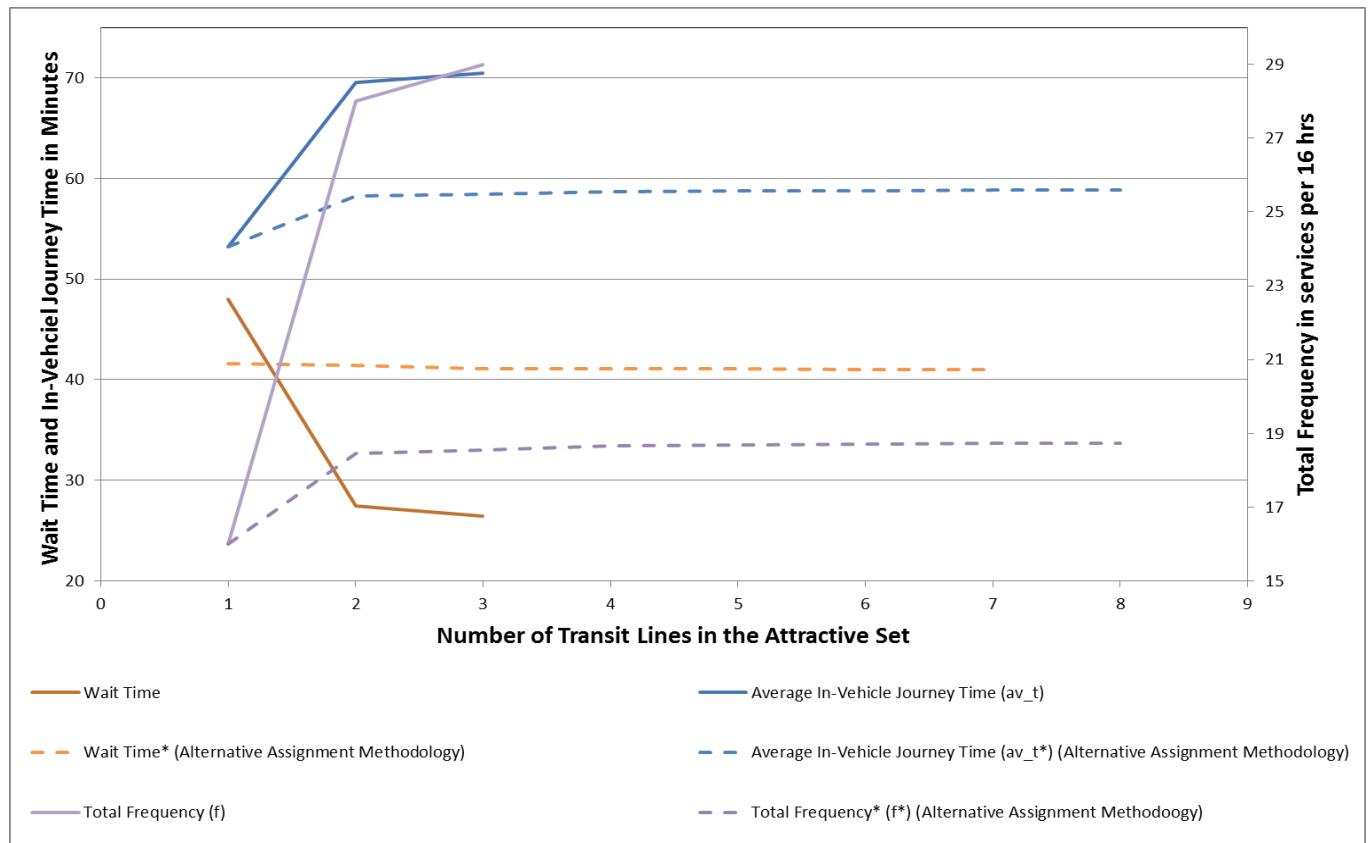
4.2.17   Therefore, while the effect of the alternative assignment method is slight in the absence of the high speed service H, it plays a much more important (and more in line with intuition) role when H is introduced. On this basis, the FJTS assignment approach has been used for PLD. This approach is also more consistent with other models in the rail industry: the ORCATS model for allocating Rail revenues between TOCs is based on a combination of frequency and journey time, and MOIRA has used a frequency and journey time assignment for many years. However, as already noted, the standard (frequency-based) algorithm has been retained for the Regional PLANETs.

4.3   The treatment of crowding

4.3.1   The rail assignment is further complicated by the need to take account of crowding. The initial allocations to services (' transit lines') produced by the algorithms just discussed assume IVT in line with scheduled times. However, the result of this allocation may be that some services receive a loading which, relative to their capacity, implies a level of crowding between one station and the next[23]. The way this is dealt with is, in essence, to weight the IVT for that section by a factor greater than 1: this may be viewed as an additional amount of 'disutility' measured in units of IVT to account for the discomfort associated with crowding[24]. In practice, the algorithm maintains IVT at its scheduled value, but introduces a further element to GJT referred to as 'crowded time'.

4.3.2   PDFH5[25] specifies (Table B6.2) 'Recommended Value of Time Multipliers' under conditions of crowding. The multipliers depend on the level of demand relative to capacity, whether the passenger is seated or standing, and the broad spatial categories of NSE[26], Regional and Intercity services: they relate to the loading on a specific train. The multipliers are close to 1 as long as the loading (demand per train) is at or below the train's seating capacity. Thereafter they rise quite rapidly, in line with passengers per square metre of standing room. This latter measure allows the procedure to take account of different rolling stock configurations in which seating capacity is traded against standing space.

4.3.3   For each rolling stock type, the area of standing room is calculated by dividing the total standing capacity by 2.5 (see Section 9.5).

---

[23] Note that passengers are assumed to experience a level of discomfort from crowding **before** all the seats are taken up and they are required to stand (see figure 4-3)

[24] An alternative interpretation, in line with PDFH, is that the "value of time" (more strictly, the value of travel time savings) is increased under conditions of crowding.

[25] The latest version of PDFH is now v5.1 (issued July 2013); however, crowding factors have not changed from v5.

[26] Network South East

4.3.4 Figure 4-3 illustrates how the multiplier for the per train crowding penalty varies with demand, for the intercity case. Up to seating capacity, the demand is measured as 'load factor' [LF] calculated as the ratio of passengers to seats, and it is assumed that there is no standing. Thereafter, demand is measured as (standing) passengers per m² of standing room. The blue line indicates how the seated penalty rises and the pink line shows the standing penalty, which only begins at LF = 1. The red line indicates the proportion of passengers who are seated[27], which is 100% up to LF = 1, thereafter falling: this is used to estimate the average penalty, shown in yellow. It can be seen that this rises more or less linearly from the point at which LF = 1 and standing pass/m² = 0.

Figure4.3: Illustrative crowding penalties (Intercity Example)
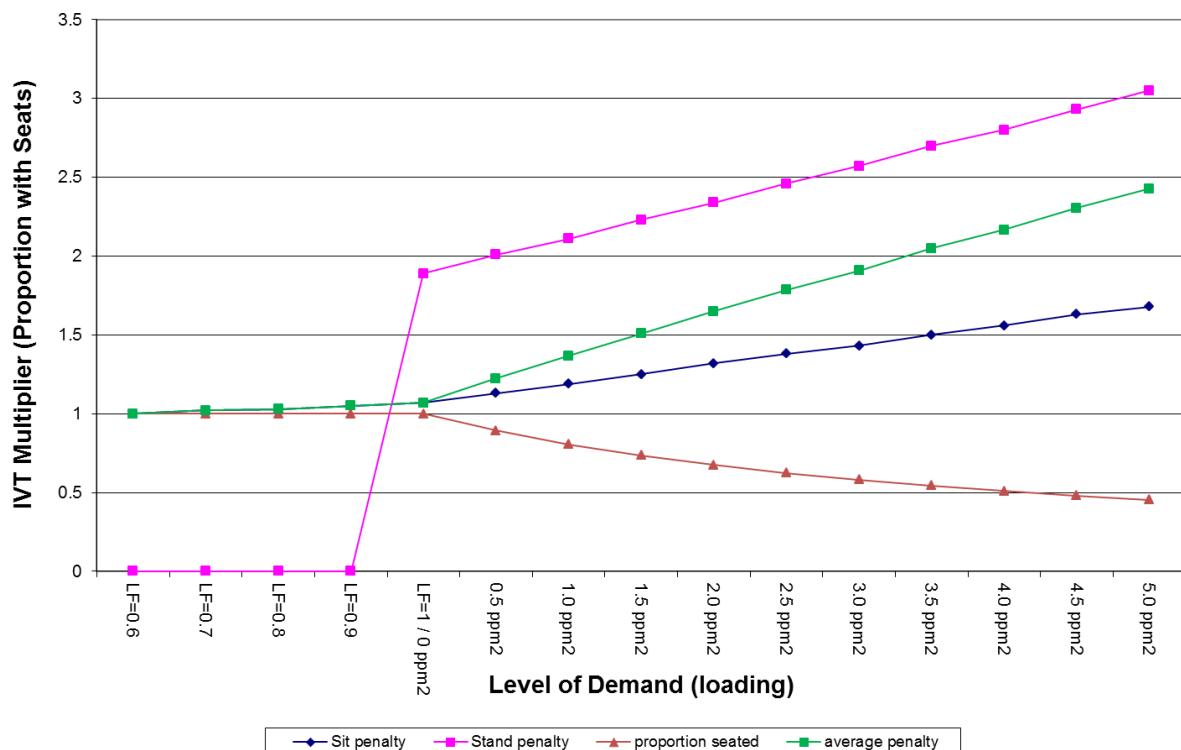
4.3.5 These crowding penalties are for a particular train with a given loading on a particular section of the line. The assignment model produces the loading for an average train on each section, and it can be expected that loadings vary over the course of the period modelled (for PLD this is the 16 hour day). Since the overall pattern is not linear over the whole range of loading, it is necessary to take account of these variations. For this purpose, the average crowding penalty function, as calculated for each rolling stock type, is approximated by a piece-wise linear formulation, as a function of the load factor [LF], as follows (where α is the value of the load factor at which crowding penalties commence, A and C are intercepts, and B and D are slope parameters for the fitted lines):

---

[27] This calculation requires an explicit assumption about the intercity configuration (in terms of standing room and seats): for this purpose the "IEP 8 car" stock has been assumed. Thus, while the standing and seating penalty curves apply to all intercity configurations, the average penalty curve is stock-dependent.

Equation 4-2

> If LF $\leq \alpha$, then Crowding Penalty CP = 1
> If α < LF $\leq$ 1, then CP = A + B.LF
> If 1 < LF, then CP = C + D.LF

4.3.6   In other words, up to a certain load (where demand = α times the seating capacity), there is no effect due to crowding: thereafter there is a modest, assumed linear, effect until seating capacity is reached [LF = 1], and then there is a different relationship (again assumed linear) beyond seating capacity.

4.3.7   The key parameter D, which gives the slope of the penalty function above the point where all seats are occupied, is obtained by determining the value of LF at 'Total Capacity' (corresponding to standing density of 2.5 passengers per m²), and using the PDFH5 recommendations to calculate the crowding penalty both at that level and at the point where all seats are occupied (LF = 1).

4.3.8   It is then further assumed that, for a predicted average loading factor μ on a specific section of a line, the variation in LF over the course of the day can be approximated by a normal distribution with mean μ and an appropriately chosen standard deviation σ. On this basis, the average crowding factor $\overline{CP}$ can be calculated by the following formula, where X (the variable of integration) represents the loading factor LF:

Equation 4-3: Average crowding factor

$$\overline{CP}_r(\mu) = \frac{1}{\sqrt{2\pi\sigma_r{}^2}} \int_{-\infty}^{\alpha_r} e^{\frac{-(X-\mu)^2}{2\sigma_r{}^2}} dX + \frac{1}{\sqrt{2\pi\sigma_r{}^2}} \int_{\alpha_r}^{1} (A_r + B_r X).e^{\frac{-(X-\mu)^2}{2\sigma_r{}^2}} dX + \frac{1}{\sqrt{2\pi\sigma_r{}^2}} \int_{1}^{\infty} (C_r + D_r X).e^{\frac{-(X-\mu)^2}{2\sigma_r{}^2}} dX$$

The formula is implemented by means of an EMME macro.

4.3.9   In this equation, r represents a particular 'TOC' group, and is thus associated with the rolling stock. Both the crowding function[28] parameters and the σ parameter representing the variation in loading over the course of the day are specific to 'r'. The values of the coefficient of variation [CV, = σ/μ] for each TOC group have been obtained by analysis of guard counts on arrival and departure from key stations.

4.3.10   This average crowding factor is then applied as a multiplier to the IVT for the section of line to which it relates and the difference between the result and (unweighted) IVT is defined as 'crowding time'.

---

[28] In previous versions of PDFH, the crowding function parameters also varied with purpose: however, since PDFH5 this is no longer the case.

4.4 The implications of crowding for the assignment

4.4.1 Starting with an assumption of no crowding, the EMME assignment procedure constructs the attractive set for each station to station pair, and allocates the demand for that station to station pair among the transit lines in the attractive set. Doing this for every station to station pair provides the total loading on every section of every transit line, and hence allows the average load factor μ per section to be calculated. Since the CV for the transit line is known, this allows the value of σ to be inferred. These values (μ, σ) are then used, via equation (4-3) above, to obtain the average crowding penalty for each section.

4.4.2 Because of the crowding penalty, the values of GJT(C) for affected routes will change, and this will have impact on the construction of the attractive sets and hence the allocation of demand. This therefore requires an iterative process to achieve convergence.

4.4.3 In order to control the iterations, a damping method known as the Method of Successive Averages [MSA] is used. Suppose V^n is the section loading in iteration n, and W^n is the implied revised loading taking account of the changed crowding levels. If just W^n is used as the next estimate of V for iteration n+1, the procedure may 'oscillate' without reaching a stable solution. Using MSA, before calculating the average loading factors μ, the loadings are averaged according to the following formula:

Equation 4-4:

$$V^{n+1} = \omega W^n + (1 - \omega).V^n$$

where ω takes the value 1/n.

4.4.4 The number of iterations is set at 10 and the following assignment statistics are output to monitor convergence:

- the overall network wide passenger kilometres is calculated for each iteration, and the percentage and absolute changes between iterations are compared.

- the segment with the minimum and maximum flow difference between each iteration is identified. (A segment is a specific transit line on a specific link.)

- The total network GJTC is calculated for each iteration (on a matrix basis) and the percentage and absolute changes between iterations are compared.

4.5 Dealing with User Classes

4.5.1 As noted in Section 2.3, PFM recognises different categories of demand. Insofar as these need to be recognised within the assignment, this is done by means of the concept of 'user classes'. The definitions used will be discussed in chapter 5.

4.5.2   Section 2.2 noted the four key procedures within an assignment model. Two of these – the choice of path, and the supply response – are potentially different by user class. Separate demand matrices are needed for each user class, and are treated independently in the assignment, though the contribution to section loading (and hence the level of crowding) is calculated by adding across all user classes.

4.5.3   The choice of path may differ by user class because the generalised cost specification is different. However, in the current version of PLD assignment, the generalised cost weights (see Table 2-2) do not vary by demand segment (except for the value of time, but this is not used in PLD assignment). In addition, while in previous versions of PDFH the supply response (which here relates to crowding) was different by purpose, the PDFHv5 based crowding response no longer varies by purpose. Hence there is in fact no requirement for different user classes: the demand matrices should be aggregated prior to the assignment.

4.5.4   In practice, for historic reasons, this is not currently done: nine user classes are used, corresponding to the different segments in the SCM, as described in chapter 5. Although computationally wasteful, this will have no impact on the outcome.

4.5.5   Hence, for the sake of completeness and to allow easier updating, the overall summary in the next section makes allowance for the possibility of user classes [denoted by the symbol 'u'].

4.6   Overview of PLD assignment

4.6.1   As will have been seen, the PLD assignment is highly complex. Because of the impact of crowding on the choice of transit lines, it involves an iterative structure with MSA damping. The crowding function itself needs to take account of the different rolling stock used on the various transit lines, and to convert between an average loading to take account of variation over the course of a 16-hour day. Finally, the route choice is achieved by means of an "attractive set" of transit lines definition: for this purpose, a recently developed algorithm has been used referred to as the "Frequency and Journey Time Strategy" [FJTS]. This is considered more appropriate than the standard frequency-based approach to deal with the special circumstances of HS2.

4.6.2   At the end of the assignment, the algorithm produces the number of passengers (by user class) on each section of each transit line, and also, by the process known as "skimming", the average GJT(C) values (also by user class) for each station-to-station movement, to be used in the Station Choice Model.

4.6.3   In order to bring the various aspects together, the box below provides a general account of the procedures carried out: while it is not intended to be a literal account of the EMME algorithm, it aims to convey the key sequences of calculations.

4.6.4   The following notation is used (in line with Table 3-1):

- R, S are stations, u is a user class, T^u_{RS} is the demand matrix
- L refers to a ' transit line' or combination of transit lines, and 'a' is a 'link' between

two adjacent stations on a transit line

- IVT_a is the scheduled time for link a, and F_L is the service interval for transit line L

- $T_{uRS,L}$ is the time taken to travel between R and S on transit line L for user class u, including allowance for interchange and crowding

- CP^u_a is the crowding penalty applied to in-vehicle time on link a for user class u

- V_a is the passenger loading on link a, and W_a is an intermediate estimate (for MSA)

4.6.5   The procedures are described using a pseudo-code which makes the looping structure clear. It also indicates the sections where the various procedures have been described.

Outline of PLD rail assignment algorithm

Read demand matrices T, network data, line seating capacities Cap_L
Set $CP^u_a = 1 \; \forall \; a, u$
Set $V^{iter}_a = 0 \; \forall \; a$
FOR iter = 1 to MaxIter DO
       WHILE (NotConverged) DO
            Set $\omega = 1/iter$
            Set $W^{iter}_a = 0 \; \forall \; a$
            {Assignment}
            FOR EACH u DO
                FOR EACH RS pair DO
                    {Determine Attractive Set A^u_{RS} using CP^u_a and $t^u_{RS,L}$ }   **§4.2**
                    {Allocate T^u_{RS} to each relevant link a}
                    {Update W^{iter}_a}
                END RS pair DO
            END u DO
            {Volume Averaging – MSA}   **§4.4**
            FOR EACH a DO
                $V^{iter+1}_a = \lambda \cdot W^{iter}_a + (1-\lambda) \cdot V^{iter}_a$
            END a DO
            {Crowding}   **§4.3**
            FOR EACH a DO {a $\in$ L}
                Calculate Load Factor $\mu_a = V^{iter+1}_a / Cap_L$
                Calculate $\sigma_a = CV_L \cdot \mu_a$
                FOR EACH u DO
                    Calculate CP^u_a
                END u DO
             END a DO
         {Test convergence}
         END WHILE
END iter DO
{Skim costs}   **4.6.2**
FOR EACH u DO
       FOR EACH RS pair DO
            {Calculate GJT component costs $C^{u,s}_{RS}$ using Attractive Set A^u_{RS} and CP^u_a }
       END RS pair DO
END u DO
Output costs $C^{u,s}_{RS}$ and loadings V^{iter+1}_a

5 Station choice model

5.1 Introduction

5.1.1 As has been seen, the PLD assignment requires a matrix of station-to-station demands T_{uRS}. However, the Demand Model (see chapter 6) works in terms of zone-to-zone movements.

5.1.2 The function of the SCM is to act as a 'bridge' between these two modules, by taking account of the access and egress possibilities between zones and stations. The SCM models passengers' combined choice of access mode and station. More precisely, it represents passengers' choice of access mode, first station and last station. Access mode is either by car or public transport[29].

5.1.3 Thus it takes a matrix of rail demand (from the Demand model described in chapter 6) and, taking account of the different times between different stations (GJTC, obtained from the assignment model described in chapter 4) as well as access and egress, allocates the demand to a pair of stations, which is then passed to the assignment model. The choice model is of the nested logit form, and has been specifically estimated for PFM.

5.1.4 The model is run separately for each purpose (Business, Other, Commuting). In addition, to allow for possibilities of car access, a distinction is made between car-available [CA] and non-car available [NCA] segments within each purpose. In what follows 'p' is used to denote purpose and 'c' to denote car availability.

5.1.5 As noted in Section 2.3, PFM works with matrices in OD, rather than P/A format. This causes some issues in relation to car availability, since the car would only be expected to be available at the home ('production') end of the trip. For this reason, additional segments CAF, CAT are defined for CA travellers, in line with the rail matrices described in Section 9.2. Direction (outward/return, or F/T)is denoted by 'd'. This distinction is not required for NCA travellers, since they have no choice of access mode.

5.1.6 The result is that the SCM works with the following 9 segments[30] [based on arguments pcd]:

- business car available from origin (thus origin is the production end) [CAF];

- business car available to destination (thus destination is the production end) [CAT];

- business car not available [NCA][31];

- other car available from origin [CAF];

- other car available to destination [CAT];

- other car not available [NCA];

- commuting car available from origin [CAF];

---

[29] Those reporting use of taxi are treated as having access by car.
[30] The combination of p c and d is equivalent to the 'user classes' discussed in the previous chapter (section 4.5).
[31] It is currently assumed that all business rail travellers have car available, so this segment is in fact empty

- commuting car available to destination [CAT]; and

- commuting car not available [NCA].

5.1.7   There is potential confusion, when working with OD matrices, in the terms 'access' and 'egress'. These therefore are define on a P/A basis, so that 'access' always relates to the production end, where there is an issue of car availability, and 'egress' always relates to the attraction end, where it is assumed there is no mode choice to the final destination.

5.1.8   In section 5.2, it was discussed the more detailed level of zoning which is required for the SCM. This then leads on, in section 5.3, to a discussion of catchment areas, and the way that the access and egress to and from stations possibilities are specified (section 5.4). In the remainder of the chapter it is discussed both how the model has been estimated and how it is applied within PFM. Beginning with the model estimation details (section 5.5). Different models have been estimated for production zones in London and those outside London. As the non-London model is more complex, it is presented first. The models make extensive use of data from the National Rail Travel Survey[32] (NRTS).

5.1.9   Because of the more detailed zonal system, a procedure is required whereby the PLD zonal demand matrices can be disaggregated. This is described as a 'gravity model' and has been estimated as a separate model, discussed in section 5.6. Finally the model application is described (section 5.7) and an overall summary of the calculations is given.

5.2   Zonal detail

5.2.1   In the context of station choice, a particular problem is caused by the size of the PLD zones (section 2.3). For this reason, the SCM considers the same geographic area as PLD but adopts a finer zoning system ('mzones') for what is defined as the core area of interest for HS2, broadly covering North West England, Yorkshire, West Midlands, East Midlands and Greater London. Figure 5.1 below shows the SCM's mzone coverage, corresponding with the core area.

5.2.2   For the core area outside London, the SCM operates at the Middle layer Super Output Area (MSOA) level, whereas Transport for London's Railplan[33] zoning system is used for the Greater London area. For the rest of Britain, the PLD zoning system is retained.

5.2.3   The above zonal disaggregation results in 3,962 mzones, comprising 2,608 MSOA zones, 1,211 Railplan zones and 143 PLD zones. A link between the 3,962 mzones and the 235 PLD zones is maintained such that data at mzone level can be aggregated to provide outputs at PLD zone level.

5.2.4   In what follows, I and J are used to denote PLD zones, and i within I, j within J to denote the mzones.

---

[32] See paragraph 5.5.4 for further details of the survey.
[33] Railplan is Transport for London [TfL]'s standard public transport assignment model.

Figure5-1: SCM zone coverage
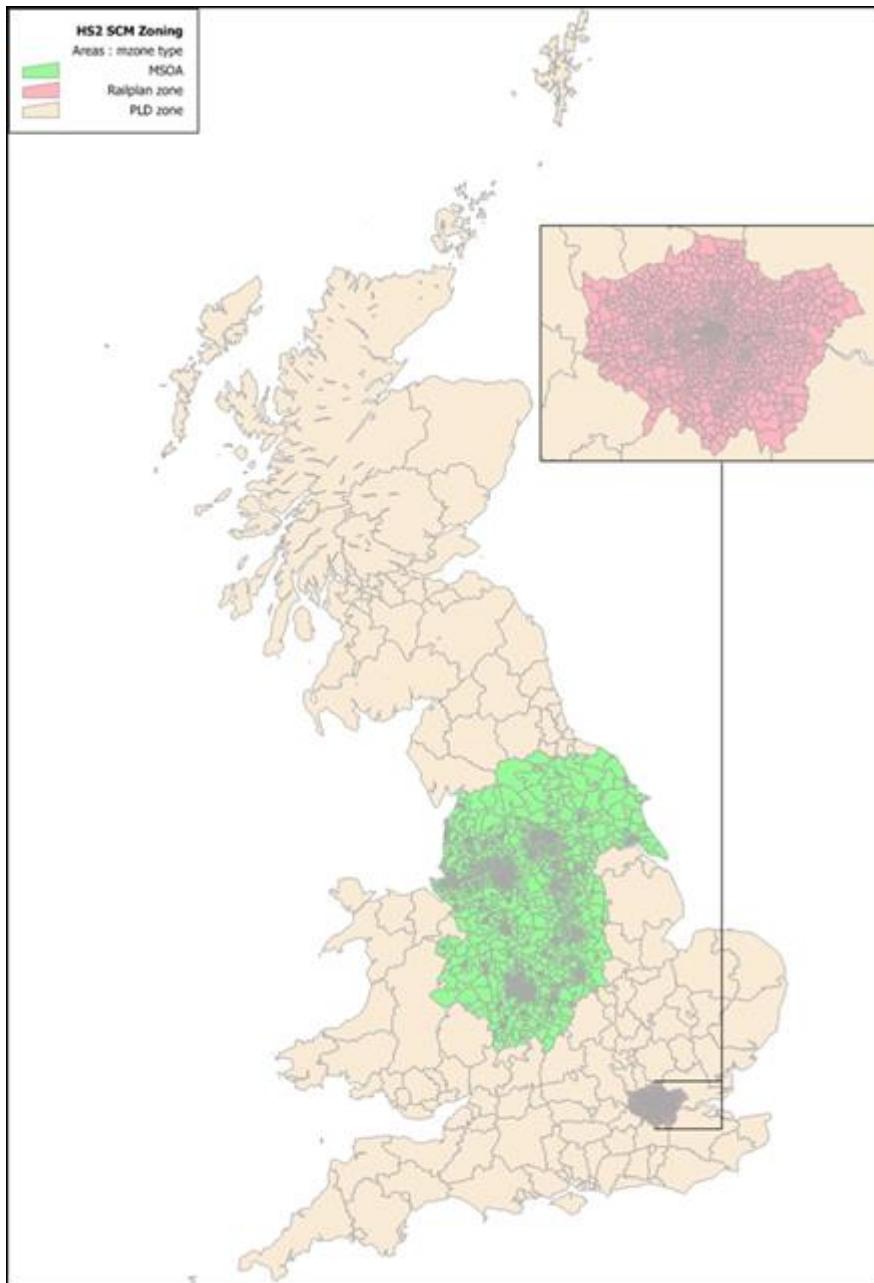
5.2.5   Although it is straightforward to aggregate the mzones up to PLD zones, an
         additional procedure is required whereby the PLD zonal IJ demand matrices can
         be disaggregated to the mzone ij level. The 'gravity model' developed for this
         purpose is discussed separately in section 5.6.

5.3   Definitions of 'catchment areas'

5.3.1   An important assumption for the model is to define the realistic choice set of
         station pairs [RS] for any i-j movement. Although it might be proposed that the
         SCM could model station use without any such constraint, this would have a
         serious impact on model run time.

5.3.2   In standard terminology, the catchment area of a station is the set of zones from which it derives potential demand. With PFM it is used in the inverse sense – for any mzone, it is the set of stations that might be accessed (on the production side), or which might serve as the alighting station for the final destination (on the attraction side).

5.3.3   In addition, the SCM does not aim to represent rail as an access mode, except where a local rail station is not represented in the PLD rail assignment[34] (described in the previous section). In other words, if a station [R or S] is included in the PLD rail assignment, it cannot be used in the SCM as an access point to another station. The only public transport [PT] access trips permitted in the SCM are:

- rail trips from a local station not in PLD rail assignment; and
- other non rail trips.

5.3.4   As a result of this restriction the 'catchment area' for PT in the SCM is generally smaller than that for highway. However, highway access is not allowed for London zones.

5.3.5   The following general principles were followed in determining catchment areas, though in key areas, particularly where HS2 are considering alternative station locations, catchment areas were examined individually to best match observed behaviour. The source of the access and egress cost data is discussed in section 5.4 below.

5.3.6   For highway access (NB for zones outside London only) the set of stations is the same for all mzones i within a given PLD zone I, and all PLD stations within 60 minutes of the zone centroid are potentially included. Not more than 20 stations are allowed in application, and in practice the number is usually much smaller. Where there is a potential need to restrict the number, this is done in favour of:

- nearer stations; and
- stations with a larger number of services.

5.3.7   For estimation purposes, the set is limited to those stations where there are observed NRTS flows between the station and the PLD zone.

5.3.8   Note that new HS2 stations are set up using these same principles. In this case consideration is given to the dominance of a station in connection with the other HS2 stations in the area.

5.3.9   Public transport 'catchment areas' are typically a subset of those for highway. In principle PT passengers are assumed to access the PLD rail network via their local station, and where the local station is within the PLD model this forms the focus for PT access trips, using bus or light rail, as well as walking.

---

[34] As described in chapter 9, the PLD rail network is intended to be a strategic network, and therefore does not include all possible stations.

5.3.10   PT access by classic rail is generally only used where the local station is not in PLD. Because of these considerations, it is possible for the actual choice set to vary across mzones within the PLD zone. However, every mzone will always have at least one PLD station to which PT access is permitted.

5.3.11   In the SCM model estimation, there are never more than 10 first/last station pairs available for a given PLD zone: hence, taking into account the two access modes, passengers have at most 20 alternatives to choose from. The situation is different in the application of the SCM, where more stations are taken into account and where the catchment areas can have up to 20 stations.

5.3.12   The catchment areas have been re-assessed in the light of the application of the SCM, to ensure that no unlikely choices are being made.

5.4   Defining access and egress

5.4.1   Within Greater London, TfL's Railplan model was used to provide generalised cost of PT access between Railplan model zones and stations. For this purpose, the Strategic Railplan 4.0 zoning system within the GLA area was used. Following model development, an update was made to Regional Railplan 6.2, which has around 4,000 zones within the GLA area. A conversion process has been developed to generate cost skims appropriate to the SCM, using AM peak values, averaged by direction.

5.4.2   Railplan adopts the following weighting factors for generalised cost elements of access times:

- walk time: 2.0

- wait time: 2.5

- in-vehicle time: 1.0

5.4.3   However, in the interests of maintaining general consistency of element weights throughout the PLD model, the weight for wait time was re-set to 2.0 to keep it the same as for the non-London model. The same approach was used for egress costs at the London end.

5.4.4   Outside London, public transport and highway access times and distances between mzones and stations are obtained from the National Accessibility Model[35] (NAM). NAM is the model used by the Department for Transport (DfT) to calculate travel times to essential services (the Core Accessibility Indicators). The model was adapted by Derek Halden Consultancy Ltd. to obtain access times between MSOA zones and stations for HS2 station choice modelling outside London (the routeing algorithm allows a maximum highway distance of 200 km and a maximum public transport access time of 120 minutes).

---

[35] See DfT report '2008 Core National Local Authority Accessibility Indicators', Final Report, November 2009 (prepared by Derek Halden Consultancy Ltd. et al.)

5.4.5   On the highway side, travel times are based on average journey times from
           Trafficmaster data[36] covering the period September 2008 to August 2009, for
           England only. Where there was no Trafficmaster link time data (a very small
           amount in England, and all links in Wales and Scotland) the following speeds were
           used:

- Motorways 100 km/hr

- A roads 70 km/hr

- B roads 60 km/hr

- Minor roads 50 km/hr

5.4.6   The public transport data used is the collective database assembled for the
           calculation of the 2009 core indicators (NPTDR - National Public Transport Data
           Repository archive - 31 October 2009).

5.4.7   The following modifications were made to the data as output from NAM:

- to avoid some very short highway times, 10 minutes was added to all highway
  times to allow for getting into the car, parking and getting to the station;

- it was assumed that for distances of up to 1km, people would walk instead of using
  public transport. The walk times came from NAM data and are based on a walking
  speed of 4.8km/h;

- for trips longer than 1km, a boarding penalty for public transport [PT] access was
  added to the public transport in-vehicle times. This varies according to distance,
  starting from five minutes and increasing linearly up to 30 minutes - 30 minutes is
  equivalent to the boarding penalty in the assignment part of the PLD model and is
  applied for trips over 30km ;

- 10 minutes have been added to PT wait time to allow for the initial wait time that is
  not included in the NAM dataset;

- after checking the access times from high NRTS demand zones to stations of
  interest against the national journey planning website 'Transport Direct', some
  further manual adjustments were made to PT journey times. Stations where
  adjustments were made to PT access times include Manchester Piccadilly,
  Manchester Airport, Warrington Bank Quay, Runcorn, Wigan North Western,
  Stockport, Macclesfield, Sheffield, Meadowhall, Nottingham, Derby, York and
  Leeds; and

- to maintain balance between station choice and mode choice, an adjustment was
  made to the car access costs in key areas of interest, such as Manchester,
  Liverpool, Sheffield and East Midlands.

5.4.8   NAM is run once to provide the required access times/distances inputs, as
           described below. The outputs from NAM are used as a set of static inputs to the
           SCM; the outputs are only changed in future years to reflect expected future local
           infrastructure schemes and to allow for the introduction of new stations. This is
           further described in chapter 11.

---

[36] http://assets.dft.gov.uk/statistics/series/the-standards-we-work-to/dft-statement-jrny-time-data.pdf

5.4.9   The outputs from NAM are:

- highway access time (in minutes);

- public transport access time - walk time (in minutes);

- public transport access time - wait time (in minutes);

- public transport access time - in-vehicle time (in minutes); and

- access distance (in metres).

5.5   SCM Model estimation

5.5.1   The SCM is a choice model based on a nested logit formulation. The estimation of a choice model requires data about the choices actually made, and explanatory variables (times, costs, etc) relating to each possible option. Both access/egress costs/times and times (GJT, including crowding) within the rail network are required. Separate model parameters are estimated for the three journey purposes: Commute, Business and Other. In addition, as is standard, Alternative Specific Constants (ASC) are used in some places to proxy for intrinsic preferences for some of the options.

5.5.2   Model estimation for the station choice at the London end was carried out separately (as it was expected that travel behaviour in London would be notably different from the rest of Britain). Beginning with the choice of station at the non-London end, and then consider the choice at the London end.

Non-London SCM

5.5.3   For station choice at the non-London end, the estimation is confined to trips that have London as one of the trip ends (i.e. attraction). For persons with no car available at the non-London end, access mode choice is not dealt with – for these records, only the station choice is relevant. For "car available" journeys where the production zone is not London, the estimation involves both station choice and access mode.

5.5.4   The choice data was taken from the NRTS[37] and represents 2004/05 demand. NRTS is an on-mode paper-based survey, designed to obtain the pattern of rail passenger travel. Raw survey data has been expanded to match station counts. A subset of the data relating to long-distance rail journeys (>50 miles) where there is a choice of station/route to use for a particular zone, travelling to/from London, was used.

5.5.5   Because the model relates to station and access mode choice outside London, egress at the London station is ignored in the generalised cost specification. As noted, access costs were obtained from the National Accessibility Model (NAM).

5.5.6   The following weights have been used:

- walk time [β_{wk}]: 2.0

---

[37] https://www.gov.uk/government/publications/national-rail-travel-survey-overview-report

- wait time [β_{wt}]: 2.0

- PT in-vehicle time [β_{PT}]: 1.0

- HW in-vehicle time [β_{HW}]: 2.0

5.5.7   All stations in the catchment area of a PLD Zone are considered available by car, whereas (as noted) some stations may not be considered accessible by public transport. Passengers that do not own a car can only choose between the stations that are accessible by PT.

5.5.8   Although long-distance rail fares are not included in the SCM, the monetary cost of access and egress is. For highway access, this is petrol cost and notional parking cost, while for PT access it is fixed and variable components of fare. These were 2005 values (consistent with the NRTS demand data) in 2010 prices (consistent with WebTAG and other components of PFM).

5.5.9   Petrol cost is calculated using WebTAG formulae from Unit 3.5.6 (October 2012 version)[38] based on fuel price, car efficiency, petrol/diesel car proportions and inflation rates (see Table 5-1). It was impossible to obtain sufficiently accurate data for parking charges in 2005 for each station, and a fixed parking cost was set to a notional amount of £13 (2009 prices), converted to 2010 prices. As a consequence, parking charge does not have any impact on people's choices, and merely affects the ASC for the highway access mode in the model formulation.

5.5.10   Public transport fares were calculated using a simple distance-related formula applied to an analysis of 2008 fares data, with a basic fare (fixfare) and an additional amount per kilometre:

Equation 5-1

$$-PT\_fare = pt\_fixfare + pt\_fare. \, Distance$$

5.5.11   These are 2005 values in 2010 prices using the DfT's local bus fares index[23] and the Retail Prices Index (RPI)[39]. The values are given in Table 5-1 below.

5.5.12   To convert the monetary costs of access/egress to time units, values of time (VoT) are required. These are consistent with the assumptions made in the PLD Demand model (Chapter 6): values were calculated at the mean distances for each purpose, taken from the NTS LD data. More information is given in Annex B.

5.5.13   The assumptions made for all monetary elements of access/egress are shown in table 5-1:

Table 5-1: Assumptions for monetary elements of generalised time

Parameters	2005 values/2010 prices
Highway cost parameters	
hw_petrol (pence/km)	6.15

---

hw_parking (pence)	1359.80
PT cost parameters	
pt_fixfare (pence)	94.67
pt_fare (pence/km)	11.77
Value of time (from PLD Demand Model)	
VoT Business (pence/min)	70.13
VoT Leisure (pence/min)	18.42
VoT Commute (pence/min)	25.14

5.5.14   Finally, for the rail GJTCs, skims were extracted from a base year (2010) run of the PLD assignment model. The skims relate to the following elements:

- in-vehicle time ['IVT'];
- additional GJT due to crowding ['Crowd', in IVT units];
- walk time ['Walk'];
- wait time ['Wait']; and
- boarding/interchange ['Board', average number]

Note that, in the context of PLD rail assignment, 'Walk' relates only to walking associated with interchange (which, in some cases, means moving between stations).

5.5.15   To ensure consistency with the GJTC weights used in the assignment, these are combined into a single quantity GJTC as follows:

Equation 5-2

$$GJTC = IVT + Crowd + 2.Walk + 2.Wait + 30.Board$$

It will be seen that the element weights are consistent with those in

5.5.16

Table2, as used in the PLD Assignment.

5.5.17   Prior to model estimation, directional differences in these costs are accounted for by averaging the costs for each direction.

5.5.18   For the three journey purposes, the utilities associated with the choice of a station pair RS for a journey[40] between mzones i and j are specified as follows (note that egress – at the London end – is omitted):

- Highway access

---

[40] As noted, I is treated as the production zone and j as the attraction, so that the "i–R"' leg is "access"' and the "S–j"' leg is "egress".

Equation 5-3

$$U^{\mathrm{P}}_{RS,Hwy|ij} = -\lambda^{Acc}_{\;p} \cdot \left(\mathrm{HWTime}_{\mathrm{iR}} \cdot \beta_{HW} + Petrol\,\frac{\mathrm{Dist}_{\mathrm{iR}}}{\mathrm{VoT}^{\mathrm{p}}} + \frac{Parking}{VoT^{\mathrm{p}}} + GJTC_{\mathrm{RS}} \right) + City_{\,p} \cdot \delta_{R=city} + Pway_{\,p} \cdot \delta_{R=Pway} + Other_{p} \cdot \delta$$

- PT access

Equation 5-4

$$U_{RS,PT|ij} = -\lambda^{Acc}_{\;p} \cdot \left(\mathrm{PTTime}_{\mathrm{iR}} \cdot \beta_{PT} + \mathrm{Walk}_{\mathrm{iR}} \cdot \beta_{wk} + \mathrm{Wait}_{\mathrm{iR}} \cdot \beta_{wt} + Fare\,\frac{\mathrm{Dist}_{\mathrm{iR}}}{\mathrm{VoT}^{\mathrm{p}}} + \frac{Fixed\ Fare}{VoT^{\,p}} + GJTC_{\mathrm{RS}} \right)$$

where:

- λ^{Acc} is the scaling parameter on generalised cost for access mode choice (utils/min);
- GJTC is the rail generalised journey time including crowding between the first and last stations (min);
- VoT is the trip-purpose-specific value of time (pence/min);
- Dist is the distance to the station (km);
- HWTime is the highway access time (min);
- Petrol is the price of petrol (pence/km);
- Parking is the parking charge at the station (pence);
- City, Pway and Other are the alternative specific constants (ASCs) for city, parkway and other stations (utils);
- δ_{City}, δ_{Pway} and δ_{Oth} are dummy variables (=1 for station types 'City', 'Parkway' and 'Other' station types respectively);
- PTTime is the public transport access time (min);
- Walk is the average walk time to the station (min);
- Wait is the average wait time en-route to/from the station (km);
- Fare is the average public transport fare per km (pence/km);
- FixedFare is the fixed component of the public transport fare (pence);
- β_{HW} is the highway access time coefficient;
- β_{PT} is the public transport access time coefficient;
- β_{wk} is the walk time coefficient; and
- β_{wt} is the public transport wait time coefficient;

5.5.19   ASCs have been used only for the car mode, and they are station-type-specific. Three station types were defined for the model calibration: 'City', 'Parkway' and 'Other', for city centre stations, parkway stations, and other stations, respectively. These ASCs and station definitions capture the components of the car access costs not included in travel time, notably parking cost and congestion in city centres. Where a station type was not clear, then the type was determined as that which gave the best fit to the observed choice of access mode.

5.5.20   For all purposes, the preferred nesting structure is for the choice of station pair [RS] to be nested below access mode choice [Highway vs PT], as shown in Figure 5-2.

Figure 5-2: Station choice structure
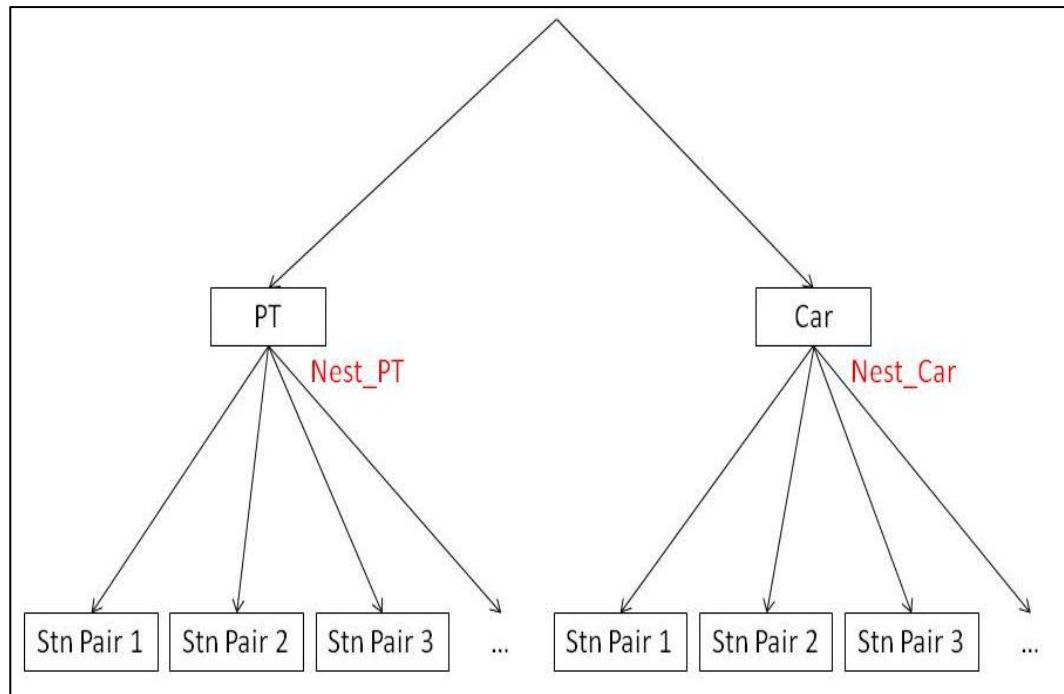
5.5.21   The parameters to be estimated are λ^{Acc}, City, Pway and Other, as well as the nesting parameter "Nest", which was constrained to be the same for both nests. Model estimations were carried out separately for the three purposes using the statistical estimation software Biogeme v1.8[41] .The results are set out in Table5-2:

Table5-2: Estimated coefficients for station choice model, by purpose

	Business	Other	Commute
Null log-likelihood	-10,873.1	-4,151.07	-5,131.42
Final log-likelihood	-5348.8	-2021.26	-2,297.86
Adj. rho square	0.508	0.512	0.551
Parameters (t-stat)			
City (utils)	0.772 (9.99)	0.527 (3.41)	1.74 (6.06)
Other (utils)	1.15 (24.9)	1.16 (11.76)	1.73 (21.32)
Pway (utils)	2.01 (13.85)	1.84 (7.73)	2.41 (13.83)
λ^{Acc} (utils/min)	0.0247 (14.7)	0.0156 (7.82)	0.0359 (15.3)
Nest = 1/θ [t-stat w.r.t 1]	2.6 (8.69)	3.86 (5.62)	1.61 (5.33)

---

[41] http://transport.epfl.ch/transport (Bierlaire, 2003). Accessed 1 October 2013.

5.5.22   Note that according to the conventions of Biogeme, the scaling parameter λ^{Acc} relates to the upper nest (i.e., the choice of access mode). The scaling parameter for the lower nest (choice of station pair RS) is obtained by multiplying λ^{Acc} by the "Nest" parameter[42]. With this convention, the Nest parameter must be greater than 1.0, and the t-statistics in the table measure its significance relative to 1.0 rather than zero, from which it is clear that the parameter in all cases is significantly greater than 1. This implies that the choice of station pair is considerably less random (lower error variance) than the choice of access mode.

London SCM

5.5.23   For station choice within London, estimation was carried out separately using observed demand data of movements to/from London where station choice exists at the London end. Unfortunately, there are no Intercity flows where there is a choice of station that is not also significantly affected by the train service or fare; it was therefore necessary to use South East flows from NRTS. The following three key movements were chosen:

- London to/from East Croydon/Brighton (choice of London Bridge, Victoria, etc);

- London to/from south-west London (choice of Waterloo or Vauxhall); and

- London to/from the South East (choice of Charing Cross, Cannon Street, etc).

These are all relatively short-distance journeys.

5.5.24   For the three key movements identified above, Railplan was used to provide generalised cost of access, with, as noted earlier, a modification to the weighting factor for wait time, while NRTS was used to provide the demand. For example, for a respondent reporting a journey from Waterloo to Wimbledon, the utilities associated with the two station alternatives (Waterloo or Vauxhall) for this respondent would be:

Equation 5.5

$$U = \alpha . \text{access time} + \beta . \text{IVT}$$

where α is the parameter associated with the access times, and β is the parameter associated with rail IVT. Note that for London station choice, there is assumed to be no choice of access mode, so nesting is not required.

5.5.25   Model estimations were carried out by journey purpose (Business, Leisure and Commuting) and also for all purposes combined. The estimations produced access time parameters α ranging between -0.13 and -0.17, depending on journey purpose. For all purposes combined, the parameter was estimated at -0.157. The estimated value by journey purpose, taking into account +/- two standard errors, did not appear to be significantly different from the overall value of -0.157. The resulting model, for all purposes combined, is shown in Table 5-3.

---

[42] An alternative convention, used in the Demand model estimation (Chapter 6), scales the parameters at the lowest level, and uses nesting or structural parameters θ (where $0 \leq \theta \leq 1$) for higher levels. The Biogeme estimates of the "Nest" parameters are equivalent to $1/\theta$ in this case.

Table 5-3: SCM logit model for the London end – All purposes

	All Purposes	All Purposes
Null Log-Likelihood	-174555	
Final Log-Likelihood	-105648	
Adjusted rho-square	0.395	
Parameter	Value	t-test
α (Access)	-0.157	-27.89
β (rail in-vehicle-time)	-0.0287	-3.26

5.5.26   Due to the characteristics of the short-distance rail journeys used in the calibration, it was felt that the estimated parameter β for the rail in-vehicle time component was not directly applicable for the long-distance rail journeys to be modelled in SCM. Hence it was decided that only the London parameter on access time should be used, and that in other respects the SCM should be the same as that estimated for non-London zones. In the case of London access times, therefore, a "London correction factor" is introduced to re-weight the PT access time so that the effective coefficient is equal to -0.157.

5.5.27   For the non-London model, the implied coefficient on PT access time (at the lower, station choice, level) is $-\lambda^{Acc}.Nest.\beta_{PT}$ (where $\beta_{PT} = 1$). Hence, for each purpose this coefficient needs to be further weighted in the London case so that it has the value of 0.157. Taking the example of the commute model (see Table5-2), the unadjusted coefficient has the value -0.0359 *1.61 = -.0578. In this case, a further weight of 2.72 is therefore applied to PT access time.

5.6   SCM Zonal disaggregation

5.6.1   It is necessary to disaggregate the PLD demand to the mzone level for the purpose of operating the SCM. The aim is, for each IJ movement, to produce fractions π_{ij}[IJ] which allocate the total zone to zone demand T^{pcd}_{IJ} to movements between the constituent mzones T^{pcd}_{ij}. The approach assumes that π_{ij}[IJ] is modelled as $f^{ipcd}.g_j^{pcd}$ (where f, g, are proportionate sub-zone allocations of the production and attraction zones, respectively). The calculation of f_i and g_j is done in various ways.

5.6.2   For zones outside the defined core area (see Figure 5.1), no disaggregation is carried out (PLD zones are used). For zones in London, Railplan demand matrices are used to apportion the total PLD zonal demand across all purposes (including Heathrow and Hillingdon) to the corresponding Railplan zones. For "key HS2 locations" (defined as PLD zones: Birmingham, Leicester, Derby, Nottingham, Sheffield, Leeds, Macclesfield, Stockport, Manchester and Warrington), the mzone distribution of productions and attractions is based on observed long-distance (over 50km) trips by purpose in NRTS.

5.6.3　For the remainder of the core area, the NRTS data is insufficient to allow observed factors to be used directly; hence "gravity" models are calibrated separately for the production [f_i^{pcd}] and the attraction [g_j^{pcd}] ends, and for each purpose, again using NRTS data for long-distance travel. The calibration dataset was restricted to those zones where the model is to be used. Calibration was undertaken using the 'R' software package[43].

5.6.4　In carrying out the disaggregation, there are two separate effects which need to be taken into account. The first, and most straightforward, is the 'size' effect due to the composition of the mzone relative to the PLD zone which contains it: this relates to the selection of appropriate zonal variables such as population, employment etc. However, because the SCM is only disaggregating rail demand, there will be a tendency for the demand to be concentrated – other things being equal – in zones with better access to the network. In order to take this into account, an intermediate model was estimated which also took account of station use by zone.

5.6.5　From NRTS, for all stations considered to be within the catchment area of a particular PLD zone, the proportion of demand at a given station associated with each mzone was used as the dependent variable. Thus, considering for example the production end, it should be first noted, for each PLD zone I, the ("catchment area") set of stations {R[I]}, and then for each station R, the NRTS weighted sample flow Q_{iR} originating from each mzone I is obtained. These flows were then taken as the dependent variable in a constrained regression on a set of zonal attributes for i, including the access time between i and station R, taken from NAM. For this purpose, the minimum of highway time and public transport weighted total time was used. The model was estimated separately for each purpose, with the total number of trips in each PLD zone constrained to match the observed. This requires a set of (PLD) zonal constants to be estimated.

5.6.6　For productions, the rail demand between a given mzone and station was found to be dependent on the population, income or number of high level managerial jobs in the mzone, and access time to the station. The following model was estimated:

Equation 5-6

$$\hat{Q}_{iR} = K_I \, P_i \, Y_i^{\beta_Y} \, W_i^{\beta_W} \, AT_{iR}^{\beta_{AT}} \quad \text{subject to} \quad \sum_{i \in I} \sum_{R \in \{R[I]\}} \hat{Q}_{iR} = Q_I$$

where, as usual,

- R relates to a station
- i relates to an mzone
- I relates to a 'PLD zone'

and

- P is the population at MSOA from ONS 2008 mid-year population estimates
- Y is the income at MSOA from ONS 2007/8 model-based income estimates
- AT is the access time from NAM
- W are the number of high-level managerial jobs from ONS 2001 Census data

---

[43] 'R' is a free software environment for statistical computing and graphics (www.r-project.org).

- K_s are constants specific to each "PLD zone"
- β_Y is the elasticity with respect to income
- β_{AT} is the elasticity with respect to access time
- β_W is the elasticity with respect to high-level managerial jobs

5.6.7   For business and commuting, it was found that income and access time gave the best fit in estimating the distribution of trips.

5.6.8   The Other purpose was split between home-based and non-home-based. For home-based Other, it was again found that income and access time gave the best fit; however, for non-home-based Other, the best model consisted of the number of high-level managerial jobs and access time, and the population variable was removed. The estimation results are provided in Table 5-4.

Table 5-4: Gravity model for Productions – t-statistics in brackets

	Business	Home-based Other	Non-home-based Other	Commuting
Elasticity to income β_Y	1.651 (34.9)	1.489 (28.2)		2.294 (49.2)
Elasticity to number of high-level managerial jobs β_W			0.9276 (103.5)	
Elasticity to access time β_{AT}	-1.831 (-60.0)	-1.578 (-51.5)	-0.7694 (-32.4)	-2.023 (-65.1)

5.6.9   It would be expected that a positive relationship exists to income and high-level managerial jobs as well as also expecting to have a negative relationship to access/egress time - people are more likely to travel long distances by train if they have a short access time to the station. The model estimates are therefore of the expected sign and magnitude. They are statistically significant at the 95% level (t-stats well above 2.0).

5.6.10   For the attraction end, the rail demand between a given mzone and station was found to be dependent on the number of high-level managerial jobs, and access time. Along similar lines, the following model was estimated:

Equation 5.7

$$\hat{Q}_{jS} = K_J W_j^{\beta_W} ET_{jS}^{\beta_{ET}} \quad \text{subject to}$$

$$\sum_{j \in J} \sum_{S \in \{S[J]\}} \hat{Q}_{jS} = Q_J$$

where, as usual,

- S relates to a station
- j relates to an mzone
- J relates to a "PLD zone"

and

- W are the high-level managerial jobs from 2001 Census
- ET is the egress time from NAM

- K_s are constants specific to each "PLD zone"
- β_w is the elasticity with respect to high-level managerial jobs
- β_{ET} is the elasticity with respect to egress time

5.6.11   Calibration was again carried out by journey purpose. Table 5-5 shows the model estimates. All estimates are of the expected sign and the ordering of the elasticities to managerial jobs being lowest for Other and highest for Business seems reasonable, as most business trips are made by those in high-level managerial jobs.

Table5.5: Gravity model for Attractions – t-statistics in brackets

	Business	Other	Commuting
Elasticity to number of high managerial jobs β_w	1.744 (70.7)	0.6799 (79.2)	0.9726 (99.9)
Elasticity to access time β_{ET}	-0.3389 (-8.9)	-1.197 (-38.2)	-2.167 (-54.0)

5.6.12   Ignoring the zonal constants (K in Eq 5-6 and 5-7) and the access time element AT, the model form can be viewed as giving an index of attractiveness for each mzone i, separately for each purpose p. The access time element then modifies this, separately for each station for PLD zone i. In order to produce an overall index H, the access time contributions were weighted by the contribution r of each station to the zone (also from NRTS). In other words, taking the production side as an example:

Equation 5-8

$$- \quad H_i^{prod,p} = P_i Y_i^{\beta_Y^p} W_i^{\beta_W^p} \sum_R r_{R|i} . AT_{iR}^{\beta_{AT}^p}$$

where the summation is over those stations R in the catchment area of PLD zone I, and $r_{R|i}$ is the proportion of demand from zone i which uses station R. Note that for the productions this quantity $r_{R|i}$ is taken as the same for all purposes, but for the attractions it is calculated separately for each purpose.

5.6.13   For the Other purpose, the separate "indices of attractiveness" for home and non-home were simply added, implying the formula:

Equation 5-9

$$H_i^{prod,p=oth} = \sum_R r_{R|i} . \left(P_i Y_i^{\beta_Y^{p=oth(home)}} AT_{iR}^{\beta_{AT}^{p=oth(home)}} + W_i^{\beta_W^{p=oth(nh)}} AT_{iR}^{\beta_{AT}^{p=oth(nh)}} \right)$$

5.6.14   The proportionate allocation f_i is then given as:

Equation 5-10

$$f_i = \frac{H_i^{prod}}{\sum_{i \in I} H_i^{prod}}$$

5.6.15   Corresponding calculations for H^{att} are carried out at the attraction end to obtain the proportionate allocation g_j.

5.6.16　　　These "gravity model" allocations f_i and g_j are applied once for each purpose (separately for production and attraction ends) for the [core area] non-key PLD zones, to obtain the required proportions $\pi_{ij}[IJ]$ of mzone level demand. These proportions are then retained as static inputs to the SCM: thus any changes resulting from population changes or new jobs over time are not taken into account in the disaggregation process.

5.7　　Model application

5.7.1　　　The operation of the model is described in figure 5-3. It should be read from bottom to top, going up the right-hand side and down the left, though in practice the interface with the other models means that there is an element of iteration. Figure 5-3 ignores demand segments, but in practice the model needs to be run separately for the 9 [pcd] segments described earlier.

5.7.2　　　Note that while the figure describes the essential logic of the model, the overall model algorithm incorporates an additional loop between the SCM and the rail assignment, as will be described in Chapters 10 and 11.

5.7.3　　　The model operates at the mzone level (ie for a movement between i and j). For any such movement, the set of available station pairs[44] {RS} for access mode x is defined as Z^x_{ij}. For the lower choice in the model, separately for each access mode, the total generalised cost, apart from the rail fare, for getting from i to j via R and S is needed.

---

[44] Note that although the set is notated as if it relates to the mzone movement i-j, the same set is actually used for all $i \in I$ and $j \in J$, for a given PLD zone to zone movement I-J.

Figure 5-3: The SCM within the Rail Demand and Supply system
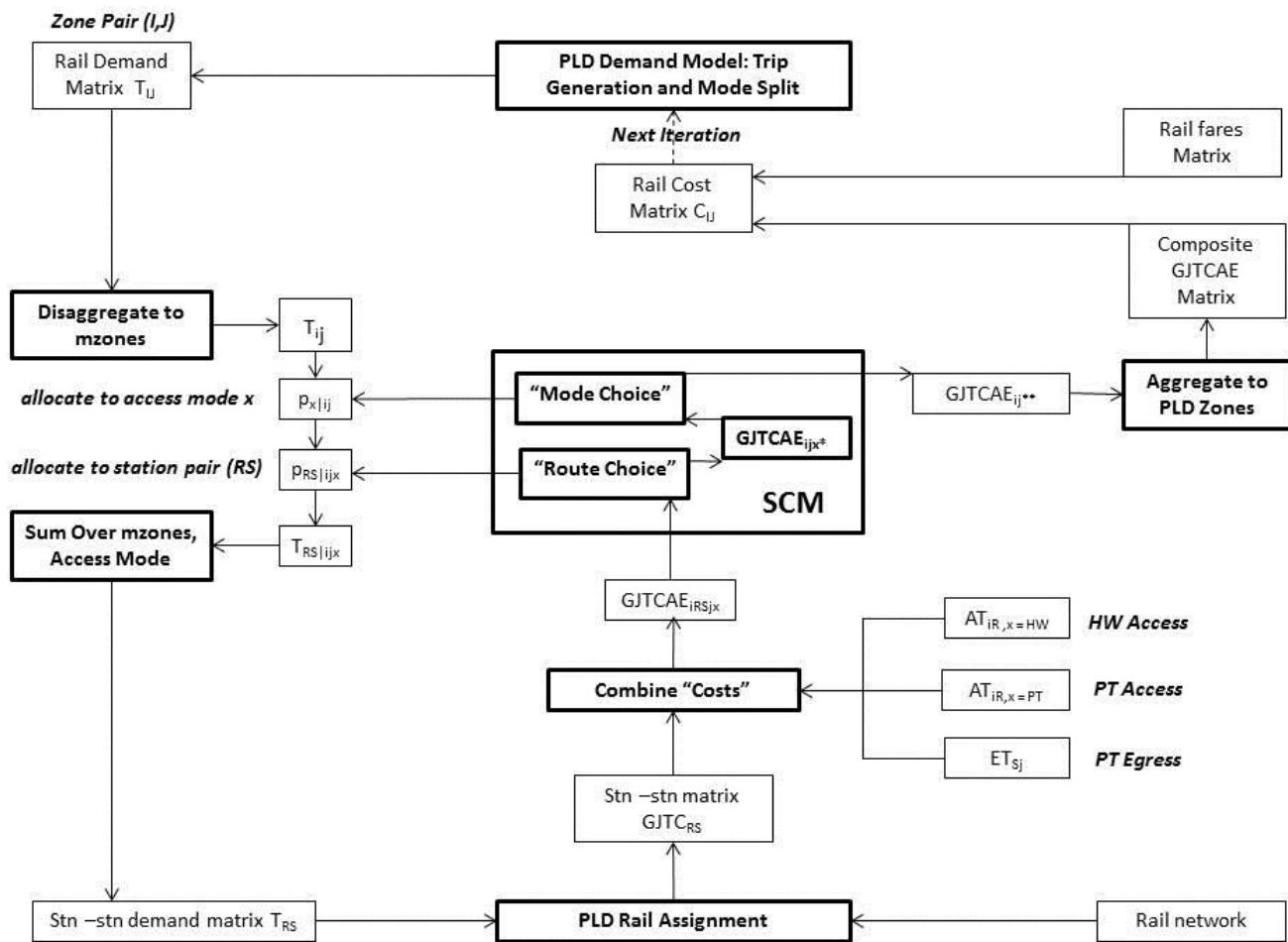
5.7.4   The diagram should be read from the bottom up, starting with the rail assignment. When the demand model is reached, the direction changes to come down the left-hand side.

5.7.5   Given the station-to-station GJTC, the first stage is to introduce the access and egress costs. For CAF travellers, access is defined as the movement from mzone i to station R, and egress as the movement from station S to mzone j. Conceived on a P/A basis, it is assumed that there is no mode choice for egress – public transport must be used. At the 'production' [home] end, access mode choice is available to those who have a car. For CAT travellers, access and egress are effectively reversed. Since NCA travellers have no mode choice, it is immaterial which end is which.

5.7.6   In point of fact, most long distance rail journeys will be round trips. The cost skims from the PLD assignment can vary significantly with direction, in part due to the definition of the 'attractive set', but also because for some movements the scheduled journey time can be significantly different by direction (up to ten minutes). For these reasons the SCM uses the average of the PLD rail GJTC in the two directions.

66

5.7.7   For reasons of computational efficiency, the calculations are carried out only when the demand matrix cell is non-zero.

5.7.8   Hence, using the same terminology as in the previous section, but now introducing egress costs explicitly, the weighted time GJTCAE (in units of IVT minutes) of a CAF movement i-R-S-j can be written as:

- for the highway nest

Equation 5-11

$$GJTCAE_{iRSj,Hwy}^{p[c=CA][d=F]} \big|_{} = \text{HWTime}_{iR} \cdot \beta_{HW} + Petrol \frac{\text{Dist}_{iR}}{\text{VoT}^p} + \frac{Parking}{VoT^p} + \tfrac{1}{2}\left(GJTC_{RS} + GJTC_{SR}\right)$$

$$+ \text{PTTime}_{Sj}.\beta_{PT} + \text{Walk}_{Sj}.\beta_{wk} + \text{Wait}_{Sj}.\beta_{wt} + Fare \frac{\text{dist}_{Sj}}{\text{VoT}^p} + \frac{FixedFare}{VoT^p}$$

$$+ \frac{1}{\lambda^{Acc}_{p}}\left(City_p \cdot \delta_{R=city} + Pway_p \cdot \delta_{R=Pway} + Other_p \cdot \delta_{R=Oth}\right)$$

- and for the PT nest

Equation 5-12

$$GJTCAE_{\big|iRSj,PT}^{p[c=CA][d=F]} = \text{PTTime}_{iR} \cdot \beta_{PT} + \text{Walk}_{iR}.\beta_{wk} + \text{Wait}_{iR}.\beta_{wt} + Fare \frac{\text{dist}_{iR}}{\text{VoT}^p} + \frac{FixedFare}{VoT^p} + \tfrac{1}{2}\left(GJTC_{RS} + G\right.$$

$$+ \text{PTTime}_{Sj}.\beta_{PT} + \text{Walk}_{Sj}.\beta_{wk} + \text{Wait}_{Sj}.\beta_{wt} + Fare \frac{\text{dist}_{Sj}}{\text{VoT}^p} + \frac{FixedFare}{VoT^p}$$

5.7.9   The choice of station pair RS (notated "Route Choice" in Figure 5-3) for access mode x (= Hwy or PT) is then given by the logit model form:

Equation 5-13

$$p_{RS|ij,x}^{pcd} = \frac{\exp([\lambda^{Acc}_{p}/\theta_p].GJTCAE_{iRSj,x}^{pcd})}{\displaystyle\sum_{[R'S']\in Z_{ij}^x}\exp([\lambda^{Acc}_{p}/\theta_p].GJTCAE_{iR'S'j,x}^{pcd})}$$

where $p_{RS|ij,x}$ is the proportion of those travelling from i to j and using access mode x who choose station pair [RS].

5.7.10   By the standard properties of the nested logit model, this yields a 'composite GJTCAE' ['logsum'][45] over all available [RS] pairs, separately for the two access modes (where available):

Equation 5-14

$$GJTCAE_{ijx*}^{pcd} = \frac{1}{[\lambda^{Acc}_{p}/\theta_p]}\ln\left(\sum_{[R'S']\in Z_{ij}^x}\exp([\lambda^{Acc}_{p}/\theta_p].GJTCAE_{iR'S'j,x}^{pcd})\right)$$

5.7.11   This composite GJTCAE is then used in the higher-level choice between highway and PT access[46] (notated "Mode Choice" in Figure 5-3), again using the logit form:

---

[45] see Section 3.3 for an explanation.
[46] This model is not needed for the NCA segments, since they are assumed "captive" to public transport access.

Equation 5-15

$$p_{x|ij}^{pcd} = \frac{\exp(\lambda^{Acc}{}_p . GJTCAE_{ijx*}^{pcd})}{\sum\limits_{x' \in \{Hwy, PT\}} \exp(\lambda^{Acc}{}_p . GJTCAE_{ijx'*}^{pcd})}$$

where p_x|ij is the proportion of those travelling from i to j who use access mode x.

5.7.12   Again, this yields a 'composite GJTCAE' ['logsum'], this time over the two access modes:

Equation 5-16

$$GJTCAE_{ij**}^{pcd} = \frac{1}{\lambda^{Acc}{}_p} \ln\left(\sum\limits_{x' \in \{Hwy, PT\}} \exp(\lambda^{Acc}{}_p . GJTCAE_{ijx'*}^{pcd}) \right)$$

5.7.13   This is therefore the composite cost (ignoring fares) of rail travel between mzones i and j, allowing for choice of access mode and station pair. Note that for NCA segments, GJTCAE_{ij**} = GJTCAE_{ij(PT)*}, as derived in Eq (5-14).

5.7.14   However, since the PLD Demand model (see Chapter 6) operates on PLD zones rather than mzones, we need to average over all i ∈ I, j ∈ J. This is done by weighting by the demand proportions π_{ij}[IJ] between mzones i and j and summing over all relevant mzones. Hence the composite GJTCAE matrix (GJTC plus access and egress costs) at the PLD zone level is obtained as follows:

Equation 5-17

$$GJTCAE_{IJ}^{pcd} = \sum\limits_{i \in I, j \in J} \pi_{ij}^{pd}[IJ] GJTCAE_{ij**}^{pcd}$$

5.7.15   Since the Demand model does not recognise the 'directional' segment (To/From), which exists only for the rail mode, the result is averaged over the two directions to give:

Equation 5-18

$$GJTCAE_{IJ}^{pc*} = \frac{1}{2}\left(GJTCAE_{IJ}^{pc[d=F]} + GJTCAE_{IJ}^{pc[d=T]} \right)$$

5.7.16   After adding in the fares matrix divided by VoT, this is passed as generalised cost to the PLD Demand model (Chapter 6).

5.7.17   The demand calculations now begin at the top of figure 5-3 , starting with an estimate T_{IJ,rail}^{pc} of total PLD zone to PLD zone rail movements (by segment pc) from the Demand model. For CA travellers, this needs to be split by direction, and this is done on the basis of the pivot[47] Matrices B, so that:

Equation 5-19

$$T_{IJ,rail}^{pcd} = \frac{B_{IJ,rail}^{pcd}}{\sum\limits_{d' \in [F,T]} B_{IJ,rail}^{pcd'}} T_{IJ,rail}^{pc}$$

---

[47] See Section 6.4 for further explanation.

5.7.18   The first step is to disaggregate this demand to the mzone level, giving the demand matrix[48] for the SCM as:

Equation 5-20

$$T_{ij}^{pcd} = \pi_{ij}^{pd}[IJ].T_{IJ}^{pcd}$$

5.7.19   Using the choice proportions previously calculated (Eqq 5-15 and 5-13), this demand is then allocated, first to access mode:

Equation 5-21

$$T_{ijx}^{pcd} = p_{x|ij}^{pcd}.T_{ij}^{pcd}$$

and then to station pairs:

Equation 5-22

$$T_{RS|ijx}^{pcd} = p_{RS|ij,x}^{pcd}.T_{ijx}^{pcd}$$

5.7.20   The required station-to-station matrix T_{RS} by user class [u = pcd] for the PLD rail assignment is obtained by summing over mzones and access modes:

Equation 5-23

$$T_{RS}^{pcd} = \sum_{I}\sum_{J}\sum_{i \in I}\sum_{j \in J}\sum_{x} T_{RS|ijx}^{pcd}$$

5.7.21   This completes the description of the SCM. As noted above, within the overall iterative system some of the steps may be omitted. This is further discussed in Chapters 10 and 11.

---

[48] henceforth in this section we drop the "rail" subscript since all quantities are now for the rail mode

6 PLD Demand Model (mode split and generation)

6.1 Introduction

6.1.1 The function of the PLD Demand model is to predict the demand for long distance [LD] travel as a result of changes in costs. It is constructed as a hierarchical (nested) logit model dealing with mode choice (between car, air and rail) and generated (new) travel. Note that, in this context, 'new' travel refers not just to wholly new trips but also those that might have been abstracted from other destinations (since destination choice is not represented in the model) and other modes that are not treated explicitly (such as coach).

6.1.2 The model has been specially constructed for PFM, and builds on the experience of estimating a similar (though more complex) demand model in the context of the DfT's Long Distance Model [LDM][49]. It is estimated using observed data on respondents' long-distance[50] travel choices from the 2002-2010 NTS surveys, using adjustments to take account of changes in travel costs over the 2002-2010 period.

6.1.3 An important aim of the model is to determine the relative sensitivity to generalised cost of mode choice and generation. In application, it 'pivots' off a base estimate of demand in response to changes in generalised cost delivered from the modal networks and, in the case of rail, from the SCM.

6.1.4 This chapter describes how the model was estimated and then how it is applied within PFM.

6.2 Model estimation

6.2.1 As with the SCM, the estimation of a choice model requires data about the choices actually made – in this case, mode choice and the frequency of LD trip making, and explanatory variables (times, costs, etc) relating to each possible option.

6.2.2 The LD data is provided by both the standard one-week diary survey used in the NTS to record information on trips of all lengths, and trips collected from the dedicated LD travel 'recall' survey[51]. Both the recall and diary data are used to estimate the models, but for the frequency model estimation, the recommended NTS weights are used to take account of the impact of recall error in total (LD) trip making in the recall survey.

---

[49] C Rohr; J.Fox; A. Daly; B. Patruni; S. Patil; F. Tsang (2010) "Modelling Long-Distance Travel in the UK", European Transport Conference, Glasgow.
[50] Trips with a one-way distance of at least 50 miles.
[51] For more information, see section 9.3 (highway matrices)

6.2.3   The NTS uses a 406 district-level zone system to code the home end, and a 146 Unitary Authority (UA) and County zone system to code the non-home end. The districts can be aggregated to the 235 PLD zones, but at the non-home end, a process was required to convert between PLD zones and the more aggregate 146 UA-County system in order to use the generalised cost data available within PLD. Three of the UA-County zones are external to the area modelled in PFM (Orkney, Shetland and Western Isles), and the correspondence for the remaining UA-County zones is as follows:

- 79 PLD zones have a 1:1 mapping between the PLD and UA-County zoning;

- 16 PLD zones have a one-to-many mapping to 30 different UA-County zones, with up to five UA-County zones mapping to a single PLD zone; and

- 140 PLD zones have a many-to-one mapping to 34 UA-County zones, with up to nine PLD zones mapping to a single UA-County zone.

The way these last two cases were resolved for the generalised cost data is described below.

6.2.4   In line with the approach taken for the LDM, the modelling unit for the PLD demand model estimation is the full LD home-based 'tour', defined as a series of linked long-distance journeys starting and finishing at the same home-location. Each individual trip has to be over 50 miles in length to be recorded in the NTS LD data. A primary destination is identified for each tour which defines the purpose of the most distant destination visited. Subsequently the units were converted to trips for implementation in PFM, as discussed below.

6.2.5   The detailed purpose codes recorded in the NTS data have been aggregated into the three model purposes as follows:

- Commute tours are tours made to work primary destinations (NTS purpose code 0);

- Business tours are tours made to primary destinations visited in the course of work (NTS purpose code 1); and

- Other travel tours are tours made to all other primary destinations including education (all other NTS purpose codes)

6.2.6   To determine the 'main' mode used for tour legs involving two or more LD trips, the following mode hierarchy has been applied across the modes used for each of the LD trips made during the tour leg:

- Air

- Rail

- Bus/coach

- Car

- Other

6.2.7   Thus, if car is used to access a public transport mode (for example, driving to an airport to catch a flight), the public transport mode is represented as the main mode.

6.2.8   Tours with main mode 'bus/coach' and 'other' are excluded. In addition, with a view to ensuring that the estimated model was tailored to the scope of PLD, a significant number of tours were dropped because they are made entirely within one of the PLANET regional models, and therefore not relevant to PLD. Furthermore, as PFM is a weekday-only model, weekend tours (where both the outward and return legs are made at the weekend) were excluded from the estimations. A significant fraction of 'other purpose' tours depart on a weekday and return over the weekend, or vice versa: these have been retained in the estimations with a weight of 0.5 applied to strip out the weekend demand. Note that because rail is the only mode available to NCA individuals (since, by assumption, neither air nor car are available), all NCA records are dropped from the mode choice estimations: records as which have no choice cannot contribute to the model estimation.

6.2.9   On this basis table 6-1 summarises the final sample sizes available from the NTS for model estimation.

Table 6-1: Mode choice model final sample sizes for estimation

Business	Other	Commute
2,581	4,626	12,589

6.2.10   Analysis demonstrated that rail mode shares were higher in the 2006-2010 data compared to the 2002-2005 data, particularly for commuting. For this reason, separate sets of mode constants were estimated for the 2002-2005 and 2006-2010 periods. In addition, since the available level of service (LOS) data reflects travel conditions in the base year at the time of model estimation (FY2010/11), while the choice data spans the 2002 to 2010 period, these separate sets of mode constants make some allowance for significant changes in LOS over this period; in particular enhancements to the rail network such as the West Coast Main Line upgrade.

6.2.11   Composite 'Logsum' rail generalised journey time (GJTCAE) measures for the base year were supplied by the SCM, thus ensuring consistency in the generalised cost element weightings, and separate fare matrices were made available. As was noted in paragraph 5.7.7, the SCM only calculates GJTCAE at the PLD zone level if there is non-zero demand in the base rail matrices B_{IJ}. This means that where no rail trips are observed in the Base Matrices, the rail mode has to be treated as unavailable in the model estimations.

6.2.12   For car, distance and time skims have been supplied from the PLD highway assignment model separately for business, other and commute purposes. No information on tolls, such as for travel on the M6 Toll or one of the Dartford crossings, is available from the skims.

6.2.13   For air, in-vehicle time, frequency, wait time, access and egress time and fare skims have been supplied from the PLD air assignment model separately for business and other purposes. Air is not modelled for commute travel and, as noted, all air trips are assumed to be in the 'car available' segment.

6.2.14   In most cases, the coding of the NTS destination permitted the direct use of the LOS data. For the 'one-to-many' cases, the LOS for the more aggregate PLD zone in which each UA-County zone lies was used for all the coded destinations. For the 'many-to-one' cases, LOS to the UA-County zone was calculated as a weighted average of the LOS to each PLD zone that lies within the UA-County zone, using as weights total employment for commute and business, and total population for the other travel purpose. Employment and population information for 2010 was extracted from TEMPRO version 6.2, and aggregated to the PLD zone level.

6.2.15   The main objective of the mode choice estimations is to estimate the sensitivities to generalised time at the lowest level in the mode choice structure – the λ values – and then to estimate structural parameters (θ) that define the relative sensitivity of higher-level choices. In addition to the generalised time parameters, the mode choice models incorporate mode-specific and other constants. Cost damping[52] was incorporated by making adjustments to the values of time (VoTs) that are used to convert costs into generalised time units. For all three purposes, cost damping improves the fit of the model to the data and increases the magnitude of the generalised time parameter, implying higher mode choice sensitivities.

6.2.16   The specification used in the mode choice models is as follows, where 'm' denotes mode:

Equation 6-1

$$C_{mIJ,y}^{pc[tour]} = GT_{mIJ}^{pc[tour]} + \mathbf{M}_{mIJ,y}^{p[tour]} / VoT_y^p (D_{IJ})$$

where:

- $C_{m,y}$ is the generalised cost for mode m in year y (minutes)
- GT_m is the overall journey time component of generalised cost for mode m, including both in-vehicle time and out-of-vehicle time components with appropriate weightings (minutes) – see below
- $M_{m,y}$ is the monetary cost of the mode (pence in 2010/11 prices) in year y
- $VoT_y[D]$ is the value of time (pence/minute in 2010/11 prices) in year y, for a trip of distance D [cost damping effect] – see below

Note that for estimation purposes, all these quantities are defined for the tour (round trip).

6.2.17   For rail, GT is given by the overall GJTCAE logsum (eg 5-18) provided by the station choice model (minutes). For car, GT is simply the in-vehicle time taken from the highway network. For air, GT is given by the following equation:

Equation 6-2

$$GT_{Air} = Air\_IVT + Air\_Wait + 2 * Air\_aceg$$
where:

---

[52] Cost Damping is a feature in some travel demand models by which the marginal disutility of (monetary) cost, and possibly of time, declines as journey lengths increase: see WebTAG Unit M2 §3.3.

- Air_IVT is the air in-vehicle time (minutes);
- Air_Wait is the air wait time (0.4 * headway) (minutes); and
- Air_aceg is the access/egress time (minutes).

6.2.18   The rail and air fares provided are defined in FY2010/11 values in 2002 prices[53]: note that for reasons of timing these are based on an earlier version of the rail fares than those described in section 9.6. After inflating to 2010/11 prices using the annual RPI CHAW index[54], these are adjusted in real terms to the year of the NTS observation, using data on average long-distance rail fares from the Office of Rail Regulation website[55], and from the CAA air passenger survey for air fares. This provides cost for year y.

6.2.19   For car, cost for year y was calculated using the formulae in the October 2012 release of WebTAG Unit 3.5.6[56], which calculate total car costs (both fuel and non-fuel) as a function of OD average speed in 2010 values and prices. The procedure that was followed is summarised as follows (detailed calculations are included in Annex C):

- 2010 fuel consumption values (litres/km) were calculated separately for petrol and diesel vehicles for each OD pair, as a function of the speed for the OD pair (implied from the distance and time skims);

- changes in fuel efficiency were calculated relative to the 2010 values;

- information on petrol and diesel prices (p/litre) was assembled for each year, together with information on the proportion of cars using petrol and diesel;

- this information was combined to calculate fuel cost in p/km separately for each OD pair and year;

- non-fuel costs were calculated in p/km, for business travel these also vary as a function of speed and therefore are calculated separately by OD pair;

- fuel costs and non-fuel costs were combined to give total car costs in p/km, which were multiplied by the highway distance skims to give total car costs; and

- these per-vehicle car costs are divided by an average party size to convert them into per-person costs.

6.2.20   Thus, adjustments have been made to calculated car costs for each year in the 2002-2009 period, though no account is taken of changes in speeds due to congestion over time.

6.2.21   The VoTs used to convert monetary costs into time units are also in 2010/11 prices, and adjusted in real terms to match the year of the NTS observation, as described below.

---

[53] Note that, in both cases, the PFM fare matrices are in pounds and so need to be multiplied by 100 to convert to pence.
[54] NB Again for reasons of timing, the RPI was used rather than the now recommended GDP deflator, following a change in ONS practice. It is not considered that this will have a significant impact.
[55] See http://dataportal.orr.gov.uk/displayreport/report/html/7cff3127-a5cc-4173-ac78-016db2339811. Accessed 1 October 2013.
[56] The latest values are now in Table A1.3.8 of the TAG data book: the model estimation has not been updated

6.2.22   For commute and other travel, WebTAG Unit M2, Appendix C3 provides a function that allows VoTs to be calculated as a function of distance and the household income of the traveller. Average incomes (in FY2010/11 prices) were calculated from the samples of individuals observed to make long-distance tours in the NTS data.

6.2.23   WebTAG does not provide any information about how business VoTs vary with distance. Therefore, additional analysis of the 2009 Stated Preference [SP] data collected during the development of the LDM was undertaken to develop a VoT relationship with distance, and a single in-vehicle time parameter was used across the three modes so that the implied VoTs do not vary with mode. For this purpose the distance elasticity of 0.36 from the WebTAG commute VoT relationship was used to impose a distance variation on the WebTAG all-modes VoT value. More information about the calculation of VoT is provided in Annex B.

6.2.24   Figure 6-1 illustrates the frequency and mode choice model structure, and highlights how the θ parameters[57] are used to define the sensitivities to generalised cost (the λ values) at each level in the choice structure. This structure applies to all three purposes, though for commuting the air mode is not available. The lowest level is the choice between the PT modes rail and air, and above that is the choice between car and PT. As it will be shown, the generation (frequency) model sits above the mode choice model.

6.2.25   Model estimations were carried out using the statistical estimation software ALOGIT[58] .

Figure 6-1: Frequency and mode choice estimation structure (for all purposes)
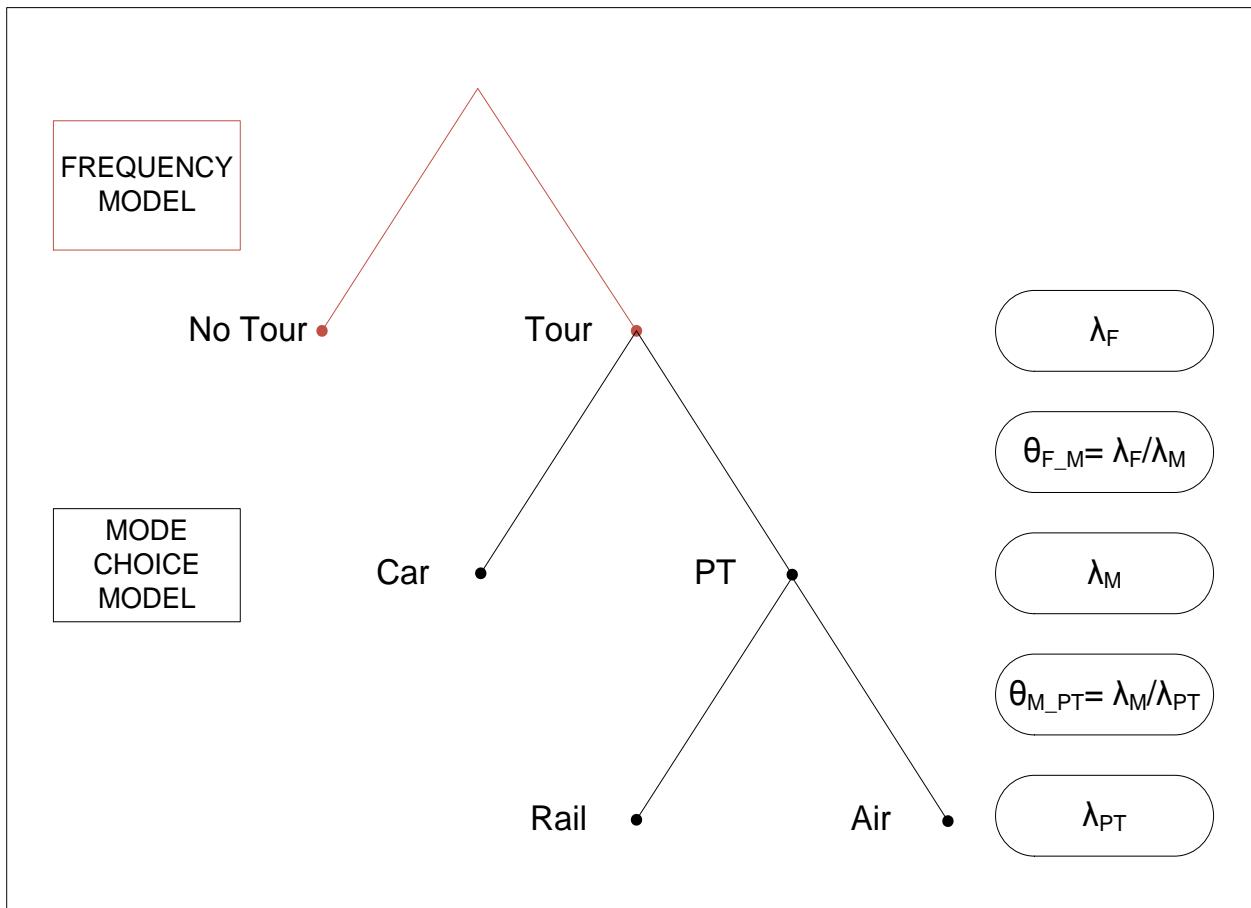
Mode Choice

6.2.26   The sensitivity to generalised cost is estimated at the lowest level in the structure (i.e. the rail versus air choice). The generalised cost parameter (GenCost)[59] defines λ_{PT} and allows the calculation of the relative sensitivities to generalised cost for the car versus public transport choice λ_M, using the estimated structural parameters θ_{M_PT}. For commute, there is no rail versus air choice, so the generalised cost parameter estimated in the model defines λ_M directly.

6.2.27   For the Other purpose, the estimated model produced sensitivities of the car versus PT (λ_M) and rail versus air (λ_{PT}) choices which were very similar, with a relative value of 0.91 (for the PT nest parameter, θ_{M_PT}).

6.2.28   The air mode share in the NTS data is just 0.9%, so there is very limited data available to identify differences between the sensitivities for the rail versus PT and the PT versus highway choices. By contrast, 24% of the leisure respondents to the 2009 SP survey were existing air users, and therefore there is much more information available from the SP data to identify a value for the PT nest parameter. Tests were therefore undertaken whereby this relative sensitivity θ is imported from the SP analysis, where the PT nest parameter was 0.72 in a comparable model specification.

---

[59] Note that these are in "tour" units, and need to be adjusted for use in PFM, as explained in section 6.4.

6.2.29　Constraining the PT nest parameter to the value from the SP analysis results in only a small loss of fit to the data of 1.5 log-likelihood points[60]. Hence this revised mode choice model was accepted and used to calculate updated mode choice logsums for the estimation of revised frequency models. Note that the SP data was not suitable for the estimation of the frequency model.

6.2.30　The results are set out in table 6-2, and an explanation of the parameters is given in Table 6-3.

Table6-2: Demand Model Estimation results (mode choice)

	Business		Other		Commute	
Observations	4653		12742		2594	
Final Log (L)	-1824.1		-2624.7		-749.4	
no. of Parameters	10		9		6	
Rho² (0)	0.504		0.691		0.578	
Rho² (c)	0.207		0.107		0.156	
Parameter estimates (see Table 6-3 for explanation)						
Crdsb0					-1.335	(-8.1)
Crdsb1	-2.363	(-7.9)	-1.534	(-8.0)		
Crdsb2			-1.119	(-5.8)		
Rail_0205	0.08997	(0.5)	-1.64	(-9.7)	0.08039	(0.3)
Rail_0610	0.509	(2.7)	-1.424	(-7.8)	1.448	(5.0)
Air_0205	-0.616	(-2.6)	-2.651	(-11.7)		
Air_0610	-0.9177	(-3.7)	-2.634	(-10.4)		
RL_male			-0.5191	(-4.7)		
Car_male	0.8124	(5.5)			0.691	(4.2)
Car_ptwrk	-0.3337	(-1.6)				
ge2cars			0.3958	(3.7)	1.06	(7.6)
RL_ge2cars	-0.5609	(-4.6)				
GenCost	-0.00606	(-10.5)	-0.00414	(-16.3)	-0.00489	(-8.1)
θ_{M_PT}	0.8731	(9.0)	0.7179	(*)	n/a	

*(constrained to the value identified from analysis of the 2009 SP data)

Table6-3: mode choice model parameter definitions

Parameter	Definition	Bus	Other	Comm
Crdsb0	Distance term on car for one-way distances of ≥75 miles,	n/a	n/a	✓

[60] For an explanation, see Section 3.3.

	introduced gradually from 75 to 100 miles, then constant, using function: min(1,max(0,(dist-75)/25))			
Crdsb1	Distance term on car for one-way distances of ≥100 miles, introduced gradually from 100 to 150 miles, then constant, using function: min(1,max(0,(dist-100)/50))	✓	✓	n/a
Crdsb2	Distance term on car for one way distance of ≥150 miles, introduced gradually from 150 to 200 miles, then constant, using function: min(1,max(0,(dist-150)/50))	n/a	✓	n/a
Rail_0205	ASC on rail, 2002-2005 records (2002-2005 car is the base mode)	✓	✓	✓
Rail_0610	ASC on rail, 2006-2010 records (2006-2010 car is the base mode)	✓	✓	✓
Air_0205	ASC on air, 2002-2005 records (2002-2005 car is the base mode)	✓	✓	n/a
Air_0610	ASC on air, 2006-2010 records (2006-2010 car is the base mode)	✓	✓	n/a
RL_male	Term reflecting lower probability of males choosing rail	n/a	✓	n/a
Car_male	Term reflecting higher probability of males choosing car	✓	n/a	✓
Car_ptwkr	Part-time workers less likely to choose car	✓	n/a	n/a
ge2cars	Individuals from households with 2+ cars more likely to choose car	n/a	✓	✓
RL_ge2cars	Individuals from households with 2+ cars more likely to choose rail	✓	n/a	n/a
GenCost	Sensitivity to generalised cost for return tour	✓	✓	✓
θ_{M_PT}	nest parameter defining relative sensitivity of the car vs PT and rail vs air choices	✓	✓ (*)	n/a

*Since air is not available for Commute, the parameter strictly relates to the car vs rail rather than car vs PT choice.

6.3   Frequency model

6.3.1   The frequency model predicts the binary choice between not travelling and making a long-distance tour on an average weekday. The NTS LD data (recall plus diary) covers a two or four-week period, hence a period of 10 or 20 weekdays. On each weekday, if an individual does not travel, they contribute a 'no tour' observation. If a tour is made where one leg is made on a weekday, and the other leg on a weekend, then it is included in the counts of weekday tours made with a weight of 0.5. Tours where both legs are made on a weekend are excluded from the tour counts.

6.3.2   To estimate the frequency models, mode choice logsums are required for both CA and NCA individuals. As noted earlier, the mode choice model parameters are estimated from CA individuals only. Further, only individuals who make a tour are included. This requires some ingenuity to provide the required quantities for the frequency estimation.

6.3.3   In what follows, the model specification conventions relate to units which differ somewhat from the standard 'generalised cost' approach. Rather than the cost for a one-way trip, the units are the 'utility' (U) for a round trip or 'tour', consistent with the approach used for the LDM. To maintain the integrity of the description, it has not been attempted to harmonise them, though appropriate commentary is provided. However, in the following section on model application, the conventional notation and terminology is reverted to.

6.3.4   For CA and NCA individuals who make a tour, mode choice logsums U* are calculated for the IJ journey actually made. For the CA segment, the following formulae are used:

Equation 6-3

$$U*(CA)_{PT,PA} = \ln\left[\exp(U_{Air,PA}) + \exp(U_{Rail,PA})\right]$$

Equation 6-4

$$U*(CA)_{M,PA} = \ln\left[\theta_{M\_PT}.\exp(U_{Car,PA}) + \theta_{M\_PT}.\exp(U*(CA)_{PT,PA})\right]$$

where:

- U*(CA) $_{PT,PA}$ is the logsum for the PT nest for the chosen P/A pair (in utility units);
- U*(CA) $_{M,PA}$ is the overall mode choice logsum for the chosen P/A pair (in utility units);
- $U_{Air,PA}$ is the utility of air for the chosen P/A pair;
- $U_{Rail,PA}$ is the utility of rail for the chosen P/A pair;
- $U_{Car,PA}$ is the utility of car for the chosen P/A pair; and
- θ_{M_PT} is the relative sensitivity of PT mode and main mode choices (see figure 6-1 above).

The 'utility' values U are equivalent to $-\lambda_{PT}*C^{tour}$, plus mode-specific and other estimated constants, as discussed in Section 3.3.

6.3.5   For the NCA segment, where only rail is available, the logsums are generated using the rail LOS for the NCA segment:

Equation 6-5

$$U*(NCA)_{M,PA} = \log\left(\theta_{M\_PT} \exp(U_{Rail,PA})\right)$$

6.3.6   Different approaches were investigated to incorporating non-travellers (i.e. those who were not observed to make a tour) in the estimation of the frequency models, in terms of the attraction zones they might have travelled to had they made a tour.

6.3.7   The preferred approach was to calculate logsums from the mode choice models for individuals who make tours, and then for each production zone, calculate average logsums from the sample of tours made from that production zone for individuals who make no tours. Note that this approach requires that the mode choice and frequency models are estimated sequentially, rather than simultaneously.

6.3.8   Therefore, for non-travellers, average P/A logsums for their production (home) zone are calculated separately for CA and NCA individuals over the set of attraction zones visited by all individuals (i.e. both CA and NCA) who do make a tour from that production zone, using the following formulae:

Equation 6-6

$$U*(CA)_{M,P} = \frac{1}{n}\sum\nolimits_{n} U*(CA)_{M,PA}$$

Equation 6-7

$$-\quad U*(NCA)_{M,P} = \frac{1}{n}\sum\nolimits_{n} U*(NCA)_{M,PA}$$

where:

- U*(CA) $_{M,P}$ is the average CA logsum for production zone P over chosen P/A pairs;
- U*(NCA) $_{M,P}$ is the average NCA logsum for production zone P over chosen P/A pairs; and
- n is the total number of tours (CA plus NCA) observed from production zone P, made to the set of P/A pairs observed from production zone P: for a given production zone there may be multiple tours to the same attraction zone.

6.3.9   On this basis, for the no tour versus tour (frequency) choice, it is possible to calculate the sensitivity of travel frequency to changes in generalised time λ_F, using the structural parameter θ_{F_M}, which is multiplied by the mode choice logsum U*$_M$. This is the key output from the frequency model estimation as it defines the relative sensitivity of the frequency and car vs. PT choices.

6.3.10   In line with the constrained treatment of the "Other" segment in the mode choice estimation, different values of the θ_{F_M} parameter were tested and the impact on the model fit was assessed. On the basis of the change (reduction) in goodness of fit and the resulting elasticity values, it was judged acceptable to constrain the parameter to a minimum value of 0.6 (this compares with values for commute and business which are 0.56 and 0.47, respectively). The higher value of θ_{F_M} that was originally estimated resulted in cross-elasticities that were considered low in comparison with external evidence.

6.3.11   The models estimated on this basis are presented in table 6-4, and an explanation of the parameters is given in table 6-5.

Table6-4: Demand Model Estimation results (frequency)

	Business	Other	Commute
Observations	83229	185954	73452
Final Log (L)	-28477.3	-67819.7	-17963.1
Parameters	7	9	7
Rho² (0)	0.963	0.96	0.982
Rho² (c)	0.056	0.032	0.064

Parameter estimates (see Table 6-5 for explanation)

Fage2959	-0.6831	(-17.4)			-0.5681	(-11.3)
Fagelt29			0.2347	(9.9)		
Fretired			0.1123	(4.3)		
Funemply			-0.3169	(-4.6)		
Fhhwcld			0.3036	(13.5)		
Fmale	-1.316	(-35.4)	-0.03231	(-1.7)	-1.602	(-29.6)
LDT0610CA	5.531	(91.0)	3.962	(176.4)	7.441	(102.9)
LDT0610NCA	5.43	(29.2)	2.725	(38.6)	7.278	(32.8)
LDT0205CA	5.763	(95.3)	3.926	(193.3)	7.101	(103.9)
LDT0205NCA	5.32	(33.3)	2.627	(49.1)	7.014	(40.5)
θ_{F_M}	0.5066	(30.1)	0.6	(*)	0.5731	(26.3)

*(constrained by analogy with the results for Business and Commute)

Table6-5: frequency choice model parameter definitions

Parameter	Definition	Bus	Other	Comm
Fage2959	Individuals aged 30-59 less likely to make no tours, i.e. make more tours than individuals aged <30 or 60+	✓	✓	n/a
Fagelt29	Persons aged under 30 more likely to make no tours	n/a	✓	n/a
Fretired	Retired persons more likely to make no tours	n/a	✓	n/a
Funemply	Unemployed persons less likely to make zero tours, i.e. more likely to travel	n/a	✓	n/a
Fhhwcld	Households with children more likely to make no tours	n/a	✓	n/a
Fmale	Males less likely to make no tours, i.e. make more tours than females	✓	✓	✓
LDT0205NCA	Zero tour constant, 2002-2005 data, no car available individuals	✓	✓	✓
LDT0205CA	Zero tour constant, 2002-2005 data, car available individuals	✓	✓	✓
LDT06210NCA	Zero tour constant, 2006-2010 data, no car available individuals	✓	✓	✓
LDT0610CA	Zero tour constant, 2006-2010 data, car available individuals	✓	✓	✓
θ_{F_M}	Relative sensitivity of frequency and car vs. PT choices	✓	✓	✓

6.4   Model application

6.4.1   The combined mode choice and generation model is applied incrementally, pivoting from Base Matrices. In future years, these are the 'do minimum' [DM] demand estimates (Chapter 10). 'B' is used to denote these, regardless of the year to which they apply.

6.4.2   Associated with the base or DM pivot matrices in any given year will be a set of generalised cost matrices, obtained by loading the base or DM matrices on the relevant networks. These matrices are denoted by G_0: section 10.9 explains how they are derived in practice.

6.4.3   The demand model works with generalised cost, and the specification is compatible with Eq (6-1), except that for the application, everything is on a trip – rather than a tour – basis. Therefore:

Equation 6-8

$$C_{mIJ,y}^{pc} = GT_{mIJ,y}^{pc} + \mathbf{M}_{mIJ,y}^{p} / VoT_{y}^{p}(D_{IJ})$$

6.4.4   The cost elements are calculated for the year y to which they relate. In the case of rail, GT is obtained as GJTCAE from the SCM, while for car and air, GT is derived from the corresponding networks.

6.4.5   The first requirement is to calculate the difference ∆C between C and the pivot costs G^0:

Equation 6-9

$$\Delta C_{mIJ,y}^{pc} = C_{mIJ}^{pc} - G_{mIJ}^{0pc} \qquad \forall I,J,p,c,m$$

6.4.6   Note that in standard future-year application, the fares for rail and air will be the same in the 'do minimum' [DM] and 'do something' [DS] scenarios, and therefore fares have no impact on the generalised cost differences. Nevertheless, for consistency the fares are still included as part of the generalised cost calculation. Similar considerations apply to all the elements of GT for the air mode. In most circumstances, ΔC will be zero for the air mode.

6.4.7   In calculating C for highway, it was noted in the section on estimation that the WebTAG formulae for fuel consumption were applied using the actual network speed for each IJ movement. In model application, however, it was considered that implementing this detailed procedure would introduce significant complexity and, after further investigation, it was considered acceptable to work with a single average speed across the network. Hence overall average car cost per kilometre values (on a per-person basis) are calculated using these average speed values, together with the advice given in WebTAG A1.3 for forecasting changes in car costs over time, and divided by occupancy.

6.4.8   The detailed calculations for the application of the demand model are now set out, with the operation summarised in figure 6-2. This should be read from bottom to top, going up the right hand side and down the left, though in practice the interface with the other models means that there is an element of iteration. Figure 6-2 ignores demand segments, but in fact the model needs to be run independently for each purpose and car availability [pc] combination.

6.4.9   At the lowest level, starting from the input of the SCM and the Air assignment, there is a 'PT choice' between rail and air (for Commute, this choice does not apply, since the air mode is not considered available).

Figure 6-2: The Demand model
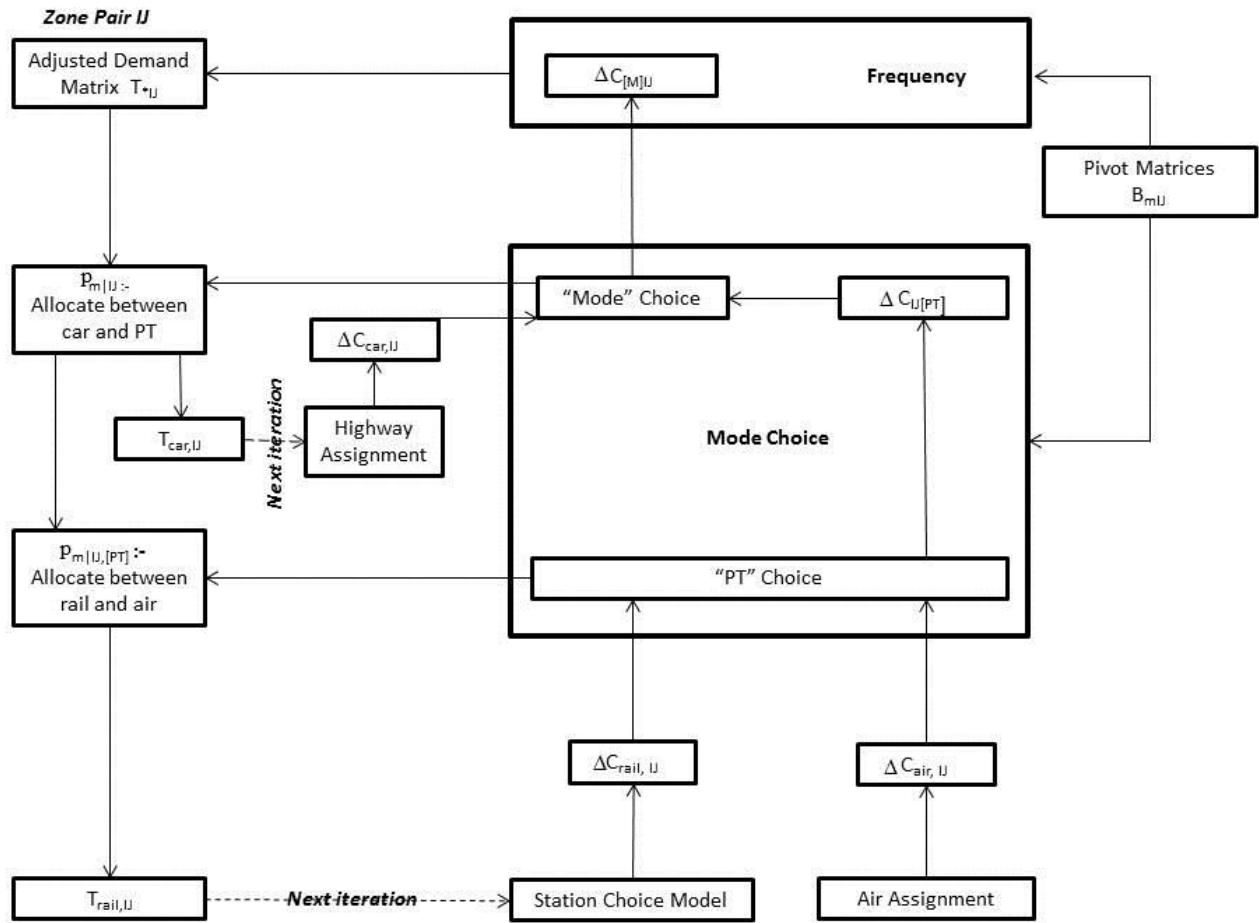
PT choice

6.4.10   To calculate the choice between rail and air, the following calculations are made:

Equation 6-10

$$p_{m|IJ}^{pc} = \frac{B_{mIJ}^{pc}.\exp(\lambda^{p}{}_{PT}.\Delta C_{mIJ}^{pc})}{\sum\limits_{m'\in\{rail,air\}} B_{m'IJ}^{pc}.\exp(\lambda^{p}{}_{PT}.\Delta C_{m'IJ}^{pc})}$$

6.4.11   By the properties of the incremental logit model, this allows the calculation of the incremental 'composite' cost $\Delta C_{[PT]IJ}$:

Equation 6-11

$$\Delta C_{[PT]IJ}^{pc} = \frac{1}{\lambda^{p}{}_{PT}} \ln\left(\frac{\sum\limits_{m'\in\{rail,air\}} B_{m'IJ}^{pc}.\exp(\lambda^{p}{}_{PT}.\Delta C_{m'IJ}^{pc})}{\sum\limits_{m'\in\{rail,air\}} B_{m'IJ}^{pc}} \right)$$

This is then passed up to the next choice level – 'mode choice' between car and PT.

Mode choice

6.4.12   The choice between car and public transport is calculated by:

Equation 6-12

$$p_{m|IJ}^{pc} = \frac{B_{mIJ}^{pc}.\exp(\lambda^{p}{}_{M}.\Delta C_{mIJ}^{pc})}{\sum\limits_{m'\in\{PT,car\}} B_{m'IJ}^{pc}.\exp(\lambda^{p}{}_{M}.\Delta C_{m'IJ}^{pc})}$$

Again, this allows the calculation of the incremental 'composite' cost $\Delta C_{[M]IJ}$:

Equation 6-13

$$\Delta C_{[M]IJ}^{pc} = \frac{1}{\lambda^{p}{}_{M}} \ln\left(\frac{\sum\limits_{m'\in\{PT,car\}} B_{m'IJ}^{pc}.\exp(\lambda^{p}{}_{M}.\Delta C_{m'IJ}^{pc})}{\sum\limits_{m'\in\{PT,car\}} B_{m'IJ}^{pc}} \right)$$

6.4.13   This is then used in the Frequency model, where a slightly different formula is applied to provide an estimate of revised total travel (by all modes) between I and J for segment pc:

Equation 6-14

$$T_{*IJ}^{pc} = B_{*IJ}^{pc}.\exp(\lambda^{p}{}_{F}.\Delta C_{[M]IJ}^{pc})$$

where the asterisk indicates summation over all three modes.

Revised Demand calculations

6.4.14   Given the mode choice probabilities, the revised demands T are now calculated by 'going down the tree':

Equation 6-15

$$T_{mIJ}^{pc} = T_{*IJ}^{pc} \cdot p_{m|IJ}^{pc} \ , \quad m' \in \{PT, car\}$$

and

Equation 6-16

$$T_{mIJ}^{pc} = T_{[PT]IJ}^{pc} \cdot p_{m|IJ}^{pc} \ , \quad m' \in \{rail, air\}$$

6.4.15   In this way, the output demand for each mode is derived, which can be passed, in the case of car, to the highway network, and in the case of rail, to the SCM. Note that the air network is not sensitive to demand, so further assignment is not strictly required for this mode[61].

6.4.16   It can be seen that apart from the pivot demand matrices and the generalised cost matrices, the only other requirement for the model is the set of λ values. These are derived from the model whose estimation was described in the previous section.

6.4.17   Note that the values presented in this section are applied to both NCA and CA segments. The λ values define the model sensitivities to generalised cost changes in minutes, and are presented here in trip units as required for the PFM implementation. This requires the estimated λ_{PT} values to be multiplied by 2, and then to be multiplied by the estimated values of θ_{M_PT} and θ_{F_M} in order to calculate λ_M and λ_F. Table 6-6 summarises the lambda values used to implement the frequency and mode choice models in incremental form.

Table 6-6: λ values for Demand Model Application

Choice		Business	Other	Commuting
Frequency	λ_F	-0.0054	-0.0036	-0.0056
Car v PT	λ_M	-0.0106	-0.0059	-0.0098
Rail v air	λ_{PT}	-0.0121	-0.0083	n/a

---

[61] although, as will be seen in Chapter 10, it is in fact carried out.

Part 3 - Other component models

In this part we discuss the other main model components of PFM: the 3 Regional PLANETs (PLANET South, Midland, and North) in Chapter 7, and the spreadsheet application for Heathrow - the Heathrow Airport Model (HAM) - in Chapter 8.

7 Regional PLANETs

7.1 Introduction

7.1.1 This chapter describes the three Regional PLANET models - PLANET South, Midlands and North, noting their general similarities and the specific aspects where they are different. These models complement the PLD rail network and ensure that all rail services are represented in key areas of interest. By means of the 'control matrices', it is ensured that any one PLD zone-to-zone movement is represented in only one of the four PLANET models.

7.1.2 PLD has priority as the representation of strategic rail, and the control matrices have been devised with this in mind. However, to represent crowding properly, the overlaps between the models need to be carefully handled. This is generally done with "pre-loads", but a different method, referred to as 'wormholes', is used for representing the impact of PLD flows on the PS network.

7.1.3 The Regional PLANETs are important, not only for the representation of potential crowding relief on local services brought about by HS2, but also for allowing a representation of revised local services to take account of released capacity.

7.2 Overview of Regional PLANET Models

7.2.1 The Regional PLANET models (South, Midlands, North) are all AM peak period rail passenger models covering the full three-hour period between 07:00 and 10:00 on a typical weekday. The zoning systems were described in Section 2.5.

7.2.2 PLANET South is the oldest, and differs in some respects from the other two, which share development history, functionality and data sources. Within PFM, all three models use the EMME network and assignment platform. The 'transit lines' are built using the same approach as for PLD (this is described in Section 9.5). The assignment uses the standard (frequency-based) Optimum Strategy[62] algorithm, and the MSA damping method to control convergence, as described in Chapter 4.

7.2.3 There is a full representation of National Rail services within the area covered by each model, which implies that this includes some services within the region also represented in the PLD model. However, as noted in Section 2.3, whether OD movements are dealt with in one of the Regional PLANETs or in PLD Demand is indicated by means of a "Control Matrix". Demand matrices for each model are initially provided covering all movements, and then the control matrices remove demand for non-selected movements.

7.2.4 Interfaces are required:

> a) to export to the PLD Assignment the level of passenger loadings from intra-Regional demand forecast on longer distance services; and

---

[62] Given that the emphasis of these regional AM models is more on high frequency peak services than on fast services, there is no significant value in considering the FJTS version of the algorithm.

b) to import to the Regional PLANETs the level of passenger loadings from PLD demand forecast on Regional services.

7.2.5 The general approach to the interface makes use of 'pre-loads', but in the case of b) for PS, the 'wormholes' method is used to import demand from PLD. The interfaces are a key aspect of PFM, and are described in Section 7.6.

7.2.6 Unlike the PLD assignment, which is strictly station-to-station, the Regional PLANETs deal with zone-to-zone movements, so that access and egress are explicitly modelled: again, there are some differences in this respect between PS and PM/PN.

7.2.7 In addition, the models use an elasticity-based approach to represent the change in demand at the OD level caused by changes in service specification, crowding or fares. For the purpose of applying elasticities, based on PDFH guidance, changes in generalised time are calculated by comparing two scenarios or cases: a "do minimum" (base) and a "do something" (test). While the elasticities generate changes in rail demand, there is no capability to understand where that demand has come from or identify modal shift or the impact of changes in competing modes, such as the use of the private car, as there is no comparable representation of the costs of car travel.

7.2.8 The Base Year Demand is compatible with LENNON station-to-station data for FY 2014/15. The distribution to zones of ultimate origin and destination uses the postal sector information reported in the 2004-5 National Rail Travel Survey (NRTS) and, in the case of PS, LATS. Demand is split by the three purposes - Business, Other, Commute - and, for PM/PN, further by car availability [CA/NCA]. This is described in detail in Section 9.2.

7.2.9 The models offer:

- rail route choice where reasonable route choices exist;
- demand responses (elasticities) varying with travel purpose (business, other (leisure) and commuting);
- response to congestion on the rail modes (crowding);
- a range of outputs to inform scheme development and decision making:
- statistics on passenger flows – such as number of passengers, travel time and distance travelled;
- passenger flows by route;
- levels of train passenger crowding anticipated;
- fare revenue by operator group; and
- economic benefit calculations

7.2.10 The network models do not take into account the fare paid by the passenger and it is considered that their route choice is unaffected by the fare paid. In terms of demand responses, the elasticities applied have been have been calibrated to GJT, without any fare term. The impact of changes in fares is captured exogenously using EDGE, as described in chapter 10.

7.2.11 In the following three sections, the individual models are described in somewhat more detail.

7.3   PLANET South

7.3.1 The PS model has been in use since the 1990s for modelling forecast crowding on the London & South East rail network and associated impacts on London Underground lines. The model represents morning peak period rail movements within an area which covers the former Network South East area, with less detailed representation beyond those areas: it is a three-hour period assignment. PS was adapted for PFM to include only trips within London, South East and South West areas.

7.3.2 In addition to National Rail services, the London Underground (LUL) and light rail (DLR) services are represented, based on data provided by Transport for London. This is less detailed than for the National Rail system, but ensures the model is able to adequately represent the role of LUL services as a feeder for longer distance travel, or as a competitor where appropriate. Similarly, but to a lower level of detail, it includes a representation of the feeder system provided by London Bus services within Greater London. This additional network representation ensures passenger dispersal within Greater London is adequately represented and evaluated.

7.3.3 Demand data is presented on a zone-to-zone basis, and each zone centroid is connected to one or more appropriate stations. There are 1,372 zones.

7.3.4 For each station pair (average) GJTC is calculated using the frequency-based (optimal strategy) described in chapter 4, and then for any zone-to-zone pair, the minimum cost route, including access/egress, is calculated over all possible station pairs available to the zone pair. As a result of MSA damping, this may in practice lead to multiple station allocations for any one zone pair. The base year network for PS is shown in figure 7-1 below:

Figure 7-1: PLANET South Base Year Network
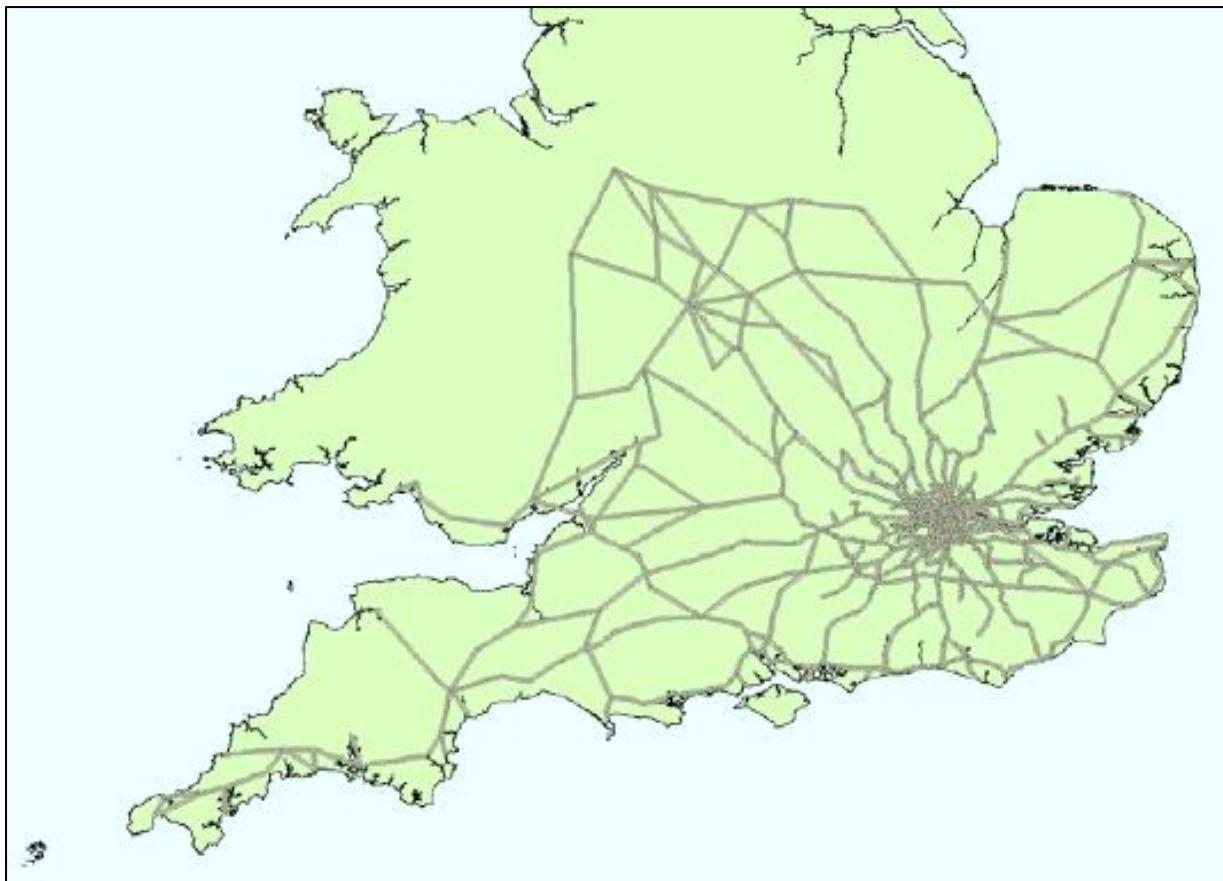
7.3.5   Generalised cost weights were reported in table 2-2. These are generally in line with WebTAG recommendations, and based on PDFH. They have not been specifically changed for PFM, given the long standing of the model.

7.4   PLANET Midlands

7.4.1   Together with PN, this was developed in 2009 and is centred on Birmingham with the East Midlands (Derby, Nottingham and Leicester) also within the core area.

7.4.2   The zoning system is based on aggregations of MSOA zones. PM has 1146 zones (59 of them external), and 259 stations. The network for PM is shown in figure 7-2 below:

Figure7-2: PLANET Midlands Network
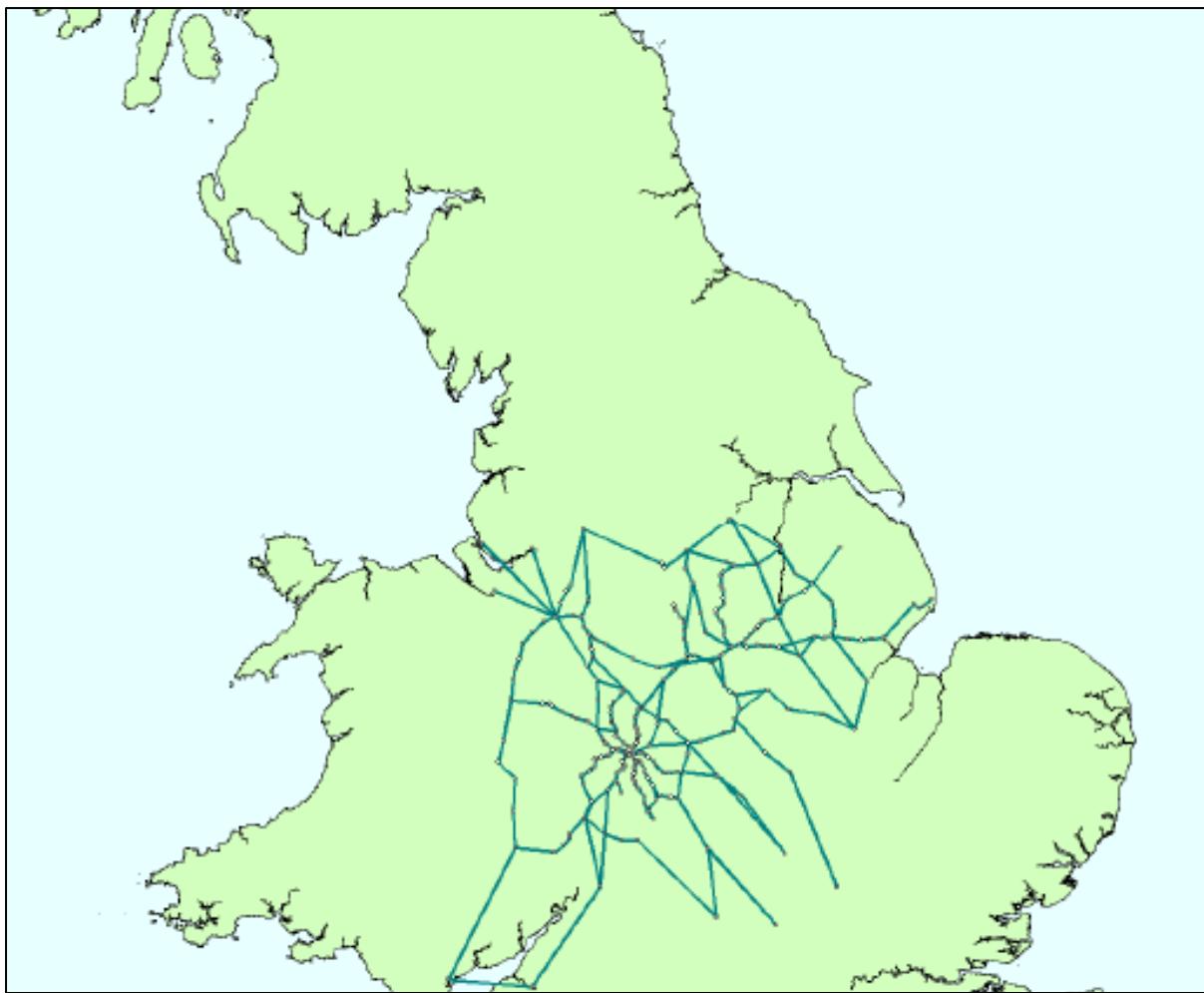
7.4.3   Generalised cost parameters (common between PM and PN) were shown in table 1-2.

7.4.4   In contrast to PS, PM has the zone-to-zone demand matrix for the three purposes further segmented by CA/NCA. There is then a simplified version of the PLD SCM to allocate to stations, as are now described.

7.4.5   For each zone, a maximum of 5 origin stations and 5 destination stations (i.e. a matrix of 25 possible routes) is allowed. In addition, zone-station-station-zone combinations with less than 0.1 journeys per AM peak (or less than 1 journey per fortnight) are excluded. The result is that there is a pre-allocation of zone-to-zone movements to a station-to-station matrix, which is then assigned. Access times to rail stations are based on the NAM, which calculates the travel time from each Census Output Area [COA] population centroid to each of the five nearest stations (in terms of travel time) for car available and non-car available households. The following assumptions have been made:

- Only one time of day has been modelled, with arrival at the station for 08:00 hours. Early arrival of up to 30 minutes is allowed without a time penalty so that the choice of this single time can be representative of AM peak journey times.

- Car available (CA) journeys assume that the car driver can park adjacent to the

station without a time penalty searching for a car park space.

- Non-car available (NCA) journeys use walk and bus but exclude rail and cycle.

- Access times are capped at 90 minutes.

7.4.6   The station choice model then makes use of the incremental logit formulation, where the allocation of station-station journeys between origin and destination zones pivots off NRTS data, with separation of CA and NCA access and egress.

7.4.7   For each zone pair [IJ], the model predicts the allocation between up to 25 (5*5) RS routeings (the set denoted as Z^c_{IJ}) when rail service interventions, or changes to access/egress, affect the relative attractiveness of alternative stations/routes.

7.4.8   The model has the following structure:

Equation 7-1

$$p^{pc}_{RS|IJ} = \frac{q^{pc}_{RS|IJ}.\exp(\lambda.\Delta GT^{pc}_{IRSJ})}{\sum_{R'S'\in Z^C_{IJ}} q^{pc}_{R'S'|IJ}.\exp(\lambda.\Delta GT^{pc}_{IR'S'J})}$$

where:

- p_{RS} is the proportion of people choosing to travel via stations R and S in the test scenario;
- q_{RS} is the proportion of people choosing to travel via stations R and S in the base scenario;
- Z^c_{IJ} is the choice set (dependent on car availability) of station pairs for those travelling between zones I and J;
- λ is the logit dispersion parameter. For PM this took the value -0.1
- $\Delta GT_{RS} = GT_{RS} – GT^o_{RS}$; where
    - GT_{RS} is the Generalised Time of travelling via stations R and S in the test scenario;
    - GT^o_{RS} is the Generalised Time of travelling via stations R and S in the base case.

7.4.9   To obtain a first approximation of the AM peak journeys that would be abstracted from other stations by the opening of a new station, a hybrid logit is used. For each affected flow, the hybrid logit uses a standard multinomial formulation to estimate the shares of trips attracted and produced by the new station, and then reallocates the remaining demand between the existing stations using the incremental formulation.

7.4.10   To reduce run times, the code limits these calculations to zone-zone flows where the new station would be closer than the 5th most popular CA/NCA station (according to the NRTS input), and also prevents access/egress above a threshold distance set to 10km.

7.5   PLANET North

7.5.1   As with PM, this was developed in 2009 and is centred on the M62 Liverpool/Manchester-Leeds/Sheffield corridor. The zoning system is based on aggregations of MSOA zones, and the base year demand matrix is derived in the same way. PN has 961 zones, and 511 stations. The demand data is compatible with that for PM, as are the generalised cost parameters.

7.5.2   The network for PN is shown in figure 7-3 below:

Figure 7-3 PLANET North Network
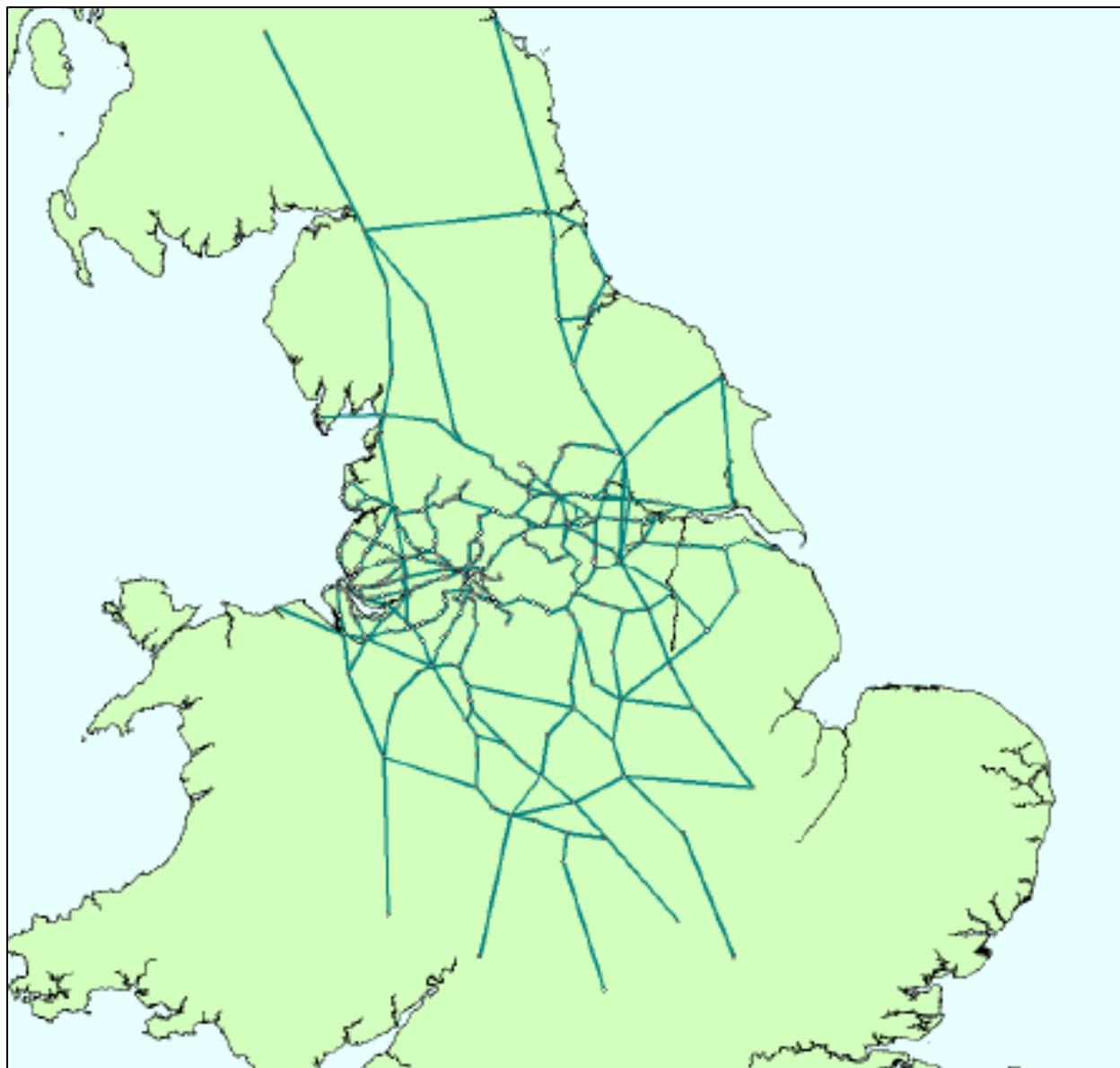
7.5.3   PN has an identical station choice procedure to that in PM, except that λ takes the value of -0.03 rather than the -0.1 for PM.

7.6   Interface with PLD

Introduction

7.6.1   It was noted in Section 7.2 that the interfaces between PLD and the Regional PLANETs are in two directions:

- to export to the PLD Assignment the level of passenger loadings from intra-Regional demand forecast on longer distance services; and

- to import to the Regional PLANETs the level of passenger loadings from PLD demand forecast on Regional services.

7.6.2   Figure 7-4 aims to illustrate the general requirements schematically. The top part of the figure illustrates, on the left-hand side, the 'Control Matrix' (see below), assuming that zones are generally numbered from north to south, with journey length increasing away from the main diagonal. According to this, each OD pair is uniquely assigned to one of the four PLANET models (PLD and the Regional PLANETs). Note that within the zonal range of PN and PM, some of the movements are in fact handled in PLD, as discussed below.

7.6.3   However, as the right-hand side shows, there is an overlap between the services represented in PLD and those in the Regional PLANETs. If the demand matrix allocated to each PLANET model were simply assigned to all services in that model, the contribution of the other demand to those services would be missed, with potential consequences for the modelling of crowding. The aim of the interface is to deal with this problem, in the manner suggested in the bottom part of the figure.

Figure 7-4: Schematic representation of interfaces between PLD and Regional PLANETs.
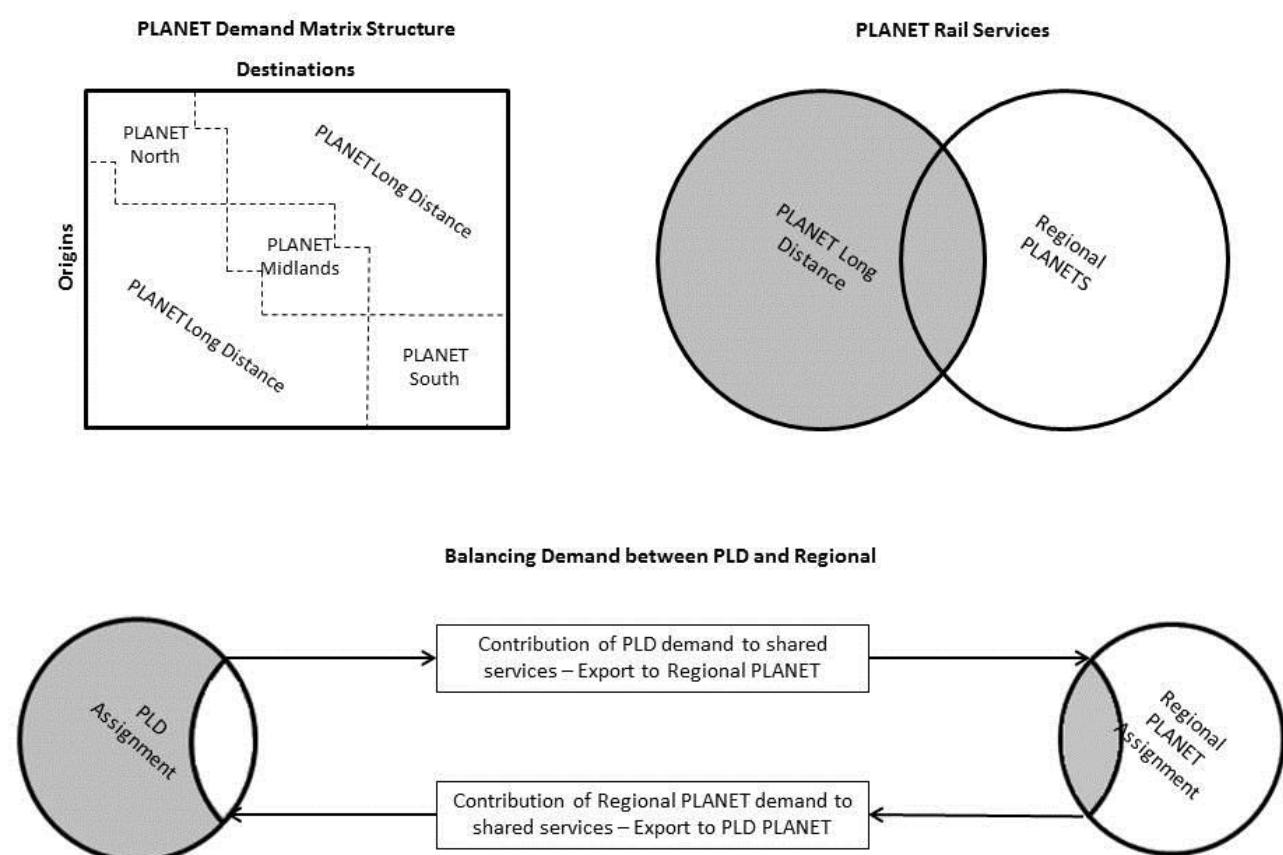
Control matrices

7.6.4   The PLANET models have initially been developed as free-standing models, so that their scope is to deal with all relevant services. Careful consideration is required in order to determine how best to use the 'strategic' PLD model and the more detailed Regional PLANETs in combination.

7.6.5   Beginning with the definition of strategic services as represented in PLD, the Control Matrix takes into account what elements of rail demand at a zone-to-zone level are best represented in PLD, bearing in mind the scope of the Regional PLANETs and the focus on the HS2 corridor, with the proposed distribution of stations. This is most straightforward in the case of PS, since apart from the London stations, there are no HS2 stations within the detailed area covered. Hence, all movements within an area bounded approximately by a line from the Severn Estuary to The Wash can be removed from PLD.

7.6.6   The Control Matrix was initially developed for the PLD zoning system. The PLD zones were then linked to their corresponding zones in PM and PN to derive control matrices for these Regional models. This approach ensures that there is no demand represented in both PLD and a regional model and ensures that all movements have a defined model within which to reside.

7.6.7   With respect to PM and PN, there are two elements to consider where the demand for a particular movement should reside – the travel to work [TTW] area for the major conurbations of interest, and local trips using key strategic corridors.

7.6.8   It is important that movements within each TTW area are represented within the appropriate regional model. The TTW areas used in the model are based on the Department for Transport TTW area definitions that aim to reflect areas where the bulk of the resident population also work within the same area.

7.6.9   Figure7-5 shows the TTW areas diagrammatically. Trips wholly within each individual TTW area are modelled within the appropriate regional PLANET model, i.e. Birmingham and East Midlands in PM, the remainder in PN.

7.6.10   Figure 7-5 shows that a number of the TTW areas overlap. For example, zones in the eastern section of the Liverpool TTW area are also in the western section of the Manchester TTW area. This necessitates building the control matrix carefully, to ensure that only those origin-destination zone pairs that constitute local trips are excluded from the PLD demand matrices.
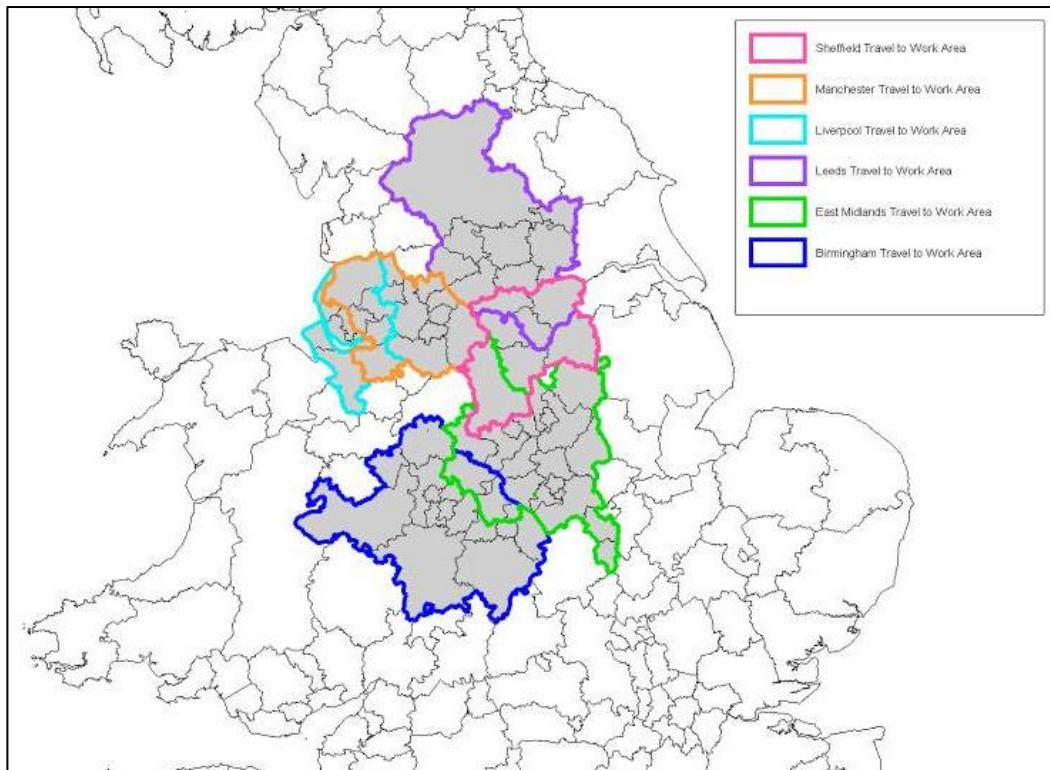
Figure 7-5: PLD travel to work areas

7.6.11   It is also important to ensure that local trips along the key rail corridors affected by HS2 are captured and modelled in the regional models. The three key corridors are the WCML (Blue), MML(Green) and ECML (Red), shown in Figure 7-6. Along each of these rail corridors the control matrix ensures that local trips are modelled in PM or PN (as appropriate). Note that the Cross Country route between Birmingham and Sheffield is dealt with in the Regional PLANETs since all adjacent OD pairs through which the line runs are within JTW areas (Birmingham, East Midlands or Sheffield).

Figure 7-6: Key Corridors
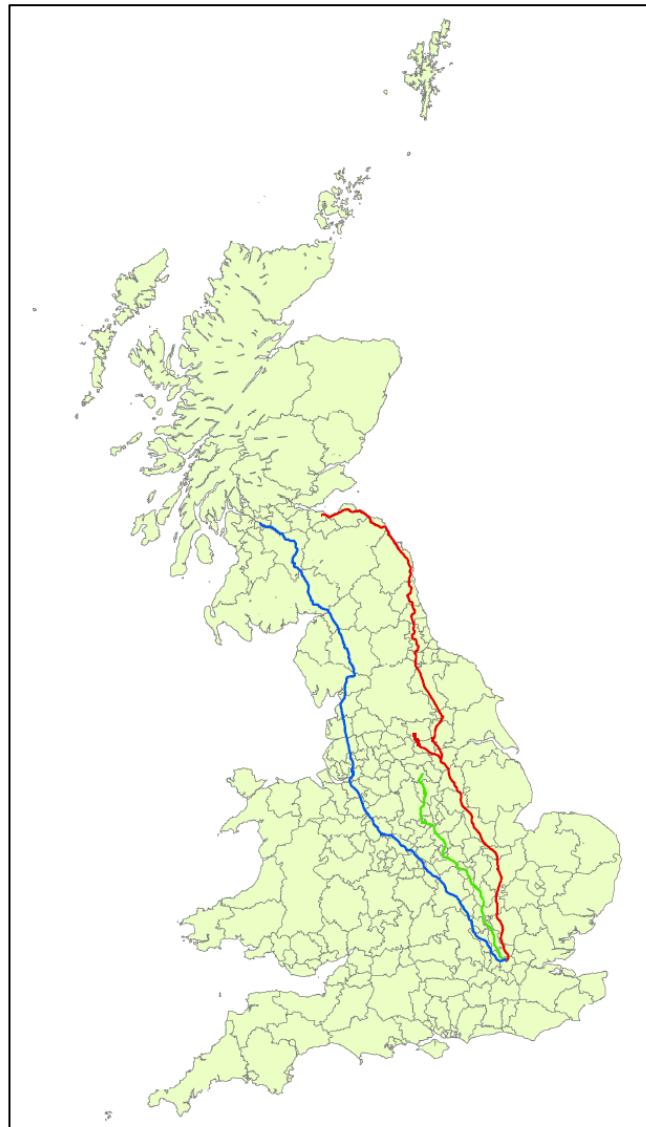
7.6.12   Local trips have been defined as trips between adjacent zones along the key corridors. This ensures that, for example, demand from Crewe to Warrington is included in PN, but demand from Crewe to Preston is modelled in PLD.

7.6.13   The control matrix removes the following trips from the scope of PLD:

- all trips wholly within defined journey to work areas for Birmingham, the East Midlands, Liverpool, Manchester, Sheffield and Leeds;

- trips between neighbouring PLD zones within key sections of the WCML, MML and ECML; and

- all trips internal to the PLANET South model area.

- all trips internal to Scotland and internal to part of Wales

7.6.14   Consequently, these movements are all catered for in the appropriate Regional PLANETs, while those movements which remain in the scope of PLD are removed from the Regional PLANETs.

Classification of services

7.6.15   In respect of transferring loadings from the Regional PLANETs to PLD [item a) in paragraph 7.2.4], the basic process is to identify the local demand in the Regional model which has been assigned to long distance services, and calculate the level of that local demand in units of passenger demand per train per hour. These pre-loads are then exported from the Regional model and imported into PLD.

7.6.16   The scope of the Regional PLANETs can be conceived in terms of boundaries or "cordons". A service is considered as long distance for a Regional PLANET model if it has at least two stops within the cordon and proceeds beyond the cordon.

7.6.17   The initial step in the process is to define a series of "dummy" services to represent the corridors of interest; then relevant services in those corridors are "flagged" and, finally, the local demand pre-load values are calculated and exported.

7.6.18   In respect of transferring loadings from PLD to the Regional PLANETs [item b) in paragraph 7.2.4], for PM and PN the demand is transferred from PLD as a passenger pre-load on selected services. This is done in terms of passengers/train/hour and represents those PLD passengers assigned to use services classified as long distance within the Regional PLANET, thus interacting with local passengers.

Pre-loads

7.6.19   In simple terms, the process for identifying and transferring pre-loads between the models relies upon two manual processes. First, the identification of the sections of train tracks (links) where pre-loads are required and, second, the identification of the train services (transit lines) which carry both 'local' and long-distance demand on those track section (links). This process enables these transit lines to be included in the pre-load process only on the applicable links, even if the transit lines traverse more than one pre-load area. This becomes even more crucial as pre-loads are included for short distance movements along strategic corridors as well as in urban areas.

7.6.20   Technically, it is feasible to match individual services in PLD and the regional models in order to directly pass actual demand to the correct service in the other model. However, the services modelled in PLD and the regional models are not necessarily consistent. For example, services may run only outside the AM peak period and therefore would not be included in the regional models. Conversely, service patterns and utilisation may be different in the AM peak when compared to an all-day model. Furthermore, the process of coding transit lines using Network Rail CIF files leads to inconsistent service numbering between PLD and the regional models, so direct matching of services is not a straightforward process. As a result, such a detailed approach is complex and labour intensive.

7.6.21   The process of producing the transit lines from CIF files for PLD and the regional models leads to a number of distinct service specifications within the PLANET models, many of which are slight variations in stopping pattern or timings rather than being substantially different services. For this reason, when transferring from the Regional PLANETs to PLD, similar transit lines are grouped into "packets", so that pre-load demand is passed between similar services in the local models and PLD. Each packet is given a unique code that is present in both the local model and PLD, and pre-load transfer occurs between these sets of transit lines at a packet level.

7.6.22   Before the process can take place, certain manual tasks need to be undertaken, namely:

- defining which services are eligible for pre-load transfer in the Regional PLANETs and PLD;

- defining which packet to each of these services belongs; and

- defining the AM Peak to all-day factor for each TOC.

7.6.23   Note that HS2 services do not fall into the category of "eligible for pre-load transfer" due to the restriction on travel between Old Oak Common, Heathrow and Euston.

7.6.24   First, the links that are within the core areas of the regional models are identified. In PM and PN these are largely within the TTW areas (for example Manchester to Stoke-on-Trent). Links in PS are defined as those on strategic corridors (for example, the WCML). This is because services that are not on strategic corridors will not be carrying long-distance demand. This shows that choice of links and services are inter-related.

7.6.25   Second, the subset of services which are eligible to be used in the pre-load process are defined. These services necessarily need to be able to transport strategic demand and local demand. Therefore, the service needs to have rail service origin and rail service destination in different TTW areas (for strategic demand criteria, i.e. if they were both in the same area, then the control matrix would mean all demand would be modelled in the designated regional model); and have at least two stops in a given TTW area (for local demand included in regional model to be able to make use of service).

7.6.26   Services are then grouped together into 'packets' in PLD and regional models if they are similar in terms of the stations they call at and the markets they cater for. This enables pre-load allocations to take account of journey times, stopping patterns and individual services. To remove the tidality of AM Peak demand, the demand is averaged between the outward and the reverse link in the Regional PLANETs.

7.6.27   As PLD is an all-day (0700-2300) model and the regional models are all AM peak period (0700-1000) models, factors (F) are required to convert between time periods, both AM Peak period to all-day and all-day to peak period, to provide a representation of crowding impacts over the appropriate time periods. To transform peak flows to all-day and vice versa, demand profiles from MOIRA have been used to create suitably robust factors, given in paragraph 7.6.31.

7.6.28   Factoring demand is important, as demand profiles vary considerably over the course of a day. For example, consider demand on long-distance West Coast services to and from Manchester. Long-distance demand on the West Coast Mainline is heavily skewed towards London in the early morning and a corresponding return peak in the late afternoon/early evening. Demand on these services modelled in PLD is balanced in both directions as the PLD model is an all-day model. When transferring long distance demand on these services from PLD to the regional model (PN in this case), demand profiles become important because in the AM peak period covered by PN, long-distance demand on services to London is much higher than on services from London.

7.6.29   Conversely, local demand using the same long-distance services also requires factoring when transferred from a Regional PLANET to PLD. For example, AM peak flows modelled in PN (such as Stockport – Manchester or Macclesfield to Manchester) are heavily biased towards Manchester in the AM peak, but for inclusion in the all-day PLD such flows are directionally balanced and should be factored before transfer to PLD.

Pre-Load Factors From Regional PLANETs to PLD

7.6.30   The following equation is used to calculate the pre-load factors from regional PLANET to PLD:

Equation 7-2

$$X_{PLD} = \frac{X_{RP} + X_{RP}{}^{r}}{2}.F.\frac{T_S}{T_P}$$

where:

- X_{PLD} = Pre-load in PLD
- X_{RP} = Regional PLANET packet flow
- X_{RPr} = Regional PLANET reverse packet flow
- F = AM Peak – All-day factor
- T_S = Trains per day
- T_P = Trains per day in PLD packet

7.6.31   The factoring mechanism allows for factors converting PM, PN and PS AM peak demand (local short-distance flows) to PLD all-day demand to be disaggregated by TOC. However, analysis of local demand using MOIRA data suggests that there is little variation between TOCs. Therefore a factor of 2.67 is applied to all TOCs in the PM and PN areas, while in the PS area a factor of 3.2 was found to be more appropriate. This factor converts two-way AM peak segment flows on a link for a specific service to an all-day flow (thereby ensuring pre-loads in PLD are directionally balanced).

From PLD to Regional PLANETs

7.6.32   The factor to convert PLD demand to regional model demand varies by the type of movement of the transit line, as there are strong tidal flows into London in the AM peak. As such, factors are provided for Non-London, To London and From London services, as shown in Table. These factors are used to apply directionality to the PLD to PM and PN Regional model pre-loads. An alternative 'wormhole' approach is used in the PS area, as discussed below.

Table7-1:All-day to peak period conversion factors

Direction	Factor
Non-London	0.22
To London	0.34
From London	0.05

Source: MOIRA demand profiles

Wormholes[63]

7.6.33   "Wormholes" are effectively dummy zones at or near the PS cordon: there are 10 of them related to existing services, and a further three to deal with HS2. By means of a "select link" analysis[64] at each cordon point in PLD, origin-destination matrices of demand passing through the point are established. After scaling from all-day to peak-only levels, these are converted to PS zones.

7.6.34   The demand data for each cordon point is then accumulated into a single matrix for all the cordon points, and the resulting demand is assigned onto long-distance services in PS, with the PS assignment routines allowing the demand to reach their final destinations within the PS area using any other rail or underground services.

---

[63] The wormhole terminology is used to represent the process of the artificial demand transfer from PLD to PS
[64] A 'select link analysis' is an option within an assignment procedure which allows the demand matrix contributions to the load on a specified link to be identified.

7.6.35   This has the advantage of enabling the long-distance demand to be assigned in a realistic manner on the more detailed PS network. The 'wormhole-based' demand is allocated to a separate 'user class', and is not subject to the standard demand (elasticity) responses in PS. In addition, it is not evaluated in the economics, as it is already included in PLD, and is purely included in PS to provide suitable crowding levels in the model, and to enable the patterns of long distance demand dispersal to be better represented and understood. It will, however, through the crowding mechanism, affect the demand responses for "local" trips within PS.

7.6.36   The PS cordon is shown in figure 7-7. For existing services, 10 links crossing this cordon are identified in PLD: Newport - Cardiff, Newport - Hereford, Gloucester - Cheltenham, Moreton - Evesham, Banbury - Leamington, Wolverton - Rugby, Wolverton - Northampton, Bedford - Wellingborough, Peterborough - Leicester, Peterborough - Grantham. A wormhole zone for PS is associated with each link: note that the links are bi-directional.

Figure 7-7: Location of PS Cordon

7.6.37   For HS2 there is a further wormhole, but in the output this is split into three matrices – one for Euston trip ends, one for Old Oak Common trip ends, and one for all others. Provision can also be made for a possible HS2 link to Heathrow.

7.6.38   For each of these links, in both directions, a select link analysis is performed in PLD assignment – separately by user class. So for link a, user class u, a (sub-) matrix $T_{a,}{}^{u}{}_{RS}$ between PLD stations R and S is obtained, representing all the movements using link a, in either direction.

7.6.39   These PLD matrices needs to be factored to allow for the fact that PS is only for AM peak. This is done using MOIRA2 profiles by departure time - there are 240 of these, defined on:

- three journey purposes (business, commute, leisure);

- five flow categories (To London, From London, To Blue[65] , From Blue, Other);

- two journey legs (Outward, Return); and

- eight journey time bands (0-20, 21-50, 51-100,101-140, 141-180, 181-270, 271-360 and 361-999 minutes).

Each profile implies a specific factor to convert from full day to AM peak.

7.6.40　For PFM purposes, some of these are combined. Given that for PS the majority of demand is to/from London, the distinction between To and From 'Blue' has been dropped. Since the MOIRA2 profiles are based on departure times, they can be used directly for the 'From London' movements, but for the 'To London' movements an adjustment has to be made to determine which trips arrive in London within the AM Peak period, using the average journey time to London for the given journey time band.

7.6.41　Separately by purpose and direction, each RS pair is allocated to one of the profiles[66], and hence the relevant factor t^{pd}_{RS} is obtained: these factors multiply the select link matrices $T_{a,}^{u}{}_{RS}$.

7.6.42　In the standard case, one of R or S will be within the remit of PS: in some cases, however, neither station is (eg, South Wales to North East England). In principle, the end which is outside PS will be converted to the corresponding wormhole zone 'a', while the end within PS needs to be disaggregated to a PS zone, as the zone systems differ between PLD and PS. This is done by aggregation correspondence lists held within EMME. The aggregation correspondence lists associates PLD stations with their PLD geographic zones, so that the select matrices are output from PLD in geographic zone format. These are then disaggregated into PLANET South zones where appropriate by another ensemble.

7.6.43　For non-London zones, the disaggregation was simply based on what proportion of trips in PS made up each PLD zone, making allowances for where PS zones cut across PLD zone boundaries. For London zones, the Railplan distribution (based on the distribution at the London end of trips to Manchester) was found to be much more suitable, and this distribution was conveyed to the Greater London zones in PS via a simple correspondence.

7.6.44　Note also that where R and S are both outside PLANET South, trip ends are re-aggregated to the wormhole zones by an automated spreadsheet process which uses the distribution of select demand flows across the PS cordon to distribute trips across the wormhole zones. This effectively replicates the assignment of the select demand but truncated at the PS cordon.

7.6.45　The resulting factored matrices can be combined (by direction, user class and wormhole) and passed as input to PS assignment.

---

[65] In MOIRA, 'Blue' refers to a major commuting destination station excluding London.
[66] NB because the profiles are defined on journey time (defined as In-vehicle time plus the Number of Interchanges multiplied by 30 minutes, derived from PLD model skims for IVT and Number of Boardings minus 1.), it is possible that the allocation could be scheme-dependent, though this will not normally be the case.

Implementation

7.6.46   The way in which the interfaces are achieved within the overall PFM structure is described in Chapters 10 and 11.

8 Heathrow Airport Access model [HAM]

8.1 Introduction

8.1.1 In this chapter the Heathrow Airport Access model [HAM], which is a spreadsheet adaptation of the free-standing LASAM[67] model is described. It is a specialist application to deal with two specific segments of HS2 passenger demand: diversion to HS2 of current Heathrow surface access trips in the HS2 corridor – excluding trips from London; and diversion to HS2 of air passengers that take a domestic flight to/from Heathrow connecting with an international flight to/from Heathrow.

8.1.2 HAM modifies the demand for rail trips to Heathrow in the light of changes in the rail network, and specifically the introduction of HS2. The way in which HAM interacts with the other components of PFM is explained in more detail in Chapters 10 and 11.

8.2 Overview

8.2.1 HAM deals with two specific segments of HS2 passenger demand that cannot be easily represented in PLD:

- diversion to HS2 of current Heathrow surface access trips in the HS2 corridor – excluding trips from London[68]; and

- diversion to HS2 of air passengers that take a domestic flight to/from Heathrow connecting with an international flight to/from Heathrow.

8.2.2 These two market segments are modelled using a spreadsheet mode choice model, drawing upon knowledge from LASAM; this is described in Section 8.4.

8.2.3 LASAM has been adapted to a simplified spreadsheet format so that it can be used to predict the mode choice made by air passengers to access Heathrow Airport. Because LASAM is only concerned with surface access, an additional access mode, Domestic Air, has been included to deal with the second segment mentioned above. This relates to the domestic legs of passengers on international air trips (interliners) rather than domestic passengers, who are dealt with in PLD.

8.2.4 One of the key simplifications is that HAM only represents air passengers that originate from non-London areas. Base and forecast cost component skims for rail, car and air are taken from PLD and are the key inputs to the HAM spreadsheet, as described in Section 8.8.

---

[67] London Airports Surface Access Model v2, created by SKM for BAA. BAA has given permission for the use of LASAM parameters for this project.
[68] There are already three rail options (and a fourth planned) for travel between London and Heathrow: this is not a market that HS2 is targeting.

8.2.5   Cost skims for other modes, such as coach and taxi, are provided from LASAM as a fixed input for each forecast year. HAM has been set up to allow different HS2 routes to be tested, although it was calibrated based on the assumption of HS2 passing through the West Midlands up to Manchester.

8.2.6   HAM produces forecasts of air, car, rail, and coach demand by zone, business and leisure passenger segments and direction for an annual average weekday.

8.3   Background

8.3.1   An analysis of CAA air passenger surveys from 2007 and 2008 at Heathrow Airport reveals the mode of transport used to access the airport[69]. Table contains the main mode[70] shares for all surface access trips to the airport. Car is the dominant mode for (Non-London) areas close to the airport, while rail gains a greater mode share as distance from the airport increases.

Table 8-1: Heathrow Airport Annual Surface Access Mode Shares, 2007/08 CAA Data [71]

Main Mode Share	All UK
Bus/Coach	8.4%
Rail	26.7%
Taxi	25.6%
Park and Fly	11.1%
Kiss and Fly	23.0%
Charter Coach	2.4%
Other	2.9%
Total Demand (Over 2 Years)	85,456,697

8.3.2   In principle, improved rail access to Heathrow Airport, especially from Northern Regions could be expected to attract significantly improved rail mode share. At the moment rail journeys from Heathrow to these regions often require complex interchanges via central London. High speed rail will also compete against the domestic air market where domestic air passengers transfer at Heathrow Airport for international destinations.

---

[69] Although more recent data is now available, the LASAM model has not been updated, and this would be a substantial undertaking.
[70] The CAA survey captures up to the final three modes used to access the airport, and the published data usually refers to the 'final mode'. However, LASAM was developed as a model of the choice of 'main mode', based on a careful analysis of the final three modes. For London origins, there is very little difference between the two definitions, but there are more significant differences for the longer non-London origin trips where the surface access journey is through Central London. For these journeys, the 'main mode' of travel may be defined as the mode used to access London; this particularly applies to coach or rail trips into London with a different final mode to the airport.
[71] CAA expansion, mode shares recalculated to reflect main mode rather than the final mode reported in CAA publications.

8.4 Methodology

Adaptation of LASAM

8.4.1 HS2 has been modelled along with all existing access modes to Heathrow Airport. To facilitate the choice of mode from each origin zone, LASAM v2 has been adapted to a spreadsheet model, known as the Heathrow Airport Access Model [HAM] with the following key features and simplifications:

- a focus on the study corridor: London – West Midlands – North West (excluding the London area);

- retaining the same hierarchical mode choice structure as LASAM;

- removing Heathrow Express, Underground, RailAir Coach and Airport Transfers as main mode options as they are only relevant to trips from London;

- adding Air as a public transport sub mode;

- retaining the same sensitivity parameters as LASAM;

- selecting an appropriate modal constant for Air; and

- using one zone to represent Heathrow. The central terminal area is used as a reference for level-of-service.

8.4.2 As the aim is to deal with international passengers using Heathrow who originate in the study corridor, international model coefficients and economic assumptions were adopted from LASAM rather than the domestic equivalents.

8.4.3 The HAM mode choice model is used to forecast the change in mode shares from a current situation and can therefore be referred to as an incremental model.

8.4.4 The HAM's structure, including all data inputs, is shown in figure 8-1.

Figure 8-1: Heathrow Airport Model Structure
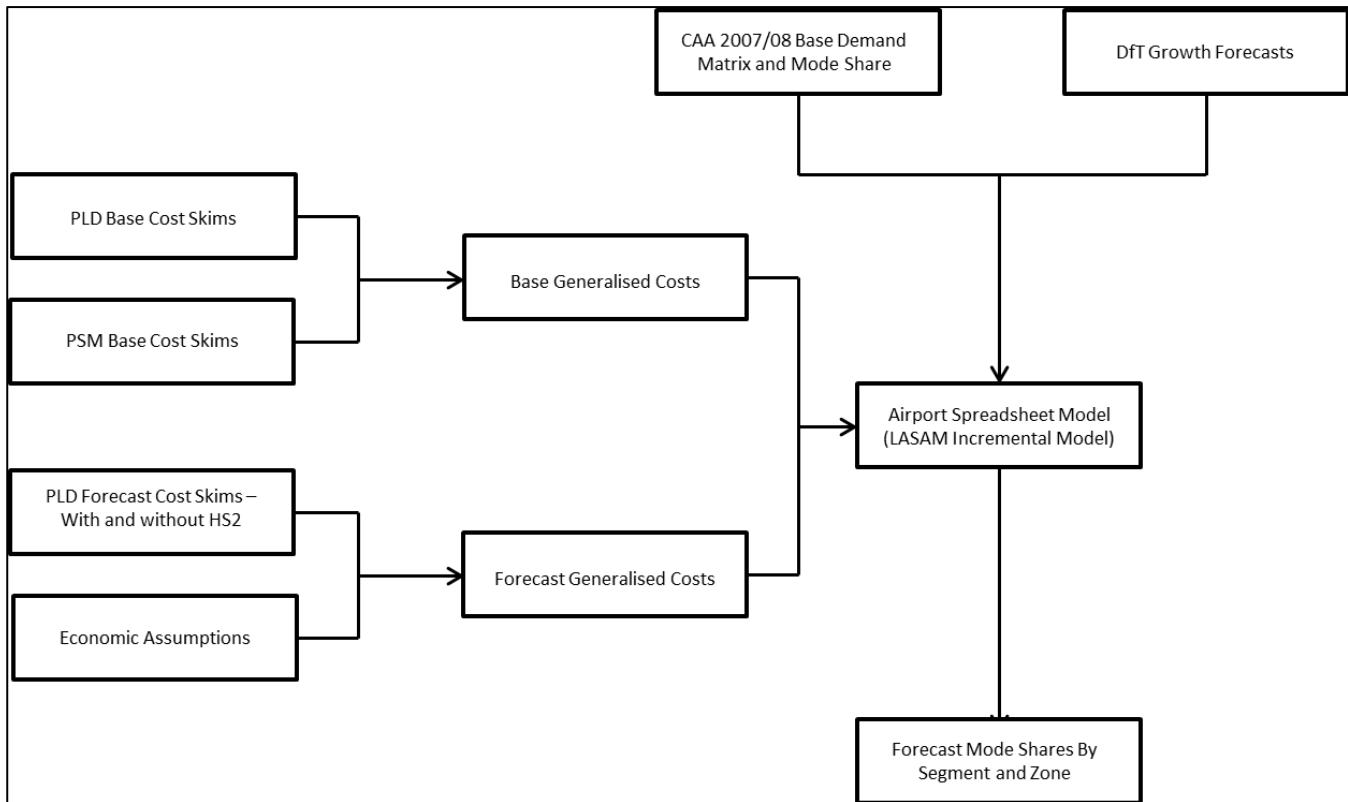
8.4.5   The base year was originally 2007/8, but this has been updated to 2014/15.

8.5   Air passenger segmentation

8.5.1   LASAM segments air passengers at Heathrow into 6 passenger segments. To limit the segmentation of data, and to remain consistent with PFM, HAM has aggregated these into the following four major segments[72]:

- UK Business – air passengers residing in the UK and travelling on business;

- Foreign Business – air passengers residing outside the UK and travelling on business;

- UK Leisure – air passengers residing in the UK and travelling for leisure purposes; and

- Foreign Leisure – air passengers residing outside the UK and travelling for leisure purposes.

8.5.2   PFM, being focused on UK rail journeys, segments passengers differently to LASAM. Table 8-2 shows the assumed equivalence between PFM and LASAM passenger segments. PFM also provides highway and air cost skims, the passenger segments of these differ for each mode and are described in table 8-3 and table 8-4.

---

[72] LASAM further splits the UK market segments into domestic and international destinations

Table 8-2: Equivalent segments of rail passengers

PFM Segment	HAM Equivalent Segment	Reason
Business	UK Business Foreign Business	Car available at home/workplace origin for departing air passengers
Other	UK Leisure Foreign Leisure	Car available at home/workplace origin for UK departing air passengers. Foreign travellers have option of being dropped off by friends/relatives (Kiss and Fly)

Table 8-3: Equivalent segments of road passengers

PFM Segment	HAM Equivalent Segment	Reason
Business	UK Business Foreign Business	Same time and distance skim for all air passengers, higher Vehicle Operating Cost (VOC) for business segment
Other	UK Leisure Foreign Leisure	Same time and distance skim for all air passengers, lower VOC

Table 8-4: Equivalent Segments of Air Passengers

PFM Segment	HAM Equivalent Segment
Business	UK Business Foreign Business
Other	UK Leisure Foreign Leisure

Time periods

8.5.3   PFM matrices represent an annual average weekday (16 hours) whereas LASAM models annual air passengers by four separate time periods, one representing the weekend and three to represent different time periods within a weekday. Details of the weekday time periods and how they relate to the CAA air passengers surveys are shown in Table 8-5: LASAM time periods (weekdays). The overall proportion of trips by time period is for combined data for 2007 and 2008, noting that the CAA air passenger survey is for departing air passengers and then scaled to represent all air passengers.

8.5.4   To be consistent with PFM, HAM does not distinguish between time periods. In order to use LASAM cost skims they are averaged using the weights listed in table 8-5.

Table8-5: LASAM time periods (weekdays)

Time period	Airport entrance	CAA survey time	2007/08 proportion
AM Peak (weekdays)	0700-1000	0900-1200	22%
PM Peak (weekdays)	1600-1900	1800-2100	17%
Interpeak (weekdays)	Rest of the day	Rest of the day	61%

8.5.5   The Heathrow Airport Model uses base data covering a full year. To convert this into an annual average weekday, for output to PFM, the CAA data for 2007 and 2008 was analysed to calculate the most appropriate factor. It was found that on average 121,800 air passengers access the airport by a surface mode on weekdays, compared to 107,900 on the weekend. Taking the average weekday total and dividing by the annual total gives a conversion of 0.28% of the annual air passengers on an average weekday.

8.6   Base year data

8.6.1   A base year matrix of annual air passengers by segment, origin and mode was created by combining surface access modes and domestic air passengers as described below. The base matrix represents all people that could switch to HS2 in order to access Heathrow Airport.

Surface access

8.6.2   CAA surface access mode shares and the overall total of air passengers at Heathrow Airport were found to be very consistent between 2007 and 2008. Using the same process as in LASAM, each air passenger was allocated a main mode of surface access based on the combination of modes used to access the airport as stated in the CAA survey. The resulting mode shares, excluding "other" modes such as walking and bicycle, are shown in table 8-6.

Table 8-6: Surface Access Main Mode Shares, Excluding Other Modes

Mode	2007	2008	2007/08 Average
Bus/Coach	8.3%	9.0%	8.7%
Rail	26.5%	28.2%	27.4%
Taxi	26.5%	25.7%	26.1%
Park and Fly	12.3%	10.3%	11.3%
Kiss and Fly	23.0%	24.0%	23.5%
Charter Coach	2.7%	2.1%	2.4%
Airport Transfer	0.7%	0.6%	0.7%
Total PAX (million)	42.48	41.14	41.81

Domestic air access

8.6.3   Manchester, Edinburgh, Glasgow and Newcastle were the only airports considered for inclusion in the model.

8.6.4   CAA surveys on air passengers travelling between Manchester and Heathrow Airports were analysed to find out the proportion of number passengers transferring to another flight at Heathrow. Table 8-7: Domestic Air Passengers, Manchester to Heathrow shows that on average in 2007/08, 65% of air passengers on the Manchester-Heathrow route transfer to another flight at Heathrow. A small number also connect at Manchester or both airports, these trips are ignored along with the point-to-point trips.

Table 8-7: Domestic Air Passengers, Manchester to Heathrow

Connection Type	2007	2008	2007/2008 Average
Connect at MAN	23,375	8,645	16,010
Connect at Heathrow	642,759	575,207	608,983
Connect Both Ends	11,911	4,091	8,001
Point-to-Point	300,075	319,602	309,839
Total	978,120	907,544	942,832

8.6.5   The CAA survey data was analysed to see if any of these trips should be excluded based on characteristics that would suggest they would be unlikely to switch to HS2. Reasons may include:

- packaged flight deals which include the UK domestic leg at (or close to) zero fare - although it may be possible that airlines could arrange for the domestic leg to be provided instead by train – similar to Air France;

- transit passengers that do not have to leave the plane at Heathrow; and

- transfer passengers that have a simple connection at Heathrow, either with the same airline or a codeshare airline.

8.6.6   The analysis proved inconclusive with the following findings:

- "all inclusive packages" are on a steady decline, from 23% of transfer passengers in 2005 to 12% in 2008; and

- only two airlines fly between Manchester and Heathrow (British Airways and British Midland)[73]. These two airlines account for 48% of all flights in/out of Heathrow, implying that a high proportion of transfer passengers will naturally (rather than by specific choice) fly the domestic and international leg of their journey with the same airline.

8.6.7   For the other airports less information was available. It was therefore decided to include all transfer passengers in the analysis.

---

[73] This was the position in 2008. The two airlines were merged in October 2012.

8.7   Mode choice hierarchy

8.7.1   The airport spreadsheet model has adopted the same tree structure as LASAM with the following modifications:

- RailAir Coach (overall 0.6% mode share) treated as standard rail;

- Passengers arriving at the airport by London Underground or Heathrow Express are modelled as standard rail;

- Air added to PT (or equivalent) nest;

- Charter Coach fixed at the 2007/8 mode share by zone (overall 3.0% mode share);

- Other modes (3% mode share) ignored; and

- Air Transfer[74] not modelled as it is not valid within the catchment area.

8.7.2   The resulting mode choice hierarchy for each passenger segment is shown in figures, 8-3, 8-4 and 8-5. The added mode (Air) is highlighted in each diagram.

Figure 8-3: UK Business Mode Choice Hierarchy
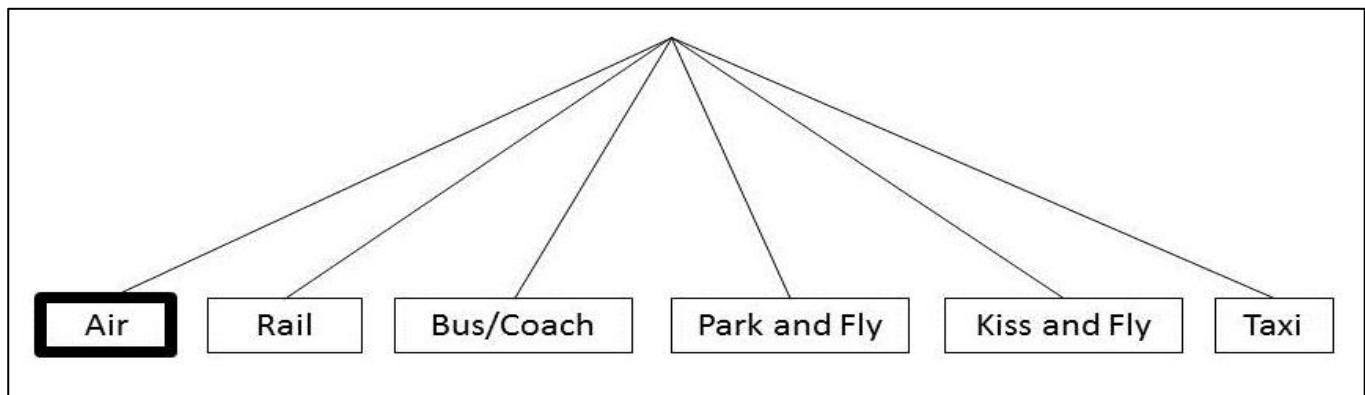
Figure 8-4: Foreign Business Mode Choice Hierarchy

---
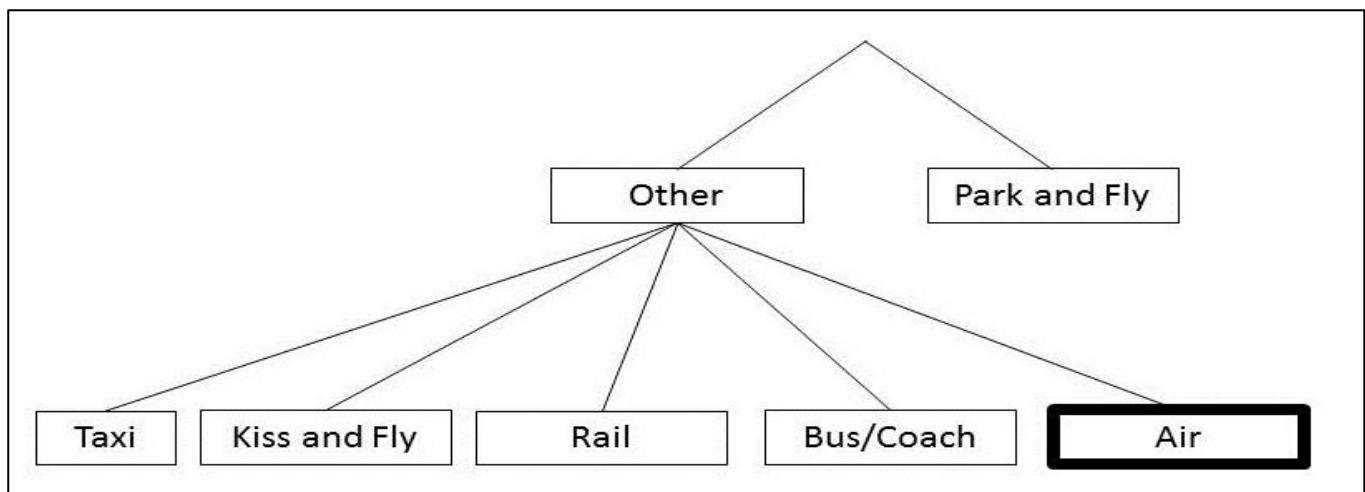
[74] The Air Transfer mode refers to air passenger transfers by designated coach between Heathrow, Gatwick and Stansted airports.

Figure8-5: Leisure Mode Choice Hierarchy
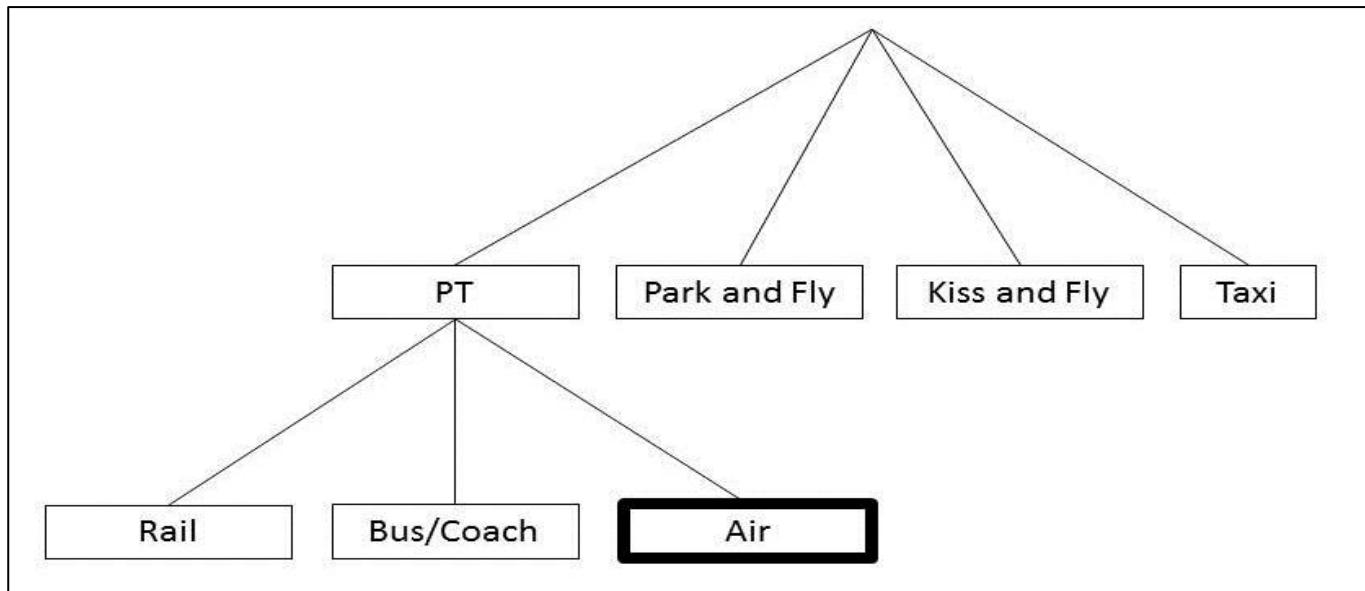
8.8 Generalised Cost data

8.8.1 To ensure the Heathrow Airport Model is as compatible with PFM as possible, where available, generalised cost skims from PLD are used in preference to those from LASAM. Details of the assumptions relating to which PLD skims are used are set out in Annex D.

8.8.2 Table D-3 of Annex D shows that the PLD Air skims do not include the check-in time. This item represents a considerable amount of "waiting time" which needs to be included in the generalised cost equation. LASAM applies a distribution of lead times to simulate the time it takes an arriving air passenger between entering the terminal entrance and the plane departure time.

8.8.3 Separate distributions are applied for business and leisure passengers, leisure passengers typically arriving at the airport earlier. To simplify this procedure, the average lead time has been extracted from LASAM and used in the Airport Demand Model. The implemented values are shown in table 8-9.

Table 8-9: Assumed Check-In Times

Segment	Departure lead time
Business	1hr 45min
Leisure	2hr 15min

8.8.4 Since transfer passengers are being modelled, there is a possibility that this "check-in" time is being double-counted for some passengers who have a streamlined check-in at Heathrow Airport.

8.8.5   As noted in Chapter 6, the PLD mode choice does not include coach, but coach travel is included in HAM, as this is an important access mode for Heathrow. Although, as shown in Table 8-1, the overall share of coach is only 8.4%, while that for rail is 26.7%, the pattern is very different if it is confined to the HAM 'study area' in which HS2 competes. The share of coach is now twice as high (16%), and almost on a par with the share of rail trips from the study area (17%).

8.8.6   Coach cost skims are derived from LASAM by aggregating time periods and applying the following assumptions:

- 2008 coach level of service from LASAM used as a base;

- no changes assumed to coach services in the catchment area in 2021/31;

- headway, access time and number of interchanges remain unchanged;

- base coach fares grown to forecast year using growth rates agreed with DfT for the Stansted Airport Planning Application;

- base coach IVT grown to forecast year using growth in highway times by zone; and

- where the PLD zoning system is more detailed than LASAM, the same cost is allocated to each PLDd zone. Where LASAM is more detailed, the costs from the most populous LASAM zone is applied.

8.8.7   PLD outputs highway times and vehicle operating costs, but it does not provide information on associated charges such as taxi/minicab fares and airport parking charges. Parking charges, parking duration and group size are applied by passenger segment as per LASAM.

8.8.8   Taxi/Minicab fares are extracted from LASAM as follows:

- it is assumed that no one uses the more expensive black cab from Non London zones (as LASAM), only Minicab;

- fare is the same regardless of time of day; and

- the base Minicab fares collected in August 2008 by SKM and used in LASAM are adopted.

8.9   Generalised cost equations

8.9.1   The components of generalised cost described in section 8.8 are combined to form the generalised costs by mode, segment and zone using the following equations:

UK Business Passengers Generalised Cost Formulae

Equation 8-1: Rail (L,S,X)

$$\frac{\alpha_p Time + \beta Wait + \delta Walk + \phi Access + \eta\ Fare}{(D)^\theta} + Z_R * (Interch.)$$

Equation 8-2: Bus/Coach

$$\frac{\alpha_p Time + \beta Wait + \phi Access + \eta\ Fare}{(D)^\theta} + Z_B * (Interch.)$$

Equation 8-3: Taxi

$$\left[\alpha_p Time + \eta\ \frac{Fare}{N}\right] \bigg/ (D)^\theta$$

Equation 8-4: Park and Fly

$$\left[\alpha_p Time + \eta\ \frac{(PCost + VCost)}{N}\right] \bigg/ (D)^\theta$$

Equation 8-5: Kiss and Fly

$$\left[\alpha_p Time + 2\eta\ \frac{(PCost + VCost)}{N} + \frac{\alpha_d Time}{N}\right] \bigg/ (D)^\theta$$

Equation 8-6: Air

$$\frac{\alpha_p Time + \beta Wait + \phi Access + \eta\ Fare}{(D)^\theta}$$

where *D* = Highway Distance, θ = 0.5, and *N* = Group size. (Applies to equations 8-1 to 8-6)

UK Leisure Passengers Generalised Cost Formula

Equation 8-7: Rail (L,S,X)

$$- \quad \frac{\alpha_p Time + \beta Wait + \delta Walk + \phi Access + \eta\ Fare}{(D)^\theta} + Z_R * (Interch.)$$

Equation 8-8: Bus Coach

$$\frac{\alpha_p Time + \beta Wait + \phi Access + \eta\ Fare}{(D)^\theta} + Z_B * (Interch.)$$

Equation 8-9: Taxi

$$\left[\alpha_p Time + \eta\ \frac{Fare}{N}\right] \bigg/ (D)^\theta$$

Equation 8-10: Park and Fly

$$\left[\alpha_p Time + \eta\ \frac{(PCost + VCost)}{N}\right] \bigg/ (D)^\theta$$

Equation 8-11: Kiss and Fly

$$\left[\alpha_p Time + 0.35\eta \; \frac{(PCost + VCost)}{N} + \frac{\alpha_d Time}{N} + \psi \, Time^2 \right] \Big/ (D)^\theta$$

Equation 8-12: Air

$$\frac{\alpha_p Time + \beta \, Wait + \phi \, Access + \eta \; Fare}{(D)^\theta}$$

where D = Highway Distance, θ = 0.6 and N = Group Size (Applies to equations 8-7 to 8-12)

Foreign Business Passengers Generalised Cost Formulae

Equation 8-13: Rail (L,S,X)

$$- \;\; \frac{\alpha_p Time + \beta \, Wait + \delta \, Walk + \phi \, Access + \eta \; Fare}{(D)^\theta} + Z_R * (Interch.)$$

Equation 8-14: Taxi

$$- \;\; \frac{\alpha_p Time + \beta \, Wait + \phi \, Access + \eta \; Fare}{(D)^\theta} + Z_B * (Interch.)$$

Equation 8-15: Park and Fly

$$\left[\alpha_p Time + \eta \; \frac{(HireCost + VCost)}{N} \right] \Big/ (D)^\theta$$

Equation 8-16: Kiss and Fly

$$\left[\alpha_p Time + 2\eta \; \frac{(PCost + VCost)}{N} + \frac{\alpha_d Time}{N} \right] \Big/ (D)^\theta$$

Equation 8-17: Air

$$\frac{\alpha_p Time + \beta \, Wait + \phi \, Access + \eta \; Fare}{(D)^\theta}$$

where D = Highway Distance, θ = 0.5 and N = Group Size (Applies to equations 8-13 to 8-17)

Foreign Leisure Passengers Generalised Cost Formulae

Equation 8-18: Rail (L,S,X)

$$- \;\; \frac{\alpha_p Time + \beta \, Wait + \delta \, Walk + \phi \, Access + \eta \; Fare}{(D)^\theta} + Z_R * (Interch.)$$

Equation 8-19: Bus/Coach

$$- \;\; \frac{\alpha_p Time + \beta \, Wait + \phi \, Access + \eta \; Fare}{(D)^\theta} + Z_B * (Interch.)$$

Equation 8-20: Taxi

$$- \left[\alpha_p Time + \eta \ \frac{Fare}{N} \right] \Big/ (D)^\theta$$

Equation 8-21: Park and Fly

$$- \left[\alpha_p Time + \eta \ \frac{(HireCost + VCost)}{N} \right] \Big/ (D)^\theta$$

Equation 8-22: Kiss and Fly

$$- \left[\alpha_p Time + 0.35\eta \ \frac{(PCost + VCost)}{N} + \frac{\alpha_d Time}{N} + \psi \ Time^2 \right] \Big/ (D)^\theta$$

Equation 8-23: Air

$$- \frac{\alpha_p Time + \beta Wait + \phi \ Access + \eta \ Fare}{(D)^\theta}$$

where D = Highway Distance, θ = 0.6 and N = Group Size (Applies to equations 8-18 to 8-23)

8.9.2   The generalised cost parameters used in the formulae are given in table 8-10 below for the base year.

Table 8-10: Generalised Cost Parameters

	2008			
	UK Business	UK Leisure	Non UK Business	Non UK Leisure
Value of time (Heathrow) p/min	73.6	27.01	64.77	26.97
Vehicle Operating Cost p/km	11.79	5.39	5.39	5.39
Time Coefficient α(p)	0.18	0.2	0.22	0.25
Wait Coefficient β	0.49	0.55	0.47	0.66
R_Walk Coefficient ϭ	0.17	0.25	0.22	0.3
Access Coefficient ψ	0.55	0.96	0.93	1.17
Rail Interchange Coefficient	0.81	0.61	0.44	0.74
Bus Interchange Coefficient	1.63	0.9	0.44	1.09
K&F time Coefficient 2 α(d)	0.13	0.22	0.02	0.1
K&F time Coefficient3 Ψ	-	0.001	-	0.002
Distance exponent θ	0.4	0.5	0.4	0.5

The variable η is used to define the coefficient of monetary cost (inverse of the Value of Time). In LASAM this variable is income segment specific, but for application purposes an average value is calculated, implying the following values of time (£/hr, for 2004):

UKB: 66.14, UKL 24.79, FB: 58.21, FL: 24.75.

8.9.3   These have been updated to 2010 values for the HAM:

UKB: 70.00, UKL: 25.94, FB: 61.61, FL: 25.90.

Part 4 - PFM model assembly and forecasting

In this part a description of how the networks and the base demand matrices were built up is provided. Following this, in two steps the use of the model for forecasting is explored. The first step relates to the "do minimum" case, where the growth of demand over time, as well as the expected changes to the networks is described. The second step relates to the introduction of HS2 and other associated changes to the rail network.

9 Base year demand and networks

9.1 Introduction

9.1.1 The Base Matrices represent the demand in the base year (FY 2014/15), and, with the exception of the air mode, are used as the foundation for forecasting the future year demand in specific years. Thus, they represent an essential component in the overall PFM.

9.1.2 Extensive work has been carried out to derive consistent and reliable base year matrices. The emphasis of the work relates to the scope of the PLD matrices (see Section 2.3). The three modes (rail, car and air) were derived independently, but cross-checks have been carried out to ensure modal consistency.

9.1.3 Separate matrices are required for the three purposes (Commuting, Business and Other), and as noted in Section 2.3, they are segmented between car available and non-car available (though for both the car and air modes the matrices are assumed to relate only to car available travellers – see section 9.4).

9.1.4 The following three sections provide an overview of how the matrices were developed. As was noted in 2.2, the matrices have generally been prepared on an OD basis, though the methodology has been designed to provide them on a P/A basis as well. They relate to average weekdays. To convert to total annual demand (for the purposes of appraisal) a set of mode-specific annualisation factors are used, reported in Chapter 12.

9.1.5 Section 9.5 then describes how the base year rail network was derived, both for PLD and the Regional PLANETs. PLD Rail Assignment was described in detail in Chapter 4, and this applies – with some minor modifications noted in chapter 7 – to the regional PLANETs. Section 9.6 describes how the base year rail fares were derived. For the remaining two modes (car and air), the base networks and assignment methods are discussed together (Sections 9.7 and 9.8).

9.1.6 As a result this chapter presents all the necessary input to the base year models. The way in which the demand matrices are modified for future growth is described in Chapter 10, where changes to the Do-Minimum networks are also described. The additional network changes for the Do-Something case are described in Chapter 11.

9.2 Base year rail matrices

9.2.1 In this section it is described how the base year rail matrices have been derived. In contrast with the car and air modes, where only long-distance movements are needed, base rail matrices are required not only for PLD but also for the three Regional PLANETs. While general principles are followed, there are some differences in approach which reflect both the context and the stage within the modelling at which the work was done. The greatest attention is paid to the PLD matrices, as these are of most importance to the assessment of HS2.

9.2.2   Note that in model application, not all the elements in these matrices are in fact used because of the need to ensure that each flow is only represented once. This is managed by means of the 'control matrices' (see section 7.6).

9.2.3   In all cases, the primary source of data is the 'LENNON' ticket sales database, which provides total annual ticket sales (by ticket type) and revenues, for every combination of origin-destination stations. However, this data does not provide a purpose breakdown nor an association with ultimate origin and destination (zone), and for this purpose the National Rail Travel Survey [NRTS], described in section 5.3, is used. In addition, it should be noted that for trips to (and within) Travelcard areas such as London many of the trips in the LENNON data will be to Travelcard zones and will not identify individual stations.

9.2.4   The LENNON data relates to all trips during the year, while PFM only models travel choices/patterns within a single weekday (or, in the case of the regional PLANETs, the weekday AM peak). Hence, the production of PFM rail demand matrices requires the LENNON journeys database to be 'deannualised', removing travel at weekends, and, in the case of the regional PLANETs, estimating the proportion of weekday trips occurring in the AM peak[75]. Factors are dependent on flow distance and based on ORCATS (Operating Revenue Computer AllocaTion System, a data source maintained for ATOC as part of the work of Rail Settlement Plan) assumptions for Season tickets.

9.2.5   While all the rail matrices are compatible with LENNON 2014/15, the production of the Regional PLANET matrices was carried out in a more simplified way, and in the case of PLANET South made partial use of earlier work based on LENNON 2007/08. The four rail models are therefore discussed separately, starting with PLD.

PLD rail matrices

9.2.6   The first stage in constructing the PLD Base Matrices distributed the de-annualised weekday LENNON 2014/15 station-to-station data between zones of ultimate origin and destination using the postal sector information reported in the 2004-5 National Rail Travel Survey (NRTS). This was done for all PLD zone pairs (235 * 235). In doing this it was noted that the distance travellers are prepared to go to access a station are related to the length of the journey being undertaken, and hence separate allocations to zones were made for the shorter trips < 20 kms, for the middle length trips of 20 to 40kms and for the longer trips over 40kms.

9.2.7   Then, because of significant changes in ticket types introduced in 2008/09, the purpose splits were controlled to the NRTS data at a regional level, without making use of ticket types. On the basis of careful testing, it was concluded that this was a better method than using the most recent PDFH5 recommendations on ticket type to purpose conversion. Note that the definitions of purpose are compatible with those shown below in Table 9-1: Trip Purposes assumed for NTS Car Matrix Analysis for highway trips: the PDFH5 definitions are slightly different in terms of the treatment of longer distance commuting and education trips.

---

[75] As shown in chapter 12, allowance for these "missing" trips is made within the appraisal, by means of the "annualisation factors".

9.2.8   A problem is that while NRTS remains the best source for converting station-to-station data to zone-to-zone flows, the ticket types changed significantly in 2008/09. Although it is possible to operate at a more aggregate ticket type level, as recommended in PDFH5, doing this produced significant changes in purpose composition in 2010/11 compared with NRTS, especially at a region-to-region level. By contrast, analysis of the National Passenger Survey (NPS) data suggested that the profile of travellers by purpose had remained fairly stable over the period 2004-11. In addition, the PDFH5 recommendations linking purpose to ticket type are global: the approach adopted for PFM makes better allowance for regional variation.

9.2.9   The NRTS was also used to segment passengers between car available and non-car available, but incorporating an allowance for the growth in car ownership between 2004/05 and the base year of 2014/15.

9.2.10   The outcome is a set of matrices which in aggregate agree with the LENNON data, but which have incorporated additional data and analysis to produce a best estimate of zone-to-zone rail demand by purpose and car availability.

9.2.11   Primarily to inform the access possibilities in the SCM, car available matrices are split between outward movements (CAF – 'Car Available From', where a car is available at the origin) and return movements (CAT – 'Car Available To', where a car is available at the destination).

9.2.12   Apart from intra-zonal movements, the PLD matrices have been constructed for all possible movements, regardless of journey length. However, by means of the "control matrices" (section 7.2), many of the cells are masked out in the model application.

Regional PLANETs

9.2.13   For the Regional PLANET v4.3 matrices, the starting procedures of de-annualisation and allocation to zones were done on the basis of LENNON 2007/08. The matrices were then subsequently updated to comply with LENNON 2010/11. While this approach was retained for PS, updating the previous matrices to comply with LENNON 2014/15, a different approach was taken for PM and PN, more in line with that for PLD.

9.2.14   Because the zoning is more detailed in the regional models, many station-station pairings in LENNON are absent in NRTS. This led to the adoption of a two-stage process whereby if the station-station flow featured in NRTS, the distribution to ultimate access/egress zones was applied directly according to the postal sectors where passengers set out / ended up. The allocation between CA and NCA was based directly on the NRTS proportions.

9.2.15   If the station-station flow was absent from NRTS, considering CA and NCA flows separately, then access and egress were looked at separately; distribution between access zones was allocated according to all flows from the origin station, and similarly between egress zones according to all flows from the destination station. As the NCA incidence in NRTS is relatively small, this approach was more likely to be used for NCA than for CA flows.

9.2.16   Note that in all cases the NRTS data were organised on a P/A basis, recognising (a) that car availability only applies at the home end of home-based trips, and (b) that as the Regional PLANETs are AM peak models, all travel is assumed to be Production to Attraction. For non-home based trips, car availability is still applied on the basis of household car ownership.

9.2.17   In the case of PM and PN, the process started with the 2014/15 LENNON data, also using the NRTS data to:

- (a) Divide travel to/from joint stations (e.g. 'Birmingham stations') between individual stations (e.g. New Street, Snow Hill, Moor St;) and

- (b) Distribute tickets with PTE zonal destinations between individual destination stations.

9.2.18   The PLANET North and PLANET Midland model values have been informed by re-calculated Journey Purpose and De-annualisation factors, based on NRTS data focused on the masked matrix areas (See section 7). The procedure is as follows:

- LENNON Data conditioning, in particular dealing with PTE stations and the use of travelcards etc.

- Allocation of demand to Regional Model zones (in line with the PLD work)

- Journey Purpose Allocation

- De-annualisation.

9.2.19   In terms of the zonal allocation, it should be noted that in PM and PN the zones are built up from MSOAs, allowing the use of smaller zones in areas of focus (e.g. in metropolitan areas of Manchester and Birmingham). This could have resulted in NRTS origins/destinations being extended over a greater number of zones than our analysis could handle. The proportion of demand with origins/destinations beyond the top 15 zones was reallocated proportionally across the top 15 zones within each market segment at each station.

9.2.20   Stations which have opened since the date of the NRTS have been considered individually and the majority have had all their demand allocated to the zone in which the station lies. This is appropriate as in most cases the new stations are stations created to serve local markets outside the 'masked' regional matrix area or have low demand.

9.2.21   However, for four new stations of more significance [Buckshaw Parkway and Liverpool South Parkway (for PN) and Coleshill Parkway and East Midlands Parkway (for PM)] the demand was spread over wider catchment areas, based on local geography, the presence of nearby existing stations, and zonal population and employment. For these stations NCA/CA splits have been informed by data for adjacent stations represented in NRTS. The approach was further modified for East Midlands Parkway station to reflect its characteristics as both an airport-access and a 'parkway' station with good trunk road accessibility.

9.2.22   The journey-purpose split by ticket type was informed by the NRTS flows within the masked area of the regional matrices. While there had been proposals to re-classify educational trips as leisure trips (as has been done for PLD), on inspection of the NRTS data it was clear that such trips, at the short distances dominating in the regional matrices, remain day trips to and from schools or colleges and as such they were retained within 'commuting'.

9.2.23   The data was also filtered to include only records completing their station-station journey between 0700-0959 and is thus better aligned towards the am peak period forming the scope of the models.

9.2.24   The de-annualisation task first breaks the total down to a standard weekday, and then determines the portion expected to travel in the PLANET morning peak. Significant components include:

- Identifying that a significant proportion of purchased season ticket / season Travelcard journeys (15%) are not made in the morning peak and so fall outside the models.

- Following the previous (v4.3) approach of including an allowance for ticketless travel, by uplifting the volume of travellers nominally on 'Single' tickets while respecting the overall proportions of ticketless travel of 6% (PM) and 9% (PN) from publicly-available regional survey evidence. This leads to a significant uplift of the modelled 'single' ticket type.

9.2.25   The application of these changes suggested a reduction in the previously-estimated volumes of passengers in the morning peak period of overall around 8%, driven to a significant extent by the reduction in regularity of season-ticket use in the peak, noted above, and a significant increase in the number of am peak journeys (modelled as if using single tickets) associated with the increase in allowance for ticketless travel. Overall, this suggests a higher proportion of users on ordinary tickets in PN, and fewer on season tickets (whether point-to-point or PTE zonal) than in PM. This may reflect a wider opportunity to make use of day or reduced tickets, and a greater prevalence of open stations and paytrain operations in the PN area than in the PM area.

9.2.26   For PS, LENNON data does not capture Travelcard and Oyster Pay As You Go trips and hence it has inherent weaknesses within the London Travelcard area. The factors to update the 2007/8 PS matrices to 2010/11 matrices and subsequently to 2014/15 were derived and applied differently for the three identified flow categories:

- Southeast to Southeast (i.e. not London Travelcard area) flows

- Southeast to/from London Travelcard area flows

- Within London Travelcard area flows

9.2.27   For the first group, growth factors could be calculated directly (allowing for the fact, as noted, that not all station pairs may be represented in the data). For flows between the South East and the London Travelcard Area, the growth factors were calculated for the station in the South East only. While this method provides no indication of differential patterns of growth at the attractor end of the trips, it was considered that over the three years between 2007/8 and 2010/11 employment growth patterns between different parts of Central London had not changed significantly and so would not impact on the distribution of trips in the London area. This assumption has been maintained in applying growth to 2014/15.

9.2.28   For flows within the London Travelcard area, data provided from Office of Rail Regulation [ORR] was used for national rail stations in London together with data with TfL's Rolling Origin Destination Survey [RODS] data for London Underground stations: in both cases both 2008 and 2011 data was used. The RODS data consisted of the number of boarders and alighters at each of the stations across the LUL network. Each LUL station was allocated to a PS zone, using GIS mapping software. Separate growth factors were calculated for boarders and alighters. The ORR station usage data was used in a similar way to provide the growth factors for any PS zones with no LUL station.

9.2.29   The uplifts between 2010/11 and 2014/5, by borough, were combined (by 'furnessing') to derive an output uplift matrix between London Boroughs, incorporating the following key features:

- Leaving Docklands Light Railway (DLR) and Croydon Tramlink out of the in-scope demand (as well as other public transport demand such as buses and taxis) focused the calculation on longer-distance movements relevant to the National Rail services which are the target for PS modelling.

- Station ridership was obtained for national rail using the annual total entrances and exits published by ORR for the year, and for LU using the RODS 'All day' weekday access and egress figures.

- Interchange stations between national rail and London Underground could potentially have received increased weighting due to making use of throughputs for both modes. To avoid this an adjustment was made by subtracting the component of the Underground demand identified in the RODs dataset as having accessed the Underground station by 'NR/DLR/Tram' (with further allowance made for the likely DLR or tram components at the few stations where this was also relevant).

9.2.30   Overall, the models show a reasonable level of validation both for strategic movements represented in PLD and the more local movements into key centres represented in the Planet Regional models. The following observations, focus on the 'regional model' findings.

PLANET South – London screenlines

9.2.31   In Planet South validation has been undertaken across a screenline into London, representing the peak flow direction in the AM peak, and covering all of the north facing stations which is appropriate to understand would be for HS2. Both the overall screenline and each individual TOC meet the WebTAG validation criteria, indicating a good level of validation. The findings indicate an improved level of validation compared with PFMv4.3, with all corridors passing the criteria.

PLANET Midland

9.2.32   The Planet Midland validation was undertaken on screenlines around Birmingham, Leicester and Nottingham in the AM peak. The Birmingham cordon does not meet the validation criteria, however the key corridors of Coventry and Wolverhampton do validate. Both Leicester and Nottingham screenlines met the WebTAG criteria of 15%.

PLANET North

9.2.33   The Planet North validation has been undertaken for cordons around the key centres of Manchester, Leeds and Sheffield for flows inbound to these centres in the AM peak period, reflecting the peak flow direction. The validation at all three stations is good and the results meet the validation criteria set out by WebTAG.

9.3   Base year highway matrices

9.3.1   The spacing of stations under HS2 proposals is generally such that it is not expected that highway trips below 50 miles will divert to HS2, hence the matrix scope is generally restricted to trips greater than 50 miles. The only exception is the inclusion of the following district pairs which are considered as possibly attractive for HS2:

- Nottingham – Sheffield;
- Derby – Sheffield;
- Leeds – Sheffield;
- Liverpool – Manchester;
- Preston – Manchester;
- Birmingham – Nottingham;
- Sheffield – York;
- Sheffield – Manchester;
- Leeds – York; and
- Glasgow – Edinburgh (for Scotland).

9.3.2   There are significant difficulties in establishing the pattern of long-distance (LD) car travel, given that only about 2.5% of car journeys are 50 miles or more[76]. In earlier versions of PFM, use was made of data from various regional models, but for PFMv4 a fresh approach was made.

---

[76] according to data from the NTS

9.3.3   The primary source of representative data relating to LD car travel is the National Travel Survey (NTS), and it was decided to make the maximum possible use of this data source. However, even after combining data for a number of years (2002-2010), the sample sizes are only sufficient to give reliable information for region-to-region movements for all purposes combined. While this provides an essential foundation to the matrices, it is therefore supplemented by other data in order to produce matrices by purpose at the PLD 235 zone level.

9.3.4   The NTS contains two sources of LD travel – diary data and "recall" data. The diary data is collected as part of the overall week's travel diary completed in respect of trips of all length by all respondents. In addition, "recall" data is collected, in which respondents are asked to recall any trips of over 50 miles they had made in the previous week[77].

9.3.5   There are some known biases in the "recall" data, and a weighting procedure has been designed by NTS statisticians to correct for this[78], though some deficiencies remain, particularly in terms of the trip length distribution. Subject to sample size limitations, the matrix development has tended to rely more on the diary data.

9.3.6   The NTS data contain 23 trip purposes for each trip origin and destination, which have been combined into five purposes: Home, Work, Employer Business, Education, and Other. While "Escort education" has been treated as education, the remaining escort trips are classified as "Other".

9.3.7   In line with the National Trip End Model (NTEM) definition the OD trip purpose for home-based trips is the non-home end purpose while for non-home-based trips the main purpose is always the destination purpose.

9.3.8   On this basis, table 9-1 shows the 6 OD trip purposes, namely: home-based work (HBW), home-based employer business (HBEB), home-based education (HBEd), home-based other (HBO), non-home-based employer business (NHBEB) and non-home-based other (NHBO). The 6 OD purposes were then further aggregated into four purposes: HBW, HBEB, HBO, and NHB, with HBEd becoming part of HBO and NHBEB and NHBO combined into a single NHB purpose.

Table 9-1: Trip Purposes assumed for NTS Car Matrix Analysis

		Destination				
	Purpose	**Home**	**Work**	**Employer Business**	**Education**	**Other**
	Home	-	HBW	HBEB	HBEd	HBO
	Work	HBW	-	NHBEB	NHBO	NHBO
Origin	**Employer Business**	HBEB	NHBO	NHBEB	NHBO	NHBO
	Education	HBEd	NHBO	NHBEB	NHBO	NHBO
	Other	HBO	NHBO	NHBEB	NHBO	NHBO

---

[77] The recall period was three weeks for the 2002-2005 NTS surveys; this was reduced to one week for the 2006-2010 surveys due to concerns about recall error over a three-week period.
[78] See Section 5.4.3 of National Travel Survey 2011 Technical Report (Dec 2012).

9.3.9   At the whole GB level, the purpose splits (average weekday daily car person demands, diary only LDJ trips) for HBEB, HBW, HBO and NHB are 18%, 14%, 50% and 18%, respectively.

9.3.10   The sample was expanded in line with the NTS sampling rates to give aggregate control totals. As far as possible (with due consideration to sample size), matrices were developed at the 11 region level.

9.3.11   To add detail at the zonal level, use was made of the Department for Transport's Long Distance Model[79] [LDM]. On the highway side, synthetic matrices of car demand had been constructed, and these had been assigned and compared with a substantial number of traffic counts. Details are given in Chapter 4 of the LDM Final Report[80]. Corresponding adjustments had been made to improve the validation, in line with WebTAG recommended procedures, and the resulting matrices were therefore used as a reliable base for development of zonal PLD matrices, provided they were compatible with the more aggregate NTS data.

9.3.12   Since the matrices were only available for the AM (0700-1000) and Inter Peak (1000-1600) periods, these matrices were adjusted to create a set of 24 hour O-D matrices for person car trips, by conforming with NTS data as far as possible. The major adjustments related to the purpose split and the trip length distribution. At a later stage, the relevant shorter distance movements (see paragraph 9.3.1) were sourced from appropriate regional Traffic Models[81] and included in the matrices.

9.3.13   At the final stage of the matrix development, the NHB matrix was split into NHBW, NHBEB and NHBO since PFM does not distinguish between home-based and NHB movement. This was done by using NTS factors derived at the 11 region level (origin only).

9.3.14   For reasons of forecasting demand growth (see Section 11.4), the matrices were also made available on a P/A basis, based on a correspondence established between the two matrix formats.

9.3.15   The objective of the update to the PLD highway matrices was to:

- Update the existing long distance car passenger matrices with the best and most reliable information available;

- Rebase the matrices to a base year of 2014/15

9.3.16   The updates to the highway matrices focussed on refreshing the existing trips in the matrix to 2014/15 without the infill process. The v4.3 matrices, constructed for 2010/11, were updated for 2014/15 using TEMPro growth factors.

[79] Rohr, C., J. Fox, A. Daly, B. Patruni, S. Patil, F. Tsang (2010) Modelling Long-Distance Travel in the UK, European Transport Conference, Glasgow.

[80] Modelling Longer Distance Demand for Travel, Phase 3 Final Report - Volume 1 Main Report, Contract Ref: PPCA 09060, February 2011.

[81] The models used were M1 J28-31 (Atkins), SEMMMS (MVA - Greater Manchester), M6 J11-19 (URS), LATIS (MVA - Scotland), PRISM (Mott MacDonald - West Midlands), LTM (AECOM - Leeds) and SYSTM+ (AECOM - South Yorkshire).

9.4 Base year air matrices

9.4.1   Since the PLD Demand model is not run in the base year, the base year matrices for air are not in fact required for PFM, only the levels of future demand. Hence, we merely discuss the general methodology. More details regarding the future year matrices are given in Section 11.5

9.4.2   While there is a separate PFM module dealing with international air travel making transfers at Heathrow, the requirement in terms of PLD is for domestic airport-airport non-transfer movements. Thus, the geographical scope of movements is confined exclusively to mainland Great Britain and therefore excludes:

- movements to/from Northern Ireland, Isle of Man, etc.; and

- interlining trips (i.e. those movements with two or more legs, where the first leg is within Great Britain but the second and any subsequent legs are international; similarly for inbound journeys, only legs wholly within Great Britain are included).

9.4.3   Within the wider context of the DfT aviation model it should be noted that its focus is much more on the international than the domestic market sector .

9.4.4   The main source of data is the CAA passenger survey, which provides information about the ultimate origin and destination of (departing) passengers, together with total passenger data which allows the data to be expanded.

9.4.5   This has been analysed to provide an airport choice model, separately for business and leisure passengers (commuting is not a significant purpose for this mode). It is assumed that all air travellers have car availability[82], so the matrices are not split between car available and non-car available. The matrices are produced on an annual basis, but converted to average weekday by means of the annualisation factors discussed in Chapter 12.

9.4.6   For v4.3 the DfT provided matrices for 2010 using the same methodology as that for the future years. In order to derive 2014 Base year air demand matrices, the new matrices were interpolated between 2010 and 2026.

9.5 Base rail (PLD) network

9.5.1   The PLD rail network represents "strategic" corridors, such as the main trunk lines across Britain (such as the West Coast Main Line, East Coast Main Line, Midland Main Line, Great Western Main Line and cross-country and trans-Pennine routes). The network (which derives from an earlier version known as the 'Planet Strategic Model' developed for the Strategic Rail Authority [SRA] in 2002) does not cover local commuter rail lines, unless they are part of a strategic corridor[83]. Key features of the rail model include:

- strategic route choice for rail trips across mainland Britain;

---

[82] There is little evidence on the car availability of air passengers (the standard CAA questionnaire does not ask about it). A very small proportion (0.02%) of all trips in the NTS (2002-2010) are coded as having main mode air, and of these, 91% are from car owning households. Thus, the assumption is probably slightly conservative (since non-car available travellers are probably more inclined to switch to rail).

[83] For instance, the majority of the London, Birmingham, Manchester or Glasgow rail commuter networks are excluded from the model, but those services sharing tracks with the various strategic routes would be included.

- the all-day service frequency and stopping patterns of trains;

- representations of wait time and interchange time; and

- "shadow services" incorporated to absorb local trips in the demand matrices (see below).

9.5.2   The original network was defined by the following criteria for inclusion of services:

- Existing rail links that link the 30 largest conurbations within England, Scotland and Wales, together with Gatwick, Stansted and Manchester Airports.

- Routes that form an integral part of the principal long-distance lines (including ECML, WCML, MML)

- Passenger rail routes included in the Trans-European Network.

It was developed in order to forecast the impact of rail network improvements on strategic passenger rail movements, defined as passenger rail movements between major traffic generators (including large cities, airports and international rail termini) of over around 100 km in distance.

9.5.3   The rail networks are supplied as:

- Multiple CIF[84] (common interface format) files of train movements (timetable data); and

- Multiple CSV (comma separated variables) files of train formations and capacities.

9.5.4   Typically there is one Train Operating Company (TOC) per CIF file. The CIF files are a comprehensive data source containing rail services scheduled on the national network. Within each rail movement, the route is described in detail in terms of arrival and departures at station stops, times of passing certain timing point locations (TIPLOCs) which may include signals, junctions, freight yards; and the activity occurring at each location, such as picking up or setting down passengers, adding or removing vehicles.

9.5.5   The starting point for the rolling stock capacities is MOIRA2. This uses the Total Seating Capacity and Standard class standing capacity defined as at 2.5 passengers per square metre. In a few cases there were concerns over the MOIRA2 values (for example the ratios of seating and standing capacity), and these were then revised based on appropriate rules of thumb between Seating and standing capacity (e.g. 45% for intercity stock based on advice from DfT and independent measurement of the rolling stock layout on, for example, Pendolinos). Note that where no MOIRA2 capacity exists for a stock (and no similar type of stock is listed), the currently used model figure (from DfT's National Modelling Framework) [NMF]) is retained.

---

[84] CIF = Common Interface Format. The full specification is at: http://www.atoc.org/aboutatoc/rail-settlement-plan/data-feeds/types-of-data.

9.5.6   The key fields for matching in the formation and capacities data are illustrated in table 9-2. The "Total Capacity" minus the number of seats gives a notional figure for the number of standing places available. In practice, it is assumed that this figure relates to a standing density of 2.5 passengers per square metre. Further information was given in section 4.3 on crowding.

Table 9-2: Example of Train Formation CSV data

Field	Example
Unique Identifier (UID)	L12455
Formation	1*TL (12-car)
Seats	659
Total Capacity	1770

9.5.7   Underground Transit Line Data was supplied by TfL in the form of an extraction of all transit lines for the Underground from Railplan[85], and was combined with vehicle type data extracted from Railplan.

9.5.8   The construction of the transit lines for the EMME networks made use of a Perl[86] script to process the CIF and CSV files, together with Lookup lists to convert the Network Rail data to PLANET format. EMME macros were then used to import the transit lines, and interpolate journey times between non-stopping nodes.

9.5.9   The process is run separately for the four PLANET Models PS, PM, PN and PLD. For each model, TOCs wholly outside the model area are disregarded.

9.5.10   Once the "relevant" transit lines with "relevant" stops are established, the transit times between stops are calculated by subtracting the times between each relevant node. These values are then stored for import into PLANET. Lines with identical stopping patterns and boarding and alighting patterns are aggregated and given a combined headway (service interval).

9.5.11   The CSV file with vehicle type, seated capacity and total capacity data is imported and matched by the Perl script with the CIF data.

9.5.12   For the underground data the method adopted to convert into PLANET coding is significantly simpler than that for national rail services. There are two elements used to perform the conversion:

- A CSV file to lookup between the Railplan and PLANET nodes; and
- Perl script to read the Railplan data and output it in terms of the PS node numbers.

9.5.13   In practice the services relate to the 16-hour period 0600-2200.

9.5.14   The PLD Rail Network is shown in figure 9-1:

---

[85] Railplan is Transport for London (TfL)'s standard public transport assignment model
[86] Practical Extraction and Reporting Language. See http://www.perl.org

Figure 9-1: PLD Rail Network
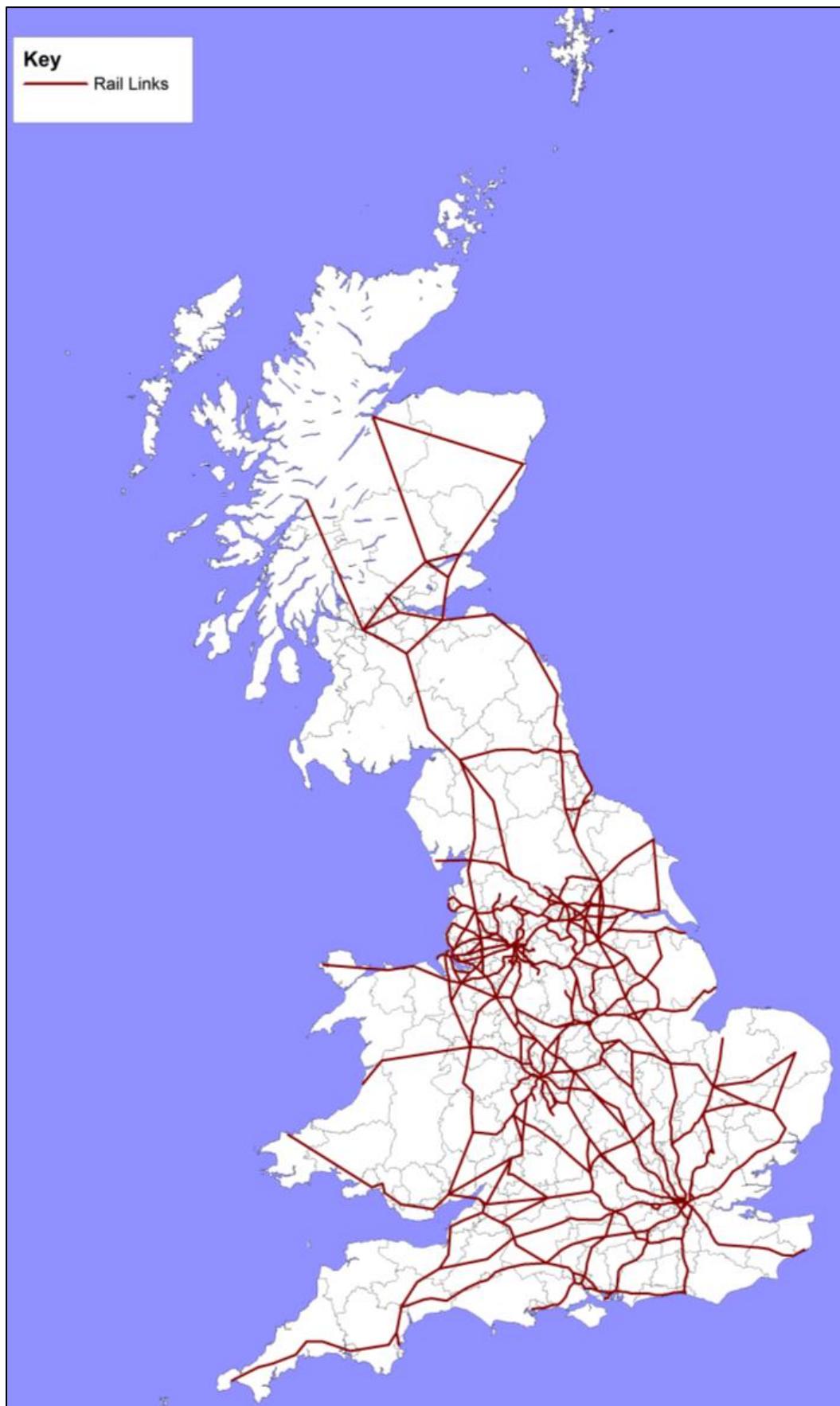
9.6 Base year rail fares

9.6.1 Although fares are not used in the rail network assignment, they are used in the demand model as well as the appraisal module, and it is therefore convenient to describe them here.

9.6.2 While ideally the approaches taken to develop the trip matrices and the fares matrices would be consistent, the approach of defining purpose for the trip matrices without recourse to ticket type means that total trips between a zone pair are split into commuting, business and leisure. If the revenues were processed with the same set of proportions the ratio of revenue to trips would be the same for each purpose – so there would be no fares differentiation by purpose. Much of the differentiation by purpose is expected to be due to the different mix of purposes travelling at weekends and in the peak and off-peak periods and is thus related to ticket type.

9.6.3 Accordingly, it was decided that the revised ticket type to purpose mapping adopted in PDFHv5 should be used to derive the fares matrices.

9.6.4 Section C0 of Part C of PDFH 5.0 (version dated March 2011) was used to obtain the ticket type to purpose conversions for the geographic areas covered by PLD – i.e. not for travel within London or the South East.

9.6.5 For each market segment two tables are presented in C0: an unadjusted set of factors based on NRTS data for weekdays and an adjusted set of factors for an average day (or week) with an adjustment for long distance commuting (including travel to university accommodation) reclassified as leisure travel.

9.6.6 For PLD the factors would ideally be based on weekdays and have all education related travel reallocated to leisure trips. Since this was not possible the intermediate position of using the weekday 'unadjusted' factors with a different adjustment to deal with the long distance commuting/university education travel reallocated to leisure was adopted – i.e. a third table in between the two tables presented in C0.

9.6.7 Four adjustments were made to the following "unadjusted" tables using the approach set out for each table in PDFH 5.0. The assumptions used in the approach are:

- Table C0.9 (Outside South-East to/from London, 100+ miles) – commuting demand on off-peak tickets reduced by 75%;

- Table C0.13 (Outside South-East 20-100 miles) – commuting demand on off-peak tickets reduced by 75%; and

- Table C0.15 (Outside South-East, 100+ miles) – commuting demand on off-peak tickets reduced by 75%, leisure demand on season tickets reduced by 75%.

9.6.8 Having made these adjustments to the tables the figures were normalised to bring the sum over purposes and ticket types back to 100%. The splits into commuting, business and leisure for each ticket type (full, reduced and season) were then derived. The resulting figures are shown in table 9-3 below.

Table 9-3 : Ticket Type to purpose factors based on PDFH5.0 (March 2011)

	PDFH Source Table and Ticket Type					
	To and From London			Outside South East		
	Co.3 Rest of SE to/From London			Co.11 Outside SE <20 miles (excl within PTE areas)		
Purpose	Full	Reduced	Season	Full	Reduced	Season
Commute	44.86%	33.44%	92.76%	66.79%	48.63%	91.85%
Business	34.58%	23.08%	4.21%	7.50%	7.98%	2.82%
Leisure	20.56%	43.48%	3.03%	25.71%	43.39%	5.33%
	100%	100%	100%	100%	100%	100%
	Co.7 Outside SE to/From London <100 Miles			Co.13 Outside SE 20 -100 miles		
Purpose	Full	Reduced	Season	Full	Reduced	Season
Commute	40.17%	32.24%	92.59%	43.93%	11.89%	92.34%
Business	36.75%	24.01%	4.31%	22.59%	19.46	3.83%
Leisure	23.08%	43.75%	3.10%	33.47%	5.00%	3.83%
	100%	100%	100%	100%	100%	100%
	Co.9 Outside SE to/From London 100+ Miles			Co.15 Outside SE 100+ miles		
Purpose	Full	Reduced	Season	Full	Reduced	Season
Commute	6.09%	3.34%	73.53%	8.70%	3.76%	19.05%
Business	76.73%	40.84%	16.67%	44.35%	22.64%	38.10%
Leisure	17.17%	55.82%	9.80%	46.96%	73.59%	42.86%
	100%	100%	100%	100%	100%	100%

9.6.9 The straight line distances between station pairs were calculated using Pythagoras to derive the distance bandings shown in table 9-3 above. Trips to/from PLD zones 117 to 123 inclusive were defined as trips to/from London, the rest as Non London.

9.6.10   The appropriate purpose splits shown in table 9-3 taking into account combinations of origin/destination and trip length, were applied to both the numbers of journeys and the total revenue for each ticket type (full/reduced/season) and zone pair for the CA trips[87]. The results were aggregated to give the total revenue and total trips by purpose for each zone pair.

9.6.11   For each station-to-station movement, the fares for a single journey are calculated by dividing the total LENNON revenue by the total number of LENNON trips. In cases where there are insufficient observations or the implied fare is significantly different from the average, a series of checks and adjustments are made to avoid extreme fares, along the following lines.

9.6.12   The average yield per kilometre (total revenue/total trips/distance) is calculated for each zone pair and trip purpose. These are used to give an average and acceptable range (min/max) for the yield per kilometre for each purpose as follows:

- Average: median yield per kilometre for purpose across all zone pairs for purpose; and

- Minimum/maximum yield per kilometre: median ± standard deviation of yields per kilometre across all zone pairs for purpose.

9.6.13   A series of thresholds were also defined:

- Small flows <0.05 trips per weekday for zone pair (summed across all trip purposes);

- Large flows >50 trips per weekday for zone pair (summed across all trip purposes); and

- Minimum fare: £2.00

9.6.14   The fares were then calculated as shown below:

- For zone pairs with large flows (>50 trips/weekday) and intra-regional flows (wholly within GORs): Fare = Maximum [Total revenue/total trips, £2.00].

- For all other zone pairs (inter-regional with flows ≤ 50 trips per weekday):

  - If average yield per kilometre not within defined range, or volume of trips is small or initial fare is less than minimum (£2.00): Fare = Maximum [average yield per kilometre * distance for OD, £2.00].
  - Otherwise: Fare = Maximum [Total revenue / total trips, £2.00]

---

[87] The processing of the trip matrix is carried out for car owners and non-car owners separately. To use the same processing tools, the car owning stage (being the majority of the trips) is run for revenue as well as trips to derive the fares for all trips whatever their car ownership.

9.7   Base highway network and assignment

9.7.1   The highway network represents the UK strategic road network, with notional access links (centroid connectors) from model zones. Generally, the model includes motorways and the primary route network, with infill in certain areas where the primary route network is sparse. Key features of the highway model include:

- Strategic route choice for car trips across mainland Britain;

- All-day representation of demand, converted to hourly demand to be compatible with hourly speed / flow relationships, based upon COBA[88];

- Strategic demand only, supplemented by local pre-loads; and

- Three trip purposes (user classes), plus a single pre-load for goods vehicles and local car trips

9.7.2   Road types were identified using Ordnance Survey mapping and verified for the Highways Agency trunk road network. Once the road type was identified, the relevant Volume Delay Function [VDF] code was added to the link information. The number of lanes on motorway links were checked using internet available imagery and on other link types using Ordnance Survey mapping.

9.7.3   The PFMv4.3 (2013) highway network was updated using a list of schemes provided by the DfT, based on built, under-construction and committed Road Investment Scheme Period 1 (RIS1) infrastructure programs included in the National Transport Model (NTM). This was provided initially in December 2014, with an updated version following in December 2015. In addition, three additional network changes have been incorporated into the highway network update. These include new link additions on the A1 (Morpeth) and M62 (Liverpool), together with a minor centroid connector road type changes.

9.7.4   38 RIS1 schemes were excluded as the PFM highway network structure did not require any action to be taken. These included junction upgrades and areas that are not modelled in detail in PFM. The proposed updates were submitted to DfT, and the returned final list of approved schemes was coded into a model run.

9.7.5   The total number of two-way links in the 2014 base highway network is 3,674 including airport link roads and zone centroid connectors. Highway network link types are defined as follows:

- Motorway;

- Dual Carriageway;

- Single Carriageway; and

- Other (reserved for centroid connectors and airport links only).

9.7.6   The road types are shown in figure 9-2.

---

[88] The Cost Benefit Analysis program developed by DfT/HA for road schemes.

Figure 9-2: Link Types in Highway Network
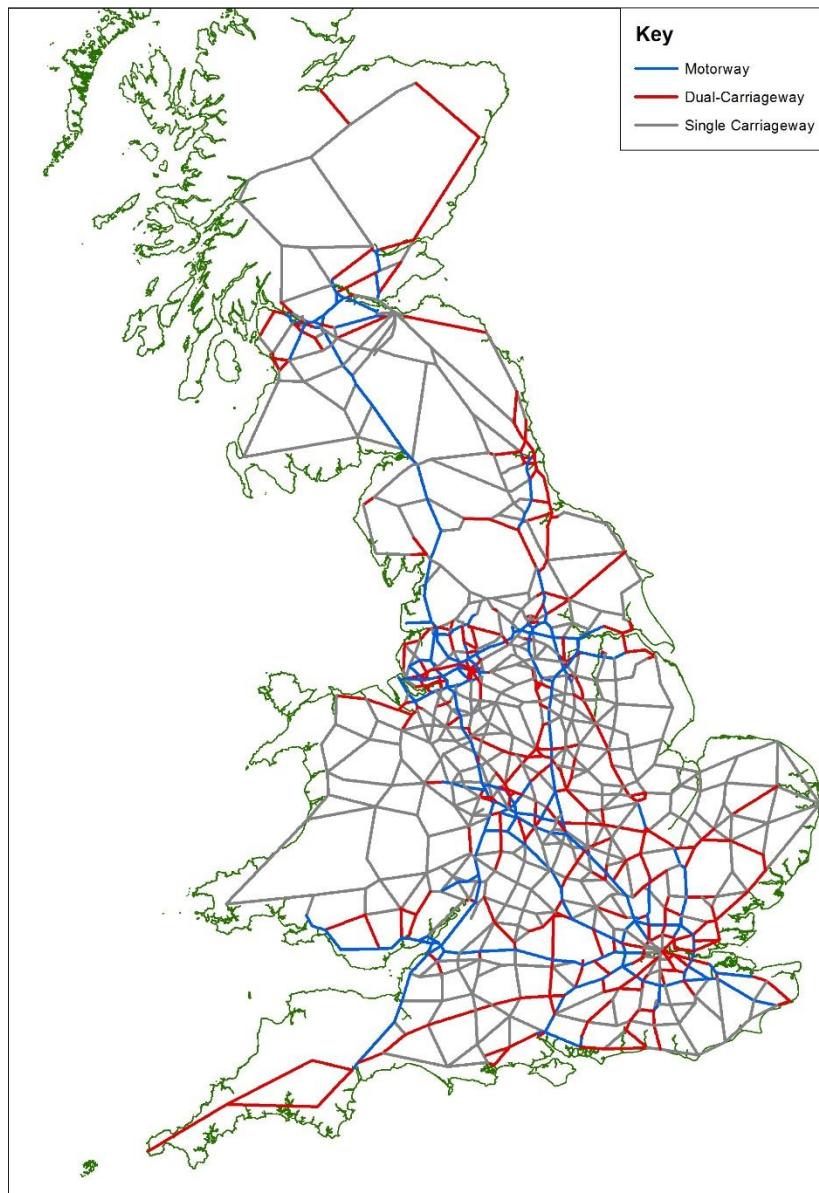
9.7.7   The majority (99.5% of highway links have been assigned with VDF type 11, 14 and 15 for motorway, dual carriageway and single carriageway respectively; with centroid connectors and airport links assigned fixed speeds. The VDFs for these link types match COBA curves, as shown in table 9-4. They allow the speed to be calculated for different levels of flow. According to standard practice for rural roads, three sections are recognised (see WebTAG M3.1): in the first section, the speed of vehicles reduces slightly as flow increases until a critical flow level "break point" (Q_B) is reached, after which we move to the second section in which the rate of speed reduction becomes greater until capacity (Q_C) is reached. The third section (above capacity) employs a formula based on queuing theory. These are compatible with the recommendations in WebTAG M3.1: note that the speeds for flows over capacity are given by the Advice Note 1A relationship.

Table 9-4: VDF Parameters for PLD Highway Network

Flow Level	Motorway (type 11)		Dual carriageway (type 14)		Single carriageway (type 15)	
	Flow/lane /hour (F)	Speed (km/h)	Flow/lane /hour (F)	Speed (km/h)	Flow/lane /hour (F)	Speed (km/h)
less than Q_B	<1200	116-0.006*(F)	<1080	116-0.006*(F)	<1280	72.1 - (0.015+0.00027*5) *(F)
> Q_B, < Q_C	>1200, <2250	105-0.033*(F-1200)	>1080, <2180	101.5 -0.033*(F-1080)	>1280, <1600	51.2-0.05*(F -1280)
> Q_C	>=2250	70.4/(1+70.4*(F-2250)/ (8*length*2250))	>=2180	65.2/(1+65.2*(F -2180)/ (8*length*2180))	>=1600	35.2/(1+35.2*(F -1600) / (8*length*1600))

where:

- Q_B is the flow at which the speed/flow slope of light vehicles changes (veh/hour/lane);
- Q_C is the maximum capacity (veh/hour/lane); and
- Length is the link length in km.

9.7.8   The PLD Highway demand only contains long distance movements, as discussed in Section 9.3. As such, assigned traffic volumes on the highway network should be lower than observed counts and the difference between the two sets of traffic volumes is assumed to be "local traffic".

9.7.9   "Pre-loads" are calculated in the base year and represent the difference between count data and the assigned flows from the PLD demand matrices. These are attached to links to reflect local traffic. These pre-loads include all vehicle types (LGVs, HGVs, PSVs as well as cars).

9.7.10   In v4.3 the highway preloads were mainly derived from TRADS data but as that data set was no longer maintained an alternative source for the data was required. The data set chosen was the DFT's traffic counts, rather than the direct TRADS replacement.

9.7.11   The DFT traffic counts have the benefit of being split by vehicle type, requiring no further processing and full road coverage for the UK. The disbenefit is that they are undertaken using manual counts once every five years,and are therefore less accurate, but for the purpose of the highway model in PLANET this was deemed to be satisfactory.

9.7.12   Count data was obtained from the DFT's major round two way flow count set[89]. The full data set was downloaded and the 2014 values used for the comparison to the modelled flows (the latest year available at the time of the exercise).

9.7.13   In addition an initial base year model run was undertaken using the new matrices with the highway preloads set to zero. This provided flows from a pure un-congested assignment to be compared against the DFT counts.

---

[89] https://data.gov.uk/dataset/gb-road-traffic-counts

9.7.14   The counts and the flows were matched with the PLD links through use of GIS software. All the counts and the base year PLD highway network were plotted in ArcGIS and each count was then assigned to any link within 300m.

9.7.15   This meant that some links were connected to multiple counts and in rare occasions counts were connected to multiple links. However, when rationalised, this resulted in a correspondence list between PFM links and DFT count site IDs.

9.7.16   This correspondence list was used to connect the count data, GIS data and the flow data from the assignment and calculate the preloads. The stages below provide an overview of the process adopted:

- For each link in the model the road name(s) and two-way count data were collated;

- The average flow by vehicle type is then calculated, and combined with GIS outputs to bring together the link flow from the base year assignment and the count data.

- The differences are then calculated as the preload value.

9.7.17   The DfT Transport Statistics data is available in Average Annual Daily Traffic (AADT) form. AADT was converted to Average Annual Weekday Traffic flows [AAWT] by applying a factor of 1.07, derived by comparing the AAWT and AADT flows from the existing TRADS counts in year 2010.

9.7.18   A summary of the conversion factors by road type is given in table 9-5. The counts were then allocated to their appropriate link and then applied equally to both directions. In total, over 900 TRADS counts and a further 1,950 DfT counts were used.

Table 9-5: AADT to AADW Factors in 2010

Road Type	AADT to AAWT factor	Count of Site
Motorway	1.08	257
Dual Carriageway	1.08	153
Single Carriageway	1.07	161
Grand Total	1.07	571

9.7.19   Whereas the longer distance traffic is an output of the demand model, the local traffic is assumed not to be responsive. The traffic volume has an impact on modelled journey times, which are governed by the VDFs described above. If traffic volume is reduced (due to mode shift to HS2 for example), less delay occurs, adding to the HS2 scheme de-congestion benefits. Local road congestion is not included in the model (as the nodes and links and detailed zone structure do not exist): the model is only interested in change in journey times on the strategic leg of the highway journeys.

9.7.20   Of the included strategic roads, junction delays are not modelled due to the relatively low proportion of delay attributable to junctions for long distance trips. Since the mode choice model is incremental, the main function of the highway network is to provide robust strategic journey times.

9.7.21   The model uses a multiclass (business, other and commuting) generalised cost assignment algorithm. It uses the standard EMME highway equilibrium assignment algorithm to achieve convergence, with up to 50 iterations permitted. Convergence is reached at 0.01% best relative gap or 0.01 mins normalised gap[90].

9.8   Base air network and assignment

9.8.1   The air model covers most domestic air services in mainland Britain. Key features of the air model are:

- demand data for domestic UK, non-interlining trips;

- two journey purposes (business and other, not commuting);

- a network that represents the timetabled paths for a particular year supplemented by a transit line timetable file giving the routes and number of flights;

- representation of fare, service frequency, wait and journey time;

- "unconstrained" (i.e. no congestion); and

- car access and egress assumed to and from airports. The choice of airport is partly related to the air service characteristics, such as price and frequency, but also to the access and egress.

9.8.2   Air services are represented on a simple basis, with individual "transit lines" representing flights operating between different UK airports, as shown in Figure 9-3. Congestion is not modelled on air routes, as congestion on the air network is less about aircraft seating capacity than a shortage of runway slots at congested airports. If there is no congestion then the airline response will be to increase frequency or the size of plane. This is not a problem in practice, as air services are more able to respond to demand with pricing mechanisms in the short term, and re-allocation of aircraft and routes in the medium term. There is also no allowance for reliability.

9.8.3   Air passengers have car access at both ends of their air trip. This is consistent with airports outside London, where public transport access is often poor, and the strategic network represents very few actual airport rail links. In addition, air passengers tend to have a higher value of time and are more likely to use taxi if a car is not available for that leg.

---

[90] from the EMME user manual, these terms are defined as follows:

"The best relative gap is an estimate of the difference between the current assignment and a perfect equilibrium assignment, in which all paths used for a given O-D pair would have exactly the same time. The relative gap is the difference between the total travel time on the network and the total travel time on the shortest paths for the current iteration, divided by the total travel time on the network.

"The normalized gap, or trip time difference, is the difference between the mean trip time of the current assignment and the mean minimal trip time. The mean trip time is the average trip time on the paths used in the previous iteration; the mean minimal trip time is the average time computed using the shortest paths of the current iteration. The relative gap decreases strictly from one iteration to the next, whereas the trip time difference does not necessarily have this property. In a perfect equilibrium assignment, both the relative gap and the normalized gap are zero."

9.8.4   Interliners[91] are not covered by the PLD air model, though a separate module (HAM) is included in PFM to deal with access to Heathrow for international flights.

9.8.5   The supply data for air has been provided by DfT Aviation. The base year domestic air fare matrix, from the DfT Aviation Model, provides air fares between all modelled airports in 2008 prices and values. These are adjusted to the 2014 base year using the index of changes in real domestic business and leisure fares supplied by the DfT - see Table 9-6. The fare matrix is based on a distance function which has been developed for each individual airport with domestic flights.

Table 9-6 Real Fare Index Factors (2008 = 1.00)

Year	Business	Leisure
2014	0.968	0.975

9.8.6   The air assignment is a two user class assignment (business and other), with no crowding process, for the reasons given earlier. It should also be borne in mind that the air model exists to provide the PLD mode choice model with the generalised costs of the air product, as a competitor to rail.

9.8.7   Because the assignment operates on a zone-to-zone basis, rather than airport to airport, airport choice is implicitly accounted for, using the mechanism of the 'attractive sets' and frequency allocation (see Section 4.2 for a more detailed discussion in the rail context). In most cases it may be expected that the nearest airports will be used for any given movement between PLD zones (subject to a service being available).

---

[91] These are international air passengers using domestic air services for part of their journey.

Figure 9-3: Air Network
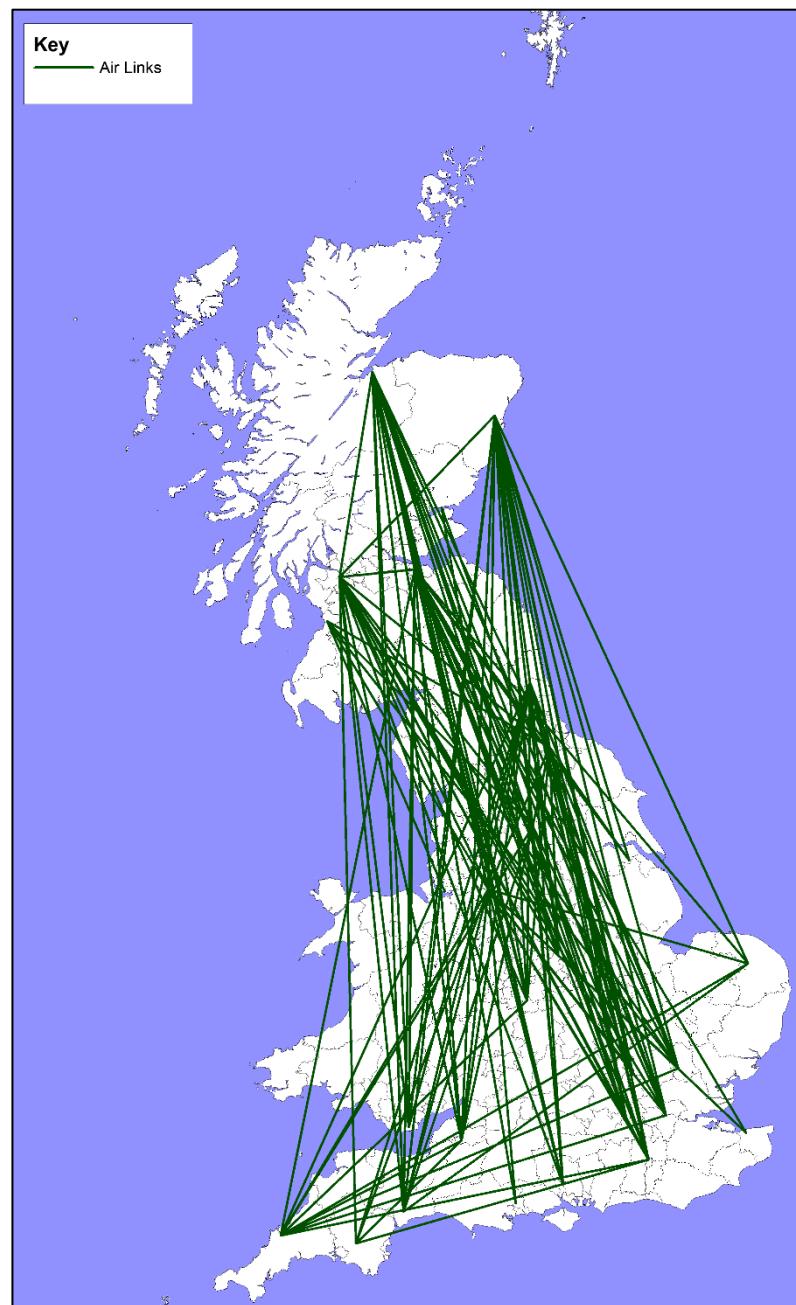
10 Forecasting the "do minimum"

10.1 Introduction

10.1.1 To construct the do-minimum [DM] forecast for a given future year Y, PFM requires forecasts of exogenous growth from external sources. These are obtained for each mode separately, but using a consistent set of "drivers", in particular common forecasts of population and GDP.

10.1.2 In the case of car and rail, growth factors are obtained which can be applied to the base year matrices described in chapter 9. For the air mode, use is made directly of the DfT aviation model, described in section 10.5. The resulting multi-modal matrices of demand T_Y are used as the "pivot" for predicting the impact of the do-something case, as discussed in chapter 11.

10.1.3 Both for the purposes of operating the demand model and for appraisal, the costs associated with the DM are also needed. To obtain these, the forecast T_Y is assigned to the (future) network and the resulting costs (C_Y) are taken as representative of the DM for that year. The principles for constructing the future year networks are set out in section 10.7.

10.1.4 The rest of the chapter describes the detail of the growth forecasts. These are carried out for two years: the opening year (currently set as 2026/27), and the cap year, which is identified using the method described in Section 10.8.

10.2 Economic Assumptions

10.2.1 As is explained in the following three sections, forecasts are made separately for the three modes (rail, car, air), and the required input variables (demand "drivers") vary between them, with the rail forecasts being the most detailed. Common to all three modes is a consistent set of assumptions relating to Socioeconomic drivers:

- Population;
- Employment; and
- GDP per capita.

10.2.2 In addition, because the car ownership forecasts are based on household income growth, there is a need for a consistent set of demographic data relating to Households.

10.2.3 In line with this, both rail and car forecasts are sensitive to Car availability.

10.2.4 There are then the specific drivers related to Intermodal competition, both relating to prices (fuel costs, fares) and level of service, as follows:

Intermodal competition drivers[92]:

- National Rail fares;

---

[92] Note that in the documentation of the EDGE model used for rail forecasting, the terms "Bus cost" and "Air cost" are used, and these are maintained here. However, it is clear from PDFH Section B1.4 that these are in fact Fare variables.

- London Underground fares;

- Car time;

- Fuel cost;

- Bus cost;

- Bus time;

- Bus headway;

- Air cost;

- Air headway;

- Air passengers.

10.3 Forecasting "do minimum" rail demand

10.3.1 The rail demand forecasting process is undertaken using the DfT's Exogenous Demand Growth Estimation Tool (EDGE v1.4.0.393 This is an implementation of the elasticity model recommended in PDFH5. Four separate elements are considered: External Environment, Inter-modal Competition, Fares and Journey Time, Frequency and Interchange. PDFH makes a distinction between the rail-specific components [Fares (F) and Journey Time, Frequency and Interchange (J)] and "external" components (E). The latter include both the socio-economic drivers ("External Environment") and the drivers for other modes ("Inter-modal Competition").

10.3.2 For the external components, the model is:

Equation 10-1

$$
I_E = \left(\frac{GDPpercapita_{new}}{GDPpercapita_{base}} \right)^g \times \left(\frac{POP_{new}}{POP_{base}} \right)^p \times \exp\left(n(NC_{new} - NC_{base}) \right) \times
$$

$$
\left(\frac{FUELCOST_{new}}{FUELCOST_{base}} \right)^f \times \left(\frac{CARTIME_{new}}{CARTIME_{base}} \right)^c \times \left(\frac{BUSCOST_{new}}{BUSCOST_{base}} \right)^b \times \left(\frac{BUSTIME_{new}}{BUSTIME_{base}} \right)^t \times
$$

$$
\left(\frac{BUSHEAD_{new}}{BUSHEAD_{base}} \right)^b \times \left(\frac{AIRCOST_{new}}{AIRCOST_{base}} \right)^a \times \left(\frac{AIRHEAD_{new}}{AIRHEAD_{base}} \right)^r
$$

Where :

I_E is the external factors index for the change in volume between the base and the new periods; and

The parameters are all elasticities, with the exception of n that determines the non car-ownership elasticity

10.3.3 These growth factors are then further multiplied by the growth brought about by changes in rail fares.

---

[93]See http://www.dft.gov.uk/webtag/documents/expert/pdf/u3_15_4-rail-passenge-demand-forecasting-120723.pdf

Equation 10-2

For a uniform change in all fares or a change in the average fare, the formula used is:

$$I_F = \left(\frac{F_{new}}{F_{base}} \right)^{f_t}$$

where:

- I_F is the index for the change in volume due to changes in fare and fare related factors

- F_{new} is the new average fare and F_{base} is the base average fare, and the ratio of the two is the uniform fare change

- Both F_{base} and F_{new} should be expressed in real terms (see Chapter D1)

- f_t is the overall fares elasticity.

10.3.4   Hence the overall growth factor is the product $I_E . I_F$.

10.3.5   The latest guidance from WebTAG (unit M4, Table 1) recommends that PDFH5.0 or 5.1 elasticities are used for all demand drivers, with the exception of National Rail fares, where PDFH4 values should be used. For the Intermodal cross-elasticities, Table B2.7 from PDFH5 is used, where the elasticities are disaggregated by journey purpose and type of journey.

10.3.6   PDFH5 provides elasticities for the remaining 'external' components (GDP/capita, Population and car availability) separately for season tickets and other ticket types, by 6 types of journey:

- Table B1.1   London Travelcard Area [LTA]
- Table B1.2   South East
- Table B1.3   Rest of Country to and from LTA
- Table B1.4   non-London Interurban (over 20 miles)
- Table B1.5   non-London short distance flows (20 miles or less)
- Table B1.6   Airport access

Note that season ticket demand growth is driven by employment growth, not GDP.

10.3.7   The fare elasticities used in EDGE are the recommended values from PDFH4, Section B2.10, for forecasting separately by ticket type (as opposed to forecasting across all ticket types). Elasticities are specifically taken from Tables B2.1 – B2.7. There are separate elasticities for each flow category and in some instances they vary by trip distance. The fares elasticities vary by ticket type (full, reduced, season), with the exception of travel within the London Travelcard Area, where elasticities vary by journey purpose (business, commute, leisure), as recommended in the guidance.

10.3.8   In the development of the EDGE model the GDP elasticity parameter was estimated using GDP with a deflator akin to the RPI, rather than the currently recommended deflator [CPI – see WebTAG M4 para 8.3.3]. The OBR has estimated that the new deflator increases real GDP growth by approximately 0.2% per annum; the real GDP growth forecasts have therefore been reduced by 0.2% per annum to ensure the growth rates are consistent with the elasticities that are applied to them.

10.3.9   EDGE is applied both to the base year PLD demand matrices and to the corresponding Regional PLANET matrices for the two modelled years.

10.3.10 As noted in Section 9.2, although base rail PLD matrices have been produced in P/A format, they have been converted to OD for use in the assignment model. For the car available segments, the P/A-based forecasts from EDGE are applied directly to the base CAF matrices and, in transposed form, to the base CAT matrices. For NCA, the forecasts produced by EDGE have been averaged over the two directions and then applied to the base NCA (OD) matrices to produce the required PLD forecast matrices. Thus, as far as possible, the directional variation in the growth factors is maintained.

10.3.11 In PLANET South the matrices are held separately by direction ("PA" and "AP"), and subsequently summed to form a full OD matrix before assignment. This means that the growth from EDGE can be applied directly (transposing for the AP direction). In PLANET Midlands and North, only OD matrices are available, so an average growth from EDGE over the two directions is used.

10.4   Forecasting "do minimum" highway demand

10.4.1   The forecasting of the PLD DM car matrices relies generally on TEMPRO methodology, as recommended in WebTAG M4 §7.The TEMPRO forecasts assume no changes in highway generalised cost, and provide growth factors by Production and Attraction zones (and also by Origin and Destination).

10.4.2   The starting point was the set of 2014/15 base year matrices described in Section 9.3. Although PLD works with OD matrices, for the explicit purpose of forecasting, the basic three purposes were expanded to the following:

- Home-based work (HBW) daily person P/A matrix;

- Home-based employers' business (HBEB) daily person P/A matrix;

- Home-based other (HBO) daily person P/A matrix;

- Non-home-based work (NHBW) daily person P/A matrix;

- Non-home -based employers' business (NHBEB) daily person P/A matrix;

- Non-home -based other (NHBO) daily person P/A matrix; and

The purposes of these matrices are compatible with the purposes present in TEMPRO.

10.4.3   P/A growth factors for the following purposes were then derived from TEMPRO, where the non-home based (NHB) factors were used for all disaggregated non-home based matrices (NHBW, NHBEB and NHBO):

- Home-based work (HBW);

- Home-based employers' business (HBEB);

- Home-based other (HBO); and

- Non-home-based (NHB).

10.4.4   These growth factors are then applied using the Furness[94] procedure until a stable (P/A) matrix is obtained for each purpose. A further elasticity-based adjustment is made to allow for any difference between the economic forecasts being used for PFM and those implicit in TEMPRO.

10.4.5   Finally, the future year DM matrices are converted back to OD format using the derived P/A to OD factors (Section 9.3). For non-home-based matrices P/A and OD matrices are identical.

10.4.6   In the following sub-sections these various procedures are described in more detail.

Derivation of Furness targets from TEMPRO

10.4.7   Version 6.2 of the TEMPRO dataset was used to derive factors to 'Furness' the 2014/15 daily highway P/A base matrices to the two forecast years. The TEMPRO options chosen were to calculate Trip Ends (on a P/A basis) for an Average Weekday for all areas and sub-areas, for all individual purposes, and for the two modes 'Car driver' and 'Car passenger'. A correspondence list was applied to aggregate the combined car driver and car passenger modes data from TEMPRO unitary authorities to the 25 zone sector system shown in figure 10-1. Finally, the 15 TEMPRO purposes were aggregated to HBE, HBEB, HBO and NHB.

10.4.8   This process was undertaken for both Production and Attraction totals and for the base year and forecast year data separately. The final stage was to divide the forecast year aggregated totals by the base year to produce a set of eight (four purposes by Production and Attraction) Row and Column factors to apply to the base year P/A matrix to produce Furness targets.

---

[94] This refers to the procedure for successively multiplying a matrix by a series of row and column factors until a converged solution is obtained (see for example Ortúzar & Willumsen, Modelling Transport, 2011).

Figure 10-1 : 25 sector System
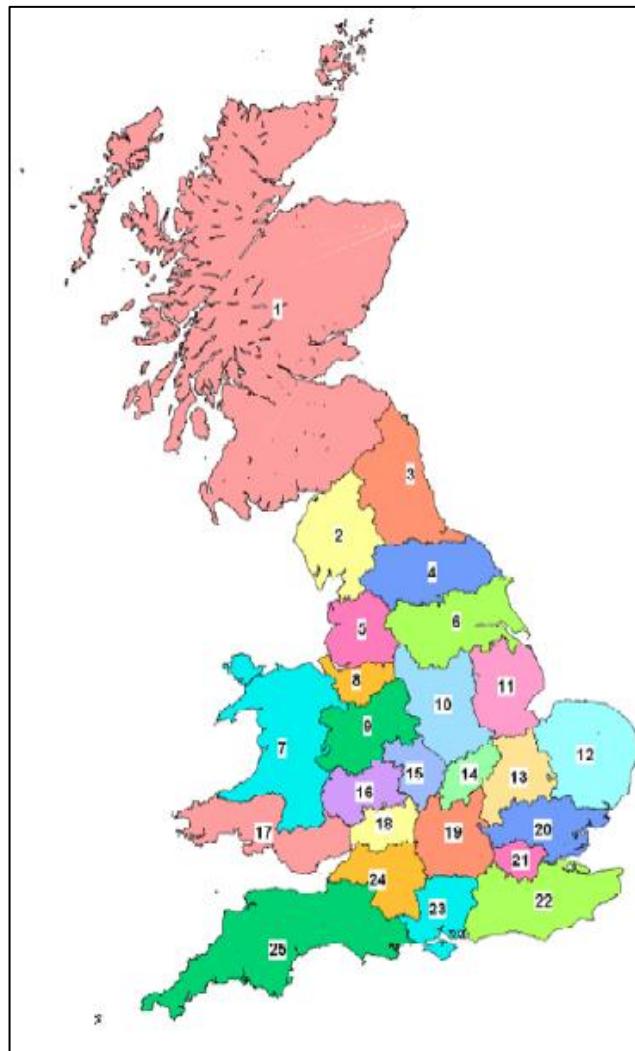
Furness calculations

10.4.9   The 2014/15 daily highway P/A base matrices in the model are factored by the row and column P/A factors, which are summed to row and column totals to produce Furness targets. The Furness calculations are then implemented, scaling to origin totals. This step produces P/A matrices for the six purposes (HBW, HBEB, HBO, NHBW, NHBEB and NHBW).

Application of GDP elasticity

10.4.10   A derived GDP elasticity (global factor) is applied to the output Business, Other and Commuting matrices to correct for the GDP discrepancy between TEMPRO and the latest OBR forecasts that have been used in the forecasting of rail demand.

10.4.11   There are two ways in which a change in GDP over time will affect the forecasts of car traffic: one is through a change in car ownership and therefore a change in the total number of trips by car, and the other is through a change in the value of time (VoT) which will change the value of the money cost component of generalized cost, potentially affecting both the number and length of trips by car.

10.4.12   The elasticity of car trips to GDP was derived from earlier work[95] for HS2 Ltd and the elasticities are shown in table 10-4. These were derived by running alternative sets of GDP forecasts through the DfT's NTEM procedure (both car ownership and trip ends). The work suggests that the elasticity is very low (in the order of 0.125), and lower than the 0.16 vehicle kilometres value specified in WebTAG M4 para 7.4.14 that also includes the VoT effects just noted.

Table 10-1: Implied elasticity of highway demand to GDP derived from NTEM outputs

Purpose	Business	Other	Commute
Implied Elasticity	0.151	0.147	0.087

10.4.13   The elasticities shown in table 10-4 were applied to the relative growth in GDP and the resulting factors were applied globally to the forecast P/A matrices to correct for the change in GDP forecast. The correspondence used to map these purposes to the six matrices was the following:

- HBW = Commute
- HBEB = Business
- HBO = Other
- NHBW = Other
- NHBEB = Business
- NHBO = Other

PA to OD conversion

10.4.14   The final process to create the future daily highway OD matrices uses the P/A to OD factors noted earlier (paragraph 9.3.15). These were input to the process at the 25 sector level, and applied to the home based purposes to convert them to OD format. The required PLD purposes were then obtained by aggregating the 6 OD purposes using the following equations:

- Business = HBEB + NHBEB;
- Other = HBO + NHBO + NHBW; and
- Commuting = HBW.

---

95 PLANET Long Distance and Long Distance Model Comparison, Phase Zero Report, High Speed Two Ltd., March 2012.

Factoring 2014/15 pre-loads to future years

10.4.15   As well as the forecast car matrices it is also necessary to adjust the pre-loads in the Highway Assignment model (Section 9.7). The pre-loads are calculated for the model forecast years using the NTM traffic forecast component of the Road Transport Forecasts 2015 (RTF15)[96]. The key input assumptions to RTF15 are:

- Population and employment data - based on the NTEM dataset which incorporates ONS and OBR projections

- GDP Forecasts based on the OBR short and long run GDP forecasts between 2010 and 2040.

- Fuel Prices taken from the DfT's Fuel Price Forecasting Model, which uses DECC oil price projections, planned VAT and fuel duty and the OBR predicted GDP deflator.

10.4.16   NTM forecasts traffic levels by region and road type, and the flows for the years required were derived using interpolation and extrapolation from Table 4.3 from Road Transport Forecasts 2011, as shown in table 10-2. The link pre-loads were uplifted using the following assumptions:

- As the projections from the National Transport Model have a broad order of magnitude they possess a significant range of uncertainty. As this uncertainty is likely to be greater for more disaggregate results, a single factor was calculated to be applied globally to all regions.

- The values calculated apply to England only; it is assumed that Wales and Scotland have the same growth factors; and

- As the nature of the network modelled is predominantly major roads, the only road types to be considered in the calculation of the growth factors are Motorway, Trunk and Principal.

Table10-2: Road Transport Forecast

Bn Vehicle Miles	Year	Motorway	Trunk	Principal	Other Roads	All Roads
Car	2010	39	24.2	67.8	77.6	208.6
	2035	55.6	33.9	91.6	104.7	285.8
LGV	2010	6.7	4.1	10.9	14.2	35.9
	2035	12.6	7.7	20.4	26.7	67.3
HGV	2010	6	2.8	3.5	1.8	14.1
	2035	8.7	4	4.9	2.5	20.1
Bus and Coach	2010	0.2	0.2	0.9	1.4	2.7
	2035	0.2	0.1	0.8	1.3	2.4
All Traffic	2010	51.9	31.3	83.1	94.9	261.2
	2035	77.1	45.7	117.7	135.1	375.6

---

[96] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/411471/road-traffic-forecasts-2015.pdf

10.5   Forecasting "do minimum" air demand

10.5.1   The approach for base year and forecast year air demand is to adopt the DfT Aviation Model forecasts of supply and demand. This approach ensures a consistent approach to forecasting domestic air passenger demand and aviation supply between the base and forecast years. Forecasts are also required for the Heathrow Airport Model.

10.5.2   The domestic air passenger demand provided by the DfT came from the DfT Aviation Model forecasts January 2013. The data provides future year unconstrained end-to-end, non-transfer demand by trip purpose (employers business and other).

10.5.3   This section includes a brief summary of the DfT Aviation Model, more details of which can be found in the DfT publication of UK Aviation Forecasts, January 2013, before presenting the forecast data.

DfT aviation model

10.5.4   The DfT Aviation Model forecasts the number of passengers passing through UK airports ('terminal passengers') each year. This covers UK and foreign residents travelling to, from or within the UK. The PLD model covers only those trips made exclusively within Great Britain and therefore excludes movements to/from Northern Ireland, Isle of Man etc. and excludes interlining trips (international movements where, for outbound journeys, the first leg of the trip is within Great Britain but the second and any subsequent legs are international). Within this context of the wider aviation model it should be noted the internal domestic market sector accounts for approximately 15% of the passengers in the model, although the proportion varies over time.

10.5.5   The DfT's aviation forecasts are primarily prepared to inform long-term strategic aviation policy rather than provide detailed forecasts at every individual airport. The airport and specific market sector level forecasts, such as those used in PLD, are therefore only generated as an intermediate output of the forecasting approach.

10.5.6   Passenger forecasts are generated for each forecast year in two steps:

- The first step is the 'Unconstrained' national air passenger demand forecasts which are generated using the National Air Passenger Demand Model (NAPDM). This combines time-series econometric models with projections of key driving variables, to forecast national air travel demand assuming no UK airport capacity constraints.

- The second step includes the likely impact of future UK airport capacity constraints, allocation of passengers to airports and translation of passengers into air transport movements is modelled with the National Air Passenger Allocation Model (NAPAM).

10.5.7 The 'unconstrained' demand forecasts from the NAPDM can also be converted to airport-level 'unconstrained' passenger demand forecasts using NAPAM. This is achieved by switching off the airport capacity constraints used in NAPAM, showing how UK air passenger numbers would grow if there were no UK airport capacity constraints. It is these unconstrained forecasts that have been used in the PLD model.

10.5.8 Figure 10-5 provides an overview of the framework used to produce forecasts of UK air passengers.

National unconstrained demand forecasts

10.5.9 NAPDM combines a set of time-series econometric models of past UK air travel demand with projections of key driving variables and assumptions about how the relationship between UK air travel and its key drivers change into the future. The key drivers vary by market sector. In the leisure sector consumer spending and air fares have been identified as the key drivers, whilst in the (domestic) business sectors GDP and fares were shown to be the main drivers. While it is capable of producing forecasts to 2080, only the forecasts up to 2050 have been used for the unconstrained demand forecasts input to NAPAM.

10.5.10 The domestic air passenger demand provided by the DfT came from the DfT Aviation Model's January 2013 forecasts. The data contained future year unconstrained end-to-end, non-transfer demand by trip purpose (employers business and other) and accompanying aviation supply from the DfT Aviation Model. These matrices are in origin to destination (OD) form.

10.5.11 The DfT Aviation Model matrices represent average annual demand. As such, the assumption is that over the course of a year demand should have similar levels of origin and destination trip totals. Any asymmetry found between origins and destinations as a result of the production of exportable matrices from the DfT Aviation model was removed by creating a transpose of the matrix and averaging the two matrices.

Figure 10-5: overview of the framework used to produce forecasts of UK air passengers

* these can be on a constrained or unconstrained basis
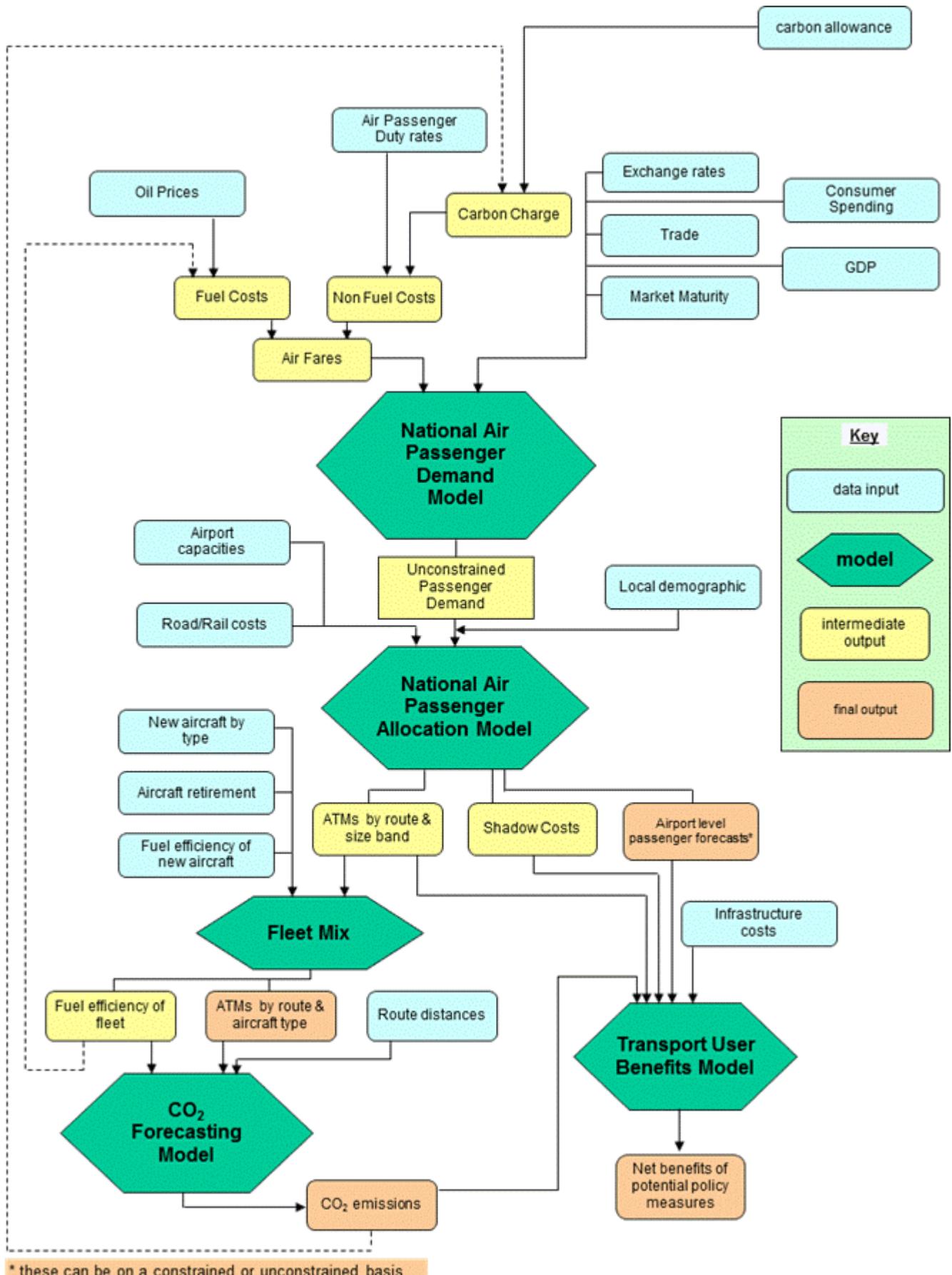
National Air Passenger Allocation Model (NAPAM)

10.5.12   NAPAM comprises several sub-models and routines which are used in combination and iteratively:

- the Passenger Airport Choice Model forecasts how passenger demand will split between UK airports;

- the Air Transport Movement (ATM) Demand Model translates the passenger demand forecasts for each airport into air traffic movements; and

- the Demand Allocation Routine accounts for the likely impact of future UK airport capacity constraints on air transport movements (and thus passengers) at UK airports.

10.5.13   One of the key features of the model is the ability of the ATM Demand Model to project the availability of routes from each modelled airport. The model assumes that, in line with mainstream economic theory, supply will respond to demand as long as the market is commercially viable. The ATM Demand Model simulates the introduction of new routes by testing in each forecast year whether sufficient demand exists to make new routes viable from each airport. The test is two-way, so routes can be both opened and withdrawn. Also, airports are tested jointly for new routes, allowing them to compete with each other. This is one of the reasons why the air supply is updated as the same time as the demand in the PLD model using the adopted DfT forecasts of supply and demand.

10.6   Heathrow Airport model

10.6.1   The HAM requires a forecast matrix representing all people that could switch to HS2 in order to access Heathrow Airport, created by combining surface access and domestic air passenger trips.

10.6.2   For Surface Access, DfT provided forecasts for the number of non-transfer air passengers at Heathrow, segmented by zone and segment . For Domestic Air Access, they provided forecasts for the volume of domestic air travel from UK airports to Heathrow. Only air passengers from Manchester, Edinburgh, Glasgow and Newcastle that transfer to another flight at Heathrow were included.

10.7   Future networks for "do minimum"

Rail

10.7.1   In addition to the pure network changes regarding future services, it is also necessary to make provision for future HS2 stations in the DM network. This is for reasons of appraisal, rather than demand forecasting, and is discussed in more detail in Chapter 13. While it has little impact on the DM demand, the actual procedure is briefly described below.

10.7.2   In general the assumptions made for the future networks are in line with DfT assumptions for timetables and stock. Note that network changes are required for both PLD and the Regional PLANETs. No changes are assumed between 2026/27 and the "cap" year.

New HS2 stations

10.7.3   In connection with the appraisal of the DS, any proposed new stations associated with the scheme need to be included in the DM as well. This is a technical requirement for the appraisal in connection with the SCM, and is discussed in some detail in Chapter 12. These new stations will, of course, not have any direct services associated with them in the DM, but it must nonetheless still be possible to reach the desired destination, so they need to be connected into the DM rail network. In cases where the new station is close to an existing station, this could be done by means of a walk link, but in other cases a transit link will be required. These links should be realistic rather than merely notional.

10.7.4   It is expected that the proportion of passengers allocated to these 'new stations' in the DM will be very small.

10.7.5   As an indication of what is required for a particular version of the HS2 scheme, we note in Table 10-6 how this has been done for the HS2 stations Old Oak Common, Birmingham Interchange, Birmingham Curzon Street, Toton, Meadowhall, Manchester Interchange and Leeds.

Table 10-6 Summary of DM connections to new HS2 stations

HS2 Station	Do-Minimum connection to network
Old Oak Common	Transit Link to Paddington IVT 15 minutes Service Frequency 10 tph TOC PVLK
Birmingham Interchange	Walk Link to Birmingham International, length 1.08 km
Birmingham Curzon St.	Walk Link to Birmingham New St., length 1.2 km Walk Link to Birmingham Moor St, length 0.4 km
Toton	Tram Link to Nottingham IVT 12 minutes Service Frequency 4 tph TOC PVLK
Manchester Interchange	Walk Link to Manchester Airport, length 2.0 km
Meadowhall HS2	Walk link to Meadowhall, length 0.0 km
Leeds HS2	Walk Link to Leeds, length 1.25 km

Highway

10.7.6   Information relating to the proposed enhancements to the highway network between 2014 and 2026 has been provided by the DfT's National Transport Model's list of schemes; which was reviewed against lists on the Highways Agency's Road Projects website (and Welsh and Scottish equivalents) and also included in the National Infrastructure Plan 2011 and subsequent DfT announcements since then.

10.7.7   The update to the future year HS2 PLD highway network follows TAG Unit M4 guidance on scheme uncertainty. The guidance states that an uncertainty log should be created that includes an assessment of the uncertainty of each individual input by placing it into one of the four categories Near Certain, More than Likely, Reasonably Foreseeable, and Hypothetical.

10.7.8   The starting point for the creation of the uncertainty log was the DfT's National Transport Model's list of schemes. The list was reviewed against lists on the Highways Agency's Road Projects website (and Welsh and Scottish equivalents) and also included in the National Infrastructure Plan 2011 and subsequent DfT announcements since then. The schemes were then assigned an uncertainty criterion, which was reviewed by DfT.

10.7.9   TAG Unit M4 guidance states that all the inputs categorised as 'near certain' will be included in the core scenario, it is also expected that those inputs categorised as 'more than likely' will be included. This approach is consistent with that adopted for rail forecasting.

10.7.10   The list of schemes provided and subsequently reviewed by the DfT included schemes marked as open since 2010 and also on site and these were included in the future year PLD highway networks. Following the TAG Unit M4 guidance, only schemes considered as near certain and reasonably foreseeable were included in the future year DM HS2 PLD highway networks.

10.7.11   A number of schemes in the reviewed DfT list were not included. Reasons for excluding schemes were:

- maintenance or structural schemes;
- junction schemes (not applicable in the PLD link only highway network);
- small scale improvements that would affect only a fraction of the modelled link;
- safety schemes; and
- schemes on the fringes of the network

10.7.12   The majority of the schemes included in the model amount to improvements to existing links, so no additional links were required for those schemes. However, the number of lanes and VDF were amended accordingly.

Air

10.7.13   The forecast yearly number of flights, aggregated to twelve regions, for 2026/27 and for the cap year (2037) is taken from the DfT Aviation Model. These will be dependent on the economic growth assumptions in the demand matrices, since the DfT Aviation Model simulates the introduction of new air routes, and the removal of existing air routes, in each forecast year model run. Hence services will be introduced, or removed, based on forecast demand, subject to assumptions on minimum loadings on services.

10.7.14   The 2008 Fares matrices described in Section 9.8 are factored by the relevant year index from table 11-13 to adjust the fare to 2026 and 2037 values.

Table 10-7:Real Fare Index Factors

Year	Business	Leisure
2014	0.962	0.975
2026	0.989	1.160
2037	0.999	1.213

10.8   Treatment of demand cap

10.8.1   Paragraph 2.3.1 of the latest WebTAG Unit A5.3 (December 2015) states that:

"*When forecasting demand, it is important that, to ensure consistency between appraisals, a demand cap is used. In the majority of cases, demand growth should be capped after a 20 year period from the year the appraisal is undertaken (e.g. an appraisal carried out in 2013/14 would be on the basis of demand capped in 2033/34). Sensitivities of 10 and 30 years of growth should be presented. Under exceptional circumstances, such as with long-term infrastructure projects, it may be appropriate to use a different demand cap. In these cases advice should be sought from DfT.*"

10.8.2   It should be noted that there is no implication that demand growth will cease in the assumed year: the practice merely aims at simplifying the modelling task and making some allowance for uncertainty surrounding the pattern of demand a long way into the future. The concept of cap year is essentially an artificial construct and there is no standard methodology for its calculation.

10.8.3   In accordance with the above statement, the PFM now operates over a fixed 20 year period, meaning that the forecast years are now 2026/27, and 2037/38. This is a change from earlier model release versions which calculated the second forecast year based on a perceived limit at which long distance rail trips would cease to grow.

10.9   Future year "do minimum" costs

10.9.1   Although the foregoing assumptions all imply that the DM demand will not be modified for possible supply-side interaction, it is still necessary to obtain an estimate of the costs associated with this demand, both for appraisal purposes (Chapter 13) and to provide the necessary 'pivot' information for the demand model. At this stage, however, the demand model is not invoked – it is only required for the DS forecasts (Chapter 12).

10.9.2   Essentially, what is required is to run the DM demand through the separate modal networks to obtain converged (generalised) costs. For the air network this is straightforward, as no supply-side effects are involved. For the highway network, the forecast DM matrices, together with the associated pre-loads, are loaded on to the future year network, and an equilibrium assignment is carried out.

10.9.3   For rail, the process is more complex, because of the need to take account of a) the SCM and b) the interface with the Regional PLANETs. The HAM demand is also included within the procedure, but – like the PLD demand model itself – is not changed. The operation of the algorithm for producing the DM costs is outlined in the Box below, with particular attention to rail costs.

10.9.4   It should be noted that the current version of the algorithm reflects its historical development. Since the EMME 'macros' which implement it are quite complex, there has been a tendency to make as few changes as possible, and the outcome is that some redundancy remains. For example, since the air assignment is not demand-sensitive, it does not need to be run more than once. However, for historical reasons, 'assignment' within PLD refers to a process which carries out a separate assignment for each of the three modes every time it is called. While this is computationally inefficient, it does not impact on the final results.

10.9.5   Note that even though demand is not changing, three iterations are required to deal with the interface between PLD and the Regional PLANETs. A first estimate of the impact of PLD rail on the Regional PLANETs is produced at the end of Iteration 0, and after running the Regional PLANETs, the corresponding impact on PLD is obtained. However, this will lead to a revision of the PLD crowding and hence the impact on the Regional PLANETs, requiring a further iteration before the system can be considered stable. After this, the DM costs can be generated.

10.9.6   At the end of this process, the DM generalised cost matrices $G^{0\,pc}_{mIJ}$ are produced for each mode m. Note that for the purposes of appraisal, these are also split into separate elements, as described in Chapter 12.

Outline of Algorithm for DM costs

(*Iteration 0*)
Assign Unit PLD rail matrix to generate uncrowded Rail GJTs
Read PLD demand matrices T^{pc}_{IJ} (including HAM), Reg PLANET demand matrices
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs and impacts on Regional PLANETs
Run HAM
Run Regional PLANETs including PLD loads: produce pre-loads for PLD Rail
(*Iteration 1*)
PLD Assignment:
      Highway/Air
      Rail (10 iterations) with pre-loads to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs and impacts on Regional PLANETs
Run HAM
Run Regional PLANETs including PLD loads: produce pre-loads for PLD Rail, DM demand and GJTC
(*Iteration 2*)
PLD Assignment:
      Highway/Air
      Rail (10 iterations) with pre-loads to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air   Skim Costs
      Rail (10 iterations) to update Rail GJTCs
Run SCM to produce PLD DM station-to-station demand and DM GJTCAE*

Output DM costs G^0 (including Rail Fares)

11 Forecasting the impact of HS2

11.1 Introduction

11.1.1 As with the DM, any version of the scheme ("Do Something") needs to be run for two years – the opening year and the cap year. In this chapter, we describe how this is carried out. For illustrative purposes, some scheme details are noted, but it should be kept in mind that the procedure is appropriate for any version of the scheme.

11.1.2 The procedure for running the model for a DS scheme is more complex than in the case of the DM, because of the need to incorporate the demand model which predicts modal shift as well as generated travel. As described in Chapter 6, this is an 'incremental' model which pivots off the DM demand matrices, according to the predicted change in costs relative to the DM costs derived in Section 10.9.

11.1.3 The change in costs is, as would be expected, predominantly seen on the rail side. In fact, the air costs are assumed not to change, while changes in PLD highway demand will have only a small effect on generalised cost. In addition, rail fares are assumed not to change, so that it is only rail service patterns (including times, frequency and seating capacity) that impact on costs, and hence on demand.

11.2 Specifying the Do Something scheme

11.2.1 The do something [DS] scheme involves changing the do minimum network to represent the scheme to be tested. In the case of HS2 this requires not only coding both the phase 1 and phase 2 service patterns, but also the changes in the service specification in the rest of the network to take account of released capacity.

11.2.2 As a result, the coding of the DS rail network is a substantial task, and extends to the Regional PLANETs as well as the PLD network. No changes are made to the air and highway networks.

11.2.3 As noted in the previous chapter, any new stations associated with HS2 need to have been included in the DM network. In the DS network, these now become properly connected to HS2 services. Note that in the case of the phase 1 HS2 scheme, phase 2 stations which are not part of the scheme need to have the same treatment as in the DM.

11.3 Predicting the impact of the scheme

11.3.1 In essence, the changes in the rail network - with the introduction of new services - impact firstly on the station-to-station GJTC values, and this leads to changes in the station choice allocation as well as the zone-to-zone GJTCAE, along the lines of Figure 1.3. This in turn leads to new rail demand (both generated and abstracted from other modes), which is then assigned with consequent "second round" effects on crowding (and, to a limited extent, highway congestion, though not air). To obtain an "equilibrium" result, iteration is necessary, subject to convergence measures discussed in section 11.4.

11.3.2   Additional complexity is introduced by the need to interface with the Regional PLANETs and the HAM. The required outputs are the DS demand and the DS costs, separately by mode and demand segment. The operation of the algorithm for producing this output is outlined in the box below, with particular attention to rail costs. Iteration 0 is carried out to produce a first estimate of the necessary pre-loads: the demand model is not invoked at this stage.

11.3.3   As was noted for the DM version of the algorithm, the current version reflects its historical development, and some redundancy remains. Again, for historical reasons, 'assignment' within PLD refers to a process which carries out a separate assignment for each of the three modes every time it is called. While this is computationally inefficient, it does not impact on the final results.

11.3.4   At the end of this process, the DS generalised cost matrices C^{pc}_{mIJ} and demand matrices T^{pc}_{mIJ} are produced for each mode m (the rail matrices are also produced for the three Regional PLANETs). Note that for the purposes of appraisal, the cost matrices need to be split into separate elements, and the way this is done is described in Chapter 12.

Outline of Algorithm for DS

(Iteration 0)
Assign Unit PLD rail matrix to generate uncrowded Rail GJTs on DS network
Read PLD DM demand matrices $T^{o\,pc}_{IJ}$ (including HAM), Reg PLANET demand matrices
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs and impacts on Regional PLANETs
Run HAM
Run Regional PLANETs including PLD loads: produce pre-loads for PLD Rail
Apply station-to-station proportions to DM PLD rail demand matrices
(Iterations 1&2)
PLD Assignment:
      Highway/Air
      Rail (10 iterations) with pre-loads to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs
Run SCM to produce PLD station-to-station proportions and rail GJTCAE*
Run Demand Model based on cost changes: produce revised PLD demand matrices
Apply station-to-station proportions to revised PLD rail demand matrices
(Iteration 3)
PLD Assignment:
      Highway/Air
      Rail (10 iterations) with pre-loads to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air
      Rail (10 iterations) to update Rail GJTCs and impacts on Regional PLANETs
Run SCM to produce PLD station-to-station proportions and rail GJTCAE*
Run Demand Model based on cost changes: produce revised PLD demand matrices
Apply station-to-station proportions to revised PLD rail demand matrices
Run HAM
Run Regional PLANETs including PLD loads: produce updated pre-loads for PLD Rail
(Iterations 4-6 as Iterations 1-3)
(Iterations 7-12 as Iterations 1,2)
(Iteration 13)
PLD Assignment:
      Highway/Air
      Rail (10 iterations) with pre-loads to update Rail GJTCs
Run SCM to produce PLD station-to-station demand
PLD Assignment:
      Highway/Air   Skim Costs
      Rail (10 iterations) to update Rail GJTCs
Run SCM to produce PLD DS station-to-station demand and DS GJTCAE*
Output DS costs C (including Rail Fares)
Output DS demand T

11.4 Model convergence

Theoretical MSA approaches

11.4.2 To achieve an equilibrium between supply and demand, an averaging process is required. For this purpose, the method of successive averages (MSA) has been used, whereby at each iteration the latest result is combined with the "rolling average", such that oscillation will be reduced. The procedure is in principle the same as that described in paragraph 4.4.2, but here we are dealing with iterations for the 'outer loop' between supply and demand.

11.4.3 The averaging can be applied to either the cost skims input to the demand model, or to the output demand from the demand model. The equations below show an example of each of these two options.

Type 1: Demand Averaging

11.4.4 Averaged demand is given as follows:

Equation 11-1

$$- \quad X_{ijcm}^{n+1} = \omega^n . D\left(C(X_{ijcm}^n)\right) + (1 - \omega^n) . X_{ijcm}^n$$

where:

- X_{ijcm} Current Demand (Averaged);
- $C(X_{ijcm})$ is Cost for Current Demand;
- $D(C(X_{ijcm}))$ is New Demand resulting from Cost for Current Demand; and
- i is Origin, j is Destination, c is User Class, m is Mode, and n is the iteration number.

11.4.5 This implies that each new iteration of demand to be passed to the assignment is a proportion (ω^n) of the current iteration's demand as output from the demand model, added to the complement proportion ($1 - \omega^n$) of the rolling average of the previous iteration's demand.

Type 2: Cost Averaging

11.4.6 Averaged Cost is given as follows:

Equation 11-2

$$- \quad C_{ijcm}^{n+1} = \omega^n . Y\left(D(C_{ijcm}^n)\right) + (1 - \omega^n) . C_{ijcm}^n$$

where i, j, c, m and n are as before, and:

- C_{ijcm} is Current Cost (Averaged)
- $D(C_{ijcm})$ is Demand resulting from Current Cost
- $Y(D(C_{ijcm}))$ is New Cost for Demand resulting from Current Cost

11.4.7   This implies that each new iteration of the assignment costs to be passed to the demand model is a proportion (ω^n) of the current iteration's costs as output from the assignment, added to the complement proportion ($1 - \omega^n$) of the rolling average of the previous iteration's costs.

11.4.8   In both cases ω^n, described as the Step Length at iteration 'n', is calculated as $\omega_n = \frac{1}{n}$ (though a constant weight of ½ was also tested, with only marginally worse results).

11.4.9   Cost Averaging (Type 2) was preferred because it could be implemented easily into the PLD model whilst maintaining the general structure and functionality of the model.

Measurement of demand and supply convergence

11.4.10   WebTAG M2, paragraph 6.3.4 sets out the following recommendation[97] for measuring convergence:

Type 1 Gap: demand averaging

Equation 11-3

$$- \; RelGAP \; (Demand \; Averaging) = \frac{\sum_{ijcm} C(X_{ijcm}).\left|D\big(C(X_{ijcm})\big)-X_{ijcm}\right|}{\sum_{ijcm} C(X_{ijcm}).X_{ijcm}}$$

11.4.11   This represents a cost-weighted demand change as a proportion of the total cost-weighted demand, with demand averaging as the input. Although not specifically stated in WebTAG, this measure is suitable only for situations where demand is being averaged. Hence it was not possible to use in PFM.

Gap type 2: cost averaging

11.4.12   Given that cost averaging is being used, the corresponding Gap measure should be written as below:

Equation 11-4

$$- \; RelGAP \; (Cost \; Averaging) = \frac{\sum_{ijcm} D(C_{ijcm}).\left|Y\big(D(C_{ijcm})\big)-C_{ijcm}\right|}{\sum_{ijcm} D(C_{ijcm}).C_{ijcm}}$$

11.4.13   This represents a demand-weighted cost change as a proportion of the total demand-weighted cost, with cost averaging as the input.

11.4.14   WebTAG suggests 0.1% (i.e. 0.001 or 10-3) is an achievable target. The adjustments made to the model as part of this process allows us to measure this easily, therefore making the PFM model convergence more readily measured and transparent.

---

[97] Note that in relation to the WebTAG recommendation, the time period subscript "t" has been omitted.

Practical implementation of MSA into PLD

11.4.15 The basic approach is to undertake the following operations within the PLD process:

- store the 'rolling average' set of costs prior to assignment and skimming;

- assign and skim all modes; and

- weight the resultant costs from (step 2) with (step 1) using MSA.

11.4.16 This ensures the appropriate averaging of costs before the next input to the mode choice model, thereby ensuring better convergence in line with the (revised) WebTAG criterion.

11.5 Assessment

11.5.1 The various model components discussed in earlier Chapters have all been brought together in a single algorithm which allows for route and station choice, demand effects (mode shifts and generation, as well as HAM), and the impact on the various networks (including the Regional PLANETs). The interactions between the different rail networks, as well as those between demand and supply in general, are all allowed for.

11.5.2 The algorithm has been developed in a heuristic way, rather than using more complex optimisation methods. While it could probably be streamlined, it is not likely that this would have any impact on the results. The convergence monitoring is in line with WebTAG recommendations, and the level of convergence at the end of the fixed set of iterations meets the WebTAG criteria.

Part 5 - Using the model for appraisal

In this part we discuss how the model has been used to produce the necessary inputs for the appraisal of HS2, and describe the calculations that are then carried out.

12 Appraisal

12.1 Introduction

12.1.1 The aim of the appraisal is to compare key quantities with and without the HS2 scheme. Cost-Benefit Analysis (CBA) is the mainstay of project appraisal for the Department for Transport (DfT) and other Government departments and agencies, although other elements also form part of the overall appraisal. An essential component of CBA is discounted cost flow (DCF) analysis, which calculates a 'Present Value' of both Benefits (PVB) and Costs (PVC), using an agreed Discount Rate, in order to obtain a Net Present Value, as well as other indicators such as benefit-cost ratios.

12.1.2 The high level reporting contains the PVB calculations shown in Table 12-1, separately for the Phase 1 (London to Birmingham) and Phase 2 (Birmingham to Manchester and Leeds) schemes:

Table 12-1: High Level Appraisal Reporting

Transport User Benefits (Business)
Transport User Benefits (Other)
Other quantifiable benefits (excl. Carbon)
Loss to Government of Indirect Taxes
Net Transport Benefits (PVB)
Wider Economic Impacts
Net Benefits including WEIs
Revenues
Costs
Net Costs to Government (PVC)
BCR

12.1.3 Additional information on demand levels for HS2, and the source of this demand, is also provided.

12.1.4 In more detail, the information in Table 12-2 is produced.

12.1.5 A further breakdown by rail GJTCAE elements and purpose is also available.

12.1.6 The focus here is on describing the economic benefit and revenue calculations. Guidance on CBA in appraisal can be found in WebTAG units A1.1,A1 3 and A5.3 and PFM follows these.

12.1.7 The following sections concentrate on the main items of benefit and revenue, and describe the input data and calculations.

Table 12-2: Format for Appraisal Table

Benefits (£m)	All Outputs	In 2011 Prices (2011 Discount Base Year)	
Summary of Benefits	**Total**		
Noise	(1)		
Local air quality	(2)		
Accidents (incl. safety)	(3)		
Consumer users	(4)		
Business users and providers	(5)		
Loss of indirect tax	(6)		
HS2 Noise	(7)		
HS2 Carbon Impacts	(8)		
HS1 Link	(9)		
PVB (sum of all benefits)	$\sum$ (1) to (9)		
Split of user benefits	**Total**	**Road**	**Rail**
Consumers user benefits			
- travel time saving			
- Vehicle opcost		-	
- user charges	-		
- during construction & maintenance	-		
Net	= (4)		
Business			
User benefits			
- Travel time			
- Vehicle opcost	-	-	
- user charges	-		
- during construction & maintenance	-		
Net	= (5)		
Private sector provider impact			
- revenue			

12.2   Inputs and outline of calculations

12.2.1   To carry out an appraisal, four model runs are required:

- Opening Year DM;

- Opening Year DS;

- Cap Year DM; and

- Cap Year DS.

12.2.2   For each of these, the appropriately segmented demand matrices T^{pc}_{mIJ} both from PLD (including HAM) and the regional PLANETs are required, though the distinction between CA and NCA (car availability) is not used in the appraisal. We also require the corresponding generalised cost matrices C^{pc}_{mIJ}, further split by elements "k" (see below).

12.2.3   The DfT CBA procedures are standardised according to the rules set out in WebTAG Unit A1.3, the tabular presentation of the "Transport Economic Efficiency" results (TEE table), and the TUBA Manual[98]. This section focuses on the main elements in the TEE table of relevance for modelling – User Benefits [S] and Revenues [R]. For notational simplicity the following mathematical description ignores distinctions of purpose and other possible "segmentations", but in practice they need to be made.

12.2.4   In terms of user benefit there are a number of generalised cost elements that need to be distinguished. In particular the TEE table identifies the following items: travel time, vehicle operating costs, user charges (including fares) and operator revenues: we denote these by k. The contribution of element k to the overall user benefit associated with mode m in year Y is given as:

Equation 12-1

$$S^{Y}_{m,k} = -\tfrac{1}{2}\sum_{IJ}\left(T'^{Y}_{IJm} + T^{Y}_{IJm}\right)\left(C'^{Yk}_{IJm} - C^{Yk}_{IJm}\right)$$

where

- T is demand, C is (generalised) cost in money terms, and the prime (') denotes the "after" (with scheme) case;
- I and J are zones, and
- m is mode.

---

[98] Transport User Benefit Appraisal program. This incorporates the principles outlined in WebTAG A1.3. However, it is not compulsory to use it, and given the complexity of the HS2 appraisal, custom-built appraisal software has been developed.

12.2.5   This calculation, generally referred to as the 'Rule of a Half' [RoH], is a very widely used approximation to the true 'Consumer Surplus'. However, as is noted in the TUBA Manual, the approximation deteriorates when the cost changes become large. In these circumstances, a better approximation – referred to in the TUBA Manual as "Numerical Integration" [NI] is recommended. Because there are some large changes associated with HS2, NI has been used throughout. To avoid notational complexity, the formulae in this section make use of the standard RoH methodology, but the approach to NI is set out in Annex E.

12.2.6   Slightly different formulae apply to the revenue calculations, which only apply to the monetary elements:

Equation 12-2

$$- \quad R_{m,k}^{Y} = \sum_{IJ} \left(T_{IJm}'^{Y} . C_{IJm}'^{Yk} - T_{IJm}^{Y} . C_{IJm}^{Yk} \right)$$

12.2.7   The following generalised cost elements [k] (varying by mode) are required:

Table 12-3: Breakdown of User Benefits for Appraisal Reporting

GENERALISED COST ELEMENTS FOR RAIL	GENERALISED COST ELEMENTS FOR HIGHWAY
Uncrowded Journey Time	Journey Time
Crowded Journey Time	Vehicle Operating Costs (Fuel)
Access/Egress Times and Costs	Vehicle Operating Costs (Non-Fuel)
Wait Time	
Walk Time (for interchange)	
Boarding Penalty	
User Charge (Fares)	

12.2.8   Note that although corresponding cost elements for the Air mode are available, they do not change between the DM and DS cases, and hence - as can be seen from Eq (12-1) - they do not contribute to the benefit calculations.

12.2.9   For the most part, the calculations can be carried out at the zonal level (e.g., for PLD at the "IJ" level). For the rail mode, however, the available cost at this level [GJT(AE)] is a composite cost which originates from the SCM, (see Eqq 5-14 to 5-18 in Chapter 5). Thus, it is not a simple (weighted) summation of the constituent elements. This leads to some difficulties which we discuss in more detail below.

12.2.10   The weightings for the rail elements in the various models were set out in Section 2.6: they do not vary by purpose. Table 12-4 repeats this information and also includes the appraisal values, which have been agreed with DfT as being consistent with WebTAG. The shaded values indicate where the appraisal values are different from the assumptions in PLD (though it should be noted that, for historic reasons, the boarding time penalty assumptions are also different in the Regional PLANETs). Note that for some categories, WebTAG requires different values for the Business purpose.

Table 12-4: Generalised Cost Element weights for Rail

Rail	Model Values (all purposes)			Appraisal Values		
	PLD	PS	PM &PN	Business	Other	Commute
IVT (uncrowded)	1.0	1.0	1.0	1.0	1.0	1.0
IVT (crowded)	1.0	1.0	1.0	1.0	1.0	1.0
Wait Time	2.0	2.0	2.0	1.0	2.5	2.5
Access/Egress Time	2.0	2.0	2.0	1.0	2.0	2.0
Board Time Penalty (mins)	30.0	3.5	20.0	30.0	30.0	30.0

12.2.11   As noted, in PLD the available costs at the zonal level are composite. There is very little in WebTAG which relates to the use of composite costs. The main discussion is in M3.2 section 5.2, relating essentially to PT assignment. Unit A1.3 suggests that the formula can 'be extended to cover network appraisal with many modes and origin/destination pairs.' [para 2.1.8]. There are references to the TUBA manual where para 12.4.5 says:

> "There are some difficulties associated with skimming costs. Theoretically the skimmed costs should be consistent with the choice model used in the assignment. It has been suggested that the use of composite costs would be better than flow-weighted averages. However, there are problems with this approach and the Department is currently considering the issues. In the meantime the recommendation is to skim costs as the passenger-weighted average across submodes and routes."

12.2.12   To adopt a procedure in line with current guidance it is necessary to ensure that the choice sets do not change between DM and DS, and to apply the RoH at the most detailed level in the model. This implies that, for PLD rail, the benefits should be calculated as:

Equation 12-3

$$S_{IJ}{}^{k} \approx -\tfrac{1}{2} \sum_{i \in I,\, j \in J} \left(\sum_{RSx} \left(T'_{IJ} . \pi_{ij}[IJ].p'_{RSx|ij} + T_{IJ} . \pi_{ij}[IJ].p_{RSx|ij} \right) \Delta C_{ijRSx}{}^{k} \right)$$

where, to recap:

- R, S denote PLD stations;
- i and j are mzones, and I and J the PLD zones which contain them;
- x denotes access mode (highway or PT);
- k denotes a generalised cost element; and
- the quantities $p_{RSx|ij}$ represent the proportions of demand choosing station pair RS and access mode x, given mzone pair ij.

Once again, the prime (') denotes the "after" (DS) case. As throughout this Chapter, segments related to purpose and car availability are suppressed in the notation.

12.2.13   Because $\pi_{ij}[IJ]$ does not change between DM and DS, this can be simplified to:

Equation 12-4

$$S_{IJ}{}^{k} \approx -\tfrac{1}{2} \sum_{i\in I,\, j\in J} \pi_{ij}[IJ] \left(\sum_{RSx} \left(T'_{IJ} \cdot p'_{RSx|ij} + T_{IJ} \cdot p_{RSx|ij} \right) \Delta C_{ijRSx}{}^{k} \right)$$

12.2.14   When k relates to access/egress elements, $\Delta C_{ijRSx} = 0$, since the access costs to any given station are not affected by the scheme. In addition, all other rail cost elements are independent of the mzones and access mode, since they are only concerned with the movement between stations R and S. This allows further simplification to:

Equation 12-5

$$S_{IJ}{}^{k} \approx -\tfrac{1}{2} \sum_{RS} \left(T'_{RS|IJ} + T_{RS|IJ} \right) \Delta C_{RS}{}^{k}$$

where

-   $T_{RS|IJ}$ is the total demand for station pair RS deriving from PLD zone to zone pair IJ, and
-   $\Delta C_{RS}{}^{k}$ is the change in cost element k for station pair RS.

This is the calculation which has been implemented for the PLD Rail User Benefit calculations. The results are summed over all IJ combinations to yield the required estimate $S^{Y}_{m,k}$. With the exception of the money elements, all other elements are (at this stage) in time units.

12.3   Annualisation

12.3.1   The demand variables relate, in the case of PLD, to an average weekday, and in the case of the Regional PLANETs, to the AM Peak of an average weekday. To convert them to an annual basis, annualisation factors are required, and the following values are used for PLD:

Table 12-5: PLD Annualisation Factors

	Rail	Air	Highway
Business	255	313	275
Leisure	428	313	361
Commuting	264	n/a	282
Average	316	313	306

12.3.2   These factors were derived by applying NTS-derived journey purpose splits for non-weekday demand to estimates of total non-weekday demand derived from the LENNON deannualisation process used in developing the PLANET matrices. The NTS long-distance data from 2006-2010 was used at a national level giving the following results (Table 12-7):

Table 12-7: Proportion of total weekly rail demand over 50 miles by journey purpose and weekday/non-weekday from NTS data

	Weekday	Weekend
Business	22%	1%
Leisure	35%	20%
Commuting	22%	1%
Total	79%	21%

12.3.3   It is assumed that, on average, there are 245 working week days per year based on 260 calendar weekdays per year, 8 bank holidays and an additional reduction to account for atypical, reduced demand in the Christmas and New Year period, particularly in the week between the two. This factor was used in the approach used to deannualise LENNON data by ticket type and has been confirmed to be appropriate through analysis of available data such as guard counts and MOIRA data.

12.3.4   On this basis, the equation used was:

Equation 12-6

$$A^p = 245 * \frac{T^{wd}.s^{p,wd} + T^{we}.s^{p,we}}{T^{wd}.s^{p,wd}}$$

where

T^{wd} = Total weekday demand for all purposes derived from LENNON deannualisation

T^{we} = Total non-weekday demand for all purposes derived from LENNON deannualisation

$s_{p,wd}$ = Share of weekday demand for journey purpose p (derived from NRTS data as part of PLANET matrix development)

$s^{p,we}$ = Share of non-weekday demand for journey purpose p (derived from NTS)

12.3.5   Using these annualisation factors, the sum of annualised demand will equal the total demand reported in the LENNON database. In carrying out the calculations, the demand from LENNON was restricted to all trips over 50 miles excluding those covered by the regional PLANETs. This is more consistent with the NTS data used in generating the factors, and with the main benefits of HS2 (which applies to long distance flows).

12.3.6   Separate factors are derived for PLANET South/Midlands/North.

Table 12-6: Annualisation factors for Regional Planets

	7AM to 10 AM	10AM to 4PM	4PM to 7PM	Total (incl. Weekend)
	User benefit annualisation			
Business	304	539	365	1,376
Commuting	278	86	260	697
Leisure	303	1,181	602	2,562
	Crowding annualisation factors			
Business	253	0	304	557
Commuting	253	0	237	490
Leisure	253	0	503	756

12.4   Interpolation for non-modelled years

12.4.1   Ideally, the model would be run for every year Y of the benefit stream, but in practice it is not feasible. For this reason assumptions are required about the path of benefits for those years when explicit model runs are not carried out. As we have seen, the PFM is run for two years which we denote as A and B, where A is the opening year and B is the 'cap' year (see Section 12.2).

12.4.2   Running the transport model in years A and B allows the calculation of the quantities S_k^Y for these two years (from now on these terms are assumed to include the annualisation factors). It is clear from the definition of these quantities in Eq (12-1) that they require both demand estimates (T) and (generalised) cost estimates (C), separately for the DM and the DS. Changes in the demand estimates over time are partly exogenous to the transport model (through population changes and income changes, the latter also affecting car ownership), and partly endogenous, due to supply-side effects. Similarly, changes in the cost estimates over time are also partly exogenous (eg fuel prices/efficiency, fares), and partly endogenous (supply-side effects such as congestion and crowding).

12.4.3   Given only a limited number of years with model runs, the approach taken used in the purpose-built appraisal software developed for PFM, which is in line with TUBA, is as follows[99] . Keeping the benefits in their natural units (i.e. time for all elements apart from the money costs), for A < Y < B linear interpolation is carried out between S_k^A and S_k^B to get a value of S_k^Y

Equation 12-7

$$S_k^Y = S_k^A + \frac{Y - A}{B - A}.(S_k^B - S_k^A) \qquad \text{for A < Y < B}$$

12.4.4   After the cap year (B) a zero growth assumption in the volume of benefits is made. Hence the implied formula is:

---

[99] Note that this is not intended as a description of the actual calculations in TUBA, but rather of the underlying principles

Equation 12-8

$$S_k^Y = S_k^B \qquad\qquad\qquad \text{for Y > B}$$

12.4.5   This produces an estimate of annual benefits for each generalised cost element for every year.

12.5   Weighting the Elements

12.5.1   The various time elements now need to be weighted (to take account of the different weights α_k set out in Table 12-4, which vary by purpose) and then converted to money terms by multiplying the weighted time elements by an appropriate Value of Time [VoT]. VoT varies by purpose and by year, in line with WebTAG Unit A1.3. This can be written as:

Equation 12-9

$$B_k^Y = VoT^Y.\alpha_k S_k^Y$$

for all time elements k

12.5.2   To give an estimate of total user benefits in any one year, the elements need to be summed over all elements k. However, for presentational purposes the breakdown by elements is retained as well.

12.6   Calculating the Present Value of Benefits

12.6.1   If the benefits in year Y are written as B^Y, then the present value of benefits (PVB) is given as:

Equation 12-10

$$PVB = \sum_Y \frac{B^Y}{(1+r)^{Y-W}}$$

where r is the discount rate (in line with WebTAG Unit A1.1, para 2.7.5), and W is the base year.

12.6.2   While in theory the calculation could be summed over an infinite number of years into the future, in practice this is limited to a finite 'appraisal period'. This raises the question as to whether any unaccounted benefits at the end of the appraisal period (technically referred to as the 'residual value') need to be taken into account. WebTAG A1.1 also provides guidance on this.

12.6.3   Of course, benefits cannot be generated before the opening year of the scheme. According to WebTAG A1., Y should range over 60 years starting with the scheme opening year. Because the opening years are different for Phase One and Phase Two, the appraisal includes the benefits between the two opening years, as well as the 60 years from the opening of Phase 2.

12.6.4   A three-year "ramp-up" period is included for both phases. Prior to applying the discounting, the benefits associated with each phase are reduced by 20% in the opening year, 10% in the next year, and 5% in the following year.

12.7   Closing remarks

12.7.1   The foregoing discussion relates to the general calculation of the user benefit elements, with special attention to the rail mode. In addition, the appraisal spreadsheet requires the calculation of revenue (separately by Purpose), indirect tax (business/non-business), and highway externalities (congestion, accident, local air quality, noise), as well as benefits from the International Rail Model and Carbon.

12.7.2   Additional Wider Economic Impact Benefits are calculated in line with WebTAG A2. 1. The factors taken into account are: Agglomeration benefits (WB1), Imperfect competition (WB3) and Labour Market Impacts (Exchequer consequences of increased GDP (WB4) due to Increase in labour force participation (GP1)).

12.7.3   All the calculations are done separately for PLD and the three Regional PLANETs, in line with the "control matrices" discussed in Sections 2.2 and 4.5. The highway externality and indirect taxation calculations are in line with the requirements of WebTAG A5.4, based on changes in car-Km. Revenue calculations are straightforward, using Eq (12-2) given earlier. All these elements are linearly interpolated as described in above between the modelled years.

12.7.4   The calculations are carried out partly by means of EMME macros and partly within a purpose-built spreadsheet, which has a pre-processing step written in VB.net and uses outputs from the SCM. Data is imported for the four model runs, separately for "Day1C" [Phase 1] and "Y" [Phase 2a, 2b] schemes, for the two modelled years. Interpolation and Extrapolation is carried out in line with the discussion in Section 13.3, following the principles embodied in TUBA. Generalised cost element weights and values of time all derive from WebTAG.

12.7.5   The outcome is that the output from the various PFM model components is all brought together in a series of appraisal tables which represent best CBA practice following the principles of WebTAG.

13 Annexes

13.1 Annex A: EMME documentation of transit assignment strategies

13.1.1 The standard transit assignment algorithm implemented in EMME is a multi-path algorithm based on the concept of strategies and optimal strategies. In EMME a strategy is a set of rules that allows a traveller to reach their destination. Due to the waiting time at stops in a transit network, a traveller may select from a more complex choice set than just a simple path toward a destination, so this strategy could constitute a single path using a single transit line or a number of paths each involving one or more transit lines. In a strategy, the traveller chooses a set of paths before embarking on the trip, and at each node where there is waiting, boards the first vehicle to arrive from any of the attractive lines. On boarding a vehicle the traveller knows where he will get off.

13.1.2 The optimal strategy between each origin and destination zone is the one that has the least overall travel time (including access, waiting, in vehicle time etc.). Therefore the optimal strategy will only include that combination of paths and related transit lines that result in the least overall travel time. It follows that if any other paths are considered as part of the strategy they would increase the overall travel time. Those transit lines included in the optimal strategy are called the attractive lines.

13.1.3 The assignment process operates in two stages:

- Calculate the optimal strategy i.e. the set of attractive routes and lines that minimises the overall journey time.

- Assign demand according to that strategy.

13.1.4 In calculating the optimal strategy and assigning demand according to that strategy, the following rules apply:

- Waiting time at a node is related to the combined frequency of all attractive transit lines at that node (combined linearly)

- It is assumed that a traveller's path is determined by the transit line that first arrives at each node

13.1.5 At a node, the probability of a particular transit line arriving first and the proportion of passengers assigned to each attractive route is based on the service frequency of each transit line in relation to the combined frequency of all attractive transit lines at that node.

13.1.6 As the wait time is calculated based on the combined frequency of all attractive transit lines at a node, it is possible that including a transit line that is slower than the current attractive lines could reduce the overall journey time i.e. the wait time saving as a result of the change in combined frequency outweighs the additional journey time of the slower service.

13.1.7   Note that the travel time includes all journey time elements e.g. access time, wait time, in vehicle time, boarding time, and their associated assignment weights e.g. wait time factor, wait time weight, boarding penalty etc. The way that the overall journey time is calculated taking into account multiple paths and combined service frequencies is outlined below.

13.1.8   The algorithm used for the optimal strategy is shown in Table A1 and a working example is included below. The algorithm is applied to each origin destination pair. It is applied from the destination zone working back towards the origin zone. The algorithm starts from the nearest node connected to the destination zone.

- Firstly it determines attractive transit lines outgoing from that node to the destination.
- Then it looks at services incoming to that node (this helps determine whether alighting at the node are attractive options compared to remaining on a service).
- It then looks at the next nearest node.

Table A1: Optimum strategy algorithm

Determination of attractive lines in optimal strategy
A. Lines outgoing from node Y - Processing of the line with smallest time to destination, t1
A1. Calculate combined frequency: f = f1
A2. Calculate waiting time: wait = headway fraction / f
A3. Calculate average time to destination: average_t = t1
A4. Calculate expected total time to destination: u = wait + average_t
B. For each line, in increasing order of time to destination, tl
If tl > u line l and the following are not attractive
If tl ≤ u line l is attractive
B1Calculate combined frequency: f = f + fl
B2. Proportion of demand for each attractive line: pl' = fl' / f
Proportion using Blue
B3. Calculate waiting time: wait = headway fraction / f
B4 Calculate average time to destination: average_t = Sum pl' * tl'
B5. Calculate expected total time to destination: u = wait + average_t
C. Repeat for Lines Incoming to node Y
D. Repeat for other nodes in increasing order of travel time

13.1.9   The process of determining attractive lines and hence the optimal strategy, using a simple example based on a path between two nodes is outlined below.

- Starting with the transit line with the smallest travel time, assuming it arrives immediately (i.e. ignoring wait time)
- Calculates the total journey time (including wait time)

- Select the next fastest transit line, assuming it arrives immediately (i.e. ignoring wait time)

- Compares travel time of this line with the total journey time using the fastest transit line.

- If the travel time it is less than total travel time, then the line is considered attractive as it will reduce the overall total travel time.

- If the travel time it is less than total travel time, the total travel time is recalculated taking into account the combined frequency and the proportion of demand using each of the attractive transit lines.

- The process is repeated using the next fastest transit line.

13.1.10   At any point in the process if the next fastest path is slower than the exiting total travel time then this line and any slower lines are not attractive and the set of attractive lines has been identified.

13.1.11   The next step is to load demand on to the attractive transit lines according to that strategy. The proportion of demand is allocated to each transit line based on the frequency of individual transit lines in relation to the overall combined frequency. Demand is loaded on starting from the origin zone working towards the destination zone.

13.1.12   The assignment algorithm based on frequency and journey time is an alternative public transport assignment algorithm that provides the facility to distribute flow between attractive lines based on a combination of frequency and travel time. The weighting of frequency may be modified on a global, node (stop) or transit line (rail service) basis.

13.1.13   The assignment algorithm for this option works in a similar way to the Optimal strategy, with a few key differences. These relate to the calculation of combined frequency, overall travel time and the allocation of demand to each transit line.

13.1.14   In calculating the combined frequency, an adjustment is made to the frequency of the next fastest transit line being considered to reflect the difference in journey time between it and the current attractive lines. Modification to the calculation of combined frequency has a knock on impact on the calculation of journey time and the allocation of demand to individual services.

13.1.15   The assignment algorithm is presented below; a worked example is given at the end of this Annex.

Table A2:   Assignment Algorithm Based on Frequency and Journey Time

Determination of attractive lines
A. Lines outgoing from node Y - Processing of the line with smallest time to destination, t_1
A1. Calculate combined frequency: $f = f_1$
A2. Calculate waiting time: wait = headway fraction / f
A3. Calculate average time to destination: average_t = t_1
A4. Calculate expected total time to destination: u = wait + average_t

Determination of attractive lines
B For each line, in increasing order of time to destination, tl
If tl > u line l and the following are not attractive
If tl ≤ u line l is attractive
B2 Calculate p_adj, where p_adjl = 1 – (t1 - average_t)/wait
B3 Calculate combined frequency: f = f + p_adjl * fl
B4 Calculate Proportion of demand for each attractive line:
pl' = p_adjl' * fl' / f
B5. Calculate waiting time: wait = headway fraction / f
B6 Calculate average time to destination: average_t = Sum pl' * tl'
B7. Calculate expected total time to destination: u = wait + average_t
C. Repeat for Lines Incoming to node Y
D. Repeat for other nodes in increasing order of travel time

13.1.16   The initial steps of the algorithm are identical to the Optimal strategy. When a second line is identified as being attractive the calculation of the combined frequency is different to that in the Optimal strategy. In the optimal strategy the combined frequency is simply the sum of the frequency of the individual services. In the frequency and journey time assignment an adjustment factor is calculated for the line to be added to the set of attractive lines. The adjustment factor effectively reduces the frequency of the next fastest transit line by the proportion of additional journey time compared to the frequency of the existing attractive lines. This adjustment factor is applied to the frequency of the next fastest transit line when calculating the combined frequency, reducing the frequency of the transit line and effectively increasing headway and wait time. This reduces the combined frequency and increases the headway and wait time compared to the Optimal strategy.

13.1.17   As the combined frequency is used to calculate the wait time, and the wait time makes up part of the overall journey time, the overall journey time is different to that calculated in the Optimal strategy. The adjustment factor is also to service frequency in the calculation of demand by services effectively reducing the proportion of demand allocated to the slower service.

13.1.18   Note that compared to the optimal strategy, the frequency and time based strategy is sub optimal, i.e. the calculation of minimum costs results in greater minimal costs than the optimal solution. This is because during the calculation of combined frequency the frequency of the next quickest route which is being looked at is factored to reflect the increased journey time compared to the optimal solution. For an identical network this can change the number of attractive lines.

13.2 Annex B: Allowing for Distance effects on Value of Time in the Demand Model

13.2.1 For each OD pair, VOTs need to be calculated in order to convert monetary costs into generalised time. These VOTs vary as a function of the one-way trip distance in miles, as measured by the highway skims for the purpose in question.

Commute and other travel

13.2.2 The cost damping relationship used for commute and other travel has been taken from WebTAG Unit 3.10.2, Modelling Road Pricing (February 2013, Draft for Consultation). Appendix A of the unit gives the following formulation for the calculation of VOTs in 2010 p/min:

Equation B-1

$$VoT = G^{\alpha} K \left[\frac{\beta_t}{\beta_c} \right] \left(\frac{Inc'}{Inc'_0} \right)^{\eta_{inc}} \left(\frac{D}{D_0} \right)^{\eta_c}$$

where:

- G is the real growth in GDP/capita relative to 2010, and α = 0.8 is the recommended elasticity of VOT to GDP/capita for non-work travel;

- hence, $G\alpha$ = 1.280 is a factor to account for real terms growth in GDP between 1994 and 2010;

- K is a correction factor for inflation between 1994 and 2010 (1994 being the date of the VoT study), which is calculated as being 1.429;

- Inc' represents the household income in £'000 p.a. based on local data;

- Inc'o is set equal to K' multiplied by 35 (which was the average household income from the sample);

- K' is a correction for inflation between 1994 and the year in while the local data is collected, which can be calculated from the GDP deflator in TAG Unit 3.5.6 in the relevant year divided by the same quantity for 1994; and

- D is the one-way trip distance in miles from the local data and Do is set to 7.58

13.2.3 For future years, VoT is further increased by $G\alpha$, where G is the assumed growth in real GDP/capita from 2010 onwards.

13.2.4 In point of fact, K' has been based on the RPI. Table B1 summarises the K' values used to implement this formula. Since all costs have been calculated in 2010/11 UK financial year prices we use a K' factor defined on that basis.

Table B1: K inflation factors

Year	RPI (CHAW Index)	K'
1994	144.1	1.000
2010/11	226.5	1.572

Source: Table 20, Annual Average Consumer Price Indices, May 2012, Office for National Statistics. Downloaded from: http://www.ons.gov.uk/ons/index.html, June 2012.

13.2.5   The remaining parameters in the VoT formula are defined in Table B2.

Table B2:   Value of time function parameters

Parameter	Commuting	Other
βt (time coefficient)	-0.10098	-0.082918
βc (cost (distance) coefficient)	-0.024729	-0.022275
Inc0	35 × K'	35 × K'
D0	7.58	7.58
ηinc (income elasticity)	0.358773	0.156806
ηc (cost (distance) elasticity)	0.421305	0.314727

Source: Table A3, WebTAG Unit 3.10.2C (February 2013, Consultation Status).

13.2.6   Distances are expressed in miles, incomes in thousands of pounds. For models with an income segmentation, we calculate the income term Inc using the mid-point of the household income band for the tour record. For models without income segmentation, we calculate an overall average income across the tour records for the model purpose. These average incomes in 2010/11 prices are £60.091k for commute and £45.583k for other travel. The corresponding average distances were 104.05 miles for commute and 150.76 for other.

Business travel

13.2.7   Results have been obtained for a model using a distance-damped relationship with a distance elasticity of 0.36 calibrated to the LDM SP data. The implied VOTs for application in the RP models can be calculated as:

-

Equation B-2

$$VoT = (1.054).G\frac{\beta_{IVT}}{\beta_{cost}D^{\eta}}$$

-

where:

- G is the real growth in GDP/capita relative to 2010;

- 1.054 is a factor to convert the SP VOTs into 2010/11 prices;

- βIVT is the in-vehicle time parameter (utils/min);

- βcost is the cost parameter (utils/pence);

- D is the one-way distance in miles (from the highway network); and

- η is the distance elasticity, fixed to -0.36.

The values for the cost and in-vehicle time parameters estimated from the SP data are summarised in Table B3.

Table B3: SP distance-damped VOT parameters

Parameter	Value
βIVT	-0.00638
βcost	-0.00073

For business travel, the average distance from the NTS LD data was 154.94 miles.

13.3   Annex C – WebTAG Car Cost Calculations in the Demand Model

13.3.1   This Annex documents the car cost calculations that have been made using the procedures set out in WebTAG Unit 3.5.6, Values of Time and Vehicle Operating Costs (October 2012). The Demand Model estimation uses NTS choice data covering the 2002-2010 period, and therefore car cost information for each individual year in this period is required for the model estimations. Information is also required for the SCM estimation, which uses 2005 NRTS data.

Fuel costs

13.3.2   Fuel consumption is calculated using a function of the form:

$L = a/v + b + c.v + d.v2$
    where:

- L is consumption, in litres per kilometre

- v is average speed in kilometres per hour

- a, b, c, d are parameters defined for each vehicle category

13.3.3   The values for the consumption parameters a, b, c and d are summarised in the following table.

Table C1: Fuel consumption formula (l/km, 2010 prices and values)

Vehicle category	a	b	c	D
Petrol car	0.964022581	0.041448033	-4.54163E-05	2.01346E-06
Diesel car	0.437094041	0.058616489	-0.00052488	4.12709E-06

Source: Table 10, WebTAG Unit 3.5.6 (October 2012).

13.3.4   Actual vehicle efficiency improvements for 2006-2010 are available from the latest version of WebTAG. For the 2002-2006 period, actual fuel efficiency improvements given in an earlier version of WebTAG were used. The following table summarises the information, and shows how it has been combined to calculate efficiency changes relative to 2010.

Table C2: Vehicle fuel efficiency improvements

Period	Change in efficiency		Year	Factor relative to 2010	
	Petrol	Diesel		Petrol	Diesel
2001-2002	n/a	n/a	2002	1.08183	1.10049
2002-2003	-0.73	-1.15	2003	1.07393	1.08783
2003-2004	-0.71	-1.19	2004	1.06631	1.07488
2004-2005	-0.67	-2.07	2005	1.05917	1.05263
2005-2006	-1.03	-0.99	2006	1.04826	1.04221

2006-2007	-0.42	-0.49	2007	1.04385	1.03711
2007-2008	-1.05	-1.07	2008	1.03289	1.02601
2008-2009	-1.78	-0.92	2009	1.01451	1.01657
2009-2010	-1.43	-1.63	2010	1.00000	1.00000

Source: Table 13, WebTAG 3.5.6 (March 2010, in draft): 2002-2006 changes in efficiency. Table 13, WebTAG 3.5.6 (October 2012), 2006-2010 changes in efficiency.

13.3.5   Historical information on fuel prices, the levels of VAT levied on fuel, is available from WebTAG and this information is used to calculate the price of petrol and diesel for each modelled year. This information is summarised in Table C3.

Table C3: Petrol and diesel fuel prices by year (2010 prices)

Year	Resource cost		Duty		VAT
	Petrol (p/l)	Diesel (p/l)	Petrol (p/l)	Diesel (p/l)	%
2002	20.39	22.46	55.92	55.92	17.5
2003	22.39	24.06	55.02	55.02	17.5
2004	24.87	26.31	54.80	54.80	17.5
2005	30.76	34.41	53.62	53.62	17.5
2006	33.91	37.30	52.06	52.06	17.5
2007	34.08	36.21	52.68	52.68	17.5
2008	42.83	51.75	52.92	52.82	17.3
2009	33.16	37.01	55.95	55.95	15.0
2010	42.57	44.31	57.19	57.19	17.5

Source: Table 11a, WebTAG 3.5.6 (October 2012)

13.3.6   Consistent with the guidance in WebTAG 3.5.6, VAT is not applied to the fuel cost calculations for business travel because businesses can reclaim VAT. However, fuel duty cannot be reclaimed and therefore the fuel duty is included in the car cost calculations. For commute and other travel, the cost is taken as resource cost plus duty plus VAT.

13.3.7   The proportions of the car fleet using petrol and diesel are available for 2004 and 2010 from WebTAG 3.5.6, following the guidance in WebTAG values for intermediate years are determined using linear interpolation. For 2002 and 2003, values were taken from an earlier version.

Table C4: Proportion of cars using petrol and diesel

Year	Petrol	Diesel
2002	0.848	0.152
2003	0.758	0.242
2004	0.7328	0.2672
2005	0.7095	0.2906

2006	0.6861	0.3139
2007	0.6628	0.3373
2008	0.6394	0.3606
2009	0.6161	0.3840
2010	0.5927	0.4073

Source: Table 12, WebTAG 3.5.6 (March 2010, in draft): 2002 and 2003 values. Table 12, WebTAG Unit 3.5.6 (October 2012): 2004 and 2010 values, and then values for 2005 to 2009 determined by linear interpolation.

13.3.8   Fuel consumption is calculated separately for petrol and diesel vehicles using the information from Table C1, and this is combined with information on changes in efficiency from Table C2 and fuel prices from Table C3 to calculate car costs for petrol and diesel vehicles for the year that is being modelled. Finally, the fleet proportion information from Table C4 is used to calculate the costs for an average vehicle. All of these costs are calculated as p/km in 2010 prices.

Non-fuel costs

13.3.9   Non-fuel costs include oil, tyres, maintenance, depreciation and vehicle capital saving (only for vehicles in working time). Non-fuel costs are calculated using a function of the form:

- $C = a_1 + b_1/v$

where:
- C is cost in pence per kilometre travelled
- v is average link speed in kilometres per hour
- a_1 is a parameter for distance related costs defined for each vehicle category
- b_1 is a parameter for vehicle capital savings defined for each vehicle category (only relevant to working vehicles)

13.3.10   Table C5 summarises the non-fuel cost parameters for cars used in work and non-work time.

Table C5: Non-fuel cost function parameters (2010 prices and values)

Vehicle category	a1 (pence/km)	b1 (pence/hr)
Work car	4.966	135.946
Non-work car	3.846	0.000

Source: WebTAG 3.5.6 Table 15

13.3.11   Non-fuel VOCs by fuel/energy type are assumed to remain constant in real terms over a forecast period. Following the same logic, it has been assumed that these 2010 values can be applied to model NTS choice data over the 2002-2009 period without adjustment.

13.4 Annex D – further information on HAM treatment of costs

13.4.1 A detailed list of the rail cost skims and how they are used in the LASAM generalised cost equations is provided in Table D-1. Similarly, highway cost skims are described in Table D-2 and air cost skims in Table D-3.

Table D-1: Elements of Generalised Cost - Rail

PLD Cost Element	Description	LASAM Equivalent
Rail Fare (£)	Average yields by journey purpose produced in Atkins EDGE model based on inputs from NMF (DfT) revenue and journey data	Rail Fare - converted to pence
In Vehicle Time (mins)	Time spent on train	In Vehicle Time (mins)
Auxiliary Transit Time (mins)	For Heathrow trips the auxiliary transit time includes car access time to the station or PT access time to the station (it also includes tube transfer times between terminals in London). It also potentially includes PT transfer times at the destination end i.e. the distance from the station to the airport terminals, or requirement to transfer	Access time + Walk Time
Total Wait Time (mins)	40% of headway	Increased to 50% to be consistent with LASAM, capped at 40min
Rail only Boardings	This is the average number of trains required to get from A to B. Using the tube to transfer between stations is included in the 'aux transit time', and not counted as a boarding	Interchanges = rail only boarding -1
Bus Add Crowd Time	Skim of the PDFH crowding function (Minutes)	Not included

Table D-2: Elements of Generalised Cost - Highway

| PLD	Description	LASAM Equivalent
Vehicles Operating Cost	A combination of fuel and no fuel operating costs, related to distance and average speed	Vehicle Operating Cost
Auto Times (mins)	Time spent in car	Time
Auto Distance (kms)	Highway distance	Distance

Table D-3: Elements of Generalised Cost - Air

| PLD	Description	LASAM Equivalent
Air Fares (£)	One way fares	Air Fare
In Vehicle Time (mins)	Time spent in plane	In vehicle time
Auxiliary Transit Time (mins)	Car Time + Park/Access Penalties +VOCs	Access
Wait Time (mins)	Time spent in airport waiting	Wait Time

13.5 Annex E – Benefit Calculation by Numerical Integation

13.5.1   When calculating benefits at the station to station level we can encounter large changes in costs and demand between the Do Minimum and the Do Something situation. This can result in a breakdown of the rule of a half [RoH], which assumes that the demand curve can be treated as a straight line between the two points: in practice the assumption of linearity may not be justified for large changes, implying that the RoH may be overestimating the benefits.

13.5.2   A better estimate of the benefits can be made using a numerical integration approach (ie calculating the area under the demand curve between the Do Minimum and the Do Something). This method involves creating a series of steps between the two points, and requires only the calculation of the demand at each intermediate cost point, with the RoH then applied separately to each step. The distribution of steps need not be even between the end points, and in particular there may need to be a higher density of stages near the Do Minimum as we expect the demand curve to be less linear at this point (i.e. the error from a linear approximation to the curve will be at its greatest). The sum of the benefits for all of these steps gives the most reliable estimate of benefits and revenue.

13.5.3   The theory behind the rule of half is set out in section 3 of the TUBA guidance note (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/492792/tuba-general-guidance-and-advice.pdf). The following graphs from the TUBA guidance illustrate the principle. The Figure below shows a demand curve and a supply curve that shifts between the do-minimum and do-something as a result of implementing a transport scheme. The shaded area represents the change in the quantity known as the consumer surplus. If we approximate the demand curve as a straight line then this area can be calculated using the rule of a half:

$$Benefit = \frac{1}{2}\left(T^0 + T1\right)\left(C^0 - C^1\right)$$
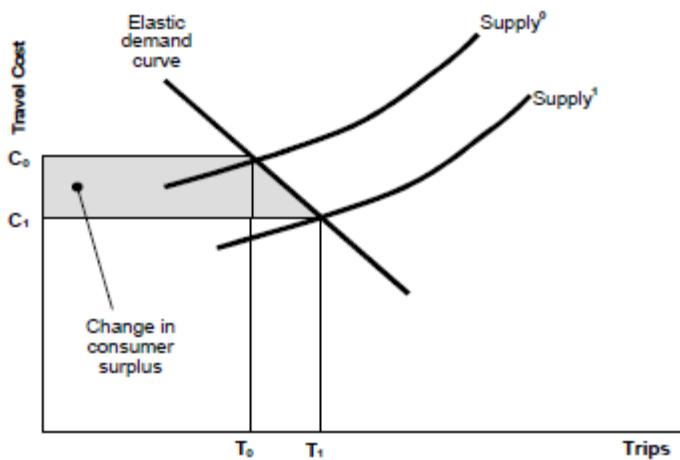
13.5.4   The figure below from the TUBA Guidance shows what can happen to the standard benefit calculation when cost changes are large:
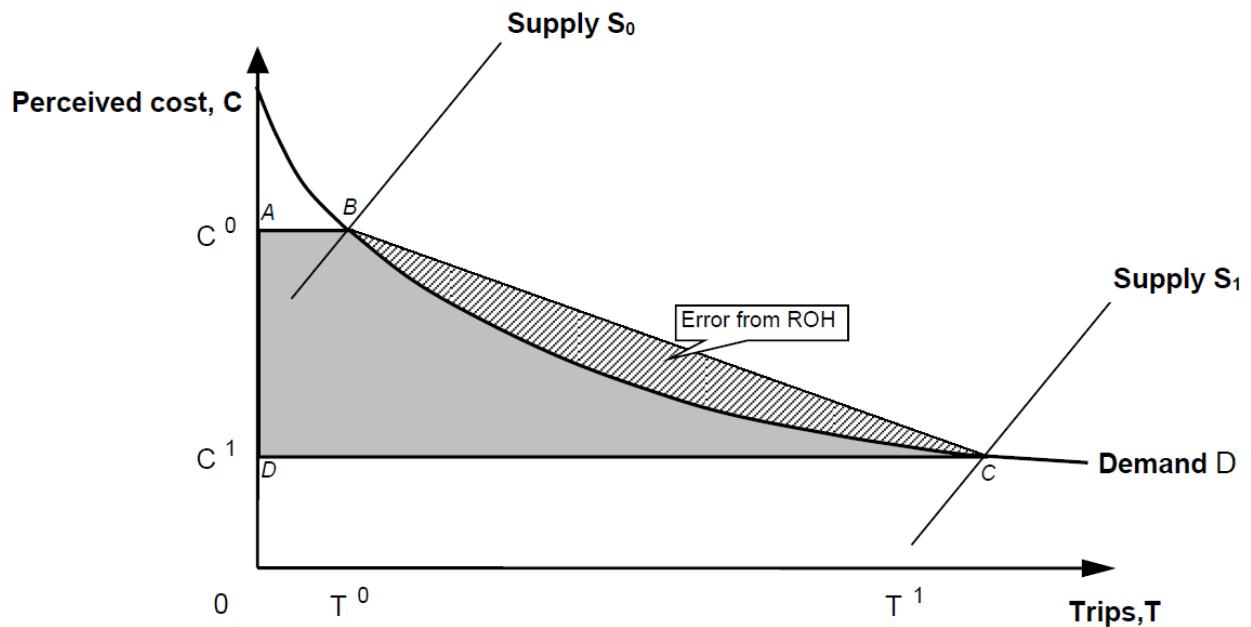
13.5.5   The solution recommended in the guidance to deal with large cost changes is to create a series of intermediate points between the do-minimum and do-something and apply the rule of a half to each pair of points in sequence. In effect, this is approximating the demand curve as a sequence of straight lines rather than a single straight line, as shown in the figure below.
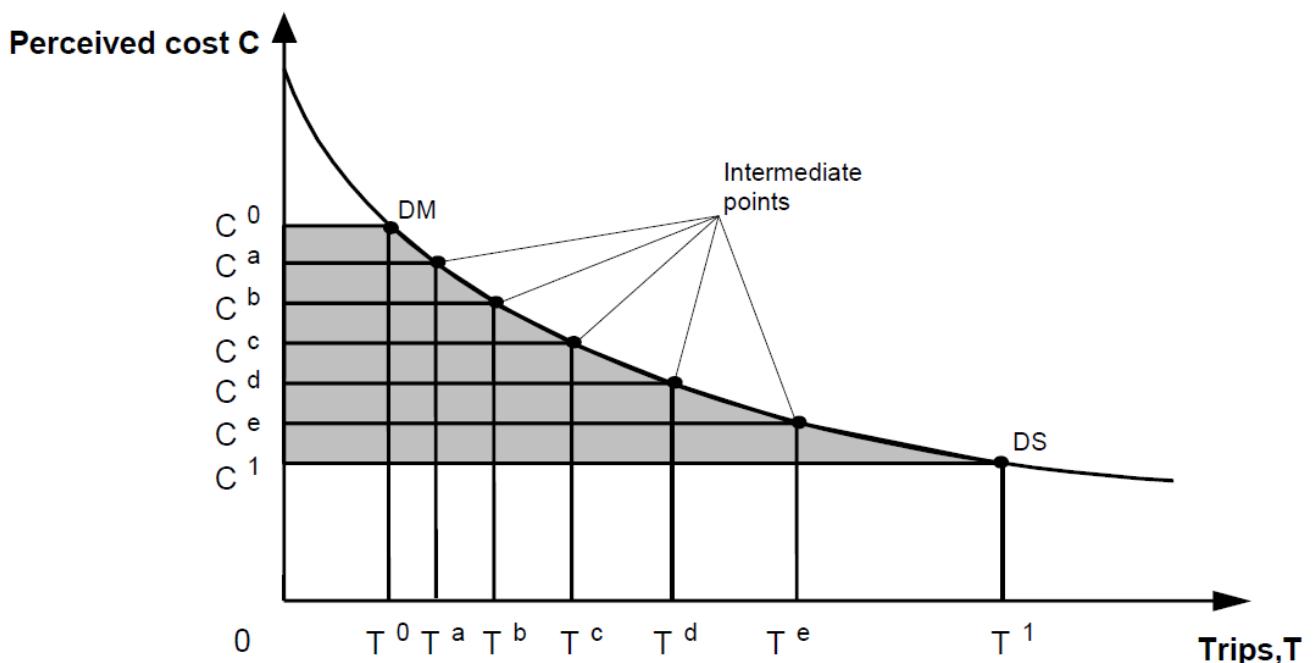
13.5.6   For the intermediate points we are only interested in what comes out of the demand model for a given set of costs, i.e. we are trying to understand the shape of the demand curve. There is no need to run any assignments, and therefore no need to code fictitious network scenarios for these points.

13.5.7   The Numerical Integration Process macro consists of the following steps:

- Run the Station Choice Model with a pre-determined set of costs based on the appraisal do-minimum demand for the step;

- Transfer the output costs skims to the demand model;

- Run the demand model with the new cost skims to get a new set of demand;

- Run the ADTM (Heathrow Model) with the skims from the SCM and the new demand to get a distribution of the Heathrow International trips in the right matrices;

- Run the SCM to get the do-something station to station demand based on the new do-something PLD to PLD demand;

- Run the pre-processing step for the appraisal.

13.5.8   The macro is then run for each intermediate point of the numerical integration process and each of the forecast years.

13.5.9   The points chosen for the final Numerical Integration results are:

- Do-minimum
- 85% of do-minimum, 15% of do-something costs ie is $C_{dm} + 0.15\,(C_{ds}\text{-}C_{dm})$
- 67% of do-minimum, 33% of do-something costs ie is $C_{dm} + 0.33\,(C_{ds}\text{-}C_{dm})$
- 50% of do-minimum, 50% of do-something costs ie is $C_{dm} + 0.50\,(C_{ds}\text{-}C_{dm})$
- 34% of do-minimum, 66% of do-something costs ie is $C_{dm} + 0.66\,(C_{ds}\text{-}C_{dm})$
- 20% of do-minimum, 80% of do-something costs ie is $C_{dm} + 0.80\,(C_{ds}\text{-}C_{dm})$
- 10% of do-minimum, 90% of do-something costs ie is $C_{dm} + 0.90\,(C_{ds}\text{-}C_{dm})$
- Do-something.