ISSN 1648-4142 / eISSN 1648-3480

VILNIUS TECH Vilnius Gediminas Technical University

TRANSPORT

2023
Volume 38
Issue 2
Pages 77-86
https://doi.org/10.3846/transport.2023.19407

Original Article

APPLICATION OF EXPERT EVALUATION METHODS IN DETERMINING THE SIGNIFICANCE OF CRITERIA FOR USABILITY OF RAILWAY TRACTION ROLLING STOCK

Lijana MASKELIŪNAITĖ[1]✉, Laurynas MEILUS[2], Henrikas SIVILEVIČIUS[3]

[1,2]*Dept of Mobile Machinery and Railway Transport, Vilnius Gediminas Technical University, Lithuania*
[3]*Transport and Logistics Competence Centre, Vilnius Gediminas Technical University, Lithuania*

Highlights:
- suitability of railway traction vehicles is determined by a significance of criteria's;
- the significance of the criteria's is determined by expert evaluation methods;
- when the experts' opinions are consistent, their average is taken as the result;
- 3 expert evaluation methods allow obtaining more reliable criteria weights;
- the most important criteria's: safety, comfort and ecology of rolling stock.

Article History:
- submitted 16 August 2022;
- resubmitted 15 November 2022, 28 December 2022;
- accepted 22 January 2023.

Abstract. Railway rolling stock must meet the requirements related to its use in the transportation process. The significance of these requirements can be determined using expert testing methods. The current research offers a framework of 9 criteria, which have been developed by the authors of the study, and which contribute to a comprehensive assessment of their importance and priority in relation to each other using expert evaluation methods. The normalised weights of the criteria were determined using Average Rank Transformation Into Weight Linear (ARTIW-L), Average Rank Transformation Into Weight Non-linear (ARTIW-N) and Direct Percentage Weight (DPW) methods. The criteria were given ranks and percentage weights by 18 experts with consistent opinions, which made it reasonable to consider the average of the experts' opinions as the outcome of the task. The normalised weights of the criteria have shown that the most important issues for the experts included passenger and crew safety (0.1619), passenger and train staff ride comfort (0.1330) and environmental protection (0.1201). The least important criteria for the experts cover the range per one electric charge or full tank of fuel (0.0776), the dynamic performance of the traction rolling stock (0.0849), and the purchase price, the rebate system, the duration of the warranty period (0.0911). The other 3 criteria are of medium importance. The outcomes of this study can be used in deciding on the best alternative for rail traction rolling stock.

Keywords: rail transport, traction rolling stock, serviceability criteria, weights, experts, priority.

✉Corresponding author. E-mail: *lijana.maskeliunaite@vilniustech.lt*

Notations

AHP – analytical hierarchy process;
ARTIW-L – average rank transformation into weight linear;
ARTIW-N – average rank transformation into weight non-linear;
CFT – conventionally fuelled train;
CFV – conventionally fuelled vehicle;
CO_2 – carbon dioxide;
DPW – direct percentage weight;
ET – electric train;
FCH – fuel cell and hydrogen;
HT – hydrogen train;
JSC – joint-stock company;
MCDM – multiple-criteria decision-making.

Introduction

Quality of life depends on efficient transport that increases people's mobility. Travel is not only by road, but also by rail. Nowadays, for travel people can use a variety of means of transport, and they usually choose the one that is the most suitable and reliable for them. The choice of one mode of transport as an alternative to another is subjective. It is usually based on passengers' views regarding the advantages and disadvantages of certain means of transport. The methods and results of the study on passengers' preferences for sustainable train travel as an alternative to travelling by bus (Maskeliūnaitė, Sivilevičius 2021) and by air (Sivilevičius, Maskeliūnaitė 2018) are presented.

To this day, rail transport continues to be popular all over the world, with improvements in railway track and rolling stock, the better quality of passenger service, and increased safety of travel, which allow rail transport to compete with other modes of transport (Maskeliūnaitė 2021). The assessment of the importance of energy quality (Danwen *et al.* 2021) and energy efficiency (García-Garre, Gabaldón 2019) in electrified railways is carried out. Not only passengers but also freight is carried by rail.

Alternative traction is a major focus for researchers and rolling stock companies. The efficiency of energy, diesel, and hydrogen traction as well as CO_2 emission reduction in rail transport is being investigated (Hoffrichter *et al.* 2012). Due to environmental requirements, diesel traction is being replaced by environmentally friendly alternatives in many parts of the world. Rail lines are being electrified rapidly; however, it is unlikely that all routes will be electrified in the near future. In addition, passenger numbers carried by rail will not justify the high infrastructure costs. Some lines are likely to have a combination of hydrogen and electric traction. The railway traction drives will need to be modified to accommodate these new power supplies. When evaluating new hybrid electric systems, a detailed overview of the available traction motors and drives is provided (Polater, Tricoli 2022).

The EU's ambitions include the reduction of environmental pollution and finding alternative means of transport. European passenger carriers employ alternative traction – batteries or hydrogen. This is done with the view of moving away from diesel engines. France and Germany have started introducing hybrid trains. The UK is also testing hydrogen and battery traction. Norway is considering replacing diesel engines with battery-powered trains. The *White paper on Transport* (EC 2011) provides the reduction of CO_2 emissions: by 2030 – 30%, by 2035 – 50%, and by 2050 – 90%. Priority is given to electrification and the use of sustainable sources of electricity. Many countries have significant electrification programmes in place, notably Germany, which has a target of 70% electrification by 2030, up from 52% today. However, electrifying railways is a costly process that is not always "green", depending on the source of electricity. In recent years, the electrification process has been problematic in UK, where diesel continues to be the main form of traction on many important lines. Other European networks with significant non-electrified sections and potential for alternative traction include the Scandinavian countries, Belgium, the Netherlands, Spain, Ireland, and Italy (Smith 2020).

In North America, despite some alternative traction tests, railways continue to run diesel trains. The Northeast Corridor is electrified, as are some commuter train lines, although most cities in the US and Canada use diesel traction. Australia also uses diesel traction, as does Brazil and much of the Russian network, as well as railways in Asia, Latin America, and Africa. In North America and Australia, the willingness of rail operators to move away from diesel and the interest in alternative traction is very high (Smith 2020).

Suppliers can offer a range of traction solutions to meet the demand. It can be a hybrid battery solution and traditional electric traction with a range of 100...120 km on batteries, or hydrogen fuel cells that can operate over a range of 600...1000 km. Germany is the first country to enter major contracts for battery-ETs. Government subsidies are not yet in place in other countries, so demand for diesel rail transport will remain stable, especially outside Europe. Alternative traction currently accounts for less than 2% of total orders. Moreover, diesel technology itself is becoming much cleaner. Generation-V engines limit particulate matter to 0.015 g/kW·h per rail car. This is a 40% reduction compared to the previous emission limit of 0.025 g/kW·h per rail car for Generation–IV engines. In some highly polluted areas, the air emitted by the Generation-V engine is cleaner than the surrounding atmosphere. Hybrid power packs also help reduce fuel consumption and CO_2 emissions by up to 25%. Moreover, there could still be good developments in the diesel traction market (Smith 2020).

Diesel and 30% of electricity are mostly used for rail traction. Concerns about environmental pollution and public health are driving the research of traction alternatives. Electrification is a traditional way to avoid fuel supply problems and environmental pollution. However, it requires significant investment. As a result, diesel traction, which is not dependent on the electrification of the railway line, is sometimes the only option. Hydrogen as secondary energy can be produced from a variety of raw materials such as fossil fuels, nuclear energy, and renewable energy sources. This reduces greenhouse gas emissions. Studies have shown that hydrogen-powered rail traction is technically feasible. It reduces energy consumption and greenhouse gas emissions and is not dependent on oil (Hoffrichter 2013). The studies aim to assess the safety, adaptability, and efficiency of traction alternatives for rail transport. One alternative does not have to be the best in all respects. Studies on hydrogen fuel cell technology for rail propulsion have been executed (Ehrhart *et al.* 2021). Researchers focus on the potential of FCH technologies in railways and identify technical and non-technical barriers to market entry (Ruf *et al.* 2019).

ETs are the most energy-efficient means of transport. They are powered directly by electricity. Other vehicles need to store and carry electricity to use it. This is mostly done by batteries; however, HTs can also store electricity. This is because hydrogen can be produced by electrolysis. It uses electricity to split water into hydrogen and oxygen. HT fuel cells can then use it to generate electricity, a process that effectively replaces the electrolysis, although storing electrical energy in this way is more complicated and less efficient than using batteries. *Siemens* comments that the efficiency of the different types of traction, ranging from the primary energy source to power at the wheels, is as follows: line electrification 80%, battery traction 65%, hydrogen 25% and diesel 25%.

A HT can travel longer distances than a battery train because the energy density of hydrogen is twice that of a battery. If diesel engines are discontinued in the future, hydrogen is the only self-propelled propulsion that can power a passenger train for 1000 km/day. Hydrogen energy density is 1/8 that of diesel and its fuel tanks need to be 8 times larger. This makes it unsuitable for freight and high-speed passenger trains (Shirres 2020).

Interest in hydrogen-powered rail vehicles has gradually increased around the world in recent decades due to global pressure to reduce greenhouse gas emissions. Increase access to technology and energy supply. Scientific research and development have focused mainly on light rail and regional trains. Recently, there has been a growing interest in hydrogen-powered freight and heavy trains. The technical feasibility has been established through project studies and experiments. Several hydrogen-powered rail vehicles are currently in operation or are the subject of experimental programmes (Sun *et al.* 2021).

ETs and HTs are considered zero emission at the point of use. True emissions are dependent upon non-tailpipe sources, primarily in energy production. UK CO_2 operating emission model outputs for CFTs, ETs and HTs between 2017 and 2050 under 4 *National Grid* electricity generation scenarios, are presented in research by Logan *et al.* (2020). Results indicate by 2050 at 100% capacity CFTs produce a fifth of the emissions of CFVs per kilometre per person. Under 2 degree generation scenario, by 2050 ETs produced 14 times and HTs produced 5 times less emissions than CFTs. Policymakers should encourage shifts away from private vehicles to public transport powered by low-carbon electricity. The development of an integrated hybrid train simulator for the *Piedmont Intercity Railway* has been introduced in North Carolina (US). A case study is used. The study includes 6 train configurations, 9 hydrogen supply options, without diesel supply. The results show that a hybrid option is not only feasible, but a low-carbon hydrogen supply chain may be possible (Madovi *et al.* 2021).

In the development of railway transport, it is necessary to constantly renew the traction vehicles. These vehicles can use different types of energy. Which type of traction most accurately economic, ecological, traffic safety and other requirements can be determined by applying expert research methods. The general opinion of the experts allows to determine the significance of the criteria by which the type of traction of the rolling stock is chosen.

MCDM is a research area that involves the analysis of various available choices in the situations or researches, which spans daily life, social sciences, engineering, medicine, and many other areas. MCDM methods, also, are used to solve transport scientific problems (Boghani *et al.* 2021; Hamurcu, Eren 2022; Li *et al.* 2020; Rostamzadeh *et al.* 2020). In the MCDM methods used to select the most suitable alternative, the subjective weights of the criteria are usually determined at the beginning (Mastrocinque *et al.* 2020; Şahin 2021; Sivilevičius, Maskeliūnaitė 2014). The theory of criteria determination and evaluation is presented and suggested a suitable order of criteria (Bajec, Tuljak-Suban 2020). Discussion are developed of criteria for evaluating different methods of group decision-making that range from the strictly technical, to the psychophysical, social and to the logical and scientific (Peniwati 2007).

The aim of this article is to research the types of railway transport traction rolling stock and develop a system of criteria that determine the suitability of railway traction rolling stock for use. Using the opinions of rail transport experts, determine the significance of these criteria using expert evaluation methods.

1. Criteria for railway traction rolling stock

Each author of the article wrote down all, in his opinion, important factors that indicate the necessary (required) characteristics of rolling stock. These criteria were discussed and formulated by the authors. Their descriptions are provided in the questionnaire. We aimed for a smaller number of criteria, which allowed us not to divide them into a hierarchical structure. The original form of the 9 criteria questionnaire compiled by the authors was adapted to evaluate the significance of the criteria by means of ranks, percentage weights and the AHP method, filling in the criteria pairwise comparison matrix. The prepared questionnaire was submitted to several specialists performing practical activities of railway transport in order to make sure that the wording of all criteria is understandable and that no important criteria have been omitted. There were no significant comments. The 9 criteria provide a comprehensive assessment of the properties that are relevant in choosing the best type of rolling stock. The study compares the following criteria, giving them ranks and percentage weights:

- A – operating costs (labour costs, cost of energy resources, maintenance costs: cost of lubricants, filters, and other replaceable items) per defined mileage, i.e. length of the road travelled (usually for 100 km);
- B – purchase price, rebate system, length of the warranty period;
- C – environmental protection (emissions, external noise);
- D – passenger and train crew ride comfort (internal noise, vibrations, temperature, air speed, air pressure, relative humidity, lighting, high acceleration, and deceleration);
- E – safety of passengers and staff (in the event of a derailment, collision with animals and other vehicles, fire resistance, effectiveness of evacuation and medical measures);
- F – distance travelled on a single electric charge or full fuel tank;
- G – durability and reliability (sustainable performance, service life before overhaul, survival time);
- H – infrastructure installation and maintenance costs (rail track, power supply facilities, railway stations, equipment, repair depots, crossings);
- I – dynamic indicators of the traction unit (maximum tractive force of the locomotive or powered wagon, its engine power, torque, speed).

The indicated criteria were set out in a questionnaire, which started with an address to an expert and the procedure for completing the questionnaire. An expert gave the criteria a ranking from 1 to 9 and percentage weights, which had to add up to 100%. There could be no duplicate ranks and percentage weights.

2. Experts

The description of the experts to whom the questionnaires were administered is presented in Table 1. There were 18 experts who shared their opinion.

One of the most important steps in the study was the selection of a panel of experts in the field. The opinion of a group of experts is always more accurate than that of a single expert. The composition and number of experts are selected on the basis of their expertise, the required reliability of the evaluations and the cost of the resources. The issue to be researched requires the involvement of specialists from a wide range of disciplines in the subject area. The minimum number of experts can be calculated by assessing the variation in their opinions (the principle of sample size) for the required confidence and accuracy (Maskeliūnaitė, Sivilevičius 2021).

3. Methods for determine the significance of criteria

The significance of criteria, expressed as their subjective normalised weights, can be determined using different methods (algorithms). None of them has a theoretical advantage over the other methods. However, the general principle of all these algorithms is the same: the most important criterion must be given the highest weight. The sum of the significance weights of all criteria must be equal to 1, i.e., they are normalised.

To increase the reliability of the research result, 3 expert evaluation methods were chosen. Initially, the rank correlation method presented by Kendall (Kendall, Gibbons 1990) has been employed. From the researched criteria ranks, applying ARTIW-L and ARTIW-N methods (Sivilevičius 2011; Šakalys *et al.* 2019; Maskeliūnaitė, Sivilevičius 2021), the normalized criteria weights have been calculated. The experts also gave the criteria percentage importance estimates, which were used to calculate the total normalised weights of their subjects using the DPW method.

The ARTIW-L and ARTIW-N methods were chosen to determine the weights of the criteria, which are relatively simple and accurate at the same time, as they allow the weights to be calculated from the averages of the ranks. DPW method is also not difficult because it is convenient for the expert to give higher percentage weights to more important criteria.

3.1. ARTIW-L method

This method calculates the average rank $\bar{R}_i$ given to each criterion (i = 1, 2, ..., m) in the study by the experts (j = 1, 2, ..., n):

$$\bar{R}_i = \frac{\sum_{j=1}^{n} R_{ij}}{n}, \quad (1)$$

where: R_{ij} is the rank given to the ith criterion by the jth expert; n is the number of experts.

The subjective normalised weight ω_i of each criterion is calculated using the ARTIW-L method (Sivilevičius 2011):

$$\omega_i = \frac{(m+1) - \bar{R}_i}{\sum_{i=1}^{m} \bar{R}_i}, \quad (2)$$

where: $\bar{R}_i$ is the average rank of the ith criterion; m is the number of criteria describing the object of study.

The criteria weights ω_i calculated by the ARTIW-L method are linearly correlated with the average ranks of these criteria $\bar{R}_i$.

To take the average of the opinions of a group of experts, expressed in ranks, scores, or weights, as the result of a task, it is necessary to check that the opinions of all the experts are not contradictory. Only when the assessment of all the experts is consistent (not contradictory) is the average a reasonable solution. For this purpose, Kendall's rank coefficient of concordance W (Kendall, Gibbons 1990), which can range from 0 to 1, is calculated to indi-

Table 1. Characterization of the experts who have assessed the significance of railway traction rolling stock

No	Expert	Number of respondents
1	Director-General, JSC "LTG Link" (the railway passenger company of Lithuania)	1
2	Heads of Departments, JSC "LTG Link"	3
3	Heads of Department Divisions, JSC "LTG Link"	3
4	Project managers, JSC "LTG Link"	2
5	Engineers, UAB "LTG Link"	6
6	Professor and Associate Professor, Faculty of Transport Engineering, Vilnius Gediminas Technical University	2
7	BSc Student of Railway Transport Engineering, Faculty of Transport Engineering, Vilnius Gediminas Technical University	1
	Total number of respondents:	18

cate the consistency of the experts' opinions:

$$W = \frac{12 \cdot S}{n^2 \cdot \left(m^3 - m\right)} = \frac{12 \cdot S}{n^2 \cdot m \cdot \left(m^2 - 1\right)}. \quad (3)$$

The sum S of the squares of the sum of the deviations of the ranks $\sum_{j=1}^{n} R_{ij}$ of each mth criteria from the average rank $\bar{R}$, is calculated from the formula:

$$S = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} R_{ij} - \frac{1}{2} \cdot n \cdot \left(m+1\right) \right)^2, \quad (4)$$

where: m is the number of criteria (i = 1, 2, ..., m); n is the number of experts (j = 1, 2, ..., n).

The calculated value W is compared with its minimum value W_{min}, which depends on the chosen significance level α (assumed α = 0.05 or more stringent α = 0.01) and the number of degrees of freedom $\nu = m - 1$ (Sivilevičius 2011):

$$W_{min} = \frac{\chi^2_{\alpha,\nu}}{n \cdot \left(m-1\right)}, \quad (5)$$

where: $\chi^2_{\alpha,\nu}$ is the Pearson criterion statistic, found in the mathematical statistics appendix table (Montgomery 2012; Čekanavičius, Murauskas 2000).

The integrity of the experts' opinions can be checked by calculating a random variable:

$$\chi^2 = W \cdot n \cdot \left(m-1\right) = \frac{12 \cdot S}{n \cdot m \cdot \left(m+1\right)}, \quad (6)$$

which follows a χ^2 (chi-squared) distribution with degrees of freedom $\nu = m - 1$.

According to the chosen significance level α (0.05 or 0.01) from χ^2 the distribution table (Čekanavičius, Murauskas 2000), the critical value $\chi^2_{\alpha,\nu}$ is found with a degree of freedom $\nu = m - 1$. If the value χ^2 calculated according to Equation (6) is greater than $\chi^2_{\alpha,\nu}$, then the experts' judgements are considered to be in agreement (flawless, similar).

The compatibility (consistence) coefficient k_c, which is calculated from the Equation (7), shows how many times the calculated concordance coefficient W is greater than its minimum (critical) value W_{min}, and how many χ^2 times it is greater than its critical value $\chi^2_{\alpha,\nu}$:

$$k_c = \frac{W}{W_{min}} = \frac{\chi^2}{\chi^2_{\alpha,\nu}}. \quad (7)$$

When the opinions expressed by the experts are in agreement, then $k_c > 1$. Otherwise (when $k_c < 1$), the opinions expressed by the experts differ significantly, i.e., they are not harmonised.

3.2. ARTIW-N method

The criteria normalised weights ω_i can be calculated by employing another average rank $\bar{R}_i$ transformation into weights method. The criteria weights ω_i calculated according to Equations (8) and (9) are related to the criteria average ranks $\bar{R}_i$ by a non-linear inverse correlation (functional). Therefore, this method is named ARTIW-N method by the authors of this research (Maskeliūnaitė, Sivilevičius 2021).

By applying the ARTIW-N method, at the beginning the ratio of the most important criterion u_i with the lowest average rank $\min_i \bar{R}_i$ to the average ranks $\bar{R}_i$ of all the other ith criteria is calculated:

$$u_i = \frac{\min_i \bar{R}_i}{\bar{R}_i}. \quad (8)$$

After normalizing the values of each criterion u_i, their subjective significances (weights) ω_i are calculated:

$$\omega_i = \frac{u_i}{\sum_{i=1}^{m} u_i}. \quad (9)$$

The criteria weights ω_i calculated by the ARTIW-N method are non-linearly correlated with the average ranks of these criteria. This correlation shows that the significance of the most important and the least important criteria is "amplified" by reducing the significance of the criteria of medium importance.

3.3. DPW method

This is the most common method used in practice. The DPW method is as clear and logical as the ranking (rank correlation) method for indicators criteria, but its accuracy is much higher. For the DPW method, the sum of the weights of all the evaluations of the research criteria for each of the expert's objects shall be equal to 100%. With percentage evaluations, they are divided by 100. It is easier to rank the criteria in percentages if they have been ranked at the start: a lower rank corresponds to a higher weight.

The sums $\sum_{j=1}^{n} p_{ij}$ of the percentage weights p_{ij} of all n experts (j = 1, 2, ..., n) for each ith criterion and their normalized weights ω_i are calculated:

$$\omega_i = \frac{\sum_{j=1}^{n} p_{ij}}{100 \cdot n}, \quad (10)$$

where: p_{ij} is the percentage weight given to the ith criterion by the jth expert; n is the number of experts.

When assigning percentage weights to criteria, it is preferable that there are no identical (duplicate) weights. As a result, they may be given to the nearest tenth of a percentage.

The criteria weights ω_i calculated by the DPW method are non-linearly correlated with the average ranks $\bar{R}_i$ of these. Most of the time, "non-linearity" is decided by the "increased" significance of the most important criteria.

3.4. Criteria average weights of the 3 expert evaluation methods

The principle of mathematical statistics is applied, which states that the arithmetic mean of the subject parameter $\bar{x}$ of a sample consisting of N samples is closer to the population mean μ than the value x_i of any individual sample. It is likely that the arithmetic mean of the weights of each criterion determined by 3 methods of research is closer to the actual weight of this criterion than the weight determined by other method.

Expert evaluation methods that have become classical or generally accepted that produce close or different criteria weights have no theoretical advantage over each other. The average results of several methods can be used to increase the reliability of the study outcomes. The criteria weights arithmetic mean calculated by 3 expert evaluation methods is computed from Equation (11), with no preference (advantage) given to any of them:

$$\bar{\bar{\omega}}_i = \frac{\sum_{k=1}^{r} \omega_{ik}}{r} = \frac{\omega_i^{\text{ARTIW-L}} + \omega_i^{\text{ARTIW-N}} + \omega_i^{\text{DPW}}}{3}, \quad (11)$$

where: the weight ω_{ik} of the ith criterion is calculated by the kth expert evaluation method (k = 1, 2, ..., r); r is the number of expert evaluation methods used in the study (r = 3).

It is likely that $\bar{\bar{\omega}}_i$ is closer to the population average, which would be obtained by interviewing a very large number of experts, than the values of the criteria weights calculated by any of the 3 methods used in the study.

4. Results and their analysis (discussion)

Each of the 18 experts read the questionnaire, studied, and understood the description of the 9 criteria detailed in the questionnaire. With reference to their personal expertise, knowledge, experience, and intuition the experts ranked them from 1 (the most important criterion) to 9 (the least important criterion) (Table 2).

These criteria were then ranked by the experts, so that the most important criterion was given the highest percentage weight and the least important criterion the lowest percentage score (Table 3).

In order that the arithmetic mean $\bar{R}_i$ of the ranks R_{ij} given to each criterion by all the experts could reasonably be taken as the overall (average, collective) opinion of the group of experts, the consistency of their opinions was checked, as indicated by the concordance coefficient W. Calculating W (Equation (3) the sum of the squares deviations was found to be S = 3628 (Equation (4), n = 18, m = 9):

$$W = \frac{12 \cdot S}{n^2 \cdot m \cdot (m^2 - 1)} = \frac{12 \cdot 3628}{18^2 \cdot 9 \cdot (9^2 - 1)} = 0.1866.$$

The minimum value of the concordance coefficient $W_{\min}$ (Equation (5)) at which the opinions of the experts are considered to be in agreement (non-contradictory), and when the level α = 0.05 of significance and the number of degrees of freedom $\nu = m - 1 = 9 - 1 = 8$, is:

$$W_{\min} = \frac{\chi^2_{\alpha,\nu}}{n \cdot (m-1)} = \frac{20.09}{18 \cdot (9-1)} = 0.1395.$$

The consistency of the experts' opinions was verified by another proposed method (Kendall, Gibbons 1990), which is suitable for practical use when more than m = 7 criteria are compared. For this purpose, a χ^2 statistic has been calculated (Equation (6)):

$$\chi^2 = \frac{12 \cdot S}{n \cdot m \cdot (m+1)} = \frac{12 \cdot 3628}{18 \cdot 9 \cdot (9+1)} = 26.87.$$

The statistics χ^2 = 26.87 calculated from the survey data are higher than the critical value $\chi^2_{0.05;8} = 20.09$, thus, the expert opinions are said to agree.

The concordance coefficient W = 0.1866 is higher than $W_{\min}$ = 0.1395, and the chi-square statistic χ^2 = 26.87 is higher $\chi^2_{\alpha,\nu} = 20.09$. Than the compatibility coefficient k_c, which is a measure of their ratio, which is marginally higher than 1 (Equation (7)):

$$k_c = \frac{W}{W_{min}} = \frac{\chi^2}{\chi^2_{\alpha,\nu}} = \frac{0.1866}{0.1395} = \frac{26.87}{20.09} = 1.338.$$

This result shows that the opinions of 18 experts in assessing the significance of the 9 criteria of railway rolling stock are aligned. The authors of the study expected a better concordance of expert opinions. Experts with a good knowledge of rail transport also differed in their views because they work in the areas assigned to them, thus being involved in the research related to these areas. Other issues that are not the expert's concern, for which they have little responsibility, become less important.

Each criterion average ranks $\bar{R}_i$ were used to calculate normalised subjective weights for these criteria using the ARTIW-L and ARTIW-N methods (Table 2). Calculated according to Equations (2), (8) and (9). The highest weights calculated by both expert evaluation methods are for criterion E $\left(\omega_E^{\text{ARTIW-L}} = 0.1580,\ \omega_E^{\text{ARTIW-N}} = 0.1813\right)$, which indicates the safety of passengers and crew. The weights for criterion D $\left(\omega_D^{\text{ARTIW-L}} = 0.1321,\ \omega_D^{\text{ARTIW-N}} = 0.1291\right)$, which is in the second place and which measures the ride comfort of passengers and train crew, are slightly lower.

The average opinion of the expert group shows that criterion F $\left(\omega_F^{\text{ARTIW-L}} = 0.0790,\ \omega_F^{\text{ARTIW-N}} = 0.0813\right)$, which assesses the range covered by a single electric charge or full tank of fuel, is the least important to them. Criterion I $\left(\omega_E^{\text{ARTIW-L}} = 0.0815,\ \omega_E^{\text{ARTIW-N}} = 0.0827\right)$, which assesses the dynamic performance of traction rolling stock, is also of little relevance to them.

The percentage weights given to the criteria by the experts, as calculated by the DPW method (Table 3), show that, in the same way as in the ARTIW-L and ARTIW-N methods, the highest weights are given to criteria E and D $\left(\omega_E^{\text{DPW}} = 0.1464,\ \omega_D^{\text{DPW}} = 0.1377\right)$. Criteria F and I $\left(\omega_F^{\text{DPW}} = 0.0727,\ \omega_I^{\text{DPW}} = 0.0906\right)$ are the least important for the experts.

Table 2. Significance in ranks R_{ij} of the indicators determining the serviceability of railway traction rolling stock

Expert, j = 1, 2, ..., n	Factor (criterion), i = 1, 2, ..., m									The sum of the ranks
	A	B	C	D	E	F	G	H	I	
E1	3	1	5	8	6	4	2	7	9	45
E2	7	8	3	2	1	5	4	9	6	45
E3	5	6	1	3	2	8	4	7	9	45
E4	3	5	4	2	1	7	8	6	9	45
E5	4	8	2	3	1	9	5	6	7	45
E6	4	8	7	1	2	6	5	9	3	45
E7	9	7	4	6	5	8	3	2	1	45
E8	6	9	7	5	1	3	8	2	4	45
E9	6	9	2	3	7	4	1	8	5	45
E10	9	7	4	2	1	6	8	5	3	45
E11	7	2	3	1	4	5	6	9	8	45
E12	1	7	8	2	3	6	5	9	4	45
E13	6	7	4	5	3	9	2	1	8	45
E14	5	7	8	6	1	9	4	2	3	45
E15	1	3	4	5	2	7	9	6	8	45
E16	2	7	4	8	3	6	5	1	9	45
E17	3	1	2	7	4	8	6	5	9	45
E18	1	8	7	4	5	6	2	3	9	45
$\sum_{j=1}^{n} R_{ij}$	82	110	79	73	52	116	87	97	114	810
$\bar{R}_i = \frac{\sum_{j=1}^{n} R_{ij}}{n}$	4.56	6.11	4.39	4.06	2.89	6.44	4.83	5.39	6.33	45
$\sum_{i=1}^{n} R_{ij} - \frac{n \cdot (m+1)}{2}$	–8	20	–11	–17	–38	26	–3	7	24	0
$\left(\sum_{j=1}^{n} R_{ij} - \frac{1}{2} \cdot n \cdot (m+1)\right)^2$	64	400	121	289	1444	676	9	49	576	3628
ARTIW-L $\omega_i = \frac{(m+1) - \bar{R}_i}{\sum_{i=1}^{m} \bar{R}_i}$	0.1210	0.0864	0.1247	0.1321	0.1580	0.0790	0.1148	0.1025	0.0815	1
Priority ARTIW-L	4	7	3	2	1	9	5	6	8	45
ARTIW-N $u_i = \frac{\min_i \bar{R}_i}{\bar{R}_i}$	0.63	0.47	0.66	0.71	1.00	0.45	0.60	0.54	0.46	5.516
$\omega_i = \frac{u_i}{\sum_{i=1}^{m} u_i}$	0.1150	0.0857	0.1193	0.1291	0.1813	0.0813	0.1084	0.0972	0.0827	1
Priority ARTIW-N	4	7	3	2	1	9	5	6	8	45

Table 3. Significance in percentage weights p_{ij} of the indicators determining the serviceability of railway traction rolling stock

Expert, j = 1, 2, ..., n	Factor (criterion), i = 1, 2, ..., m									The sum of the weights
	A	B	C	D	E	F	G	H	I	
E1	10.1	30.0	9.9	4.9	5.2	10.0	20.0	5.1	4.8	100.0
E2	6.0	5.0	15.0	20.0	24.0	9.0	10.0	4.0	7.0	100.0
E3	9.9	5.2	25.0	15.0	20.0	5.0	10.0	5.1	4.8	100.0
E4	14.1	10.0	14.0	16.0	17.0	7.1	7.0	9.9	4.9	100.0
E5	10.2	9.8	15.0	10.3	15.1	9.6	10.1	10.0	9.9	100.0
E6	10.1	5.0	5.1	30.0	15.0	7.9	10.0	4.9	12.0	100.0
E7	3.0	6.0	13.0	9.0	12.0	4.0	16.0	18.0	19.0	100.0
E8	9.0	7.8	8.0	9.2	25.0	10.1	7.9	13.0	10.0	100.0
E9	10.0	3.0	17.0	14.0	9.0	12.0	20.0	4.0	11.0	100.0
E10	4.0	6.0	13.0	18.0	21.0	7.0	5.0	10.0	16.0	100.0
E11	5.0	25.0	10.0	30.0	9.9	5.3	5.1	4.8	4.9	100.0
E12	23.0	6.0	3.0	18.0	14.1	7.0	12.0	2.9	14.0	100.0
E13	8.0	6.0	12.0	10.0	15.0	2.0	20.0	22.0	5.0	100.0
E14	10.0	7.5	6.1	8.5	20	5.9	12.0	16.0	14.0	100.0
E15	30.0	10.0	9.9	9.1	10.1	8.9	5.0	9.0	8.0	100.0
E16	20.0	6.9	10.0	6.0	12.0	7.0	8.0	25.0	5.1	100.0
E17	10.2	25.0	15.1	9.8	10.1	5.0	9.9	10.0	4.9	100.0
E18	20.1	7.9	8.0	10.0	9.0	8.1	17.1	12.0	7.8	100.0
$\sum_{j 1}^{n} p_{ij}$	212.7	182.1	209.1	247.8	263.5	130.9	205.1	185.7	163.1	1800
$\bar{p}_i$	11.817	10.117	11.617	13.767	14.639	7.272	11.394	10.317	9.061	100
ω_i	0.1182	0.1012	0.1162	0.1377	0.1464	0.0727	0.1138	0.1032	0.0906	1.0000
Priority DPW	3	7	4	2	1	9	5	6	8	45

Table 4. Comparison of the criteria weights ω_i, their maximum differences $\omega_{i\max} - \omega_{i\min}$ and their averages calculated by the different methods

Indicator	Factor (criterion), i = 1, 2, ..., m									Total
	A	B	C	D	E	F	G	H	I	
$\omega_i^{ARTIW\text{-}L}$	0.1210	0.0864	0.1247	0.1321	0.1580	0.0790	0.1148	0.1025	0.0815	1.0000
$\omega_i^{ARTIW\text{-}N}$	0.1150	0.0857	0.1193	0.1291	0.1813	0.0813	0.1084	0.0972	0.0827	1.0000
ω_i^{DPW}	0.1182	01012	0.1162	0.1377	0.1464	0.0727	0.1138	0.1032	0.0906	1.0000
$\omega_{i\max} - \omega_{i\min}$	0.0060	0.0155	0.0085	0.0086	0.0349	0.0086	0.0064	0.0060	0.0091	0.1036
$\bar{\bar{\omega}}_i$	0.1181	0.0911	0.1201	0.1330	0.1619	0.0776	0.1123	0.1010	0.0849	1.0000
Priority	4	7	3	2	1	9	5	6	8	45

The criteria weights decided by all 3 methods show almost the same priority of the criteria: E $\succ$ D $\succ$ C $\succ$ A $\succ$ G $\succ$ H $\succ$ B $\succ$ I $\succ$ F (Table 4, Figure).

The difference between the criteria weights $\omega_{i\max} - \omega_{i\min}$ calculated by the different expert evaluation methods varied from 0.006 to 0.0091 (criteria A, C, D, E, E, H, and I). This difference was higher for criteria B and E: 0.0155 and 0.0349, respectively. The purchase price of railway traction rolling stock, the rebate system and the length of the warranty period were more important to the experts in terms of percentage weights than in terms of ranks. The highest weight difference of 0.0349 between the different methods employed, is for the most important criterion E, as it has the highest significance $\omega_E^{ARTIW\text{-}N} = 0.1813$, which was calculated by the ARTIW-N method. This indirect method "upscales" the most important and least important criteria, while "downscaling" the moderately important criteria.

Although the experts opinions are harmonized (coefficient of compatibility k_c = 1.338), a higher degree of compatibility was expected. The railway transport specialists who assessed the significance of the criteria are usually responsible for a separate area in the company, for which the criteria they assess seem most important to them.

The outcomes of this study will be used in deciding on the best alternative for rail traction rolling stock. A future study will compare 4 traction alternatives: diesel, electric contact, battery-electric and hydrogen. Research data can be used to develop a strategy for the renewal of railway rolling stock in the country. JSC "LTG Link" specialists and managers are interested in them.

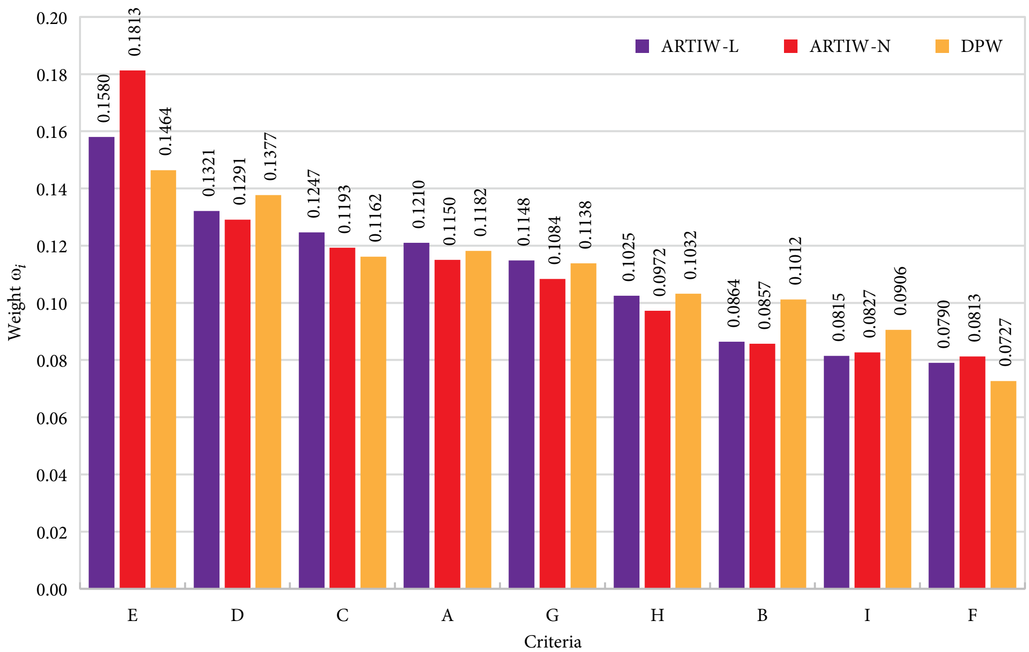

Figure. The subjective normalised weights and the priority of significance (serviceability indicators) of railway traction rolling stock, which have been decided by different methods

Conclusions

Conclusions of the carried-out research are:

1. to select the best rail traction alternative, firstly the factors (criteria), which determine the serviceability of rolling stock are identified. The article has formulated 9 original criteria, which in terms of ranks and percentage weights were assessed by 18 experts. To increase the reliability of the results, 3 expert evaluation methods (ARTIW-L, ARTIW-N, DPW) were employed to calculate each criterion subjective normalised weight. The experts' opinions in assigning ranks to the criteria were harmonised (non-contradictory), which allowed each criterion average ranks and average percentage weights to be used as the result of the issue under analysis;
2. the average normalised weights of the criteria, calculated by the 3 expert evaluation methods, show that passenger and crew safety (0.1619), passenger and train staff ride comfort (0.1330) and environmental protection (0.1201) are the most important issues for the experts. The least important criteria for the experts cover the range per one electric charge or full tank of fuel (0.0776), the dynamic performance of the traction rolling stock (0.0849), and the purchase price, the rebate system, the duration of the warranty period (0.0911). The other 3 criteria are of medium importance: They include operating costs (0.1181), durability and reliability (0.1123) and the cost of installing and maintaining the infrastructure (0.1010);
3. although the experts opinions are harmonized (coefficient of compatibility k_c = 1.338), a higher degree of compatibility was expected. The railway transport specialists who assessed the significance of the criteria are usually responsible for a separate area in the company, for which the criteria they assess seem most important to them. The opinions of the group of individual rail transport specialists did not always coincide. Therefore, when determining the significance of the criteria, it became necessary to attract specialists from various fields to the expertise.

References

Bajec, P.; Tuljak-Suban, D. 2020. A framework for detecting the proper multi-criteria decision-making method taking into account the characteristics of third-party logistics, the requirements of managers, and the type of input data, in F. Pedro Garcia Marquez (Ed.). *Application of Decision Science in Business and Management*, 1–16. https://doi.org/10.5772/intechopen.87222

Boghani, H. C.; Ambur, R.; Blumenfeld, M.; Saade, L.; Goodall, R. M.; Ward, C. P.; Plášek, O.; Gofton, N.; Morata, M.; Roberts, C.; Dixon, R. 2021. Sensitivity enriched multi-criterion decision making process for novel railway switches and crossings – a case study, *European Transport Research Review* 13: 6. https://doi.org/10.1186/s12544-020-00467-x

Čekanavičius, V., Murauskas, G. 2000. *Statistika ir jos taikymai*. I knyga. Vilnius: TEV. 240 p. (in Lithuanian).

Danwen, Y.; Guanglei, L.; Weixu, W.; Qingyu, W. 2021. Power quality pre-evaluation method considering the impact of electrified railway, in *2021 6th Asia Conference on Power and Electrical Engineering* (*ACPEE*), 8–11 April 2021, Chongqing, China: 1556–1559. http://doi.org/10.1109/ACPEE51499.2021.9436913

EC. 2011. *White Paper on Transport – Roadmap to a Single European Transport Area: Towards a Competitive and Resource Efficient Transport System*. European Commission (EC), Directorate-General for Mobility and Transport, Brussels, Belgium. 32 p. https://doi.org/10.2832/30955

Ehrhart, B. D.; Klebanoff, L. E.; Mohmand, J. A.; Markt, C. 2021. *Study of Hydrogen Fuel Cell Technology for Rail Propulsion and Review of Relevant Industry Standards*. Final Report No DOT/FRA/ORD-21/20. US Department of Transportation, Federal Railroad Administration, Office of Research, Development and Technology, Washington, DC, US. 20590. 59 p. Available from Internet: https://railroads.dot.gov/sites/fra.dot.gov/files/2021-06/Study%20of%20Hydrogen%20Fuel%20Cell%20Tech.pdf

García-Garre, A.; Gabaldón, A. 2019. Analysis, evaluation and simulation of railway diesel-electric and hybrid units as distributed energy resources, *Applied Sciences* 9(17): 3605. https://doi.org/10.3390/app9173605

Hamurcu, M.; Eren, T. 2022. Applications of the MOORA and TOPSIS methods for decision of electric vehicles in public transportation technology, *Transport* 37(4): 251–263. https://doi.org/10.3846/transport.2022.17783

Hoffrichter, A. 2013. *Hydrogen as an Energy Carrier for Railway Traction*. PhD Thesis. University of Birmingham, UK. 364 p. Available from Internet: https://etheses.bham.ac.uk/id/eprint/4345/9/Hoffrichter13PhD1.pdf

Hoffrichter, A.; Miller, A. R.; Hillmanser, S.; Roberts, C. 2012. Well-to-wheel analysis for electric, diesel and hydrogen traction for railways, *Transportation Research Part D: Transport and Environment* 17(1): 28–34. https://doi.org/10.1016/j.trd.2011.09.002

Kendall, M.; Gibbons, J. D. 1990. *Rank Correlation Methods*. 5th Edition. Oxford University Press. 272 p.

Li, C.; Xu, C.; Li, X. 2020. A multi-criteria decision-making framework for site selection of distributed PV power stations along high-speed railway, *Journal of Cleaner Production* 277: 124086. https://doi.org/10.1016/j.jclepro.2020.124086

Logan, K. G.; Nelson, J. D.; McLellan, B. C.; Hastings, A. 2020. Electric and hydrogen rail: potential contribution to net zero in the UK, *Transportation Research Part D: Transport and Environment* 87: 102523. https://doi.org/10.1016/j.trd.2020.102523

Madovi, O.; Hoffrichter, A.; Little, N.; Foster, S. N. Isaac, R. 2021. Feasibility of hydrogen fuel cell technology for railway intercity services: a case study for the Piedmont in North Carolina, *Railway Engineering Science* 29(3): 258–270. https://doi.org/10.1007/s40534-021-00249-8

Maskeliūnaitė, L. 2021. Railways in Lithuania: from tsarist Russia to Rail Baltica, *Transport* 36(4): 364–375. https://doi.org/10.3846/transport.2021.16086

Maskeliūnaitė, L.; Sivilevičius, H. 2021. Identifying the Importance of criteria for passenger choice of sustainable travel by train using ARTIW and IHAMCI methods, *Applied Sciences* 11(23): 11503. https://doi.org/10.3390/app112311503

Mastrocinque, E.; Ramírez, F. J.; Honrubia-Escribano, A.; Pham, D. T. 2020. An AHP-based multi-criteria model for sustainable supply chain development in the renewable energy sector, *Expert Systems with Applications* 150: 113321. https://doi.org/10.1016/j.eswa.2020.113321

Montgomery, D. C. 2012. *Introduction to Statistical Quality Control*. 7th Edition. Wiley. 768 p.

Peniwati, K. 2007. Criteria for evaluating group decision-making methods, *Mathematical and Computer Modelling* 46(7–8): 935–947. https://doi.org/10.1016/j.mcm.2007.03.005

Polater, N., Tricoli, P. 2022. Technical review of traction drive systems for light railways, *Energies* 15(9): 3187. https://doi.org/10.3390/en15093187

Rostamzadeh, R.; Esmaeili, A.; Sivilevičius, H.; Nobard, H. B. K. 2020. A fuzzy decision-making approach for evaluation and selection of third party reverse logistics provider using fuzzy ARAS, *Transport* 35(6): 635–657. https://doi.org/10.3846/transport.2020.14226

Ruf, Y.; Zorn, T.; Akcayoz De Neve, P.; Andrae, P.; Erofeeva, S.; Garisson, F. 2019. *Study on the Use of Fuel Cells & Hydrogen in the Railway Environment. Report 3: Overcoming Technological and Non-Technological Barriers to Widespread Use of FCH in Rail Applications: Recommendations on Future R & I*. Shift2Rail Joint Undertaking and Fuel Cells and Hydrogen Joint Undertaking. 50 p. Available from Internet: https://rail-research.europa.eu/wp-content/uploads/2019/04/Report-3.pdf

Şahin, M. 2021. Location selection by multi-criteria decision-making methods based on objective and subjective weightings, *Knowledge and Information Systems* 63(8): 1991–2021. https://doi.org/10.1007/s10115-021-01588-y

Shirres, D. 2020. Hydrogen trains coming soon?, *Rail Engineer*, 16 December 2020. Available from Internet: https://www.railengineer.co.uk/hydrogen-trains-coming-soon

Sivilevičius, H. 2011. Application of expert evaluation method to determine the importance of operating asphalt mixing plant quality criteria and rank correlation, *The Baltic Journal of Road and Bridge Engineering* 6(1): 48–58. https://doi.org/10.3846/bjrbe.2011.07

Sivilevičius, H.; Maskeliūnaitė, L. 2018. Multiple criteria evaluation and the inverse hierarchy model for justifying the choice of rail transport mode, *Promet – Traffic & Transportation* 30(1): 57–69. https://doi.org/10.7307/ptt.v30i1.2417

Sivilevičius, H.; Maskeliūnaitė, L. 2014. The numerical example for evaluating the criteria describing the quality of the trip by international train, *E+M Ekonomie a Management* 17(2): 73–86. https://doi.org/10.15240/tul/001/2014-2-006

Smith, K. 2020. Do hydrogen and battery trains mean the end for diesel traction?, *IRJ: International Railway Journal*, 2 April 2020. Available from Internet: https://www.railjournal.com/in_depth/hydrogen-and-battery-trains-end-for-diesel-traction

Sun, Y.; Anwar, M.; Hassan, N. M. S.; Spiryagin, M.; Cole, C. 2021. A review of hydrogen technologies and engineering solutions for railway vehicle design and operations, *Railway Engineering Science* 29(3): 212–232. https://doi.org/10.1007/s40534-021-00257-8

Šakalys, R., Sivilevičius, H., Miliauskaitė, L., Šakalys, A. 2019. Investigation and evaluation of main indicators impacting synchromodality using ARTIW and AHP methods, *Transport* 34(3): 300–311. https://doi.org/10.3846/transport.2019.9718