Subject	AO/037: Review of Network Availability Alternative Metrics- Final		
Date	8 March 2013	Job No/Ref	223767-08

1 Introduction

This technical note reviews the alternative metrics that have been proposed by Network Rail to measure the level of passenger and freight disruption caused by engineering possessions. These metrics have been developed by an industry wide workgroup and have been described in a series of documents referenced at the end of this technical note.

The ORR and Network Rail have asked us to give our opinion on the methodology to calculate these metrics, specifically to answer:

1. How accurately do the new metrics report network availability for passenger and freight?
2. Is there evidence of buy-in from Network Rail's customers?
3. How accurate is the proposed input data (including passenger numbers) and what impact could poor quality input data have on the high level?

Timescales have not permitted a detailed investigation. Instead our opinion is based on:

- a meeting held with Network Rail on the 6th February 2013,
- a meeting with ORR on the 5th March 2013 primarily to discuss the appropriateness of using the station footfall data in the passenger disruption metric,
- telephone interviews with two TOC and one FOC representatives, and
- a review of the supporting spreadsheets and documentation provided by Network Rail.

2 Proposed Metrics

2.1 Passenger Disruption Measure

The general aim of the proposed measure is to compare the Working Timetable (WTT) run on a particular day against a base WTT which is unaffected by possessions.

The choice of a base timetable is unclear in the documentation. References are made to using an annual WTT, the London Olympics WTT and the T-18 WTT. In our meeting with them, Network Rail wanted to make sure that the base was operational and did not contain too many schedules that could not be covered by rolling stock and crew resources.

The disruption of the daily WTT is measured in three ways against the base WTT:

1. The number of cancellations,
2. The number of stations missed on the remaining schedules, and
3. The extended origin station – destination station journey times of the remaining schedules.

The formula is shown below, taken from the document 'CP5 Network Availability Metrics methodology (draft), dated January 2013'.

$$WTT\text{-}P = \frac{OF_{BASE\ TIMETABLE} + weighted OV_{T\text{-}2}}{OF_{BASE\ TIMETABLE}}$$

OV weighting = Weighted missed stations + weighted extended journey ti me

The formula expanded

$$WTT\text{-}P = \frac{\sum OF_{BT} - \sum_{OV_{T\text{-}2}} \left\{ w_m \sum \frac{f_s}{\sum_v f} + w_t \frac{EJT_v}{OJT_v} \right\} - \sum CA_{T\text{-}2}}{\sum OF_{BT}}$$

What the terms mean:

OF = Trains offered in the schedule ***OV***= Variations to train schedules

BT = The base timetable ***CA***= cancelled train schedules

EJT_v = extended journey time of this train OJT_v = original journey time of this train

w_m = weighting for missed stations f_s = footfall for station s

w_t = weighting for extended journey time ***s*** = this station ***v*** = this train

In order to try to understand how this applied, consider a simple route with the following base and operational WTTs.

Station	Footfall	Base				Operational			
		Train 1	Train 2	Train 3	Train 4	Train 1	Train 2	Train 3	Train 4
A	10,000	0700	1100	1500	1900	-	-	1300	1700
B	20,000	0730	-	1530	-	-	-	1330	1730
C	30,000	0800	1150	1600	1950	-	-	1400	1800
D	40,000	0830	1220	1630	2020	-	-	1430	1830
E	50,000	0900	1250	1700	2050	-	-	-	-

The calculations for the passenger disruption measure are as follows.

Description	Term	Train 1	Train 2	Train 3	Train 4	Total
Base WTT – trains offered in schedule	OF_{BT}	1	1	1	1	**4**
Operational WTT – cancelled schedules	CA_O	0	0	1	1	**2**
Missed station (weighted by station footfall)		0	0	0.33[1]	0.33[1]	
Extended Journey Time (compared to Base schedule)		0	0	0	0.06[2]	
Variation to Operational Schedules	OV_O			0.33	0.39	**0.72**

[1] $(50{,}000 * wm) \quad / \quad (10{,}000 + 20{,}000 + 30{,}000 + 40{,}000 + 50{,}000)$

where $w_m = 1$

[2] $w_t * ((1830 - 1700) - (2020 - 1900)) \quad / \quad (2020 - 1900)$

where $w_t = 0.5$

Therefore WTT-P $= (4 - 2 - 0.72) \quad / \quad 4$

$= 32\%$

Two points to note are:

- Individual trains are compared in this calculation. This is straightforward if train headcodes can be meaningfully matched between the two timetables, i.e. the operational timetable is a variant of the base timetable (for example by removing some trains or terminating them short). It becomes less clear how this would work if the operational timetable is re-written and matching trains becomes more complicated.
- The additional station stop of train 4 at station B is not counted as a benefit.

2.2 Rail Bus Replacement

This metric aims to measure the number of rail replacement bus service schedules during engineering works. The aim is to incentivise keeping passengers on trains and minimising the deployment of replacement bus services.

The calculation of the measure is:

$$\text{Rail Bus Usage} = \frac{\sum_{SG} BR_{T-1}}{\sum_{SG} OF_{BASE\ TIMETABLE}} \text{ per day}$$

What the terms mean:
BR_{T-1} = Number of bus schedules planned
$OFT_{BASELINE\ TIMETABLE}$ = Number of trains in base timetable
SG – *Train service group*

The metric is therefore the number of bus schedules expressed as a percentage of train schedules in an annual 'clean' base timetable without any bus schedules for that day. It is calculated for each service group on a daily basis, and the intention is to produce a report on a four weekly basis.

2.3 Freight Disruption Measure

The general aim of the freight metric is to measure the % of time that a route (or its specified diversion) is open and available for use by freight trains. It is calculated on a daily basis. 100% availability means that the primary freight route is open for 24 hours' operation.

The measure is calculated for each of the 24 freight routes specified in the Joint Network Availability Plan (JNAP). For each one, a diversionary route (if available) is specified and a weighting is applied to reflect its capability at handling all the freight traffic from the primary route.

The equation to calculate the disruption measure is

$$\text{WTT-F} = \frac{\sum \text{Hours best route available} * \text{Weighting for route}}{\textit{24 (or availability hour specified for flow in JNAP)}}$$

where the best route is either the primary route (with a weighting of 1) or the diverted route (with a specific weighting of <= 1).

It does not rely on a Working Timetable but instead calculates the opening hours from an analysis of possessions. Currently this is a manual process which is time consuming and Network Rail are looking to automate it to make it a viable measure.

3 Accuracy of proposed measures

3.1 Passenger Metric

In considering accuracy, we must first define what we are trying to measure. In the 'Network Availability Metric – Revised Working Timetable Compliance metric (Passenger and Freight) Draft' document dated January 2013, the following definition is provided:

> *"Working Timetable Compliance Measure*
>
> *The timetable based metric for passengers measures network availability as the impact of possessions based on the excess journey times or train service unavailability experienced by passengers in CP5. The concept is to align the metric with current processes and metrics such as PPM, Schedule 4 Compensation and delay attribution."*

In the method illustrated in the simple example above, *train service unavailability* is measured in terms of the number of cancellations. Whilst this is straightforward to calculate, it should be noted that a cancelled train has equal weighting irrespective of its length of journey or attractiveness to passengers. It is, therefore, a simplification, the importance of which should be tested.

The *excess journey times* calculation focuses on the origin – destination journey time so again is a simplification since it does not take account of the number of passengers affected. In addition, a weighting factor of 0.5 is applied to minimise the possibility that an extended journey time receives a weighting in excessive of a train cancellation (i.e. it should be less than 1). The logic of this is sensible though the choice of 0.5 does not appear to be based on any hard evidence.

As mentioned above, the calculations are less clear when a new timetable is written which makes it difficult to compare trains. Also, additional station stops inserted into train schedules to compensate other train cancellations are not taken into account.

The method also fails to take into account impacts on passengers changing trains. The additional waiting time at an interchange station caused by a train cancellation or late running train is not included in this measure.

Overall, and within this short review, it is not possible to state an accuracy level of this measure. We have pointed out some simplifications above which could be refined in any further development by Network Rail.

3.2 Bus metric

The key assumption in this measure is that all bus schedules cause the same disruption to passengers. In practice this is not the case, and a schedule can involve more than one bus if the train that has been replaced carries many passengers. In addition, no account is taken of the length of the bus schedule, whether this is in terms of distance or journey time. It is therefore a fairly crude measure of disruption to passengers.

3.3 Freight Metric

A challenge for this metric is the identification of relevant possessions that impact a route's availability. Whilst each primary route has a specified diversionary route (with associated weighting), care will need to be taken when reviewing the location of possessions. For example, if only one of two lines is closed, it is possible that the route could still be 100% open to freight traffic without the use of the diversionary route. In contrast, a possession could close both the primary and diversionary routes. We understand that this process is still under development.

Producing a national measure from the 24 constituent routes will require consideration, in particular how to weight each of the routes.

4 Views of Network Rail's Customers

4.1 Interviews

A series of interviews were held with TOC and FOC representatives to discuss the proposed network availability measures. The purpose of the interviews was to check the level of involvement the interviewees had with the process of creating the new measures, issues they have with the current PDI measures, and their understanding and support for the proposed measures.

Interviews were held with representatives from:

- Virgin Trains,
- First Great Western, and
- DB Schenker.

4.2 WTT Compliance - Passenger

4.2.1 Involvement with the review process

First Great Western had been closely involved in the development of the metric through being a member of the working group. The Virgin Trains representative had no involvement in the development and was therefore unfamiliar with the detail of the proposed measure.

The First Great Western representative said that Network Rail had recently given a presentation to the industry Operational Planning Steering Group on their intentions. The steering group did all its work last summer and he has not since then been significantly involved.

4.2.2 Thoughts on the current metric

Both felt that the current PDI–P measure was unhelpful to them and the actual metric was very difficult to understand. Both wanted to see a simple measure akin to PPM. Virgin are very unhappy with PDI-P because they believe it does not reflect the impact of engineering work enshrined in the timetable. For example on Sunday mornings all their trains go via Northampton to allow maintenance work between Hanslope and Rugby. PDI-P assumes this is the normal route so the measure does not show it as a change to their preference to go via the main line.

4.2.3 Views on the new measure

Both broadly support the need for a "simpler" measure.

The First Great Western representative described the need for a measure that rewarded the use of a train over buses i.e. divert or use Single Line Working (SLW) rather than simply terminate and bus to destination. He described a strong desire to use available data rather than spend a long time inventing new data processes.

This fed the desire to use footfall data. This was seen as a reasonable way of ranking the importance of stations but recognised as not precise in managing to differentiate days of the week

etc. The steering group looked at using revenue data but did not pursue it because of the confidentiality issues.

Virgin felt that reflecting somehow the value of different times of the day/week would be useful. This is particularly important since the impact on the West Coast on a Sunday is much greater than on a route like Essex Thameside.

The base timetable is seen as a critical issue. Network Rail had suggested using the WTT as a base but the TOC representatives on the steering group rejected this as it embedded a lot of engineering work. The Olympic Timetable was suggested as this was free of engineering work and seen as a reasonable proxy for a disruption free network. First Great Western described this as a major debating point.

Both TOC representatives felt that the First WTT used in Schedule 4 could be used as the base. This is a timetable which is agreed between the TOC and Network Rail as representing what the TOC would run on each day of the week when there is no impact at all from engineering work. For many TOCs this simply involves agreeing an actual day's operation as the base. For other operators, particularly Cross Country this involves creating a bespoke timetable which is unlikely to ever operate in reality.

Overall the First Great Western representative was keen to see a new measure in place which can use existing data sets but recognised it will not be perfect.

Virgin want a measure that challenges Network Rail during the midweek as well as weekends, particularly as Integrated Access Planning (an RDG initiative) rolls out. It should also properly reflect the impact of major blockades which hit midweek periods.

4.2.4 Rail Replacement Bus Measure

Due to time constraints this was only discussed with the Virgin Trains representative.

He did not feel just using bus schedules would give a robust measure. He stated that typically they will run a mix of fast, slow and stand-by buses and they will not all be on a timetable schedule. He feels bus hours are more reflective but we pointed out that this relies on TOCs giving Network Rail their data and this is unlikely to happen.

He felt there should be some form of weighting – possibly based on mileage. This has a major impact on what type of buses they run and the mix of schedules. The measure ought to reflect somehow the number of passengers affected. A single taxi or 6 buses could each count as one schedule in the measure.

4.3 WTT Compliance – Freight

4.3.1 Involvement with the review process

The DB Schenker representative has been involved in the freight steering group although he said he had not had involvement since a presentation by Network Rail last October where the outputs for the Peterborough to North London Lines flows were shared.

4.3.2 Thoughts on the current metric

He described the current measure as irrelevant to their business and little if any reference was made to it.

4.3.3 Views on the new measure

He was very supportive of the use of the JNAP flows as the basis of the new measure. This will allow reasonably detailed management of the most important flows within their business, and for other operators.

The likely outputs were seen as useful to share with their customers on the flows to demonstrate what they are doing with Network Rail to manage end delivery times.

He was very keen that the measure should cover midweek nights as in the proposal. DB Schenker are concerned about the impact of the workbank in CP5 and the drive towards more midweek access and want the measure to reflect this properly. It must also be capable of reflecting Adjacent Line Open/SLW reasonably accurately.

He described JNAP flows as reasonably fixed and unlikely to change radically. This should not, therefore, pose a major risk to the continuity within the numbers.

There have been no discussions with DB Schenker to date on weighting of the 24 flows. This is likely to prove a challenge as different companies will have different views.

5 Accuracy of Input Data

5.1 Stations Footfall

The station footfall figure is used to weight the impact of changed train schedules on passengers. It is taken from ORR's publicly available data which is updated on an annual basis. The ORR's website describes the data in the following way:

> *"The latest station usage information is based on ticket sales in the financial year 2010-11 and covers all National Rail stations throughout England, Scotland and Wales. The station usage figures are subdivided by ticket type (full, reduced and season tickets), whilst information on the county and region of each station is also provided.*
>
> *Station usage data are an estimate of the number of passengers travelling to and from each station (entries and exits). They are based on ticket sales data from the national ticketing database and estimates of travel using zonal/multi-modal tickets sold by Strathclyde Partnership for Transport and English Integrated Transport Authorities (formerly PTEs). There are limitations to the dataset and these estimates should be treated with caution."*

As indicated above, the largest source of error is the approximation in the use of travelcard / multi-modal tickets which do not have specified origin and destination stations for all journeys.

There is no split between weekdays / Saturdays / Sundays. This is potentially important for the proposed passenger metric since it will not distinguish the relative importance of routes on different days of the week. For example, weekends will have equal weighting as weekdays on a commuter route. In our view this is a concern. We note that it might be possible to apply day of week factors that are assumed in MOIRA, for example, which could be investigated further.

The data separately identifies journeys that start, end and interchange at a station. ORR advise that the numbers of interchange journeys will be an under-estimate since they exclude journeys on flows below a certain threshold (to reduce the amount of processing). The spreadsheet provided by Network Rail that tests the proposed metric uses the 2009/10 start and end totals (we note that 2010/11 is now available) and ignores the interchange journeys.

The final point we note is that using the station footfall data ignores where passengers are travelling from and to. For example, on commuter routes to London, many of the passengers getting on at country end stations will be getting off at London. This means that if the route is closed at London then this will affect most passengers, more than the station footfall weighting relative to the whole route would suggest.

5.2 Timetables

It is proposed that these will be produced in ITPS. As such, they should be accurate with no major conflicts.

6 Conclusions

There is stakeholder support, at least in principle, for the passenger and freight disruption measures being proposed. It is clear, though, that further development of the measures is required.

6.1 Passenger Disruption Measure

Its benefit over PDI-P is its simplicity and clearer connection to possessions planning. There is, though, a balance to strike between simplicity and accuracy and some assumptions have been made that reduce accuracy. This includes the use of station footfall which does not distinguish weekday from Saturday and Sunday and which could be significant on some routes. One of the TOCs interviewed raised this as a concern.

There might also be practical difficulties in calculating the measure if the engineering timetable is a more of a recast than a variation to the Base timetable. Based on the documents we have reviewed, we believe that there is some scope for refining the calculations and testing their accuracy under a number of test cases.

The choice of the Base timetable is still to be decided. The TOCs suggested that the London Olympics or Schedule 4 timetables could be used, whereas Network Rail have suggested the T-18 WTT.

6.2 Bus Replacement Measure

This is a simple measure that aims to drive down the number of rail replacement services. The one TOC we interviewed on the Rail Replacement Bus Schedule measure felt that using a count of bus schedules would not be a robust measure. It does not adequately measure the number of passengers affected or the disruption they suffer. A possible addition to the measure would be some form weighting, for example mileage.

6.3 Freight Disruption Measure

The main challenge for this metric is automating the process of identifying relevant possessions and their impact on the primary and diversionary routes. If this can be solved, it would be welcomed by the FOC that we interviewed.

6.4 Our opinion

The proposed measures have merit in meeting their stated aims of measuring the end user impact of planned timetable changes and being easy to understand and interpret by the industry. There is a balance to strike between simplicity and accuracy, and we believe that all measures should be developed and tested further to better understand this trade off. In terms of input data, consideration should be given to applying a day of week weighting to the station footfall data to distinguish weekdays from Saturdays and Sundays.

7 References

- Network Availability Metric – Revised Working Timetable Compliance metric (Passenger and Freight), Draft, January 2013
- Network Availability CP5 Review, Network Rail presentation, 6th February 2013
- Joint Network Availability Plan – Freight version 18
- CP5 WTT passenger metric model – P10 2012_13.xls spreadsheet
- CP5 WTT metric passenger data P10 2012_13.xls spreadsheet