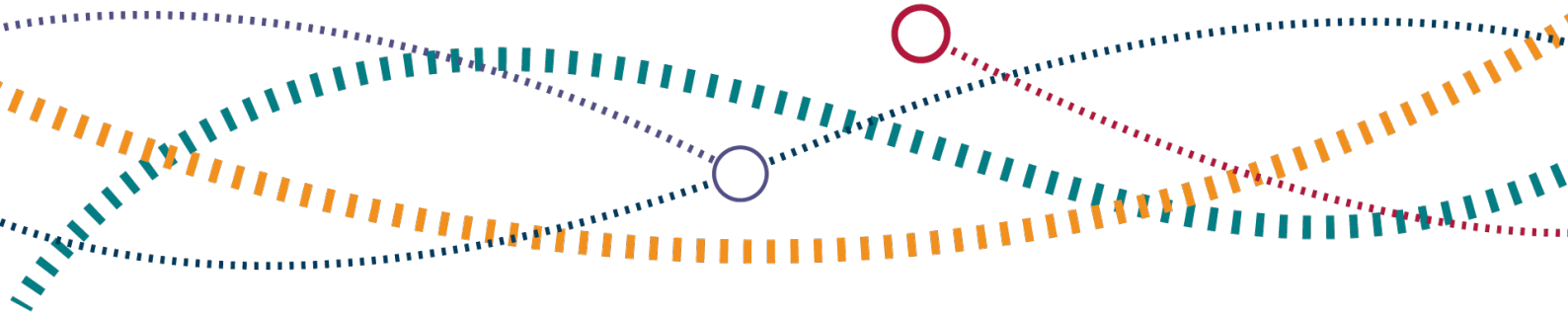
Estimates of station usage

Quality and methodology report

21 November 2024

Contact information:
Responsible statistician: D. Leese
Public enquiries: rail.stats@orr.gov.uk
Media Enquiries: Tel: 07856 279808



Contents

Introduction

This is a report on the quality and methodology used to produce the annual Estimates of station usage. It helps users to understand the quality of our statistics, and also ensures ORR is compliant with the three quality principles in the Code of Practice for Official Statistics - Q1: Suitable data sources, Q2: Sound methods, and Q3: Assured quality. This report also provides information on the methodology and data sources used to produce the statistics.

This report covers the following areas:

- Data sources, methodology and definitions – detail on the various data sources, methodology used to compile the statistics and definitions;
- Historic background – background to the time series (summary of methodological improvements contained in annex 1);
- National Statistics accreditation – the assessment of these statistics by the Office for Statistics Regulation;
- Relevance to users – the users of the statistics, and our engagement;
- Accuracy and reliability – the accuracy, data coverage and quality assurance of the statistics;
- Timeliness and punctuality – our timescales for the production and publication of the statistics;
- Accessibility and clarity – the format of our statistics and where they can be found;
- Coherence and comparability – similar statistics published elsewhere and the degree in which the statistics can be compared over time;
- Annex 1 – Details of methodology changes over time.

ORR commission Steer to produce the Estimates of station usage dataset covering the latest financial year (1 April to 31 March). More detailed information on methodology and changes each year are available in two reports produced by Steer: *Estimates of Station Usage 2023/24: Methodology Report* and *Station Usage & Origin Destination Matrix 2023/24: Historical Methodological Changes*.

Data sources, methodology and definitions

The Estimates of station usage dataset consists of estimates of the total numbers of people:

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

The estimates of entries and exits are further split by ticket type (full price, reduced price and season tickets). There is also a range of station attribute information included too, e.g. geographic data.

Time series of entries and exits and interchanges by station are available from April 1997.

Data sources

Estimates of station usage are primarily based on sales data from LENNON, the rail industry's ticketing and revenue system. This is supplemented with some local ticketing data.

Below is a list of all the data sources used to create the Estimates of station usage:

Entries and exits:

- LENNON, Transport for London (TfL) data and train operator data (Gatwick Express and Stansted Express) as an input to the MOIRA2.2 base matrix
- Local ticketing data from Passenger Transport Executives (PTEs)
- Manual station counts
- Heathrow Express ticketing data

Interchanges:

- Central Allocations File (CAF)

Station attribute data:

- Geographic data – Office for National Statistics geoportal

- National Rail Enquires
- Rail Delivery Group (RDG) fares database
- Transport for London

Methodology

These statistics on station usage are **estimates** primarily based on tickets sales, sourced from LENNON, the rail industry's ticketing and revenue system and local ticketing data. These data sources and the methodology used provide the best approach possible given Great Britain does not have a fully gated rail network or robust count data for every station. However, this data does have weaknesses when utilised for this purpose and, although some of these are catered for in the methodology and we continue to seek improvements to address identified issues, the user should be aware of these acknowledged limitations and bear these in mind when using the data. The key **limitations** are detailed in the '*Accuracy and reliability'* section of this report.

LENNON contains the majority of National Rail tickets purchased in Great Britain. However, it excludes some tickets sales e.g. London Travelcards. These ticket sales, together with LENNON data are used to derive a base matrix of journeys and revenue which is an input to the MOIRA2.2 rail planning tool. Steer take this MOIRA2.2 base matrix and supplement it with updated local ticketing data for Passenger Transport Executive (PTE) areas. Various adjustments are made to the data to deal with a range of issues (see below) to create a comprehensive matrix of passenger flows throughout Great Britain, the Origin Destination Matrix (ODM). The ODM is used to derive the number of entries and exits at each station in Great Britain.

A column ("Data source/adjustments") is included in the Estimates of station usage dataset to summarise the basis of each station's estimate, e.g. if the data is supplemented by PTE data or an adjustment has been made to improve the estimate. Data source and adjustments are only flagged within this field where they change the usage estimate for an individual station by 0.5% or greater. These adjustments are described below with the abbreviation used in the dataset shown in brackets:

Infills and adjustments:

- PTE infills ("**PTE**") – local ticketing data provided by PTEs and prepared by Steer (for West Midlands) and Mott MacDonald (Greater Manchester, Merseyside, South Yorkshire, Tyne and Wear, West Yorkshire, Strathclyde). These infills are subject to annual improvements, which normally represent a simple update, but some years contain a step change in the methodology;

- Ranger/Rover infills ("**Other**") – Rover and Ranger products, e.g. Ride Cornwall, Cambrian Coaster, Anglia Plus, are tickets with non-geographic destinations so estimates are developed to represent passenger flows on a selected number of these tickets;

- London adjustment ("**London**") – Allocation of demand associated with tickets sold to 'London BR' between the London terminals. This is for tickets that have an origin and destination on the ticket of 'London Terminals', which is then reallocated between the component stations based on known sales between specified origins/destinations from LENNON;

- Allocation of demand between individual stations within station groups ("**Group**") – Many products are sold with the origin or destination as a group of stations, e.g. London BR, Birmingham BR. Current industry data does not distinguish between the component stations and therefore a split between these stations must be estimated. For example, where tickets are sold to/from 'Birmingham BR' it is necessary to estimate how these journeys are distributed between New Street, Snow Hill and Moor Street stations. For London BR this is done using the Central Allocations File to model assumed routes a passenger would take. For other group stations, this is done by apportioning to individual stations based on known ticket sales between specified origins/destinations from LENNON. A consequence of this adjustment is that all stations in a group will have the same annual percentage change, unless the station split assumptions change in a given year;

- Manual counts used to allocate demand between individual stations within 21 station groups outside central London ("**Counts**") – Steer undertake a programme of manual counts on an annual basis at selected group stations to calculate the splits where the above 'Group' adjustment isn't representative. For example, where tickets are sold to/from 'Dorking BR' it is necessary to estimate how these journeys are distributed between Dorking West, Dorking and Dorking Deepdene stations. A consequence of this adjustment is that all stations in a group will have the same annual percentage change, unless an updated count adjustment is applied in a given year. Since April 2012 progressively more station groups have a count based methodology for apportioning total demand amongst its member stations;

- Season ticket adjustment ("**Season**") – There are a number of cases where adjustments are made to selected stations to account for specific known issues. For example, adjustments at a number of stations are made to reflect circumstances where there are significant numbers of season tickets sold at a

particular station (where the passenger travels from) for travel to London that allow for travel to/from a different origin station to provide flexibility. This leads to a situation where station usage, as estimated by using the origin and destination on the ticket can be under- or over-estimated and journeys involving those stations needs to be adjusted to reflect actual usage. This adjustment also applies to stations around Cardiff;

- The 'Digby and Sowton' adjustment ("**D&S**") – relates to journeys associated with a season ticket product for students which are being made to Exeter Central and Exeter St. David's on tickets with a recorded destination of Digby and Sowton. Adjustment uses data provided by Great Western Railway;

- Caledonian Sleeper ("**CS**") adjustment – relates to journeys on Caledonian Sleeper services that are not included in the MOIRA2.2 base matrix. To compensate for this an adjustment is made to add these journeys using additional analysis of LENNON ticketing data.

- Heathrow adjustment ("**HEx**") – not all ticket sales to/from Heathrow stations are included in LENNON, i.e. some Heathrow Express tickets. Therefore an adjustment is made to add in the non-LENNON ticketing data supplied by Heathrow Express. In addition, gateline data is used to apportion usage between the three Heathrow stations.

- Elizabeth line ("**EL**") overstatement adjustment – Since the opening of the central section of the Elizabth line, there has been a known issue with LENNON overestimating contactless and Oyster Pay As You Go (PAYG) journeys on the Elizabeth line. To correct for this we have used data supplied by Transport for London (TfL) as a direct replacement for these ticket types in the LENNON data.

- Same origin and destination ("**Same OD**") adjustment – This is used to account for journeys in the MOIRA2.2 base matrix that have the same origin and destination. This issue mainly occurred in the Oyster/CPAY area where passengers either did not tap in or tap out. The adjustment reallocates these journeys to other flows based on the underlying distribution of Oyster/CPAY journeys at the station

- AutoTOD ("**TOD**") adjustment – This is used to exclude journeys that were paid for but never used. These are tickets that were not collected from 'Ticket on Departure' (TOD) machines.

Further details relating to the infills and adjustments outlined above can be found in Chapter 3 of Steer's methodology report.

Methodology changes

The methodology to produce the Estimates of station usage is reviewed annually and enhancements are implemented to address known issues. Often these enhancements utilise new sources of data that were not previously available and improve the estimates.

Consistency with past datasets is important to enable comparisons to be made over time. Nonetheless, stakeholders have indicated that they are keen to see improvements, even where this reduces consistency with historic data, provided any changes are clearly explained. ORR has worked with Steer to scope and implement methodological enhancements to address identified issues and utilise new data as it is made available whether this is from primary data collection (e.g. passenger counts at stations), or industry systems such as TfL's Oyster Clicks Model (OCM).

A number of improvements to the methodology have been implemented over recent years (see Annex 1). These changes should be taken into account when considering year-on-year changes in journeys for some stations as it may not reflect an actual change in demand. A column ("Quality limitations") is included in the Estimates of station usage dataset to highlight if data are not comparable to the previous year, and/or if there are any other limitations or quality considerations.

In the April 2022 to March 2023 dataset the following methodological improvements have been implemented:

- An adjustment to account for split ticketing, where a passenger completes one journey using two or more tickets.
- A change in how entries and exits are allocated to stations for tickets to London Terminals, using a more up to date data source.
- Updates to the distribution of entries and exits across West Yorkshire Metro stations.
- Updates to the distribution of entries and exits across Heathrow stations.
- Inclusion of a number of ticket types in the 'Other' infill, which had not previously been included.

Interchanges

An estimate of the number of people interchanging at each station is included in the Estimates of station usage dataset. This is obtained by combining the number of journeys made on each flow (from the ODM) 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 2023 CAF is used in the creation of the April 2022 to March 2023 Estimates of station usage dataset.

An overview of the ORCATS allocation process can be found in in Appendix B of Steer's methodology report.

Definitions

- LENNON – 'Latest Earnings Networked Nationally Over Night' is the rail industry's ticketing and revenue system. It contains information on the majority of national rail tickets purchased in Great Britain. However, it excludes some tickets sales.

- MOIRA2.2 base matrix – produced by Resonate as an input into the MOIRA2.2 rail panning tool, it provides an estimate of journeys on the Great Britain rail network for the duration of a financial year. It includes all journeys associated with point to point flows and includes overlays ("infills") to reflect travel using tickets not included in LENNON (e.g. London Travelcards and some specific tickets to/from airports and multi-modal and zonal products sponsored by PTEs).

- Origin Destination Matrix (ODM) - a comprehensive matrix of passenger flows throughout Great Britain.

- Passenger Transport Executive (PTE) - There are six metropolitan counties in England. These are Greater Manchester, Merseyside, South Yorkshire, Tyne and Wear, West Midlands, West Yorkshire. Formerly, each of these areas had a Passenger Transport Executive (PTE), which was a local government body with public transport responsibilities. They were accountable to Integrated Transport Authorities (ITAs), which have now been reformed into Combined Authorities, some with a larger geographic coverage than the ITA they replace. Some

Combined Authorities (Greater Manchester, Merseyside, North East, South Yorkshire) continue to have a free-standing transport executive, whilst in others (West Midlands and West Yorkshire) the transport executive has been incorporated within the Combined Authority. In Scotland the Strathclyde Partnership for Transport is the equivalent body covering the region of Strathclyde. For convenience, in this report we continue to refer to these seven areas as PTEs.

- Ticket Types:
  - Full: all walk-up **undiscounted** single or return tickets, whether or not issued with a status discount (child, railcard, etc);
  - Reduced: all walk-up **discounted** single or return tickets, whether or not issued with a status discount (child, railcard, etc). All advance-purchase tickets are also included in the category;
  - Seasons: all multi-use tickets.

For more detailed information on the methodology see *Steer's Estimates of Station Usage 2023/24: Methodology Report*.

Historical background

April 1997 to March 1998 to April 2002 to March 2003:

Estimates of station usage were calculated from CAPRI (Computer Analysis of Passenger Revenue Information) which was the rail industry's former central ticketing system.

April 2003 to March 2004:

No estimates produced.

April 2004 to March 2005 to April 2022 to March 2023:

From April 2004, LENNON (Latest Earnings Networked Nationally Over Night), which is currently the rail industry's central ticketing and revenue system, has been the basis for calculating these statistics.

Several improvements to the methodology have been implemented over the years. A summary of methodology improvements since April 2006 can be found in Annex 1. Also see Steer's historical methodological changes report.

Steer has been contracted by ORR to produce these statistics annually since April 2011. Prior to that (since April 2005) DeltaRail (now known as Resonate) were the contractor.

In December 2020 the statistics received National Statistics accreditation.

National Statistics accreditation

During April to November 2020 Estimates of station usage were assessed by the Office for Statistics Regulation (OSR) against the Code of Practice for Statistics for National Statistics designation. National Statistics status means that official statistics meet the highest standards of trustworthiness, quality and public value.

OSR identified several ways in which we should develop the statistics in order to achieve National Statistics status. These are described in the assessment report published on 3 November. All the requirements were addressed for the April 2019 to March 2020 publication and OSR confirmed in a letter on 1 December that these statistics have been designed at National Statistics.

In their report OSR said: *Estimates of station usage provide a valuable, trusted source of information about the use of the rail network across Great Britain. They are used by a broad range of people, who appreciate the granularity of the data and the ability to use these statistics for further analysis.*

As part of their assessment, OSR sought and received feedback from a wide range of users of our statistics[1]. This feedback informed their requirements and was considered by us in developing an improvement plan for estimates of station usage statistics. All the actions in the plan were successfully implemented for the April 2019 to March 2020 publication. For example, replacing the factsheet with a more informative statistical release; a new quality and methodology report; inclusion of two new columns of information in the data table, the first one highlighting any data quality issues or methodology changes, the second listing the data sources/adjustments used to estimate usage at each station.

[1] Feedback was received from the following users: DfT, Welsh Government, RDG, Network Rail, RSSB, Transport Focus, London TravelWatch, Transport for Wales, TfL, Transport For West Midlands, Cambridgeshire County Council, BBC, All the Stations, Community Rail Network, TSUG, academics.

Relevance to users

The degree to which the statistical product meets the user needs in both coverage and content.

Uses and Users

Estimates of station usage are the only source of published information on the use of all stations in Great Britain. They are used by a broad range of people, e.g. by government departments, local transport authorities and community groups, and transport watchdogs. The statistics are also of interest to journalists, academic researchers and members of the public.

One of the main strengths of Estimates of station usage is that they provide an annual data series going back to April 1997. This means that users can explore trends over time and combine the data with their own local knowledge to understand the impact of infrastructure projects or changes to the usage of the rail network.

User engagement during the assessment by the Office for Statistics Regulation in 2020 identified the following uses of Estimates of station usage:

- Monitoring station use to understand capacity and identify potential issues
- Informing business cases for station or service developments
- Planning, monitoring and evaluating infrastructure projects
- Used in the production of other official statistics, such as those from the National Rail Passenger Survey
- National and local media outputs that inform public interest
- To facilitate academic research
- Personal interest

How these statistics can and cannot be used?

- Monitoring the number of annual entries and exits or interchanges at individual stations e.g. to understand demand
- Monitoring how usage at individual stations changes over time (subject to methodology changes) and insights as to why
- Comparing the relative usage of stations within local areas, regions or across the whole of Great Britain
- To gauge the use of different ticket types e.g. season vs reduced

- Monitoring passenger rail usage at a national level, by train operating company or by ticket type (refer to Passenger rail usage statistics)
- Monitoring the number of passenger journeys between and within regions (refer to Regional rail usage statistics)
- Volume of entries compared to exits at an individual station (methodology makes these equal)

When using Estimates of station usage, it is important to be aware of:

- Methodological improvements made to the dataset over time which can impact consistency between years;
- Limitations of the data and specifically factors, e.g. some ticket sales not being included, that may mean that demand on particular flows and at stations is underestimated or overestimated; and
- Factors which can affect reporting of entries and exits, e.g. infrastructure improvements; temporary line or station closures.

User satisfaction

A number of changes and improvements were made to the April 2019 to March 2020 statistics following feedback from OSR and users. These are summarised in *the 'National Statistics accreditation'* section above.

ORR's last user survey took place from mid-January to mid-April 2020. The aim of the survey was to gather feedback on ORR's new data portal; this includes statistical releases, data tables and other supplementary material. There were 42 responses to the survey. ORR created an implementation plan following the 2020 user survey.

More detailed information on users of ORR statistics and meeting the needs of users is available on our user engagement webpage.

Accuracy and reliability

The proximity between an estimate and the unknown true value.

These statistics on usage are **estimates** based primarily on tickets sales using the methodology described above. This methodology is the best approach possible given Britain does not have a fully gated rail network or comprehensive and robust count data at every station. The methodology's national coverage makes it suitable as a basis for the production of these National Statistics accredited statistics.

Data coverage

The data presented in this release are for all open mainline stations in Great Britain, i.e. those with a timetabled train service.

Limitations

A number of improvements to the methodology have been implemented in recent years. These changes should be taken into account when considering changes in usage between years, as it may be a result of improved methodology, rather than reflecting an actual change in demand at stations. These improvements, and the reasons for them are documented in Steer's historical methodological changes report.

As these statistics are primarily based on ticket sales, there are a number of limitations that users should be aware of which are summarised below. There is more detail on these in Chapter 5 of Steer's methodology report.

Stations with known under-estimates or no estimates - Eurostar ticket sales are not in the rail industry's ticketing system (LENNON). Therefore our estimated usage at St. Pancras, Ashford International and Ebbsfleet stations will not be a true reflection of the total usage at these stations. There are two other stations which we do not have estimates for as tickets sales are not recorded in LENNON: Bishops Lydeard and Corfe Castle.

Concessionary travel - TfL and most PTEs subsidise some form of free travel for certain types of users including those over a certain age, students and those with disabilities. This creates a substantial additional element of demand which is very difficult to include in these estimates as information on the level and distribution of journeys associated with these free travel products is not recorded and will not even have point of sale information. The current approach to this in the ODM is to include this demand where data has been made available by TfL/PTEs which would generally be estimates based on surveys.

Currently concessionary travel data are included for London, Greater Manchester, Merseyside, South Yorkshire, Tyne and Wear, West Midlands and West Yorkshire areas. No information is available for Strathclyde, therefore there will be some under-estimates of usage at some stations in these areas. Any concessionary travel for under 5's is not included in Estimates of station usage statistics.

Ticketless travel - As the estimates of station usage are based on ticket sales, journeys associated with ticketless travel are not included in the data. This is more likely to be an issue on some flows and where ticketless travel is significant. It should also be noted that levels of ticketless travel may have changed during the pandemic and those changes may vary substantially by station. As more stations have been gated over time and train operators focus on revenue protection activities, this is likely to be less of an issue than in the past. It can be argued that it is not appropriate to include ticketless travel in the dataset as its purpose is to record genuine journeys on the rail network. The inclusion of ticketless travel could distort business cases for new investment where these are reliant on the estimates of station usage data. Ticketless travel also includes an element of individuals who are legitimately travelling for free, such as the British Transport Police or some rail industry employees.

Group stations - Many products to major destinations are sold with the origin or destination as a group of stations (e.g. London BR Terminals, Manchester BR). Current industry data does not distinguish between the component stations and therefore a split between these stations has to be estimated. These estimates are currently apportioned to individual stations based on a combination of known sales between specified origins/destinations, survey and/or count data.

Season ticket fare zones - In some areas, multiple stations have identically priced season tickets to London. As a result, London season tickets are generally sold as being from the furthest station, regardless of the actual origin of travel, giving the passenger additional flexibility for no increase in fare. This means that the ticket sales data shows that there are more people travelling to/from this station than is the case. Therefore for stations where this is a known issue the data is adjusted to reflect actual usage at each station, for example, stations along the Southend Victoria and Central branches. There may be similar issues at stations where we don't make adjustments to the data.

Season ticket journey factors – Ticket sales in LENNON are converted into an estimate of the number of journeys made by applying a series of ticket type journey factors. Therefore station usage estimates are based on an assumed number of journeys made based on the ticket type sold. The journeys factors used for the main season tickets are as follows:

Season ticket validity	Journey factor
Weekly	10.3
Monthly	45
3 monthly	135
6 monthly	270
Annual	480

Interchanges on the central section of the Elizabeth line - For April 2023 to March 2024, we have not published estimates for interchanges taking place at London Liverpool Street and London Paddington. This is because we know that interchanges taking place to or from Elizabeth line services at these stations have not been included in the December 2023 CAF file used to derive our interchange estimates. We would expect the magnitude of each of these interchanges to be large, but we are unable to confidently estimate this. It is possible that interchange estimates at other stations on the Elizabeth line route have been overestimated as they include some modelled interchanges that are likely to have actually taken place at London Liverpool Street or London Paddington.

Quality assurance

ORR has a quality assurance framework for Estimates of station usage. This sets out the steps and methods used to quality assure these statistics throughout the production process by both Steer and ORR. Also included is a risk log detailing any high risk points around data quality during the process and how these are mitigated.

A summary of the quality assurance framework is below:

- Steer produce draft station usage estimates following a prescribed set of processing steps. Quality assurance checks are undertaken at each processing step including various comparisons of the draft estimates with other data sources for validation purposes. These comparisons are shared with ORR and subsequently published in Steer's methodology report. Once a draft dataset is ready Steer and ORR meet to discuss the draft estimates and any emerging issues are investigated and resolved before the dataset is formally passed to the lead statistician in ORR.

- ORR then share the draft estimates with the stakeholders who had provided supplementary ticketing and usage data, e.g. the PTEs. These recipients were asked to review the estimates for the stations in their area, provide reasons for large changes in usage at individual stations or to flag any estimates where they had quality concerns. Following the external review outlined above, estimates are reviewed internally by analysts and experts within ORR. Colleagues are asked to use their own knowledge or investigate reasons for large changes in usage and to flag any estimates where they had quality concerns. Any estimates where the internal or external reviews raise quality concerns are shared with Steer and undergo further investigation.

- The final data are then prepared for publication by the lead statistician. The process includes quality assuring the tables and charts produced and providing supporting commentary regarding the key trends. These are subject to peer review by another analyst following a well documented process which is followed for all ORR statistical releases.

- The final stage of the quality assurance process is a sign off by the Head of Profession for Statistics confirming the statistical release and associated outputs, e.g. data tables, meet quality standards and are fit for publication.

Revisions policy

ORR's statement on orderly release and revisions policy outlines ORR's revision policy. Details of any revisions are available in the revisions log. Further information on revisions and data series breaks can also be found in the data tables.

Timeliness and punctuality

Timeliness refers to the time gap between publication and the reference period. Punctuality refers to the gap between planned and actual publication dates.

ORR aims to publish the Estimates of station usage as soon as possible after the end of the financial year, which is currently around eight months after in November. It is important to take time to implement the detailed methodology, collect additional data (e.g. from PTEs) and carry out a series of quality assurance checks to ensure the final data are as accurate as possible.

ORR will continue to work with our consultants and stakeholders to shorten the time between the reference period and publication.

The publication schedule available on the data portal outlines the publication dates for National Statistics quarterly and annual statistical releases and other official statistics up to 12 months in advance.

ORR is committed to releasing its statistics in an open and transparent manner that promotes confidence.

Accessibility and clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

Statistics need to be presented in a clear and understandable form. All our rail statistics data tables can be accessed free of charge on our data portal. Commentary about the statistics and trends are provided in the statistical releases. Interactive dashboards (PowerBI) are also available.

Our data portal and its content meet the accessibility standards of the Public Sector Bodies Website Accessibility Regulations. We support our users by providing the information they need in a way that is clear and accessible. Our statistical releases use plain language, and any technical terms, acronyms and definitions are clearly defined and explained when this is appropriate, to ensure that the statistics can be used effectively. Our data tables are available at the highest level of detail that is practical and in accessible formats. All data tables are available in OpenDocument Spreadsheet (.ods) format. We can also provide data in csv format on request.

Please see our accessibility statement for further details, including any non-accessible content.

Data tables and other outputs

All tables and other outputs associated with this release can be found on the Estimates of station usage theme page.

- Estimates of station usage statistical release (1 April 2023 to 31 March 2024)
- Passenger entries and exits and interchanges by station (1 April 2023 to 31 March 2024) – Table 1410 (ods and csv)
- Time series of passenger entries and exits and interchanges by station (April 1997 to March 1998 to April 2023 to March 2024) – Table 1415 (ods)
- Interactive dashboard (Power BI)
- Animated graphics (MP4) and infographics (PDF)
- Frequently Asked Questions (FAQs) document (PDF)
- Quality and methodology report (PDF)
- Steer's methodology report (PDF)

- Steer's historical methodological changes report (PDF)

Coherence and comparability

Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain.

Related statistics

Passenger rail usage (ORR):

Quarterly statistics reporting the volume of passenger journeys, kilometres and revenue on the mainline network in Great Britain. Statistics are presented by ticket type, sector, and train operator. Long-running time series on passenger journeys (Table 1220) and passenger kilometres (Table 1230) are updated annually.

Origin and destination Matrix (ORR):

The ODM contains the estimated number of journeys between each pair of mainline stations in Great Britain. Annual datasets going back to April 2018 to March 2019 are available on the Rail Data Marketplace.

Regional Rail Usage (ORR):

Annual statistics providing passenger journeys data for each region of Great Britain, including journeys between regions, within regions and between England, Scotland and Wales. These estimates are also produced by Steer based on the ODM.

Station footfall (Network Rail):

Quarterly data on station concourse footfall at 18 of Network Rail's managed stations only. These data are collected using a different method to Estimates of station usage and include all people using the stations e.g. visiting shops and restaurants who may not make a rail journey.

Passenger numbers and crowding (Department for Transport):

Rail passenger numbers and crowding statistics provides information on the number of passengers travelling by rail into and out of major city centres in England and Wales. The statistics represent passengers on National Rail services on a 'typical' weekday.

Transport use during the coronavirus (COVID-19) pandemic (Department for

Transport):

Daily data on the use of transport by mode in Great Britain presented as a percentage compared to a pre pandemic baseline.

Comparability

Consistency with past datasets is important to enable comparisons to be made over time. Nonetheless, stakeholders have indicated that they are keen to see improvements, even where this reduces consistency with historic data, provided any changes are clearly explained. See *Methodology* section above and Annex 1 for more information.

Length of comparable time series

Measures	Start of time series	Any break in time series
Estimates of entries, exits and interchanges	April 1997	April 2003 to March 2004 (no data) April 2006 to March 2007 (the methodology was improved by adding additional estimates for rail travel using Transport for London (TfL) sold travel cards) April 2008 to March 2009 (the methodology was improved by adding estimates of rail travel in Passenger Transport Executive (PTE) areas) April 2015 to March 2016 (the methodology was improved using by Oyster Clicks Model (OCM) data used to allocate journeys made wholly within the London Travelcard Area to individual London stations rather than based on a survey from 2001) April 2022 to March 2023 (two large methodology changes impacted a large number of major stations. These were the split ticketing adjustment and reallocation of journeys to stations in the London Terminals group) Methodology changes each year affect some stations

Annex 1 - Details of methodology changes over time

Summary of methodology improvements each year since April 2006

April 2006 to March 2007:
Additional estimates for rail travel using TfL sold travelcards and airport links were included.

April 2008 to March 2009:
The generation of the ODM was integrated with the demand matrix in MOIRA, a software tool used by the industry to model the impact of timetable changes on the rail market. In addition to having LENNON data, MOIRA also provided more robust estimates of rail travel on TfL sold travelcards and airport links. It also included estimates of rail travel in PTE areas which had previously been excluded from the ODM due to a lack of data.

April 2009 to March 2010 and April 2010 to March 2011:
From January 2010, rail travel using Oyster pay-as-you-go (PAYG) was included in LENNON so these data were included in the ODM from 2009-10 with the first full year of data being April 2010 to March 2011.

April 2011 to March 2012:

- Improved estimates of travel in the West Midlands (Centro) PTE area included.
- Estimates of rail travel made using a small number of Rover and Ranger products included. The tickets included were: St Ives Day Rangers, Valleys Night Rider, and Cambrian Coaster Ranger. Whilst volumes of travel on these products are relatively small, in the specific area of use they can be significant.

April 2012 to March 2013:

- An improved PTE infill was included for two more PTEs – West Yorkshire (WYPTE) and Greater Manchester (GMPTE/TfGM).
- Estimates of rail travel using TfL's concessionary product, the 'Freedom Pass', were included for the first time.
- A further five Rover and Ranger products were included: Anglia Plus, Devon Day Ranger, Devon Evening Ranger, Ride Cornwall, and Freedom Travel Pass (West of England product).

April 2013 to March 2014:
A number of changes were made to improve the representation of journeys on PTE-sponsored tickets in South Yorkshire, Merseyside and Strathclyde.

April 2014 to March 2015:

- An improved infill for the Tyne and Wear PTE area was included.
- An adjustment process was made to account for the change in LENNON treatment of PAYG journeys to make the statistics more consistent with previous years. This adjustment was a one off as in April 2015 to March 2016 it was included in the MOIRA base matrix.
- An adjustment was required due to changes in journey patterns as a result of the London Bridge works. Data from Transport for London's (TfL's) Oyster Clicks Model (OCM) was used to estimate the number of journeys 'to London Bridge' and the number of journeys 'to London Terminals'.
- Journeys using a season ticket product for students have been redistributed to Exeter Central and Exeter St. David's from Digby and Sowton to better reflect actual journey destinations.

April 2015 to March 2016:

- London (In-boundary) Travelcard Methodology - Oyster Clicks Model (OCM) data used to allocate journeys made wholly within the London Travelcard Area to individual London stations rather than based on a survey from 2001.
- London Terminals Demand Allocation - improved due to MOIRA base matrix now disaggregated by individual London Terminal where possible, such as where a ticket is bought to a specific terminal rather than to the generic 'London Terminals.'
- St. Ives Branch Line Counts - Passenger counts were carried out at all five stations on the St. Ives Bay line (St. Erth to St. Ives) in August 2016 and the results of these counts were used to produce a more accurate allocation of entries and exits from sales of ranger or rover tickets across the stations.
- Season Ticket Journey Adjustment (Southend) – An adjustment to the allocation of usage at stations around Southend was made to account for season tickets issued for travel to/from Southend Victoria which were actually being used to travel from alternative stations on the branch, as the price of a season ticket is the same.

April 2016 to March 2017:

- London BR Allocation Update - Reallocation of some journeys for Kensington Olympia due to previous overestimates.

- Season Ticket Journey Adjustments (expanded number stations) - In the production of the April 2014 to March 2015 and April 2015 to March 2016 statistics, some adjustments were made to account for situations where passengers buy season tickets for travel to/from a station other than the one they generally travel from, in order to allow additional flexibility. For the production of the April 2016 to March 2017 statistics additional LENNON analysis was conducted and discussions with train operators to identify and include additional stations in the adjustment to better reflect their usage.

- Updated Demand Allocation at Group Stations – In order to validate and improve the allocation of journeys between stations within groups (e.g. Worcester BR), passenger counts were carried out at selected group stations on the network. These counts were carried out in Autumn and Winter 2016 and have informed the allocation of demand at the following station groups: Dorchester BR, Newark BR, Southend BR, Warrington BR, Wigan BR and Worcester BR.

April 2017 to March 2018:

- Season Ticket Journey Adjustments - Similar to previous years, adjustments were made to account for situations where passengers buy season tickets from a station other than the one they generally travel from. The analysis underpinning this reallocation was updated with April 2017 to March 2018 LENNON data.

- Updated Demand Allocation at Group Stations – Passenger counts were carried out in Autumn 2017 and have informed the allocation of demand at the following station groups: Bicester BR, Birmingham BR, Farnborough BR, Southend BR, Warrington BR, Wigan BR and Worcester BR.

April 2018 to March 2019:

- Concessionary travel in Greater Manchester - Concessionary ticketing data were available for Greater Manchester PTE for inclusion in the ODM. This led to a total increase of 3.6m journeys, or 7.2m entries and exits, across Greater Manchester.

- Season Ticket Journey Adjustments - Similar to previous years, adjustments were made to account for situations where passengers buy season tickets from a station other than the one they generally travel from. The analysis underpinning this reallocation was updated with April 2018 to March 2019 LENNON data.

April 2019 to March 2020:

- Merseyside PTE – Off network sales (commercial retailers, non-commercial retailers and Merseytravel centres) of Saveway and Trio tickets included. Previously only sales at stations and on trains were included.

- South Yorkshire PTE - Concessionary tickets (senior and disabled) included.

- Season Ticket Journey Adjustments - analysis underpinning this reallocation was updated with April 2018 to March19 LENNON data.

- Updated Demand Allocation at Group Stations – Passenger counts were carried out in Spring 2020 and have informed the allocation of demand at the following station groups: Dorchester BR, Edenbridge BR, Warrington BR and Worcester BR. In addition, updated splits for the Manchester BR group stations were provided by TfGM and implemented.

- Estimates for the three Heathrow stations were included for the first time. This addition also improved the estimates for London Paddington station and other local stations.

April 2020 to March 2021:

- Inclusion of concessionary tickets for West Yorkshire PTE and updated methodology used to estimate journeys using local tickets across the West Midlands and South Yorkshire areas.

- Updated season ticket journey allocation adjustments, including new adjustments for some stations in Wales.

- Inclusion of Caledonian Sleeper journeys.

- Updated allocation of journeys between selected Group stations based on passenger count surveys.

April 2021 to March 2022:

- An additional 98,200 entries and exits have been added to stations across the Tyne and Wear PTE area as a result of additional tickets being included in the dataset for the first time. This increased usage estimates across the Tyne and Wear PTE area by 1.3%.

- This year an updated distribution has been used for allocating entries and exits for West Yorkshire PTE tickets. The new distribution uses latest year (April 2021 to

March 2022) data (replacing the April 2019 to March 2020 data), better reflecting post-pandemic travel patterns. The sample size is also greater than the previous data, covering a full year rather than one week.

- Gateline data has been supplied by Heathrow Airport. This has been used to update the distribution of entries and exits across the three Heathrow stations and enabled a more accurate estimate of usage at each one.

- Following a review of tickets included in the estimates, a number of products were identified for inclusion for the first time. This includes the Derbyshire Wayfarer, Brighton and Worthing Unizone cards and more derivatives of the Bristol Freedom Pass product. This led to an addition of 616,000 entries and exits.

April 2022 to March 2023:

- Split ticketing is where a passenger completes a single journey using two or more tickets. Total entries and exits at some individual stations will be significantly over estimated due to some passengers not boarding or alighting at the split point and instead staying on the train.

  From April 2022 to March 2023 year we have introduced an adjustment to account for split ticketing in the LENNON data. Rail Delivery Group have developed an algorithm to detect split tickets, and we have incorporated this into our data processing. Therefore, in the latest year (April 2022 to March 2023) we only count an entry and exit at the start of the journey and end of the journey, whereas previously an entry and exit would have been recorded at each split point along the way.

- This year an updated distribution has been used for allocating entries and exits for passengers who have bought non-travelcard tickets with a destination of London Terminals, for example Oxford to London Terminals. The new distribution uses modelled data from the latest CAF data to allocate entries and exits to each station based on possible travel routes. Previously entries and exits were allocated using results from the 2001 London Area Travel Survey (LATS). While many allocations to central London stations will be similar to when the LATS took place, some will have changed significantly (e.g. due to the impacts of Thameslink). This represents a significant methodological improvement.

- This year an updated distribution has been used for allocating entries and exits for West Yorkshire Metro tickets. The new distribution uses latest year (April 2022 to

March 2023) data (replacing the April 2021 to March 2022 data), better reflecting post-pandemic travel patterns.

- The methodology for calculating concessionary demand in the Greater Manchester region has been updated. A new data source was available (rail surveys undertaken in August 2023) which covered all ticket types (including concessions on radial Manchester rail routes). Previously survey data from 2017 was used. This data is to estimate the proportion of concessions across all ticket types to account for concessionary travel that would otherwise not be covered by the ticket purchase data.

April 2023 to March 2024:

- A methodological adjustment was used to account for journeys in the MOIRA2.2 base matrix that have the same origin and destination. This issue mainly occurred in the Oyster/CPAY area where passengers either did not tap in or tap out. The adjustment reallocates these journeys to other flows based on the underlying distribution of Oyster/CPAY journeys at the station.

- A methodological adjustment was used to exclude journeys that were paid for but never used. These are tickets that were not collected from 'Ticket on Departure' (TOD) machines. This removed 850,000 journeys.

- An updated distribution was used for allocating journeys using Strathclyde Partnership for Transport tickets. The updated distribution used April 2023 to March 2024 data (replacing pre-pandemic data), better reflecting post-pandemic travel patterns. This change did not affect the total number of journeys but it did impact their distribution across the network.

See Steer's historical methodological changes report for more detail.

This publication is available at orr.gov.uk

Any enquiries regarding this publication should be sent to us at orr.gov.uk