Evaluation of changes in corridor railway traffic in the Czech Republic during the pandemic year 2020

Michal KUČERA, Zdena DOBEŠOVÁ

Abstract: *Railway traffic was significantly affected by the global Covid-19 pandemic in 2020. The article's main objective is to evaluate the changes in rail transport on rail corridors in the Czech Republic in 2020. The traffic from that year was compared with the values from previous years. The Czech Statistical Office provided the data base for general information about railway transport. This contains details of freight and passenger transport and the performance of that transport for the whole country, as reported by the Ministry of Transport. The presented evaluation of changes in rail traffic is based on traffic parameters obtained from the railway infrastructure administrator of the Railway Infrastructure Administration. The changes were evaluated from three hundred points located on the corridor railways. The first part of the article describes the evaluation of changes in each traffic parameter. In 2020 passenger traffic was characterised by a significant decrease in the total number of trains on most corridors. Cargo trains took up the railway traffic capacity released by the drop in demand for passenger trains, and the number of cargo trains increased as a result. Passenger trains were shortened due to the low demand for transport during the pandemic. This fact was reported in the parameter of average train length. The second part contains a comprehensive assessment based on clustering methods. The clustering enabled the detection of those sections of the rail corridors that had similar changes in rail traffic in all the monitored parameters in 2020.*

Keywords: *railway, corridors, Covid-19, traffic, Czech Republic*

Introduction

The worldwide pandemic of Covid-19 disease noticeably influenced mobility including railway traffic. Reduced population mobility and industrial closures have led to a decline in the need to transport people and freight. The presented study evaluates the change in domestic railway transport in the pandemic year 2020 in the main corridor railway network in the Czech Republic. The initial expectation was about the total traffic decline at all sections of the railway corridors.

Our study aims to investigate the character of changes in Czech railway corridors in detail scale during the pandemic year 2020. The main objective was to analyse traffic changes and to identify sections of the railway network with similar traffic changes. First, an assessment of rail transport was carried out based on the Czech Statistical Office data. Subsequently, a detailed analysis was prepared using data from the Railway Administration.

Literature review

Several studies addressed transport changes during the pandemic year 2020. Pászto, Burian and Macků (2020) evaluate data sources and map applications. The mobility reduction was visible in all types of transport like air, car, railway, maritime transport and personal moves like cycling and walking.

DOI: https://doi.org/10.33542/GC2023-1-03

The global pandemic has also affected air travel. Declining passenger numbers in International intra-EU air passenger transport are visible by the end of 2019. The most significant decrease then occurred in March 2020, from 21.6 million to 0.2 million passengers carried compared to the same month in 2019. At that time, the planes were grounded and air traffic was suspended. Air passenger numbers have been gradually returning to pre-peak levels, but these were not yet reached in 2021 (Borucka et al. 2022).

Maritime transport was affected too. Throughout the lockdown period, part of cargo capacity at key ports in India remained unused. Major 12 ports handled more than half of India's maritime transport, which fell 10.5% to 414 million tonnes between April and November 2020 compared to the previous year (Saxena and Yadav 2022).

In case of railway traffic, a 40-60% reduction in commuters and occasional travellers was reported by performance analysis on six railway lines in northwest Italy (Grechi and Ceron 2022). A similar situation occurred in Croatia, where domestic passenger numbers decreased by 33.6% and international passenger numbers dropped by 57.3% (Jugović et al. 2022). The total number of passengers in Poland declined by 38% between 2019 and 2020 (Ojdana-Kościuszko 2022).

Xin et al. (2021) examined the impact of Covid-19 on daily ridership on urban rail systems in 22 cities in Asia, Europe, and the United States. They reported that most Chinese cities experienced about a 90% reduction in ridership with some variation among different cities. Seoul and Singapore experienced a minor decrease compared to Chinese cities. Chinesse cities recorded the outbreak of Covid-19 earlier in January 2020 than Europe and American cities, where influence on urban rail systems is reported in from March 2020.

There was also a shift in the transport of cargo from road to rail due to the restrictions on truck drivers that were in place in some countries (Germany) (Barczak 2022). In Poland the increasing vaccination rate among the population has encouraged travellers to return to rail transport. This has mitigated the impact of further pandemic waves (Ojdana-Kościuszko 2022).

The pandemic has also affected car transport. According to Apple Mobility Data, traffic started to decline in early 2020 and reached its lowest level in April. Subsequently, it began to grow again. Further developments have always been country-specific, depending on the measures in place and the evolution of the pandemic. The decline in traffic has also led to a reduction in road deaths in most countries. In a few cases, however, the opposite trend was observed, where road clearance allowed traffic speeds to increase and subsequently led to several serious accidents (Great Britain, Australia, Netherland) (Wegman and Katrakazas 2021).

The Google COVID-19 Community Mobility Report (Google 2021) contains data about the decrease of 50-60% in personal mobility reported in the category of Transit Stations for the spring period of 2020 in the Czech Republic. A microstudy of personal behaviour before and during the lockdown also reports a big change in personal mobility (Pászto, Burian and Macků 2021).

The total weight of cargo transport in the Czech Republic and the neighbouring countries decreased, especially in the 2nd and 3rd quarters of 2020. Eurostat (2022) reported that the dip in rail passenger transport performance in 2020 was particularly significant in the second and fourth quarters (-74% and -54% compared with the same quarters in 2019) in whole EU. Despite the slight recovery in 2021, the performance of the EU rail passenger transport was still below the levels observed before the pandemic. The rail situation improved after the initial reaction to the pandemic calmed to some extent. The subsequent reduction in restrictions led to traffic levels increasing nearly to pre-Covid values in some countries.

Firstly, it was necessary to suggest the frame and method for data comparison for our research. The literature review reveals that the most studies assessing changes in traffic due to the Covid-19 pandemic are based on a two-year comparison. Specifically, 2019 and the following year that was affected by the 2020 global pandemic. Two years seems to be a logical comparison. However, the impacts identified may be distorted if there have already been extraordinary changes in 2019 (local decrease or increase). E.g., changes in using Taipei Metro compared data from January to March 2020 against January to March 2019 based on passenger's cards. No older selected months from 2018 or 2017 years were taken into account (Mützel and Scheiner 2022). For railway timetable comparison on the Prague-Pilsen traffic line in the Czech Republic was also taken in the only pre-Covid year 2019 (Surmařová et al. 2022).

The inspiration for considering longer time comparisons was taken from literature about time series processing (Hyndman and Athanasopoulos 2018) where the assessment of trend, seasonal and periodic parts are taken from long time series. Saxena and Yadav (2022) used the ARIMA modelling to predict the revenue lost loss and volume of rail freight by India Railways by Covid effect in from April 2020 to September 2020. They use the monthly data from April 2012 to November 2021. Comparison of nearly ten years' time series with and without Covid era in 2020 estimated the total loss of freight volume at 149.08 million tonnes. India had draconic lockdown with passenger trains traffic prohibited from March 2020 until 30 July 2020. Only parcel trains for the transport of essential goods were operated (Saxena and Yadav 2022).

Our railway traffic study used the median of traffic parameters determined from several pre-Covid years (specifically 2016-2019). This eliminates extraordinary effects that may be caused, for example, by the reconstruction of railway line sections. In the case of a year-on-year increase in traffic, the use of older data may be perceived negatively. However, this impact is reduced by using just the median value.

The railway network in the Czech Republic

The railway network of the Czech Republic consists of 9,377 km of tracks in total (Railway Administration 2021). The network has the highest density in Europe, about 0.12 km/km^2.

The Czech corridors are parts of the European transport corridor system. The arrangement of corridors in the Czech Republic is shown in the map in the middle of the fig. 1. Some corridors are partially simultaneous (Děčín-Praha, Praha-Ostrava). There are four main transit corridors in the Czech Republic (Ministry of Transport of the Czech Republic 2020):

1st transit corridor East – Mediterranean
Děčín – Praha – Pardubice – Česká Třebová – Brno – Břeclav

2nd transit corridor Balt – Jadran
Petrovice u Karviné – Ostrava – Přerov – Břeclav

3rd transit corridor Rhein – Donau
Cheb – Plzeň – Praha – Ostrava; side branch Plzeň – Domažlice

4th transit corridor Stockholm – Dresden
Děčín – Praha – Tábor – Veselí nad Lužnicí – České Budějovice – Horní Dvořiště

Data and methodology

The evaluation of changes could be based on various sources. The official statistical data contains financial data and numbers of passengers and goods (Czech Statistical Office 2022b). The following source data could be technical data from the Railway Administration about the number of trains set. Both sources describe the same situation but in different ways. The direct join of information is not possible due to different character of data and different level of detail. Performance statistical parameters describe the situation more globally for the whole state. Technical data describe the transit in detail on corridor lines.

The change in railway traffic can be calculated from the official statistical data provided by the Czech Statistical Office (2022b) as a public database. It publishes long-term data and updated time series by topics, including a topic Transport, Information and Communication activities. The statistical reporting unit is the quarterly period of each year. The census contains the number of carried passengers (in thousands of passengers), the transport performance (in passenger-kilometres) and the average transport distance (in kilometres). Cargo transport is characterised by the quantity of transported goods, the transport performance (in tonne-kilometres) and the average value of the transport distance in the census by the Czech Statistical Office. Tab. 1 presents the statistical data in mentioned parameters from 2019 to 2021. Passenger traffic decreased significantly by 33% in 2020. The number of passengers slightly increased in 2021, but it is still under level 2019. Cargo traffic went down at the beginning of 2020 but then grew up. At the end of 2020 and 2021, cargo traffic was at the same level as in previous years. Cargo has not been as significantly affected by the pandemic as passenger transport.

Tab. 1. *Values of indicators of passenger and cargo railway transport from 2019 to 2021*

Year	Passenger transport			Cargo transport		
	Passenger transp. (thous.)	Transp. performance (mil. pers-km)	Average transp. dist. (km)	Total goods transp. (thous. tonnes)	Transport performance (mil. tonne-km)	Average transport distance (km)
2019	194 211	11 070	57	98 527	16 070	163
2020	129 142	6 484	49	90 863	15 137	167
2021	133 821	7 219	50	97 216	16 126	166

Czech Statistical Office (2022b)

The source data for detail scale evaluation originated from Railway Administration. The data are traffic values on monitoring points SR70 along the railway network. They are located very densely on all railway tracks. The total number of points is 3,350, defined by the internal infrastructure manager document. The SR70 monitoring points used in this study are limited to the location on rail corridors (only 311 points).

Traffic on the SR70 point is reported in several parameters:

- *type of trains (passenger and cargo trains)*
- *number of trains*
- *total weight of train*
- *the average length of trains*
- *the average weight of train*
- *the average number of axles of the train set*

Parameters were aggregated to year sum for each monitoring point.

The change in traffic was determined based on the difference between the parameter's value in 2020 and the median value from 2016 to 2019 (Formula 1). This calculation was chosen to minimise the impact of drop values caused by some emergencies like reconstructions and accidents on the network. The antecedent drops could significantly affect the evaluated values from the 2020 years. The median well balance the base for comparison. Changes in parameters were set for SR70 corridor points separately for passenger and cargo traffic. They are expressed as a percentage of change.

$$v = \frac{h_{2020} - median(h_{2016}, h_{2017}, h_{2018}, h_{2019})}{median(h_{2016}, h_{2017}, h_{2018}, h_{2019})} \times 100 \quad (1)$$

where

v	*percentage change of parameter [%]*
h_{2020}	*parameter value in 2020*
$h_{2016,2017,2018,2019}$	*parameter values in 2016, 2017, 2018, and 2019*

In the first analysis, the change for each parameter was evaluated using Formula 1. The following analysis contains more comprehensive view of the changes caused by the 2020 pandemic restrictions. The evaluation of the changes was based on cluster analysis, where the main aim was to find sections of railways with similar traffic changes. The process involved testing multiple clustering methods and comparing their results. The k-Means, Self Organising Map (SOM) and hierarchical clustering methods were tested. The k-Means method showed low silhouette scores for individual clusters, therefore the optimal number of clusters was not found. The SOM method resulted in an inappropriately high number of clusters, including single-element clusters. As the most appropriate approach, dimension reduction was first performed using the t-Distributed Stochastic Neighbour Embedding (t-SNE) method, and then clusters were formed using the hierarchical method. The t-SNE method reduces the input parameters to two new components V1 and V2.

The t-SNE algorithm consists of two main stages. First, t-SNE constructs a probability distribution over pairs of high-dimensional objects so that similar objects are assigned a higher probability, while dissimilar points are assigned a lower probability. Second, t-SNE replicates the same probability distribution on lower dimensions iteratively, and until the Kullback-Leibler divergence is minimised.

The t-SNE respects the local variance (Vidiyala 2020). The t-SNE method was used mainly because of its suitability for nonlinear spatial data and its robustness in relation to outliers.

The parameter settings for the t-SNE algorithm must be adapted to the input data, and it is advisable to test several variants to obtain the correct result (van der Maaten and Hinton 2008). One of the most important is the Perplexity parameter, which expresses the strength of the link between local and global aspects of the evaluated data. Its value can be considered an estimate of the predicted near neighbours of each point. In the case of its lower value, there is a greater emphasis on nearby points, resulting in a more compact distribution of points. At the same time, the mutual differentiation of input values is reduced. A lower value is then more appropriate in situations where it is desirable to resolve smaller differences in the input data set (Wattenberg, Viégas and Johnson 2016). For the purpose of this study, the Perplexity parameter was set at 30. This value ensures that an optimal number of clusters were created with a sufficient resolution of the mutual difference between the input data.

The t-SNE method reduced the input parameters (dimensions) to two newly created components; V1 and V2. The t-SNE method was processed in Geoda software (Anseline 2022). Subsequent processing was done in Orange software (University of Ljubljana 2023). The Euclidean metric was used to calculate the distance of objects in the new components. Subsequently, hierarchical clustering was performed. The calculation of the inter-cluster distance (linkage) was done using Ward's method in hierarchical clustering (Ward 1963) (Murtagh and Legendre 2014).

Results

Traffic change in corridor monitoring points

The selection of parameters from the dataset for analysis was based on the evaluation of the level of correlation between them. At the same time, their suitability for analysis was assessed. The level of correlation between parameters in the dataset was determined. The total number of trains and average train length had the lowest correlation (0.8). Therefore, they were used to assess the impact of the pandemic on rail transport in detail. Both parameters were evaluated separately. The correlation of the other parameters exceeded 0.9 or were not assessed as suitable for the analysis, e.g. the average number of axles of the train set. The result of evaluations of the traffic change in corridor monitoring points was depicted on the four maps in fig. 1.

The total number of passenger trains dropped in 61% of monitoring points in 2020 compared to the median of previous years (fig. 1a). The drop was particularly significant on the line from Prague to Ostrava (1st and 3rd corridors). There is dominant traffic in the commercial risk of the carrier. In the case of commercial risk operations, carriers directly cancelled some train connections. The state also orders personal transport on the corridors. The number of train connections contracted with the state must be guaranteed. This ensured the transport of people working in critical state infrastructure such as health, trade, supply and police during the pandemic. Therefore, the train connections reduction was not as high as by commercial carriers. The total number of passenger trains was significantly reduced near the state border. International passenger traffic has been reduced.

Carriers had to shorten trains to minimise losses due to low demand. The average length of passenger trains was shorter on most corridor lines in 2020 than in previous years (fig. 1c). This change was more pronounced near border crossings. This situation might be related to restrictions on travelling abroad, including the closure of border crossings during the pandemic. International trains normally reach longer lengths. Their limitation has a more significant impact on the average train length parameter. As already mentioned, passenger transport is operated at the commercial risk of the carriers on the parts of the corridor network. As a result, cancelling connections was the commercial decision of private carriers. Their absence influenced the evaluated train's average length parameter, especially near the state border. A significant increase of both parameters (the average length and the total number) was identified on the 4th corridor line from Prague to the Austrian border. The increase is mainly due to the gradual modernisation of this corridor, which improves the quality of travel and increases transport capacity. Almost 66% of the total number of evaluated points showed a decrease in 2020 in the parameter of the average length of passenger trains.

We detect that the railway sections with a low change in the total number of passenger trains showed the most significant shortening of train sets.

The total number of cargo trains decreased at 42.3% of all points on the corridor lines in 2020, compared to the median of the previous four years (fig. 1b). An increase was recorded for 57.7% of all points. The value of the total number of cargo trains showed an increase of 820 trains in 2020. The rise in cargo trains might result from capacity freed up by cancelling passenger trains. A decrease was recorded on the 2nd corridor from Břeclav to Ostrava. The decrease was caused by the reduction of industrial production (Czech Statistical Office 2022c) and international transport due to the pandemic (Czech Statistical Office 2022a).

There is an increase in the average length of cargo trains at most points in 2020 (fig. 1d). In order to maintain the efficiency of rail cargo transport, it is necessary to run long trains. This is probably the reason, why there was no significant decrease in average train length. It even increased in 74 % of the monitored points.

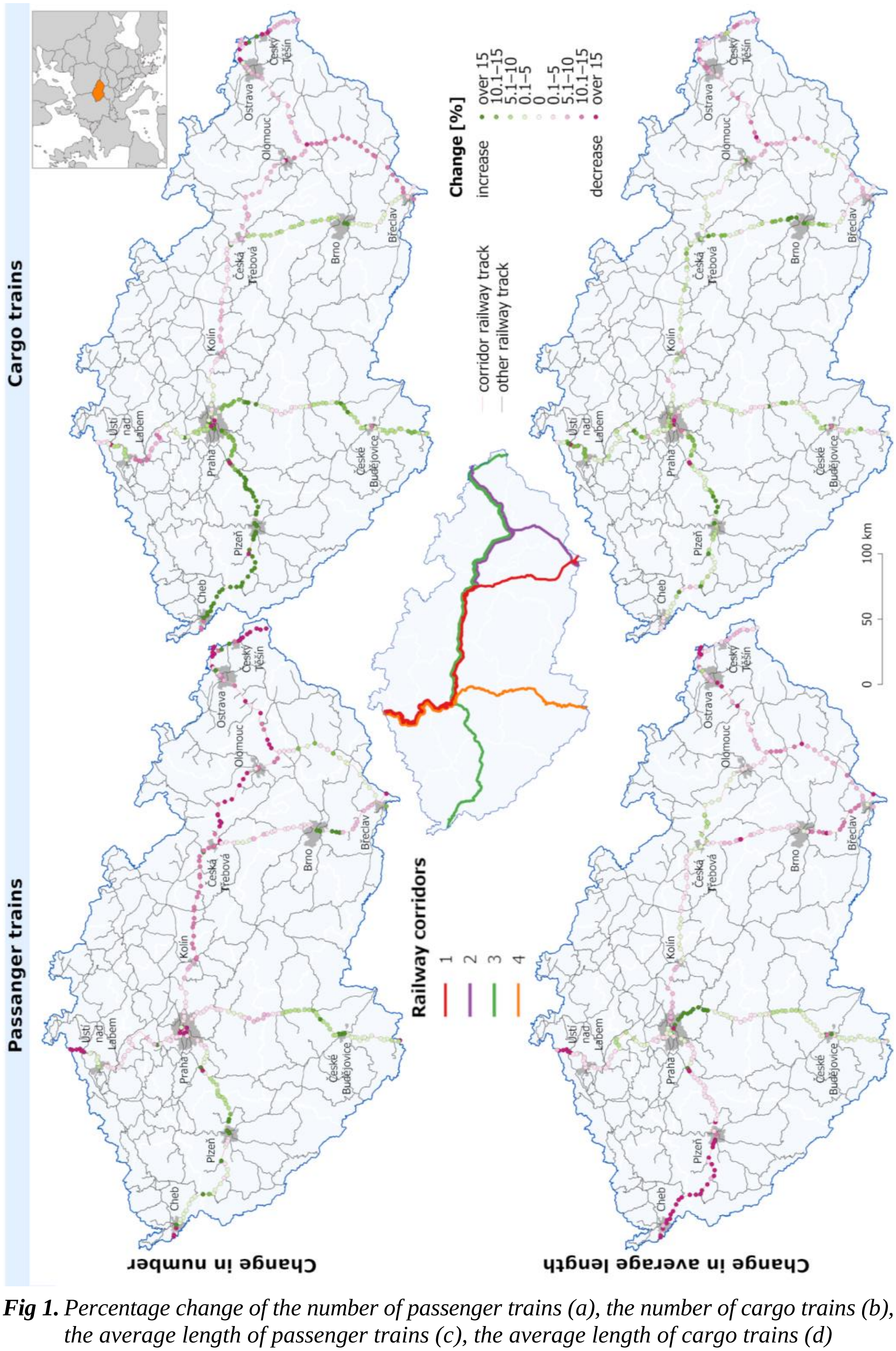

Fig 1. *Percentage change of the number of passenger trains (a), the number of cargo trains (b), the average length of passenger trains (c), the average length of cargo trains (d)*

The results provide information on the spatial location of the observed trends in the assessed attributes that characterise the rail corridors in the pandemic year 2020. The identified changes coincide with the census data of the Czech Statistical Office. The maps provide a comprehensive overview of the distribution of locations where there have been significant changes in the total number of trains or their average length. The results contain information on cargo and passenger transport separately. Change in quarterly census values and the detailed situation on the whole railway network are presented in the article (Kučera and Dobešová 2022).

Hierarchical clustering of railway segments with similar changes

The first analysis presented in the previous chapter evaluates passenger and cargo traffic changes separately. A simultaneous evaluation of the parameters of both traffic types gives a more comprehensive view of the changes caused by the 2020 pandemic restrictions. This second analysis aims to locate corridor railway segments that had similar changes in traffic characteristics. This process will give a complete result based on all input parameters. For this second analysis, clustering methods were used. However, interpreting the results – the clusters, is more difficult, and has to be based on an evaluation of the partial parameter changes in the source data for each cluster.

Transport attributes from the Railway Administration (2022) were used as the input data, with low-level correlation calculated using the Pearson correlation coefficient (Institute of Biostatistics and Analyses FM MU 2022). The results are based on the percentage change in the total number of trains, total train weight and average train length for passenger and cargo trains. Each SR70 significant point was thus associated with data that showed the percentage change in selected transport attributes in 2020, relative to their median value determined from the previous four years (Formula 1).

Six percentage change values thus characterise each SR70 point. Three values describe passenger traffic and three values describe cargo traffic. From the previous separate sub-analyses of passenger and cargo traffic, it appears that the passenger and cargo traffic interacted on some corridor sections and acted as internal factors of traffic change. A decrease in passenger traffic freed up the throughput capacity on the corridor lines and allowed for a possible increase in cargo trains in some segments of lines. However, the freed-up capacity may not have been used in all cases due to the drop in rail orders.

The analysis in this section is based on the t-SNE method and hierarchical clustering. The t-SNE method was used to reduce the dimension of the input data and hierarchical clustering was used to aggregate SR70 points with similar traffic changes. Parameters set for the t-SNE method are shown in tab. 2.

Tab. 2. *Parameters set for the t-SNE method*

Parameter	Value	Parameter	Value	Parameter	Value
Perplexity	30	*Final momentum*	0.8	*Distance function*	Manhattan
Theta	0.5	*Iteration switch momentum*	250	*Transformation*	Standardise
Max. iteration	5 000	*Learning rate*	200	*Momentum*	0.5

Hierarchical clustering produced a dendrogram where SR70 points with similar traffic changes were displayed in the clusters. Six main clusters were identified from the dendrogram. These clusters cover corridor line segments that show similar changes in the parameters of rail traffic intensity. The resulting clusters are shown as a scatter plot in the reduced dimensions determined by the t-SNE method (fig. 2). The values of the constructed components V1 and V2 are plotted on the X and Y axes. It can be seen that the points naturally cluster into six clusters without significant overlap.

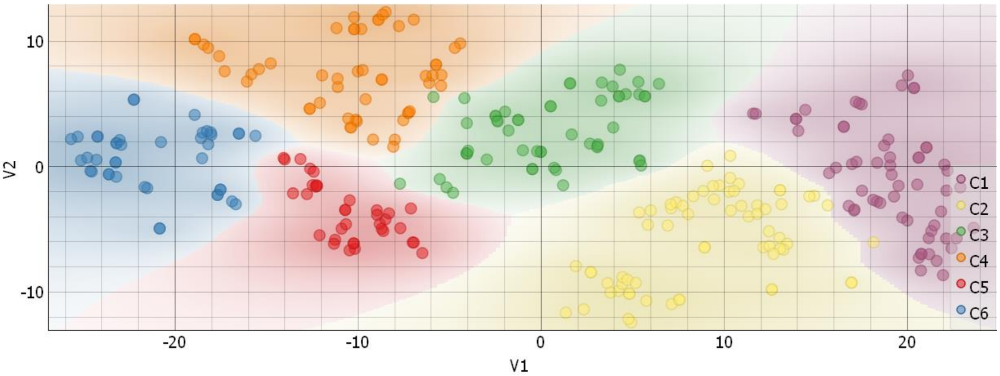

Fig. 2. *Clusters of SR70 points in dimensions V1 and V2 established by the t-SNE method followed by hierarchical clustering*

The quality of the distribution of points into clusters was further evaluated using the Silhouette plot (fig. 3) (Rousseeuw 1987). Most points in the clusters have a positive silhouette index value. The number of points with a negative silhouette index value, i.e. points that are not well classified in the clusters, is only 12. The mean silhouette value is 0.45 and the median is 0.49. Thus, the applied procedure using the t-SNE method and hierarchical clustering can be evaluated as the most appropriate. The resulting clusters were visualised in coordinates of the SR70 traffic corridor points on the map of the Czech Republic (fig. 4).

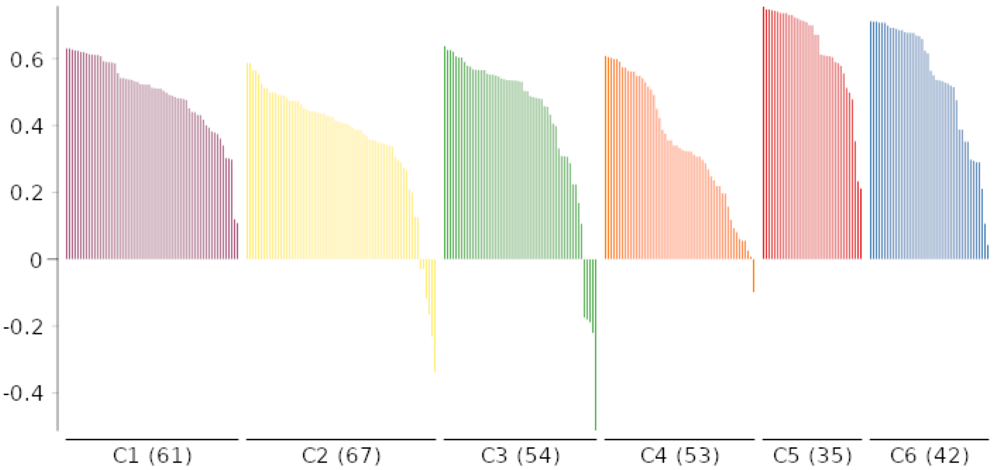

Fig. 3. *Silhouette plot of resulting clusters*

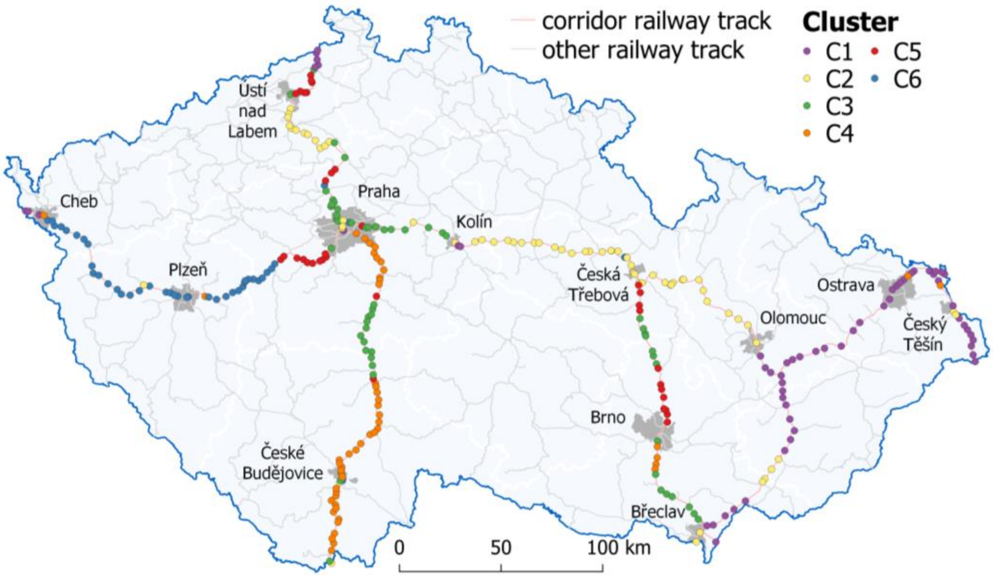

Fig. 4. *Sections of corridor railways with similar changes in passenger and cargo rail traffic*

The generated clusters can be characterised by the various trends of the input traffic parameters (tab. 3). The combinations of changes in traffic parameters highlight six resulting clusters. The clusters were interpreted by comparing the boxplots of the selected cluster and the boxplots of the remaining clusters for each parameter step by step. This comparison was performed in Orange software as that made it possible to rank the relevant parameters for the selected clusters (Orange 2022).

Tab. 3. *Cluster characterisation based on changes in evaluated traffic parameters*

Parameter		Cluster					
		C1	C2	C3	C4	C5	C6
Total count	Passenger	▼▼	▼▼	▼	▲▲	►	▲
	Cargo	▼	▼	▲	▲	▲	▲▲
Average length	Passenger	▼▼	►	▼	▲	▼	▼▼
	Cargo	▼	▲	►	►	▲▲	▲
Total weight	Passenger	▼▼	▼▼	▼	▲▲	▼	▲
	Cargo	▼▼	▼	▲	▲	▲	▲▲

▼▼ significant decrease | ▼ slight decrease | ► unchanged | ▲ slight increase | ▲▲ significant increase

Cluster C1 shows the most significant decreases in all parameters. These occur mainly in the eastern part of the corridor network (part of the 2nd transit corridor) and near the border (fig. 4). Both passenger and cargo traffic were reduced on these parts of the rail system. This may be specifically related to the reduction in cross-border transits during the pandemic in 2020. At the same time, the length of trains was reduced. The shortening trend was not as intense for cargo trains as it was for passenger trains.

Cluster C2 shows a similar negative trend to cluster C1. However, the average length of passenger trains did not change and there was a slight increase in the average length of cargo trains. The points of this cluster mainly cover the middle sections of the rail corridor, from Prague to Ostrava and from Prague to Ústí nad Labem (part of 1st transit corridor).

A positive development in cargo transport is characteristic of the C3 cluster, which saw an increase in the total number of trains. However, the total weight of cargo trains decreased slightly, and the average train length remained the same. Passenger traffic shows a slight decrease in all parameters.

The C4 Cluster covers lines that had a significant increase in the total number of passenger trains and a slight increase in cargo trains. The total weight of trains shows a similar trend. The points are mainly on the line from Prague via České Budějovice to the Austrian border (south part of 4th transit corridor). The positive development can be linked to the modernisation of this corridor line, which has increased in transport capacity and speed. The average train length has increased for passenger transport, while it has not changed for cargo transport.

A positive trend in cargo traffic and lack of change in passenger traffic is characteristic of the C5 cluster. These points cover shorter line sections, especially near the capital city of Prague and on the line from Česká Třebová to Brno (part of the 1st transit corridor). There was a significant increase in the average length of cargo trains and a slight increase in their total weight. This development can be mainly linked to the increase in their total number. Passenger trains, on the other hand, were shortened, decreasing their total weight.

In addition to cluster C4, a positive development in both types of traffic was also recorded for cluster C6. However, in comparison to cluster C4, cluster C6 shows a more significant increase in the total number of cargo trains and a slight increase in passenger trains. Cluster C4 shows the opposite trends. The average length of passenger trains shows a significant drop. Thus, shorter trains were dispatched, but in higher numbers. The points in this cluster only occur on the western corridor line from Prague to Cheb (4th transit corridor), and they cover almost its entire length.

Railway junctions represented by distinct points often showed a different cluster affiliation than the remains part of the corridor line on which they are located. These are, e.g. the Brno, Kolín, České Budějovice, Plzeň, and Cheb junctions (fig. 4). This significant difference is caused by the different nature of traffic at the railway line junctions, as it is here where the lines branch into other national and regional lines. The junctions also often function as transhipment points for cargo traffic. This different nature of traffic at railway junctions is a possible reason for including a single point on a corridor in a different cluster to other points located in the same corridor section.

Discussion

The novelty and contribution of the presented evaluation is the newly proposed procedure for calculating changes in railway traffic. For each SR70 point on the corridors, a simple comparison of the 2020 traffic parameter value with that of the previous year was offered initially. However, rail traffic was greatly affected by engineering work on the sub-sections and the resulting technical closures and traffic redirections. This simple difference in the calculation comparing 2019 and 2020 would not provide a correct expression of change. Therefore, a new procedure was implemented. The change was calculated as the percentage difference between the 2020 values and the median of the four previous years. Formula 1 describes the calculation used, taking the values from 2016 to 2019. This procedure is particularly suitable for evaluating a specific mode of transport, such as rail traffic, when examining sub-sections of the rail network.

As mentioned in the introduction, several studies only compare the data from 2019 with those from 2020 (Mützel and Scheiner 2022, Surmařová et al. 2022, Eurostat 2022). It could be a big disadvantage and weakness of the presented results of other studies concerning pandemic influence in 2020. The base line for comparison could be a source of biasing data. We agree with Kačírková (2022), who presented the limited time base in the case of COVID-19 Community Mobility Reports. The baseline days of COVID-19 Community Mobility Reports represented a normal value for that day of the week. They were the median value over the five weeks from January 3rd to February 6th 2020 (Google 2021). Kačírková mentioned that comparing the individual values of the days for the whole year to the short five-weeks winter period might show an exaggerated influence of weather, holidays and seasons. People's mobility is also influenced by weather and season (tourism in summer). Only the right way is a comparison of the same year seasons (summer with summer etc.). On the opposite side, our methodology uses the median of the previous four-years as a base. It is more sophisticated than a simple one-year or short-time comparison. Evaluations also often suffer from the fact that they are sample data that may not be representative of the entire study problem. Surmařová (2022) used timetables and a questionnaire for a small group of 27 passengers that had limited representativeness. Using trains in Israel was also monitored by two surveys completed by 237 participants and 149 participants respectively, in 7-month span. Questionnaire surveys revealed reasons for the decision to use or not to use train transport (Elias and Zatmeh-Kanj 2021). COVID-19 Community Mobility Reports have another limitation. It represents only users that use the Google application and agree with the monitoring of movement. On the contrary, our presented research processed complete objective traffic data on monitoring SR70 points. The results have valuable objective testimonial capability.

The data analysis also reveals findings about the capabilities and limits of commercial and state carriers. State carriers guarantee basic passenger transport according to contracts and are a certain guarantee of passenger transport operations. In contrast, private carriers can respond more flexibly to sudden situations like pandemics. The data on the Prague-Ostrava corridor reported flexible shortening of train sets where most private carriers operate. This finding is also important information that can guide the state, policymakers and its government including regional government in future decisions about the structure of carriers, where a heterogeneous environment is more resilient in sudden situations. The attraction of travellers and good transport from roads to railways is an important task to make Earth sustainable.

Conclusion

The article presents two approaches to a detailed evaluation of railway traffic changes on corridor lines in the Czech Republic. The first approach evaluates passenger and cargo traffic separately in two parameters, namely the total number of trains and the average train length (fig. 1). Other traffic parameters such as weight were not evaluated due to their high correlation with the parameter of the total number of trains at traffic points SR70. A significant decrease of more than 15% was recorded in the number of passenger trains on the corridor from Prague to Ostrava. It is also remarkable that the significant decrease in the average length of passenger trains was caused by the requirement to reduce costs because of the unused capacity of passenger trains. The increase in traffic during the pandemic year is due to the modernisation of railway tracks and new contracts drawn up between the state and carriers in 2019 as preparation for higher passenger transport volumes in subsequent years. Cargo traffic was not so significantly affected by the pandemic restrictions, and an increase in the number of cargo trains was even detected on some sections of the line. A decrease was recorded near the state borders.

The second comprehensive analysis found similar corridor railway segments. A hierarchical clustering method over reduced dimensions (t-SNE method) was used for this second analysis. The resulting six clusters; C1 to C6 on the rail segments of the corridor were created by varying six traffic parameters in the data source (fig. 4). The clusters represent different combinations of positive and negative changes in the traffic parameters (tab. 3). Only the fourth transit corridor in the eastern part of the Czech Republic had a significant or slight decrease in all traffic parameters, and it was therefore placed in the C1 cluster. Cluster C3 shows a combination of slight decreases and slight increases in parameters. Cluster C6 shows an unusual situation where sections of the corridors have experienced an increase in freight parameter values and, in contrast, a decrease in the average length of passenger trains.

In conclusion, it can be stated that the pandemic led to not only a decrease in rail traffic, but also an increase in volume on some corridor sections and in some parameters. Partial increases may be due, among other factors, to the interaction between passenger and cargo traffic on corridor lines limited by technical capacity. On the one hand, the cargo traffic is more influenced by temporary customer orders, the volume of which can fluctuate over time. Passenger traffic, on the other hand, is more tightly determined by government contracts with the state carrier Czech Railway. So, the latter is more stable in the parameter of the number of trains. It was only private carriers that partially influenced the dynamic of passenger traffic. The private carriers could react to pandemic restrictions more flexibly in consideration of their economic effectiveness.

The contribution of this study is the methodology to reveal the impact of the abnormal situation on the intensity of rail traffic. This methodology could be applied for further research. The knowledge and experience could be very useful for future cases. The analysis of traffic intensity in the time of the Covid pandemic is part of a comprehensive assessment of the development of railway transport in the Czech Republic. The main objective is to identify locations with positive and negative developments. Based on this knowledge, it would be possible to target resources for development and identify weaknesses in rail infrastructure. The study used data information that is primarily intended for infrastructure charging purposes only. This provides interesting information about the traffic changes which will be provided to the railway infrastructure manager. The results are delivered to Railway Infrastructure Administration, state organization, for the internal purpose and future utilization.

References

ANSELINE, L. 2022: An Introduction to Spatial Data Science. *GeoDa*. Retrieved from: https://geodacenter.github.io/.

BARCZAK, A. 2022: Impact of the Covid-19 Pandemic on Rail Freight Transport in Selected EU Countries. *Scientific Journals of the Maritime*, 72, 144.

BORUCKA, A., PARCZEWSKI, R., KOZŁOWSKI, E., ŚWIDERSKI, A. 2022: Evaluation of air traffic in the context of the Covid-19 pandemic. *Archives of Transport*, 64(4), 45-57. DOI: https://doi.org/10.5604/01.3001.0016.1048.

CZECH STATISTICAL OFFICE 2022a: *Public database: Cross-border movements of goods*. Prague (Czech Statistical Office). [cit. 2022-02-01]. Retrieved from: https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=statistiky#.

CZECH STATISTICAL OFFICE 2022b: *Public database: Goods and passenger transport, transport outputs*. Prague (Czech Statistical Office). [cit. 2022-02-01]. Retrieved from: https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=statistiky#katalog=31028.

CZECH STATISTICAL OFFICE 2022c: *Public database: Sales from industrial activity: Data adjusted for seasonal and calendar effects*. Prague (Czech Statistical Office). [cit. 2022-02-01]. Retrieved from: https://vdb.czso.cz/vdbvo2/faces/index.jsf?page=statistiky&&katalog=30835&&akt=3658.

GOOGLE 2021: *COVID-19 Community Mobile Report*. Mountain View (Google). [cit. 2021-05-01]. Retrieved from: https://www.google.com/covid19/mobility/.

GRECHI, D., CERON, M. 2022: Covid-19 Lightening the Load Factor in Railway Transport: Performance Analysis in the North-West Area of Milan. *Research in Transportation Business and Management*, 43, 100739. DOI: https://doi.org/10.1016/j.rtbm.2021.100739.

ELIAS, W., ZATMEH-KANJ, S. 2021: Extent to which COVID-19 will affect future use of the train in Israel. *Transp. Policy*, 110, 215-224. DOI: https://doi.org/10.1016/j.tranpol.2021.06.008.

EUROSTAT 2022: *EU rail passenger transport: partial recovery in 2021*. Luxembourg (EUROSTAT). [cit. 2023-03-11]. Retrieved from: https://europa.eu/!JNN67v.

HYNDMAN, R. J., ATHANASOPOULOS, G. 2018: *Forecasting: Principles and Practice, 2nd ed*. Melbourne (Monash University).

INSTITUTE OF BIOSTATISTICS AND ANALYSES FM MU 2022: *Pearson Correlation Coefficient. Applied analysis of clinical and biological data*. Brno (Faculty of Medicine, Masaryk University). [cit. 2022-06-12]. Retrieved from: https://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat--biostatistika-pro-matematickou-biologii--zaklady-korelacni-analyzy--pearsonuv-korelacni-koeficient.

JUGOVIĆ, A., AKSENTIJEVIĆ, D., BUDIĆ, T., OBLAK, R. 2022: Impact of COVID-19 Pandemic on Passenger Mobility in National and International Railway Traffic of the Republic of Croatia. *Pomorstvo*, 36(1): 147-54. Rijeka (University of Rijeka). DOI: https://doi.org/10.31217/p.36.1.17.

KAČÍRKOVÁ. T. 2022: *Impact of the COVID-19 pandemic on international mobility in Europe [Diploma thesis]*. Olomouc (Palacký University).

KUČERA, M., DOBEŠOVÁ Z. 2022: Railway Traffic in 2020. In Růžičková, K., Horák, J., Růžička, J. eds. *Symposium GIS Ostrava 2022: Smart City - vize a realita*. Ostrava (Technical University of Ostrava). p. 12. DOI: https://doi.org/10.31490/9788024846071-161.

MINISTRY OF TRANSPORT OF THE CZECH REPUBLIC 2020: *Transit Railway Corridors*. Prague (Ministry of transport). [cit. 2022-05-01] Retrieved from: https://www.mdcr.cz/Dokumenty/Drazni-doprava/Zeleznicni-infrastruktura/Tranzitni-zeleznicni-koridory.

MURTAGH, F., LEGENDRE, P. 2014: Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31(3), 274-295.

MÜTZEL, C. M., SCHEINER, J. 2022: Investigating spatio-temporal mobility patterns and changes in metro usage under the impact of COVID-19 using Taipei Metro smart card data. *Public Transport*, 14(2), 343-366. DOI: https://doi.org/10.1007/s12469-021-00280-2.

OJDANA-KOŚCIUSZKO, M. 2022: The impact of the covid-19 pandemic on the rail transport network in Poland. *Zeszyty Naukowe UPH seria Administracja i Zarządzanie*, 58(131), 5-12.

ORANGE 2022: *Explaining Clusters, Introduction to Data Science Series*. Ljubljana (University of Ljubljana). [cit. 2022-01-10]. Retrieved from: https://youtu.be/3SkjU2eBzNY.

PÁSZTO, V., BURIAN, J., MACKŮ, K. 2020: COVID-19 Data Sources: Evaluation of Map Applications and Analysis of Behavior Changes in Europe's Population. *Geografie*, 125(2), 171-209. DOI: https://doi.org/10.37040/geografie2020125020171.

PÁSZTO, V., BURIAN J., MACKŮ, K. 2021: Changing Mobility Lifestyle. *International Journal of E-Planning Research*, 10(2), 66-79. DOI: http://doi.org/10.4018/IJEPR.20210401.oa6.

RAILWAY ADMINISTRATION 2021: *Statistical Yearbook 2020*. Prague (Railway Administration).

RAILWAY ADMINISTRATION 2022: *Railway traffic*. Prague (Railway Administration).

ROUSSEEUW, P. 1987: Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.

SAXENA, A., YADAV, A. K. 2022: Examining the Effect of COVID-19 on rail freight volume and revenue using the ARIMA forecasting model and assessing the resilience of Indian railways during the pandemic. *Innovative Infrastructure Solutions*, 7, 348. DOI: https://doi.org/10.1007/s41062-022-00929-2.

SURMAŘOVÁ, S., VRÁNA, M., ILÍK, J. 2022: Travel behaviour changes during the pandemic: Prague-Pilsen rail case study. *Review of Economic Perspectives*, 22(4), 317-332. DOI: https://doi.org/doi:10.2478/revecp-2022-0015.

UNIVERSITY OF LJUBLJANA 2023: *Orange. Data Mining*. Ljubljana (University of Ljubljana). [cit. 2022-09-12]. Retrieved from: https://orangedatamining.com/.

VAN DER MAATEN, L., HINTON, G. 2008: Visualizing Data Using T-SNE. *Journal of Machine Learning Research*, 9, 2579-2605.

VIDIYALA, R. 2020: What, Why and How of t-SNE. *Towards Data Science*. Retrieved from: https://towardsdatascience.com/what-why-and-how-of-t-sne-1f78d13e224d.

WARD, J. H. 1963: Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301): 236-244.

WATTENBERG, M., VIÉGAS, F., JOHNSON, I. 2016: How to Use T-SNE Effectively. *Distill*. DOI: http://doi.org/10.23915/distill.00002.

WEGMAN, F., KATRAKAZAS, C. 2021: Did the COVID-19 pandemic influence traffic fatalities in 2020? A presentation of first findings. *IATSS Research*, 45(4), 469-484. DOI: https://doi.org/10.1016/j.iatssr.2021.11.005.

XIN, M., SHALABY, A., FENG, S., ZHAO, H. 2021: Impacts of COVID-19 on urban rail transit ridership using the Synthetic Control Method. *Transp. Policy*, 111, 1-16. DOI: https://doi.org/10.1016/j.tranpol.2021.07.006.

Acknowledgement: This paper was created within the project of the Erasmus+ Programme of the European Union, Jean Monnet Module (Project No. 620791-EPP-1-2020-1-CZ-EPPJMO-MODULE, and the project "Analysis, modelling, and visualisation of spatial phenomena by geoinformation technologies II" (IGA_PrF_2023_017 with the support of the Internal Grant Agency of Palacký University Olomouc).

Authors´ affiliation

Michal Kučera
Department of Geoinformatics
Faculty of Science, Palacký University on Olomouc
Railway Administration (Správa železnic)
Jeremenkova 103/23, 779 00 Olomouc
Czech Republic
kuceramic@spravazeleznic.cz

Zdena Dobešová
Department of Geoinformatics
Faculty of Science, Palacký University on Olomouc
17. listopadu 50, 771 46, Olomouc
Czech Republic
zdena.dobesova@upol.cz