Analysing Rail Traffic Patterns in Operational Data

Stephan Zieger [a], Wiebke Lenze [a], Nils Nießen [a]
[a] Institute of Transport Science, RWTH Aachen University
Mies-van-der-Rohe-Str. 1, 52074 Aachen, Germany

Abstract
Reliable estimations of traffic loads are essential for dimensioning rail infrastructure adequately. Since the number of trains varies, the classification of typical flow profiles helps to specify the necessary capacity aim. Traffic flow profiles are already used for dimensioning highways, but not yet in railways. Therefore, we investigate the variation of train numbers for German rail infrastructure empirically. This paper analyses train numbers and delays using clustering methods. In this way, we demonstrate the benefits and limits of applying cluster methods on rail traffic while also creating awareness for the variation of train numbers which is the basis for infrastructure dimensioning.

Keywords
Cluster analysis, Flow profiles, Infrastructure development, Train numbers, Traffic pattern

1 Introduction

Reliable traffic load estimates are essential for strategic and operational planning of transport networks. The estimated number of vehicles is the basis for dimensioning transport infrastructure and the accuracy of this input parameter determines whether the infrastructure can be utilised adequately.

Based on historical data, typical flow profiles for roads have been established in several countries. Weijermars and van Berkum (2005) investigated the determination of historical traffic patterns and classified five clusters that show distinct flow profiles on one location of a highway. After a pre-classification into working days and non-working days, they clustered daily traffic profiles. Working days were distinguished into four types: Mondays, core week days, Fridays and days within vacation periods. They suggest using the resulting patterns as a basis for traffic monitoring, forecasting, and modelling.

Pinkofsky (2005) analysed the results of approx. 1200 permanent automatic counting stations on German highways and divided them into seven cluster types. Typical yearly, weekly and daily flow profiles were distinguished.

For traffic forecasting, Caceres et al. (2012) developed a methodology for estimating traffic flows using road features as clustering variables. They used a factor related to the characteristics of nearby areas to cluster road sections and derive typical flow profiles.

The measured or estimated daily number of vehicles over a year is relevant for dimensioning infrastructure. As a basis for designing highways, traffic volumes of the 50th peak hour are commonly used in Europe. The United Nations Economic Commission for Europe accordingly proposes to use the 50th highest hourly traffic volume to determine their level of service (UNECE (2002)) and includes the 50th highest hourly traffic volume in their transport statistics. German highway capacity manual (Forschungsgesellschaft für Straßen- und Verkehrswesen (2015)) also defines the 50th highest hour as the design hour for highways.

Traffic volumes vary not only on highways but also on rail infrastructure. To some extend this is accounted for by clustering railway lines. European Technical Specifications for Interoperability (TSIs) include categories of railway lines that are classified based on gauge, axle load, line speed, train length and usable length of platform(European Commission (2014)). German infrastructure manager DB Netz AG developed 10 standard lines that are characterised by design speed, main purpose and utilisation rate (DB Netz AG (2020)). These standards distinguish passenger traffic lines, mixed traffic lines, freight traffic lines and regional lines with different design speeds each. For example, passenger traffic line P 230 should be planned for maximum 70 long-distance and 50 short-distance passenger trains per direction and day with a design speed between 161 and 230 km/h while freight train line G 50 is planned for maximum 20 freight trains per direction and day with a design speed of 50 km/h. The directive 413 of DB Netz AG (DB Netz AG (2020)) gives recommendations for designing those lines. Not only the design speed but also characteristics such as lengths of block sections and types of train control systems are recommended.

Classifying train lines into different categories takes into account that line characteristics such as design speed, train control systems and block section lengths affect the capacity of lines. However, the traffic load's variability is not explicitly included.

To the best of the authors knowledge, investigations of empirical train numbers using clustering methods on a whole rail network has not been published so far. Liu et al. (2018) merely analyse train capacity utilisation in China and use clustering techniques to show the unique characteristics that affect the passenger load factor.

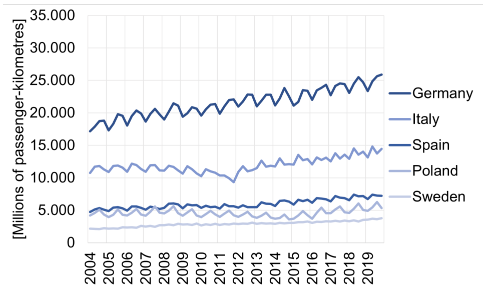
With the objective of clustering rail lines' traffic loads, this paper analyses the number of trains and their yearly, weekly and daily variation. Section 2 gives an overview on the known and expected variation in rail traffic and sets up working hypotheses. In Section 3, data of train numbers and delays for one year are analysed using clustering methods. As an example, the analysis is performed on trains operating on German railway infrastructure. A global analysis examines the variation in traffic loads of all lines and helps to identify lines with special characteristics. Afterwards, examples for train number variations are shown for train lines with mainly long-distance traffic, local traffic, S-Bahn (suburban trains) and freight traffic. Finally, Section 4 summarises the results and gives an outlook on further research possibilities.

This paper aims to contribute to the awareness of train number fluctuations that need to be considered when evaluating the capacity. Furthermore, it demonstrates the benefits and limits of applying cluster methods on rail traffic.

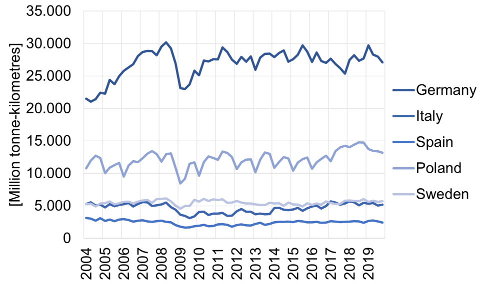
2 Variation in Rail Traffic

It is known that the number of trains varies over the course of a year. Based in data by Eurostat, Fig. 1 shows quarterly traffic loads on rail infrastructure for 15 years (2004-2019). Exemplary for five countries, Fig. 1(a) illustrates the transport performance measured in millions of passenger-kilometres and Fig. 1(b) illustrates the development of freight transportation in million tonne-kilometres.

Over the last years, the transportation of passengers has increased. The quarterly data also shows distinct yearly variations with the first quarter of each year usually having the fewest passenger-kilometres.



(a) Passenger transported



(b) Goods transported

Figure 1: Quarterly transport performance data based on Eurostat (Eurostat (2021a,b))

Fig. 1(b) shows the variation in freight transportation. The yearly fluctuations are less regular than those in passenger transportation. Apparently, there are more factors than the time of year relevant for the amount of freight transported. For example, a distinct reduction in freight transportation can be seen around 2009 as the effect of the global financial crisis.

These variations in passenger-kilometres and tonne-kilometres lead to the conclusion, that the number of trains on rail infrastructure also changes significantly over a year. Of course, transport measured in passenger-kilometres and tonne-kilometres do not necessarily correlate with the number of trains. Theoretically, only the number of passengers or tonnes per train could be changing. Nevertheless, it is a realistic assumption that, e.g. around Christmas, not only more passengers are travelling but also that transport companies increase the number of trains to meet the demand.

Expected number of trains are essential input parameter for dimensioning rail infrastructure. Infrastructure capacity is defined as the number of trains being able to operate on a given infrastructure during a period of time at a defined level of service. The capacity is based on its utilisation, in particular on the interdependencies between the four aspects number of trains, average speed, stability and heterogeneity (UIC (2004)). This paper focuses on the aspect number of trains.

A well-known capacity analysis method is the so-called compression method by UIC (2013). UIC Code 406 defines the occupancy time rate as occupancy time per defined time period. The occupancy time is measured by compressing a timetable's blocking time stairways, therefore this value is clearly impacted by the number of trains. UIC (2013) proposes occupancy time rates for different types of lines to achieve an acceptable level of service. Limiting the permitted capacity consumption and therefore the number of trains is necessary to ensure acceptable operational quality.

The compression method (UIC (2013)), however, has the disadvantage that it only evaluates the utilisation of a line for trains scheduled in the timetable for a chosen day. This number can be too high or too low on a frequent basis. As a result, a line may be overdimensioned, which is bad from an economic point of view, or, on the contrary, underdimensioned. The latter results in conflicts during daily operations. The possible difference between planned and actual number of trains is a relevant aspect to consider when dimensioning rail infrastructure. Unlike road transport, trains are operated using scheduled train paths.

The operational quality is usually measured by occurring delays or rather punctuality. Several studies have indicated the significance of punctuality for customers - both in passenger and freight transportation. Requested by the European Commission, TNS Opinion & Social (2014) conducted a survey investigating how quality of rail transport is perceived. More than 33 % of the respondents answered that lack of reliability and punctuality were the most serious problems affecting rail transport in their country.

There are many different influences that cause delays in daily operations. Primary delays are caused by disruptions such as high passenger traffic volume, extreme weather conditions, malfunctioning of infrastructure or rolling stock (see e.g., Økland and Olsson (2021) and Wen et al. (2017)).

Primary delays can cause a cascade of secondary delays or so-called knock-on delays. Especially in highly used networks with small buffer times between trains paths, a delayed train is likely to disrupt other trains. The number of trains is one of the reasons that determine the occurrence of (secondary) delays. Thus, their distribution might indicate the expected level of service.

Reliably estimating the expected number of trains and infrastructure occupation helps to adequately plan infrastructure and reduce the risk of supplying too much or too little capacity. To support choosing the correct number of trains, this paper analyses the variation of train numbers over a year. Based on the data shown above, yearly variations are expected. Certain events such as constructions, extreme weather events or Christmas holidays impact the available capacity of infrastructure. Furthermore, we expect weekly and daily changes due to changing demands. Flow patterns for highways also show distinct yearly, weekly and daily variations. On working days, peak hours in the morning and afternoon can be seen. Weekends show more regular distribution of traffic (cf. Pinkofsky (2005)).

For rail infrastructure, we propose the following hypotheses:

- Railway lines' utilisation are heterogeneous. Some lines are highly utilised whereas only few trains operate on other lines.
- The number of trains varies during a year, week and day.
- Clustering rail lines on the basis of train numbers and their variations is possible.
- A growing number of trains on a line increases occurring delays.

3 Analysis - Use-Case German Rail Network

To examine the posed hypotheses, we analyse number of trains and delays on the German rail network. The procedures can, however, be implemented for any regional or national data set. After giving details on the underlying data, a global analysis of train numbers on all examined lines gives an overview on the data set. With cluster analysis tools, lines with similar characteristics are identified and selected exemplary lines are investigated in detail. Finally, the hypotheses are checked and results are discussed.

3.1 Data

The underlying data for the analysis includes train data of a major part of the German rail network, which consists of more than 33,000 km railway lines (Deutsche Bahn AG (2020)). Especially smaller lines with (partially) manual recording data might not be included. The data for the period from 01.07.2015 to 30.06.2016 is analysed.

Each day is classified in separate files, depending on the purpose of investigation. For this analysis, the date, train ID, train category, timing point location, event, time stamp for planned and actual time, and planned and actual line number are relevant. The events can be the ”start”, ”end”, ”arrival”, ”departure” or ”passing” of a train in a timing point location. The data set consists of approx. 8,400 timing point location which are usually railway stations and will be denoted as such in the text. These are in some cases divided in passenger and freight station. The time stamps for planned event times are derived from the timetable. On the contrary, the time stamps for the actual events are usually derived from the signal states on a train route, i.e. they depend on the type of interlocking that is present on the line. Each weekday, approx. 45,000 train IDs and approximately 1,300,000 events are recorded.

For confidentiality reasons, the data and corresponding results can only be presented in an anonymised form.

3.2 Global Analysis

The global analysis' premise is to analyse the data to the extent that the general train movement distribution is captured and the validity is assessed whenever possible with external sources. Due to the amount of data, it makes sense to choose a station for each line that represents it. For shorter lines, little variation in utilisation is to be expected. For longer lines, changes may well occur, but these are accepted in favour of automation.

Some assumptions are necessary for the analysis. For example, in main stations, a line number is assigned to each track. However, trains of different lines may run over these, so that the data often cannot be clearly interpreted at this point. Furthermore, the traffic load in large nodes is usually significantly higher than on the surrounding lines. Therefore, we choose a station with the median load as the representative for each line. For each of the line numbers available in the data set, the planned and actual train number is counted up for each day on the representative station.

Several tables can be created from the data with this approach. First, the tables containing planned and actual train number are of size 366 (number of days between 01.07.2015 and 30.06.2016; leap year) times 1665 (number of lines). Officially, 1717 lines were registered in the network in 2015 (DB Netz AG (2015a)) such that the data set includes most of them. Each line consists of one or two tracks. If there are three or more tracks in parallel, two or more distinct line numbers would be allocated for this section. Second, the delay data for the chosen station is recorded for all days and lines, yielding the same size. Third, the train category is recorded. In the data set, trains are classified into various categories, e.g. InterCity or long-distance freight train, and further sub categories, e.g. regular transport, empty running or shunting, to determine the exact type of movement. These fine categories were condensed into five main categories: Long-distance passenger train, local passenger train, S-Bahn (urban railway), freight train and other. This results in a 366 (number of days) times 1665 (number of lines) times 5 (number of categories) table.

Fig. 2 shows a histogram over the average daily planned and actual number of trains on the railway lines. For better readability a few outliers above 320 average daily trains are omitted. These are in the region of 900 average daily trains and are high-volume S-Bahn lines with a two to three minute headway time which are nearly completely saturated during day time. The lines with the smallest traffic volumes are often (partially) single-tracked lines whereas the lines with high traffic volumes are double-tracked (nearly) everywhere. The

German railway network consists of many branch lines which are often short and increase the number of lines accordingly due to their tree structure.

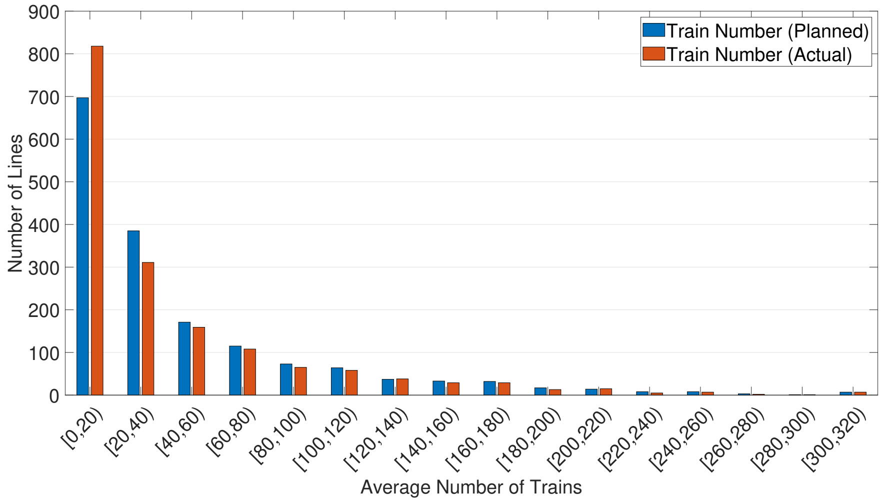


Figure 2: Average daily planned and actual number of trains on the railway lines

A large discrepancy between planned and actual average daily number of trains is apparent. Generally, both follow an exponential distribution with a high amount of lines with low volume and only few lines with high throughput. The resulting number for the planned trips seems to overestimate the summarised number of train movements slightly, but seems to be rather accurately when compared to the individual passenger and freight train classes provided by Statistisches Bundesamt (2017) – the Federal Statistical Office. This observation is helpful to explain the difference between the evaluation and other data sources. When using the median station as a representative of a line it can happen that these are dedicated passenger or freight stations and therefore not all trains on the line are recorded. Although the data provided by the Federal Statistical Office is structured differently, e.g. it is not clear how they determined the number of trains on a whole line and for the complete year of 2015, it generally fits the data set results.

According to the German infrastructure manager (Deutsche Bahn AG (2019)), 26% of the interlockings are still mechanical and these cover 15% of the network. As these are unlikely to record train movements, the actual number of trains in the data set is massively decreased on the corresponding lines. However, these lines make up a large share of the network, but only a very small share of the transport performance. Furthermore, these lines are often not interesting as the traffic volume is low (a regional train every one or two hours on weekdays or occasional freight traffic) and the infrastructure is usually adequately dimensioned. Hence, the general reliability for lines with low traffic volumes is smaller and planned and actual traffic have to be looked at separately and cannot be compared directly. To circumvent these effects, only actual data will be discussed in the following.

The second axis of observation is not to check the deviation between planned and actual on the basis of the railway lines, but to look at the course of the month. For this purpose, the daily average for actual train numbers is shown and broken down by train type for each month in Fig. 3. Overall it can be seen that, in correspondence to Fig. 1, not only the transport performance alters throughout the year, but this is also correlated to the number of used trains.

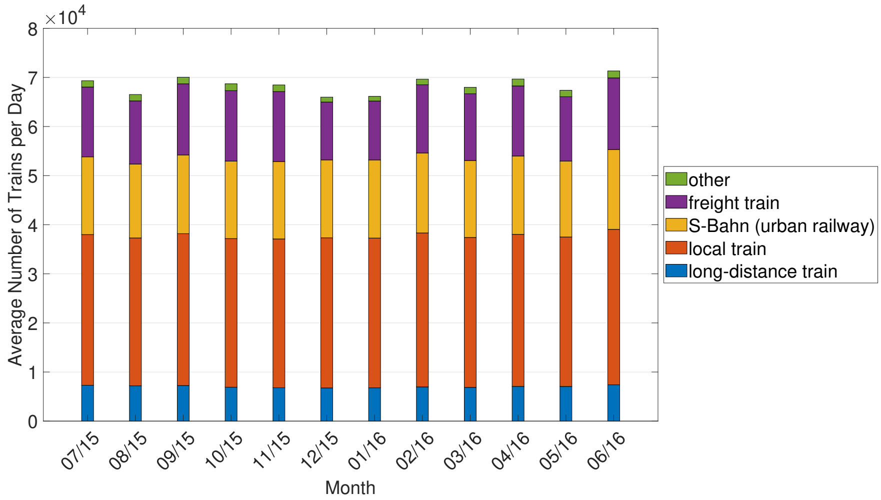


Figure 3: Average daily actual number of trains per month over all lines

Not surprisingly, the train numbers for local passenger trains and S-Bahn are very constant over the months. These provide a contractually agreed service, usually covered by Public Service Contracts (European Commission (2020)), that is characterised above all by its regularity. The long-distance trains and freight trains operate demand-driven and are planned accordingly with monthly fluctuations. In long-distance transport, these fluctuations can be explained primarily by holidays, additional trains or known major events over the long term. In freight transport, there is a certain basic load for regular transports, for example for chemical or mining companies, but also many short-term or cyclical transports, such as foodstuffs.

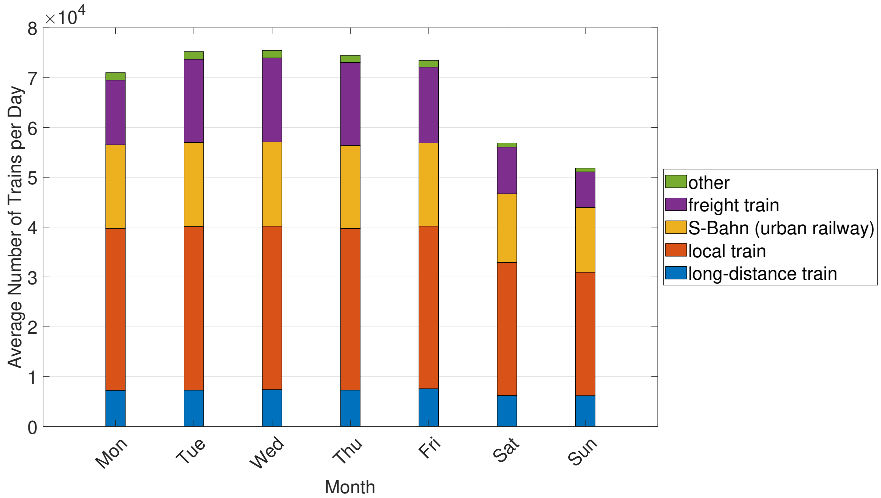


Figure 4: Average daily actual number of trains per day over all lines

Fig. 4 gives an overview of average train numbers by day of the week. Saturday and Sunday are less busy and are not usually decisive for the capacity utilisation of individual lines. The services provided in passenger traffic are reduced by about 20-25% at weekends. Freight traffic declines by about 40-60% on weekends. On Mondays, the average transport performance is also lower compared to the rest of the week, which is mainly due to freight transport, but also partly due to long-distance transport. There are on average 5% more long-distance trains and about 8% fewer freight trains on Fridays compared to the period from Tuesday to Thursday. Hence, the weekends as well as Monday and Friday are generally not suitable for assessment and the period from Tuesday to Thursday should be chosen for line capacity assessment.

3.3 Identification of Lines with Similar Characteristics

In the first part, the question of whether it is possible to identify lines that have similar characteristics will be examined. Lines are classified as similar if they have a comparable annual pattern in terms of the number of trains. For this purpose, the rows of the table with the actual train numbers will be clustered. The actual train numbers are preferred in the following as they reflect the actual train movements.

An apparent approach is the use of cluster analysis methods. There are various possibilities, such as the use of centroid-based, density-based, hierarchical or distribution-based clustering. The first attempt to use the probably most widespread centroid-based method, the k-means clustering algorithm or one its variations such as k-means++ (Endo and Miyamoto (2015)), was not successful. It turns out that the clusters formed are very heterogeneous due to the number of outliers and the different utilisations of the lines (see Fig. 2). A significant problem is that clusters have to be of similar size.

The attempt to use the density-based method DBSCAN (Ester et al. (1996)) provides the following findings. As soon as the number of minimum points in a cluster ($minPts$) is set to more than 1, all lines are identified as outliers. If $minPts = 1$ is set and the distance as a similarity measure (ϵ) to the other lines is chosen small or moderate, 300-400 clusters are identified. If ϵ is chosen larger, the number of clusters decreases very quickly and only one cluster is output for all lines. Furthermore, the produced clusters often have a high range in train numbers on the involved lines and are therefore not really similar. The results of the k-means and DBSCAN indicate that a clustering should probably have clusters of different sizes and that the gap between the clusters will be high.

Since standard methods do not seem to be very effective, we use our own procedure for clustering, which is a fusion of hierarchical, density-based and centroid-based concepts. First, a distance matrix of size (number of lines) $\times$ (number of lines) is initialised. In the next step, the distances between the line vectors are determined calculating the Euclidean norm of the difference for each vector pair. Then an ϵ is to be chosen, in this case $\epsilon = 100$ is set, and all distances greater than ϵ are set to 0 in the distance matrix. As a result, only pairs of distances that have a certain similarity remain in the matrix. Subsequently, the line with the largest number of non-zeros in the distance matrix can be chosen as the first cluster centroid and all associated lines can be included in this cluster. Then, the entries belonging to the selected lines are set to zero and the iteration continues until the last line is assigned to a cluster.

This procedure, like the others, has certain drawbacks. On the one hand, a distance metric and an ϵ have to be chosen, which influence the design of the final clusters. On the other hand, shorter major events, such as a line closure, can increase the distance between two lines in such a way that they are no longer combined in a cluster, even though they are very similar in the rest of the year. Dealing with noisy data is often a challenge in the development of cluster algorithms.

Fig. 5 shows the first 21 clusters out of 841 clusters formed according to the method. From cluster 44 on, they contain only single lines.

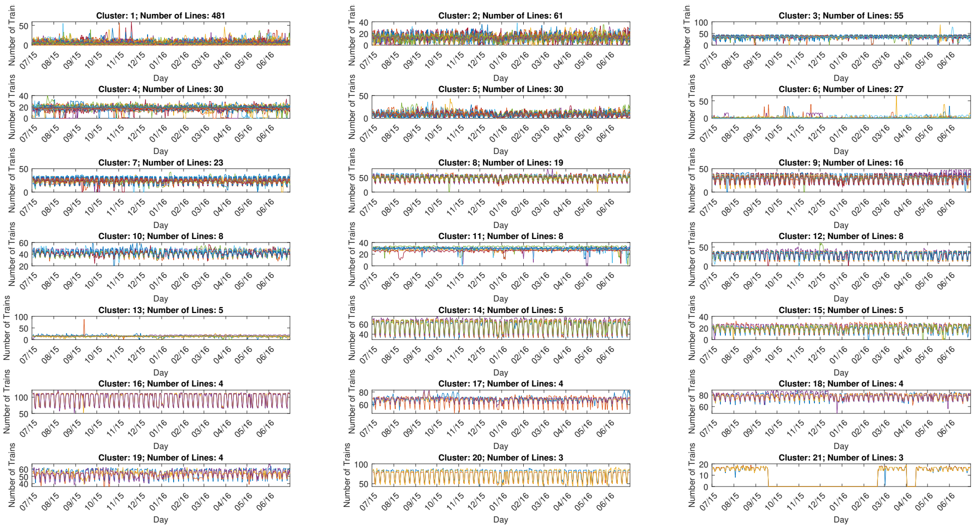


Figure 5: 21 largest line clusters produced with the described method

Each graph shows the number of trains for each line (colour) through the course of the one year long investigation period. It can be observed that the number of lines in the identified clusters decreases very quickly, which supports the hypothesis of strong overall heterogeneity. Furthermore, the mixed clustering approach leads to the result that in fact not only lines with similar load factors but also similar operating schedules are identified, e.g. lines with very constant load factors and always the same drops in train numbers at weekends. Single events occur in nearly all of the clusters, but they usually do not provoke a substantial offset between two line vectors. Strongly positive deviations from the standard utilisation can be a result of diverted traffic, additional trains for holiday season or large events, or reallocation of rolling stock, so that the utilisation of these lines is significantly higher than usual on certain days. Spikes in the other direction can be observed as well which are, in the regular case, the weekends and, in the single events, correspond to maintenance work, major disruptions, holiday seasons or similar.

Due to the high number of lines with an average of only 0-20 trains per day, many of these lines are sorted into the first and largest cluster, as expected. This cluster combines lines that generally have a very low load factor, but some single events can also be identified. The clusters 2, 4, 5 and 6 have similar structures compared to the first cluster. Clusters 2, 4 and 5 just have another base load of the included lines, but otherwise a similar traffic pattern. Cluster 6 seems only to consist out of lines which have basically no traffic, but in seldom occasions.

Many of the other clusters are characterised above all by the greatly reduced utilisation at weekends. The strong dips can be seen very well in clusters 8, 9, 12 14, 15, 16 and 17. There are very regular deviations here, which could be crucial for the similarity check. Cluster 21 impressively demonstrates the problem in clustering a year of train recordings.

As soon as disruptions such as longer construction works result in a strong deviation of the train numbers the distances between the lines become too large and they are identified as dissimilar.

If only the working days were considered, significantly larger clusters would result, as the distances between the vectors would then be smaller. However, the overall behaviour of the lines is definitely relevant and it is an operational difference whether the same number of trains run at the weekend as during the week or not. In general, it is evident how the similarity in the clusters increases significantly with smaller size.

A second, but more coarse, approach is not to cluster the lines according to their traffic patterns, but according to the type of traffic. Therefore, the table of train categories is evaluated again. Each line should be classified as either a long-distance train (high-speed), local train, S-Bahn, freight train or mixed traffic line. For this purpose, the different train types are added up over the year and classified. With the total number, a line with a chosen barrier of 80% of one type is sorted in a special category and is otherwise identified as a mixed traffic line. This approach has the drawback that the chosen representative station on the line might have traffic on the station tracks assigned with their line number which is then not operating on the line itself. Furthermore, the chosen station might not be representative for the type of transport on the line. However, as an approximation this approach yields sufficient precision.

The classification method results in 23 long-distance train, 517 local train, 100 S-Bahn, 325 freight train and 700 mixed traffic lines. This result is anticipated as dedicated high-speed lines in Germany are scarce and the few are often very long. Some of these lines are also only transition curves, which receive their own line number. The result might not be accurate for each line as only one representative station is chosen. This station, though, might not be representative for the train mix on the whole line. The resulting clusters are illustrated in Fig. 6.

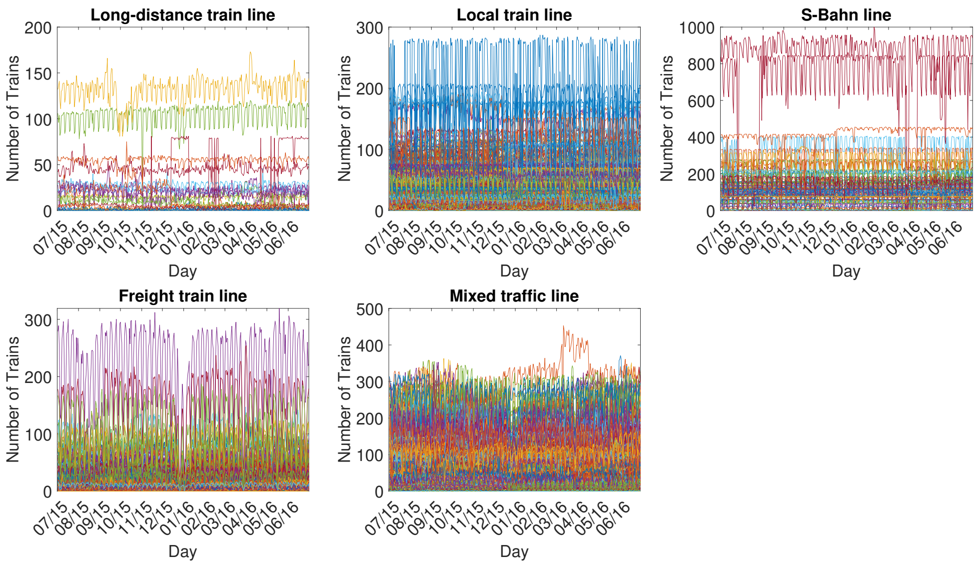


Figure 6: Clusters according to line specification

The classification of the lines according to the predominant train type shows how different the capacity utilisation is on the lines, but also how similar or dissimilar the lines are over the course of the year.

First of all, it is noticeable that local train and S-Bahn lines generally have a very constant capacity utilisation throughout the year, which is only temporarily reduced by events such as construction sites or line closures. An increase in utilisation rarely occurs, except for timetable changes, such as the third line from the top in the S-Bahn chart, or temporarily during major local events. It is also noteworthy how different the utilisation is at the weekend. In some cases it drops by about a third and in other instances it is almost as high as during the working days. The two lines with very high daily train numbers result from highly optimised and utilised double-track railway lines in very large cities with nearly all-day operation and headway times of two to three minutes.

In long-distance traffic, the major long-distance lines stand out. Particularly prominent are the high fluctuations over the course of the year on some lines. Due to the low availability of dedicated high-speed lines, these are particularly susceptible to events such as changes in traffic demand or diversions due to closures. The high variance makes the selection of the day particularly relevant for the capacity assessment of the line very difficult.

Freight transport is characterised by a high degree of flexibility. As a result, the utilisation levels fluctuate, sometimes massively, during the week, but strong fluctuations can also be seen across the weeks. The drop around the Christmas holiday is particularly striking, when operations seem to be significantly reduced.

In mixed traffic, different train categories interact. Here, the more regular local trains form the base load and long-distance traffic and freight traffic are mainly responsible for the variation in train numbers. Due to the high number of mixed traffic lines, they also cover a very large utilisation spectrum. As a result of the strongly varying loads, the correct dimensioning of these lines is particularly important and the choice of investigation period is especially delicate.

Conclusively, it is difficult to identify common behaviour in these groups and the high degree of variation is particularly bad for meaningful pattern recognition. Nevertheless, some particularities of the line types are investigated in next section.

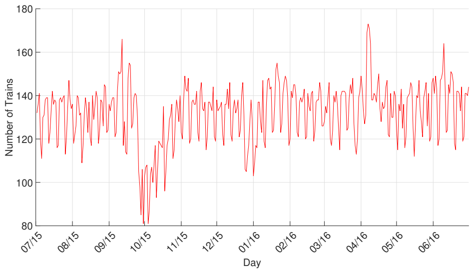
3.4 Exemplary investigation of railway lines

In the following, one representative line will be presented and examined in more detail for each of the line categories. In doing so, different characteristics of line types will be dealt with and special features in the process will be discussed. For these lines, the train numbers are also compared with the delays that occurred during operation in order to check the influence of the choice of the investigation day. The amount of delays is clearly determined by many potentially overlapping influences. Assuming some continuity of the other influences the number of trains is assumed as sole and at least significant input parameter in the following. The delay of early trains is set to zero which does not change too much in passenger transport, but might have significant effects on the freight transport evaluation which is discussed in the appropriate sections. In addition, existing outliers or errors in the data set might affect the result.

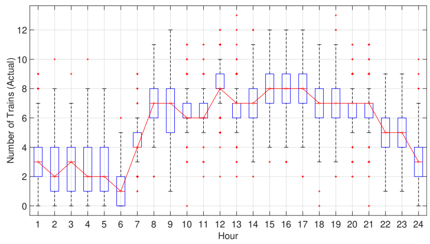
Long-distance train line
The number of dedicated high-speed lines in Germany is small. Therefore, a line with a relatively high traffic load is chosen in order to obtain as much information as possible and

to be able to better evaluate single events. The chosen representative station on the line has a long-distance traffic share of about 90%. However, only long-distance traffic operates on the line itself. The other recorded trains just use the station tracks associated with the line number from time to time.

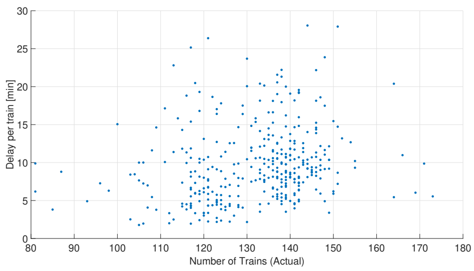
Fig. 7(a) shows the actual train numbers over the course of a year for a long-distance train line. In general, the line shows significant variation in and over the weeks despite the high regularity of the traffic. There are some particularities in the course of the year. Around the beginning of October, the train numbers decrease significantly. The period and the restrictions suggest that there were holidays and construction work on the line at this time. The second dip at the end of December is likely caused by the holidays around Christmas.



(a) Actual daily train numbers over the course of the year



(b) Distribution of the hourly actual train numbers over the course of a day



(c) Delay per train depending on the utilisation of the railway line

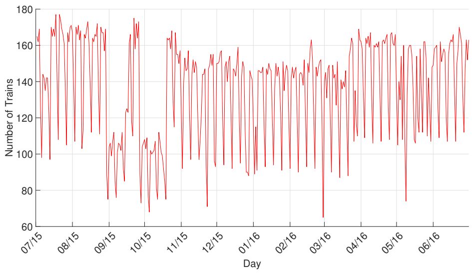
Figure 7: Traffic and delay pattern for a long-distance train line

The occupancy rate throughout the day (cf. Fig. 7(b)) shows that the days were very irregular and that the occupancy rate also fluctuates greatly between morning and evening. This may be due to the position of the station, as this has a significant influence on when the long-distance trains, which usually perform very long trips, arrive at the station. As long-distance traffic is demand-oriented, there are almost no movements at night, but some freight trains which sometimes use the tracks allocated to long-distance traffic during this time. The line's behaviour shows that it is difficult to choose a representative day or hour for the capacity assessment of the line.

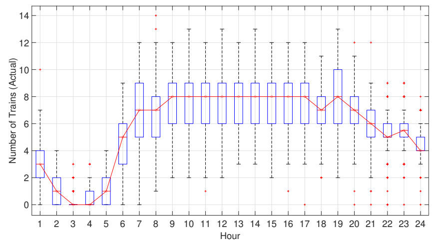
This assumption is also confirmed if the delay experienced in operation is taken as a quality measure. A clear trend emerges, showing that delays do not only grow linearly with increasing train numbers. The assumed volatility of train numbers is very clearly shown in the scatter plot in Fig. 7(c). Two clusters around 120 trains (weekend) and 140 trains (working day) are recognisable. With the knowledge gained before, this part of the line should be able to handle 140-150 trains per day with reasonable quality.

Local train line

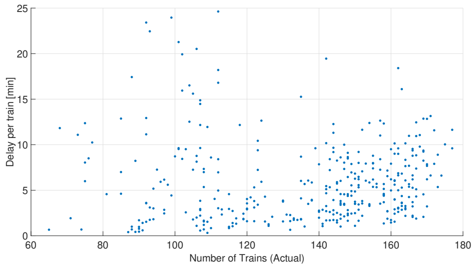
Local train lines are characterised by their large variety. These lines range from single-track lines in rural areas with a service every 2 hours up to densely utilised double-track lines with mixed-traffic regime and long-distance trains as well as freight trains on the line from time to time. Consequently, the operation can be homogeneous or heterogeneous which has strong implications, e.g. in terms of delay propagation.



(a) Actual daily train numbers over the course of a year



(b) Distribution of the hourly actual train numbers over the course of a day



(c) Delay per train depending on the utilisation of the railway line

Figure 8: Traffic and delay pattern for a local train line

Fig. 8(a) shows the actual train numbers over the course of the year for a local train line. It is noticeable that several prolonged variations can be observed over the course of the year. First of all, the difference between weekdays and Saturdays, and to an even greater degree Sundays, is very significant. In addition, there are several longer reductions in capacity utilisation due to 5 weeks of construction work, which reduce capacity utilisation by almost half. Interestingly, between November 2015 and March 2016, one of the services appears to no longer use the line due to adjacent construction work. This effect is observed over a longer period of time, so even looking at a line for several weeks can lead to erroneous conclusions about load factors if no further information is available.

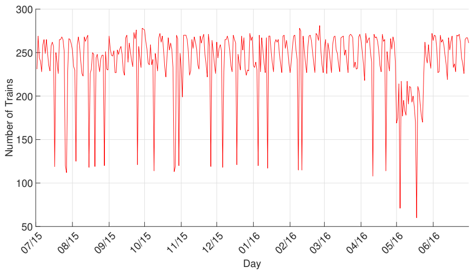
The hourly distribution of train numbers during a day is presented in Fig. 8(b). It clearly indicates the distinction of day and night hours. During the daytime, the number of trains is constant whereas the number during the night hours decreases and a rest period is observable. These hours without traffic are imminent for the recovery of potential delays from the previous day and for the clean set up for the next day.

Furthermore, delays depending on the number of trains are illustrated. The delay data from Fig. 8(c) show a strong trend. The delays clearly increase in a super-linear manner for larger numbers of trains on the line. The strong scattering of train numbers over the year is also particularly apparent. Although the clusters for weekdays and weekends are recognisable, the dispersion is nevertheless very high. For a regional line with a relatively dense frequency, the delays obtained are very high and the line does not seem to be sufficiently dimensioned for the high utilisation rates. Although the majority of days have less than 160 trains, the development of delays should be examined to see if the line can be upgraded to handle at least 160 trains on a regular basis, allowing for future increases in transport capacity.

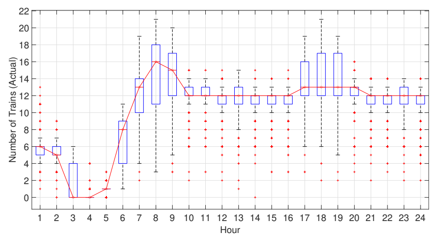
S-Bahn line

S-Bahn lines are characterised by their homogeneous traffic. Typical for passenger traffic is the recurring number of trains every week. The number of trains on week days is clearly higher than on weekends.

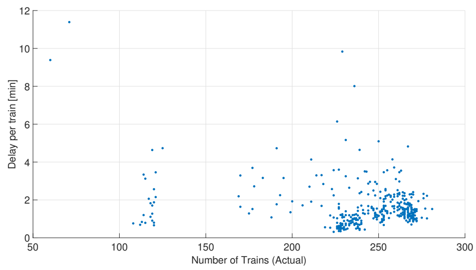
Fig. 9(a) shows the actual train numbers over the course of the year for a S-Bahn line. The traffic load on S-Bahn lines such as this one is particularly high. With more than 250 trains on a double-tracked line, the number of trains on S-Bahn lines is substantially higher than on mixed traffic lines. This is possible due to the homogeneous traffic with similar train types that have high acceleration values. Furthermore, the infrastructure is usually optimised for short distances between trains.



(a) Actual daily numbers over the course of a year



(b) Distribution of the hourly actual train numbers over the course of a day



(c) Delay per train depending on the utilisation of the railway line

Figure 9: Traffic and delay pattern for a S-Bahn train line

Some variations in train numbers can be observed on this line. Contrary to expectations, train numbers also vary significantly across weekdays. The line continues to be affected by construction work on a very regular basis during the investigation period, in 2015 mostly on weekends and 2016 then for several weeks.

Fig. 9(b) shows the distribution of train numbers during a day. At night between midnight and 6 in the morning, the hourly number of trains is very low. Here, a rest period is observable as well and can help reorganising in case of larger disturbances during the previous day. During the day, at least 12 trains per hours are operated with a peak in the morning and afternoon. These peaks correspond to the passenger demand peaks, mostly for people on their way to or from work and pupils. The operators deploy additional trains during these hours to increase their transport capacity.

Fig. 9(c) illustrates the delay on this particular line depending on the number of trains. The clusters for weekdays and weekends as well for the construction works are well observable. The majority of trains is delayed by less than three minutes which is usually the maximum for high-density S-Bahn traffic. Usually, these trains get cancelled when the delay exceeds some threshold as these trains would massively propagate the delays. However, in the region of 250 trains onward the average delay shows a clear increasing trend.

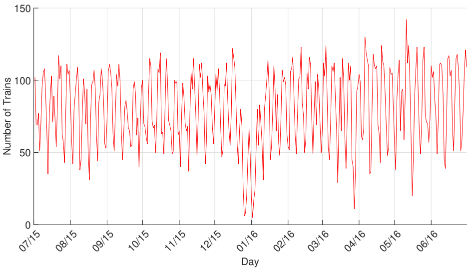
Freight train line

Train lines with a majority of freight trains are expected to have heterogeneous traffic loads during each day and also during the year. The operation is very demand-driven and depending on the line the traffic can be very regular or be even seasonal with (nearly) no usage of the line for several months.

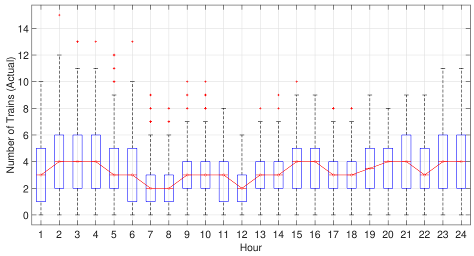
Fig. 10(a) shows the actual train numbers over the course of the year for a freight train line. The number of trains varies daily. The lowest number of trains can be noted for Sundays. Usually, Wednesdays and Thursdays have the highest traffic loads. The graph for the yearly utilisation indicates that the dimensioning of such a line is rather difficult as the load varies drastically.

Fig. 10(b) illustrates the hourly traffic load distribution during a day. In contrast to the previously presented train lines with passenger traffic, this freight train line displays no difference between daytime and nighttime. This is actually an advantage on this line, as the load is very evenly distributed and there is no need to dimension according to the peak hour. If there are one or more peak hours, this often leads to a very strong build-up of delays or the infrastructure is oversized for most of the day. In the case of a very even load, the quality is typically equally high throughout the day.

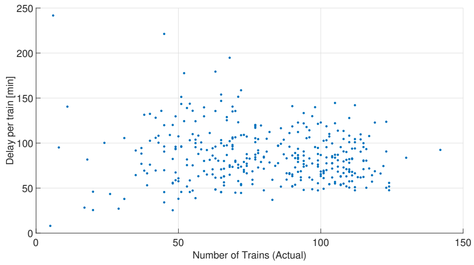
Just as the number of train varies so do the delays. Fig. 10(c) shows the delays depending on the number of trains. For freight trains conflict-free slots are reserved in the timetable for legal reasons. However, freight trains can move up to 3 hours early and up to 20 hours late compared to their reserved slot as regulated in the Network Statement (DB Netz AG (2015b)). For strategical reasons, freight trains often operate late on purpose as they can then usually move without further stops due to not being early and put aside. This leads to a high unreliability in terms of the delay data and a high flexibility in freight train operation. Hence, it is difficult to read off a trend and set infrastructure recommendations for freight lines on the basis of absolute delay data only and the delay changes in the course of the journey could be considered.



(a) Actual daily train numbers over the course of a year



(b) Distribution of the hourly actual train numbers over the course of a day



(c) Delay per train depending on the utilisation of the railway line

Figure 10: Traffic and delay pattern for a freight train line

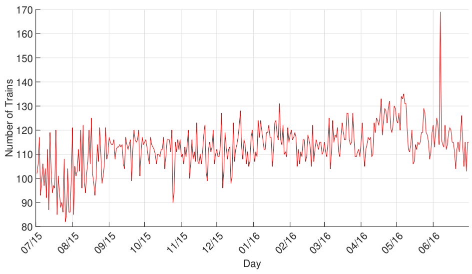
Mixed traffic line line

Mixed traffic lines can be very diverse and are by definition characterised by very heterogeneous traffic. Depending on which modes of transport are involved, the situation is very different. In the case of the line under investigation, only local and long-distance traffic is involved. No mixed traffic line with freight trains was chosen, as with these the correlation between train numbers and delays would be difficult to interpret, as in the previous subsection.

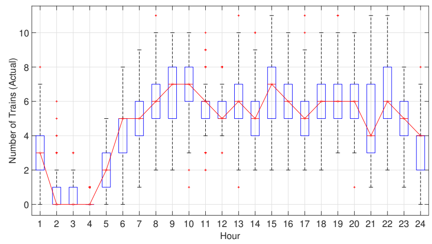
Fig. 11(a) shows the actual train numbers over the course of the year for a mixed traffic line. The traffic on this line is very volatile, as the local trains set the base load, but the long-distance trains sometimes run very irregularly and also spatially differently depending on the day. Particularly noticeable is the day in June when the line was used as an alternative route and many trains were diverted over it.

The daily pattern in Fig. 11(b) shows the passenger traffic pattern that was already observed before. During the day, the load fluctuates somewhat and at night very few or no trains run. The peaks in this case are generated by the long-distance trains, which run at different times of the day on this line due to their long running distances and irregular pattern. The volatility during the day is also very high.

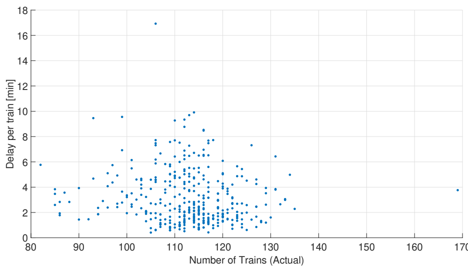
The average of delays in Fig. 11(c) is significantly influenced by long-distance trains. Due to their long running distances, these trains have a significantly higher probability of building up larger delays and then carrying them onto the route. In general, the line seems to be well utilised and the delays are kept within limits. Particularly interesting and worth highlighting here is the day on which the utilisation rose to almost 170 trains and this number of trains could be processed on the line at all with apparently even low average delays.



(a) Actual daily train numbers over the course of a year



(b) Distribution of the hourly actual train numbers over the course of a day



(c) Delay per train depending on the utilisation of the railway line

Figure 11: Traffic and delay pattern for a mixed traffic line

In fact, at that time there were already plans to run more services over the line, as it was already dimensioned for higher traffic volumes.

3.5 Discussion of the results

Based on the conducted analysis, we check the proposed hypotheses. The exemplary analysis of trains on German rail infrastructure over one year has clearly shown that the lines' utilisation differs significantly. While on most lines, less than 50 trains run per day, a small number of double-track lines have daily train numbers above 200. When interpreting these results, it needs to be considered that the examined lines have different characteristics. Depending on infrastructure, signalling systems and operational usage, the capacity of lines differ greatly. When looking at the variation of train numbers of specific lines, some time periods include noticeably less trains than others. Over a year, the average monthly traffic loads vary with the fewest trains running in December. This can be explained by the Christmas vacations. For some lines, prolonged variations of train numbers can be observed over the course of the year due to construction works or when lines are closed due to other reasons. Fluctuations during a week are distinctive with the busiest days being Tuesdays to Thursdays. On weekends, especially on Sundays, the traffic load is reduced. This has already been considered in most rules for capacity assessments. German infrastructure manager DB Netz AG for instance recommends the days Tuesday to Thursday as days with average traffic load for capacity assessment (DB Netz AG (2008)). These rules are, however, often solely based on planned traffic and ignore possible daily variations. These vary from consistently high utilisation throughout the day to distinct day and night times with well-defined peaks throughout the day.

Based on the number of trains and their variation, clustering rail lines is possible to a limited extent. Difficulties have occurred due to the heterogeneity of lines. A large number of lines operate with limited number of trains whereas other parts of the network are very busy. These lines usually have different track-side equipment and operation including the variation of train numbers can only hardly be compared. Furthermore, single events greatly influence the clustering. Time periods with unusual train numbers can be caused by, e.g. construction works or other external effects. This causes lines to be clustered in separate groups. Excluding the influence of single events can be done by pre-processing the data. That way, although complicated on its own, the generation of better results is expected for most clustering methods.

The number of trains seems to impact the occurring delays on lines. The analysis of the five exemplary lines shows that average delays on a line usually increases on days with high traffic loads. Since early arrivals were set to zero, the calculated delays might differ from the average delays in operation. When interpreting delays at a specific section of the network, it needs to be considered these delays might have been caused elsewhere. Interdependencies in rail networks complicate the search for the cause of delays and bottlenecks. For this purpose it is necessary to examine the infrastructure, timetable and operation in detail.

4 Conclusion

Knowing traffic loads and their variations is a an essential element of dimensioning infrastructure appropriately. For roads, typical flow profiles have been established that characterise the variation in traffic loads. These are used e.g. when dimensioning highways. The aim of this paper was to examine past traffic loads on railway infrastructure and identify typical characteristics of different rail lines using clustering methods.

For this purpose, different clustering methods have been tested. Each cluster method has its own advantages and disadvantages, which makes it difficult to recommend a general method. The evaluation of the data along the time axis for a dedicated station looks promising in principle. A natural extension would be not only to look for lines with similar absolute number of trains, but also for lines which share a similar progression of the number of trains. For this purpose, not the absolute number of trains could be compared, but the daily difference in train numbers compared to the previous day, either absolute or relative. Furthermore, looking at the number of trains over the entire length of the line could open up further opportunities. However, further pitfalls must be avoided here, as the length of the route or the number of stations now becomes relevant and the question arises as to how two lines are then compared. In addition, the methods are always dependent on the quality of the data basis and these data have some peculiarities.

Analysing German rail infrastructure has shown great variation of traffic loads on lines. A large number of lines are utilised by less than 20 trains per day while some lines have extremely high traffic loads. Examining long rail lines should be done in different sections since the lines' properties such as the number of trains or their composition might vary drastically. The separate analysis of lines has shown monthly, daily and hourly variation of train numbers. Distinct variation can be seen for different line types. Since the number of trains is an important input parameter when dimensioning infrastructure it is crucial to take this variation into account and chose a representative value. The provided capacity and the traffic load have direct impact on the service quality.

As further research, the improvement of the application of clustering methods for rail infrastructure is recommended. It is necessary to filter unusual events in the input data. For example, construction or extreme weather events temporarily reduce the capacity and might hinder a clustering of otherwise similar lines. When a representative week has been found, it could be compared with data from different years to analyse yearly variations. Furthermore, it could be tested to scale the traffic load in order to cluster lines with different traffic loads but same traffic patterns.

Up to now, each individual line is often analysed on a very small scale for capacity issues in the railway sector. These evaluations are usually related to a reference day, which, as this article has shown exemplarily, can lead to incorrect results under circumstances. If different lines are better comparable, then one way could be to use similar procedures as in road transport. There, the flow profiles are ordered according to the utilisation in a decreasing manner and the 50th hour is usually taken as the decisive representative. Using similar approaches would allow capacity studies to be carried out in a much more general and perhaps even higher quality manner.

Acknowledgements

This work is partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 2236/1.

References

N. Caceres, L. M. Romero, and F. G. Benitez. Estimating traffic flow profiles according to a relative attractiveness factor. *Procedia - Social and Behavioral Sciences*, 54:1115–1124, 2012. doi: 10.1016/j.sbspro.2012.09.826.

DB Netz AG. Richtlinie 405 - Fahrwegkapazität, 2008.

DB Netz AG. Who we are, 2015a. URL `https://fahrweg.dbnetze.com/fahrweg-en/company/db_netz_ag/aboutus-1394546`.

DB Netz AG. Network Statement 2015, 2015b.

DB Netz AG. Richtlinie 413 - Handbuch für betriebliche Infrastrukturplaner, 2020.

Deutsche Bahn AG. Vom Stellhebel zum Mausklick: Wie die Bahn in ihrem Schienennetz täglich 40.000 Züge steuert. Technical report, 2019. URL `https://www.deutschebahn.com/resource/blob/1295806/b424db51a0b9a4d51226c1e424b17297/2015-07-Themendienst-Stellwerke-data.pdf`.

Deutsche Bahn AG. Trassenportal DB Netz AG (TPN) - Stammdatenaktualisierung, 2020. URL `https://www.dbnetze.com/infrastruktur-de/Kundeninformationen/2015{_}21{_}TPN{_}Stammdatenaktualisierung-1441650`.

Y. Endo and S. Miyamoto. Spherical k-means++ clustering. In *International Conference on Modeling Decisions for Artificial Intelligence*, pages 103–114. Springer, 2015.

M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

European Commission. Commission Regulation (EU) No 1299/2014 of 18 November 2014 on the technical specifications for interoperability relating to the 'infrastructure' subsystem of the rail system in the European Union, 2014.

European Commission. Transforming Europe's Rail System. Technical report, 2020.

Eurostat. Passengers transported (detailed reporting only) - Quarterly data, 2021a. URL https://ec.europa.eu/eurostat/databrowser/view/rail_pa_quartal/default/line?lang=en.

Eurostat. Goods transported (detailed reporting only) - Quarterly data, 2021b. URL https://ec.europa.eu/eurostat/databrowser/view/RAIL_GO_QUARTAL__custom_601063/default/table?lang=en.

Forschungsgesellschaft für Straßen- und Verkehrswesen. Handbuch für die Bemessung von Straßenverkehrsanlagen: HBS, 2015.

F. Liu, Z. Sun, P. Zhang, Q. Peng, and Q. Qiao. Analyzing Capacity Utilization and Travel Patterns of Chinese High-Speed Trains: An Exploratory Data Mining Approach. *Journal of Advanced Transportation*, pages 1–9, 2018. doi: 10.1155/2018/3985302.

A. Økland and N. O. E. Olsson. Punctuality development and delay explanation factors on Norwegian railways in the period 2005–2014. *Public Transport*, 13(1):127–161, 2021. doi: 10.1007/s12469-020-00236-y.

L. Pinkofsky. *Typisierung von Gangliniender Verkehrsstärke und ihre Eignung zur Modellierungder Verkehrsnachfrage*. PhD thesis, TU Braunschweig, 2005.

Statistisches Bundesamt. Betriebsdaten des Schienenverkehrs 2015. Technical report, 2017. URL https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Transport-Verkehr/Unternehmen-Infrastruktur-Fahrzeugbestand/Publikationen/Downloads-Betriebsdaten-Schienenverkehr/betriebsdaten-schienenverkehr-2080210187004.pdf?__blob=publicationFile.

TNS Opinion & Social. *Quality of transport: Report*, volume 422a of *Special Eurobarometer*. Publications Office, Luxembourg, 2014. ISBN 978-92-79-44436-4.

UIC. Code 406 - Capacity, 2004.

UIC. Code 406 - Capacity, 2013.

UNECE. TEM Standards and Recommended Practice, 2002.

W. Weijermars and E. van Berkum. Analyzing highway flow patterns using cluster analysis. In *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005*, pages 831–836. IEEE, 2005. doi: 10.1109/ITSC.2005.1520157.

C. Wen, Z. Li, J. Lessan, L. Fu, P. Huang, and C. Jiang. Statistical investigation on train primary delay based on real records: evidence from Wuhan–Guangzhou HSR. *International Journal of Rail Transportation*, 5(3):170–189, 2017. doi: 10.1080/23248378.2017.1307144.