



Cluster Analysis as a Basis for the Development of an Application Assessing the Reliability of Transport Infrastructure

Piotr Trojanowski¹ · Jozef Husár² · Stella Hrehová² · Lucia Knapčíková¹

Accepted: 15 April 2024
© The Author(s) 2024

Abstract

Increased demand for transport services and increased mobility of citizens can lead to a reduction in the level of reliability of transport systems. This in turn increases the demand for ways to assess the reliability of road infrastructure by both transport companies and individual users. The article presents the substantive basis of an application used to determine the reliability of transport infrastructure. Our approach was based on grouping information into clusters – based on the author’s proprietary clustering method. Its basis is the detailed analysis of the road infrastructure in terms of errors occurring on it, divided into conceptual, design and operational errors. The methodology consists of three stages of clustering (1) creating a database of sections with assigned errors, (2) determining the initial clusters, (3) creating a final database of clusters, and then assessing the reliability of the road infrastructure of the transport system on their basis. The assumption is that the application will remain open-ended – i.e. the database will be developed by users. The proposed methodology was verified on the example of the selected route in Poland (between Kalisz and Szczecin). Based on the results obtained during the experiment on the selected route, errors in the road infrastructure were determined. This, in turn, allowed us to find that there are a number of errors in the road infrastructure, including errors with a high frequency of occurrence i.e. the so-called permanent errors, which further confirms the need to create an application to assess the reliability of the road infrastructure.

Keywords Road network · Clustering · Reliability of road infrastructure · Transport system · Applications

1 Introduction

The transport system consists of several areas, including: material and technical related to road infrastructure and means of transport, economic and organizational, institutional and legal, spatial and functional, and others. Therefore, determining a universal measure of reliability for the

entire system is a very complex challenge that is threatened with numerous disruptions.

Such disruptions include, for example, road accidents and collisions [1], the occurrence of which is stochastic in nature [2–4]. It is not possible to accurately determine the place, time and type of an accident or road collision. However, it is possible to indicate places within the road infrastructure, where there is an increased risk of adverse events [5]. These places, depending on the research methodology, are defined as high-risk zones [6], danger zones [7], black sections, black spots [8–10], etc.

There were many methods developed to determine the location of individual elements of the transport system [11], including black spots, whether on urban roads [12] or extra-urban roads [13]. The database on the basis of which the research is conducted is the main disadvantage of these solutions. It must contain information on adverse events that occurred in the area covered by the analysis. This makes it possible to designate hazardous zones only after a certain period of road use and following a certain number of accidents and collisions [14]. In addition, with data

✉ Jozef Husár
jozef.husar@tuke.sk

Piotr Trojanowski
piotr.trojanowski@zut.edu.pl

Stella Hrehová
stella.hrehova@tuke.sk

Lucia Knapčíková
lucia.knappcikova@tuke.sk

¹ West Pomeranian University of Technology in Szczecin, 17, Piastów Ave., 70-310 Szczecin, Poland

² Faculty of Manufacturing Technologies with a seat in Prešov, Department of Industrial Engineering and Informatics, The Technical University of Košice, Košice, Slovakia

only available on the number of participants in accidents and collisions at a specific time without accurate reports on the nature and type of adverse events, we remain at risk of erroneous designation of such zones. This occurs when, as a result of, for example, a weather anomaly, several accidents occur on 1 day or many vehicles participate in a single road event [15], which in turn will lead to exceeding the threshold number of road events and designating a high-risk zone at this place. Another problem that is often not taken into account is the complexity of factors affecting the occurrence of an adverse event, e.g. related to the reliability of means of transport [16].

In the result of such situations, the risk of non-performance of the transport task increases. Therefore, improving transport safety is one of the most important objectives of the EU transport policy, which is aimed at meeting the expectations of modern society, related to increasing mobility and improving the quality of life, especially following the COVID-19 pandemic [17].

Making decisions on shaping the transport system requires information support [18–20]. Correct utilization of collected road safety information [21] can support the design and subsequent operation of the transport system, ensuring a high level of its reliability [22–25].

In connection with the above, there is also the need to revise the existing state of knowledge and develop a methodology for designing and analyzing the operational reliability of the transport system, which will create the possibility of a systemic review of safety issues [26].

Road Infrastructure Safety Management (RISM) [27–29], which determines the operational reliability of the transport system [30] by separating dangerous road sections, is therefore of strategic importance.

Directive 2008/96/EC of the European Commission is a document that sets out the directions for action consisting in uniform and integrated management of road infrastructure safety in the EU [31]. The abovementioned Directive identifies four tools recommended for use in RIS Management procedures:

1. Assessment of the impact of the planned road on road safety in the network of cooperating roads (Road Safety Impact Assessment),
2. Identification of hazardous sections (Safety Ranking Audit),
3. Identification of road sections with a high accident concentration (AC Classification) and road network safety ranking (NS Ranking),
4. Road infrastructure safety inspection (Safety Inspections).

The requirements contained in the Directive form the basis for road safety tests, which include:

1. Selection of methods for identifying dangerous sections on the existing road network (black spots and black sections) [32].
2. Identification of dangerous spots and/or sections.
3. Risk assessment for selected infrastructure segments [33].
4. Ordering of dataset on dangerous segments of infrastructure by grouping them.
5. Analyzing the results of the ordering performed.

The methodology of operational reliability analysis of transport infrastructure is of particular significance, as it allows for conducting research in the scope of irregularities (errors) occurring in the infrastructure and grouping individual sections into clusters [34–36].

There are various known methods for analyzing and processing information in the field of transport systems [37]. The effectiveness of cognitive processes aimed at studying possible events and phenomena that are able to disrupt the course of the transport process, including, for example, in cargo processes in freight transport, is increasing significantly [38]. Every attempt to broaden the information [39] is based on the procedures of data collection and processing in order to organize (structure) them according to certain parameters, e.g. the level of threats to traffic flow, probability of road accidents, etc. In order to organize the collected data, they are clustered [40–43] or classified [44–47], which allows to take into account the requirements for information processing needed for the subsequent design of the application.

The presented considerations lead to the development of a concept that will allow for the assessment of the reliability of road infrastructure not only for the use of the authorities managing it, but also for transport companies and individuals. This concept constitutes the substantive basis for the development of an application that will enable the assessment of a selected route (divided into sections) in terms of its reliability.

The next part of the article consists of the methodology, which includes an excerpt from the road infrastructure error database and a three-step procedure scheme. The next part, Results and Discussion, contains a verification of the methodology presented on the basis of the conducted field experiment. The last part of the paper contains conclusions drawn from the conducted research, limitations of the proposed methodology and directions for further research.

2 Methodology

The concept of the application for assessing the reliability of road infrastructure was based on an open - editable database. This database should contain the set of routes, along with an

analysis of errors occurring on them, creating an image of the entire transport system in material, technical and spatial, and functional terms [48, 49].

The basic version of the database should be created empirically in a limited area and then gradually expanded by the users of the application. Errors occurring on the route are divided into three categories [50]:

- conceptual errors,
- design errors,
- operating errors

The list of errors is also open, i.e. it is not definitively defined. This is related to the changes that take place in the functioning of transport systems, both from the infrastructure side, as well as issues related to the means of transport, organization and management, etc. Table 1 presents a fragment of the database of identified errors [51].

The creation of a database with routes, and then with errors assigned to them, constitutes the basis for the further clustering process of infrastructure sections. The clustering was developed on the basis of three stages [48].

2.1 Stage I – Creating a database of sections with errors assigned to them

This stage starts with selection of the route T , based on the route database. This route should then be divided into sections of equal length O_m .

$$T_p = \{O_m : m = 1, 2, \dots, M\}, \tag{1}$$

where:

- T_p the analyzed route,
- O_m the section of the route where $m = 1, \dots, M$,
- M number of route sections

Due to the significant differences between the individual classes of roads, we proposed to divide them into motorways along with express roads and other roads of a lower class. This is due to a significant difference in both the types of errors occurring and the average number of errors occurring on different road classes.

The subsequent stages of clustering are carried out independently for the two distinguished road classes, however, they are carried out according to the same principles. After

Table 1 Fragment of the database on road infrastructure errors

Error category	Reference number	Type of error
Conceptual error (x_k)	x_{k1}	No shoulder/emergency lanes
	x_{k2}	Vegetation in the immediate vicinity of the road
	x_{k3}	No bypasses/transit roads passing through built-up areas
	x_{k4}	Excessive number of merging roads
	x_{k5}	Too narrow lanes
	x_{k6}	Generators of traffic located in the immediate vicinity of transit roads
	x_{kn}	...
Operating error (x_e)	x_{e1}	Poor condition of the road surface, e.g.: ruts, bumps, cracks, incorrectly made patches, cavities on the pavement, etc.
	x_{e2}	Vegetation overlapping the roadside
	x_{e3}	Damaged/destroyed elements of technical infrastructure
	x_{e4}	Damaged/destroyed horizontal markings
	x_{e5}	Damaged/destroyed vertical traffic signs
	x_{e6}	Road soiling – sand, mud, leaves, etc.
	x_{en}	...
Design error (x_p)	x_{p1}	Wrong road geometry
	x_{p2}	Poor visibility on arcs
	x_{p3}	Insufficient turning radius
	x_{p4}	The proximity of technical infrastructure elements, among others, elements of bridges, trestle bridges, high-voltage and telecommunications lines, sewage, gas heating and water supply pipes, drainage devices, etc.
	x_{p5}	Poor visibility at nodes/intersections – obstacles
	x_{p6}	Cross-section of the Road that is not adapted to the function it performs
	x_{pn}	...

Source: Authors' own elaboration

the division into classes, errors are determined x on the selected route, which forms the basis for creating databases of sections with assigned errors.

$$XO = XP \cup XK \cup XE, \tag{2}$$

where:

- XO the set of errors of the route section,
- XP the set of design errors of the route section,
- XK collection of conceptual errors of the route section,
- XE the set of operating errors of the route section.

$$XP = \{xp_i : i = 1, 2, \dots, LB_{xp}\}, \tag{3}$$

where:

- xp_i design error, i -th, where $i = 1, \dots, LB_{xp}$,
- $XK = \{xk_j : j = 1, 2, \dots, LB_{xk}\}, \tag{4}$

where:

- xk_j conceptual error, j -th, where $j = 1, \dots, LB_{xk}$,
- $XE = \{xe_k : k = 1, 2, \dots, LB_{xe}\}, \tag{5}$

where:

- xe_k operating error, k -th, where $k = 1, \dots, LB_{xe}$.

2.2 Stage II – Designation of preliminary clusters

The second stage of clustering consists in determining preliminary clusters, which are later transformed into the final clusters. Each cluster (both preliminary and final) consists of an initiating section O_j , and sections attached to it. Its construction is carried out as follows:

1. The number of errors that occur on each section of the route is calculated, and then the frequency of identified errors is determined according to the following formula

$$C_s = \frac{LB_s}{M}, \tag{6}$$

where:

- C_s the frequency of error occurrence within the route [–],
- LB_s the number of errors of the same type occurring on all analyzed sections,
- M number of route sections.

This frequency is determined for every group of errors.

2. After performing the calculations, the condition is checked whether there are errors with frequency on a given section of the route that is greater than or equal to 0.8. If so, the set of such errors (single or multi-element) is defined as a set of constant errors. Constant errors are not taken into account in the subsequent part of the clustering,
3. The next stage is to create a list of hierarchically arranged sections of the route in relation to the number of errors on the respective section, excluding constant errors. The section with the highest number of errors occupies the highest place on the list. If there is more than one section with the same number of errors, the section placed higher on the list is the one with the higher sum of the error frequencies that were identified on it.

$$\sum_{x=1}^{LB_s} C_s = \sum_{xp_i=1}^{LB_{xp}} C_{xpi} + \sum_{xk_j=1}^{LB_{xk}} C_{xkj} + \sum_{xe_k=1}^{LB_{xe}} C_{xek}, \tag{7}$$

where:

- C_{xpi} the frequency of design error occurrence within the route,
- C_{xkj} the frequency of conceptual error occurrence within the route,
- C_{xek} the frequency of operating error occurrence within the route,
- LB_{xp} number of design errors of the same type identified within the route,
- LB_{xk} number of conceptual errors of the same type identified within the route,
- LB_{xe} number of operating errors of the same type identified within the route.

4. After the list is created, a section (the highest ranking) is selected O_j , which is a potential section initiating the creation of a cluster. After selecting the initiating section, there is a procedure for appending subsequent sections of the route to the newly formed cluster. For each section of the route, the coefficient of error overlap (WPB) and the level of correlation (PK) in relation to the initiating section are calculated.

The coefficient of error overlap is the ratio of the number of errors occurring simultaneously on the appended section and on the initiating section to the number of all errors of the appended section.

The method of determination WPB is presented in Table 2. The number of errors occurring simultane-

Table 2 WPB Calculation example

No.	Type of section	Errors								
		x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₈	x ₉
1.	Initiating section	1	1	1	1	1	1	1	1	1
2.	Appended section	1	1			1	1		1	

Source: Authors' own elaboration

ously on the appended and the initiating section (marked in green) is 4, while the total number of errors on the appended section is 5 (the error that does not occur in the case of the initiating section is marked in red). Thus, the WPB of the appended section is $\frac{4}{5} = 0,8$.

The level of correlation PK is determined on the basis of the following formula:

$$PK = \frac{LB_d}{LB_l} \tag{8}$$

If the section meets the criteria of $WPB \geq 0,75$ and $PK \geq 0,5$ it is selected as a section capable of co-creating a cluster.

- After analyzing all sections of the route, a decision is made to create a preliminary cluster. The condition for its creation is the addition of at least two sections of the route to the initiating section. If this condition is not met, a two-element cluster shall be dissolved. The initiating section goes to the cluster of dispersed sections, while the attached section returns to the database of route sections. If the preliminary cluster creation condition is met, the preliminary cluster is added to the preliminary clusters database.

As long as there are more than two sections remaining in the database, the procedure is repeated. However, if there are no more sections in the database or their number is less than three, the process ends and any remaining sections are transferred to the cluster of dispersed sections.

2.3 Stage III – Final cluster database

The third stage of clustering aims to create the final database of clusters. To this end, the following steps are carried out:

- A list of sections in the preliminary clusters is created, excluding initiating sections and sections from the dispersed sections cluster.
- Then the process of relocation of sections is carried out. It consists in rechecking all the sections from the list in relation to the initiating sections according to the altered

criterion PK, while maintaining the previously defined criterion concerning $WPB \geq 0,75$. In the relocation process, for each sections from the list, PK is calculated in relation to all initiating sections. The section joins the initiating section for which its PK is the highest, while meeting the condition of $WPB \geq 0,75$.

- After analyzing all sections from the list, the list of newly created clusters is created. Each of the clusters is checked according to the condition from the second stage (the creation of the cluster is conditioned by adding at least two sections of the route to the initiating section). If this condition is met, the cluster is placed in the database of preliminary clusters, while if the condition is not met, the remaining sections are transferred to the cluster of dispersed sections.
- After analyzing all newly created clusters, the final cluster database is created.

Determination of the reliability class for individual clusters is carried out according to the following principles. The first step is to create a list of initiating sections of the considered clusters. Then, for each section, the reliability level is calculated according to the following formula:

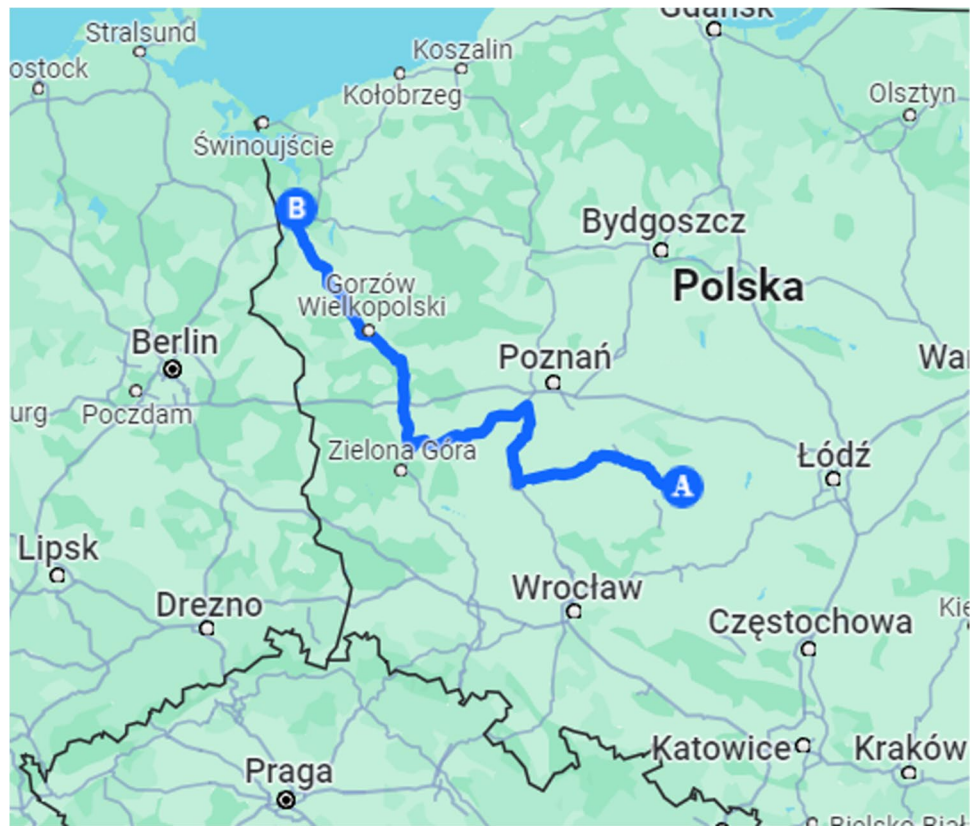
$$PN_{OI} = 1 - \frac{LB_s}{\sum x} \tag{9}$$

where:

- PN_{OI} reliability level of the calculated initiating section,
- LB_s number of errors identified within the route section,
- $\sum x$ the sum of all conceptual, design and operational errors.

After calculating the reliability level of the initiating segment, its value is assigned to the entire cluster it represented. Then, the cluster reliability level value is compared with the designated confidence level interval, on the basis of which a reliability class is assigned to the cluster. Once the process of assigning the reliability levels covers all clusters from the database, the final step is to create a final cluster list with the calculated reliability levels.

Fig. 1 Route used for verification of the proposed method (Kalisz-Szczecin)



3 Results and Discussion

The verification of the proposed method was based on a 460 km route (in Poland, between Kalisz and Szczecin – Fig. 1), which was then divided into 92 sections of 5 km each. These sections were assigned to 11 groups. The first 10 groups consist of 9 sections each, and the last group consists of the remaining two sections. According to the methodology, the route was divided into motorways and expressways, and other roads. The routing of the route and individual sections is based on the criteria adopted, i.e. the route should include a minimum of three classes of roads including highways and expressways, the road must be used for freight transport, the road must be of significant economic importance for the regions through which it passes [48].

A fragment of the database containing identified errors on the route is presented in Table 3.

On the section shown in Tables 3, 6 errors in road infrastructure were identified, while on the entire analyzed route the total number of errors was 346. In a situation if the same error occurred several times on the analyzed section (5 km) it was counted as one.

In the result of the clustering process, three sections initiating the cluster were designated for sections containing roads of the class of motorways and expressways, along with eight initiating sections (Table 4) for sections containing other roads (i.e. other than motorways and expressways).

Table 3 Fragment of the database concerning identified errors on the selected route section

No.	Section FROM	Section TO
1.	Wyspiańskiego DK 25, Kalisz Distance from the start: 0 km	DK12, Kościelna Wieś Distance from the start: 5 km
2.	Types of errors identified:	
	1. x_{p4} – Proximity of technical infrastructure elements	
	2. x_{k2} – Vegetation near the roadside	
	3. x_{e7} – Presence of driver distractions – advertisements	
	4. x_{k1} – No shoulder/emergency lanes	
	5. x_{e2} – Vegetation overlapping the roadside	
	6. x_{k3} – No bypasses/transit roads passing through built-up areas	

Source: Authors’ own elaboration

On the basis of calculations regarding the frequency of errors on the analyzed route, constant errors for different road classes were determined. The results of the calculation of the frequency of constant errors for sections containing other roads (other than motorways and expressways) are presented in Fig. 2 [48].

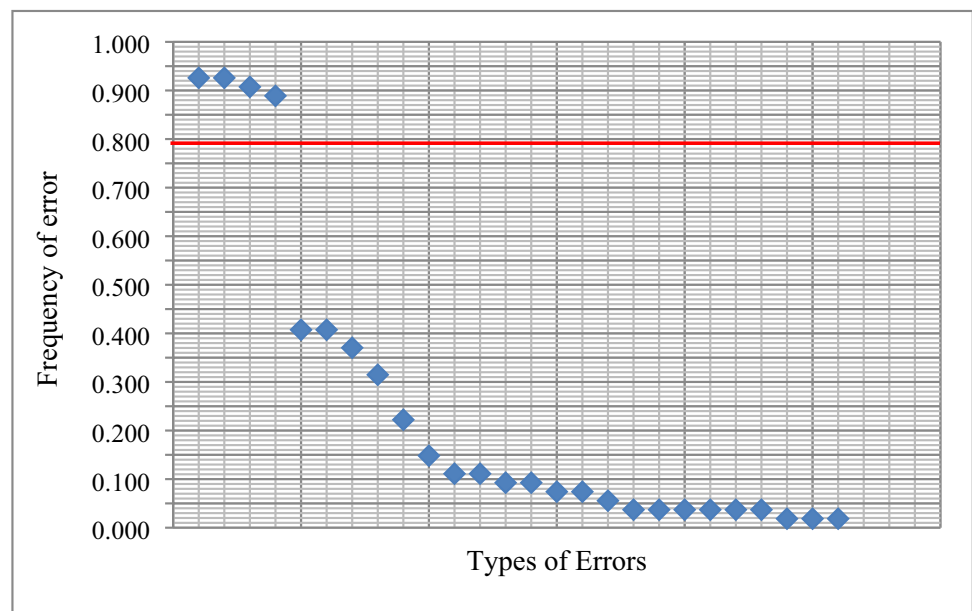
In the case of motorways and expressways, none of the errors reached the level of 0.8, so constant errors were not found. For roads with a class lower than motorways and expressways, four errors were found (x_{k1} , x_{k2} , x_{k3} , x_{e2}).

Table 4 Initiating sections for roads other than motorways and expressways

No.	Road class	Section no.	Distance of the beginning of the section from the starting point	Distance of the end of the section from the starting point
1.	Roads other than motorways and expressways	I. 5.	20	25
		I. 1.	0	5
		V. 2.	185	190
		II. 5.	65	70
		VI. 5.	245	250
		IV. 2.	140	145
		III. 3.	100	105
		I. 7.	30	35

Source: Authors' own elaboration

Fig. 2 Graph of the frequency of errors on sections with a class lower than motorways and expressways. Source: Authors' own elaboration



As a result of the procedure of connecting sections of the route to the initiating cluster, 13 clusters were created, of which 4 with sections representing roads of the class of motorways and expressways, including one cluster of dispersed sections and 9 clusters containing sections of the lower class, including one cluster of dispersed sections. An example of clusters designated for sections with a class lower than motorways and expressways is presented in Table 5.

The presented example of a cluster concerns roads that are not motorways nor expressways, for which four fixed errors x_{k2} , x_{k3} , x_{k1} , x_{e2} (marked in blue) were been determined with frequencies of 0.93, 0.93, 0.91 and 0.89, respectively. The starting section was the one with the

number V.2. located 185 km from the starting point. Three sections were joined with it:

- V.9. o $WPB = 1$ i $PK = 1$,
- III.1. o $WPB = 1$ i $PK = \frac{1}{2}$,
- IV.8. o $WPB = 1$ i $PK = \frac{1}{2}$.

3.1 Assessment of the reliability of the road infrastructure for the transport system

The reliability of the route was assessed in relation to the designated clusters. For this purpose, reliability classes with assigned reliability levels were elaborated.

Table 5 Example of a designated cluster

No.	Road class								
	Roads other than motorways and expressways								
1.	Section no.	Beginning of the section from the starting point	End of the section in relation to the start point	x_{k2}	x_{k3}	x_{k1}	x_{e2}	x_{e4}	x_{e1}
	V.2.	185	190	1	1	1	1	1	1
	V.9.	220	225	1	1	1	1	1	1
	III.1.	90	95	1	1	1	1	1	
	IV.8.	170	175	1	1	1	1	1	

Source: Authors' own elaboration

Table 6 Infrastructure Reliability Assessment

No.	Cluster number	Calculated cluster reliability level	Reliability class
1.	1.1	0,917	Cl.5
2.	1.2	0,875	Cl.7
3.	1.3	0,875	Cl.7
4.	1.4	0,854	Cl.8
5.	1.5	0,813	Cl.10
6.	1.6	0,813	Cl.10
7.	1.7	0,813	Cl.10
8.	1.8	0,793	Cl.10
9.	1.9	0,833	Cl.9
10.	2.1	1000	Cl.1
11.	2.2	0,958	Cl.3
12.	2.3	0,958	Cl.3
13.	2.4	0,896	Cl.6

Source: Authors' own elaboration

Based on the guidelines presented in the methodology, the reliability classes for individual clusters were developed. The results are presented in Table 6. Reliability classes were determined on the basis of [46], where Cl.1 is the highest reliability class.

The results presented in Table 6 allow us to conclude that there is a very wide range of designated reliability classes for the clusters along the analyzed route. This is understandable given that the studied route consists of roads of different categories. The determined classes from points 10 to 12 apply to clusters in which highway and expressway sections were concentrated. The remaining clusters include roads of lower categories, resulting in a visibly lower reliability class.

4 Conclusion

The substantive basis of the application for assessing the reliability of road infrastructure presented in this publication was based on the sections clustering process. The advantage

that the clustering process offers is that it facilitates the analysis of the operational reliability of the infrastructure. This is due to the fact that it eliminates the need to analyse all sections of the route in order to determine its reliability level. Instead of individual sections, only the initiating section of the respective cluster is analyzed, and the sections connected to it, in accordance with the adopted matching criteria, are deemed identical. This type of abridging of the computational process affects the speed of operation of the application based on the proposed methodology.

The proposed approach, consisting in grouping sections into clusters and then determining risk measures in relation to the respective clusters, instead of determining one global risk value for the entire route, allows for a decision-making process aimed at reducing the risk of failure to perform a transport task.

The conducted experiment consisting in recording of the route and its subsequent analysis confirmed the implementation potential of the proposed methodology, which was the objective of the present work. The obtained results indicate that there are a number of errors on the road infrastructure, including errors with a high frequency of occurrence, i.e. no shoulder/emergency lanes, vegetation in the immediate vicinity of the road, no bypasses/transit roads passing through built-up areas, vegetation overlapping the roadside. We performed an assessment of the reliability of road infrastructure on the basis of the designated clusters.

The key achievement of the work is the implementation of cluster analysis, the so-called clustering, for assessing the reliability of transport infrastructure. The proposed method makes it possible, among other things, to assess the reliability of transport routes as a whole, as well as by section, which increases the flexibility of the application's use in commercial purposes.

The limiting element of the proposed methodology is the fact that it concerns a constantly evolving research area, i.e. the transportation system. Consequently, there is a cyclical need to supplement databases. This may affect, among other things, the fixed errors defined in the paper, which will require verification over time.

In further research, the authors will focus on extending the proposed methodology to include further criteria relevant to the reliability of transport infrastructure, e.g. the occurrence of traffic incidents, congestion or infrastructure modernization works.

Acknowledgements This work was supported by the projects VEGA 1/0268/22, KEGA 038TUKE-4/2022 granted by the Ministry of Education, Science, Research and Sport of the Slovak Republic.

Author contributions P.T., S.H and J.H. wrote the main manuscript text, L.K. prepared Figures 1 and 2 and L.K. prepared tables. All authors reviewed the manuscript.

Funding Open access funding provided by The Ministry of Education, Science, Research and Sport of the Slovak Republic in cooperation with Centre for Scientific and Technical Information of the Slovak Republic. This paper is part of a project that has received funding from the Ministry of Education, Science, Research and Sport of the Slovak Republic under grant agreement KEGA 038TUKE-4/2022.

Data availability No datasets were generated or analysed during the current study.

Code availability Not applicable.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Trojanowski P, Trusz A, Stupin B (2022) Correlation between accidents on selected roads as fundamental for determining the safety level of road infrastructure. In: Design, simulation, manufacturing: the innovation exchange. Springer, Cham, pp 104–113. https://doi.org/10.1007/978-3-031-06025-0_11
2. Rolison JJ, Regev S, Moutari S, Feeney A (2018) What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. *Accid Anal Prev* 115:11–24. <https://doi.org/10.1016/j.aap.2018.02.025>
3. Ashraf I, Hur S, Shafiq M, Park Y (2019) Catastrophic factors involved in road accidents: underlying causes and descriptive analysis. *PLoS One* 14(10):e0223473. <https://doi.org/10.1371/journal.pone.0223473>
4. Szumska E, Frej D, Grabski P (2020) Analysis of the causes of vehicle accidents in Poland in 2009–2019. *LOGI Sci J Transp Logistics* 11(2):76–87. <https://doi.org/10.2478/logi-2020-0017>
5. Lee AS, Lin WH, Gill GS, Cheng W (2019) An enhanced empirical bayesian method for identifying road hotspots and predicting number of crashes. *J Transp Saf Secur* 11(5):562–578. <https://doi.org/10.1080/19439962.2018.1450314>
6. Mukherjee S, Jain T (2021) Do the mobility patterns for city taxicabs impact road safety? *IISE Trans* 53(12):1324–1336. <https://doi.org/10.1080/24725854.2021.1914879>
7. Mane SS, Shah SAK (2017) Object tracking and analysis for detecting various situations at level crossings of railway-road. In: In 2017 international conference on computing, communication, control and automation (ICCUBEA). IEEE, pp 1–5. <https://doi.org/10.1109/ICCUBEA.2017.8463973>
8. Elena K, Sergei E, Jarosław R (2020) Potential for improving the procedure of inspecting road traffic accident black spots. *Archit Eng* 5(3):56–62. <https://doi.org/10.23968/2500-0055-2020-5-3-56-62>
9. Ghadi M, Török Á (2019) A comparative analysis of black spot identification methods and road accident segmentation methods. *Accid Anal Prev* 128:1–7. <https://doi.org/10.1016/j.aap.2019.03.002>
10. Debrabant B, Halekoh U, Bonat WH, Hansen DL, Hjelmborg J, Lauritsen J (2018) Identifying traffic accident black spots with Poisson-Tweedie models. *Accid Anal Prev* 111:147–154. <https://doi.org/10.1016/j.aap.2017.11.021>
11. Trojanowski P, Filina-Dawidowicz L (2021) Diagnostic and repair centers locating methodology for vehicles carrying sensitive cargo. *Transp Res Proc* 55:410–417. <https://doi.org/10.1016/j.trpro.2021.07.004>
12. Fan Z, Liu C, Cai D, Yue S (2019) Research on black spot identification of safety in urban traffic accidents based on machine learning method. *Saf Sci* 118:607–616. <https://doi.org/10.1016/j.ssci.2019.05.039>
13. Iqbal A, Ur Rehman Z, Ali S, Ullah K, Ghani U (2020) Road traffic accident analysis and identification of black spot locations on highway. *Civil Eng J* 6(12):2448–2456. <https://doi.org/10.28991/cej-2020-03091629>
14. Adamczak M, Kolinski A, Trojanowska J, Husár J (2023) Digitalization trend and its influence on the development of the operational process in production companies. *Appl Sci* 13:1393. <https://doi.org/10.3390/app13031393>
15. Drosu A, Cofaru C, Popescu MV (2020) Influence of weather conditions on fatal road accidents on highways and urban and rural roads in Romania. *Int J Automot Technol* 21(2):309–317. <https://doi.org/10.1007/s12239-020-0029-4>
16. Trojanowski P, Trojanowska J (2021) Reliability of road transport means as a factor affecting the risk of failure—the transport problem case study. In: Design, simulation, manufacturing: the innovation exchange. Springer, Cham, pp 253–261. https://doi.org/10.1007/978-3-030-77719-7_26
17. Nguyen T D, Gupta S, Andersen M, Bento A, Simon K I, Wing C (2020) Impacts of state reopening policy on human mobility (No. w27235). National Bureau of Economic Research. <https://doi.org/10.3386/w27235>
18. Zhu L, Yu FR, Wang Y, Ning B, Tang T (2018) Big data analytics in intelligent transportation systems: a survey. *IEEE Trans Intell Transp Syst* 20(1):383–398. <https://doi.org/10.1109/TITS.2018.2815678>
19. Kurganov V, Gryaznov M, Dorofeev A (2018) Management of transportation process reliability based on an ontological model of an information system. *Transp Res Proc* 36:392–397. <https://doi.org/10.1016/j.trpro.2018.12.113>
20. Alic AS, Almeida J, Aloisio G, Andrade N, Antunes N, Ardagna D et al (2019) BIGSEA: a big data analytics platform for public transportation information. *Futur Gener Comput Syst* 96:243–269. <https://doi.org/10.1016/j.future.2019.02.011>
21. Pal C, Hirayama S, Narahari S, Jeyabharath M, Prakash G, Kulothungan V (2018) An insight of World Health Organization

- (WHO) accident database by cluster analysis with self-organizing map (SOM). *Traffic Inj Prev* 19(sup1):S15–S20. <https://doi.org/10.1080/15389588.2017.1370089>
22. Frehaileab AG, Szpytko J (2019) Reliability assessment of the transport system, Addis Ababa case study. *J KONBiN* 49(4):27–36. <https://doi.org/10.2478/jok-2019-0073>
 23. Szkoda M, Kaczor G, Satora M (2019) Reliability and availability assessment of a transport system using dynamic fault tree and Monte Carlo simulation. In IOP conference series: materials science and engineering (Vol. 664, No. 1, 012028). IOP Publishing. <https://doi.org/10.1088/1757-899X/664/1/012028>
 24. Smolarek L, Ziemska M (2019) Modelling of transport system operational reliability a Markov approach. *J KONES* 26. <https://doi.org/10.2478/kones-2019-0112>
 25. Yang K, Wang R, Yang L (2020) Fuzzy reliability-oriented optimization for the road-rail intermodal transport system using tabu search algorithm. *J Intell Fuzzy Syst* 38(3):3075–3091. <https://doi.org/10.3233/JIFS-191010>
 26. Husar J, Knapcikova L, Balog M (2019) Implementation of material flow simulation as a learning tool. In: Ivanov V et al (eds) *Advances in design, simulation and manufacturing*. DSMIE 2019, Lecture notes in mechanical engineering. Springer, Cham. https://doi.org/10.1007/978-3-319-93587-4_4
 27. Persia L, Usami DS, De Simone F, De La Beaumelle VF, Yannis G, Laiou A et al (2016) Management of road infrastructure safety. *Transp Res Part 14*:3436–3445. <https://doi.org/10.1016/j.trpro.2016.05.303>
 28. Budzyński M, Kustra W, Okraszewska R, Jamroz K, Pырchla J (2018) The use of GIS tools for road infrastructure safety management. In: In E3S web of conferences, vol 26. EDP sciences, p 00009. <https://doi.org/10.1051/e3sconf/20182600009>
 29. Kustra W, Michalski L (2017) Tools for road infrastructure safety management in Poland. In: In MATEC web of conferences, vol 122. EDP sciences, p 02008. <https://doi.org/10.1051/mateconf/201712202008>
 30. Budzynski M, Jamroz K, Kustra W, Michalski L, Gaca S (2017) Road infrastructure safety management in Poland. In IOP conference series: materials science and engineering (245, No. 4, 042066). IOP Publishing. <https://doi.org/10.1088/1757-899X/245/4/042066>
 31. Directive 2008/96/EC of the European Parliament and of the Council of 19 November 2008 on road infrastructure safety management
 32. Ge H, Dong L, Huang M, Zang W, Zhou L (2022) Adaptive kernel density estimation for traffic accidents based on improved bandwidth research on black spot identification model. *Electronics* 11(21):3604. <https://doi.org/10.3390/electronics11213604>
 33. Valentin V, Naderpajouh N, Abraham DM (2018) Integrating the input of stakeholders in infrastructure risk assessment. *J Manag Eng* 34(6):04018042. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000638](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000638)
 34. Elena K, Sergei E, Grigory G (2020) Systemic indicators of road infrastructure at accident clusters. *Archit Eng* 5(1):51–58. <https://doi.org/10.23968/2500-0055-2020-5-1-51-58>
 35. Ghadi M, Török Á, Tanczos K (2019) Integration of probability and clustering based approaches in the field of black spot identification. *Period Polytech: Civ Eng* 63(1):46–52. <https://doi.org/10.3311/PPci.11753>
 36. Hu L, Wu X, Huang J, Peng Y, Liu W (2020) Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. *Saf Sci* 127:104710. <https://doi.org/10.1016/j.ssci.2020.104710>
 37. Trojanowska J, Varela MLR, Machado J (2017) The tool supporting decision making process in area of job-shop scheduling. In: *World conference on information systems and technologies*. Springer, Cham, pp 490–498. https://doi.org/10.1007/978-3-319-56541-5_50
 38. Knapčíková L, Martiček M, Husár J, Kaščák J (2022) Intelligent monitoring of loading and unloading process in Enterprise transport system. In: *International conference on future access enablers of ubiquitous and intelligent infrastructures*. Springer, Cham, pp 193–203. https://doi.org/10.1007/978-3-031-15101-9_14
 39. Filina-Dawidowicz L, Cernova-Bickova A, Semenov I, Moždreń D, Wiktorowska-Jasik A, Bickovs D (2020) Information support of cargo ferry transport: case study of Latvia. *Procedia Comput Sci* 176:2192–2201. <https://doi.org/10.1016/j.procs.2020.09.256>
 40. Gharib Z, Bozorgi-Amiri A, Tavakkoli-Moghaddam R, Najafi E (2018) A cluster-based emergency vehicle routing problem in disaster with reliability. *Sci Iran* 25(4):2312–2330. <https://doi.org/10.24200/sci.2017.4450>
 41. Pettet G, Nannapaneni S, Stadnick B, Dubey A, Biswas G (2017) Incident analysis and prediction using clustering and bayesian network. In: In 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & big Data Computing, internet of people and Smart City innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). IEEE, pp 1–8. <https://doi.org/10.1109/UIC-ATC.2017.8397587>
 42. Mohanty A, Mahapatra S, Bhanja U (2019) Traffic congestion detection in a city using clustering techniques in VANETs. *Indones J Electr Eng Comput Sci* 13(2):884–891. <https://doi.org/10.11591/ijeecs.v13.i3.pp884-891>
 43. Iwańkiewicz RR, Rosochacki W (2014) Clustering risk assessment method for shipbuilding industry. *Ind Manag Data Syst* 114(9):1499–1518. <https://doi.org/10.1108/IMDS-06-2014-0193>
 44. Soilán M, Riveiro B, Martínez-Sánchez J, Arias P (2017) Segmentation and classification of road markings using MLS data. *ISPRS J Photogramm Remote Sens* 123:94–103. <https://doi.org/10.1016/j.isprsjprs.2016.11.011>
 45. Ghebleh R (2018) A comparative classification of information dissemination approaches in vehicular ad hoc networks from distinctive viewpoints: a survey. *Comput Netw* 131:15–37. <https://doi.org/10.1016/j.comnet.2017.12.003>
 46. Guido G, Haghshenas SS, Haghshenas SS, Vitale A, Gallelli V, Astarita V (2020) Development of a binary classification model to assess safety in transportation systems using GMDH-type neural network algorithm. *Sustainability* 12(17):6735. <https://doi.org/10.3390/SU12176735>
 47. Bengag A, El Boukhari M (2018) Classification and comparison of routing protocols in VANETs. In: In 2018 international conference on intelligent systems and computer vision (ISCV). IEEE, pp 1–8. <https://doi.org/10.1109/ISCV.2018.8354025>
 48. Trojanowski P (2019) Methodology of designing and analyzing the operational reliability of the land-water transport system for sensitive cargoes to polish seaports, dissertation, west Pomeranian University of technology in Szczecin. Faculty of Maritime Technology and Transport, Szczecin
 49. Hrehová S, Husár J, Knapčíková L (2022) The fuzzy logic predictive model for remote increasing energy efficiency. *Mobile Netw Appl*. <https://doi.org/10.1007/s11036-022-02050-1>
 50. Kozłowski E, Antosz K, Mazurkiewicz D, Sęp J, Żabiński T (2021) Integrating advanced measurement and signal processing for reliability decision-making. *Eksploatacja i Niezawodność – maintenance and reliability* 23(4):777–787. <https://doi.org/10.17531/ein.2021.4.20>
 51. Kascak J, Baron P, Torok J, Pollak M, Teliskova M (2019) Macrostructure digitalization of roadway surface profiles. *MM Sci J*:2839–2844. https://doi.org/10.17973/MMSJ.2019_03_201875

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.