Office of Rail Regulation and Network Rail

AO/039: Review of Performance Measures

Interim Report

Issue | 2 July 2013

Job number 223767-21

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

ARUP

Contents

Page

1 Executive Summary 1

2 Introduction 4

2.1 Background 4
2.2 Objectives and Scope of Review 4
2.3 Audit Methodology 5
2.4 Structure of Report 5

3 Progress on Recommendations 7

3.1 April 2011 Review (Q4 2010/11) 7
3.2 Mandate AO/033 – Right Time Performance Review (December 2012) 13

4 Performance KPIs 16

4.1 Introduction 16
4.2 Data Management Processes & KPI Production 16
4.3 Data Assurance in the Context of the Right Time Performance Report – PPM & CaSL 22
4.4 Confidence Ratings 28

5 Delay Attribution Review 30

5.1 Introduction 30
5.2 Methodology 30
5.3 The use of FO/TO Delay Cause Codes 31
5.4 The use of Z* Delay Codes 33
5.5 Rates of Commercial Takeback 34
5.6 Recoded Delays 36
5.7 Delays in Dispute on Day 2 37
5.8 The level of Roll Up Delays 38
5.9 Management Incident Levels 39
5.10 Overview 41
5.11 Recommendations for Phase Two 42
5.12 Other possible reviews 42

Appendices

Appendix A

Mandate of Review

Appendix B

Spot Checks of Sample TOC PPM & CaSL Data

Appendix C

Data Assurance in the Context of Right Time Performance Matrix

1 Executive Summary

This commission requires the Reporter to review a series of KPIs and measures produced by Network Rail for the ORR, to ensure their correctness. Also, the Right Time Performance Measure study (RTP) in December 2012 raised concerns about the reliability of an undocumented process for manually recorded data. The report stated that there could be implications for the reliability of PPM and CaSL data, and this review seeks to establish whether this is the case. ORR also requested that the Reporter team carry out an investigation into whether the quality of delay attribution had deteriorated during CP4.

Recommendations from the April 2011 Performance report, and the December 2012 Right Time Performance Measure report, have been reviewed and are substantially complete.

For cancellations data, NR has implemented a revised reconciliation procedure for TOC-provided data, to ensure that differences between NR and TOC data are always positively investigated.

In respect of both PPM and CaSL, the Reporter team have reviewed the end-to-end calculation process through the spreadsheets provided by Network Rail. All linkages remain formulae based, with the only area of hard coding remaining as the input of TOC supplied cancellation data. A series of spot checks, based on a sample of TOCs, have been carried out on the data flow from extraction from PSS through to final figures, to ensure this is accurate. These spot checks are summarised in Appendix B and show no concerns. On the basis of these checks, the Reporter team are content that the data is still flowing through correctly.

Following the publication of Arup's Right Time review[1], we have assessed the implications of the identified timing errors on PPM and CaSL. This follow-up review uses the key findings from the Right Time review and performs a statistical analysis to measure the impact on PPM and CaSL. The results from the statistical analysis show that the expected impact is negligible. The conclusion is that the identified errors in the timing recordings have an immaterial impact on the national PPM and CaSL metrics.

The ratings for the performance KPIs are outlined in the table below, demonstrating any changes since the last review presented in the 2011 Quarter 4 report.

KPI	Original Score	New Score	Comments
5a: PPM	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. As noted previously, with the exception of the use of TOC data, the process is fully automated. The findings of the Right Time study have negligible impact on PPM, and so provide no reason to downgrade this rating.
5b: CaSL	Reliability A Accuracy 2	Reliability A Accuracy 2	No Change. Network Rail are still reliant on TOC

[1] 223767-07 Mandate AO/033 - Right Time Performance Measure Final Report (Complete)

KPI	Original Score	New Score	Comments
			cancellation data for reporting purposes. The provision of TOC data provides confidence in terms of enabling a reconciliation exercise between data held within PSS and by TOCs, but there do remain differences in the datasets. NR only investigate and reconcile differences of over 0.1%. The findings of the Right Time study have negligible impact on CaSL.
5c: Network Rail Delay Minutes to TOCs	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. The impact of the Adjusted Data Series calculation has added greater confidence in the accuracy of these figures (see section 4.2.2 of this report).
5d: Network Rail Delay Minutes to FOCs per 100 train kms	Reliability A Accuracy 3	Reliability A Accuracy 3	No Change. The accuracy rating reflects the on-going issues with the collation of the freight mileage data used as the normaliser. NR have now devised a more accurate methodology for calculating freight mileages but have taken the decision to introduce it in April 2014 (see update on recommendation 2010.5.4a in Section 3.1). Once this is fully implemented the measure should be able to rise to A1.
6a/6b: Asset Management (Track / Non Track Delay Minutes)	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. This dataset is a direct derivative of Network Rail delay minutes, and so is reflective of the KPI score for 5c.

In regard to Delay Attribution, ORR requested that the Reporter team carry out an investigation into whether the quality of delay attribution has deteriorated during CP4. The request arose from concerns that reductions in staffing levels or other causes, such as the impact of Alliances between NR and TOCs, may have led to deterioration in the quality of attribution. Data analysis was undertaken for a period before changes were perceived to have occurred (2009/10) and the data from corresponding periods in 2012/13.

The view of the Reporter Team is that the overall levels of delay attribution show *an improvement* between the two review periods. There are no obvious national trends within the data that highlight significant country wide issues of concern. Most of the measures show a positive improvement since 2009/10. This reflects well on the greater focus on data quality now carried out by NR and increased levels of data integrity checks being carried out both at Route and through the central team as defined in the Performance Measurement Manual. The use of periodic data quality checks has assisted the Routes in improving data quality, and the figures shared with the Reporter Team show an improving trend across the country with areas of concern being tackled.

However, the Reporter team have identified certain issues affecting individual routes and operators that are worth further examination. In general these should

be investigated by NR and reported back to ORR where they have not already given details noted within this report.

Given the overall improvements in the areas investigated the Reporter team do not feel it would be sensible use of resources to undertake a detailed follow up (Phase 2) to the specifics highlighted within this report, as originally envisaged at the initial launch meeting.

Although the overall quality of delay attribution has improved, we have identified variations in practice between the Routes. We therefore suggest that the Reporter team explore the emerging arrangements across three Routes alongside their lead TOCs, to understand the drivers behind the variations being observed. It would also give us the opportunity to explore some of the specific changes on these routes observed during CP4 that we have identified in our analysis. The suggested routes for these discussions are:

- Scotland and ScotRail;
- Sussex and Southern; and
- Wessex and SWT ***or*** LNE and Northern.

2 Introduction

2.1 Background

Arup was appointed by the Office of Rail Regulation (ORR) and Network Rail in 2009 to undertake the role of Independent Reporter (Part 'A'). This commission requires the Reporter to review a series of KPIs and measures produced by Network Rail for the ORR, to ensure their correctness. These reviews are undertaken as part of a rolling programme and are reported to the ORR in a series of Quarterly Reports. This report covers the Reporter's data assurance review of Performance Measures, the mandate for which is included in Appendix A.

The last Reporter review of NR's performance data took place in March/April 2011. On Public Performance Measure (PPM), and Cancellations and Significant Lateness (CaSL), the review assessed the confidence ratings as (respectively) A1 and A2. Since that review, the Reporter has undertaken a separate review of Right Time Performance data (RTP) in December 2012. This study raised concerns about the reliability of an undocumented process for manually recorded data. The report stated that there could be implications for the reliability of PPM and CaSL data, and this review seeks to establish whether this is the case.

ORR also requested that the Reporter team carry out an investigation into whether the quality of delay attribution had deteriorated during CP4. The initial request was due to concerns that reductions in staffing levels or other causes, such as the impact of Alliances between NR and TOCs, may have led to deterioration in data quality. At the initial project tripartite meeting it was agreed that this review should be addressed in two phases. The first phase would carry out empirical comparisons between selected data from early in CP4 and recent periods, to test if there are any identifiable trends either at national or Route level that indicated a deterioration in data quality.

The Reporter team will then identify potential hypotheses if the data shows any worsening trends and recommend further, more detailed analysis in a second phase of investigation.

2.2 Objectives and Scope of Review

This review is required to include:

1. PPM – the implications of the findings of Arup's Right Time review[2] (e.g. berthing offsets, manual reporting) on PPM;
2. CaSL – the implications of the findings of the Right Time review (as above) as well as implications on missed intermediate stations;
3. Delay Attribution – ORR would like assurance on whether an apparent reduction in resources available to 'level one' delay attribution (undertaken on shift, in real time) has impacted the quality of delay minute data. To that end, ORR would like a better understanding of the delay datasets (adjusted and unadjusted) and confidence that the data received from Network Rail is of a high standard.

[2] AO/033 – Review of Right Time Performance (RTP) Data

The Reporter should briefly review PPM, CaSL and delay minutes KPIs and:

- comment on the reliability, quality, consistency, completeness and accuracy of the reported data;
- present a confidence grade for each KPI and comment on the change since last reviewed in April 2011; and
- report on progress against recommendations made in both the April 2011 review and the December 2012 Right Time Performance Measures review, and make appropriate recommendations where necessary.

This report covers all of the above objectives with the exception of Delay Attribution, where we report on phase one and recommend work for phase two.

2.3 Audit Methodology

Following the tripartite inception meeting to agree scope and timescales for this review, a number of meetings were established, and various NR train performance datasets were downloaded:

Date	Network Rail Attendees	Purpose	Location
25 March	Performance Support Analyst Performance Process & Controls Manager	Review of Right Time Performance recommendations & progress	Milton Keynes
27 March	Performance Analyst	Core KPI review	Milton Keynes
23 April	Performance Support Analyst Performance Process & Controls Manager	Delay Attribution Findings review	Milton Keynes
20 May	Performance Support Analyst Performance Analysis Manager	Delay Attribution Findings review	Milton Keynes

At these meetings, the specific areas of review were checked and then data was collected to check that the evidence presented could be verified. Analysis of the data provided was undertaken after the meeting, with any clarifications raised directly with the provider. The findings are set out in the following sections.

2.4 Structure of Report

To enable easy comprehension, the Report is structured as follows:

Section 3 - Progress on Recommendations

- April 2011 Review of Performance KPIs
- December 2012 Right Time Performance Measures Review

Section 4 - Performance KPI Review

- Data Management Processes and KPI Production
- PPM, CaSL & Delay Minutes Data Assurance in the Context of the Right Time Performance Measures Review

Section 5 - Delay Attribution Review

In accordance with Reporter review protocol, confidence ratings have been awarded in the Performance KPI Review section of the Report.

3 Progress on Recommendations

3.1 April 2011 Review (Q4 2010/11)

A review of progress on any outstanding recommendations has been undertaken as part of this 2013 review. No recommendations were made as a result of the April 2011 review and no review was undertaken in 2012. However the table below, taken from the April 2011 report, details progress from the 2010 audit, and contains an updated progress for this 2013 report.

Number	Recommendation to Network Rail	NR Data Champion	Due Date	March 2011 Update	May 2013 Update
2010.5.4a	Devise and agree a plan to resolve outstanding freight mileage data issues	Performance Analysis Manager	September 2010	Network Rail has investigated options to improve the collation of mileage data in accordance with the recommendation. They have set out options to remove dependence on legacy systems and provide data through PSS. The plan is to do this by utilising the Train Operator Data System (TODS+). The plan has been created with a target implementation date of end of April 2011. **Closed** **Reporter team to review implementation of plan in next audit.**	The programme of work to ensure that the basis of mileage in PSS and TABS (the billing system) is updated is nearing completion and will be ready for implementation by April 2014. NR decided it was not appropriate to implement in the middle of a year. Work still needs to be done to check any impact of the change on reported train mileage to the ORR data warehouse and also on the freight delay/mile as used within freight reporting – it should not affect the Schedule 8 regime as that is not calculated from PSS / Business Objects. **The original recommendation was to put in place a plan to improve the freight mileage data issues and this is complete but the completed mileage data will require checking in 2014/15** **Closed**
2010.5.4b	Devise and agree a plan to resolve outstanding freight SRT data issues	Freight Performance Team	September 2010	Network Rail has developed a Process Overview document that defines the process for changes to timetable planning rules (which will include Freight SRTs) – "Timetable Planning Rules – Changes and Amendments (Draft)", dated 10/03/11, which was provided to the Reporter team. However, at present there is no definitive evidence of a forward plan for resolving and updating outstanding freight issues, although it was advised that this issue is prioritised to be resolved on each route	A programme of updating freight SRTs is now in place and is managed industry wide through the Freight Forum. Between mid-2012 and March 2013, a total of over 31,000 missing Class 66 SRTs were calculated, validated and updated to B-Plan making them available for Train Planning purposes. On completion, a further check was done of missing Class 66 SRTs and this revealed 1801 records missing to infill all of the regularly required SRTs. The work to fill

Number	Recommendation to Network Rail	NR Data Champion	Due Date	March 2011 Update	May 2013 Update
				alongside the development of the relevant Long Term Plan timetables. **On-going**	these is now being done manually with less than 1000 left to complete. This work is resourced and expected to complete in July. This leaves a 'Business as Usual' procedure as agreed at the Freight Forum with FOCs where diversionary and rarely used routes will be infilled next. Also as part of this process SRT amendments can be proposed in the event of inaccuracies being proved and these will be captured as part of regular timetable change. A similar process is now adopted for new traction such as Class 70s. **The original recommendation was to put in place a plan to improve the freight SRT position and this is complete but a verification check should be undertaken in 2014/15.** **Closed**
2011.5.1	Network Rail should complete the formal documentation of the procedure associated with data export from PSS.	Performance Analysis Manager	September 2010	Network Rail has developed a Process Document covering the entire process calculation process, from data extraction through to result reporting. While the status of this document is currently 'work-in-progress', the content is largely complete, and contains a detailed and prescriptive step-by-step guide through the tasks required to be followed by the performance analyst. **Not complete, finish by May 2011**	Documentation is complete and in full use and shared with the Reporter team. **Closed**
2011.5.2	Network Rail should review policy towards, and the handling of,	Performance Analysis	November 2010	Network Rail has reviewed several aspects of the handling of schedules and delay	The specific issues within this recommendation have now all been fully

Number	Recommendation to Network Rail	NR Data Champion	Due Date	March 2011 Update	May 2013 Update
	severe disruption in its widest sense, including the uploading of emergency timetables.	Manager		management. 1) Day A for Day B – ITPS now has the capability to store contingency timetables which allows emergency timetables to be uploaded on Day A for Day B. 2) Inputting of VSTP's / Emergency Timescales – an enhanced system (Integrale) is in the process of being rolled out to all Network Rail Control's which will enable quicker input of VSTPs / Emergency timescales by Network Rail Control staff. This will be a phased roll out which is scheduled to start in May 2011 and be completed by the end of the year. 3) Managing delay attribution in times of significant perturbation – a prioritised list of responsibilities/tasks for Train Delay Attribution staff to concentrate on during times of severe perturbation has been produced and implemented. As such, Level One staff (Network Rail staff responsible for initial 'real-time' delay attribution) adopt a "phased withdrawal" from certain workload types, to concentrate on the core elements which deliver the best possible outputs for performance reporting purposes. **Not complete, finish by December 2011**	implemented. 1) Day A for Day B is now implemented and is an industry standard requirement. 2) Integrale is now in use across the industry 3) A new attribution process targeted at managing attribution during severe weather was agreed with the Delay Attribution Board and issued as a bulletin across the industry in time for winter 2011/12. **Closed**
2011.5.3	Network Rail should produce a full register of local attribution agreements with TOCs, and work to remove	Network Rail	January 2011	A full register of agreements has been produced. Network Rail has identified 42 such agreements which it has grouped into different types. These are being reviewed systematically to look at which ones should be included in the	The register is still in place but NR is looking to abolish local agreements through negotiation with TOCs. However, there is an increasing pressure to create local practices as a result of devolution and the creation of

Number	Recommendation to Network Rail	NR Data Champion	Due Date	March 2011 Update	May 2013 Update
	them, as part of an effort to reduce data discrepancies.			DAG, which will be required longer term etc. The Performance Manual is being updated to give clear instruction on how agreements should be documented in the May 2011 update of the Performance Manual. NR has complied with the requirement of the recommendation to produce a full register of agreements and is in the process of formalising these into industry documentation (i.e. the DAG). **Closed** **Reporter team to confirm implementation of these actions in next audit.**	Alliances and this is discussed further in section 5 of this report. **Closed**
2011.5.4	Network Rail should review its staffing levels for the management of delay attribution across the network to address the resource imbalance noted in some Routes, ensuring that staff are fully briefed and that briefing records are kept up-to-date, and protect the much improved levels of data integrity seen by the Reporter team.	Network Rail	January 2011	A very detailed review of comparative TDA staffing levels has been undertaken by Network Rail. This has looked at both the volume of TRUST incidents created by Route against staff and quality indicators based on audit criteria. This has helped to produce an overall view on what the optimum staffing requirements are to deliver an acceptable level of quality attribution. The analysis has only recently been completed. The next phase is to decide what will be done with TRUST attribution resources. This decision will be influenced by the overall review being held within NR on how Routes and centralised teams will be staffed in the future. The analysis was shared with the Reporter team but given the sensitivity of the staffing implications copies were not taken. Therefore NR has complied with the requirement of this recommendation to carry out a review, but no changes have yet been	The report that was produced in 2010 was never implemented due to a change in organisational policy. It was initially written with a view to centralising delay attribution before the decision was taken to actually devolve activities to the Routes. **Closed**

Number	Recommendation to Network Rail	NR Data Champion	Due Date	March 2011 Update	May 2013 Update
				proposed or implemented. **Reporter team to review outcomes of this review in next audit.**	

3.2 Mandate AO/033 – Right Time Performance Review (December 2012)

A review was undertaken of the recommendations made in the Right Time Performance Measure study, at a meeting on the 25th March at NR's Milton Keynes offices. The progress made is summarised in the following table.

Number	Recommendation to Network Rail	NR Data Champion	Due Date	May 2013 Update
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.	Industry Wide Acceptance Required	June 13	NR representatives confirmed that NR already has the right to 'deem' Part A of the Recording Point Change Request (RPCR) as accepted by any party failing to sign off within 28 days. Part A signifies acceptance of the revisions to offset values following on site review and recalculation. This is acknowledged in the Right Time Performance Measures (Mandate AO33) report (section 4.3.4). It continues to be difficult to progress the remaining aspects of the Berth Offset change process for a whole variety of reasons. None of the parties to the process are entirely blameless - TOCs, NR Customer Relations Executives etc. - and even references up to Director level, in NR and particular TOCs, have not resolved the problem. Although some progress has been made in clawing back the backlog of berthing offset reviews identified in the RTP report (particularly in Scotland and LNE): • The backlog in Kent has been difficult to progress with South Eastern but is now resolved. • Some TOCs are refusing to sign off any changes, under alleged instructions from DfT. • Certain PDQS are reluctant to adopt the 'Declaration of No Change' process, which allows them to avoid the need to undertake the 5 yearly offset review, even when there has been no actual change in infrastructure or rolling stock. This adds to the workload pressure and generates unnecessary difficulties with the involved TOCs. In an attempt to make progress NR has tabled a proposal at NTF-OG to establish a cross-industry working group to consider the issues, and resolve those which can be resolved (paper not seen and outcome not known). NR has not yet actively considered reference to the Delay Attribution Board or the Access Disputes Resolution Committee. **Status: On-going**
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	Performance Analysis Manager	May 13	These two recommendations both cover the process for recording and reporting arrival data at 'manual' reporting locations. The following arrangements have been implemented:- • Revision to the NR Performance Manual - Section 5.4, Data Capture - Actual Train Times; • Establishment of the template procedure forms on which the Route Performance Measurement Manager or PDQS must document the procedure for reporting trains when the Signaller observes train arrival/departure (Form MR1) and where the Signaller cannot see train arrival/departure (Form MR2); • Nomination of Designated Persons to enact the procedures (which satisfies Recommendation 2012 RTP02); and • Audit arrangements. NR HQ now also has a comprehensive list of manual reporting PPM locations, which itemises, by Route:-
2012RTP03	The means of recording manual reporting arrivals should be appropriately documented	Performance Analysis Manager	May 13	

Number	Recommendation to Network Rail	NR Data Champion	Due Date	May 2013 Update
				• Method of reporting/recording (5 categories); • 4 weekly train count (and % of Route ppm); and • Location specific comments. As the pace of resignalling, and the reduction in numbers of manual reporting points, picks up, the list will be updated and reissued periodically. The information is now stored in a shared drive and the Routes update their own records as and when changes to the infrastructure are made. NR representatives stated that, at last week's Route Performance Measurement Managers' meeting, all the RPPMs confirmed that they had either fully implemented, or were close to fully implementing the revised manual reporting point procedures. However, the Reporter team have not checked this independently. **Status: on-going**
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.	Performance Analysis Manager	March 13	At the aforementioned RPPM meeting, all Routes confirmed that they now have established appropriate cover arrangements for PDQS personnel responsible for maintaining the currency of berthing offsets. **Status: closed**
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.	Performance Analysis Manager	March 13	The revised RPPM meeting agenda, incorporating 'good practice' review is believed to satisfy the intent of this recommendation. **Status: closed**

4 Performance KPIs

4.1 Introduction

This section reviews the reliability, quality, consistency, completeness and accuracy of PPM, CaSL and delay minutes KPIs. It starts by considering the overall processes and data involved in their production, repeating the review we last carried out in April 2011. The intention, as set out in the mandate, is for this to be a high level review. We then consider the implications of the findings from the Right Time review on PPM and CaSL, in particular what impact the errors identified in some of the berthing offsets will have on their accuracy.

Finally, taking into account both the general processes and data, and the findings from the Right Time review, we grade each of the KPIs using the criteria set out in the mandate (see Appendix A of this report).

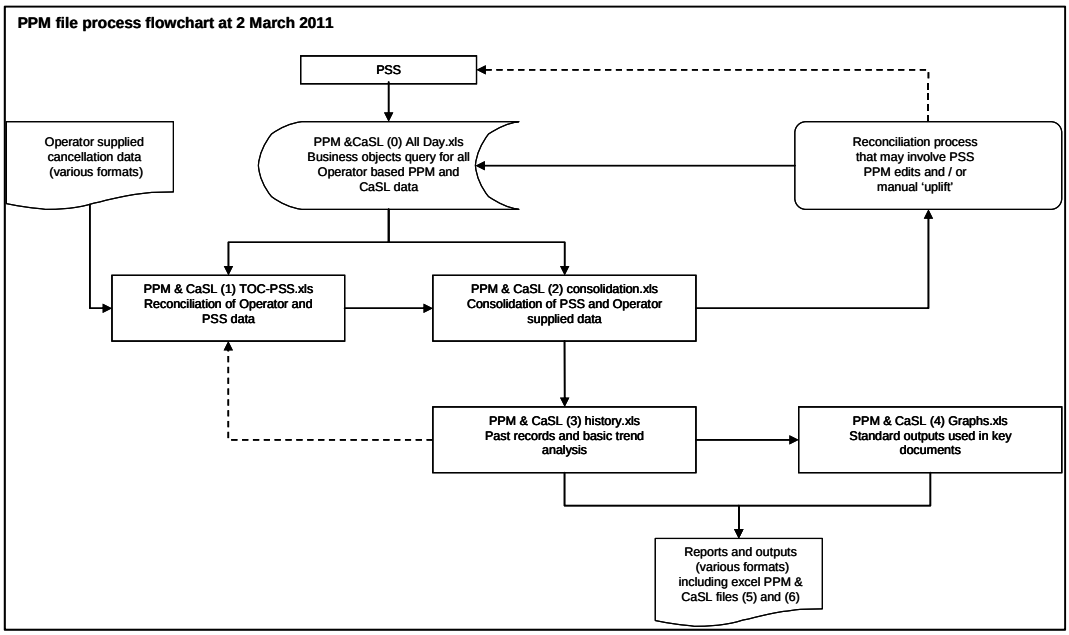
4.2 Data Management Processes & KPI Production

4.2.1 PPM/CaSL

4.2.1.1 Data Flow Process

The process for calculating PPM and CaSL is broadly unchanged to that outlined in the 2010 Quarter 1 report. That report provided a detailed outline of the process, and it is recommended that this section be read with in conjunction with that report.

The process is outlined in the flow chart below, as taken from the Network Rail's 'PPM and CaSL Industry Data Process Note' document. Only two notable amendments have been made to the process since the 2010 Quarter 1 report, which are outlined in detail below.



Reconciliation of Cancellation data

As noted in previous reports, official PPM and CaSL figures are based on cancellation data supplied by TOCs (with the exception of a small minority of TOCs who do not supply Network Rail with cancellation data - since this is not mandated - for which PSS cancellation data is used). This is due to historic concerns over completeness of cancellation data within TRUST, although this was partially alleviated with the SRP77 TRUST upgrade. The outstanding issue is trains failing to stop at a station which is not automatically detected in the data and relies on information from operators.

In the 2010 Quarter 4 previous review, the Reporter noted that the Performance Analyst in Network Rail Milton Keynes was responsible for undertaking a reconciliation check each period between cancellation data in PSS and that supplied by each TOC. Any differences would be reviewed with the TOC, thus providing a check of this data before being carried forward into final reported figures. A description of this reconciliation process was provided in the 2010 Quarter 1 report.

While the process still relies on TOC cancellation data where supplied, following devolution the responsibility for the reconciliation process now lies within the Lead Route for each TOC, rather than at Milton Keynes (with the exception of Arriva Cross Country for whom the Route Performance Manager is based in the National team).

The revised reconciliation process is now as follows in each period:

- The following steps remain as before:
  - The Performance Analyst in Milton Keynes collates cancellation and PPM data from each TOC.
  - The Performance Analyst produces a spreadsheet which compares PSS cancellation and PPM numbers with those supplied by the TOCs.
- The following steps are different to before:
  - For any TOCs where the difference in the TOC-supplied PPM and the PPM calculated from PSS is more than 0.1%, these details are passed to the Lead Route for further review (or the Performance Manager in the National team, in the case of AXC) via a template email, which was provided to the Reporter team.
  - It is the responsibility of the Route to review and reconcile this data in the same way as previously carried out by the National Performance Team (see 2010 Quarter 1 report).
  - If the Route does not respond to the initial request, a prompting email is sent with a clearly defined deadline date for response. This is also in the form of a template email which was provided to the Reporter team.
  - If the Performance Analyst does not get a response from the Route after the prompting email, the Reporter was advised that the TOC supplied cancellation figures are used for PPM reporting.

The threshold of 0.1% for investigation remains as previous. The only change is where responsibility lies for investigating and reconciling differences. This may represent a small risk in terms of ensuring that this reconciliation process always

happens when required. Before, this was a key part of the process for the Performance Analyst in the National Performance team, and from previous Reporter reviews was observed to be carried out faithfully. With the responsibility moving to the Routes, the National Performance team have less influence on ensuring that this is carried out with the same level of vigilance.

That said, it is noted from reviewing data from 2012/13 Period 11 that for 70% of TOCs (17 out of 24), the difference between TOC supplied PPM and PSS calculated PPM was less than 0.1%. Of the remaining seven TOCs, all but two were only 0.1% different. Therefore, any concerns about this process are alleviated by the fact that TOC and PSS data now appear to be fairly closely aligned pre-reconciliation. The average difference between TOC and PSS PPM figures for the 13 periods to 2011/12 Period 11 was 0.03%.

Automatic Input of Island Line Performance Data

In the previous report, it was noted that data from Island Line was provided manually since this network is not connected to the TRUST system. This data was provided separately by Island Line. The Performance Analyst was responsible for manually typing this data directly into the calculation spreadsheets.

The Reporter was advised that the data for Island Line services now feeds directly into the PSS system, and so is extracted from PSS via the extraction query as per all other TOCs.

4.2.1.2 Review of Data Flow Issues from 2011 Q4 Report

This section covers latest progress on other issues noted in the 2011 Quarter 4 report.

Enable Automation of Data Entry for TOC Cancellations

In the 2011 Quarter 4 report, it was noted "*The only area of hard coding in the process remains the manual entry of TOC Cancellation data, as provided by the TOCs*". In the 2010 Quarter 1 report it was recommended that "*Some benefits could be realised from working with the TOCs to produce a more automated process for gathering this data, e.g. through use of standardised templates, although it is acknowledged that Network Rail is unable to impose such changes upon TOC processes*".

At the time of the Reporter interviews in April 2011, Network Rail had taken a proposal to the PPM Steering Group to issue a standard template form for TOCs to fill in their cancellation data each period. This would have enabled automated data input, and so remove one of the key data-flow risks to the process. However, the Reporter was advised that this proposal had "met resistance", and as a result no changes to the process have been implemented. Therefore, as reported previously, this information continues to be provided in a variety of formats (examples provided ranged from Microsoft Excel files to Microsoft Word files and PDF files), with no standard naming convention. Thus, this remains the area of most risk to data accuracy in the process.

The Performance Analyst does keep a clear record of all data supplied by the TOCs in the "PPM TOC Data – 1213 P11.xls" spreadsheet. The Performance

Analyst populates the following information on receipt of each file from each TOC:

- File name and person who provided;
- For files provided in Excel format, the cell reference for each relevant piece of information; and
- For files in non-Excel format, the relevant value is directly typed into this spreadsheet.

This spreadsheet then compares the PPM for each TOC based on TOC supplied data and PSS data for the reconciliation process.

It is noted that this data is then manually copied from this spreadsheet into the main calculation spreadsheet individually for each TOC.

Documentation of PPM/CaSL Calculation Process (Recommendation 2011.5.1)

During the previous Reporter visit, the PPM and CaSL process documentation was underway, but not complete. This documentation has now been completed, and a copy of this guide was provided to the Reporter team for review (PPM & CaSL Industry Data Process Note" review date 07/03/2013).

This document provides a very detailed step-by-step guide to the calculation process, including:

- Timeline of required outputs;
- Data sources;
- Description of all spreadsheets used in process, and flow chart of data linkages;
- Clear and detailed step-by-step guidance for:
  - Collecting initial data from PSS;
  - Collecting and checking TOC data (reconciliation process), including details of those who historically have not provided this information; it is noted that this section outlines what to check for in reconciliation, even though this responsibility now falls to the Routes;
  - Compiling and publishing PPM and CaSL figures.

The document also provides guidance on the expected time required each period to carry out this process (~20 hours per period).

Although the process has historically been run each period by the same Performance Analyst, it was confirmed that other members of the National Performance Team have now been trained to run this process. The Reporter team was advised that two other members of the Performance Team have run this process in recent periods.

4.2.1.3 PPM / CaSL Conclusions

The Reporter team have reviewed the end-to-end calculation process through the spreadsheets provided by Network Rail. All linkages remain formulae based, with the only area of hard coding remaining as the input of TOC supplied

cancellation data. A series of spot checks, based on a sample of TOCs, have been carried out on the data flow from extraction from PSS through to final figures, to ensure this is accurate. These spot checks are summarised in Appendix B and show no concerns. On the basis of these checks, the Reporter team are content that the data still appears to be flowing through correctly.

Note, as per previous audits, no verification of the process used by TOCs to create their cancellation data has been carried out as this is outside the remit of the mandate.

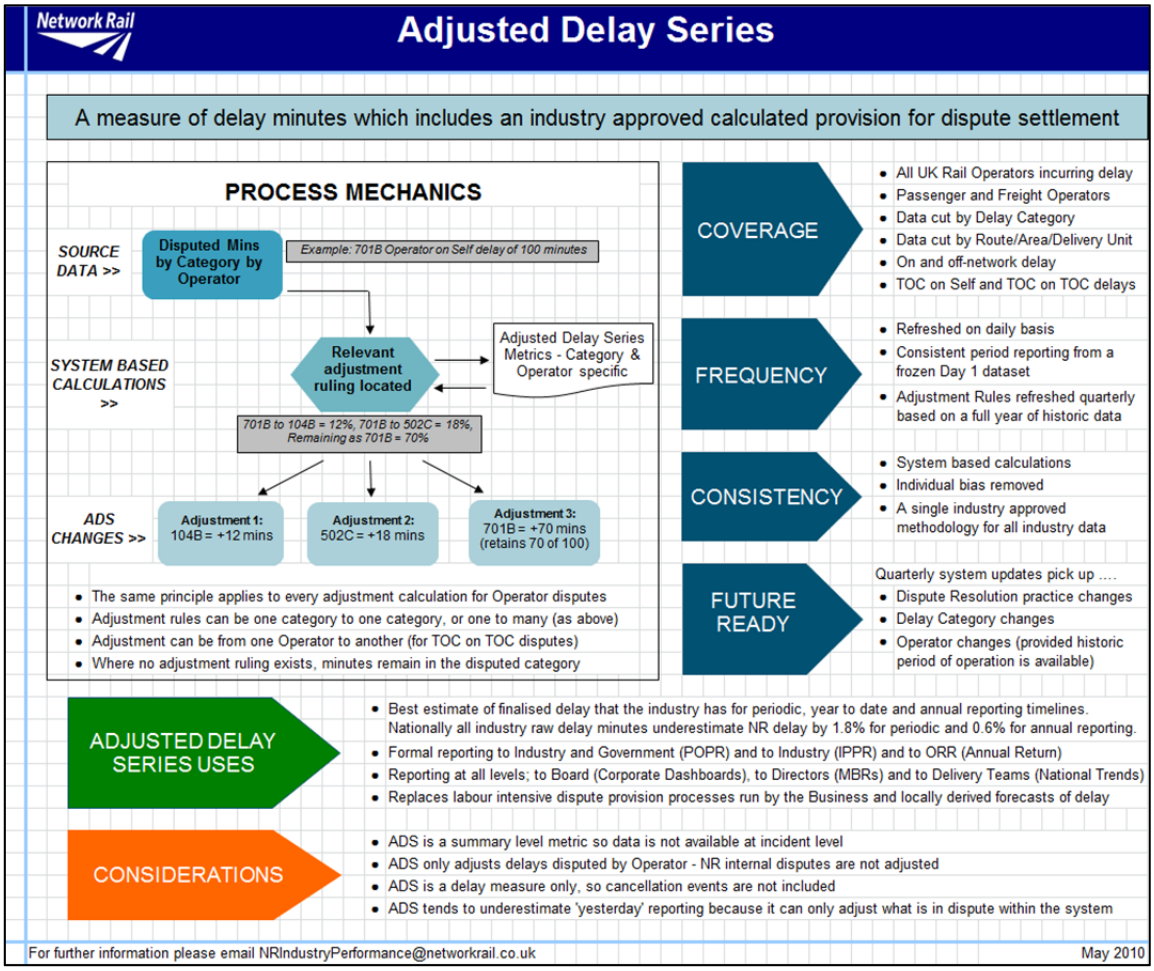
The Reporter team are satisfied that the PPM/CaSL data flow process is fully embedded and operating correctly, and that the documentation of this process is now complete and provides clear guidance on the approach, and is being used by National Performance Team members.

4.2.2 Delay Minutes

As outlined in the 2010 Quarter 1 report, delay minutes are now calculated and reported in the Industry Periodic Performance Report (IPPR) using the Adjusted Data Series (ADS) approach. This was introduced to improve the accuracy of delay minute reporting for incidents still under dispute at the time of publication; as outlined in the 2010 Quarter 1 report:

> "*The previous system only allowed the estimation of the likely outcome of disputes at whole TOC level. The revised process is designed to calculate the impact at delay category level to give a more accurate picture of the likely outcome …*
> *The ADS methodology uses historical data showing where disputed minutes, by delay category, are re-allocated to once a settlement is reached. The adjustment factors are updated every 3 months based on the most recent data.*"

This process is summarised in the flow chart below, as provided by Network Rail.



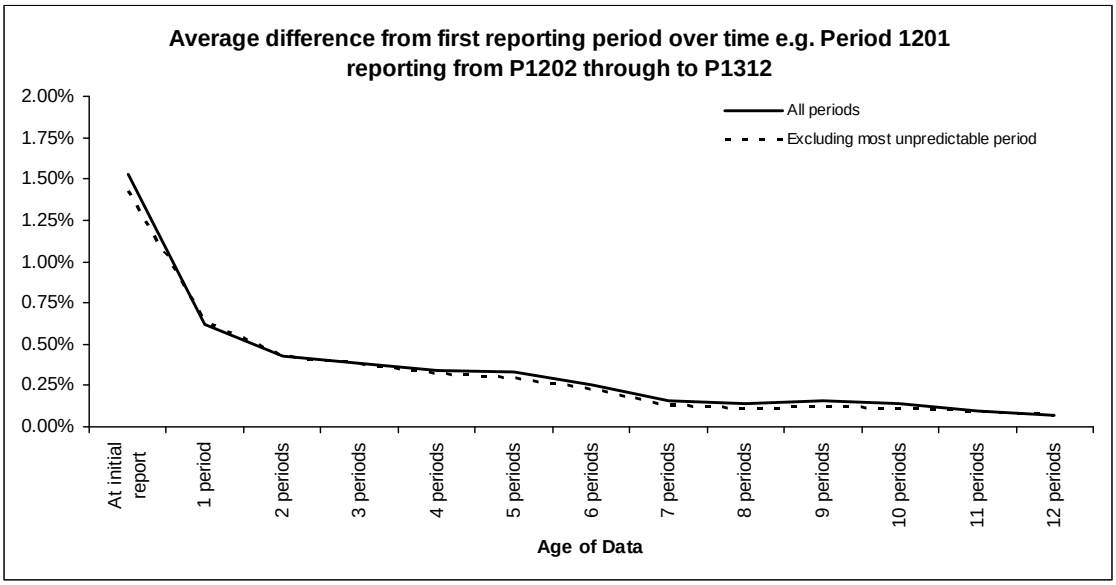
At the time of the 2011 Quarter 4 report, this process had only been in place for a couple of periods. For the purposes of this review, data has been supplied by Network Rail to demonstrate the impact of ADS on the accuracy of initial period delay minute reporting.

In the initial period of reporting, for incidents still in dispute, delay minutes are attributed based on forecasts by ADS as shown above. In subsequent periods, as disputes get resolved, delay data gets refreshed such that the 'ADS estimates' are replaced by actual final attributed delay minutes. Therefore the final attributed NR minutes for a period once all disputes are resolved are likely to differ from those reported in the initial period – with the difference representing where ADS did not accurately forecast how particular incidents would be resolved.

The chart below quantifies the size of the difference between final attributed NR minutes for a period, and the initial ADS-estimated attributed NR minutes. This is shown as the average difference for each period in 2011/12. So for example, on average the difference in the number of delay minutes finally attributed to NR in 2012/13 after all disputes were resolved compared with the figure estimated by ADS (and reported) in the initial period was 1.5%.

This indicates that the average difference in NR attributed delay minutes when reported in IPPR in the initial period compared to the final attributed figure was around 1.5%. By the subsequent period, this difference has reduced to ~0.5%. In

contrast, as estimated in the 2010 Quarter 1 report, the difference in Network Rail delay minutes was around 4% prior to ADS.



Source: Network Rail Spreadsheet ("WinderPhillips P1312 ADS Summary Analysis 20130402")

4.3 Data Assurance in the Context of the Right Time Performance Report – PPM & CaSL

Following the publication of Arup's Right Time review[3], we have assessed the implications of the identified timing errors on PPM and CaSL. This follow-up review uses the key findings from the Right Time review and performs a statistical analysis to measure the impact on PPM and CaSL.

4.3.1 Background – Summary of Initial Right Time Review Study

The Right Time review made a distinction between the two methods of capturing arrival data, namely automatic reporting and manual reporting. Under automatic reporting the time at which a train passes a known point is recorded by the signalling system and a pre-determined berthing offset is added to the overall time to allow for the time taken to stop at the platform. Under manual reporting the arrival time of the train is input directly into the train reporting system by signallers based on direct sight of the train coming to a stand, reports from the traincrew, or by adding an allowance to an observed timing.

The key findings from the review of *automatic* reporting included:

- This reporting method accounts for 77% of the final destination locations and handles 91% of train movements;
- There are well documented processes in place to support the systems used for automatic reporting and the vast majority of measurements are up to date and accurate within the limits of the system;

[3] 223767-07 Mandate AO/033 - Right Time Performance Measure Final Report (Complete)

- When infrastructure or operational aspects change on site or when a site is due a review, the berthing offset review process can be slow and difficult to follow;
- Analysis of berthing offset sample data revealed that the maximum confidence level for accuracy was on average +/- 13 seconds across all automatic locations (in this study we refer to this as the automatic reporting error band);
- Applying the error band to period 5 2012/13 arrivals indicated that 1.74% of trains were within the right time arrival threshold and could potentially be reported wrongly at the Great Britain national level; and
- When consideration was also given to lapsed berthing offsets the percentage of trains potentially impacted increased to 2.00%.

The key findings from the review of *manual* reporting included:

- There are no discernible processes in place to manage the reporting of arrivals;
- It accounts for less than 10% of the daily arrivals nationally but the percentage of manual reporting by individual routes and for certain TOCs varies significantly;
- Analysis of data from Sheffield University provided a manual reporting error band of +/- 35 seconds; and
- Again, applying this error band to period 5 2012/13 indicated that 4.93% of trains were within the right time arrival threshold and could potentially be reported wrongly at the Great Britain national level.

By taking account of the number of trains under automatic reporting and manual reporting, the combined results indicated that 2.23% of trains could potentially be reported wrongly against right time at the national level.

4.3.2 Elimination of Lapsed Berthing Offsets

At the time of the right time study, there were 70 terminating stations with automatic reporting that had out-of-date or invalid offsets. As at March 25th this number had reduced to 25. As Network Rail has made steady progress in rectifying the out-of-date offsets, this study treats all lapsed berthing offsets as having been eliminated. Although not explicitly reported previously, the combined potential right time error percentage would have been 1.99% if there were no lapsed berthing offsets.

4.3.3 Description of Datasets Used

For the data assurance review of PPM and CaSL in the context of the Right Time Performance Measure, two datasets were obtained from Network Rail following a meeting on 27th March 2013:

1. “ORR dataset AO39 P1213 to P1312.xls” provides 13 periods of arrival frequency counts aggregated by period, destination, reporting method, operator and sector for each minute ranging from 3 minutes early up to 19 minutes late, with bandings for subsequent lateness; and

2. "ORR dataset AO39 P1310 to P1312 extended.xls" provides 3 periods of similarly formatted frequency counts which also includes additional columns for 29, 30 and 31 minutes late.

The first dataset has contributed to the PPM study with the lateness frequencies around 5 minutes and 10 minutes being of interest, depending on the TOC sector. Some 7 million trains reached their destination across the 13 periods. The second dataset has contributed to the CaSL study with the additional breakdown around 30 minutes late being of interest for all TOCs. Some 1.6 million trains reached their destination across the 3 periods.

The datasets available to this study provide a better quality sample compared to those that were available during the Right Time study. In this case the time horizons are considerably longer and the arrival frequencies are disaggregated to each minute rather than comprising bandings for "Early", "On time", "Within 3 minutes", "Within 5 minutes", etc. The additional granularity in the data has permitted an improved methodology for measuring the impact of timing errors.

4.3.4 Description of Methodology

The methodology is summarised below with full details provided in Appendix C. In essence, three steps have been followed:

1. Estimate net potential errors;
2. Calculate expected net errors; and
3. Calculate the percentage impact on the PPM and CaSL metrics.

Given the error bands of +/-13 seconds for automatic reporting and +/- 35 seconds for manual reporting, there is a chance that a train recorded as arriving early may in fact be late (referred to as a false positive), or a chance that a train recorded as arriving late may in fact be early (referred to as a false negative). In Step 1, the net potential errors are calculated as the number of false positives minus the number of false negatives as the errors will tend to cancel each other out. This gives the maximum potential impact of the error bands and is the basis for the figures in the Right Time report. These represent overly pessimistic figures and err on the side of caution.

Thus, a further calculation is required to weight the potential errors according to the probability that the error is large enough to require reclassification, which has been undertaken in Step 2. In other words, Step 1 provided the net number of trains that could be recorded incorrectly, while Step 2 extends this to consider their likelihood of actually being recorded incorrectly. Probability theory is applied as trains recorded as arriving close to the threshold are more likely to be recorded incorrectly than those further from the threshold.

4.3.5 Detailed Results

Net potential errors and expected net errors have been calculated for Right Time arrivals across all locations at the national level for comparison with the previously reported error figure. In **Table 4.1**, the previously reported error figure is 1.99% (assuming no lapsed berthing offsets). Utilising the superior datasets in this study provides a comparative net potential error of 1.24%. When the

methodology is extended to include probability theory, the expected error is reduced to 0.21%.

Table 4.1: Right time impact

	Combined Locations		
	% Error (Previous)	**% Error (Potential)**	**% Error (Expected)**
NATIONAL			
Great Britain	1.99%	1.24%	0.21%

The PPM impact would be expected to be substantially reduced from the Right Time impact as fewer trains are recorded at 5 minutes late (London & South East, Regional and Scotland services) or 10 minutes late (Long Distance services) than are recorded Right Time. **Table 4.2** shows a breakdown of expected net errors in PPM for automatic and manual locations separately and for all locations combined. With the overall national error showing as 0.008% this indicates that any nationally reported PPM is marginally over stated by 0.008%. The maximum error across the combined locations is 0.013% under First ScotRail.

Table 4.2: PPM impact - expected net errors

	Automatic Locations	**Manual Locations**	**Combined Locations**
NATIONAL			
Great Britain	0.006%	0.021%	0.008%
England and Wales	0.006%	0.018%	0.007%
Scotland	0.009%	0.040%	0.013%
SECTOR			
Long Distance	0.002%	0.014%	0.003%
London & South East	0.006%	0.010%	0.006%
Regional	0.005%	0.021%	0.009%
Scotland	0.009%	0.040%	0.013%
FRANCHISED TOC			
First TransPennine Express	0.001%	0.009%	0.004%
Greater Anglia	0.006%	0.011%	0.006%
Northern Rail	0.005%	0.039%	0.011%
Heathrow Connect	-	-	-
First Great Western	0.005%	0.015%	0.008%
First Capital Connect	0.006%	0.000%	0.006%
CrossCountry	0.002%	0.011%	0.002%
London Midland	0.011%	0.010%	0.011%
London Overground	0.002%	0.000%	0.002%
East Midland Trains	0.003%	0.025%	0.005%
First ScotRail	0.009%	0.040%	0.013%

	Automatic Locations	Manual Locations	Combined Locations
East Coast	0.002%	0.005%	0.002%
Merseyrail	0.005%	0.020%	0.006%
Virgin Trains	0.002%	0.013%	0.002%
Arriva Trains Wales	0.002%	0.013%	0.006%
Chiltern	0.002%	0.013%	0.003%
c2c	0.002%	0.000%	0.002%
Southeastern	0.007%	0.000%	0.007%
Southern	0.009%	0.043%	0.009%
South Western Trains	0.007%	0.017%	0.007%
Franchised Total	0.006%	0.022%	0.008%
NON-FRANCHISED TOC			
Grand Central[4]	0.002%	-0.011%	-0.001%
Heathrow Express	0.006%	0.000%	0.006%
Hull Trains	0.000%	0.021%	0.010%
Non-Franchised Total	0.006%	0.011%	0.006%

The impact on CaSL is reduced further due to the very small number of trains recorded as 30 minutes late (all services). In fact, only 8,000 of the 1.6 million planned trains were significantly late. Furthermore, the CaSL measure is dominated by cancellations so any impact due to errors in recording significant lateness will be minimal in comparison.

There are insufficient recordings at many locations to provide reliable expected net error estimates at the TOC level, particularly for manual locations. As the sample sizes are generally too small for statistical treatment, it is more appropriate to consider the worst case errors by not taking into account the possibility of errors in the recordings cancelling each other out, as presented in **Table 4.3**. Positive and negative errors arise as so few trains are involved and in some cases there may be more trains recorded as 30 to 31 minutes late than were recorded as 29 to 30 minutes late. The maximum absolute error across the combined locations is 0.13589% under Hull Trains. There is insufficient data for London Underground and Heathrow Express to provide any estimates for manual locations.

[4] A negative error occurs on the Grand Central manual locations because the small sample size for this TOC shows more trains as 10 to 11 minutes late than were 9 to 10 minutes late.

Table 4.3: CaSL impact - worst case errors

	Automatic Locations	Manual Locations	Combined Locations
NATIONAL			
Great Britain	+/- 0.00062%	+/- 0.00265%	+/- 0.00031%
England and Wales	+/- 0.00072%	+/- 0.00129%	+/- 0.00053%
Scotland	+/- 0.00030%	+/- 0.01226%	+/- 0.00167%
SECTOR			
Long Distance	+/- 0.00120%	+/- 0.00413%	+/- 0.00059%
London & the South East	+/- 0.00030%	+/- 0.00000%	+/- 0.00029%
Regional	+/- 0.00169%	+/- 0.00116%	+/- 0.00103%
Scotland	+/- 0.00030%	+/- 0.01226%	+/- 0.00167%
FRANCHISED TOC			
First TransPennine Express	+/- 0.00614%	+/- 0.00628%	+/- 0.00619%
Greater Anglia	0.00000%	0.00000%	0.00000%
Northern Rail	+/- 0.00250%	+/- 0.01425%	+/- 0.00010%
Heathrow Connect			
First Great Western	+/- 0.00171%	+/- 0.00910%	+/- 0.00071%
First Capital Connect	+/- 0.00512%	0.00000%	+/- 0.00512%
CrossCountry	+/- 0.00581%	0.00000%	+/- 0.00563%
London Midland	+/- 0.00156%	+/- 0.00920%	+/- 0.00360%
London Overground	+/- 0.00025%		+/- 0.00025%
East Midland Trains	+/- 0.00275%	+/- 0.04641%	+/- 0.00250%
First ScotRail	+/- 0.00030%	+/- 0.01226%	+/- 0.00167%
East Coast	+/- 0.02636%	+/- 0.11438%	+/- 0.01995%
Merseyrail	+/- 0.00050%	0.00000%	+/- 0.00047%
Virgin Trains	+/- 0.00815%	+/- 0.11929%	+/- 0.01038%
Arriva Trains Wales	+/- 0.00044%	+/- 0.01269%	+/- 0.00375%
Chiltern	0.00000%	0.00000%	0.00000%
c2c	+/- 0.00083%	0.00000%	+/- 0.00083%
Southeastern	+/- 0.00058%	0.00000%	+/- 0.00058%
Southern	0.00000%	0.00000%	0.00000%
South Western Trains	+/- 0.00087%	0.00000%	+/- 0.00087%
Franchised Total	+/- 0.00064%	+/- 0.00157%	+/- 0.00044%
NON-FRANCHISED TOC			
Grand Central	+/- 0.02130%	+/- 0.24205%	+/- 0.02915%
Heathrow Express	+/- 0.00183%		+/- 0.00183%
Hull Trains	+/- 0.04224%	+/- 0.23102%	+/- 0.13589%
Non-Franchised Total	+/- 0.00162%	+/- 0.23458%	+/- 0.01393%

4.3.6 Conclusions

In respect of PPM impact, with the overall national error showing as 0.008% this indicates that the nationally reported PPM is marginally overstated by 0.008%. The largest error across the combined locations is 0.013% for First ScotRail.

In respect of CaSL impact, the breakdown of errors presented in **Table 4.3** above shows that the overall national error is essentially zero percent. The maximum absolute worst case error across the combined locations is 0.13589% at Hull Trains.

In the findings of the Right Time Performance Measure report it was stated that there were some issues with the delivery of the automatic reporting process and manual reporting was completely lacking a recorded process. Despite the misgivings of the overall process, the impact on national, sector and TOC PPM and CaSL is negligible.

4.4 Confidence Ratings

The ratings for the performance KPIs are outlined in **Table 4.4** below, demonstrating any changes since the 2011 Quarter 4 report.

Table 4.4: Ratings of the Performance KPIs

KPI	Original Score	New Score	Comments
5a: PPM	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. As noted previously, with the exception of the use of TOC data, the process is fully automated. The findings of the Right Time study have negligible impact on PPM, and so provide no reason to downgrade this rating.
5b: CaSL	Reliability A Accuracy 2	Reliability A Accuracy 2	No Change. Network Rail are still reliant on TOC cancellation data for reporting purposes. The provision of TOC data provides confidence in terms of enabling a reconciliation exercise between data held within PSS and by TOCs, but there do remain differences in the datasets. NR only investigate and reconcile differences of over 0.1%. The findings of the Right Time study have negligible impact on CaSL.
5c: Network Rail Delay Minutes to TOCs	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. The impact of the ADS calculation has added greater confidence in the accuracy of these figures.
5d: Network Rail Delay Minutes to FOCs per 100 train kms	Reliability A Accuracy 3	Reliability A Accuracy 3	No Change. The accuracy rating reflects the on-going issues with the collation of the freight mileage data used as the normaliser. NR have now devised a more accurate methodology for calculating freight mileages but have taken the decision to introduce it in

KPI	Original Score	New Score	Comments
			April 2014 (see update on recommendation 2010.5.4a in Section 3.1). Once this is fully implemented the measure should be able to rise to A1.
6a/6b: Asset Management (Track / Non Track Delay Minutes)	Reliability A Accuracy 1	Reliability A Accuracy 1	No Change. This dataset is a direct derivative of Network Rail delay minutes, and so is reflective of the KPI score for 5c.

5 Delay Attribution Review

5.1 Introduction

ORR requested that the Reporter team carry out an investigation into whether the quality of delay attribution has deteriorated during CP4. The request arose from concerns that reductions in staffing levels or other causes, such as the impact of Alliances between NR and TOCs, may have led to deterioration in the quality of attribution.

At the initial project tripartite meeting held on the 5th March it was agreed that this review should be addressed in two phases. The first phase would carry out empirical comparisons between selected delay data from early in CP4 and more recent periods, to identify if there are any trends either at national or Route level that indicated a fall in the quality of attribution. On the basis of this comparison, the Reporter team would then propose potential hypotheses to explain any deterioration of delay attribution for more detailed investigation in a second phase. These hypotheses could include reduction is staffing numbers but should cover all possible causes.

This report presents the findings of the first phase of work and sets out our recommendations for the second phase.

5.2 Methodology

For this first phase of the review, it was agreed at a meeting with NR on the 23rd April that they would provide data for comparison purposes covering the following periods:

- Periods 5,6,7,11,12 and 13 for the financial years 2009/10 and 2012/13.

This gives six periods' worth of data within both years and covers different times of the year. It also avoided the main autumn periods where the attribution arrangements with some TOCs, and the more inconsistent nature of such periods in terms of level of disruption, may have made drawing any firm conclusions on generic changes in quality of attribution more difficult.

Comparing the six periods in each year sampled, the number of all incidents of delay increased by **1.5%** to an average of 55k per period, and the number of delay minutes increased by **16%**. This is important to bear in mind when we make our comparisons.

During our meeting, the categories of delay to be checked were discussed in detail, and a specification for data extraction was agreed. The chosen categories were selected as being reasonable indicators that the TRUST delay attribution processes are functioning well and that any changes in these measures could indicate changes in the underlying arrangements. The measures reviewed were:

- The use of FO/TO codes;
- The use of Z* codes;
- Levels of commercial takeback;
- The levels of recoding;

- Incidents in dispute on Day 2;
- Roll up incidents; and
- Creation of management incidents.

These categories are explained in more detail in the sub-sections that follow. For some of them, the collation of the data by NR was not straightforward due to changes in attribution practices since 2009/10. However, solutions were identified in each case, and assumptions regarding data extraction are detailed in this report for each category of delay.

The data was provided by Route and, where appropriate, by TOC and FOC for each measure. This allowed the Reporter team to look at both a national picture and a more focused look at any changes by Route or operator.

NR also provided their own data check reports that they submit to the Delay Attribution Board (DAB) on a periodic basis as well as the latest version of the Performance Measurement Process Manual. This has been recently reissued and sets out in detail the TRUST attribution arrangements on a national basis. This includes the data check processes required to ensure data quality is kept to acceptable standards. The NR central team routinely monitor the use of most of the measures covered in this report particularly the use of Z*, QT, roll ups and the use of management incidents.

It is important to note that many of the measures in this report are at the softer edge of the rules and processes set out in the Delay Attribution Guide (DAG), and practices will vary according to agreements between NR and the relevant TOCs around areas such as the handling of small delays. This does mean that the numbers of incidents recoded across each Route does vary and the Reporter team have tried to focus on proportional changes rather than applying a direct comparison between Routes.

5.3 The use of FO/TO Delay Cause Codes

5.3.1 Findings

FO and TO are “other” codes which indicate that an incident has been allocated to a FOC or a TOC respectively, without a specific known cause. FO and TO would indicate that Level 1 staff have failed to find a NR reason for the delay and have correctly allocated to the TOC in line with the DAG. The TOC or FOC should then investigate the incident to ascertain a specific cause. Typically this will involve requesting additional information from key personnel such as drivers. Significant recoding of FO/TO back to NR could indicate a failure by the Level 1 staff to investigate NR reasons in the initial review but is more likely to reflect further information becoming available after the initial attribution had occurred. Once the TOC/FOC have accepted an incident, any subsequent recoding is undertaken by the TOC not NR.

The data provided allowed the Reporter team to compare the levels of initial FO/TO attribution in both years. **Table 5.1** shows the levels of FO/TO over the study periods with the observed percentage change. FO and TO have been shown separately.

Table 5.1: FO and TO Incidents by Route

Route	Total Day 1 TO Incidents by Route			Total Day 1 FO Incidents by Route		
	2009/10	2012/13	% Change	2009/10	2012/13	% Change
Anglia	1,414	941	-33%	1,886	1,720	-9%
East Midlands	704	446	-37%	1,504	1,790	19%
Kent	4,031	4,486	11%	1,103	1,326	20%
London North Eastern	6,360	5,124	-19%	6,829	9,307	36%
London North Western	10,558	11,773	12%	3,703	5,183	40%
Scotland	9,249	10,413	13%	2,405	2,221	-8%
Sussex	6,898	4,269	-38%	276	367	33%
Wales	450	472	5%	876	1,238	41%
Wessex	573	835	46%	1,031	1,023	-1%
Western	2,326	3,540	52%	793	1,694	114%
Total	**42,757**	**42,569**	**0%**	**20,406**	**25,872**	**27%**

Overall there has been an 8% increase in the number of incidents coded to FO/TO on Day 1 between the two years. However, the breakdown does show the picture is not a uniform one across the country.

In terms of TO incidents, overall there is no increase between the two years. This masks a high increase on Western Route where incidents initially attributed to TO, increased by just over 50%. Meanwhile, significant improvements were observed on Sussex Route.

The absolute levels of FO incidents for freight companies are high and rising and overall have increased by 29%. FO delay minutes have increased by 203% over the same period.

FO/TO delay for freight operators has increased across the country with the largest increase on Western. An analysis by freight operator shows there has been a substantial increase across all of the main freight operators as shown in **Table 5.2** below. Note, this table excludes those incidents which were initially coded to FO/TO for freight operators and were subsequently recoded, since this level of data aggregation for recoded incidents was not available. However such recoded incidents account for less than 10% of all incidents which were initially coded to FO/TO for freight operators.

Table 5.2: FO/TO Incidents by Freight Operator

TOC	Total Day 1 FO/TO Incidents by FOC			Total Day 1 FO/TO Minutes by FOC		
	2009/10	2012/13	% Change	2009/10	2012/13	% Change
Advenza Freight	6	0	-100%	37	0	-100%
Colas Freight	103	213	107%	692	1,273	84%
DB Schenker	11,138	12,877	16%	86,935	100,365	15%
DRS	533	863	62%	3,725	6,241	68%
Fastline Freight	128	0	-100%	815	0	-100%
Freightliner Int	4,477	6,901	54%	25,596	40,772	59%
GB Railfreight	1,891	2,998	59%	17,624	20,501	16%
Harsco	31	76	145%	201	640	218%
NYMR	31	28	-10%	153	292	91%
Serco Rail Ops	248	0	-100%	2,388	0	-100%
Total	**18,586**	**23,956**	**29%**	**138,165**	**170,083**	**23%**

(Note that a small number of TO incidents have been attributed to freight operators, which explains why the number of incidents is slightly different to the number of FO incidents in Table 5.1.)

5.3.2 Factors for Consideration

The levels of TO across the industry have remained broadly static although there have been substantial variations across the Routes. The numbers do not indicate any significant issues of concern.

The increase in FO use is a universal issue both in terms of freight operators and NR routes. This suggests that the investigation of freight delays to identify more specific causes has reduced over the period and therefore there has been a reduction in the overall levels of data quality. This could be caused by an increase in freight volumes over the period and some normalisation may account for some of the increase. However, there is still a significant increase.

The FOC TRUST processes beyond Level 1 are managed by eight Dispute Resolution Coordinators (DRCs) based in the Route offices but there is no overall coordination through a lead route or the centre of the level 2 or 3 processes. Level 4 is however managed by the Customer Relationship Executive (CRE) for each operator. It is also the case that freight operators have reduced their resources involved in delay attribution, and this might be a contributory factor to the increases in the levels of FO particularly as small delays have no commercial impact. The detailed reasons behind the increase should be investigated with NR and the FOCs.

5.4 The use of Z* Delay Codes

5.4.1 Findings

Z* coded incidents are unexplained delay or cancellation events where the TRUST DA has been unable for whatever reason to investigate and provide an appropriate cause. Any changes in levels of Z* codes could indicate a change in the arrangements for delay investigation.

Table 5.3 below is broken down as follows:

1. Column 1 shows the number of incidents that were initially attributed to a Z* incident and remained attributed to this code, i.e. a delay cause was never found.
2. Column 2 shows the number of incidents that were initially attributed to another code and were re-attributed to Z* following investigation at L2.
3. Colum 3 shows the number of incidents that were initially attributed to Z* and after further investigation a delay cause was found.

Table 5.3: Z coding by Route

	2009/10			2012/13		
Initial Attribution ==>	**Z* Code**	**Non Z* Code**	**Z* Code**	**Z* Code**	**Non Z* Code**	**Z* Code**
Final Attribution ==>	**Z* Code**	**Z* Code**	**Non Z* Code**	**Z* Code**	**Z* Code**	**Non Z* Code**
Anglia	318	499	8	134	495	20
East Midlands	1,024	95	19	1,305	101	6
Kent	21	118	34	138	849	6
London North Eastern	5,111	172	23	7,293	244	37
London North Western	11,089	449	27	9,540	724	91
Scotland	361	6,518	13	31	2,042	149
Sussex	111	1,680	1	3,797	1,299	4
Wales	1,382	273	5	691	277	24
Wessex	29	172	6	77	374	1
Western	2,604	862	9	1,507	1,617	12
Total	**22,050**	**10,838**	**145**	**24,513**	**8,022**	**350**

Overall the levels of Z* initial attribution increased by 12% over the study period. However, this masks a considerable variation by Routes. Sussex over this period had a 3294% increase in the initial attribution of incidents to Z* codes and a 185% increase in the levels of incidents remaining as Z* following investigation –

to the extent that it now has the third highest number across the Routes. Kent and Wessex have also both seen percentage increases but the absolute levels are much lower than Sussex and of much less concern.

The overall level of incidents subsequently recoded to Z* has fallen by 25%. This is largely driven by the change in practice introduced in Scotland where the use of TO has increased in proportion to the reduction in Z*.

The other Route to see a significant increase is LNE. Initial use of Z* has risen by 43% and the final attribution levels have increased at the same rate. Only 1% of the Z* delays on LNE are reattributed to a non Z* code. East Midlands has followed a similar pattern, but as the DA process is managed by the LNE team in York this is not unexpected.

5.4.2 Factors for Consideration

There is no discernible national trend. There are however specific issues on Sussex and LNE (and its association with East Midlands) about the increase in the levels of initial attribution.

The increases in the initial use of Z* on Sussex is associated with a trial of using ZZ when sub threshold delays created an above threshold delay. This was put to ZZ and shared between NR and Southern rather than spend a lot of time chasing delays which may ultimately have no clear cause.

5.5 Rates of Commercial Takeback

5.5.1 Findings

Commercial takeback occurs when NR and the TOC agree to split delays following a failure to attribute the delay to a specific operator or NR code. These will be attributed to the TOC initially and will normally be dealt with after Day 1. An increase in commercial takeback could point to a reduction in focus on delay quality and using bilateral discussions to resolve issues rather than focusing on root cause.

Commercial takeback can be handled in two ways:

- either the delay will be coded QT and the minutes and cancellations accepted by NR or the incident will remain coded to the original manager and be edited in PEARS, or
- NR may use a D code responsible manager to split the incident between the TOC/FOC and NR.

The Delay Attribution Guide sets out certain circumstances where incidents should be split, e.g. fire alarms or security alerts at stations. Since these are mandated in the DAG, these have been excluded from the figures shown below.

Table 5.4 sets out the levels of QT delays over the study periods.

Table 5.4: QT Incidents and Delay Minutes by Route

Route	Number of QT Incidents			Number of QT Minutes		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	339	14	-96%	15,195	379	-98%
East Midlands	91	37	-59%	2,181	645	-70%
Kent	397	115	-71%	4,044	1,289	-68%
London North Eastern	199	159	-20%	4,261	3,174	-26%
London North Western	452	335	-26%	9,223	5,355	-42%
Scotland	102	135	32%	2,687	1,238	-54%
Sussex	724	81	-89%	5,927	2,256	-62%
Wales	174	152	-13%	1,758	2,458	40%
Wessex	127	64	-50%	1,909	454	-76%
Western	377	345	-8%	5,161	3,852	-25%
TOTAL	**2,982**	**1,437**	**-52%**	**52,346**	**21,100**	**-60%**

The number of QT incidents has fallen on every route except Scotland whilst the number of delay minutes has reduced on every route except Wales which had a very small increase. Overall the number of incidents subject to QT has halved.

Table 5.5 shows the use of D coded responsible manager to split delays.

Table 5.5: D Coded Incidents and Delay Minutes by Route

Route	Number of D Coded Incidents			Number of D Coded Minutes		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	701	346	-51%	9,438	4,549	-52%
East Midlands	892	405	-55%	3,965	2,018	-49%
Kent	181	49	-73%	1,947	495	-75%
London North Eastern	4,812	3,215	-33%	16,666	12,431	-25%
London North Western	6,391	6,279	-2%	35,611	31,647	-11%
Scotland	1,787	1,050	-41%	8,088	5,807	-28%
Sussex	1,313	122	-91%	7,865	1,480	-81%
Wales	283	244	-14%	3,469	3,045	-12%
Wessex	234	508	117%	6,931	4,954	-29%
Western	1,102	2,049	86%	8,356	14,406	72%
TOTAL	**17,696**	**14,267**	**-19%**	**102,334**	**80,829**	**-21%**

The national picture shows a significant reduction in overall levels of both incidents and delay minutes using D coding. There has, though, been an increase in the use of D coding on two routes: the number of D Coded incidents on Wessex has increased by 117% but in absolute terms the numbers are still relatively low and delay minutes have actually decreased; whereas Western has seen an increase of 86% in D Coded incidents and a 72% increase in delay minutes.

5.5.2 Factors for Consideration

The use of both QT and D coding has fallen substantially over the three year period suggesting that more incidents are being properly attributed to specific causes rather than being subject to a commercial arrangement. As an example LNE and LNW have agreed with Northern to no longer D code TO delays since April 1st this year, so this will impact future numbers.

The increase in QT delays in Scotland Route was due to the introduction of a commercial agreement to deal with autumn delays which occur before the start of the autumn season national process (1st October). This is to deal with the earlier onset of autumn in some years in Scotland. If this factor were removed the overall number of QT incidents would have fallen.

The changes on Wessex had already been identified by the Route Performance Manager and targets for a reduction in D coding have been put in place and are regularly monitored.

The changes on Western are related to a number of factors. A commercial agreement was put in place to deal with the following:

- A problem with the Sectional Running Time (SRT) at Kemble which is two minutes out and means that any station overtime will trigger a threshold delay. The delay is split between the genuine overtime and the SRT deficiency. This will cease in December 2013 when the SRT is corrected.
- There has been an increase in the number of incidents that are being split where there is a conflict of information, i.e. where the standard signaller response is "no signal delay" or "heavy passenger loadings" and the standard driver response is "signals". These have risen since 2009/10.
- Dual coding incidents where there are multiple causes of delay has arisen within the last 2 years – primarily driven by workload as performance has deteriorated and the number of incidents has increased. It is also due to the fact that FGW view the incidents split as sub threshold. The Route is currently reviewing the practice with FGW and has given notice that it will cease at the end of the current franchise.

5.6 Recoded Delays

5.6.1 Findings

This is a measure of how many incident cause codes are changed from the original attribution made by the Level 1 DA staff. This gives an indication of the accuracy levels of the original Level 1 attribution process. High levels of recoding would indicate that the initial attribution processes were creating significant errors in the attribution which had to be subsequently corrected. This could be because of staffing shortages or inadequate training.

Table 5.6 below shows the levels of recoding by route over the study periods, in terms of proportion of overall incidents which were recoded after Day 1. The overall levels of recoding are shown in the first set of columns (including recoding internally, i.e. from one NR code to another NR code or one TOC code to another TOC code). The overall levels of recoding from TOC codes to NR codes is shown in the second set of columns.

Table 5.6: Recoded Incidents by Route

Route	% Overall Incidents Recoded			% Overall Incidents Recoded from TOC to NR		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	11%	15%	3%	6%	6%	-1%
East Midlands	14%	11%	-4%	6%	4%	-2%
Kent	26%	24%	-2%	6%	8%	2%
London North Eastern	15%	12%	-3%	7%	6%	-2%
London North Western	17%	18%	1%	9%	8%	-1%
Scotland	36%	23%	-13%	24%	10%	-14%
Sussex	38%	22%	-15%	21%	9%	-12%
Wales	10%	10%	0%	5%	4%	-1%
Wessex	17%	17%	-1%	5%	7%	3%
Western	18%	23%	5%	9%	13%	3%
TOTAL	**19%**	**17%**	**-2%**	**10%**	**8%**	**-2%**

This table shows that the overall level of recoding of incidents has fallen very slightly from 19% to 17%, with the proportion of incidents being recoded from TOC codes to NR codes also dropping by 2 percentage points. The proportion of incidents being recoded has dropped most significantly on Scotland and Sussex.

However, as shown in **Table 5.7** below, the overall proportion of delay minutes which were recoded after Day 1 has remained relatively static at around 14% (with around 3% being recoded from TOC to NR).

Table 5.7: Recoded Incidents (Delay Minutes) by Route

Route	% Overall Minutes Recoded			% Overall Minutes Recoded from TOC to NR		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	10%	14%	3%	3%	3%	-1%
East Midlands	12%	12%	0%	2%	3%	1%
Kent	19%	13%	-6%	4%	3%	-1%
London North Eastern	12%	11%	-1%	2%	2%	1%
London North Western	11%	15%	3%	3%	3%	1%
Scotland	22%	18%	-4%	6%	3%	-3%
Sussex	26%	24%	-2%	6%	4%	-3%
Wales	8%	11%	3%	2%	2%	0%
Wessex	16%	15%	0%	1%	2%	0%
Western	9%	12%	3%	3%	3%	1%
TOTAL	**13%**	**14%**	**1%**	**3%**	**3%**	**0%**

Note these figures are against a backdrop of a 48% increase in level of overall incidents, and 16% increase in level of overall minutes.

5.6.2 Factors for Consideration

Overall recoding of incidents has reduced slightly suggesting that initial attribution quality is better in 2012/13 than it was in 2009/10. Although the proportion of incidents being recoded has significantly dropped on Sussex, it is notable that nearly a quarter of all minutes on the route are still being recoded after Day 1.

The fact that the number of recoded incidents has fallen but the delays remain at the same level indicates that their delay per incident (DPI) reason has increased. The reason for the increase is unclear, although we have not carried out a detailed investigation at this stage. We note that nearly 50% of all recoded minutes occurs on LNE and LNW, where DPI for recoded incidents has increased by 50% and 36% respectively. One potential factor is that post DAB decisions on any changes to attribution definition and process, NR will retrospectively recode incidents, and any that happened in the review periods will be included in these figures. This can on occasions include thousands of incidents.

5.7 Delays in Dispute on Day 2

5.7.1 Findings

This is a measure of the number of incidents and minutes in dispute between NR and the TOCs at Day 2 after the initial attribution process is complete. These levels are affected by many factors and in particular by the policies and behaviours of individual TOCs. This can reflect the maturity levels of the specific relationships and can give an indication of improvements or deterioration in the bi-partite relationship. It is important, like all the measures in this report, not to

take this measure in isolation to judge the effectiveness of the process. The DA processes do drive the dispute process and some TOCs will routinely dispute incidents whilst they await staff reports since technically they cannot dispute after Day 2.

Table 5.8 shows the percentage of incidents that were in dispute on Day 2 across the study periods, and the percentage of TOC delay minutes that they represent.

Table 5.8: Incidents in Dispute at Day 2 by Route

Route	% TOC Incidents in Dispute (Day 2)			% Mins (TOC Incidents) in Dispute (Day 2)		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	24%	33%	10%	38%	43%	6%
East Midlands	24%	20%	-5%	23%	18%	-5%
Kent	19%	15%	-3%	27%	17%	-9%
London North Eastern	17%	16%	-1%	15%	16%	1%
London North Western	25%	30%	4%	27%	31%	4%
Scotland	24%	17%	-7%	29%	26%	-3%
Sussex	51%	26%	-25%	48%	32%	-16%
Wales	-	-	-	-	-	-
Wessex	13%	14%	1%	24%	19%	-5%
Western*	18%	26%	8%	22%	32%	9%
TOTAL	**23%**	**23%**	**0%**	**25%**	**26%**	**1%**

* Wales and Western have been merged on this measure

Overall the percentages of incidents in dispute remain virtually unchanged as does the percentage of delays minutes. At route level, though, there are some substantial variations with Sussex showing a substantial reduction in incidents of 25% and Anglia increasing by 10%. The Sussex reduction is primarily driven by a 29% reduction in disputes by Southern which is because of the trial referred to earlier of using ZZ. Greater Anglia is now disputing more than in 2009/10.

5.7.2 Factors for Consideration

On a national level there are no significant trends that highlight any changes in practice or process leading to an increase or decrease in disputes. Changes in TOC approach are likely to have the biggest impact in this area. The increase in disputes by Greater Anglia may be because of a different policy applied by Abellio as the new operator as opposed to the previous franchise operator. Similarly the very recent change in Scotland to not dispute incidents routinely at level 1 will have an impact on any future assessment of similar numbers.

5.8 The level of Roll Up Delays

5.8.1 Findings

At 06:00, 14:00 and 22:00 the TRUST DA goes through a roll-up process which allocates unexplained and un-attributed delays to a general unexplained status. This means that any incidents not being dealt with will go into a roll-up incident and may create duplicate delays for subsequent attribution if not managed correctly.

NR has set a target of zero over threshold delays being attributed into roll ups and the Performance Measurement Manual sets out the process to be followed to deal with any roll up incidents. An increase in the number of roll up incidents on Day 1 could indicate a resource issue with DA staff unable to cope with the volume of incidents. It could however be caused by failures of the TRUST DA system.

Table 5.9 shows the changes in the level of roll up incidents and delay minutes across the study periods.

Table 5.9: Roll Up Incidents by Route

Route	Number of Incidents			Number of Minutes		
	2009/10	2012/13	Change	2009/10	2012/13	Change
Anglia	26	22	-15%	324	307	-5%
East Midlands	12	18	50%	113	1,784	1479%
Kent	73	45	-38%	473	17,490	3598%
London North Eastern	42	45	7%	247	1,352	447%
London North Western	30	10	-67%	354	104	-71%
Scotland	283	128	-55%	2,004	518	-74%
Sussex	32	58	81%	1,716	40,387	2254%
Wales	805	108	-87%	20,697	2,105	-90%
Wessex	33	29	-12%	232	372	60%
Western	1,121	123	-89%	31,730	1,241	-96%
TOTAL	**2,457**	**586**	**-76%**	**57,890**	**65,660**	**13%**

Overall there has been a large reduction in the number of roll up incidents across the network. The major exception to this is Sussex where there has been an increase of over 80% in the number of incidents, although the number remains relatively low. Western and Wales on the other hand have made substantial improvements.

The number of delay minutes within roll ups has actually increased over the period. However this is entirely driven by events on Kent and Sussex routes. Kent had one incident in period 13 2012/13 which incurred 10.5k minutes. Sussex had 42 incidents in period 11 2012/13 which incurred 35K minutes, of which two incidents incurred a total of 15k minutes alone.

5.8.2 Factors for Consideration

The Western and Wales figures in 2009/10 were generally lots of small incidents averaging 30 and 40 minutes respectively. This would suggest that the issue was not related to DA staff being overwhelmed but more to do with a process failure which has since been rectified.

The large increase in delay minutes on Kent and Sussex appear to suggest that both routes had occasions where the DA staff were overwhelmed or there was a major TRUST outage. The Kent increase is largely down to a single event. However, the Sussex increase was caused by a series of significant weather events and the Route team failed to take appropriate action and allowed the delays to roll up rather than creating a management incident. The central team were aware of the circumstances and have discussed with Sussex how to handle future events. However, it does indicate that on these occasions the Sussex level 1 team was overwhelmed by the levels of delay suffered.

5.9 Management Incident Levels

5.9.1 Findings

Management incident levels are created when the TRUST DA staff cannot keep up with the volume of delays occurring across their area of responsibility. These generally occur during major incidents such as severe snowfall. The levels of management incidents gives an indication of whether staffing levels can cope with major incidents but will clearly be also impacted by the frequency of such events

and the numbers of trains involved. It is useful to view this measure alongside roll up delays as discussed previously.

The practices for the creation of management incidents have changed since 2009/10. Now all delays and cancellations are coded OU and so are straightforward to identify. In 2009/10 incidents were usually coded OD, OZ or ZZ and the header would contain the phrase "management incident". NR have extracted 2009/10 data from TRUST which contained these characteristics to allow a comparison to the 2012/13 data. It is possible, though, that the 2009/10 under-report the number of management incidents at the time.

Table 5.10 shows the level of management incidents by route across the study periods.

Table 5.10: Management Incidents by Route

Route	# Incidents Initially Coded to Mgtment Inc.		% Recoded	
	2009/10	2012/13	2009/10	2012/13
Anglia	2	2	50%	0%
East Midlands	0	5	-	80%
Kent	7	2	14%	50%
London North Eastern	0	4	-	50%
London North Western	2	23	50%	70%
Scotland	19	4	0%	0%
Sussex	5	2	80%	50%
Wales	0	1	-	0%
Wessex	1	9	0%	78%
Western	0	7	-	0%
Total	**36**	**59**	**19%**	**53%**

Nationally, there has been a 64% increase in management incidents – although it should be noted that these only represent 0.01% of all incidents on the network. This is despite a significant drop in management incidents in Scotland. The overall increase is substantially driven by the large increase on LNW.

There has, however, been an increase in the number of incidents subsequently reattributed after Day 1 from 19% to 53%, although again this has been driven by the drop in incidents in Scotland (where no incidents were recoded in either year – note, if Scotland was excluded, the proportion of incidents recoded increased from 41% to 57%).

Table 5.11 shows the same comparison by delay minutes.

Table 5.11: Delay minutes attributed to Management Incidents by Route

Route	# Minutes Initially Coded to Mgtment Inc.		% Recoded	
	2009/10	2012/13	2009/10	2012/13
Anglia	154	544	34%	0%
East Midlands	0	2,936	-	72%
Kent	2,469	1,118	55%	68%
London North Eastern	0	2,240	-	32%
London North Western	462	10,764	4%	86%
Scotland	1,293	932	0%	0%
Sussex	12,455	687	100%	53%
Wales	0	141	-	0%
Wessex	44	7,798	0%	89%
Western	0	2,906	-	0%
Total	**16,877**	**30,066**	**82%**	**68%**

The total number of delay minutes has increased by 78% but again with a significant variation in trends by Route. Sussex has had a 95% reduction in minutes attributed to management incidents whilst LNW has had a large increase over the same period. There have been significant increases on Wessex, Western and LNE as well.

5.9.2 Factors for Consideration

Overall the levels of management incidents have increased with a substantial increase in delay minutes. The reasons for this are unclear but the position on LNW in particular should be investigated. One factor may be that the ability to now use the OU code may have inadvertently created an easier route to creating management incidents when TRUST DA staff are under pressure. It does however make identification of such incidents easier and it is likely that the 2009/10 position understates the number of similar events due to the difficulty of identifying them.

The decrease in management incidents on Sussex is in contrast to the increase in the level of roll ups and as already stated they have been instructed that they should ensure that a management incident is created rather than allowing incidents to default to roll ups.

5.10 Overview

The view of the Reporter Team is that the overall levels of delay attribution show an improvement between the two review periods. There are no obvious national trends within the data that highlight significant country wide issues of concern. Most of the measures show a positive improvement since 2009/10. This reflects well on the greater focus on data quality now carried out by NR and increased levels of data integrity checks being carried out both at Route and through the central team as defined in the Performance Measurement Manual. The use of periodic data quality checks have been used to help focus on assisting the Routes in improving data quality and the figures shared with the Reporter Team show an improving trend across the country with areas of concern being tackled.

Table 5.12 shows a summary of the overall position across the measures discussed within this report with only two of the measures showing a negative trend.

Table 5.12: Percentage change in number of Incidents

Route	% Change in FO/TO Incidents	% Change in Initial Z* Coded Incidents	% Change in Takeback Incidents	Change in % of Incidents Recoded	Change in % of TOC Incidents in Dispute (Day 2)	% Change in Roll Up Incidents	% Change in Management Incidents
Anglia	-19%	-23%	-65%	3%	10%	-15%	0%
East Midlands	1%	26%	-55%	-4%	-5%	50%	-
Kent	13%	610%	-72%	-2%	-3%	-38%	-71%
London North Eastern	9%	43%	-33%	-3%	-1%	7%	-
London North Western	19%	-11%	-3%	1%	4%	-67%	1050%
Scotland	8%	-70%	-37%	-13%	-7%	-55%	-79%
Sussex	-35%	185%	-90%	-15%	-25%	81%	-60%
Wales	29%	-42%	-13%	0%	-	-87%	-
Wessex	16%	124%	58%	-1%	1%	-12%	800%
Western	68%	-10%	62%	5%	8%	-89%	-
Total	**8%**	**-1%**	**-24%**	**-2%**	**0%**	**-76%**	**64%**

There is no evidence that there is any reduction in quality which could be attributed to a reduction in staffing levels engaged in TRUST DA activities. In some cases it appears there are slightly more staff involved in this area now than in 2009/10.

However there are some specific areas of concerns on Routes as highlighted in the individual sections and that variability is highlighted above in Table 5.12.

One factor that is apparent from discussions with NR is that at the margins of delay attribution there are varying practices within the industry on how delay attribution is handled. There are differences between Routes and TOCs on how delay attribution is managed and the impact of devolution and the creation of Alliances might cause these to widen further.

The analysis in this section also highlights some wide variations in the numbers of incidents being allocated to certain codes between the Routes. Although we stated at the outset that we would not cross-compare Routes, this is perhaps an area worthy of further investigation. The scale of these variations is not as well understood as they could be and there are no clear boundaries as to how far they will widen. In our opinion, the pressure to save costs within the industry means that this will be an area where NR and TOCs will jointly explore cost reductions. This may well mean agreements being reached to avoid costly investigation of minor or difficult areas such as small, difficult-to-investigate delays. The purist DAG view will state that this is not acceptable but a more pragmatic view is likely to challenge this.

5.11 Recommendations for Phase Two

Within this section, the Reporter team have highlighted certain issues affecting individual routes and operators that are worth further examination. In general these should be investigated by NR and reported back to ORR where they have not already given details noted within this report.

Given the overall improvements in the areas investigated the Reporter team do not feel it would be sensible use of resources to undertake a detailed follow up to the specifics highlighted within this report as originally envisaged at the initial launch meeting.

5.12 Other possible reviews

As noted above, we have observed some wide variations in practice between the Routes. In discussion with NR, it is clear they are looking for guidance on what is acceptable in allowing some degree of variation in attribution practices across the country.

We therefore suggest that the Reporter team explore the emerging arrangements across three Routes alongside their lead TOCs, to understand the drivers behind the variations being observed. It will also give us the opportunity to explore some of the specific changes on these routes observed during CP4 that we have identified in our analysis. The suggested routes for these discussions are:

- Scotland and ScotRail;
- Sussex and Southern; and

- Wessex and SWT ***or*** LNE and Northern.

To carry out the review, the Reporter team would meet jointly with the Route and TOC Performance Managers to understand the reasons behind changes, what direction they see their arrangements taking, and discuss issues picked up in this report. It would also be appropriate to hold discussions with the Chair of the DAB and with representatives from the ORR performance team, to understand their perspectives.

The deliverable would be a report highlighting the reasoning behind the variations and how these affect, if at all, the focus on understanding the causes of delay and delivering improvements. It would set out potential future minimum standards that should be applied in managing delay data and set boundaries within which any future alliances should operate.

Appendix A

Mandate of Review

Mandate for Independent Reporter Part A – Review of performance data

Audit Title:	Review of performance measures
Mandate Ref:	AO/039
Document version:	Final
Date:	17 January 2013
Draft prepared by:	Andy Lewis
Remit prepared by:	Chris Fieldsend
Network Rail reviewer:	John Thompson

Authorisation to proceed

ORR	John Larkinson	
Network Rail	Angelique Tjen	

1 Purpose

This mandate sets out the scope of work for the Part A Independent Reporter (Arup) to review Network Rail's (NR) performance data. As a regulated target, it is critical that ORR has assurance of the quality of this data which offers stakeholders key headlines on industry performance.

2 Background

Arup last reviewed NR's performance data in April 2011. On Public Performance Measure (PPM) and Cancellations and Significant Lateness (CaSL), the review assessed the confidence ratings as (respectively) A1 and A2 (see Annex A below). Since that review (in December 2012) Arup carried out a separate review of Right Time Performance data (RTP). This study raised concerns about the reliability of an undocumented process for manually recorded data. The report stated that there could be implications for the reliability of PPM and CaSL data. ORR would like to know whether this is the case

Additionally, ORR would like assurance on whether an apparent reduction in resources available to 'level one' delay attribution (undertaken on shift, in real time) has impacted the quality of delay minute data. To that end, ORR would like a better understanding of the delay datasets (adjusted and unadjusted) and confidence that the data received from Network Rail is of a high standard. This will enable stakeholders and the public to make better informed decisions.

3 Scope

The review should include:

1. PPM – the implications of the findings of ARUP's right time review[1] (e.g. berthing offsets, manual reporting) on PPM;
2. CaSL – the implications of the findings of the right time review (as above) as well as implications on missed intermediate stations;
3. Delay minutes – level one attributions. ORR is conscious of an apparent reduction in resources committed to level one delay attribution for both Network Rail and Train Operating Companies and requires assurance that this is not impacting data quality. On that basis ORR would like ARUP to review data from areas that have recently changed their delay attribution complement (e.g. York). ORR would like to know if these resource changes could account for an increase in unexplained delay on the Sussex route.

The reporter should briefly review PPM, CaSL and delay minutes KPIs and :

- comment on the reliability, quality, consistency, completeness and accuracy of the reported data;

[1] AO/033 – Review of Right Time Performance (RTP) Data

- present a confidence grade for each KPI and comment on the change since last reviewed in April 2011; and
- report on progress against recommendations made in the April 2011 review and make appropriate recommendations where necessary.

On delay attribution, there are several ways in which a deterioration in quality or accuracy might become evident after a reduction in the amount of resource devoted to level one delay attribution, especially by a TOC. The reporter should therefore comment on:

- whether there has been an increase in the number of occasions when the delay attribution system has become overwhelmed where incidents have to be initially coded into a management incident and subsequently re-coded to their correct incident in the following days;
- whether there has there been an increase in unexplained or un-investigated delays, and what the causes of this are;
- whether there has been an increase in delays coded finally simply to TO or FO (TOC and FOC unexplained) with no information available from operator – i.e. the operator does not contest the delay but has not managed to specifically allocate it to fleet, driver, station or whatever, thereby losing the opportunity to identify root cause and take corrective action);
- whether there has been an increase in 'commercial takeback' due to delay incidents being incorrectly coded to Train Operating Companies and subsequently re-coded to Network Rail or vice versa;
- whether there has been an increase in delays coded to Network Rail and then subsequently, and without sufficient time for a reasonable investigation to have taken place, re-coded to a Train Operating Company. Special attention should be paid to incidents which are eventually resolved back to Network Rail or subject of resolution by means of a 'commercial split';

Answers to these questions should include a review of data "before and after" backed up by a few interviews (which could be by telephone rather than site visits).

4 Methodology

The Reporter should meet with relevant Network Rail employees to understand any procedural changes to the processes used to report the above KPIs since the April 2011 review. The Reporter should also review all relevant documentation and systems, and comment on their quality and fitness for purpose. The Reporter should draw on (and not duplicate) work previously undertaken in their review of RTP. The Reporter's proposal should articulate clearly how they will address each element of the above scope.

5 Deliverables

The Reporter should provide a publishable report, including findings, conclusions and recommendations (explaining the reasons for them and the benefits if implemented along with timescales for completion). 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.

6 Timescales

A fully costed proposal for this work is required by 31 January 2013. Work is expected to commence shortly after following approval by NR and ORR. A draft report is required by 29 March 2013 and a final report is required by 30 April 2013.

7 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 (including consideration of the mix of seniority and skills required) and costs.

8 Confidence grades

The Independent Reporter shall provide a confidence grade for the PPM and CaSL measures as well as for NR delay minutes to TOCs and NR delay minutes to FOCs per 100 train km. 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.

Confidence grades should be provided for:

- PPM (% MAA) England & Wales long distance
- PPM (% MAA) England & Wales London & South East
- PPM (% MAA) England & Wales Regional
- PPM (% MAA) England & Wales Total
- PPM (% MAA) Scotland (ScotRail)
- CaSL (%MAA) England & Wales long distance
- CaSL (%MAA) London & South East
- CaSL (%MAA) Regional
- Delay mins – passenger (000's) England & Wales
- Delay mins – passenger (000's) Scotland (ScotRail)
- Delay mins per 100 train km – freight

9 Conflict of interest

The Reporter should explicitly highlight any conflicts of interest.

10 ARUP quality assurance

The Reporter should describe the internal processes in place to quality assure the work delivered under this mandate.

11 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 measures under review.

Measure	Benchmark grade
PPM	A1*
CaSL	A1
Delay Minutes - Level One attribution	A1

Appendix B

Spot Checks of Sample TOC PPM & CaSL Data

This Appendix provides examples of the checks of data flow carried out by the Reporter team on Network Rail's performance data for calculating PPM and CaSL. These checks are based on 2012/13 Period 11, and are broad repeats of prior checks of the process. A sample of 5 TOCs was selected at random, although ensuring at least one TOC from each sector was included, and at least one TOC operating across more than one sector was included.

Flow of Data from 'PSS Data Export Spreadsheet' to 'TOC Reconciliation Spreadsheet'

The Network Rail spreadsheet '*PPM & CaSL (1) TOC-PSS.xls*' contains separate worksheets for individual TOCs and Sectors. The PSS cancellations data and the number of trains run are linked by formulae from the PSS export ('*PPM & CaSL (0) All Day P1311 20130311.xls*') to this reconciliation spreadsheet. Checks carried out have shown that the formulae are accurately set up, as shown in the table below.

TOC	PPM & CaSL (0) All Day P1311 20130311.xls			PPM & CaSL (1) TOC-PSS.xls			Difference	Comment
	Total Cancellation	Part Cancellation[5]	Trains Run	Total Cancellation	Part Cancellation	Trains Run		
Northern Rail	730	566	62,191	730	566	62,191	0	OK
First Great Western (LSE)	307	210	20,066	307	210	20,066	0	OK
First Great Western (Regional)	181	205	13,107	181	205	13,107	0	OK
First ScotRail	364	341	57,248	364	341	57,248	0	OK
East Coast	84	70	3,709	84	70	3,709	0	OK

Source: 'PPM & CaSL (0) All Day P1311 20130311.xls', 'PPM & CaSL (1) TOC-PSS.xls'

[5] In all tables, part cancellation includes trains which were over 120 minutes late

TOCs Cancellation Data

Spot checks on the data supplied by TOCs and that which has been reported by Network Rail were carried out for the TOCs shown in the table below. The data supplied for this review covered 2013/P11. The spot checks carried out, as listed below, indicate the TOC cancellation data has been accurately inputted into the spreadsheet.

TOC	TOC Provided		Network Rail Reported		Difference	Comment	TOC Data Source
	Total Cancellation	Part Cancellation	Total Cancellation	Part Cancellation			
Northern Rail	731	553	731	553	0	OK	_Sch 7 1 Benchmarks DfT - Northern 2012 05
First Great Western (LSE)	306	211	306	211	0	OK	PHIS Export PPM Figures-2012-13 By Day.xls
First Great Western (Regional)	184	203	184	203	0	OK	
First ScotRail	358	335	358	335	0	OK	Bwop1213 1.pdf
East Coast	84	66	84	66	0	OK	Network Rail Performance Sheet.xls

Source: 'PPM & CaSL (1) TOC-PSS.xls', TOC data as indicated above

Note that this table shows, for these TOCs, the magnitude of difference in TOC reported figures (as above) with the initial figures extracted from PSS (in table in previous subsection). This is summarised in the table below for quick reference. It is the TOC provided figures which are used for reporting, and the responsibility of the Route to review any significant differences in these figures.

TOC	PSS Extracted		TOC Provided		Difference	
	Total Cancellation	Part Cancellation	Total Cancellation	Part Cancellation	Total Cancellation	Part Cancellation
Northern Rail	730	566	731	553	1	13
First Great Western (LSE)	307	210	306	211	1	1
First Great Western (Regional)	181	205	184	203	3	2
First ScotRail	364	341	358	335	6	6
East Coast	84	70	84	66	0	4

Calculation of PPM and CaSL

Values from the PSS Data Export spreadsheet are also linked by formulae to the calculation spreadsheet (PPM & CaSL (2)). The values in the table below have been checked to ensure accuracy and have been found to be robust. Checks have also been carried out to confirm the TOC cancellation data has been correctly linked to this spreadsheet from the 'PPM & CaSL (1)' spreadsheet, and found to be consistent (although details not shown in this table).

Train Numbers	PSS Export ('PPM & CaSL (0) All Day P1311 20130311.xls')					PPM & CaSL Calculation ('PPM & CaSL (2) Consolidation.xls')				
	Northern	FGW (LSE)	FGW (Reg)	ScotRail	East Coast	Northern	FGW (LSE)	FGW (Reg)	ScotRail	East Coast
PPM Passes	56,602	18,411	12,077	53,747	2,760	56,602	18,411	12,077	53,747	2,760
Within 15	61,097	19,779	12,881	56,966	3,023	61,097	19,779	12,881	56,966	3,023
15-20 Late	457	126	98	151	203	457	126	98	151	203
20-30 Late	425	92	78	82	227	425	92	78	82	227
30-61 Late	190	57	45	42	205	190	57	45	42	205
61-120 Late	22	12	5	7	51	22	12	5	7	51

Source: 'PPM & CaSL (0) All Day P1311 20130311.xls', 'PPM & CaSL (2) Consolidation.xls'

Train Numbers	Difference (from above table)				
	Northern	FGW (LSE)	FGW (Reg)	ScotRail	East Coast
PPM Passes	0	0	0	0	0
Within 15	0	0	0	0	0
15-20 Late	0	0	0	0	0
20-30 Late	0	0	0	0	0
30-61 Late	0	0	0	0	0
61-120 Late	0	0	0	0	0

Calculation of Sector Level PPM/CaSL

The calculation spreadsheet is set up with a reference worksheet ('*Master*') which contains a list of all TOCs and the relevant Sector they operate within (and for those TOCs which operate in more than one sector, these are explicitly separated, e.g. First Great Western LD, First Great Western LSE and First Great Western Regional).

These values are then transposed onto a worksheet called '*Template*'. The TOC values for each of the measures, as summarised in each of the individual measures calculation sheets (e.g. "Within 5"), are then multiplied with the values on the '*Template*' sheet to obtain the aggregated sector values. The calculation also includes input from the sheet called '*Timelines*' which takes into account the validity of franchise dates.

Auditing of the process of amalgamating TOC figures to produce sector results focussed on two sample measures, '*PPM*' and '*CaSL'*.

The values from individual TOC sheets were checked to ensure they had accurately been fed through to the 'individual measures' worksheets. These checks are summarised in the table below for the same four sample TOCs, which showed no concerns.

TOC	From Individual TOC Worksheet		To Measures Worksheet		Difference	Comment
	PPM Passes	CaSL Trains	PPM Passes	CaSL Trains		
Northern Rail	56,602	1,496	56,602	1,496	0	OK
First Great Western (LSE)	18,411	586	18,411	586	0	OK
First Great Western (Regional)	12,077	437	12,077	437	0	OK
First ScotRail	53,747	742	53,747	742	0	OK
East Coast	2,760	406	2,760	406	0	OK

Source: 'PPM & CaSL (2) Consolidation.xls'

These figures are then fed into the Sector level figures as reported by Network Rail.

Appendix C

Data Assurance in the Context of Right Time Performance Matrix

C1 Detailed Methodology

The methodology used for the PPM and CaSL calculations is essentially the same although different datasets have been utilised. In both cases the data has been aggregated by reporting method, operator and sector across all periods. The error band of +/- 13 seconds has been applied against those services measured by automatic reporting, while the error band of +/- 35 seconds has been applied against those services measured by manual reporting. In this analysis it is assumed that the error distribution is symmetric around a mean of zero. Furthermore, the error distribution is assumed to be uniform. If the distribution was normal then the impact would be significantly less than calculated here since the errors would be clustered close to the mean of zero.

As the error bands extend less than one minute on either side of the recorded arrival times, the one minute period either side of the target (or threshold) time has been considered. With a PPM target of 5 minutes for the London & South East, Scotland and Regional services, the area of interest is arrivals between 4 minutes late and 5 minutes and 59 seconds late. For Long Distance services where the PPM target is 10 minutes, the area of interest is between 9 minutes late and 10 minutes and 59 seconds late. In the case of CaSL where the threshold is 30 minutes for all services, the area of interest is between 29 minutes late and 30 minutes and 59 seconds late.

Given the respective error bands there is a chance that a train arriving in the minute prior to the threshold was in fact late (a false positive). Conversely, there is a chance that a train recorded as arriving in the minute after the threshold was in fact early (a false negative). The false positives and the false negatives will tend to cancel each other out, so the net potential value is required to estimate the impact on PPM or CaSL.

A three-step process was used:

1. Estimate net potential errors;
2. Calculate expected net errors; and

3a. RT error percentage = expected net errors / number of trains run;
3b. PPM error percentage = expected net errors / number of trains run; or
3c. CaSL error percentage = expected net errors / number of trains planned.

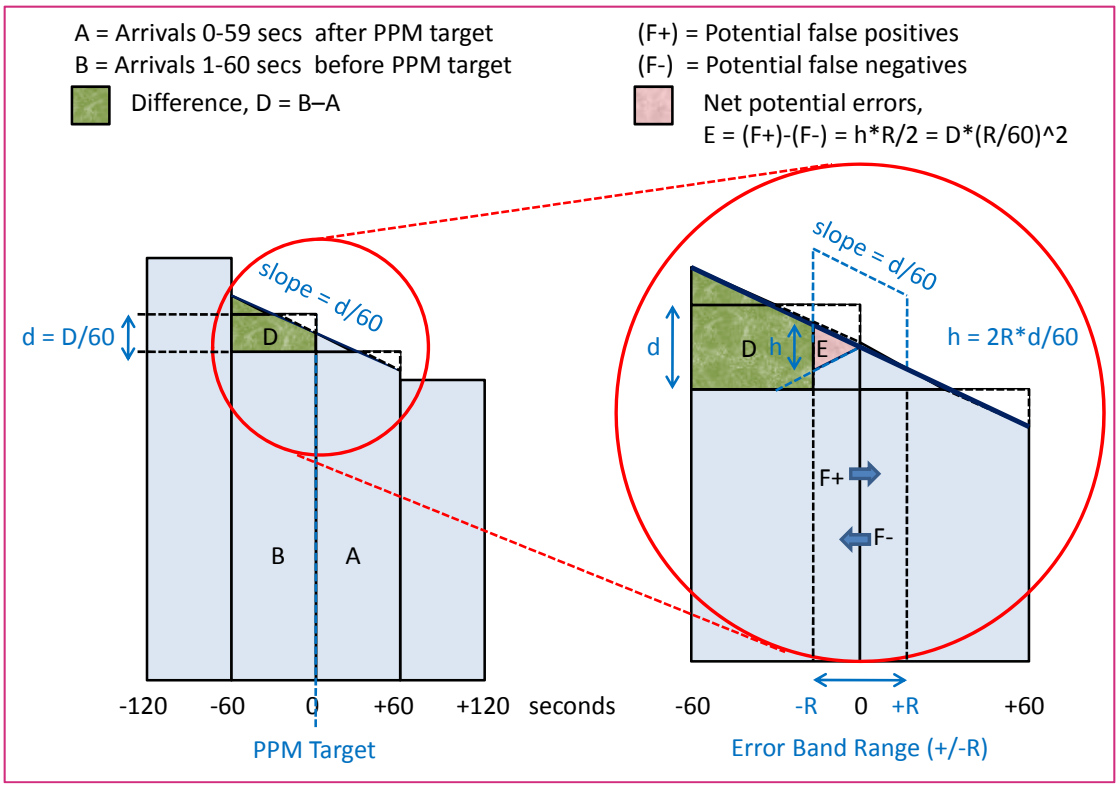
In Step 1, as illustrated below, given the shape of the arrival distribution and the timings of interest, more trains will generally arrive in the minute before the threshold than the minute after. Thus, area B is greater than area A in the illustration on the left. Area D represents the difference. Assuming the arrival rate drops gradually over the 120 second period of interest, the slope of the line is d/60, where d=D/60

In the magnified illustration on the right, the error band range of +/- R provides a bound to the potential false positives (F+) and false negatives (F-). As area B is greater than area A, more false positives are expected than false negatives and area F+ exceeds area F-. Under the assumption that the error distribution is symmetric most of the errors in recording will cancel each other out.

The net potential errors (represented by the triangular area denoted E) is equal to the number of false positives minus the number of false negatives. The error

triangle has a height (h) of 2R*d/60 and a base of R. The area of the triangle is calculated as half the base times the height or R/2*h. Using substitution, area E = D*(R/60)^2 = D*R^2/3600.

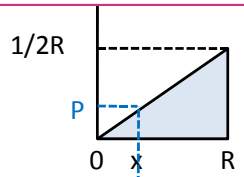
Step 1: Estimate net potential errors - PPM



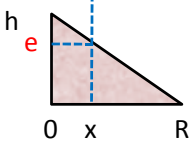
In step 2, as described below, numerical integration is used to weight the potential errors according to the probability that the error for an individual train x is large enough to require reclassification. In other words, a probability density function is used to calculate the expected number of errors. For convenience, arrival times between -R and 0 are relabelled from 0 to R to avoid an excess of minus signs in the algebra.

Step 2: Expected net errors (using numerical integration)

i) Probability density, P
= p(error exceeds gap between arrival and target)
= x/2R (assuming uniform distribution of errors)



ii) Net potential errors, e
= h * (1 – x/R)
= 2R * (D/3600) * (1 – x/R)
= 2D/3600 * (R-x)



iii) Expected net errors = Integral {Net potential errors * Probability density} , x = 0 to R
= Integral { (2D/3600) * (R-x) * x/2R }, x = 0 to R
= (D/3600) * Integral {x – (x^2/R) }, x = 0 to R
= (D/3600) * {x^2/2 - (x^3/3R) }, x = 0 to R
= (D/3600) * {R^2/2 - (R^2/3)} – {0 – 0}
= (D/3600) * (R^2/6)
= E/6

Assuming the error distribution is symmetric, at time 0 (relabelled R) there is a 50% chance of an error either way, hence a 50% chance of reclassification. An error in the opposite direction will not generate a false positive so half of the density function can be discarded as irrelevant. Assuming the distribution of errors is uniform, the probability that an error warrants reclassification will increase linearly between 0 and R seconds, reaching a maximum of 1/2R as illustrated by the upper triangle in step 2. Thus, the density p = x/2R and the area of this triangle is 0.5.

The number of net potential errors (e) as derived in step 1, is represented by the lower triangle in step 2 and although it now has a different shape with the height adjusted to start from zero, it has an equivalent area. Assuming a simple linear fit to the arrivals distribution, the number of potential errors will decrease linearly from h to zero over R seconds.

Multiplying the values of p and e gives a peak at R/2 and zero probability at 0 and R. The expression is a quadratic curve (parabola) and the area under the curve which represents the expected net error, can be calculated using numeric integration. The result is to reduce E by a factor of 6, implying that 1 in 6 potential errors are expected to be large enough to require reclassification, thus producing the expected net errors.

The worst case estimates have been obtained by ignoring the possibility of false negatives and assuming all arrivals within range R of the target are false positives. In this case the worst case errors are D*R/60.