Timetable Analysis of the Netherlands Railways (NS)
A new systematic analysis method for the
causes of timetable punctuality deviations

[1]F. Makkinga, [2]B.W.V. Stam-Van den Berg, [2]C. van de Geijn
Movares, Utrecht, the Netherlands[1]; Nederlandse Spoorwegen (NS), Utrecht, the Netherlands[2]

Abstract

This article covers the development of a new method for analysing and assessing a rail transport operator's timetable performance. Reliable and predictable timetable performance is a major factor in the service provided by a transport operator. An important measure of this performance is punctuality. Punctuality expresses the performance delivered by the transport operator in terms of trains arriving on time at their destination. The Dutch rail network is used by several passenger and freight carriers, with more than five thousand passenger trains and over five hundred freight trains running every day. The widely varying train characteristics and the sheer intensity of train transport make the Dutch rail network one of the busiest in the world. Many different things may cause a train to arrive or depart late at or from a station. The reason for investigating the situation is to shed light on what might detract from punctual operation. The output will enable analysts to determine the underlying causes with greater accuracy and speed, and find ways of improving train transport performance. The transport operator's customers will then be more likely to arrive on time at their destinations.

Introduction

The rail transport operators' customers want to arrive on time at their destinations. Reliable and predictable timetable performance is a major factor in the service provided by a transport operator. An important measure of this performance is punctuality. Punctuality expresses the performance delivered by the transport operator in terms of trains arriving on time at their destination. A train is defined as being on time if the scheduled and actual times differ by less than three minutes. Punctuality is expressed as the proportion of trains that meet this criterion. Punctuality is measured at a limited number of relatively large stations. The government negotiates annual punctuality agreements with the passenger transport operator as part of the transport concession. It is therefore in the transport operator's interest to understand why trains arrive more than three minutes late. One important distinction is whether a train arrives too late for a persistent underlying reason or because of an incident. The current analyses that are used to gain this understanding are performed manually. These analyses are awkward and labour intensive in everyday practice. The objective of the research covered in this article is to produce a system to support the analyst, making the analysis process both more reliable and faster.

The article starts with an explanation of the complexity of train transport in the Netherlands, and continues by describing the schedule and actual data available. The research and the applied analysis methodology are then discussed. A case is constructed to illustrate the use of the method in practice and the benefits for the transport operator. The article concludes with suggestions for future development.

The complexity of the transport process in the Netherlands

The Dutch rail network is used by several passenger and freight carriers. It has four thousand kilometres of double track and two thousand kilometres of single track, with over three hundred stations. Nederlandse Spoorwegen (NS) is the largest passenger transport operator in the Netherlands, carrying some one million customers on more than five thousand intercity and local passenger trains every day. The maximum passenger train speed on most sections of track is 140 kilometres an hour, with 160 kilometres an hour permitted on a few sections.

There is also daytime freight transport, with over five hundred freight trains every day. Goods are carried from the ports of Rotterdam and Amsterdam and the industrial areas to destinations within the country and abroad, including via Germany to Italy. A freight train may be up to five thousand tons and over seven hundred metres long. Freight trains travel at up to about 80 kilometres an hour.

The Dutch rail network is also used by regional and international passenger carriers. The first category operates on the secondary lines in the regions. The second category operates services to major cities in mostly Germany, Belgium and France.

The widely varying train characteristics and the sheer intensity of train transport make the Dutch rail network one of the busiest in the world.

Current transport data

Timetables for properly regulating all train movements are drawn up on multiple levels. The process starts with an overall recurring basic hourly pattern and ends with schedules for every train movement in terms of time and space. These schedules specify the paths and routes of the trains over the rail network. All the details of paths and routes, including the use of the tracks and the points, are covered. The timing of preparing train paths is also predetermined. ProRail is the official body responsible for the safe running of trains [1]. The traffic management unit ProRail-Railverkeersleiding ensures the timely operation of points and signals for scheduled passenger and freight transport.

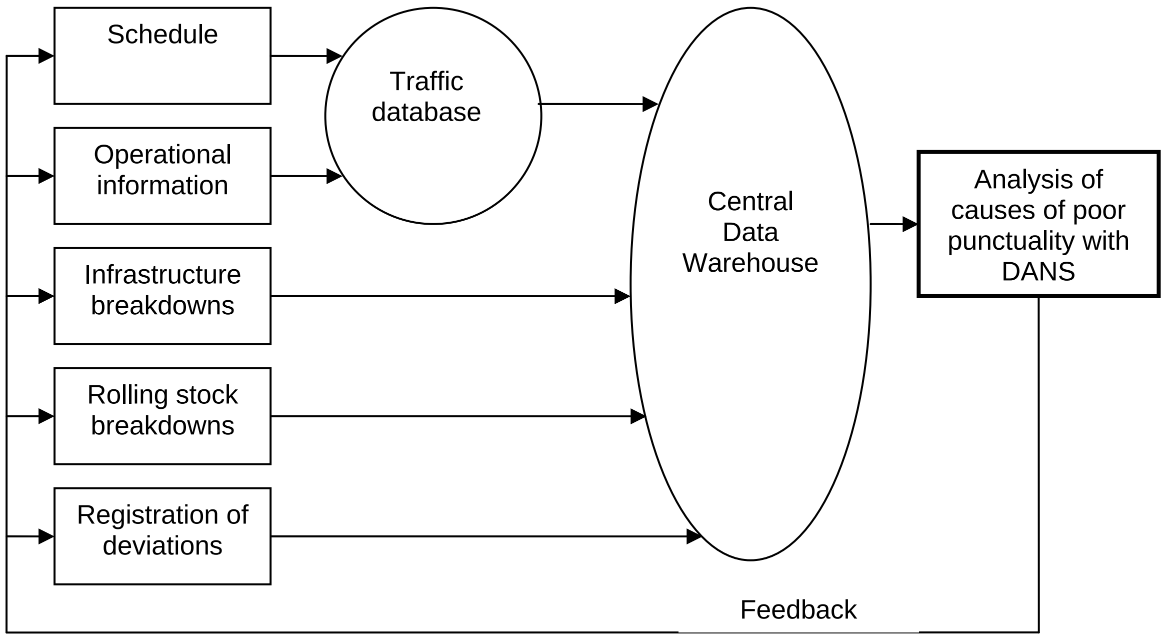


Figure 1: Input data for punctuality analysis and feedback

A completely automatic record is kept for every train of the time of passing certain measurement points. All the current operational information is stored together with the schedule data in ProRail's traffic database (Figure 1). The punctuality of the transport operators is deduced from this operational and schedule data.

The reason for studying punctuality issues

Many different things may cause a train to arrive or depart late at or from a station. They include failures in the rail infrastructure (points, tracks, signals, bridges and crossings), technical breakdowns in the dispatching or safety systems, collisions with road traffic, bad weather (snow, lightning, gales), bomb scares, defective equipment (doors, brakes and traction), delays of other trains, problems with passengers, a lack of personnel, problems with the train's electricity supply (overhead wiring), schedule conflicts, and so on.

The data of all these breakdowns and events are recorded. However, this process now occurs in separate systems (both locally and centrally, and in paper and digital form), and involves a variety of staff working for different organizations. Consequently, analysing why punctuality problems occur is complex and time-consuming. The analysts are confronted with myriad systems, multiple databases, a complex transport process and a profusion of possible causes.

The central data warehouse is now presenting new analysis opportunities. The data from the transport database plus data on breakdowns and process deviations is stored in the central data warehouse (Figure 1).

The reason for the research described in this article is to shed light on what might detract from punctual operation. The output will enable analysts to determine the underlying causes more accurately and rapidly and identify ways of improving train transport performance. The transport operator's customers will then be more likely to arrive on time at their destinations.

The method for analysing punctuality

NS is aiming to be able to complete analysis of timetable execution and to feed back the findings to production units within twenty-four hours. The organization has accordingly set up a data warehouse for the information involved (Figure 1). This means that all schedule and operational data are available centrally in digital form. Infrastructure and rolling stock breakdowns are also recorded and stored, as are all process deviations involving trains delayed by three minutes or more.

The complexity of the transport process means that poor punctuality may have extremely diverse causes. This demands an ability to analyse on various levels of detail and on multiple performance indicators. The NS timetable analysis tool known as DANS[1] was specified and developed for analysing the reasons for unpunctual trains. DANS allows the analyst to examine the operational data in three domains: trains, time and place. Selecting the 'trains' domain involves defining the train numbers or train series to be examined. The process of setting the 'time' domain determines the start and end dates and times. Selecting the 'place' domain involves defining junctions, stations, or stops to be examined.

An advantage of this analysis research method is that it rapidly provides an overall insight and then allows zooming in on specific details in one of the three domains. The analysis results can be presented in various graphical forms, which resemble the schedule graphs, except that they also include the operational data.

An example of another DANS function is a statistical analysis of the operational data, for determining the distribution of punctuality on various aggregation levels. Some examples of levels are train numbers and train series, and the arrival and departure of trains at stations, possibly in combination with the geographical location of the station. This directly points to the trains and locations implicated in persistent problems. Information can be retrieved immediately from the identified problem train, because the data are available centrally, from registered deviations and reports filed by other employees. DANS can be used for analyses both in support of incident inquiries and in determining the persistent causes of poor punctuality.

Performing an analysis with DANS

DANS supports various approaches to analysing the data in the central data warehouse. The date and time of a situation can be set, if they are known. All train movements within the set period are then gathered automatically. The analyst can continue to zoom in, and filter out train series or train numbers (e.g. the direction of travel). It is also possible to filter out specific days. The data to be analysed are presented in a table or graphically in a time-distance diagram (Figure 2). Performing an analysis of this kind, including the graphical reports, takes just a few seconds on a standard PC. The speed means that the analyst can perform multiple analyses interactively, and investigate 'what-if' scenarios. A tree view of trains and dates is shown on the left hand graphical screen, corresponding with the display on the right hand screen. The right hand screen shows the time-distance diagram, with the distance on the horizontal axis (showing the stations and stops involved) and on the vertical axis the time at which the train was at a station or stop. The red line with the diamonds is the schedule, and the blue line with the circles is the actual situation. Clicking one of the lines with the mouse selects the associated train number in the left hand screen, and vice versa.

[1] DANS stands for Dienstregeling Analyse Nederlandse Spoorwegen (NS timetable analysis)

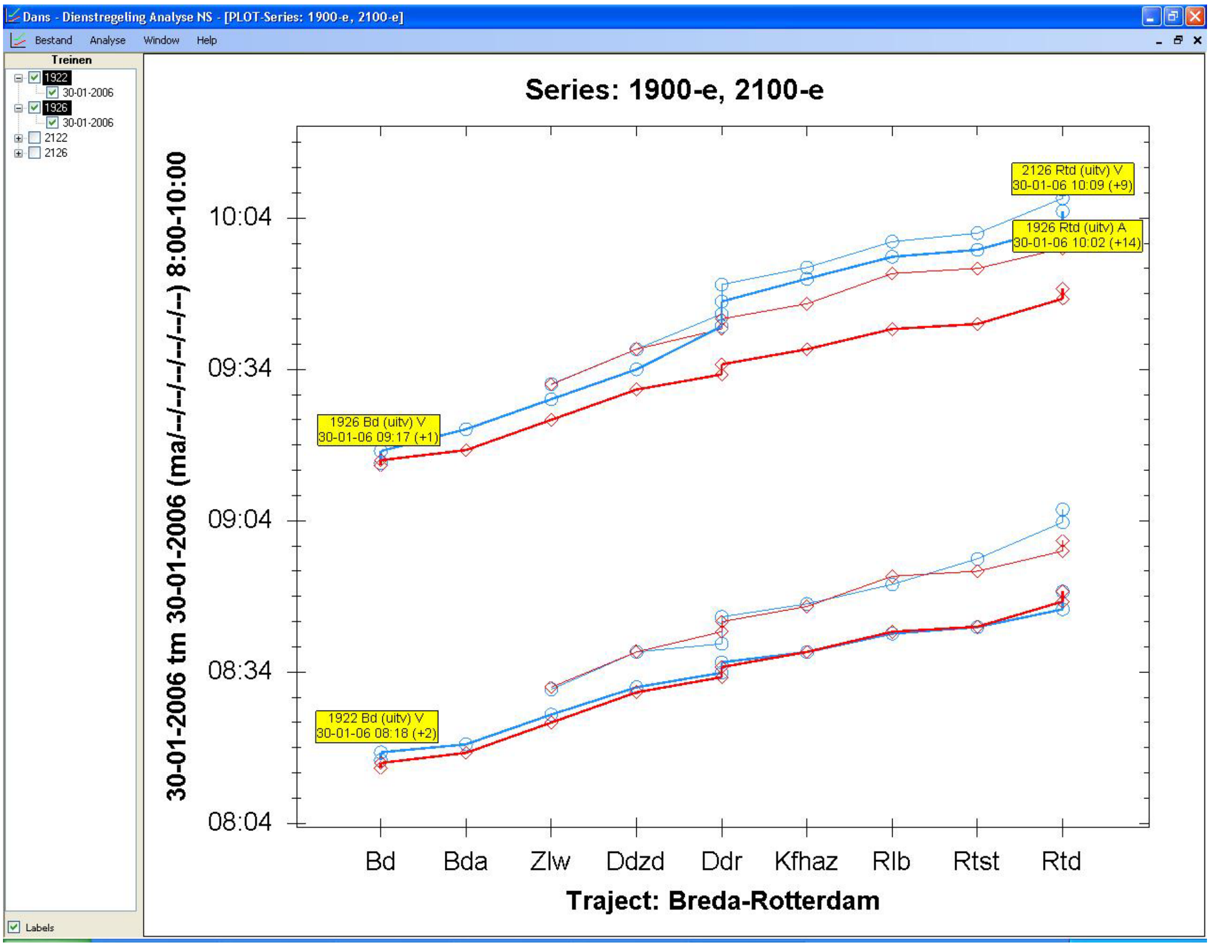


Figure 2: Graphical representation of analysis data in a time-distance diagram

As well as the filters mentioned above, trains can also be filtered on minimum and maximum amount of time early or late, which is particularly useful for analyses of persistent problems. For instance, it is possible to filter out trains that are more than 10 minutes late, which are usually because of incidents, and therefore of little interest in a persistent analysis. A clear graphical representation requires trains of this kind (i.e. those that are either on time or affected by incidents) to be excluded. The train lines can be labelled with the data of the train at a station or stop point for the purpose of analysis reporting.

The NS timetable is based on a recurring hourly pattern. The trains in a given train series run in accordance with a recurring pattern from station A to station B. For example, the trains in the 9800 series leave The Hague for Utrecht at 25 minutes and 55 minutes past every hour. DANS allows the actual timetable stability of a train series to be investigated by projecting the operational information for a lengthy period, such as a couple of weeks, onto one hour period.

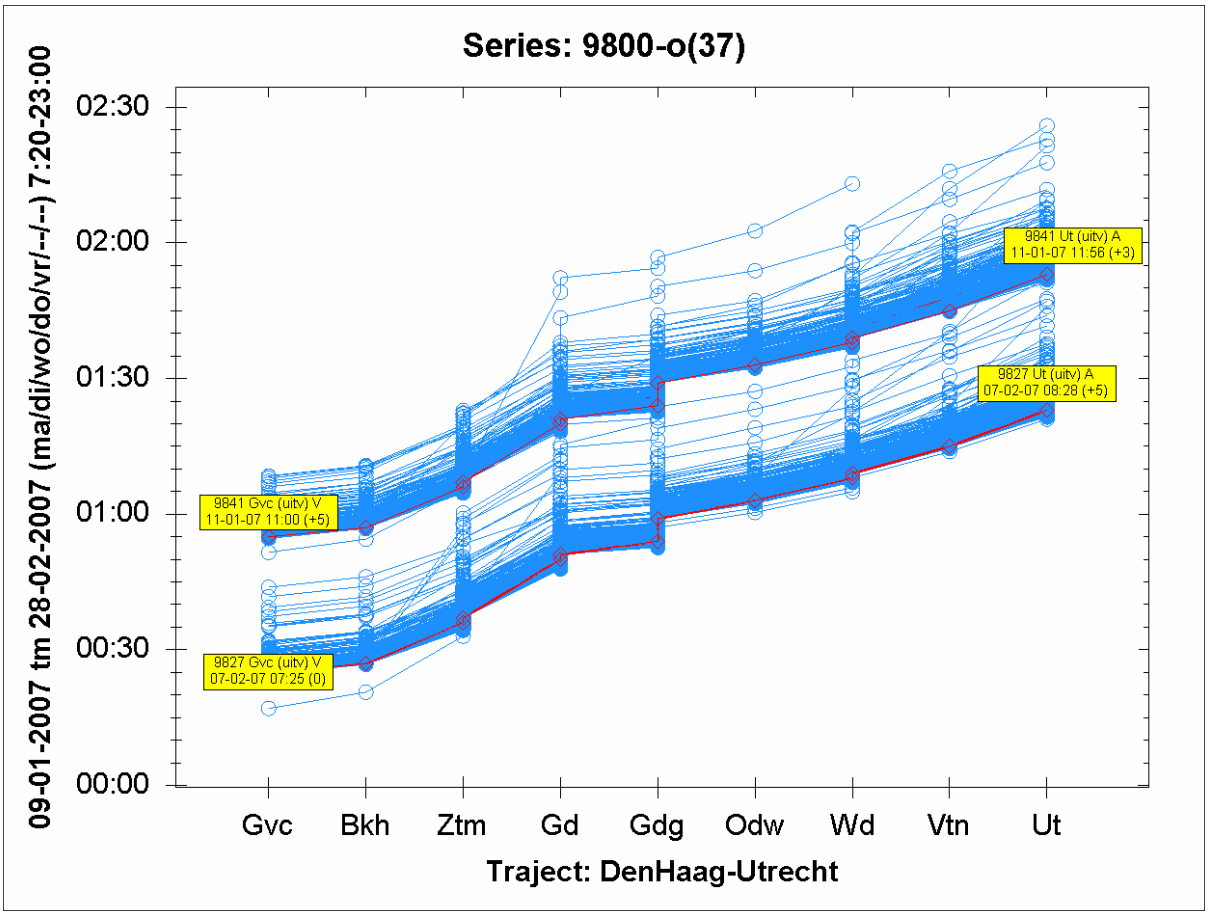
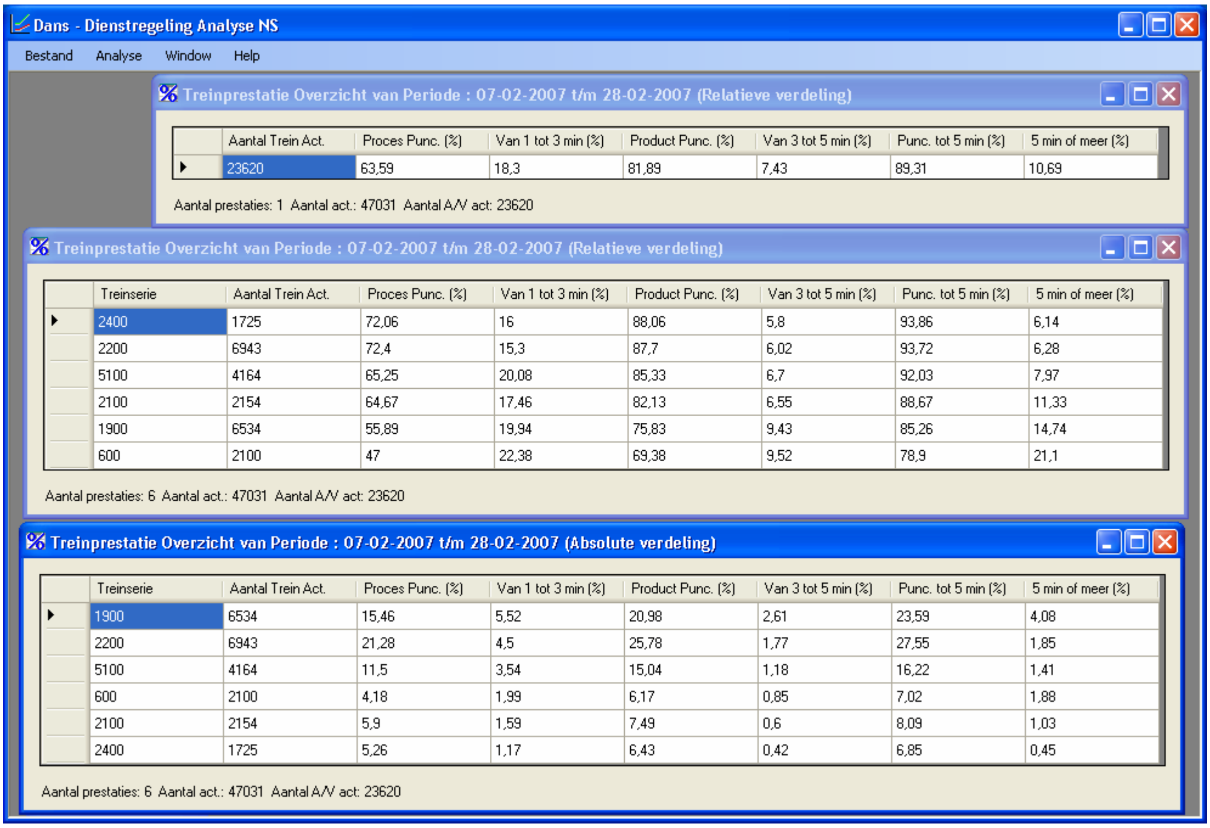


Figure 3: The operation of train series 9800 is on schedule

The near correspondence of the schedule and operational lines indicates that the operation is stable and that the train series uses no more infrastructure capacity than planned. If the train lines fan out more widely in the chart, operation is deviating from the schedule and the series is using more infrastructure capacity than planned. Furthermore, a train series of this kind would cause disruption to other train series scheduled shortly afterwards. Figure 3 shows the 9800 series (from The Hague (Gvc) to Utrecht (Ut)), which is running to schedule. Most of the train data shows an arrival delay of less than five minutes. The chart is for the period from 9 January to 28 February 2007, inclusive, from 7.20am to 11pm.

DANS also enables the analyst to investigate actual punctuality. NS uses two quality parameters for this purpose, process quality and product quality. Process quality is the proportion of trains arriving or departing less than one minute late. Product quality is the proportion of trains arriving or departing less than three minutes late.

Figure 4 shows three tables of process quality and product quality. The top table shows the overall view of the selected train series in the set period of 7 February to 28 February 2007, inclusive. The actual process quality was 63.59% and the actual product quality 81.89%. The second table shows the relative contribution of each train series to the quality parameters. This table also has a column with the proportion of the train series that is between three and five minutes late, and a column with the proportion delayed by 5 minutes or more. The trains that are just too late, which is to say by between three and five minutes, are those that could very likely be made to run on time again by taking appropriate measures. DANS has the facility to zoom in further in order to determine which train numbers are involved, the associated stations and stops, and whether arrival or departure was affected. Furthermore, as well as relative punctuality, DANS can also determine the absolute contribution to the punctuality problem. The absolute value is a weighted average, and it allows the analyst to identify the train series that is making the greatest contribution to the problem. The result is shown in the bottom table, where it can be seen that, in theory at least, attention to the 1900 series would offer the most gain. If the trains delayed by between three and five minutes were to be brought back inside three minutes, punctuality would improve by 2.61%.


Dans - Dienstregeling Analyse NS

Bestand Analyse Window Help

Treinprestatie Overzicht van Periode : 07-02-2007 t/m 28-02-2007 (Relatieve verdeling)

Aantal Trein Act.	Proces Punc. (%)	Van 1 tot 3 min (%)	Product Punc. (%)	Van 3 tot 5 min (%)	Punc. tot 5 min (%)	5 min of meer (%)
23620	63,59	18,3	81,89	7,43	89,31	10,69

Aantal prestaties: 1 Aantal act.: 47031 Aantal A/V act: 23620

Treinprestatie Overzicht van Periode : 07-02-2007 t/m 28-02-2007 (Relatieve verdeling)

Treinserie	Aantal Trein Act.	Proces Punc. (%)	Van 1 tot 3 min (%)	Product Punc. (%)	Van 3 tot 5 min (%)	Punc. tot 5 min (%)	5 min of meer (%)
2400	1725	72,06	16	88,06	5,8	93,86	6,14
2200	6943	72,4	15,3	87,7	6,02	93,72	6,28
5100	4164	65,25	20,08	85,33	6,7	92,03	7,97
2100	2154	64,67	17,46	82,13	6,55	88,67	11,33
1900	6534	55,89	19,94	75,83	9,43	85,26	14,74
600	2100	47	22,38	69,38	9,52	78,9	21,1

Aantal prestaties: 6 Aantal act.: 47031 Aantal A/V act: 23620

Treinprestatie Overzicht van Periode : 07-02-2007 t/m 28-02-2007 (Absolute verdeling)

Treinserie	Aantal Trein Act.	Proces Punc. (%)	Van 1 tot 3 min (%)	Product Punc. (%)	Van 3 tot 5 min (%)	Punc. tot 5 min (%)	5 min of meer (%)
1900	6534	15,46	5,52	20,98	2,61	23,59	4,08
2200	6943	21,28	4,5	25,78	1,77	27,55	1,85
5100	4164	11,5	3,54	15,04	1,18	16,22	1,41
600	2100	4,18	1,99	6,17	0,85	7,02	1,88
2100	2154	5,9	1,59	7,49	0,6	8,09	1,03
2400	1725	5,26	1,17	6,43	0,42	6,85	0,45

Aantal prestaties: 6 Aantal act.: 47031 Aantal A/V act: 23620

Figure 4: Analysis of relative and absolute punctuality

Successful implementation of innovation project

The NS Rotterdam regional quality office and the engineering consultancy Movares embarked on an innovation research and development project in 2005. A small and tightly knit team of four worked in a setting known as an innovation studio, which provides for close collaboration between end users and developers. The innovation team met on a weekly basis to discuss user requirements, set priorities and identify possible solutions, with the associated advantages and disadvantages. A prototype of the selected solution was implemented by the developer and made available to the end users. The process was repeated until the prototype was suitable for presentation and demonstration to colleagues. Results from the prototype justified proceeding with the DANS analysis tool development. The project was carried out by the software house InTraffic, using the DSDM[2] development methodology. Not all aspects of DSDM were applied. The elements that were important for this project were the MoSCoW[3] list, the time-boxing principle and variable scope of functionality to be implemented. MoSCoW is how DSDM categorizes and prioritizes the requirements for the project output, and it enabled InTraffic to develop the first version within three months and deliver it to the transport operators' analysts. The DANS system was made available nationally to NS analysts on 12 November 2007. This national implementation confirms the success of the innovation project.

Results of the investigation for the transport operators

The DANS system was made available nationally in the autumn of 2007. The analysts in the regions are now able to perform analyses with DANS. DANS then presents the analysed data graphically. The

[2] DSDM, Dynamic Systems Development Method, see www.dsdm.org
[3] MoSCoW, Must have, Should have, Could have, Would like to have but won't have this time around

analysis of a line can now rapidly clarify which stations or stops have persistent delays, as well as the distribution in operations and the associated use of the infrastructure. Any interaction between the trains of one or more transport operators is also clearly visible. The insight into persistent delays that DANS gives the transport operators' analysts is based on facts, and lays the foundation for targeted improvement of transport by train. The result will be a higher quality of the transport operator's services to its customers.
The use of DANS for the transport operator will be illustrated through two cases.

Case 1: Early departure

Figure 5 shows the stop train series 5000 from Rotterdam (Rtd) to Dordrecht (Ddr). All trains in the last weeks of timetable 2007 are projected on one hour period. On the stations and stops between Rotterdam and Dordrecht only the departure time is visible. The lines with a circle represent the execution of the timetable and the lines with a diamond represent the schedule.

In Barendrecht (Brd) some of the lines with a circle are below the schedule line. This means that some trains are departing before the scheduled departure time. As a result it is possible that passengers missed the train. Further research showed that early departure in Barendrecht happened more often, because of slack in the running times towards Barendrecht. As a result of this research the timetable for train series 5000 was changed in the new timetable for 2008. Departure times in Barendrecht and Zwijndrecht (Zwd) are scheduled one minute earlier. The schedule now better fits the execution. DANS provided easy insight in this case. Furthermore, the graphical representation helped to make people aware of the undesirable situation of early departure.

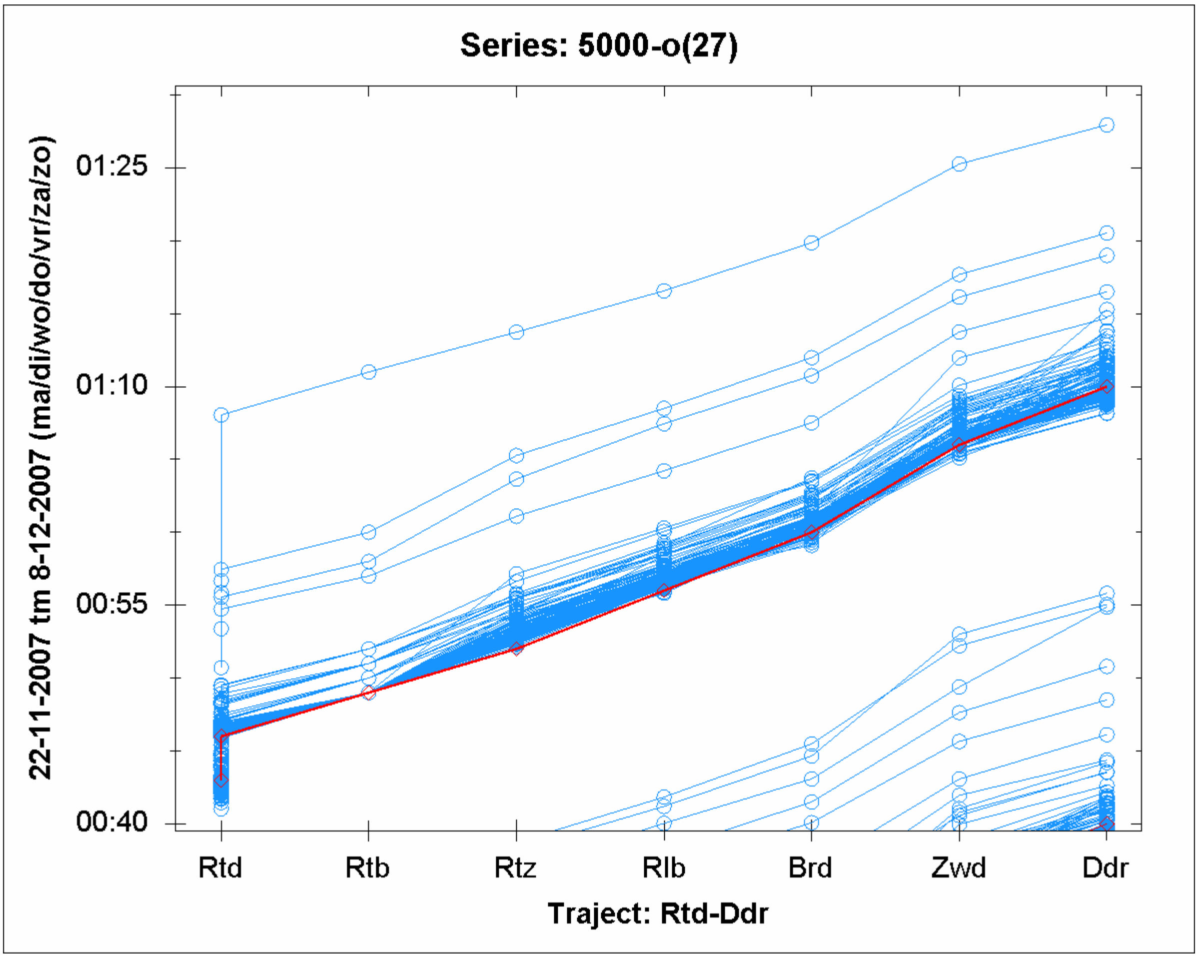


Figure 5: Stop train series 5000 projected on one hour period

Case 2: An incident, secondary delay

Figure 6 shows the track between Breda (Bd) and Tilburg (Tb), and two train series running on this track: de intercity service 1900 (bold lines) and the regional service 13600 (thin lines).
In Breda train 1961 arrives (and departs) with a delay of 4 minutes. This causes a delay for train 13662, which is scheduled to depart in the same direction 3 minutes after train 1961. Further research with DANS by projecting all trains in a few weeks on one hour easily showed this was an incident; there were no problems with the schedule.

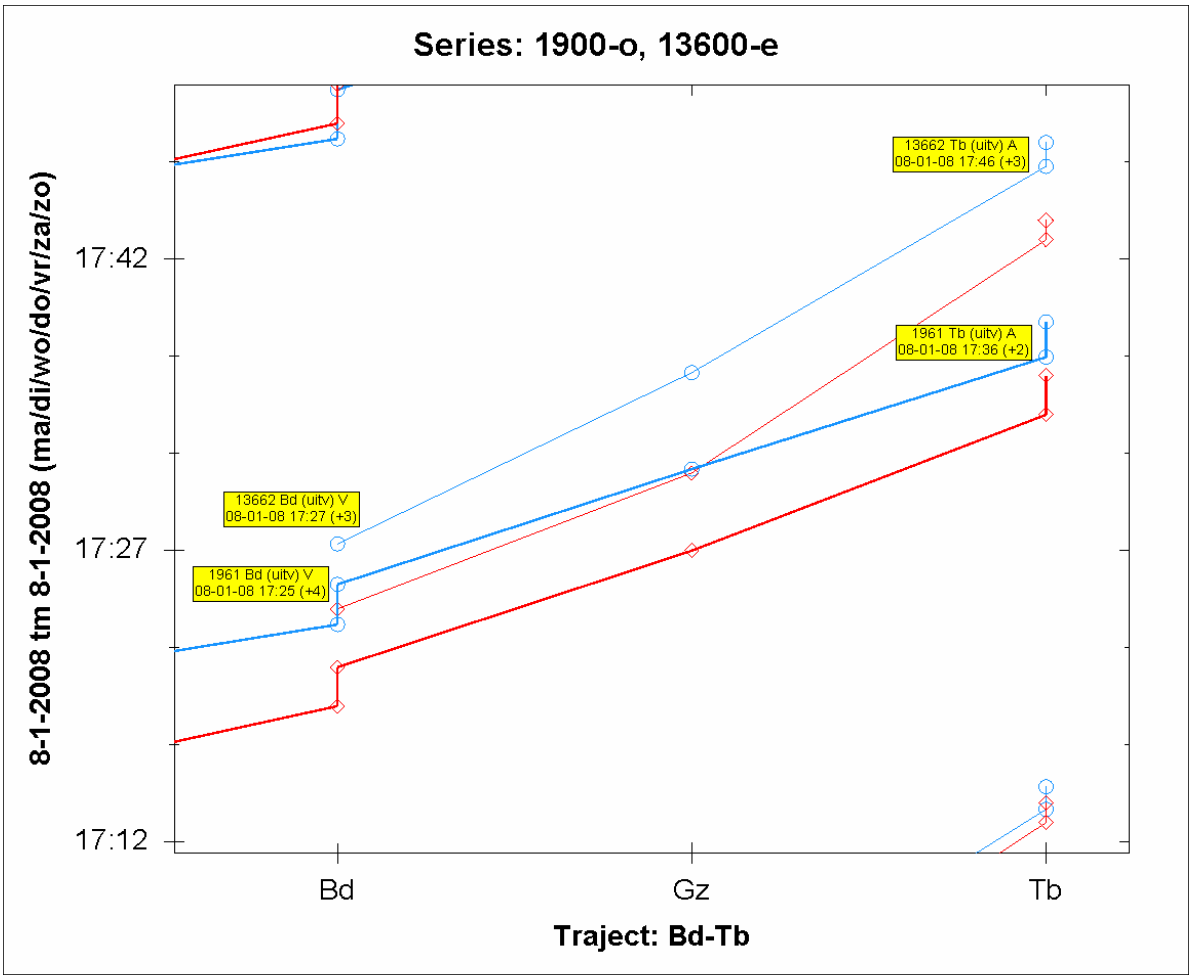


Figure 6: Analysing intercity service 1900 (bold lines) and the regional service 13600

Future developments

The current version of DANS has been used with great success by NS analysts. However, there is always room for improvement. For example, a later version will be easier to use because of an additional facility for making selections on a track map. Users will then have an easy way of selecting a line by mouse on a national map, thereby avoiding having to enter lists of stations and stops.

The application will also be adapted for use on lines with multiple tracks in each direction of travel. The current version has a built-in assumption that each line has only one track per direction of travel, which can have the undesirable effect of superimposing the operational lines of several trains in the graphical representation of multiple-track lines. In the real world the trains run alongside each other on separate tracks.

NS is also to use the VTL tool [2] for investigating underlying obstacles in the timetable. This tool supports the calculation of accurate arrival and departure times, the identification of route conflicts and the analysis of process times, such as running and dwelling times. It will be investigated how to link the VTL tool with DANS.

Finally, there will be effort on linking DANS and other data in the central data warehouse. It must be possible in future to link directly through from a delayed train in DANS to all available information on that train. Possible examples are records of causes of delay, or rolling stock and personnel data.

References

[1] Makkinga F., “A new user interface for the train traffic control system” – Proceedings of Comprail 2004, pp 725 – 732, Dresden, Germany.

[2] B.W.V. Stam - Van den Berg, V.A. Weeda, “VTL-Tool: Detailed Analysis of Dutch Railway Traffic”, ISROR (2007).