CETRA[2010]

17–18 May 2010, Opatija, Croatia
First International Conference on Road and Rail Infrastructure

A study of the performance and utilization of the Swedish railway network

A. Lindfeldt
Division of Traffic and Logistics, Department of Transport and Economics, Royal Institute of Technology, Stockholm, Sweden.

Abstract

Many lines in the Swedish railway network are heavily utilized and the demand for freight transport and passenger traffic is increasing. The Swedish National Rail Administration has been forced to declare some sections to be overloaded and with the up-coming deregulation of railway traffic, the load is going to increase even more.

This paper describes how the performance of the entire Swedish railway network has been investigated and the results. Data from several different databases (supplied by the Swedish Railway Administration) has been processed. The data consists of the scheduled timetable, data about the design of the infrastructure and operational data such as recorded delays, train weights and train lengths.

To perform the calculations, the railway network has been divided into 123 sections. The division has been made according to traffic patterns and type of infrastructure (single-track, double-track). For each line, the data has been used to calculate several descriptive parameters, e.g. delay development, number of trains/h during different parts of the day, time for the peak hour, inter-station distance, track length, train length, train weight, total mass/day (freight trains), train average speed and the heterogeneity of the train speeds.

Together, all of the parameters form a picture of the traffic flows in the Swedish railway grid and identifies the characteristics of the different lines such as: load, traffic mix, delay situation and infrastructure performance. The information can be used to detect deficiencies and propose possible solutions.

Finally, the correlation between the calculated parameters and the delay situation has been investigated.

Keywords: performance, capacity, load, traffic mix, delay.

1 Introduction

As the load on the railway network increases, so does the need to know how the network performs. The Swedish Railway Administration continuously investigates the performance of the Swedish railway with regard to capacity utilization and limitations. The results are presented in an annual report [1]. The work presented in this paper has been performed as a part of a research project financed by the Swedish Railway Administration and KTH with the aim to increase the understanding of the behaviour of congested infrastructure.

In this work a complete mapping of the Swedish railway network has been performed with respect to several different performance measures describing the infrastructure, timetable, train properties and delay situation. The railway network has been divided into smaller parts and the supplied data used to calculate several parameters describing the situation on each part. The intention is that the parameters will depict the status of the railway network, point to weaknesses and give suggestions of what can be done to solve them [2].

The information presented here is only a small selection of the work performed. The entire work is presented in [2].

2 Method

2.1 Data acquisition

The data used in the study has been taken from four different databases, all maintained by the Swedish National Rail Administration. The requested data has been supplied as text, excel, or PDF files and much effort had to be spent on importing and saving the large amounts of data to a format useable by Matlab. Below, the different databases are listed.

- BIS: Contains data describing the infrastructure. The following information about each station was exported: station name, name of the adjacent stations, inter-station distance, station coordinates, nr of tracks on the station, station track lengths, track types (main track or sidetrack) and parallel movement facility. The data covers the whole Swedish railway network and was exported on 2008-12-19.
- Tidtabellsboken: Is the timetable for the whole Swedish railway network. The data is valid for one year (2008).
- BANSTAT: Contains reported data about the trains running on the network. The exported data included the weight, length and the number of axles of all the trains that operated in Sweden in October 2008.
- TFÖR: Records the delays of all trains as they pass the stations. The exports were made only for the start and end stations of each evaluation route, see Section 2.6. The data is valid for September and October 2008.

2.2 Division of the railway network

The Swedish railway network was divided into 123 parts, evaluation routes. The division has been made with traffic patterns and type of infrastructure (single-track, double-track) in mind to get as homogenous a traffic situation as possible along the entire evaluation route, or else the calculated parameters would be less applicable. In some cases an even finer division was needed, see Section 2.4.

2.3 Infrastructure

Each evaluation route has been defined as being a single-, double- or multiple-track route. The definition is used when deciding if a station is valid or not. On single-track routes crossing stations are valid, and on double-track routes overtaking-stations are valid. To find the valid stations, conditions were applied to the station properties, in this case the number of tracks and track types.

For each valid station, the inter-station distance and maximum track length were calculated. It was also determined if the station had parallel movement facility. For the entire evaluation route, the minimum, maximum, mean and standard deviation for the inter-station distances and the maximum track lengths were calculated. In addition, the share of stations with three tracks or more and the share of stations with parallel movement facility were calculated.

2.4 Timetable

The timetable data was primarily used to count the number of trains running along the routes. To be able to search the timetable database, a particular date had to be specified. It was decided that the 9th of October was a representative date (a Thursday and no holidays close by). Furthermore, many of the evaluation routes had to be divided into several smaller sections in order to also catch trains that only travelled on a small part of the route.

For every hour and for each direction separately, the number of trains passing the border stations of the finer sections was counted. The corresponding values for the evaluation routes were then calculated as the mean of its finer sections. When the number of passing trains for each hour had been established, the following parameters were calculated: the maximum number of trains passing during one hour, time for the maximum hour, number of trains per hour during the morning rush hours (06-09), number of trains per hour during the afternoon rush hours (15-18), number of trains during the presumed afternoon rush hour (16-17), number of trains per hour during the day (09-15, 18-20), number of trains per hour during the night (20-06) and finally, the total number of trains/day.

2.5 Train properties

The average speeds of the trains were calculated using departure and arrival times at the border stations of the evaluation routes together with the calculated route lengths. The speed could only be established for the trains running the entire evaluation route and includes the time for stops at stations. The following parameters for the speed of the trains were calculated: maximum, minimum, mean, median, standard deviation, standard deviation divided by the mean and the 95 percentile divided by the 10 percentile. The last three parameters describe the heterogeneity of the train speeds. Other measures of heterogeneity are discussed in [3]. The calculations were made for three groups of train-types: passenger trains only, freight trains only, and for all trains except service trains and shunting movements.

For freight trains, the minimum, mean, maximum and standard deviation was calculated for the train weight, length, nr of axles and weight/axle. Also, the total gross weight/day of all the freight trains was calculated as the mean of all weekdays. For the passenger trains, the proportion of trains with more than 12 axles, corresponding to 3 normal cars, was calculated.

2.6 Delays

Since most trains run along more than one evaluation route, the absolute delays of the trains are not of interest when evaluating the performance of the routes. Instead the change in delay (increase or decrease) was calculated as the trains entered and left the routes. The delays have been used to calculate three different parameters: the share of the total number of trains that have received an increased delay, the median delay per kilometre for the trains with an increased delay and the standard deviation of the delay per kilometre for the trains with an increased delay. The delays in the second and third parameter were normalized by the length of the evaluation route because it would be reasonable to expect the increase in delay to be proportionate to the route length [4], and that it was important that the results were comparable between the evaluation routes. The calculations have been made for freight trains and passenger trains separately.

3 Results

3.1 Parameters

The results are presented as maps, where the values of the calculated parameters are represented by either the colour or the thickness of the evaluation routes. Only a small fraction of the maps produced are discussed in this paper. Figure 1 shows the Swedish railway network and the number of tracks of the evaluation routes: single, double or quadruple (red). Figure 2 shows the total number of trains/day in black with the total number of freight trains superimposed in green. The total number of trains ranges from 0 to 267 per day and direction. Figure 1 together with Figure 2 gives a rough idea of the capacity utilization of the network. It can be concluded that the freight traffic is dominating in the north and that the passenger traffic is dominating around the greater cities of Stockholm, Göteborg and Malmö. On the double tracks running between the greater cities, the traffic is a mix of freight and passenger trains.

In Figure 3 the mean track length of the stations have been compared to the length of the freight trains. Green means no trains on the evaluation route are longer than the mean track length, and red means that a large percentage of the trains are longer than the mean track length. The range is from 0 % to a maximum of 61 %. It is evident that some routes have insufficient track lengths.

On single track routes, the consequence is that meetings with long trains cannot be carried out at all stations, and hence the effective inter-station distance becomes longer, with an increased sensitivity to delays as a result. This is a well known fact on the route between Gällivare and Luleå (part of the red coloured route in the north) where the long ore-trains can only meet at five stations for a distance of 204 km. On double track lines, the problem consists of not being able to overtake slow freight trains that are too long.

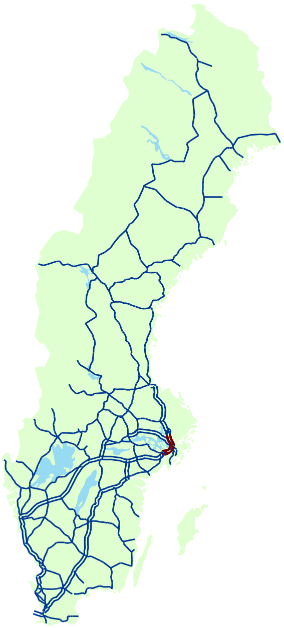

Figure 1.

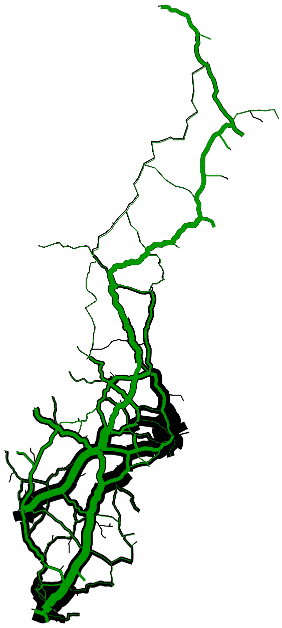

Figure 2.

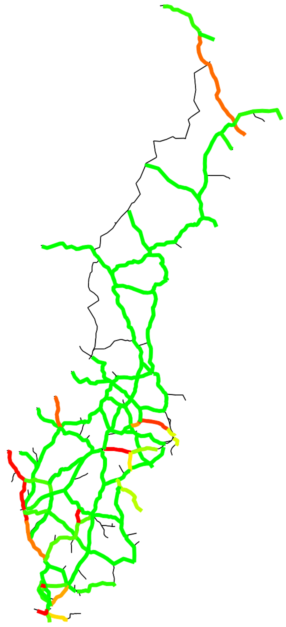

Figure 3.

Figure 4 and Figure 5 show the mean speed and speed mix of the trains. Green colour means low average speed and that the speed mix is homogenous. Red colour indicates that the speeds of the trains are high and heterogeneous. In Figure 4 the minimum speed is 23 km/h and the maximum speed is 142 km/h, in Figure 5 the speed mix ranges from 1 to 2.6 (95 percentile divided by the 10 percentile). The speed parameters are of interest since the capacity of a line is dependent on both the speeds of the trains and, especially, the heterogeneity of the train speeds [5]. The lines with the most mixed traffic are the double track lines Stockholm-Göteborg and Stockholm-Malmö. On these lines, high speed long distance passenger trains share the same track as slow freight trains. Due to the congested traffic situation on these lines, the freight trains are forced to do sidings often, which decrease their average speed even further [6].

The share of the passenger trains that have increased their delay on the routes are illustrated in Figure 6. The data ranges from 14 % (green) to 89 % (red). The highest values, above 80 %, are explained by temporary speed restrictions being active during the whole, or a large part, of the period of measurement. In these cases the timetable was not modified to compensate for the extended running time. The effect of temporary speed restrictions is further investigated in [7]. The share of delayed freight trains was in general within the same interval as for the passenger trains. However, the median delay for the freight trains with an increased delay, were in general several times higher than for the passenger trains. The explanation is the great variability of the freight trains' delay distribution.

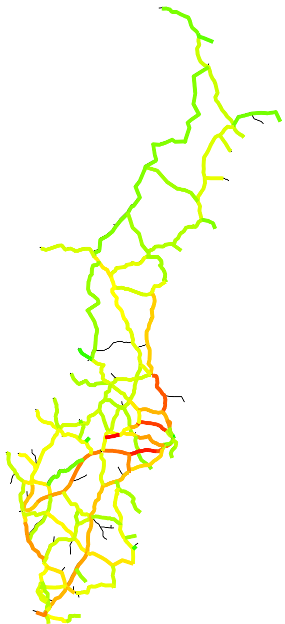

Figure 4.

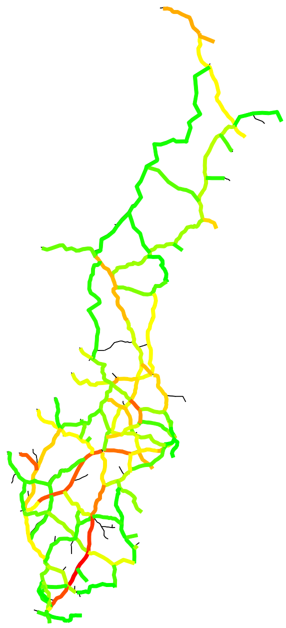

Figure 5.

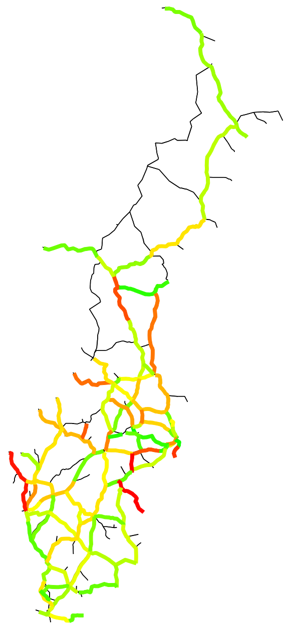

Figure 6.

3.2 Delay correlated parameters

Matlab was used to perform a stepwise multilinear regression to find out how the calculated parameters correlated to the delay parameters. The algorithm uses p-values for the F-statistic to decide which parameters to include in the model. The test was performed on single-track and double-track routes separately. To refine the test, routes with extreme conditions, such as routes with extremely low traffic load, temporary speed restrictions, partial double tracks (single-track routes), multiple-tracks or very short routes, were discarded. All in all, 32 single-track routes and 11 double-tracks remained for the test.

For the double-track routes, the test showed no significant connection between the delays and the other parameters. The reason for this was probably that 11 routes were too few to get any significance in the test. For the single-track routes, the results are summarized in Table 1. The test was performed for passenger trains and freight trains separately. The delay parameters tested were the share of the total number of trains that have received an increased delay and the median delay per kilometre for the trains with an increased delay.

The share of passenger trains that has received an additional delay showed a positive correlation to the total number of trains per day, the mean speed of the trains and the heterogeneity of the train speeds, which all seem natural. For the freight trains, the significant parameters were the total number of trains per day and the share of stations with 3 tracks or more.

Table 1. Parameter analysis, significant at the 0.05 level.

Train type	Delay type	Parameter	Slope
Passenger	Share	Total nr of trains/day	0.44
Passenger	Share	Mean speed (all trains)	0.42
Passenger	Share	Speed mix (all trains)	0.29
Passenger	Median	Route length	-0.58
Passenger	Median	Share of passenger trains with > 12 axles	0.24
Passenger	Median	Std of the inter-station distance	-0.41
Freight	Share	Total nr of trains/day	0.57
Freight	Share	Share of stations with at least 3 tracks	-0.46
Freight	Median	Route length	-0.72
Freight	Median	Freight train mass	0.37
Freight	Median	Max inter-station distance	-0.37

For the median delay, the results were less coherent. The route length seemed to be connected to the median delay for both the passenger and the freight trains, despite the fact that the mean delay had been normalized by the route length to eliminate this factor. The explanation was the low resolution (1 minute) of the delay data. It follows that the minimum accumulated delay for a train is 1 minute (given that the train has received an additional delay) and that the median accumulated delay for all trains must be at least 1 minute. This is a problem for short routes where the median delay hits the 1 minute limit, with the result that the delay/km will be very high for the shortest evaluation routes, hence the negative correlation.

A result that is surprising is that the median delay for both passenger trains and freight trains show a negative correlation to the standard deviation and maximum inter-station distance respectively, when the opposite would seem more natural.

3.3 Method validity and applicability

It is general knowledge that it is only the secondary delays that increase with an increased capacity utilization, and not the primary delays. In our case the delay data is the sum of the two. Still, correlations were found between several parameters and the calculated delays, Table 1. Several of the parameters in Table 1 are defined by the UIC as basic parameters underpinning the capacity [8]. Hence the conclusion is that some of the measured delays can be explained by the capacity utilization. The length of the passenger trains has been proved to correlate to delays [4]. This correlation is in our case represented by the parameter share of passenger trains with more than 12 axles.

Compared to the UIC method to calculate capacity utilization, the method proposed in this paper is more easily applied to a complete network where you do not have detailed data about for example signal block lengths and block occupation times. Instead of estimating the capacity utilization, the method reacts to the consequences thereof. An advantage with a method based on operational data is that the influences of unknown parameters are included. However, it cannot produce an absolute number of the capacity utilization and it cannot be used to predict how a change may affect for example punctuality.

4 Conclusions

The conclusions of this work are that it is possible to use data about the infrastructure, timetable, train properties and delays in an automated process to calculate descriptive parameters for a railway network. The parameters can be used to identify weaknesses in the network and suggest appropriate measures to solve the problems. The developed programs can easily be reused with new data as input to provide information about how the situation develops over time.

The correlation between the calculated parameters and the delays of the trains has been investigated using stepwise linear regression. On single-track lines, positive correlations have been established between the number of trains/day, the mean speed of the trains and the heterogeneity of the train speeds to the share of delayed passenger trains at the 5 % level of significance. For the freight trains, the number of trains/day and the share of stations with at least 3 tracks were significant.

References

[1] M. Grimm, M. Wahlborg. Kapacitetsutnyttjande och Kapacitets-begränsningar 2007/2008. (Capacity utilization and limitations) Banverket, 2008. (In Swedish)

[2] B-L. Nelldal, A. Lindfeldt, O. Lindfeldt, Kapacitetsanalys av järnvägsnätet i Sverige, delrapport 1-3. (Capacity analysis of the Swedish rail network, part 1-3). KTH, 2009. (In Swedish)

[3] Michiel J.C.M Vromans, Rommert Dekker, Leo G. Kroon. Reliability and heterogeneity of railway services. European Journal of Operational Research 17, 647-665, 2006.

[4] N.G. Harris, Planning Passenger Railways: A Handbook. Transport Publishing Co, UK, 1992.

[5] L-G. Mattson. Railway Capacity and Train Delay Relationships. Critical Infrastructure: Reliability and Vulnerability, Springer-Verlag, 129-150, 2007.

[6] H. Sipilä, Tidtabellsläggning med hjälp av simulering, Effekter av olika tillägg och marginaler på X2000-tågen Stockholm-Göteborg (Timetable construction by simulation, Effects of allowances and margins on the X2000-trains Stockholm-Göteborg), KTH, 2009. (In Swedish)

[7] N.O.E. Olsson, Hans Haugland. Influencing factors on train punctuality-results from some Norwegian studies. Transport Policy 11: 387-397, 2004.

[8] UIC Code 406, Capacity, International Union of Railways, 2004.