DeltaRail

In Confidence
DeltaRail-15951-005
Issue 2

Station Usage

2010/11

February 2012

DeltaRail

Title	Station Usage 2010/11
Customer	Office of Rail Regulation
Customer reference	itt_61 ORR/st/596
Confidentiality, copyright and reproduction	
File reference	LD15951 005
Report number	DeltaRail-15951-005
Issue number	Issue 2

DeltaRail
3rd Floor Central House
London
WC1H 0JN

	Name	Signature	Date
Author	Javier Morillo		07 Feb 2012
Reviewed by	Brian Ball		14 Feb 2012
Approved by	Brian Ball		14 Feb 2012

Executive Summary

This report explains the information contained within the Station Usage file (Station Usage 2010-11.xls). The report provides guidance to the methodology followed during the process of creating this file for financial year 2010/11.

Station Usage data consists of estimates of the total numbers of people:

- Travelling from or to the station (entries & exits); and
- Interchanging at the station (interchanges).

Information is given for all the national rail stations around England, Scotland, and Wales based on tickets sales data. These results are the most recent in a series DeltaRail have supplied since 1997/98. The spreadsheet is in a similar format to those previously provided.

Station Usage data is generated from the O-D Matrix. In 2008/09 the O-D Matrix was integrated with the production of MOIRA Replacement's Demand Matrix (MOIRA is the rail planning tool used by the rail industry). This has brought substantial benefits as MOIRA Replacement's Demand Matrix includes an estimate of journeys and revenue made on zonal products sold by Passenger Transport Executives (PTEs), to provide a more complete representation of travel on the national rail network. This was previously a major deficiency of the O-D Matrix and Station Usage.

A thorough programme of checks has been undertaken on the Station Usage data. Results of these checks are listed in the Station Usage spreadsheet that accompanies this report, and further details are given in Section 7. Users of Station Usage information should take note of the limitations of this dataset, outlined in Section 9.

Contents

1 Introduction 1

2 MOIRA Demand Matrix – Base Data 2

2.1 Overview 2

2.2 Underlying Base Data - LENNON 2

2.3 Ticket Type Definitions 3

2.4 Infills for London Travelcards, Major Urban Areas (PTE) & Airports 3

2.5 MOIRA Replacement Documentation 4

3 Summary of Results 5

3.1 Overview of the Entries and Exits Results 5

3.2 Overview of the Interchanges Results 8

4 Station Usage File Definition 10

5 Entries and Exits Methodology 11

5.1 Overview 11

5.2 Methodology Changes 11

5.3 PTE Infill 13

5.4 Unknown Destinations 13

6 Interchanges Methodology 14

7 Checks on Station Usage: Significant Growth or Decline .. 15

7.1 Oyster Use on Rail 16

7.2 Joint Rail & TfL Stations 17

7.3 Other Impacts on Station Usage 17

8 Regions, Counties and Districts 19

9 Limitations of the LENNON data 20

9.1 Known Problems of Data Capture 20

9.2 Travelcards 20

9.3 Return and Single Journey Tickets 21

9.4 Multiple Tickets 21

9.5 Rail Staff Passes 21

9.6 Ticketless Travel 21
9.7 Other Rail Systems 21
9.8 Factors Affecting the Data 22
9.9 Journey Factors 22
Appendix 1 – Overview of the ORCATS Allocation Process.... 24
Appendix 2 – Methodology: Non-Station Tickets 25
Appendix 3 – Data Excluded From Station Usage 29

1 Introduction

DeltaRail has provided a series of Station Usage data for the rail industry in previous years. This report accompanies the Station Usage data for 2010/11. It gives details of the process and outputs in producing the Station Usage file for financial year 2010/11, on behalf of the Office of Rail Regulation (ORR).

DeltaRail are providing the ORR with an MS Excel file, "Station Usage 2010-11.xls" containing entries, exits and interchanges made at stations throughout England, Scotland and Wales, for the financial year 1st April 2010 to 31st March 2011. For the entries and exits, figures are split into the three main categories of the available ticket products (Full, Reduced, and Season).

2 MOIRA Demand Matrix – Base Data

2.1 Overview

All the estimates of station usage, exits, entries and interchanges, are derived from the Origin Destination Matrix (O-D Matrix). This O-D Matrix, also produced by DeltaRail for the ORR, is created from the MOIRA[1] Replacement Demand Matrix.

The base data for this whole process is LENNON[2] ticket sales, with the addition of infills for London Travelcards, Airport links and zonal products sold by Passenger Transport Executives[3] (PTEs). The use of the MOIRA Replacement Demand matrix as the base input was a significant change made in 2008/09. The O-D Matrix is now consistent with MOIRA Replacement which is the rail industry's principal planning tool, and therefore includes a more complete representation of travel on the national rail network.

2.2 Underlying Base Data - LENNON

The underlying matrix of ticket sales and associated journeys and revenue used in the current MOIRA model is derived from LENNON. It is based on an extract from LENNON, produced by Atos, of total sales revenue and journeys for the year, broken down by flow (origin and destination National Location Code (NLC)), route code and by product type (CTOT). However, as there are known omissions in this data in respect of Transport for London (TfL) and PTE zonal tickets, and non-National Rail tickets on some airport services, there needs to be a "matrix infilling" exercise undertaken to estimate a complete origin-destination matrix.

Infilling is required as there are some journeys/revenue which do not appear in the underlying matrix, related to particular ticket types. There are three main such cases:

- Tickets with non geographical destinations, e.g. zonal products, Rovers
- Tickets sold at some non National Rail (RSP: Retail Settlement Plan) outlets, e.g. newsagents
- Tickets which do not appear in LENNON at all. This includes some Train Operating Company (TOC) tickets on airport flows, also for those TOCs which fall outside the Rail Settlement Plan.

---

[1] MOIRA is the rail industry's tool for forecasting the impact of timetables on passenger revenue

[2] LENNON is the rail industry's ticketing and revenue system

[3] It should be noted that for convenience and clarity, the term 'Passenger Transport Executives' (PTEs) is used in this report, though they are officially now designated Integrated Transport Authorities and Strathclyde Partnership for Transport.

Certain tickets with destination codes that are not national rail stations are included in the MOIRA Replacement demand matrices, being mapped to the corresponding rail station. These Rail Links usually include a third party element, such as to a bus zone, or tourist attraction. The MOIRA Replacement demand matrix includes the journeys and the net revenue associated with such tickets.

Data excluded from the MOIRA Replacement demand matrix is set out in Appendix 3.

2.3 Ticket Type Definitions

Within the base demand matrices, journeys and revenue have been sub-divided into the following four types, each of which is further split by First & Standard Class:

1. Full: all walk-up undiscounted single or return tickets, whether or not issued with a status discount (child, railcard etc)
2. Reduced: all walk-up discounted single or return tickets, whether or not issued with a status discount (child, railcard etc)
3. Advance: all advance-purchase tickets
4. Seasons: all multi-use tickets

2.4 Infills for London Travelcards, Major Urban Areas (PTE) & Airports

There are certain areas within the underlying matrix where demand and revenue are under-estimated, in particular:

- **Within London Travelcard area.** Whilst the underlying matrix includes an estimate of journeys made on Day Travelcards / Travelcard seasons purchased at National Rail stations, it does not include a significant number of national rail trips made using Travelcards purchased at Tube stations, travel shops and newsagents.

- **Within Passenger Transport Executive (PTE) areas.** The underlying matrix excludes virtually all rail trips made on PTE-sponsored tickets, which are usually zonal and often multimodal.

- **Trips to/from Airports.** The underlying matrix includes many trips to/from airports, but excludes all Heathrow Express journeys, and some tickets sold for Gatwick Express, Stansted Express and other airport operators.

There are also other ticket sales which are not included in the underlying matrix, but these are insignificant and disparate in nature and so the matrix infilling covers the three flow types identified above. It should also be noted that journeys with no

associated ticket sales such as staff travel, and particularly fare evaders, are not including in the origin-destination matrix.

The two major “infills” are for the London Travelcard area (sales made by Transport for London (TfL)), and for PTEs, since in both cases a substantial proportion of the rail journeys made use multimodal travelcard type of tickets.

The third infill, for Airports, estimates the significant number of rail journeys on Gatwick and Stansted Express, made on tickets sold outside of the RSP system i.e. not sold by National Rail outlets. Journeys on Heathrow Express are excluded from the MOIRA Replacement Demand Matrix.

Each of these three infills has been taken from the MOIRA Replacement Demand Matrix.

2.5 MOIRA Replacement Documentation

A full description of the process of creating the demand matrix for MOIRA Replacement is contained in the report: ‘MOIRA Replacement, Stage: MS-5-1, Demand Matrix Documentation, Version: 0.2’ dated August 2009.

3 Summary of Results

The following table gives the total number of entries, exits, and interchanges made over the whole network for 2010/11, compared with the previous year.

Table 3-1: Entries, Exits and Interchanges for 2009/10 to 2010/11 (millions)

Year	Entries	Exits	Entries & Exits	Interchanges
2010/11	1,156.90	1,156.90	2,313.79	165.53
2009/10	1,065.39	1,065.39	2,130.78	153.33

Overall, the increase in entries and exits is around 8.6% in 2010/11, compared with the previous year.

3.1 Overview of the Entries and Exits Results

In this section we set out a summary of the overall entries and exits results. The spreadsheet contains entries and exits results for 2,531 stations, compared with 2,525 last year. The tables below show the stations no longer in station usage this year, and the new stations that have been added.

In 2010/11, nine stations were removed, and fifteen stations were added, as listed in the following tables.

Table 3-2 – Stations in 2009/10 but not in 2010/11

NLC	Name
2877	Dean Lane
2652	Derker
2951	Failsworth
2901	Hollinwood
2922	Milnrow
2923	New Hey
2907	Oldham Mumps
2908	Oldham Werneth
2925	Shaw and Crompton

Table 3-3 – New Stations added in 2010/11

NLC	Name
0990	Armadale
0991	Blackridge
0992	Caldercruix
1659	Canada Water
1443	Dalston Junction
1022	Haggerston
1023	Hoxton
5786	Okehampton
1039	Rotherhithe
8724	Sampford Courtenay
1082	Shadwell
1024	Shoreditch High Street
1083	Surrey Quays
1085	Wapping
4935	Whitechapel

The table below shows data for the ten stations with the highest numbers of entries and exits for 2010/11.

Table 3-4: Top 10 stations based on the entries and exits made for 2010/11

Rank This Year	NLC	Station Name	1011 Entries & Exits	0910 Entries & Exits	Change	Rank Last Year
1	5598	London Waterloo	91,750,382	86,397,666	6%	1
2	5426	London Victoria	73,573,492	70,224,543	5%	2
3	6965	London Liverpool Street	55,769,423	51,596,155	8%	3
4	5148	London Bridge	51,478,131	48,723,068	6%	4
5	5143	London Charing Cross	37,222,298	36,459,945	2%	5
6	1444	London Euston	34,073,413	30,068,092	13%	6
7	3087	London Paddington	32,200,316	29,104,198	11%	7
8	6121	London King's Cross	26,254,644	24,817,616	6%	9
9	9813	Glasgow Central	24,950,987	23,809,949	5%	10
10	1127	Birmingham New Street	24,686,632	25,267,757	-2%	8

The total journeys made at one of the top ten stations account for a total of 452 million, 6.0% more than the 426m journeys made at the top ten stations of last year. The top ten stations account for 20% of all entries and exits, the same as in 2009/10.

3.2 Overview of the Interchanges Results

In all, around 165 million interchanges are estimated to have been made among National Rail operated services (interchanges between rail and tube or other modes are excluded). This is an increase of 8.0% compared to the 2009/10 results (153.3 million). The ten top stations are listed in the table below.

Table 3-5: Top 10 stations based on the interchanges made for 2010/11

Rank This Year	NLC	Station Name	1011 Interchanges	0910 Interchanges	Change	Rank Last Year
1	5595	Clapham Junction	20,667,636	20,520,598	1%	1
2	5148	London Bridge	7,346,732	7,060,751	4%	3
3	5355	East Croydon	7,113,300	7,120,189	0%	2
4	5598	London Waterloo	5,772,501	5,466,424	6%	4
5	5426	London Victoria	4,800,979	5,078,272	-5%	5
6	1127	Birmingham New Street	4,319,983	3,957,161	9%	6
7	3149	Reading	2,898,671	2,617,998	11%	8
8	9813	Glasgow Central	2,641,171	2,284,406	16%	10
9	2968	Manchester Piccadilly	2,624,292	1,924,480	36%	12
10	1555	London St.Pancras	2,158,903	2,080,620	4%	11

Interchanges occurred at 536 stations in 2010/11 compared to the 523 stations in 2009/10. Stations appearing for the first time in 2010/11 and those not seen this time are listed below.

Table 3-6: Changes in Interchange Stations in 2010/11.

	Interchanges 2010/11	Interchanges 2009/10
New		
Surrey Quays	685,971	
Dalston (Kingsland)	187,533	
Dalston Junction	182,753	
Wembley Central	24,821	
Boston	3,702	
Yeovil Junction	1,325	
Emsworth	1,260	
Yeovil Pen Mill	916	
Wanstead Park	658	
Corby	615	
Horsforth	382	
Ludlow	299	
Thornford	192	
Smethwick Rolfe Street	43	
Hackney Wick	16	
Altrincham	10	
Thetford	4	
Old		
Penge West		86
Upper Tyndrum		15
Dorchester West		2
Fishguard Harbour		2

The numbers in this table are estimated numbers for actual passenger interchanges made during the year. In some cases the numbers are extremely small.

We have not identified reasons for changes in the interchanging stations. However, it is important to note that interchanges can change significantly from year to year for a variety of reasons. Factors such as new service patterns and changes in journey times play a part. The number of interchanges is based on the rail industry ORCATS model, which predicts passenger choices of rail route and trains used. Refer to Appendix 1 for more information on the ORCATS allocation process.

4 Station Usage File Definition

This spreadsheet lists the entries, exits and interchanges made at stations throughout England, Scotland and Wales in the financial year 2010/11 (1st April 2010 to 31st March 2011). It also gives details about the entries and exits for different ticket categories. It contains data on entries and exits made at rail stations by passengers using the rail network. The fields included in the Station Usage file are:

Table 4-1: Station Usage File

Field	Description
Station (Name, NLC, TLC)	Station Name, NLC: National Location Code, TLC: Three Letter Code
District, County, Region, NUTS2 Code and NUTS2 Spatial Unit for the Station.	Station's geographical location.
Station Facility Owner (SFO)	The company that is the station facility owner (provided by Network Rail in 2008).
Station Group	Name of the Group where applicable. The user of this data may wish to filter on the 'Station Group' column, or create pivot tables, to investigate the results at a group level.
PTE Urban Area Station	Stations within the urban areas covered by PTE services are identified with a flag: 'PTE Urban Area Station'
London Travelcard Area	Stations within the Greater London area where a London Travelcard is valid are identified with a flag: 'London Travelcard Area Station'
London Joint Station	Joint stations which are served by both rail services and TfL services are identified with a flag: 'Joint TfL & TOC Station'
Entries (Full, Reduced, Season, Total)	Entries made at the stations split by ticket categories and in total
Exits (Full, Reduced, Season, Total)	Exits made at the stations split by ticket categories and in total
10/11 Entries & Exits	Sum of Entries and Exits for 2010/11
09/10 Entries & Exits	Sum of Entries and Exits for 2009/10
10/11 Interchanges	Total Interchanges made for 2010/11
Check	Fail if number of entries & exits combined is more than 20% higher or lower than figure in 2009/10
Check Detail	Flag indicates Growth or Decline, and where the change is ignored if less than 15,000 entries & exits
Check Reason	Identified reason(s) for failing the check

5 Entries and Exits Methodology

5.1 Overview

An estimate of the number of people entering and exiting each of the National Rail stations for the financial year 2010/11 is derived from the O-D Matrix created by DeltaRail for the ORR.

Each record in the O-D Matrix reflects an estimate of the actual passenger journeys undertaken on rail services in England, Wales and Scotland. The O-D Matrix contains the number of journeys for each flow, where a flow consisted of a unique origin, destination and route code combination.

The number of entries and exits is calculated for a particular station by summing all journeys starting at the station, and all journeys terminating at the station.

5.2 Methodology Changes

Improvements have been made to the O-D Matrix and Station Usage methodology in the last four years.

Between 2006/07 and 2008/09 the accuracy and usefulness of the O-D Matrix has been improved. This was achieved by applying new procedures on the way journeys with unknown origin and/or destination have been treated, and by including journeys that were previously excluded from the file or did not appear in the LENNON sales data. In summary, the main changes were:

- Adding in previously missing journeys, e.g. TfL sold Travelcards, and some airport link tickets (this is undertaken in the production of the MOIRA demand matrix).
- Rail Links such as PlusBus and Attractions. The rail element of these ticket sales is now included (this is undertaken in the production of the MOIRA demand matrix).
- Estimating the split of records for station groups, including London BR, into the constituent individual stations. This methodology was further refined for those groups with no ticket office at one or more stations within the group. (This processing is undertaken in the O-D Matrix),
- Via the integration with the process that creates the MOIRA Replacement Demand Matrix, PTE ticket sales are now included, in addition to TfL sold Travelcards, and some airport link tickets.
- The method for estimating passenger journeys from ticket sales has changed. This is a result of using the MOIRA Replacement Demand Matrix as a starting point. The MOIRA Replacement Demand Matrix does not disaggregate single journeys, and so when estimating passenger journeys all ticket sales have

been split equally into the two directions of travel. This will only have an impact on the O-D Matrix if there is more travel on single tickets away from a station compared to travel to the station, which is not likely to be material. Therefore in the Station Usage file, entries are the same as exits.

In 2009/10 further improvements were made:

- Adding in data for journeys undertaken by Oyster "pay-as-you-go" (PAYG) in the London area. This is undertaken within the base LENNON data, in the production of the MOIRA demand matrix. This applies to journeys made after 1 January 2010 (see Section 7.1).
- Refinement of the methodology used to calculate journeys undertaken using PTE tickets.

There were no enhancements planned for 2010/11, but a few particular amendments had to be made. These especially affected the PTE infill, which is covered in a separate sub-section (see 5.3 below). Others, which are expected to be one-off changes, are:

- A previous adjustment factor, affecting predominantly Farringdon and Elephant and Castle stations in London, was removed. This has the effect of a large increase in flows to/from Farringdon, and reduction to/from Elephant and Castle.

- Several stations in the North West were found (in the original LENNON dataset) to have large numbers of flows with origin station as the BR Stations code, whereas usually the majority of these would have been expected to show the individual station of sale as origin. These were Manchester BR, Warrington BR and Wigan BR. We are investigating the reasons for this change in the LENNON file. This has meant that, despite all the stations in the group having ticket offices, our usual methodology (see Appendix 2, Category 2a) would not work. For 2010/11 we have therefore applied the following process:
    - For Warrington and Wigan, we have used the alternative methodology 2b. The stations involved are on different lines and it has therefore been possible to estimate which station is used (or most attractive) for all the top flows.
    - For Manchester, method 2b would not have been able to split between stations, because they are not all on different lines. Hence for 2010/11 we have had to assume the same percentage changes across all stations in the Manchester BR group.

In addition, 2010/11 figures include a full year of Oyster PAYG specific journeys data which has led to some significant changes in the London area (see Section 7.1).

5.3 PTE Infill

The revenue and journeys associated with PTE Infill have been estimated based on an uplift applied to the previous year's PTE Infill figures. This uplift was taken from the National Rail Trends average change for "Franchised Regional Operators", Tables 1.2b and 1.3b of the National Rail Trends Yearbook:

- PTE Infill revenue increased by +8.0%
- PTE Infill journey increased by +4.6%

The estimated number of journeys and the distribution of those journeys by origin and destination is still based on that estimated as part of the creation of the 2008/9 MOIRA2 Demand Matrix. It has emerged that the original 2008/9 figures which were given for one PTE, West Yorkshire, were not a complete record of all the rail journeys on multimodal tickets which should have been included. A correction has therefore been made in the latest 2010/11 Station Usage figures: this involved uplifting the West Yorkshire PTE Infill, both revenue and journeys figures, by 53% (in addition to the above general increases). Note that within West Yorkshire PTE area, the majority of rail journeys are made on rail-only tickets, i.e. not PTE Infill tickets. Thus the overall effect of this correction is relatively small.

Estimation of number of journeys on PTE tickets is a recognised weakness in the overall compilation of rail demand matrices. ATOC is undertaking work on this issue: when an improved methodology is determined in the future, this may result in a significant change from the current estimates for the PTE Infill.

5.4 Unknown Destinations

Ticket sales do not always tell us where a passenger is travelling, for example where the Origin or Destination is a London Travelcard. As in previous years, we have converted unknown destinations into an estimate of the actual stations that passengers are travelling to. The full detail of this part of the methodology appears in Appendix 2.

6 Interchanges Methodology

An estimate of the number of people interchanging at each station is obtained by combining the number of journeys made on each flow (from the O-D Matrix) with the information on passenger journeys taken from the Central Allocations File (CAF).

The CAF is an output of the ORCATS system which predicts passenger choices of rail route and train used, and determines the allocation of passenger revenue between TOCs. Since ORCATS is a model, the CAF contains estimates rather than actual journeys. However, it is used throughout the rail industry, so it is an appropriate source of data to use for this purpose. Since CAFs are updated with the timetable, not with financial years, no CAF will match the ticket sales data exactly. The December 2010 CAF is used in the creation of the 2010/11 Station Usage, which relates to the timetable in operation for a substantial portion of the 2010/11 financial year.

The CAF contains:

- Origin and destination;
- Route alternatives for each origin and destination, including all interchange points;
- Ticket type data; and
- For each flow, the proportion of passengers who choose to travel on each route alternative as calculated by the ORCATS model.

An overview of the ORCATS allocation process can be found in Appendix 1.

7 Checks on Station Usage: Significant Growth or Decline

A check on the station usage figures is performed to identify all stations at which total Entries plus Exits has changed significantly since the previous year, and is contained in the station usage spreadsheet.

The overall results show an increase of 8.0% in total Entries and Exits from 2009/10 to 2010/11 when compared on like for like basis, i.e. excluding stations closed or opened between the two years. We define a significant change as a deviation of more than +/-20% from this mean: therefore we have investigated all cases where growth is more than 29.6% or there is a decline of more than 13.6% (the factors have been applied multiplicatively to give these overall limits). Only stations whose entries and exits exceeded 15,000 were considered in this check. In total 159 stations failed the check. The accompanying station usage spreadsheet shows likely reasons for the majority of these changes.

Table 7-0 shows the stations with the largest increases in total flow, not including flows of less than 15,000.

Table 7-0 – The Largest Increases in Station Usage

NLC	Station Name	1011 Entries & Exits	0910 Entries & Exits	Increase	Reason
3498	Lelant Saltings	17,224	622	>1000%	Switch to point-to-points from St Ives Bay Line Ranger
3542	Carbis Bay	55,334	7,980	593%	Switch to point-to-points from St Ives Bay Line Ranger
9586	Imperial Wharf	737,388	119,250	518%	Full year of usage for station that opened in Sep 2009
7222	Stratford International	407,497	69,436	487%	Full 12 months of services on HS1 (started Dec 2009)
0577	Farringdon	4,632,081	1,492,179	210%	Oyster PAYG accepted but not captured prior to Jan 2010
0598	Harrow-On-The-Hill	1,309,328	453,158	189%	Oyster PAYG accepted but not captured prior to Jan 2010

Table 7-1 shows stations with significant reductions in flow, again excluding total Entries & Exits of less than 15,000.

Table 7-1 – The Largest Decreases in Station Usage

NLC	Station Name	1011 Entries & Exits	0910 Entries & Exits	Decline	Check Reason
3891	Treorchy	64,980	218,766	70%	Treorchy school travel no longer goes through Lennon
9757	Drumgelloch	58,550	170,940	66%	Station relocated and closed for part of year (Airdrie-Bathgate)
3897	Ton Pentre	39,698	109,428	64%	Treorchy school travel no longer goes through Lennon
2004	Sellafield	219,394	361,974	39%	Single User Station (nuclear power station)
9968	Bowling	31,784	43,916	28%	#No obvious reason to explain change in usage
5126	Sole Street	60,348	79,800	24%	Railheading to eg Strood for HS1 (started Dec 2009)
5762	Lympstone Commando	47,660	60,558	21%	Single User Station (Royal Marine training centre)

There are a number of significant factors affecting the station usage figures in 2010/11, and these are described below. In addition there are many reasons why a station may fail the check, and these, including those provided in previous reports, have been listed below for completeness.

7.1 Oyster Use on Rail

The biggest discontinuity in the data is due to the roll-out of Oyster 'Pay As You Go' (PAYG) at National Rail stations in January 2010. Prior to this date Oyster PAYG was available on selected routes only and was not recorded (in LENNON) on a flow or station basis. After this date Oyster PAYG was available at *all* National Rail stations in the Travelcard Area and recorded by flow.

The 2009/10 data contains roughly 9 months of data prior to January 2010 and 3 months of data after, while the 2010/11 data is wholly after January 2010 when Oyster PAYG, with data capture, had been fully implemented. This can lead to some very large reported growth figures for some stations within the London Travelcard (/Oyster PAYG) area. The new figures, based on recorded use of Oyster PAYG

should be accurate, but the percentage growth may be over-represented since the old figures will be largely estimates made without the benefit of Oyster records.

7.2 Joint Rail & TfL Stations

Joint stations are stations which are served by both rail services and TfL services i.e. Underground or DLR. These stations can have both a TfL and a TOC ticket office, or they may have just a TfL or just a TOC ticket office. Special treatment of the ticket sales at these stations is important to ensure a realistic estimate of passengers using rail services. Passengers travelling on Underground or DLR services should not be included. These stations are identified on the station usage by the flag: London Joint Station = 'Joint TfL & TOC Station'.

Within the MOIRA Replacement demand matrix, an estimate of the number of travellers using rail (as opposed to other modes i.e. tube or DLR) is made. Ticket sales at the joint stations are therefore scaled down in order to provide a better estimate of actual rail journeys.

In addition, there are a number of stations that MOIRA treats in the same way as Joint Stations. These are identified by the flag: London Joint Station = 'MOIRA Joint Station', e.g. Lewisham and Greenwich. These stations are not classified by the TOCs and TfL as a 'Joint Station', but they are the same: they are served by both rail and Underground/DLR services.

7.3 Other Impacts on Station Usage

7.3.1 Gating Schemes

Installation of ticket gates can significantly affect not only the usage figures at that station, but also those at neighbouring stations. The gates help to ensure that customers purchase tickets, but customers may also alter their travel patterns to avoid gated stations. We would expect travel patterns to be most affected in the months following the installation of the gates.

7.3.2 Change in Service Pattern

Alterations in service frequency or stopping pattern would be expected to alter station usage figures. This is particularly apparent where a group of consecutive stations show similar increases or decreases. Again, this can be a long-term trend.

7.3.3 Ticket Issuing Facilities Changes or Product Changes

Some London stations have both underground and National Rail trains operating. LENNON does not capture tickets sold by London Underground, only those sold by

TOCs. Changes in ticket facilities provided by TOCs, for example the provision of ticket machines, can therefore increase the ticket sales captured by the system.

Product changes can have an effect on passengers' purchasing patterns at rail outlets thus affecting station usage data. For example, the introduction of Oyster cards at rail outlets can affect stations inside the Travelcard boundary in the London area.

7.3.4 Engineering Work

Significant engineering work can alter customers' travel patterns.

7.3.5 Tourism

Stations near to tourist attractions may show significant changes in usage as a result of weather, promotions or other factors, which affect tourists' journeys.

7.3.6 Special Stations

Some stations serve a particular activity or business. Some fluctuation in usage of such stations is reasonable. Such activities include:

- Racecourses
- Sports Events
- Exhibition Centre Glasgow
- Airports

7.3.7 Trend of Growth or Decline

For stations with a history of growth or decline, it is reasonable to expect this trend to continue. There are many possible reasons for these trends, such as demographic and employment changes.

7.3.8 Changes in the Sales of Individual Ticket Types

Miscoding of ticket information entered into LENNON can alter station usage results, although this would not be reflecting an actual change in customers' journeys.

8 Regions, Counties and Districts

For all rail stations, the District, County, Region and NUTS2 Region & Code are provided for the origin and destination to describe the geographical location.

The source of this data is:

- District or the Unitary Authority – ATOC (dated January 2008) and ORR (dated January 2008)
- District, County & Region – ONS[4] website (dated January 2008)
- NUTS2 Code and Description – ORR (dated January 2010)

[4] http://www.statistics.gov.uk/geography/geographic_area_listings/administrative.asp#04

9 Limitations of the LENNON data

The LENNON database captures ticket sales for the entire national rail network from many different input machines. It is as a consequence a very large data set. With all large data sources there will always be input errors resulting in a certain amount of invalid data. Generally such errors will be small, and are more likely to occur in the journeys rather than revenue fields.

We perform checks on the data, but due to the size and complexity of the dataset we are not able to validate each and every entry.

We have used similar information extensively in the last ten years or more, and have found the data to be reliable, particularly when examining the data at an aggregated level.

There are a number of areas where we know that LENNON does not capture the data correctly, or instances where it is not possible to derive passenger journeys from ticket sales data. These areas are expanded upon below.

9.1 Known Problems of Data Capture

The data in LENNON from which the O-D Matrix is derived is based on ticket transactions. In order for the data to be included in the O-D Matrix it must include an origin station and a destination station. However if this is not the case then the data will automatically be excluded.

Human error at the point the ticket sale is entered into the input machines will produce invalid data in LENNON.

9.2 Travelcards

As Travelcards are for multi-modal travel they allow the purchaser to make journeys on the rail system and on other modes. Equally, tickets purchased elsewhere on the local transport system will be valid for rail travel. Therefore LENNON gives only a partial picture of the rail travel in conurbation areas, such as: London, Birmingham, Glasgow, Leeds, Liverpool, Manchester, Newcastle and Sheffield.

The O-D Matrix contains reasonably robust estimates of journeys within London and other conurbation areas where travelcards are widely used. An infill for London Travelcards has been included in the O-D Matrix since 2006/07, and an infill for PTE tickets is included from 2008/09.

Refer to Section 3 of the report: 'MOIRA Replacement, Stage: MS-5-1, Demand Matrix Documentation, Version: 0.2' dated August 2009, for a detailed description of the 'London Travelcard Area Infill'.

9.3 Return and Single Journey Tickets

It is possible that on certain routes the cost of a return ticket could be lower than a single ticket. This leads to the cheaper return ticket being purchased even though the passenger has no intention of making the return journey by rail. This results in two journeys being recorded instead of one.

9.4 Multiple Tickets

It is possible to buy special cheaper tickets between certain stations for example under a promotion by one of the train companies. In these cases a local ticket may be bought to gain access to a main station and a second ticket bought for the rest of the journey. This results in two journeys being recorded in the O-D Matrix and will not accurately represent the journey undertaken.

9.5 Rail Staff Passes

Prior to the privatisation of the rail network, British Rail employees and their families were eligible to various levels of free or reduced rate rail travel. When the various rail companies were converted to private companies, this benefit often continued.

If you consider the network as a whole, the effect of staff passes is unlikely to be significant. However, it may be significant on certain routes, for example on routes out of Derby due to large concentration of companies in Derby relating to British Rail both pre and post privatisation.

9.6 Ticketless Travel

On every route on the network there will always be passengers who travel without purchasing a ticket. This is referred to as ticketless travel. As LENNON data is derived from ticket transactions it cannot reflect this travel.

9.7 Other Rail Systems

There are a number of rail systems in operation in the country that are not covered by LENNON. For Heathrow Express and Eurostar revenue and journeys data were not available.

9.8 Factors Affecting the Data

Although not relevant for this year, there are factors worth taking into account when considering generic annual data:

- Years may have been affected by industrial action such as 1994/95.
- Major incidents affecting services such as Southall, Ladbroke Grove and Hatfield.
- Adverse weather.
- Infrastructure changes e.g. ticket gating significantly increases revenue - more gates have been installed in recent years which will affect the data but which does not represent higher passenger numbers.

9.9 Journey Factors

Ticket transactions are converted into an estimate of the number of journeys made by applying a series of ticket type journey factors. Single and return tickets unambiguously translate into one and two journeys respectively, for season tickets, the factors used represent a rough historic estimate as set out in Table 9-1 overleaf.

Ticket periods of other lengths are converted to a number of journeys using a proportion of the monthly journey factor.

Therefore the journeys data in the O-D Matrix represents an assumed number of journeys made based on the ticket type sold and the above journey factors. In particular it should be noted that the journeys data has not been cross-checked against other data sources of the actual number of journeys made on the network.

These journey factors have been used within the LENNON system for a number of years at their current values. The source of the factors is unclear, and there is some indication that they were based on reasonable estimates of ticket use made in excess of fifteen years ago. It can therefore be argued that these journey factors do not provide an accurate estimate of the number of journeys that result on the rail system at present, or in any O-D Matrix.

Table 9-1: Journey Factors used in LENNON

Description	Journeys Per Issue
Single Journey Ticket	1
Return Journey Ticket	2
Return Journey 2 Persons	4
3 Day Return/ 6 Single Journeys	6
4 Day Return/ 8 Single Journeys	8
5 Day Return/ 10 Single Journeys	10
6 Day Return	12
5 Day Single	5
1.5 Journeys	1.5
Weekly Ticket	10.3
10 Day Return/ 20 Single Journeys	20
2 Weekly Ticket	22
Seasons-Variable Periods	***
Monthly Ticket	45
Not Used	0
3 Monthly Tickets	135
Not Used	0
6 Monthly Tickets	270
Summary Group Codes	***
Annual Ticket	480
8 Day Ticket	22
22 Day Ticket	44
14 Day Ticket	30
50 Journeys	50
10 Weeks	103

Appendix 1 – Overview of the ORCATS Allocation Process

This section gives an outline of the Central Allocations File (CAF), which is used in producing the interchange figures, and the ORCATS process which is used to create the CAF.

Most of the train tickets that are sold are inter-available – the customer has a choice of routes and operators. For example, when a customer buys a ticket to travel from Leicester to Leeds, that customer may travel on various combinations of East Midlands Trains, East Coast, CrossCountry Trains and Northern, and may interchange at Doncaster, Sheffield, Derby or Nottingham. LENNON captures the sale of the ticket, but unless the ticket has stringent route restrictions, the route actually taken by the customer is not recorded.

The route taken by any particular customer may never be known, but some route options are more attractive than others. The customer is more likely to choose a faster, more frequent service than a slower, less frequent one. This likelihood can be translated into the proportions of customers choosing each route option, on a particular flow. (A 'flow' represents all journeys from a given origin station to a given destination station, irrespective of the route taken.) The revenue received from all customers on that flow should be split between different operators to reflect the proportion of customers which each operator carried.

ORCATS was developed to model the choice made by the customers, and to allow revenue to be split between operators. It applies passenger choice modelling to the train timetable, to determine the relative attractiveness of different route alternatives. It then weights the results by journey mileage.

For any given timetable, ORCATS works out the possible routes between each origin and destination, and calculates the percentage of the passengers that are expected to choose each route based on the services in that timetable.

The output from ORCATS is the Central Allocations File (CAF). This lists the proportion of journeys on each flow (or origin-destination pair) estimated to be made by each route alternative. For journeys involving interchanges, each leg of the journey is listed. By combining this information with LENNON data, which contains actual ticket sales figures for all flows, the number of interchanges occurring at individual stations has been estimated.

Appendix 2 – Methodology: Non-Station Tickets

Ticket sales do not always tell us where a passenger is travelling. Ticket sales can be divided into the seven categories listed in table below. Ticket sales data has been converted into an estimate of the actual stations that passengers are travelling from/to.

The processing of ticket sales data is undertaken in the creation of the MOIRA Replacement demand matrix, and then subsequently in the creation of the O-D Matrix. For each of the flow categories, the table below states where the flow is processed: MOIRA or O-D Matrix.

Table A2-1: Categorisation of Ticket Sales in LENNON

Flow Category	Description	Processing
Category 1	Origin and Destination Stations Known	No processing required
Category 2	Origin or Destination a Group Station (excl. London BR)	O-D Matrix
Category 3	Origin or Destination is London Terminals	O-D Matrix
Category 4	Origin or Destination a London Travelcard including Zone 1	O-D Matrix
Category 5	Origin or Destination a London Travelcard excluding Zone 1	MOIRA Demand Matrix
Category 6	Origin or Destination a London Travelcard Boundary Zone	MOIRA Demand Matrix
Category 7	Non-National Rail Stations	MOIRA Demand Matrix

Category 1 – Origin and Destination Stations Known

Both the origin and destination were known stations so no further processing is required for such flows.

Category 2a – Origin or Destination a Group with all Stations Having a Ticket Office

In 2005/06 all origins or destinations that were a group station (with the exception of London BR) were changed to the major station within the group. For example, all ticket sales to or from Reading BR were recoded to Reading. This was clearly over-simplistic.

In 2006/07 the O-D matrix was based on the journeys from ticket sales to the individual stations within a group. We assumed that passengers travelling **to** the stations in a group would act in the same way as passengers travelling **from** the stations in that group. We believed that this was, in general, a valid assumption to make, and no bias would be introduced into the journey figures.

From 2007/08 onwards this process is still used where all stations in the group have ticket offices, so that the relative flows from the individual stations are credible.

For example, in 2006/07 the journeys between stations in the 'Manchester BR' group and Crewe and vice-versa are shown by the column "jnys" in the table below. First the proportion of journeys **from** each of the individual Manchester stations **to** Crewe is determined, as shown in column "%split".

Then these proportions are applied to both the 'Manchester BR to Crewe' and 'Crewe to Manchester BR' flows, giving the breakdowns to individual stations shown in column 'BR portion'. These are added to the base values to give "Total Journeys", before the 'Manchester BR to Crewe' and 'Crewe to Manchester BR' flows are deleted, to avoid double counting. The slight discrepancy between the Grand Totals is due to rounding error.

Table 9-1: Example of how we split journeys to/from a BR group of stations

Orig	Dest	Origin Name	Destination Name	Jnys	%split	BR portion	Total Jnys
2963	1243	DEANSGATE	CREWE	83	**0.32%**	85	168
2966	1243	MANCH OXF RD	CREWE	5,464	**21.03%**	5580	11,044
2968	1243	MANCH PICC	CREWE	19,733	**75.95%**	20152	39,885
2970	1243	MANCH VICT	CREWE	700	**2.69%**	714	1,414
0438	1243	MANCH BR	CREWE	26,533	**Remove**		
1243	2963	CREWE	DEANSGATE	207		1478	1,685
1243	2966	CREWE	MANCH OXF RD	2,262		97287	99,549
1243	2968	CREWE	MANCH PICC	8,017		351349	359,366
1243	2970	CREWE	MANCH VICT	343		12464	12,807
1243	0438	CREWE	MANCH BR	462,578	**Remove**		
			Grand Total:	525,920			525,918

The above methodology has been applied to all flows with more than 1,000 journeys in total leaving the individual group stations (i.e. not including the 'BR Group NLC to

destination' flow. For the smaller flows an average split is applied based on the flows with more than 1,000 journeys.

Category 2b – Origin or Destination a Group with some Stations Having no Ticket Office

For this class of stations the above process breaks down because the proportion of journeys **to** the group stations with no ticket offices will tend to be estimated as zero because the sales **from** those stations are necessarily zero. For these groups an alternative process is used which considers each origin station / group destination pair in turn and estimates the proportion of flow to each group station from the origin according to the existence or relative attractiveness of the service to each of the group stations.

Category 3 – Origin or Destination is London BR

This category contained all flows that had London BR as either the origin or destination. In order to assign an appropriate London station on flows where either the origin or destination is London BR (NLC=1072) or a London Travelcard involving Zone 1, we analysed responses from the 2001 London Area Travel Survey (LATS). For journeys from any given station, we established the percentage of passengers using each London terminus.

For example, if the flow was from Ashford International to London BR, we used our pre-generated table showing the percentage spilt between the alternative London termini for passengers starting at Ashford International. From this we apportioned the exits between London Bridge, Charing Cross, Victoria and other London termini.

Stations with small sample sizes were removed from the 2001 LATS data. Where there was insufficient data in the 2001 LATS to generate the split for a particular station, a similar process with the Non London Groups methodology was applied. Firstly for all the flows with more than 1000 journeys leaving London BR and having as a destination the particular station we used split factors as above. However, if the sum of journeys was less than 1000 we assigned to the flow the top origin from the London BR stations.

Category 4 – Origin or Destination a London Travelcard including Zone 1

All origins and destinations that were London Travelcard Zones that include Zone 1 were converted to 'London BR' under the assumption that they will travel to the same stations as point-to-point passengers and then transfer to another mode. The methodology set out above for Category 3 was then applied.

Category 5 – Origin or Destination a London Travelcard excluding Zone 1

This category contained all Travelcards that did not include Zone 1, for example Zone R2345 London.

For flows with origin or destination a London Travelcard (excluding zone 1) we use a set of assumptions based on survey responses from the 2001 LATS. They use the starting station to work out which stations it is possible for the passenger to be travelling to, and also give the proportion of passengers travelling to each of these stations. This is based on the assumption that a passenger holding a Zones 2-6 Travelcard would travel as far as Zone 2.

This processing is undertaken during the production of the MOIRA Replacement demand matrix.

Category 6 – Origin or Destination a Boundary Zone

All origins and destinations that were a London Travelcard Boundary Zone were converted to 'London Travelcard including Zone 1' under the assumption that a passenger travelling from or to a Boundary Zone will hold a Travelcard that includes Zone 1. The methodology set out above for Category 3 was then applied.

This processing is undertaken during the production of the MOIRA Replacement demand matrix.

Category 7 – Non-National Rail Stations

This final category contains all those flows in the original ticket sales data that do not fall into one of the above categories. Refer to Section 6: 'Data Excluded From the O-D Matrix' for a detailed description of this data and what has been included and excluded from the O-D Matrix.

This processing is undertaken during the production of the MOIRA Replacement demand matrix.

Appendix 3 – Data Excluded From Station Usage

Some of the LENNON data has been excluded from the MOIRA Replacement Demand Matrix, and subsequently from the O-D Matrix.

All the products that were classified into the 'miscellaneous' ticket pot were excluded. These products were:

- Car Parking
- Railcard Sales
- Penalty/Excess Fares
- Seat Reservations
- Sleeper Supplements.

Also excluded from the analysis were all the flows that had either an Origin or Destination that did not represent a geographical location (these are mainly "I codes"), e.g.

- Rover and Ranger Tickets (e.g. Anglia Plus)
- BritRail Tickets
- Gate passes usually used by staff
- Passenger Charter Discounts
- Headquarters Input Items, other than those which can be identified as TfL or PTE

Finally for flows that have either Origin or Destination a Private Settlement Code some are included and some are excluded.

- PTE tickets and TfL sold London Travelcard records from Lennon are removed, and replaced with an estimate of all rail travel using these tickets via 'infill's to the MOIRA demand matrix (refer to Section 2).
- PlusBus – all significant flows have been included since 2007/08 (refer to Section 2), and minor flows are excluded.
- Attractions – the rail element of the significant flows have been included since 2007/08, which include:
    - Bluewater Shopping Centre
    - Alton Towers
    - Whipsnade
    - Chatsworth House

All other flows involving Private Settlement are excluded, e.g. Irish Stations.