Office of Rail Regulation

Network Rail

Right Time Performance Measure

Final Report (Redacted)

Issue 5 | 22 January 2013

Job number 223767-07

Ove Arup & Partners Ltd
13 Fitzroy Street
London
W1T 4BQ
United Kingdom
www.arup.com

ARUP

Document Verification

ARU

Job title		Right Time Performance Measure			**Job number** 223767-07
Document title		Final Report			**File reference**
Document ref		Report Ref			
Revision	**Date**	**Filename**	RTP Draft Report v1.doc		
Draft 1	5th Sept 2012	**Description**	First draft		
			Prepared by	Checked by	Approved by
		Name	Keith Winder	Douglas Leeming	Stefan Sanders
Issue 1	14th Sept 2012	**Filename**	RTP Draft Report v2.doc		
		Description	Draft issue		
			Prepared by	Checked by	Approved by
		Name	Keith Winder	Ian Hood	Stefan Sanders
Issue 2	26th Sept '12	**Filename**	RTP Draft Report v5.doc		
		Description	Revised Draft		
			Prepared by	Checked by	Approved by
		Name	Douglas Leeming	Keith Winder	Stefan Sanders
Issue 3	12th Oct '12	**Filename**	RTP Draft Report v6a.doc		
		Description	Revised draft following tripartite meeting on 3rd October 2012		
			Prepared by	Checked by	Approved by
		Name	Douglas Leeming	Keith Winder	Stefan Sanders
Issue 4	20th Dec '12	**Filename**	RTP Draft Report v8.doc		
		Description	Revised draft following receipt of Manual Reporting data		
			Prepared by	Checked by	Approved by
		Name	Douglas Leeming	Stefan Sanders	Stefan Sanders
Issue 5	22nd Jan '13	**Filename**	RTP Draft Report v9.doc		
		Description	Revised draft following receipt of Manual Reporting data		
			Prepared by	Checked by	Approved by
		Name	Douglas Leeming	Stefan Sanders	Stefan Sanders
		Signature			
				Issue Document Verification with Document	✓

Contents

Executive Summary

Purpose

This report provides a review of the processes and accuracy of the data which Network Rail use to record train arrivals at final destination. The report was commissioned under Mandate AO/033 issued by the Office of Rail Regulation (ORR) and Network Rail in July 2012. The study was commissioned to allow the ORR and Network Rail to determine the quality of 'right time performance' data, which considers train accuracy to a one minute tolerance of timetabled arrival time. This is in light of the desire to publish right time alongside other performance metrics.

Background

Network Rail broadly employs two methods of capturing arrival data:

Automatic Reporting: where the arrival time is recorded automatically by the signalling system based on the time at which the train passes a known point combined with the addition of an allowance (the berthing offset) which allows for the time taken for the train to stop at the platform; and

Manual Reporting: where the arrival time of the train is input directly to the train reporting system by signallers based on either direct sight of the train coming to a stand, reports from the traincrew, or by applying an allowance to an observed timing; for example when the train passes the signal box.

During the course of the study interviews were held with the Network Rail teams in the National Centre and in eight of the Routes. In each case documentation and questions relating to the processes and data were discussed.

Automatic Reporting

Automatic reporting covers 77% of the final destination locations and handles 91% of train movements. It was found that there are well documented processes in place to support the systems used to report train at these sites and the vast majority of the measurements on most Routes (see Figure 4-2) are up to date and accurate within the limits of the system. The berthing offset review process, when infrastructure or operational aspects change on site or when a site is due a review, were in some respects slow and difficult to follow through. This was particularly the case where there was a requirement to introduce changes to the berthing offset values. The delays in the progression of such changes appeared to lie with the respective TOCs; however resource constraints and competing pressure on staff time within Network Rail had a contributory effect on the build-up of a backlog of updates to the allowances. It was noted that efforts are being made centrally to address these.

Manual Reporting

These cover the remainder of the reporting locations. There appeared to be no discernible processes in place to manage the reporting of arrivals at manual reporting locations. This was clear from all of the Route meetings. Whilst this concerns less than 10% of the daily arrivals nationally, it was found that the

percentage of manual reporting by individual Routes, and certain TOCs, varied significantly. The lack of such processes was concerning.

The study's review of the accuracy of the manually reported data relied on the output of surveys carried out by Sheffield University for certain TOCs. This provided the comparison between an observed arrival time to the second compared to the time recorded in TRUST by the signaller. Over five hundred such examples were reviewed.

Findings

The processes associated with the two regimes provided a stark contrast. The automatic reporting appeared to be well documented and a well conceived process. There were some issues with its delivery however. Manual reporting was completely without recorded process. On this basis an overall Confidence Grading for reliability of '**C**' has been awarded.

With regard to the measure of accuracy it was possible to develop methodologies which would provide an indication of how accurate both regimes were. Essentially these considered the impact of the 'errors' in terms of tipping certain arrival times over or under the one minute threshold. Results were produced for each regime by TOC and Sector and from these a combined (automatic and manual reporting) assessment was made. Table 1 summarises the combined accuracy gradings.

Table 1: Overall Accuracy Score

Measure	% Error	Accuracy Grading
NATIONAL		
Great Britain	2.23%	2
England and Wales	2.14%	2
Scotland		
SECTOR		
Long Distance	2.20%	2
London & the South East	2.11%	2
Regional	2.18%	2
Scotland		
FRANCHISED TOC		
First TransPennine Express		
Greater Anglia		
Northern Rail		
Heathrow Connect		
First Great Western		
First Capital Connect		
CrossCountry		
London Midland		
London Overground		
East Midland Trains		
First ScotRail		
East Coast		
Merseyrail		
Virgin Trains		
Arriva Trains Wales		
Chiltern		
c2c		
Southeastern		
Southern		
South Western Trains		
Franchised	2.23%	2
NON FRANCHISED TOC		
Grand Central		
Heathrow Express		
Hull Trains		
Non-Franchised	1.68%	2

Conclusion

Based on the foregoing the analysis has shown that with an approximately twenty thousand train services operating each day in the order of 440 of them may be miss-reported under the proposed regime. It is the Independent Reporter's view that this level of error is not significant enough to warrant the invalidation of the publication of the right time performance statistics based on the current regimes of train arrival recording.

1 Introduction

1.1 Background

Performance data is currently recorded by Network Rail for all passenger train services operating on the UK rail network. The timings for each train service are extensive and cover:

- journey commencement;
- a range of intermediate points; and
- journey's end.

The published measures for passenger train performance currently consider only the arrival time at the final destination of the train. These are measured against thresholds of five and ten minute allowances depending on the type of train being considered.

As part of its aspiration to improve the transparency of published information, the Office of Rail Regulation (ORR) has committed to regularly publishing Right Time Performance data.

Right Time Performance (RTP) will continue to focus on the arrival time of the service, but in this context, "Right Time" is defined as the point when the wheels of a train stop at the terminal station. To have arrived at the 'Right Time' a train must arrive at no more than fifty-nine seconds after its scheduled time of arrival as published in the public timetable. ORR proposes to publish this data at both the Train Operating Company (TOC) and Sector levels.

In moving to RTP as the measure of train punctuality there is concern that the quality of the data used to generate the measure is currently unknown. As a result the ORR and Network Rail have commissioned this study to determine the quality of this supporting data.

The requirements for this data quality audit have been specified in Mandate AO/033. A copy of the Mandate is included in Appendix A.

1.2 Scope

The scope of the study requires the assessment of the accuracy and reliability of the RTP measure. The review is to assess the quality of the measure at the following levels:

- National (Great Britain, England and Wales, and Scotland);
- Sector (London and the South East, Long Distance, Regional, and Scotland); and
- Train Operating Company (Franchised and Non-Franchised TOC).

1.3 Structure

Following this introductory section the report is broken down as described below.

Background: providing a summary of the history and description of the processes to derive the data which is currently reported by Network Rail.

Methodology: a description of how the review was undertaken considering the interviews which were held and the documentation which was reviewed.

Findings: a statement of the findings of the study considering what was discussed at the meetings and a review of the available data.

Conclusions: using the findings as its basis a brief statement is provided on the conclusions of the study.

Confidence Grading and Recommendations: provides an alpha-numeric grading of the current process and data accuracy along with recommendations going forward to improve the data quality.

2 Background

2.1 Introduction

This section of the report provides a brief description of the history of the performance measurement arrangements and how these have developed to the systems which are currently in use.

It is important to describe these arrangements since an understanding of the methodology by which performance data is captured and recorded is a fundamental prerequisite to understanding the limitations of the data.

There are two principle means of recording arrival times.

- Auto-Reporting: This relies on the application of an adjustment to the last automatically recorded timing of the train within the signalling system. This is to take account of the time difference between the time reported in the signalling and the time the wheels stop at terminating station. The arrangements by which these adjustments are derived and kept up to date have a significant bearing on the overall effectiveness of the performance monitoring arrangements for RTP.
- Manual Reporting: The second means of recording train arrivals relies on the manual observation of the train and the manual inputting of times into TRUST.

2.2 TRUST

The processes through which Network Rail gathers performance data for trains operating on the network were established prior to the privatisation of the industry. These have remained largely unchanged since then, save for the various technological "add-ons" to the TRUST train reporting system, which have improved the granularity of data and the extent of automatic (as distinct from 'manual') reporting into the mainframe system.

The TRUST system captures timing information for each train over the course of its journey. This automated input occurs when a train activates a signalling detection system (for example track-circuit or axle counter) as it passes throughout the network. When this occurs the information is linked to the identification of the train (from the train describer) and is then passed into the SMART system. This aggregates the data from all such connected train describer systems and passes it into downstream information systems. Where the details correspond to a TRUST reporting point, SMART will flag the activity accordingly and the event will be processed by TRUST. These reporting points are locations on the network where train passing or arrival times are captured by the signalling system. This data is required to fulfil the monitoring obligations as part of the contractual arrangements between the Train Operators and Network Rail. There are three types of reporting points:

- Timing Points;
- Delay Recording Points (DRPs); and
- Contractual Monitoring Points (CMPs).

It is important to note that SMART records accurate strike-in times to the second. TRUST, however, only records whole minutes. This is because TRUST ignores the seconds' element of the recorded time. Thus, a strike-in which aggregates to two minutes and fifty seconds will be recorded in TRUST as two minutes, as will an aggregate of two minutes and two seconds. This explains the rationale behind the right time definition – a scheduled arrival time of 14:22 is actually achieved if the strike-in plus the offset time aggregates to no more than 14.22:59. In this case it will be recorded in TRUST as 14:22.

The various types of reporting point have different functions. However, they all form part of a nationwide reporting and recording system for capturing train performance data measured as actual times compared to the timetable.

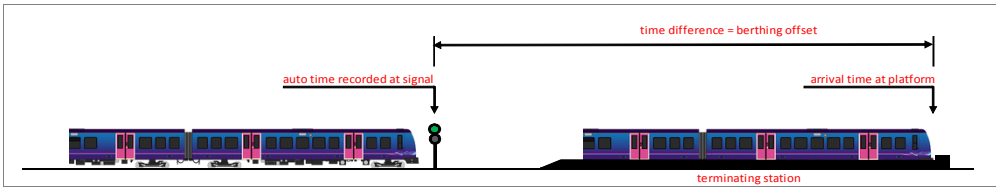
There are a number of locations in manually signalled areas where automatic reporting equipment is not available. At these locations reporting information is still required and thus an alternative reporting method must be employed. These are termed manual reporting locations. Generally these rely on the signaller inputting the arrival time of the train directly into TRUST. In some locations, where the signaller may be unsighted, traincrew or other Network Rail staff will advise the signaller of the train's arrival time.

Timings recorded at all of these points form the basis of the calculations in the contractual financial regimes between Network Rail and the TOCs (known as Schedule 8 of the Track Access Contract). Reporting at the Contractual Monitoring Points is therefore critical not only for the performance information but also for the financial relationship between Network Rail and its Train Operating customers.

2.3 Berth Offset

Away from manually reported areas the destination times recorded in the system are not generated simply by the time recorded at the final signal berth prior to the termination platform. This is because the automatically generated time does not represent the time at which the wheels of the train stop at the platform. To generate the arrival time based on the automated data an allowance, the berthing offset, must be added to the time recorded by the signalling equipment – see Figure 2.1.

Figure 2-1: Diagram of Berth Offset Allowance



As a result of the signalling system not being able to measure the 'wheels stop' time directly the berth offset must be calculated for each permutation of approach route and destination platform that is used by services on a planned basis. This is derived by undertaking a site survey of actual train times covering various combinations of rolling stock.

Complex terminating stations can have a significant number of offsets; for example at London Euston there are 182 berth offsets. It should be noted that berth offsets are also calculated for departing services where there would be a deduction made from the signal timing to count back to the time when the wheels started to roll.

The value of the berth offset is calculated by averaging the respective time difference between the signalling system and the wheel stop point through field observation. A record of every offset for every route is maintained in a Route Margin Book. This Network Rail National Centre generated document is periodically updated and reissued as and when offsets have been reviewed and updated.

2.4 Performance Measurement

These arrangements, and the requirements which sit alongside them, are detailed in the Network Rail Performance Measurement Manual (PMM). The contractual obligations for data accuracy and the interface arrangements with TOCs are described in a code of practice, the Performance Data Accuracy Code (PDAC), (version dated 24th July 2011) which is 'owned' by the Delay Attribution Board (DAB) on behalf of the industry parties.

In terms of understanding the accuracy and reliability of performance data generally, and the RTP measure in particular, an important feature of these arrangements is that for auto-reporting locations the system and the processes which support it do not record actual arrivals. They record an assumed arrival time calculated from the recorded time in the signalling system modified by the addition of a berth offset derived by sampling trains over each route into the terminating station. Given the established accuracy of the signalling system time recording, all of Network Rail's associated data quality procedures for Right Time Performance are focused on ensuring that the offsets are as accurate and current as possible. This includes the various administrative processes which support the assessment of offsets.

Inevitably in a process involving sampling data 'perfect' accuracy is impossible to achieve. At the level of PPM, whether at the five minute or ten minute threshold, the level of inaccuracy is very small. The only trains which are at risk of being incorrectly reported are the relatively small number falling at the fringe of the measure. For example under the PPM, regime those local services actually arriving at six minutes and five seconds late (one minute late compared to their schedule) but which, because of the offset, may be recorded arriving at five minutes and fifty-five seconds would therefore be recorded as right time.

In respect of the RTP measure, the number of trains reporting at the fringe of the measure is very significant - nationally, approximately 70% of passenger services are recorded as 'Right Time'. Within this a significant proportion of these will actually be arriving in the fifteen second window either side of the Right Time threshold.

As a result the accuracy of berth offsets is a very significant issue in the context of the RTP measure. The result of this is to magnify the offset accuracy's impact on the reporting RTP and could potentially be a key weakness of the system. The accuracy of the offset will also affect the results under the PPM regime but will be

less significant because of the smaller population of trains which the variation will affect.

2.5 Manual Reporting

All of the arrangements described thus far relate to auto-reporting CMPs. Additionally, there are a not insignificant number of CMPs which cannot be reported automatically, and have to be reported 'manually'. On the network there are 464 (terminating) CMPs of which 107 (23%) are reported manually. Manual reporting locations are heavily concentrated in Western & Wales, London North Eastern, London North Western and Scotland Routes. They also exist to a lesser extent in Anglia. There are no manual reporting CMPs in Kent, Sussex or Wessex. Whilst many of these locations are remote or branch line locations with manual signalling and an infrequent train service, a number are suburban or inter-urban termini with more frequent services – for example Cleethorpes on the Transpennine Express network.

The Network Rail requirement to monitor/ review/ audit manual reporting locations was deleted from the PMM some years ago. This is because it was considered by Network Rail to be impractical for the PDQS to annually maintain an adequate level of surveillance at these locations. It may however be possible for other Route based personnel to monitor the processes and accuracy at these locations. This may involve checks made by the LOM. In the meantime, given the proportion of locations covered by manual reporting this lack of checks has a significant impact on the overall reliability of the process.

2.6 Offset Reviews

As noted earlier, the arrangements by which offsets are proposed for change has a significant bearing on the overall effectiveness of the process. In summary, these arrangements are as follows;

Section 5.8 of the PMM describes a series of conditions or circumstances where the offsets recorded in the Margin Books (see Section 4.3.6 of this report) may be changed or proposed for change. Essentially the offsets at a CMP should be physically reviewed when:

- any change to infrastructure characteristics has occurred;
- when changes to rolling stock (including types of rolling stock or the train formation) have occurred;
- if stopping locations at a platform have been changed; or
- if any party to the data requests a review.

In the steady state Network Rail has adopted a regime of reviewing each CMP every five years. If there are no grounds for believing the offsets might be inaccurate, and no known changes have taken place to the infrastructure or operations, the Route Performance Data Quality Specialist (PDQS) may propose that the berth offset is retained. In such cases a "CMP/ DRP No Change Review" pro-forma is issued to the relevant TOCs and FOCs, seeking signatures to the acceptance of the current offsets without the requirement for a physical review.

Where there has been a change to the infrastructure or operation, or there is no acceptance of the status quo, then a site review must be undertaken. The PMM gives clear guidance regarding how the physical reviews of CMP offsets should be structured and organised. It describes the equipment and skills required, and the sample train service size to be used over each route into a terminating point to give a 'representative' set of data on which to calculate the 'average' offset value. Guidance is also given in the PMM concerning the need to avoid abnormal conditions (for example bad weather, leaf fall, service disruption) and to remove from the sample data any obvious 'outliers' in order to ensure the data sample provides a reasonable assessment of normal conditions. The Routes are encouraged to involve TOC representatives in the surveys, to try to engender some ownership of the process and results. This also helps to resource what can be labour intensive surveys.

When a survey has been completed, fully documented, and average offsets calculated, these are compared to the previous figures to determine if there has been a material change in any of the offsets. In this context 'material' is currently defined in the PDQS Guidance Notes (Version 2 June 2011) as three seconds. Thus where a variation of three seconds or greater is found through the field work the PDQS must initiate the documentation proposing the appropriate change(s).

The process of change is based on the circulation by the PDQS of an appropriately completed Recording Point Change Request (RPCR) pro-forma. This is sent to all of the train operators running services at the given location. It contains details of the proposed changes to the offsets and seeks their endorsement of the update.

The issue of the RPCR also initiates the commercial process for neutralising any financial impact in the Schedule 8 performance regime between TOCs and Network Rail directly caused by the changes in offset values. The means of calculating the financial effect of the offset change is undertaken in software known colloquially as the "Ready Reckoner". This work was previously managed by the Route Commercial teams but has more recently been operated by the Network Rail National Centre Performance Process and Controls teams. Once the "Ready Reckoner" has been run, the results agreed by the respective commercial teams ('General Approval'), and providing a unanimous sign off to the offset values has been received from all the affected TOCs and FOCs, the changes to the offsets can then be implemented. This is done by the Network Rail National Centre team responsible for SMART.

There are a small number of circumstances where more limited and less onerous consultation requirements apply. For instance, where major infrastructure change has taken place, the PDQS is mandated to undertake a prompt review of all the affected offsets and implement changes as soon as possible after commissioning. If the scheme has been commissioned with interim estimated offsets (to populate the system ahead of a full review) a survey should also take place as soon as possible afterwards to confirm the revised values.

2.7 Study Focus

As described above, the two key elements of the process are the mechanisms for gathering data for calculating the offsets applicable to terminating trains (covering automatic reporting), the procedure for initiating (and following through) any

required changes to the timing values in the TRUST system. It will also consider the accuracy of the manually reported locations.

The focus of this review has therefore been as follows:

- to assess Network Rail's internal processes and procedures, at both National Centre and in the Routes, to evaluate the extent to which the arrangements protect the integrity of reported data, and minimise the inherent inaccuracy; and
- to quantify the levels of reliability and accuracy which emerge from these arrangements to provide a robust grading for the RTP measure.

3 Methodology

3.1 Introduction

This section of the report provides a summary of the methodology which has been used on the study. It covers the meetings which took place at Network Rail National Centre and on the Routes. It also describes the gathering, review and analysis of the base data which was used to determine the reliability of the reporting at the levels specified in the commission brief.

3.2 Approach

We have approached this review in a similar manner to that undertaken to assess the reliability and accuracy of the regulated performance KPIs (e.g. PPM, CaSL). However, for RTP we have undertaken a more intrusive and detailed review of the processes, procedures, and systems deployed, to validate and review the arrangements in place to determine the accuracy and reliability of the data. Our review has assessed data and the methodologies for gathering it in all ten Network Rail Routes.

The information gathered from the meetings with the National Centre and Route teams forms the basis of our assessment of the reliability of the RTP data at auto reporting locations. Raw data has been sourced from Sheffield University for an assessment of the manual reporting locations. The data gathered and its subsequent analysis is the basis of our assessment of the accuracy of the results.

Our methodology is described in the following sub-sections.

3.2.1 Network Rail National Centre Performance Process and Controls Team

As an early activity a meeting was arranged with the Network Rail National Centre Performance Process and Controls team. The aim of the meeting was to review:

- definitions for “Right Time”, how these are codified and how the measure is computed;
- processes for recording and reporting data, including identification of the various methods in use across the network (auto reporting through TRUST, offsets and ‘manual’ reporting by signallers etc.);
- systems in use, and how they are calibrated and assured as accurate;
- documentation such as standards and process charts which describe the arrangements, and how they are managed, reviewed and changed;
- how records are maintained, particularly in respect of changes to the arrangements; and
- any internal audit or assurance arrangements in place for both systems and processes, and data quality.

3.2.2 Network Rail Performance Data Quality Specialists (PDQS)

Following the discussion with the National Centre team a series of meetings were arranged with the respective Network Rail Performance Data Quality Specialists (PDQS) within the Performance teams at Route level. The principle aim of these was to review the local application of the national processes, systems and arrangements. The PDQSs were asked to invite representatives from the TOCs operating on the respective Routes. The meeting sought to review:

- local arrangements, especially where these differ from those generally in place;
- data recording for locations where auto reporting does not apply, or where the signalling system prevents standard auto reporting;
- liaison between the TOCs and Network Rail Route in respect of changes to data reporting or recording arrangements;
- assurance and data quality management;
- what use is currently made of the RTP measure; and
- whether any TOC or Route is exploring new technology for the recording of performance data.

There are currently ten Routes but only eight PDQSs. This is because two of the Routes have their work undertaken on an 'agency' arrangement. Western Route covers the work and responsibilities in Wales; and London North Eastern looks after the Midlands Route. In the course of the study all eight PDQSs were visited.

3.2.3 Data Analysis

A key output from this study, as well as an assessment of the procedures and processes, was to assess the accuracy of the data.

To come to a view on this we acquired data at three levels. The first being a picture of the overall size of the information associated with RTP, derived from the National Centre team, covering:

- the numbers of CMP and whether they were auto or manually recorded;
- the number of berth offsets; and
- the number of train services associated with each offset.

The second level, in the Route, considered the status of these individual elements including:

- how many offsets are currently not up to date (for whatever reason);
- the means of the sampling to determine updated offsets; and
- what the scale of change is in the updates to the offsets.

Finally, for the manual reporting locations where no comparative data existed within Network Rail access was granted to original research undertaken by Sheffield University on behalf of certain Train Operating Companies.

It was the analysis of each of these data sets which formed the basis of the assessment of the current accuracy of the RTP data.

4 Study Findings for Auto-Reporting

4.1 Introduction

This section of the report considers the findings associated with the arrival times reported by means of the automatic systems including the use of the berthing offset. It considers the results from the National and Route levels before describing the data which was acquired from the subsequent analysis which was undertaken.

4.2 Network Rail National Process

In looking at the recent history of the process it is clear that Network Rail has identified that there are issues with the RTP data and has instigated action to address these. It was suggested at the Network Rail National Centre that the last eighteen months had been a period of consolidation. The actions taken have been designed to support the Routes in recovering a significant backlog of work which had arisen for a variety of reasons – for example RPCR forms were not being circulated, consultation with TOCs was not being completed, and TOCs were not responding to the forms. This backlog manifested itself through adherence to the requirements of both the Network Rail PMM becoming increasingly patchy. All of this is acknowledged by Network Rail.

The National Centre view was that this position has now been largely addressed. There was evidence to support the view that the Routes were certainly getting more help and that there was now stronger direction from National Centre.

Network Rail National Centre is now routinely undertaking CMP process compliance audits of the Routes. An Annual PDAC data completion review now records the extent to which the Routes are managing to recover and backfill missing data. In addition a number of amendments to the PMM are now in the process of implementation some of which will sharpen up the RPCR arrangements. These aim to improve the focus on finalising the commercial negotiations associated with offset changes. As a consequence of the reinvigorated focus from the National Centre, a number of the fundamental flaws in the processes are now being addressed.

From the ‘customer’ perspective it was noted that the National Centre Performance Process and Controls team have received no representations from TOCs or FOCs expressing any concern about the processes for data collection and the accuracy or reliability thereof.

4.3 Network Rail Route Process

4.3.1 Process

In terms of the process associated with the determination of the berth offset it clearly relies on the Route team visiting site and gathering data on actual train movements and timings. There is a reliance on the survey being a representative sample of the operations at the site on a normal day. A representative sample of each movement is required to be surveyed and the results analysed to provide an

appropriate average offset. Nevertheless there remain a number of significant flaws in the process for setting berth offsets, yet there appear to be no straightforward means of rectifying these given the systems which are currently in use. The flaws relate to:

- abnormal conditions at the final signal berth approaching a terminus; and
- the sampling regime when reviewing offset values.

Average offset values correctly do not take account of the abnormal, temporary, or transient situations which occur from time to time. Such situations almost always slow trains down in arriving at a terminus and in such situations the TRUST arrival times recorded with offset applied will generally be better than actual – in some cases quite significantly so. Such situations are far from unusual or infrequent – and will generally occur every time a train is stopped at the final signal approaching the terminus. This results in the train approaching the platform more slowly than usual. A similar situation can occur when a speed restriction is in place close to the final signal berth, during the autumn leaf fall season, and when defensive driving is in place.

The guidance with regard to the sampling given to those undertaking physical surveys to update offsets recommends twenty train reports for each route and not less than six trains at CMPs. There is no evidence that these figures have been justified from the perspective of statistical validation. They merely represent pragmatic advice to gather a sample of the movements. Given the variety of movements at some of the complex locations it is recognised that even the modest number of samples required will not be achievable in many situations. As a consequence, the average offset values cannot be robustly assured statistically, even when evaluated diligently. The impact of the sampling process is assessed in Section 4.5.2.

4.3.2 Berth Offset Reviews

As a general finding the five yearly checks on the offsets were broadly up to date. One Route believed the threshold to be three years and was reviewing at this timescale. At least one other was trying to re-survey at four years, in recognition of the time taken to implement survey changes. Where a shortfall in the survey data existed, it concerned a small number of locations which were stated to be difficult to re-survey either because of the logistics of organising the site visit (resourcing and timing the work) or because of the remoteness of the location. The issue here was due to a backlog of work which had yet to be properly recovered.

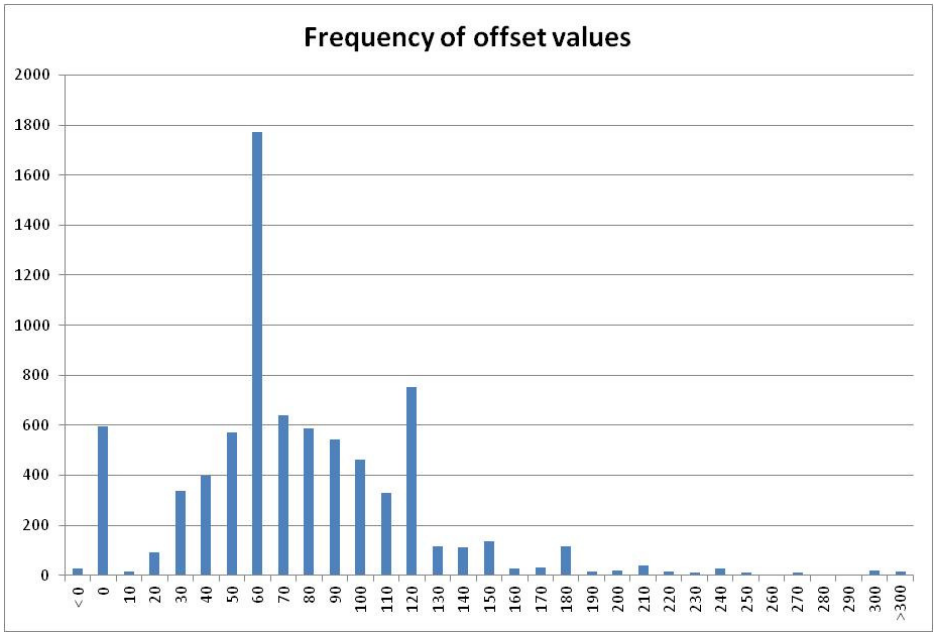
One Route had sixteen of its CMPs overdue out of a total of forty-two auto-reporting points. Some of these have been in this status since 2001. Regardless of this there was evidence from the records of the site surveys and commentaries that a recovery plan was now in place.

In general, Routes were reluctant to enact the “Declaration of No Change” process for the five yearly reviews. This was because there appeared to be a general lack of confidence in their internal arrangements for keeping up to date with changes which may have occurred. Tracking changes in the infrastructure, rolling stock or operations was well structured in one Route but more ad-hoc in some although where the PDQS was based in close proximity to the TOC this seemed to provide a reliable source of information although on an informal basis. As a result of this

lack of confidence most CMPs are being fully surveyed at their five year anniversary.

Prior to 2004 the offsets were measured to the nearest sixty seconds. As such it might be expected that these would have all been converted to the more accurate one second interval measures. From the review of the data it would appear that 19% of CMP locations have berth offsets measured to nearest sixty seconds. This is further evidenced in Figure 4-1 which shows the distribution of offset values. It is however noted that some of these will be old offsets associated with infrequently used or non-passenger train moves which may be a low priority to update.

Figure 4-1: Distribution of Size of the Offsets for RPCR Sample



4.3.3 Survey Documentation

The documentation produced by the PDQS as a result of the field work to validate the berth offsets was sampled across the Routes. The records were found to be complete in all cases. All the PDQSs had a disciplined approach to survey planning. The manner in which surveys were undertaken took account of the National Centre guidelines on the planning and conduct of surveys. The range of sample trains used to validate a berth offset varied with dominance on a sample size of six. Despite these variations there was no evidence on any Route of complaints from Operators about the manner in which surveys were being undertaken or how the results had being derived. On one Route a TOC which would be particularly affected by a proposed change had independently surveyed the offsets and verified the proposed changes. However, it was clear from our meetings with the PDQS and TOC representatives that in the majority of the cases there was a good working relationship between the Network Rail and TOC teams.

4.3.4 Berth Offset Change Process

Based on the findings from the Route interviews it was clear that the initiation of the process to amend a berth offset through circulation of the RPCR pro-forma by the PDQS was generally very good. However the overall process was far from satisfactory in almost every Route. Even where the process appeared to be working well and where the primary TOC was properly engaged in the process and understands the requirements, it can take in excess of fifteen months to achieve the required agreement. It is however noted that this time period has significantly improved recently with the intervention of the National Centre Performance Process and Controls team.

The process involves the unanimous sign-off of Part A of the RPCR (signifying acceptance of the proposed physical changes to the offsets) from affected Operators, the running of the 'Ready Reckoner', commercial sign off, and then the implementation of the proposed changes. It is noted that Part A is deemed to have been accepted by a TOC if it fails to respond within a twenty-eight day period.

On one Route fifty per cent of the CMP population is currently within the change process. In another this figure was 43%. The national average is 14% (see Table 4-3). This was taken as a clear sign of a process not working although it is recognised that work is taking place to remedy the situation.

By definition once the need for a re-survey has been identified and the review process is running the Route is potentially operating with out of date, inaccurate offsets until proven otherwise. The more CMPs that are inside the change process, and the longer this takes, the less accurate and reliable is the performance data recorded at these locations.

4.3.5 Ready Reckoner

The running of the 'Ready Reckoner' follows on from sign off to the physical timing change proposals. It is a commercial process designed to neutralise any financial movements within the Schedule 8 regime resulting from the berth offset change. The process is in accordance with the requirements of the PDAC, a document which is 'owned' by the Delay Attribution Board and managed by Network Rail.

From the study it is clear that this is almost certainly the least well understood aspect of the change process - both by the TOCs and by Network Rail's commercial teams. It was also clear that until the recent intervention by the National Centre Performance Process and Controls teams, Route Commercial teams gave the process little priority leading to a back log in the progression of this very necessary step in the process. The running of the Ready Reckoner by the National Centre Performance Process and Controls teams has dramatically improved the situation. This was evidenced by the records of agreement of changed offsets and comments from Network Rail and TOC representatives at the meetings. This is understood to be a permanent arrangement which will not involve input from the Route Commercial teams.

4.3.6 Margin Books

All Routes were able to demonstrate that they had a Margin Book. This details all the current offsets in place throughout the Route. The Margin Books are produced for each Route and are managed centrally by the National Centre Systems Support team who reissue it periodically with updates.

It appeared to be common practice in the Routes to submit the changes but not to validate that the correct updates had been captured in the revised versions of the Margin Book. However, this review of the accuracy of the Margin Book entries and updates is undertaken by the National Centre Performance Process and Controls team.

4.3.7 Staffing Levels

From the visits to the Routes it became clear that there are a number of different approaches to covering the PDQS responsibilities. The norm appears to be one or two PDQSs responsible for all aspects of data completeness, data integrity, and the routines surrounding DRPs/ CMPs; the latter comprising around fifty per cent of the workload. One Route had delegated much of the daily data completeness work to the Delay Resolution Coordinators, and had, as a result, freed up more PDQS time for the more strategic aspects of the role. A second Route had plans for organisational change with something similar in mind. Only one Route had a full time post dedicated to managing and reviewing the processes around SMART berth offsets. From the study observations only two Routes appeared to have adequate arrangements in place for covering the PDQS workload in the event of prolonged absence. A number of Routes were carrying a backlog of work because a prolonged absence had recently occurred and the responsibilities had not been adequately covered during that period.

All the Routes with a single PDQS felt the staffing / workload balance was fragile. The volume of work was considered to be manageable when things were normal but very difficult to cope when a backlog arose. This could come about through a secondment, staff sickness or non-availability, or when there is a significant workload associated with data recovery after a spell of bad weather, or system failure.

The individuals occupying the PDQS posts were generally long serving with a high of twelve years' experience in one case and all others having spent at least five years in post.

4.3.8 Perverse Incentives

A number of Routes had experience of TOCs being reluctant to progress the change process for apparent fear of this impacting negatively on their performance figures. However there was absolutely no evidence of this manifesting itself in either a refusal to participate, or as an apparent attempt to hide the real data. All Routes and TOCs acknowledged that there could be a perverse incentive, or disincentive, to update offsets which could in extreme circumstances be construed as dishonesty. Whilst understanding the point, at the meetings with the Routes and TOCs no evidence was provided that this behaviour had occurred. Subsequent comments from the National Centre Performance Process and Controls team indicated that there had been experience of TOC manipulation of

the process to their benefit in a Route where a TOC franchise was due to be renewed.

4.3.9 Use of Right Time Performance Data

During the interviews most Routes and TOCs confirmed that the use of Right Time data has increased significantly in recent years. Many TOCs now publish internally their Right Time Performance within the suite of performance data produced periodically.

In addition, many Routes and TOCs have cross-industry Right Time Railway Groups active in their areas. These generally rely on Right Time data as the focus for performance issues. A number of high performing TOCs, such as c2c, use Right Time information as their default KPI on the grounds that PPM data is too crude to provide a guide to the resolution of delay incidents. All of the users of Right Time data are frustrated at the lack of granularity in the current data, the three minute sub-threshold in TRUST, and the general lack of available management information. Several bemoaned the exhaustive efforts which have to be employed to obtain good quality data at this level of granularity, for example by cab riding etc.

4.3.10 Technological Advances in Right Time Performance Reporting

Under questioning almost all the Routes and TOCs believed that recent advances in technological applications offered the potential to improve upon the current somewhat outdated methodology for recording train arrivals accurately. However, none were confident that these applications could be linked together in a consistent, coherent and reliable manner in the medium term. There was also a strong doubt expressed at Route level regarding the will of the industry to standardise on technological applications at this time.

The best prospect appeared to lie in some form of hard link between modern On Train Data Recorders, GPS or proximity detection, and an accurate and reliable on-board clock, which will physically and reliably record wheels stop at terminus stations and pass the data automatically to Network Rail's reporting and recording system. It is considered that, in the absence of a national, all-party commitment these aspirations will take many years to realise.

4.4 Base Data

This section incorporates the key baseline data which describes the current process associated with the whole berthing offset regime.

Table 4-1 provides a summary of the distribution of the CMPs by Route which shows whether they are automatically or manually reported. It also shows the percentage of the automated reporting points which are currently not up to date through either:

- Locations which require minor changes;
- Changes made since the previous monitoring, not surveyed yet;
- Locations which have been measured but the RPCR not yet produced;

- Locations which have been measured, inaccuracies have been identified, but changes have not been implemented; and
- Locations where an RPCR has been produced and completed but changes have not been implemented.

Table 4-1: Breakdown of Terminal CMPs by Reporting Type

Route	Auto Reporting Accurate		Auto Reporting Lapsed		Manual Reporting		Total
	Number	Percentage	Number	Percentage	Number	Percentage	
Anglia	41	86%	3	6%	4	8%	48
Kent	20	50%	20	50%	0	0%	40
London North Eastern	19	25%	33	43%	25	32%	77
London North Western	69	79%	1	1%	17	20%	87
Scotland	34	59%	3	5%	21	36%	58
Sussex	30	91%	3	9%	0	0%	33
Wessex	34	97%	1	3%	0	0%	35
Western & Wales	44	51%	2	2%	40	47%	86
Total	291	63%	66	14%	107	23%	464

Figure 4-2: Breakdown of CMPs Numbers by Reporting Type

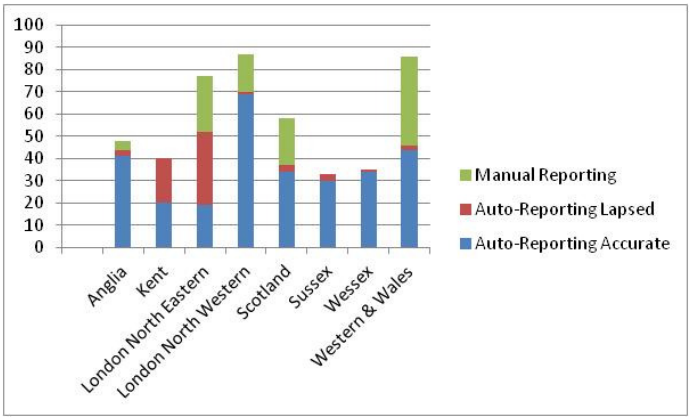
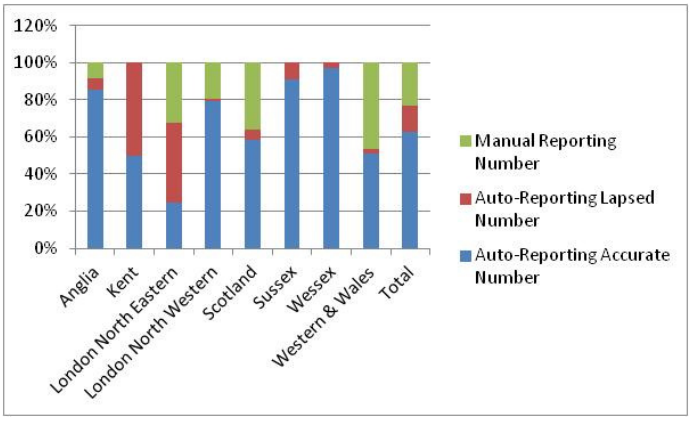
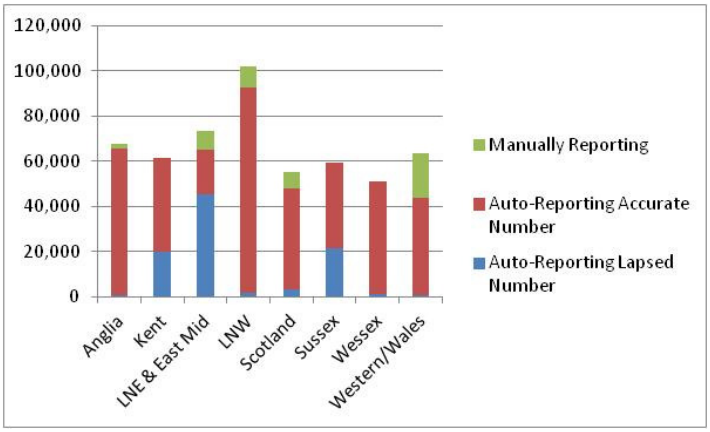


Figure 4-3: Breakdown of CMPs Percentages by Reporting Type



It is clear that not all of the CMPs have the same magnitude of impact on the overall performance data set. The following shows the data analysed by number of trains. This is based on the overall train numbers terminating at a location although without tying services to specific berth offsets within the station. Figure 4-4 shows the number of trains in a period (Period 5 2012/13) linked to the type of reporting and the whether the automatic recording point is 'accurate'.

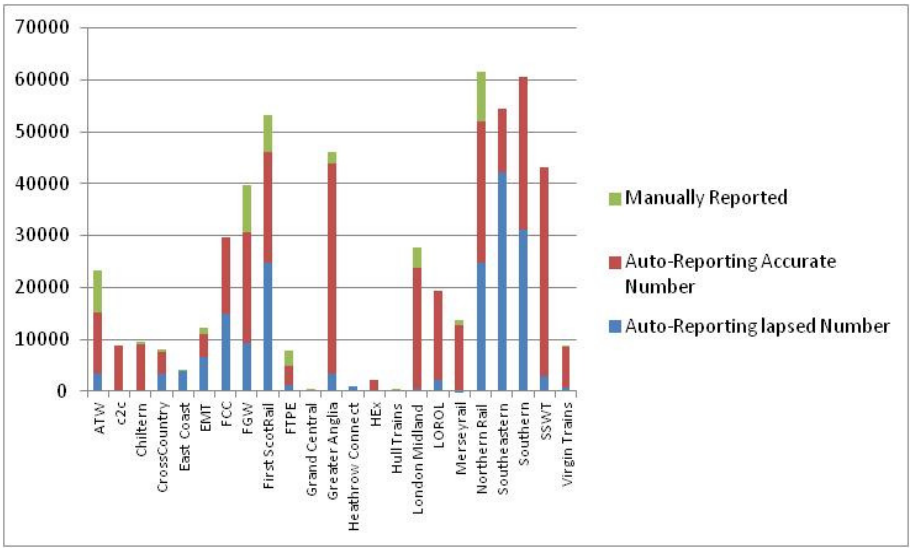
Figure 4-4: Breakdown of Route Services by Reporting Type



This data shows the there is a clear disparity between the 'accurate' and 'lapsed' ('lapsed' meaning offsets which are either beyond their five yearly review or has been subject to some operational change but not yet updated) results on the Routes. The percentage of all CMPs where the auto reporting has 'lapsed' range between 50% and 1%. Given the proposed actions by Network Rail this could be significantly reduced by 2013.

This data can also be considered by TOC.

Figure 4-5: Breakdown of Route Services by TOC



This shows that for some TOCs manually reported locations represent a significant proportion of their arrival reporting.

4.5 Operational Analysis

The fact that trains are being reported at what is termed lapsed monitoring points does not necessarily mean that the reporting is incorrect. The study has looked at this on a number of levels.

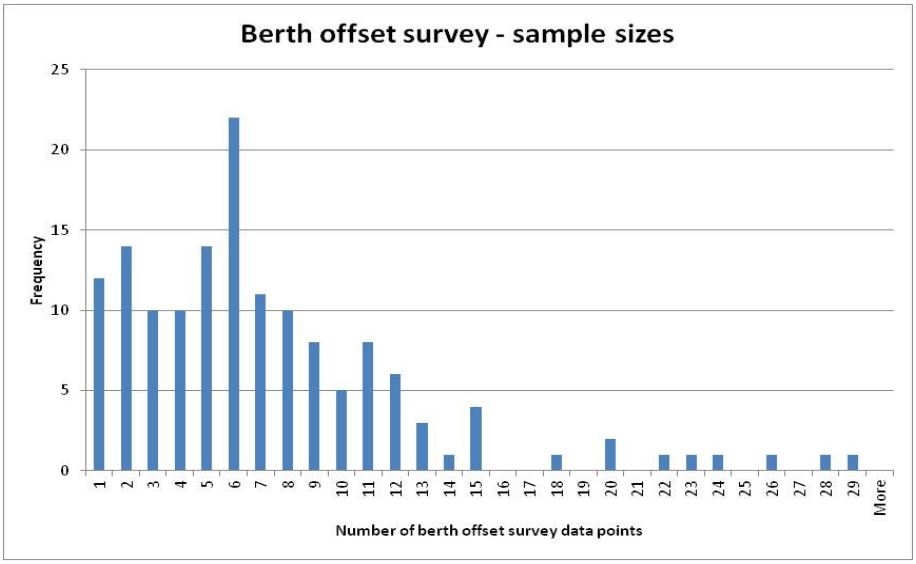
4.5.1 Proportion of Offsets at a Location

Within a location which is classified as having 'lapsed' offset values this does not necessarily apply to all offsets at a multiple offset site. This may be because a change has occurred to the rolling stock using a particular platform, or a speed increase has occurred affecting a particular approach route. Thus, again, not all trains at such a location will be using lapsed offsets. In order to come to a view on the impact this may have the study has looked at a sample of the RPCR forms covering the past year. In each case we have noted the overall number of offsets at each location and identified how many have been updated in each case. From this the proportion of offsets reviewed at each location has been identified. This averages out at 35%. Based on an assumption of equal distribution of the trains at the location there are 35% of trains utilising offsets which are the subject of reviews where a location is class as 'lapsed'.

4.5.2 Sampling Effects

As previously stated, the determination of the offsets at a location is derived through a sampling of the actual train movements on site. In order to understand the impact of the sampling process the study has considered the available data from recent RPCR forms which document the number and individual timings associated with the site surveys.

Figure 4-6: Distribution of Survey Sample Size (from RPCR Forms)



The size of the samples taken during the surveys varied with the frequency of the individual movement and its accessibility to be observed during a survey. At busy locations getting the required sample size for a particularly popular movement could be achieved with ease. However, validating the measurement for a highly infrequent offset may result in only one measurement being taken. In discussions with the PDQSs their approach relied on a combination of matching the proportions of the movements with the practicality of gaining meaningful results.

The study has examined the available data, in terms of the sample RPCR forms from the past year supplied by Network Rail, which show examples of the sampling process and their respective results. From this a distribution of the results has been produced which has then been applied to the overall train arrival distribution based on Period 5 2012/13 to identify those trains which are statistically likely to have been wrongly reported. It has been assumed, in the absence of more detailed information, that the distribution of survey results from the RPCRs that have been reviewed are typical of those applying nationally given that they represent a random sample of locations. The way in which the assessment has been done is described in Appendix B.

Table 4-2 shows the impact the sampling methodology has on the accuracy of the arrival data. This is the percentage of trains which are potentially wrongly reported.

Table 4-2: Proportion of Trains Impacted as a Result of Sampling Process at Automatic Recording Points by TOC

Train Operator	Percentage Impact
First TransPennine Express	
Greater Anglia	
Northern Rail	
Heathrow Connect	
First Great Western	
First Capital Connect	
CrossCountry	
London Midland	
London Overground	
East Midland Trains	
First ScotRail	
East Coast	
Merseyrail	
Virgin Trains	
Arriva Trains Wales	
Chiltern	
c2c	
Southeastern	
Southern	
South Western Trains	
Franchised	1.74%
NON-FRANCHISED TOC	
Grand Central	
Heathrow Express	
Hull Trains	
Non-Franchised	1.51%

Based on the relationship between TOC and Sector it is possible to convert these results to Sector level (in Table 4-3). This takes account of the proportion of services operated by each of the TOCs.

Table 4-3: Proportion of Trains at Auto-Reporting Locations Impacted by Survey Sampling Process by Sector

Sector	Percentage of Services Impacted by the Survey Sampling Process
Long Distance	1.61%
London & the South East	1.75%
Regional	1.55%
Scotland	
England & Wales	1.68%
National	1.74%

4.5.3 Right Time Impact

The final level of data analysis considers the number of services which will have their arrival time wrongly reported as a result of the process. This builds on the portion of offsets at each location which are termed as lapsed, and the level of variation which applies.

This analysis looked at the number of trains whose arrival time is recorded at a lapsed berth where the variation between the old and new offset times would be such that a train would be recorded as 'right time' when it wasn't or vice versa.

The assessment of level of variation in the offsets was derived from the sample of RPCR forms. These covered twenty-eight locations and covered the surveys of 147 offsets. In each case the lowest and highest times recorded on site were compared to the original offset and the percentage variations noted. Adopting this practice for the sample population an average variation was then calculated.

The range of variations to the original offset were between +24% and -5%

The larger list of around 1,700 offsets which encompassed all offsets from each lapsed CMPs was then examined. For each of these offsets the change in seconds from the existing offset was calculated based on the determined range. These were then averaged for the particular CMP location to give a location specific offset error (both positive and negative).

Using the Period 5 2012/13 train performance data a typical distribution of arrival times was developed for each TOC by location. This allowed an estimate of the number of services arriving at each location in steps of one minute. Using the location specific potential offset errors the number of services late when actually on time; and on time when actually late were identified. The estimated number of services at each location was then summed to give a total value for each TOC.

Table 4-4 summarises the results of the proportion of trains in each TOC which are affected by the changes. This shows the range of services impacted taking both the high and low variation extremes

Table 4-4: Proportion of Trains Impacted by Lapsed Auto Recording by TOC

Train Operator	Range of Trains Impacted
First TransPennine Express	
Greater Anglia	
Grand Central	
Northern Rail	
Heathrow Connect	
First Great Western	
First Capital Connect	
CrossCountry	
London Midland	
London Overground	
East Midland Trains	
First ScotRail	
East Coast	
Merseyrail	
Virgin Trains	
Arriva Trains Wales	
Heathrow Express	
Chiltern	
c2c	
Southeastern	
Southern	
South West Trains	
Hull Trains	

These figures are small and represent the proportion of trains at locations where there are lapsed offsets whose arrival times (based on historical performance) fall into the time period when the variation in the offset would lead to them being wrongly reported. This takes account of the 65% share of offsets at a location not requiring to be updated when a location is flagged as being 'lapsed'.

Considering these results by Sector gives the percentages in Table 4-5.

Table 4-5: Proportion of Trains Impacted by Lapsed Auto Recording by Sector

Sector	Percentage of TOC Services Impacted by the Invalid Offsets (Worst Case) and Taking Account of Proportion of Offsets Updated at Locations (35%)
Long Distance	0.22%
London & the South East	0.24%
Regional	0.13%
Scotland	
National	0.23%

To derive an overall figure for the level of accuracy which applies to the auto-reporting locations it is necessary to add the relevant percentages applicable to the 'accurate' and 'lapsed' proportions. This, in effect, adds the percentages in Tables 4-2 and 4-4, and those in Tables 4-3 and 4-5. The resulting percentages are shown in Table 7-1.

5 Study Findings – Manual Reporting

5.1 Introduction

This Section of the report provides an account of the findings in relation to the locations where the recording of train arrival is undertaken through manual reporting – that is the arrival times are entered directly into TRUST by the local signaller.

5.2 Network Rail Route Process – Manual Reporting

5.2.1 Manual Reporting of Contract Monitoring Points

The arrangements associated with the reporting of Right Time Performance where there is manual reporting vary significantly throughout the country. The level of accuracy associated with these CMPs may be highly variable. There had previously been a requirement to review the reported times of train arrivals but this requirement has now been removed from the PMM. Currently, the level of monitoring ranges from some Routes which undertake annual reviews, where this can be resourced, to others which do no monitoring or checking of any kind other than to ensure that a report is received for every train.

It was found that most of these manual reporting locations involve a signaller in some capacity or other. They are managed (generally) by a Local Operations Manager who may have formalised review arrangements with signalling staff to ensure the integrity and completeness of the manual train reporting.

Of the eight Routes interviewed:

- one had a formalised methodology for surveillance and review;
- two had aspirations to do more but were constrained by resources; and
- the remainder accepted the current offsets and had no plans to undertake any form of review.

Whilst two Routes were able to confirm that signaller briefing had recently focused on train reporting requirements and their importance, none of the Routes had visibility of the formalised reporting arrangements implemented by Local Operations Managers. Even the Route with an annual review arrangement in place was unable to confirm whether the procedures for recording and reporting train arrivals were properly documented. It may therefore be necessary to consider the nomination of an individual who would be responsible for the delivery and accuracy of the manual reporting of train arrivals.

It was thus not possible to verify the extent to which ‘local’ procedures have been formulated, and in the absence of any more robust confirmation, the assumption is that such procedures are minimal or non-existent. Thus the findings were that the manual CMPs are unmonitored and unaudited, save for ensuring that a report (however accurate) is made for each train each day.

To understand the scale of the issue Table 5-1 shows the split in the individual Routes of manually reporting CMPs and the associated terminating trains.

Table 5-1: Manual Reporting Locations and train Services by Route

Route	Percentage of Manually Reporting Locations	Percentage of Trains Terminating at Manually Reported Locations
Anglia	4%	3%
Kent	0%	0%
London North Eastern (including Midland)	23%	11%
London North Western	16%	9%
Scotland	20%	13%
Sussex	0%	0%
Wessex	0%	0%
Western (and Wales)	37%	31%
National	**23%**	**9%**

These numbers are gradually reducing as re-signalling takes effect, and the pace of change will quicken over Control Period 5 as Network Rail implements its Traffic Management strategy for signalling control. Whilst it is easy to suggest that these manual reporting points are on branch lines where train numbers are low and thus do not contribute significantly to the overall accuracy of RTP this is not always the case particularly on the Western and Wales Routes. In addition, the number of the reporting points in some Routes means that they cannot be ignored as being irrelevant to the overall accuracy of the measure. In the meantime, the apparent inadequate arrangements for ensuring data accuracy and reliability mean that overall quality of the performance data may be being compromised.

5.3 Manual Reporting

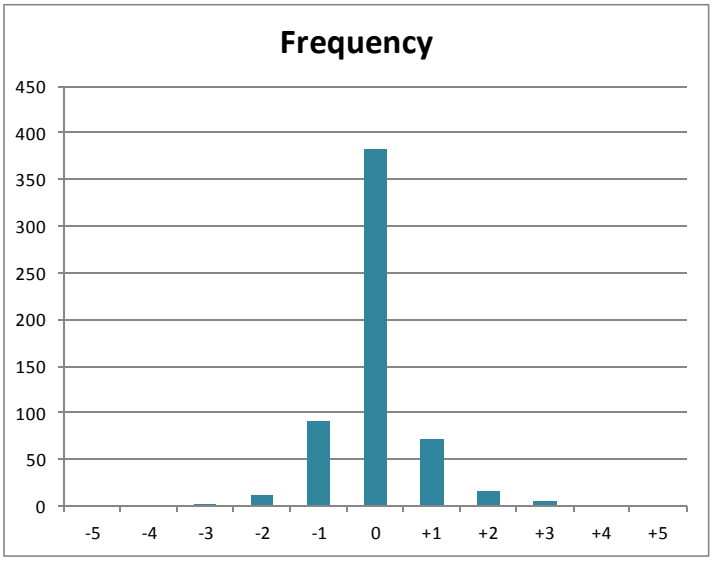
The processes associated with the use and monitoring of manual reporting locations are described in Section 5.2.

It is noted that the Margin Books provide no account of the manual reporting locations. It has been determined that in the absence of comparative data it would be necessary to undertaking an on-site assessment of the accuracy of the manual reporting of trains. This data are already in existence from work undertaken by Sheffield University as a direct commission for a number of TOCs as part of their review of train performance. The study team was granted access to such data for specifically nominated TOCs – Arriva Train Wales, First Great Western, First ScotRail, and Northern Trains. Each of these organisations approved the use of their respective data in the development of this study.

The data which were provided covered the sampling of arrival times undertaken by the University at the respective manual reporting locations. This provided some five hundred samples overall.

The two elements which made up the level of 'error' in the automated reporting locations are not present when considering manual locations. There is no berthing offset and thus no averaging or lapsed error. Instead, the review concentrated on determining the difference between the entry which the signaller applies to the train arrival compared to a rounded version of the exact arrival time (recorded on site to the second). This provides a direct comparison between the recorded time in TRUST against the time which a passenger may perceive the train arrived at. Whole minutes have been used as the measure here because it represents the limit of accuracy which the signaller could be expected to apply when inputting the data to TRUST. As a result this will produce either a zero or whole minute variation between the recorded and actual arrival times.

Figure 5-1: Distribution of Variations in Manually Reported Arrival Times



These have been averaged over the whole dataset to produce a mean variation and range which have then been applied to the overall population of arrivals (as used in the auto reporting assessment) to determine which are likely to trip over the one minute threshold.

The output of this analysis is shown in Table 7-2.

6 Conclusions

6.1 Introduction

This section uses the findings of the study to develop a series of conclusions. These are presented under what are considered to be the key headings. In turn these lead on to the Confidence Gradings described in Section 7.

The study aimed to cover the process and practicality of the derivation of Right Time Performance data across the country. In doing so, and as described in Section 4 and 5, a number of issues have emerged. The following seeks to capture the key conclusions of the study based on the findings of the various meetings and analysis which has taken place. The practical issues associated with the gathering of the data are presented, not in terms of a single Route's approach, but rather as the process is worked nationally. However, this is clearly influenced by the National Centre and individual Route findings.

6.1.1 Process

In terms of the locations where the RTP measures are based on auto-reporting it is clear that there is a well understood process which is being followed by a group of experienced individuals in the Routes.

The time taken to cycle from the identification of a need to re-survey a site to the point where it is signed-off is considered to be very long for what would appear to be a relatively straight forward procedure. The delay to any update represents a period of time when potentially the offset is wrong. This clearly compromises the integrity of the data particularly if a significant number of sites are in the review process at any one time.

It is concluded that whilst there is a clear understanding of the process and its drawbacks it was felt that it was too early to say that the issues had all been solved. Whilst work has clearly been done to advance the commercial settlement issues, there remain issues associated with obtaining the sign-off for changes to the offsets and the length of time this can take.

6.1.2 Backlog

There has been significant progress in reducing the backlog of updated berth offsets. Network Rail National Centre has set a target of April 2013 for the whole backlog to be cleared. Given the progress to date and in particular the efforts of the National Centre Performance Process and Controls team there is no evidence to suggest that this will not be achieved given the plan and revised processes which are in place and providing there is co-operation from all TOCs.

6.1.3 Resources

The issue of resources surfaced regularly during the Route meetings. It was no surprise that the Route with the dedicated individual looking after the data had the best grasp on the process and was clearly on top of the reviews and the progression of the updates. More commonly however, the work associated with

the berth offsets is part of a wider job description. In addition, because of the less immediate nature of the work to keep the offsets up to date, where there were competing pressures in the role, updating the offsets tended to get pushed down the priorities.

Without exception, the individuals carrying out these roles had been in post for several years and are knowledgeable on the subject. There was very little evidence of any form of succession planning or holiday cover in the majority of the Routes. This combined with competing pressures, means that there is a risk that keeping up to date may be a start-stop affair. Further, the reliance on one specialist means that the individual Route's compliance is vulnerable to staff moving on, or going on long-term sick leave. This increases the risk of the data being unreliable at any given time.

6.1.4 Manual Reporting

The processes associated with the recording of arrivals in manually reporting areas appear to be very poor. This is because the majority of the PDQS could not demonstrate any process, and were not able to confirm the accuracy of the means of measuring arrivals as they had not been checked for some time.

In one Route the study team saw copies of documentation covering checks which had taken place on an annual basis and this gave confidence that the relevant timings were as accurate as the process would allow. Even in this Route there was no evidence of documentation to support the means by which the reporting should take place including, for example, written instructions for the signallers. This was felt to be a significant omission leading to a conclusion that the rules governing the recording of the train arrivals appeared somewhat ad-hoc.

The removal of the need to audit the manual recording offsets has also effectively allowed the PDQSs to ignore the five yearly reviews or address any changes to the infrastructure, or train set formations. This cannot be good for data accuracy. The view that the manual reporting locations account for only a very small proportion of the train movements is not valid since on some routes they account for up to 37% of the CMP locations and up to 31% of the train movements. Clearly if these sites are misreporting then the data for the Route and TOCs operating over it are significantly compromised.

The use of data from Sheffield University and with the agreement of the respective TOCs it has been possible to come to a view on the accuracy of the manually reported locations. The analysis which has been undertaken is described in Section 7.3.2.

6.2 Summary

In summary the conclusions of the study are:

- For automatic reporting there is a documented process to develop the berth offsets but this is compromised by the length of time taken to include updates (although it is acknowledged that this has reduced significantly of late);
- There is a high risk that once the current National Centre push to clear the backlog of berthing offset updates is concluded the system could revert to type;

- The staff resourcing to maintain data quality are susceptible to competing pressures with little spare coverage; and
- Manual reporting would appear to be largely ill defined in terms of process and accuracy.

7 Confidence Grades

7.1 Introduction

There is a requirement under the Mandate covering the study to provide a Confidence Grading covering the system reliability and data accuracy associated with the generation of Right Time Performance data. This Section provides a commentary on the Confidence Grades which has been awarded and the justification for the grading.

The definition of the respective grades in the Confidence Grades is included in Appendix A.

7.2 System Reliability

7.2.1 Auto Reporting

The key factor in the determination of whether there is accurate data available on which to base RTP for auto reporting locations is the berth offset. Specifically it is the addition of the offset which determines the accuracy or otherwise of the arrival data.

As described in Section 4 there are three key issues associated with the data at present. These are:

- the delay in updating the data;
- the uncertainty surrounding the manual reporting point accuracy; and
- abnormal or disrupted working close to the final signal berth.

All these are considered to be a function of the currently defined processes. In the first case, the delay means that out of date offsets are being used when new data are available. There is also the process to determine the offset which is based on surveys and then averaged findings. Finally, in the case of abnormal working, this will invariably extend the time that trains take to reach their termination point and may result in trains being reported as 'right time' when they are actually one or two minutes late.

The significance of these issues means that the system is currently classified as Reliability Grade 'B'. This is as a result of the fact that whilst there is a strong well documented process in place there was evidence that there was a backlog of lapsed measures which were shown to be contributing albeit in a relatively minor way to inaccuracies in the reporting.

7.2.2 Manual Reporting

As described earlier the processes surrounding the manual recording of arrival times appear to be somewhat vague or missing. No evidence was produced during the course of the study to demonstrate the presence of documented processes covering manual reporting. In addition, the removal of the requirement to validate these measures, understood to be driven by the resource implications and practicality, has further clouded the governance of the processes. In the

absence of evidence to the contrary the manual reporting regime is currently classified as Reliability Grade 'D'.

7.2.3 Overall Reporting

In order to develop an overall assessment for the processes present to support RTP it is necessary to combine the grades awarded to the automatic and the manual reporting. The significant number of manual reporting locations (23%) which between them represent 9% of train arrivals means that the poor grading for the manual reporting processes cannot be ignored. However, it is clear that the overwhelming majority of the arrivals rely on the auto reporting systems. Nevertheless it has been decided to award an overall Reliability Grading of 'C'.

7.3 Accuracy Grading System

7.3.1 Automatic Reporting

The analysis which has been undertaken on the data is described in Sections 4.4 to 4.6 of the report.

Table 7.1 summarises the study's findings with regard to the various categories required in the project brief. This analysis only covers the auto-reporting locations. It is based on the data from the RPCR forms showing the sampling process and variations in offsets at the time of change married to the actual train arrival data for Period 5 2012/13. Combining these data sets allows the assessment of the trains which are on the cusp of right time which could be impacted by either the offset determination process or those offsets which are in a 'lapsed' state.

Table 7-1: Automatic Recording Accuracy Grading by Measure

Measure	% Error	Accuracy Grading
NATIONAL		
Great Britain	2.00%	2
England and Wales	1.94%	2
Scotland		
SECTOR		
Long Distance	1.84%	2
London & the South East	2.03%	2
Regional	1.72%	2
Scotland		
FRANCHISED TOC		
First TransPennine Express		
Greater Anglia		
Northern Rail		
Heathrow Connect		
First Great Western		
First Capital Connect		
CrossCountry		
London Midland		
London Overground		
East Midland Trains		
First ScotRail		
East Coast		
Merseyrail		
Virgin Trains		
Arriva Trains Wales		
Chiltern		
c2c		
Southeastern		
Southern		
South Western Trains		
Franchised	2.00%	2
NON FRANCHISED TOC		
Grand Central		
Heathrow Express		
Hull Trains		
Non-Franchised	1.56%	2

7.3.2 Manual Reporting

Based on the analysis of the data obtained from Sheffield University, taking an overall average of those results, and applying these to the Period 5 2012/13 arrival data population (as used in the auto reporting review) the accuracy gradings shown in Table 7.2 have been derived.

Table 7-2: Manual Recording Accuracy Grading by Measure

Measure	% Error	Accuracy Grading
NATIONAL		
Great Britain	4.93%	2
England and Wales	4.37%	2
Scotland		
SECTOR		
Long Distance	4.41%	2
London & the South East	4.91%	2
Regional	4.19%	2
Scotland		
FRANCHISED TOC		
First TransPennine Express		
Greater Anglia		
Northern Rail		
Heathrow Connect		
First Great Western		
First Capital Connect		
CrossCountry		
London Midland		
London Overground		
East Midland Trains		
First ScotRail		
East Coast		
Merseyrail		
Virgin Trains		
Arriva Trains Wales		
Chiltern		
c2c		
Southeastern		
Southern		
South Western Trains		
Franchised	4.63%	2
NON FRANCHISED TOC		
Grand Central		
Heathrow Express		
Hull Trains		
Non-Franchised	3.20%	2

7.3.3 Combined Accuracy

Based on the findings of the separate auto and manual reporting accuracy analysis Table 7-3 has been produced which shows the overall accuracy score. The combination of the results in Tables 7-1 and 7-2 takes account of the respective numbers of trains reported by each process.

Table 7-3: Combined Accuracy Grading by Measure

Measure	% Error	Accuracy Grading
NATIONAL		
Great Britain	2.23%	2
England and Wales	2.14%	2
Scotland		
SECTOR		
Long Distance	2.20%	2
London & the South East	2.11%	2
Regional	2.18%	2
Scotland		
FRANCHISED TOC		
First TransPennine Express		
Greater Anglia		
Northern Rail		
Heathrow Connect		
First Great Western		
First Capital Connect		
CrossCountry		
London Midland		
London Overground		
East Midland Trains		
First ScotRail		
East Coast		
Merseyrail		
Virgin Trains		
Arriva Trains Wales		
Chiltern		
c2c		
Southeastern		
Southern		
South Western Trains		
Franchised	2.23%	2
NON FRANCHISED TOC		
Grand Central		
Heathrow Express		
Hull Trains		
Non-Franchised	1.68%	2

7.4 Wider Implications

This study has focused on the current processes and accuracy of the arrangements associated with the reporting of RTP. It is clear however that some of the findings may have implications on the processes associated with other performance measures. These are likely to be the subject of further Mandates and have not been considered as part of this commission.

7.5 Comparison with Benchmark

Table 7-4 provides a summary of the current benchmarks and attained grading based on the current study.

Table 7-4: Comparison of Grading with Benchmark

Measure	Benchmark Grade	Attained Grade
National RTP	B1	C2
Sector RTP	B1	C2
TOC RTP	B2	C2

8 Recommendations

Based on the findings of the study the following recommendations are proposed:

Table 8-1: Reporter Recommendations

Number	Recommendation to Network Rail	Location in Text	N	Due Date
2012RTP01	Limitations to the length of time permitted for TOC sign-off of the RPCR should be investigated based on deemed acceptance if no response is received after a specified duration taking account of the practicality of gaining such agreement	4.3.5	Industry Wide Acceptance Required	June '13
2012RTP02	Processes should be put in place to ensure that named posts in the Network Rail organisation are accountable for the accurate reporting of train arrival times where manual reporting is required	5.2.1	Stephen Draper	May '13
2012RTP03	The means of recording manual reporting arrivals should be appropriately documented	5.2.1	Stephen Draper	May '13
2012RTP04	The Routes should identify suitable coverage for the PDQS post to maintain the RTP data in times of absence. Consideration should be given to succession planning.	4.3.7	Stephen Draper	March '13
2012RTP05	There are variations in the way in which the Routes undertake the work associated with RTP data. It is recommended that the individual PDQS meet to share and exchange views on a six monthly basis and to check the applicability of the current guidance in the PMM.	4.3.4 5.2.1	Stephen Draper	March '13

9 Glossary of Terms / Abbreviations

CMP	Contract Monitoring Point
DAB	Delay Attribution Board
DRP	Delay Recording Points
FOC	Freight Operating Company
PDAC	Performance Data Accuracy Code
PDQS	Performance Data Quality Specialist
PMM	Performance Management Manual
PPM	Public Performance Measure
RPCR	Recording Point Change Request
RTP	Right Time Performance
SMART	Signal Monitoring and Reporting of Trains System
TOC	Train Operating Company
TRUST	Train Reporting System

Appendix A

Study Scope

Mandate for Independent Reporter Part A – Review of right time performance data

Audit Title:	Review of right time performance data
Mandate Ref:	AO/033
Document version:	Draft
Date:	11 July 2012
Draft prepared by:	Chris Fieldsend
Remit prepared by:	Chris Fieldsend
Network Rail reviewer:	Stephen Draper

Authorisation to proceed

ORR	Chris Fieldsend	
Network Rail	Stephen Draper	

Purpose

This mandate sets out the scope of work for the Part A Independent Reporter (Arup) to review right time performance (RTP) data. ORR has made a commitment to publish RTP data and it is therefore critical that the quality of the data can be reported. This will enable stakeholders and the public to make informed decisions based on the accuracy and reliability of the data.

Background

The ORR believes that RTP data provides the industry and the public with one of the most transparent measure of performance. ORR made a commitment in its 2012-2013 business plan to publish RTP data, and DfT support the move to disseminate the data as soon as possible.

ORR recognises that RTP data has not been independently quality assured, and believe the accuracy and reliability of disaggregated data should be known before it is published. This view is shared by the National Task Force and the Performance Delivery Group.

Scope

This review will assess the accuracy and reliability of the RTP measure, expressed as a moving annual average. RTP measures the performance of individual trains against the published planned timetable for the day, and shows the percentage of trains ‘right time’ compared to the total number of trains planned. A train is defined as ‘right time’ if it arrives at its planned destination station early or within one minute (i.e. 59 seconds or less) of the planned arrival time against the public timetable and does not miss any of its booked stops.

The review should assess the quality of the RTP measure at the following levels:

1. National (Great Britain, England & Wales and Scotland)
2. Sector (London & South East, Long Distance, Regional and Scotland)
3. Train Operating Companies (TOC; Franchised and Non-Franchised)

The review should:

- comment on the reliability, quality, consistency, completeness and accuracy of the reported data;
- present a confidence grade for the RTP measure at each of the three levels listed above (national, sector and TOC);
- make a judgement for each TOC on a likely confidence range of reported data; and
- make appropriate recommendations for the betterment of the RTP measure which could be undertaken at minimal cost.

Methodology

The Reporter should:

- outline their proposed methodology to assess the specific requirements listed above;
- meet with relevant Network Rail employees in each route and at Milton Keynes, to understand the process for collation, storage and computation of the RTP measure, along with the frequency of which it is reviewed;
- carry out a quick review of all relevant documentation and systems (including TRUST, margins books and berthing offsets), and comment upon their quality and fitness for purpose;
- identify the likely accuracy of reporting right time at a location dependent on how the data is captured – report on contributing factors which effect the variability of the accuracy of data;
- propose a statistically robust method of reviewing the RTP measure at TOC level;
- meet with representatives from all TOCs to understand any differences in collation methods;
- review any data quality issues arising from different signalling (manual and automatic) methods across the network; and
- state the confidence that ORR / NR can have in the findings, given their proposed methodology.

The Reporter should draw on (and not duplicate) work previously undertaken in its review of PPM and CaSL. Network Rail will share existing data with the Reporter on the variability of individual train performance. The Reporter should utilise this data to avoid any unnecessary data collation.

Deliverables

The Reporter should provide a publishable report, including findings, conclusions and recommendations. The report should be prepared in draft form and sent electronically to Network Rail and ORR, at the same time. The Reporter should facilitate feedback (via a tripartite feedback session if appropriate) and provide a revised report with track changes. This should be followed by a final report for publication on ORR's website.

Timescales

A fully costed proposal for this work is required by 19 July 2012. Work is expected to commence shortly after following approval by NR and ORR. A draft report is required by 14 September 2012 and a final report is required by 12 October 2012.

Independent Reporter Remit Proposal

The Independent Reporter shall prepare a fully costed proposal for review and approval by NR and ORR on the basis of this mandate. The approved remit will form part of the mandate and shall be attached to this document.

The proposal will detail methodology, tasks, programme, deliverables, resources and costs.

Confidence Grades

The Independent Reporter shall provide a confidence grade for the RTP measure at national, sector and TOC level. The confidence grading system in Annex A should be used. For each measure, the Independent Reporter should include the:

- confidence grade for this review;
- commentary on the grade against ORR's benchmark; and
- an indication of the highest achievable grade at each level.

Annex A: Confidence Grading System

System Reliability Grading System

System Reliability Band	Description
A	Appropriate, auditable, properly documented, well-defined and written records, reporting arrangements, procedures, investigations and analysis shall be maintained, and consistently applied across Network Rail. Where appropriate the systems used to collect and analyse the data will be automated. The system is regularly reviewed and updated by Network Rail's senior management so that it remains fit for purpose. This includes identifying potential risks that could materially affect the reliability of the system or the accuracy of the data and identifying ways that these risks can be mitigated. The system that is used is recognised as representing best practice and is an effective method of data collation and analysis. If necessary, it also uses appropriate algorithms. The system is resourced by appropriate numbers of effective people who have been appropriately trained. Appropriate contingency plans will also be in place to ensure that if the system fails there is an alternative way of sourcing and processing data to produce appropriate outputs. Appropriate internal verification of the data and the data processing system is carried out and appropriate control systems and governance arrangements are in place. The outputs and any analysis produced by the system are subject to management analysis and challenge. This includes being able to adequately explain variances between expected and actual results, time-series data, targets etc. There may be some negligible shortcomings in the system that would only have a negligible affect on the reliability of the system.
B	As A, but with minor shortcomings in the system. The minor shortcomings would only have a minor effect on the reliability of the system.
C	As A, but with some significant shortcomings in the system. The significant shortcomings would have a significant effect on the reliability of the system.
D	As A, but with some highly significant shortcomings in the system. The highly significant shortcomings would have a highly significant effect on the reliability of the system.

Notes:

1. System reliability is a measure of the overall reliability, quality, robustness and integrity of the system that produces the data.

2. Some examples of the potential shortcomings include old assessment, missing documentation, insufficient internal verification and undocumented reliance on third-party data.

Accuracy Grading System

Accuracy Band	Description
1*	Data used to calculate the measure is accurate to within 0.1%
1	Data used to calculate the measure is accurate to within 1%
2	Data used to calculate the measure is accurate to within 5%
3	Data used to calculate the measure is accurate to within 10%
4	Data used to calculate the measure is accurate to within 25%
5	Data used to calculate the measure is accurate to within 50%
6	Data used to calculate the measure is inaccurate by more than 50%
X	Data accuracy cannot be measured

Notes:

1. Accuracy is a measure of the closeness of the data used in the system to the true values.

2. Accuracy is defined at the 95% confidence level - i.e. the true value of 95% of the data points will be in the accuracy bands defined above.

Benchmark Grades

As agreed with Network Rail, from Q3 2011-2012 data assurance reviews have been using this new confidence grading system. A characteristic of the new system is the introduction of a benchmark grade; the grade at which ORR believes the measure should be, given what we know about the processes and level of subjectivity in deriving it. It should be noted that the derivation and application of benchmark grades has recently been introduced, and all parties should decide how useful this element is throughout the review. The table below provides ORR's benchmark grades for the 2011-2012 data assurance review of RTP data.

Measure	Benchmark Grade
National RTP	B1
Sector RTP	B1
TOC RTP	B2

Appendix B

Sampling Impact Determination Methodology

The following described the process applied to the determination of the impact the data sampling has in the setting of the offsets.

1. For the berth changes where sample data is available we have looked at the distribution of survey results, to determine the 5% and 95% confidence levels – measured in seconds
2. For each location, made up of one or more samples, we have pulled out the **mean** confidence level, and also the **maximum** confidence level – again these are in seconds. In a few cases a single berth offset (potentially one which is not used that frequently) was significantly higher, hence taking the mean as well as the maximum for comparison.
3. Averaging these across all locations gave a mean of
    a. + / - 7 seconds (mean value of confidence levels)
    b. + / - 13 seconds (max value of confidence levels)
4. Using the Period 5 2012/13 arrival time distribution we then looked at the trains arriving either side of the right time plus one minute mark (in second intervals)
5. Based on this we then worked out the proportion arriving late/on-time which should have been classified differently in the 7s and 13s cases. This was done for each TOC.
6. These results were then added (taking account of train numbers) to create the Sector and national results.
7. In the tabulations the 95% confidence level figures are used.