Developing the UIC 406 Method for Capacity Analysis

Khadem Sameni, Melody; Landex, Alex; Preston, John

Published in:
Proceedings for 4th International Seminar on Railway Operations Research

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Abstract
This paper applies an improvement cycle for analysing and enhancing capacity utilisation of an existing timetable. Macro and micro capacity utilisation are defined based on the discrete nature of capacity utilisation and different capacity metrics are analysed. In the category of macro asset utilisation, two methods of CUI and the UIC 406 are compared with each other. A British and a Danish case study are explored for a periodic and a non-periodic timetable: 1- Freeing up capacity by omitting the train that has the highest capacity consumption (British case study). 2- Adding trains to use the spare capacity (Danish case study). Some suggestions are made to develop meso indices by using the UIC 406 method to decide between the alternatives for adding or removing trains.

Keywords
Railway capacity, UIC 406, CUI

1 Introduction

Privatization of European railways, concerns for the environment and sustainability, higher fuel costs and road congestion have resulted in enormous growth in railway passenger and freight in the past decade [3, 32]. This has not been matched by enough increase in capacity, making railway networks more and more saturated. Many European railways are struggling to accommodate necessary train services on their limited infrastructure or tackling ‘the railway capacity challenge’. In this regard, efficient management and planning for measuring capacity and enhancement measures are needed. To improve capacity utilisation and allocation, an improvement cycle is needed as shown in Figure 1.
Railway and road are two modes of transportation that face capacity constraints on their main infrastructure as well as their nodal bottlenecks; hence comprehensive overview of capacity is very much needed. Table 1 compares the status of capacity manual for these two modes of transportation. Contrary to road transportation, many aspects of railway capacity have not been systematically explored and as expressed by TRB Rail Capacity Joint Subcommittee, the need for railway capacity manual is felt [19].

This paper analyses the improvement cycle for railway capacity utilisation and makes some suggestions for possible improvements of current practices. Different definitions of railway capacity and lean capacity utilisation are explored. Strengths and weaknesses of
current capacity metrics are discussed. Two major analytical methods of UIC 406 [31] that is used in continental Europe and Capacity Utilisation Index (CUI) [10] which is used in Great Britain are compared. In the later sections methods to enhance capacity utilisation are studied in two case studies.

2 **Defining Railway Capacity**

The first step toward improving the utilisation (and exploitation) of the railway infrastructure is to define railway capacity. When railway capacity is defined, it is possible to develop methods to measure it. However, unlike other modes of transportation, there is no unique definition for railway capacity. In this section current major definitions of capacity are provided, the discrete nature of capacity utilisation is discussed and a new approach is suggested toward defining the railway capacity.

### 2.1 Definitions of Railway Capacity

Railway capacity is the outcome of close interaction between different subsystems of the railway: rolling stock, infrastructure and timetable that link these together as shown in Figure 2.

This complex interaction has been expressed through different definitions:

- “Capacity is the level of traffic (i.e. number of trains per day) that a rail line can accept without exceeding a specified limit of queuing time”. [24]
- “The ability of the carrier to supply as required the necessary services within acceptable service levels and costs to meet the present and projected demand.” [15]
“Capacity is the highest volume (trains per day) that can be moved over a subdivision under a specified schedule and operating plan while not exceeding a defined threshold.” [17]
“Line capacity is the maximum number of trains that can be operated over a section of track in a given period of time, typically 1 hour.” [30]
“Capacity is measured as the count of valid train paths over a fixed time horizon within an optimal master schedule”. [12]
“The maximum number of trains that may be operated using a defined part of the infrastructure at the same time as a theoretical limiting value is not reached in practice.” [11]
UIC [31] has concluded that: “A unique, true definition of capacity is impossible.”

UIC [31] states that railway infrastructure capacity is a trade-off between the number of trains, heterogeneity, average speed and quality of service (stability). This is due to discrete nature of capacity utilisation which is discussed in the following section.

2.2 Discrete Nature of Capacity Utilisation

Passengers and freights cannot use the railway infrastructure directly; they are packed into batches of trains. Railway capacity is used in discrete steps (as opposed to road capacity that can be continuously used till it is saturated at a standstill level). These discrete steps can be taken in various ways and different combinations of train types, speed and levels of service. This explains why International Union of Railways has concluded that “Capacity as such does not exist” and “Railway infrastructure capacity depends on the way it is utilised”[31]. However, there can be indirect measures of defining capacity and how utilising it generates added value. Value is an expression of “the relationship between function and resources where function is measured by the performance requirements of the customer (such as quality of service) and resources are measured in materials, labour, price, time, etc. required to accomplish that function” [26].

Analysis of railway network can be done at different levels of macro, meso and micro as described in detail by Erol et al.[16] and Gille et al. [13]. It is also important to consider different levels of capacity utilisation. Hereby we define two categories for it:

- Macro capacity utilisation : Quantity of discrete steps to use railway capacity (e.g. the number of trains)
- Micro capacity utilisation: Quality of discrete steps to use railway capacity (e.g. Load factor)

To efficiently utilise the railway capacity, both aspects should be considered.

2.3 Lean Capacity Utilisation

Having a purely macro or micro approach toward capacity would not lead to efficient capacity utilisation: Too much effort for micro capacity utilisation would lead to overloading and neglecting it would cause capacity challenges at macro level where the network is saturated while some trains have very low load factor.
An example of this was described by Smith [28] for some train services in the south of England:

- Overloading at peak hours (125–150 percent)
- Overall load factor: 25 percent
- Many empty seats being hauled around off-peak
- Hauling empty seats long distances to satisfy short distance demand (e.g. South West Trains from Weymouth and Exeter to meet the Woking demand)
- Trains carrying few passengers around the fringes of country while there is overcrowding in the central parts of the network.

To avoid waste of capacity, it is important to consider how value is generated by using the capacity of infrastructure and define capacity accordingly. Figure 3, schematically shows a simplified version of capacity utilisation. Therefore, railway capacity can be defined as “the ability of infrastructure to generate value by moving passengers (or freight) toward their destination”. Formulae (1) and (2) encompass this concept.

\[
\text{Lean capacity utilisation} = f(\text{macro capacity utilisation} \times \text{micro capacity utilisation}) \quad (1)
\]
\[
\text{Lean capacity utilisation} = f(n \times d \times \tan(\alpha)) = f(n.S) \quad (2)
\]

The area marked as S in Figure 3, resembles blocking stairs and how macro capacity consumption is calculated by the UIC 406 and CUI methods. This definition of capacity has also some similarities with the concept of ‘traffic energy’ as defined by Hertel [13] where he defines traffic energy as traffic flow (number of trains per unit of time) multiplied by average speed [23]. This definition of capacity suggests to add an element of micro capacity utilisation (like load factor) to the above mentioned approaches.
3 Measuring Capacity Utilisation

Railway infrastructure is a scarce and expensive resource which should be allocated and utilised as efficiently as possible. Capacity utilisation should be quantified to analyse how well it is utilised.

3.1 Analysing Metrics of capacity utilisation

Different metrics can be used for measuring capacity utilisation. Dingler [7] has recently studied current metrics for measuring capacity of a freight railway mainly for North America. Here, Dingler’s [7] approach is continued to analyse metrics of capacity utilisation for passenger services in three main categories of throughput, quality of service and asset utilisation. Strengths and weaknesses of each type of metric are briefly summarized in Table 2.

Table 2 - Analysing metrics of capacity utilisation. Based on [7] and [16]

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
<th>Description</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Throughput</strong></td>
<td>Macro</td>
<td>Number of trains, train-km</td>
<td>How many passengers can be transported over a period of time</td>
<td>Easily measurable and understandable</td>
</tr>
<tr>
<td></td>
<td>Micro</td>
<td>Number of passengers, Passenger-km, seat-km</td>
<td>Measures reliability and timeliness</td>
<td>Does not reflect quality of service</td>
</tr>
<tr>
<td><strong>Level of service</strong></td>
<td>Average delay, percentage of cancelled or late trains (e.g. Public Performance Measure in Great Britain)</td>
<td>Measures reliability and timeliness</td>
<td>Important for general public</td>
<td>Indirect measure heavily depends on how saturated the network is. Does not take scheduled waiting time and timetable supplements into account which are a waste of time for passengers</td>
</tr>
<tr>
<td><strong>Macro Asset utilisation</strong></td>
<td>Capacity Utilisation Index (CUI),</td>
<td>Estimating how saturated the network is</td>
<td>Important to estimate how efficiently the capacity is used</td>
<td>A measure of macro capacity</td>
</tr>
</tbody>
</table>
3.2 Comparing UIC 406 and CUI Timetable Compression Methods

The quality of timetable determines how well capacity is utilised and how stable the operation is. Extensive research has been carried out on different aspects of scheduling and rescheduling trains, punctuality, reliability and stability of timetables which have been reviewed by Cordeau et al. [5] Tornquist [29], Hansen and Pachl [11] and Lusby et al. [20].

The major characteristics of timetables and how they affect capacity utilisation are illustrated in Table 3. Scheduled waiting time and punctuality are respectively defined as “An artificial increase in the overall timing of a train which is caused by the resolution of conflicts during the scheduling process” and “The percentage of trains which arrive at (depart from, pass) a location with a delay less than a certain time in minutes” [11].

Heterogeneity of traffic is caused by variations in speed and stop patterns as well as variations in headways. To measure heterogeneity, speed ratio has been defined by Krueger [17] and sum of shortest headway reciprocals was suggested by Vromans et al. [33]. Landex [18] proposes to use the ratio of the headway at departure station ($h_{i}^{D}$) to the following headway ($h_{i,i+1}$) multiplied by the ratio of headways for arrival at stations ($h_{i}^{A}$ and $h_{i+1}^{A}$). To provide a formula independent of number of trains, the result is divided by the number of headways minus 1 ($h_{N-1}$).
\[ Heterogeneity = 1 - \frac{\sum \left( \min \left( \frac{h_{i,j}^D}{h_{i,j+1}^D}, \frac{h_{i,j}^D}{h_{i,j+1}^D} \right) \times \min \left( \frac{h_{i,j}^A}{h_{i,j+1}^A}, \frac{h_{i,j}^A}{h_{i,j+1}^A} \right) \right)}{h_{N-1}} \] (3)

**Table 3- Major timetable characteristics and capacity utilisation**

<table>
<thead>
<tr>
<th>Scheduled waiting time</th>
<th>Average waiting time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative sensitivity of waiting time</td>
<td>Max</td>
</tr>
<tr>
<td>Traffic energy</td>
<td>Min</td>
</tr>
<tr>
<td>Traffic flow</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heterogeneity</th>
<th>Number of trains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on [18]</td>
<td></td>
</tr>
</tbody>
</table>

[23] [13]
The less time the infrastructure is idle can be an indirect measure of how well it is utilised. Therefore, timetable compression methods try to figure out the ratio of time that infrastructure is effectively used compared to the total available time.

The UIC 406 method has successfully been applied in several European railways. Höllmüller and Klahn [14], Wahlborg [34], and Landex [18] apply the UIC 406 method to Austrian, Swedish, and Danish railway networks respectively.

Confessore et al. [4] combine a discrete event simulation approach with the UIC 406 compression method to calculate the commercial capacity of a line in Italy (measured in number of trains). The general workflow is shown in Figure 4. In the simulation phase they cover factors that are not accommodated in the optimization phase mainly stochastic traffic perturbation.

Figure 4 - Simulation-based approach to capacity assessment [4].

The CUI method is the measure used in the Great Britain for capacity analysis and is based on the minimum headways derived from Network Rail’s ‘Rules of Plan’[2]. Train paths are squeezed together up to minimum headway times. There are two types of headway times used: fast and slow. The fast headway time is used when the preceding service does not stop at that station otherwise the slow headway time should be used for timetable compression.

CUI method is the measure used in the UK for capacity analysis and it based on the minimum headways derived from Network Rail’s ‘Rules of Plan’. Train paths are squeezed together up to minimum headways. There are two types of headways used: fast and slow. Fast headway is used when the preceding service does not stop at that station otherwise the slow headway should be used for timetable compression. [2]

Figure 5 - Timetable copression methods.
The results of timetable compression according to the UIC 406 and CUI methods are compared for a small case study in the South of Great Britain between Southampton Central station and Basingstoke and for the trains that move towards London Waterloo. The timetable was compressed for the morning peak hours from 7:00 am to 10:00 am. Results from timetable compressions for the case study are illustrated in Figure 6 and Figure 7.

Comparing the results it is seen that the average capacity utilization index by the CUI method is 1.6% higher than the UIC 406 method (Table 4). This can be explained by the fact that in the CUI method, minimum headway is considered at the node, while in the UIC 406 method the minimum headway is considered for a link. However, to generalize the results, more case studies are needed. Parts of the nodal capacity constraints are roughly considered using the CUI method, e.g. longer headway times are set when there is a change from quadruple tracks to double tracks at Shawford station. Therefore, the capacity utilisation from Shawford to Winchester link is considerably higher than the UIC 406 result.

<table>
<thead>
<tr>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>UIC 406</td>
<td>31.6%</td>
</tr>
<tr>
<td>CUI</td>
<td>33.2%</td>
</tr>
</tbody>
</table>
4 Analysing and improving Capacity Utilisation

By using the UIC 406 method and the related feature in version 6 of RailSys [25], the capacity utilisation for a British and a Danish case study are analyzed. Some suggestions are made to develop the UIC 406 method as a meso capacity analysis tool for determining which train to remove for freeing up capacity and which train to add for using spare capacity.

4.1 South West Main Line case study in Great Britain: freeing up capacity

The British case study as mentioned above is South West Main Line which is one of the congested routes in Great Britain’s railway network. The route is a major commuter route to London and is also critical for freight traffic from Port of Southampton. [22]

As Figure 8 shows, current passenger loading levels are near capacity and the projected demand shows that it would be over capacity in 2030 [6]. The percentage of passengers standing during the morning peak period (7:00 am to 10:00 am) in the trains operated by South West Trains is the second highest in Great Britain (17%) [21]. Table 5 and Figure 9 illustrate how capacity utilisation adds up as more trains merge towards London Waterloo.

Figure 8- South West Main Line region according to Route Utilisation Strategy [1] and loading factor between Southampton Central and Worting during morning peak hours [4].
There is a sudden jump in capacity utilisation from Shawford to Winchester due to the change from quadruple tracks to double tracks. During the morning peak period, there are no freight trains, and due to platform restrictions at London Waterloo station, no more passenger trains can be added during the morning peak hours. An interesting fact is that average capacity utilisation per train (capacity utilisation divided by number of trains) is around 1%, although the length of line section incrementally increases. It starts from 1.04% and reaches 1.96% for the line section from Southampton Central to Worting (10 stations).

Table 5- Capacity utilisation from Southampton Central station towards London Waterloo during morning peak hours

<table>
<thead>
<tr>
<th>Station</th>
<th>Capacity utilisation from previous station to this station (Individual)</th>
<th>Capacity utilisation from Southampton Central (Cumulative)</th>
<th>Number of trains (7:00 am-10:00 am)</th>
<th>Average capacity utilisation per train</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Denys</td>
<td>30.1%</td>
<td>30.1%</td>
<td>29</td>
<td>1.04%</td>
</tr>
<tr>
<td>Swaythling</td>
<td>21.2%</td>
<td>33.5%</td>
<td>31</td>
<td>1.08%</td>
</tr>
<tr>
<td>Southampton Airport Parkway</td>
<td>20.7%</td>
<td>36.3%</td>
<td>31</td>
<td>1.17%</td>
</tr>
<tr>
<td>Eastleigh</td>
<td>24.7%</td>
<td>43.2%</td>
<td>31</td>
<td>1.39%</td>
</tr>
<tr>
<td>Shawford</td>
<td>28.6%</td>
<td>55.6%</td>
<td>39</td>
<td>1.43%</td>
</tr>
<tr>
<td>Winchester</td>
<td>30.5%</td>
<td>70.5%</td>
<td>40</td>
<td>1.76%</td>
</tr>
<tr>
<td>Waller’s Ash</td>
<td>30.9%</td>
<td>77.8%</td>
<td>40</td>
<td>1.95%</td>
</tr>
<tr>
<td>Micheldever</td>
<td>27.8%</td>
<td>81.1%</td>
<td>41</td>
<td>1.98%</td>
</tr>
<tr>
<td>Worting</td>
<td>31.6%</td>
<td>84.3%</td>
<td>43</td>
<td>1.96%</td>
</tr>
</tbody>
</table>

Figure 9- Capacity utilisation for the morning peak hours (7:00 am – 10:00 am)
If it is intended to remove a train to free up some capacity, the decision can not be solely made upon macro capacity utilisation (i.e. how much a train blocks the infrastructure). Table 6 shows some examples for capacity utilisation after omitting trains. Micro capacity utilisation (i.e. load factor) should also be considered. Therefore the following methodology is suggested:

1. Calculating capacity utilisation for each scenario of omitting a train.
2. Developing a meso capacity utilisation criteria.
3. Calculating the meso capacity utilisation index for each scenario.
4. Choosing the best alternative.

Table 6- Capacity utilisation after omitting a train

<table>
<thead>
<tr>
<th>Train Type</th>
<th>Name</th>
<th>Origin</th>
<th>Destination</th>
<th>Capacity Utilisation after omitting the train</th>
<th>Meso capacity index after omitting the train</th>
</tr>
</thead>
<tbody>
<tr>
<td>170_1x2C1754</td>
<td>2R26</td>
<td>TOTTON</td>
<td>ROMSEY</td>
<td>83.3%</td>
<td>2</td>
</tr>
<tr>
<td>220_1x4C13</td>
<td>4M71</td>
<td>SOTON</td>
<td>CLPHMJJC</td>
<td>80.7%</td>
<td>1.1</td>
</tr>
<tr>
<td>220_1x4C2105</td>
<td>3B92</td>
<td>STDENYS</td>
<td>BSNGSTK</td>
<td>81.8%</td>
<td>1.6</td>
</tr>
<tr>
<td>220_1x4C2386</td>
<td>1T26</td>
<td>PHBR</td>
<td>WATR</td>
<td>81.2%</td>
<td>1.3</td>
</tr>
<tr>
<td>442_1x5C1595</td>
<td>1B26</td>
<td>POOLE</td>
<td>WATR</td>
<td>67.5%</td>
<td>0.3</td>
</tr>
<tr>
<td>442_2x5C2276</td>
<td>1B32</td>
<td>SOTON</td>
<td>WATR</td>
<td>79.2%</td>
<td>2</td>
</tr>
</tbody>
</table>

Different meso capacity utilisation indexes can be proposed. Hereby we suggest using the following for this case study:

\[
Meso \ capacity \ index = \frac{Micro \ capacity \ lost}{Macro \ capacity \ gained \ by \ omitting \ the \ train} \quad (3)
\]

\[
Meso \ capacity \ index \ to \ free \ up \ capacity = \frac{n_{cl}}{C_b - C_a} \quad (4)
\]

- \( C_b \): Capacity utilisation before omitting the train
- \( C_a \): Capacity utilisation after omitting the train
- \( n_{cl} \): Number of carriages lost

In this regard, the numerator considers how much micro capacity would be lost and the denominator considers how much macro capacity would be freed up by omitting the train. The number of carriages for different trains varies between 2 to 10. As the time period considered was for morning peak hours, the load factor of all trains was considered high and only the number of carriages is used. However, more complicated indexes can be developed according to the distribution of loading factor during the peak hours.

Calculating this meso capacity index for all trains, it can be advised which train is the
‘weakest link of the chain’, the one with the lowest meso capacity index. For this case study, the ‘weakest link of the chain’ was the stopping train from Poole to London Waterloo (1B26).

4.2 Suburban railway network case study in Denmark: using spare capacity

The UIC 406 capacity method can not only calculate the capacity utilization but can also examine how many extra trains that can be operated. This is done by adding more trains to the timetable after the timetable compression and gives an indication of how much extra capacity will be used. However, the methodology does not describe which type of trains should be added to the timetable. To analyse the sensitivity of the UIC 406 method when adding more trains, the line section from Holte to Hellerup (Figure 10) on the Copenhagen suburban railway network is analysed.

The line section form Holte to Hellerup has two train routes – E and B – that are operated in a fixed interval timetable [27] with a 10 minute frequency each. The operation from Holte to Hellerup is heterogeneous as train route B stops at all stations while train route E skips 5 stops. The analysis covers a 12 hour period from 6:30 to 18:30 where the line section has a capacity utilization that has been calculated to 75.5%.

With no quality factor it is examined how many trains that can be added to the timetable in this 12 hour time period. The results are either:

- 62 extra B trains or
- 45 extra E trains or
- 21 B trains and 22 E trains alternately (43 trains in total) or
- 22 E trains and 21 B trains alternately (43 trains in total)

Figure 10- The analysed section from Holte to Hellerup. Based on [8]
Figure 11- Blocking time stairs of B and E trains and safety overlap for E trains

With a difference of up to 19 trains in the analysis period, the result shows that the UIC 406 method is highly sensitive to which kind of trains that is added to the compressed timetable. If the two train routes are added alternately, the fewest trains can be added due to the heterogeneity. However, the difference between the two train routes B and E when added homogeneously is more surprising as the difference is 17 trains. The reason for this big difference is that train route E arrives to Holte from Hillerød station while train route B starts from Holte. When train route E arrives at Holte station it reserves a safety overlap after the exit signal (Figure 11) resulting in a long block occupation time which is not necessary for train route B that turns around at the station (at a separate platform track).

As discussed above, for adding extra trains there are different options possible. Purely based on macro capacity utilisation, it can not be advised which to choose. Therefore a meso capacity index is needed. For the scenario of adding trains, the following meso index is suggested:

\[
\text{Meso capacity index to add trains} = \frac{\text{Micro capacity gained}}{\text{Macro capacity lost}} \quad (5)
\]

\[
\text{Meso capacity index to add trains} = \frac{n_{et}}{C_a - C_b} \quad (6)
\]

\( C_a \): Capacity utilisation after adding extra trains
\( C_b \): Capacity utilisation before adding extra trains
\( n_{et} \): Number of extra trains

Further meso capacity indices (e.g. by considering load factor) can be developed to decide how to add extra trains. Alternatives that yield higher meso capacity indices are more efficient which in this case study would be options with more B trains.
5 Conclusions

Different aspects of the railway capacity and improvement cycle, from definition to analysing, improving and controlling capacity utilisation, need to be explored more comprehensively and systematically like the Highway Capacity Manual. The nature of railway capacity is even more complicated than road capacity as it is a multidisciplinary area. In this paper, major definitions of railway capacity are studied and the concept of lean capacity utilisation is suggested by differentiating between macro and micro capacity utilisation. Ignoring each of these aspects would result in a waste of capacity and resources and/or a low quality of service. Current metrics of capacity utilisation all have strengths and weaknesses which are discussed. The UIC 406 method and the CUI method by compressing the timetable estimate how much the infrastructure as a resource is not idle. The results of these two methods are compared with each other for a case study. Both methods generate close average capacity utilisation while the CUI method generates some 1% higher average as it considers headways times at nodes which encompass part of nodal capacity constraints.

For analysing and improving capacity exploitation, two case studies are done: 1- Freeing up capacity by omitting the train that has the highest capacity consumption (South West Main Line between Southampton Central and Worting in Great Britain). 2- Adding extra trains to use the spare capacity (line section form Holte to Hellerup in Denmark). Solely based on macro capacity utilisation, the decision can not be made regarding which trains to add or to remove. Therefore some meso indices are suggested to extend the UIC 406 method as a tool for making decisions in such cases.

References

[26] SAVE, Value standard and body of knowledge. 2007, SAVE International: USA.
[28] Smith, D., Timetabling is the culprit in ‘empty seats’ problem, in Rail Professional. 2009, Rail Professional Ltd: Cambridge
Paris, France.


