Traffic & Transportation

Visual and Statistical Methods for Analysing Irregularity Data in Rolling Stocks – An Application on Turkey's Freight Wagons Fleet

Ömür AKBAYIR[1], Serdar BENLİGİRAY[2], Ahmet ONAY[3]

Original Scientific Paper
Submitted: 24 June 2024
Accepted: 9 Oct 2024

[1] omurakbayir@eskisehir.edu.tr, Eskisehir Technical University, Vocational School of Transportation
[2] serdarbenligiray@anadolu.edu.tr, Anadolu University, Faculty of Business Administration
[3] Corresponding author, ahmetonay@eskisehir.edu.tr, Eskisehir Technical University, Vocational School of Transportation

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT

Maintaining freight wagons is an essential operational process and a significant cost factor for rail transport companies. Analysing detected irregularities in the freight wagons offers distinctive and valuable insights for planned maintenance. The primary purpose of this study is to provide various techniques to shed light on the characteristics of these irregularities and identify any interrelations between them. This study also reveals the general characteristics of Turkey's freight wagon fleets and maintenance depots in relation to the detected irregularities. New-generation visualisation tools, such as heat maps and chord diagrams, were utilised in this study. To determine the relationships between pairs of irregularities based on wagon type, irregularities with a high co-occurrence rate were identified and tetrachoric correlation analyses were conducted. In the final stage, Stepwise Poisson Regression Models were tested to explain the irregularities for each wagon type. The analysis techniques exemplified in this study were proven to reveal many interrelations between irregularities. The methods proposed in this study have the potential to provide crucial information for maintenance planning, parts supply and wagon repair processes. However, their practical application requires careful interpretation and detailed consideration by expert railway managers and engineers.

KEYWORDS

freight wagon; irregularity; maintenance; data visualisation; tetrachoric correlation; Poisson regression.

1. INTRODUCTION

Reducing operational costs is critical to an organisation's success in the current climate of demanding competition in the global economy. Maintenance and repair costs are among the most significant variables determining a company's operational expenses, accounting for 18% to 30% [1]. Maintenance and repair encompass all activities undertaken to maintain and repair equipment in compliance with standardisation and at an acceptable cost [2]. Maintenance and repair consist of a series of technical and administrative tasks to ensure the equipment properly performs its intended role [3]. Perils of poor practices of maintenance are significant, leading to high levels of unexpected financial losses. These failures include internal costs due to production loss, rework, scrap machines, depleted human resources, spare part shortages and delays or customer loss due to dissatisfaction [4]. Maintenance and repair costs directly impact the organisation's budget and profitability. Furthermore, the lack of proper maintenance and repair planning shortens the lifespan of the equipment [5, 6]. For expert railway managers and engineers, managing these costs and ensuring effective maintenance planning is paramount.

The maintenance of the rolling stock plays a crucial role in providing safe, reliable and competitive transportation services. Regular inspection and maintenance are essential to ensure network availability, reliability and operational efficiency. Maintenance activities include repairing or replacing specific components at predetermined time intervals or tonnage levels measured in millions of gross tons [7]. The rolling stock also undergoes preventive maintenance programs involving inspection, testing, adjustment, lubrication and cleaning of critical components at regular time or mileage intervals.

For a railway company, maintaining freight wagons is a significant cost factor [8]. Considering the entire lifecycle of railway freight wagons, repair and maintenance expenditures account for 25% to 50% of the investments [9]. A survey among private wagon owners revealed that annual regular maintenance and repair costs per wagon range from $800 to $10,000 [10, 11]. This cost variance is primarily influenced by factors such as the type of wagon, its age and the annual mileage covered. One emerging issue in managing maintenance and repair costs is the focus on short-term periods rather than the long-term impact of planning these operations. This short-term focus can hinder the optimal utilisation of the wagon fleet. Delays in maintenance and repair or the cancellation of operations due to these activities can result in additional costs for the company.

Wagons spend a significant portion of their downtime undergoing maintenance and repair activities in workshops, leading to overhead costs (e.g. rental costs) and variable costs (e.g. storage costs). Railway vehicles remain idle in the marshalling yard for about 70% of the time. Idle time incurs additional storage costs and implies that the theoretical capacity of the wagon fleet is underutilised. Furthermore, as the fleet size increases, the designed maintenance operations may lead to exponential inefficiencies due to the need to service and inspect many idle wagons [12].

Studies indicate that the maintenance of various components of railway freight wagons remains a significant issue for railway decision-makers [12, 13, 14, 15]. Freight wagons are a crucial part of the rail transport system, and their maintenance policies affect practical capacity, operational cost, freight rate and safety. The most commonly used maintenance policies are part replacement upon failure or periodic part replacement. Issues related to part replacement upon failure are associated with higher operational costs, failure costs and safety concerns. Determining the correct design life for component replacement is challenging regarding periodic part replacement. The same parts in different types of wagons may have different design lives due to varying patterns of functioning. Replacing a component under scheduled maintenance before its lifetime can lead to higher lifecycle costs due to unnecessary maintenance. Conversely, if the replacement period exceeds the component's lifetime, many components will fail before being replaced.

The increase in the use of rail transport has led to the establishment of more maintenance depots on the network and less time allocated for wagon maintenance to prevent operational disruptions [16]. An essential problem in rolling stock repair and maintenance planning is the allocation of appropriate maintenance depots/locations. Depending on the wagon type, each maintenance unit's varying competence and adequacy pose significant limitations. The type of irregularity is a critical factor that necessitates considering the competence of the maintenance unit. Providing service to a wide variety of wagons and failure types in each maintenance unit requires extra fixed asset investments and a workforce with diverse competencies.

Based on our research on the extant literature, wagon types, irregularity types and maintenance units have not been investigated in combination so far. In addition, how to analyse such integrated data with various techniques is yet to be exemplified. Determining the interrelated irregularities and depicting the most frequent irregularity/wagon types associated with maintenance units or scanning the wagon types and some specific irregularities in relation can be synergically beneficial for a rail transport company. The primary aim of this study is to provide multidisciplinary techniques to investigate these and, thus, to explore the relationships between wagon and irregularity types and maintenance units. This study also characterises the general outlook of Turkey's freight wagon fleets and maintenance depots in relation to the detected irregularities. The potential contributions of the study can be summarised as follows:

- The interrelationships between irregularity types will be identified. This will enable an understanding of which irregularities trigger each other. During maintenance, parts causing related failures can be inspected, thereby reducing maintenance and downtime.
- The most frequent irregularities for each wagon type will be determined. This will ensure that appropriately skilled maintenance personnel and repair equipment are readily available, contributing to the optimal utilisation of labour and material resources.
- Which wagon types and irregularities maintenance depots frequently encounter will be depicted. This will provide insights to facilitate the direction of freight wagons to the most suitable workshops for maintenance.

The next section of the study defines the data set and research methodology. The results are presented in the next section. The last section includes a conclusion and suggestions for future studies.

2. METHODOLOGY

This section presents information about the characteristics of the data set and the analysis in distinct sub-sections.

2.1 Dataset characteristic

This study analyses four years of freight wagon maintenance data of the Republic of Turkey State Railways Transportation from 2019 to 2022. The data set was obtained from the Vehicle Maintenance Management System module of the company's SAP SE Enterprise Resource Planning software. In this module, irregularities are stored according to the classification in Appendix 9: Conditions for the technical transfer inspection of wagons from the General Contract of Use for Wagons [17]. Appendix 9 classifies irregularities based on components into eight groups: 1. Running gear, 2. Suspension, 3. Brake, 4. Wagon under-frame and bogie frame, 5. Buffer and draw gear, 6. Wagon body, 7. Loads and containers, 8. Special items. The relevant appendix of this document defines roughly one thousand irregularity codes.

The data set consists of 80,057 irregularity logs of the freight wagons over four years. Before conducting the analysis, the data was refined to avoid outliers that have the potential to distort statistical inference.

Total data loss after all the refinements, enabling research objectives effectively, is approximately 6%. The refinement procedure and its effect on the data is given below:

- Data on wagon types with fewer than 500 irregularity reports were removed from the data set. Thus, the number of failures decreased to 77,955, and the number of wagon types decreased from 48 to 29. This enables an analysis focused on the most frequently used wagon types.
- Four-digit irregularity codes with fewer than 100 reports were transferred to higher-degree three-digit codes. Subsequently, the minimum number of reports of a particular irregularity was limited to 100 in the dataset. The number of irregularities decreased to 75,343, and the number of irregularity types decreased from 324 to 98. This enables an analysis focused on the most frequently encountered failure types.
- Records related to maintenance workshops with fewer than 100 repairs were removed from the data set. Thus, the number of irregularity reports decreased to 75,112, and the number of maintenance depots decreased from 46 to 36 in the dataset. This enables an analysis focused on the most active maintenance workshops.

2.2 Data analysis

Statistical analyses are mainly designed to test pre-hypothesised relationships, while data visualisation tools can effectively be used to understand the nature of big data better. This study utilises data visualisation tools to gain insights before conducting any statistical analysis. Considering the characteristics of the data, particular aspects were visualised using bar charts, chord diagrams, and heat maps to exhibit the potential information. Visualisation tools are a powerful first step in data analysis. Their findings only provide meaningful insight when they are validated and supported by statistical analysis. In this study, the findings of visualisation tools were examined in depth on a comprehensive data set with advanced statistical techniques.

A chord diagram is a circular visualisation of interconnected data, with connections represented along arcs and flows between them depicted as chords [18]. This fruitful visualisation technique has been used in various fields [19, 20, 21, 22]. In this study, chord diagrams provide insight into the relationships between pairs of irregularities, indicating which types of irregularities are related to each other.

Heat maps have become increasingly popular for visualising data-rich information in two and three dimensions, thanks to advances in fast data processing and visualisation software [23]. This rewarding visualisation technique has seen widespread use in studies across various scientific disciplines [24, 25, 26, 27]. This study used heat maps to visualise the irregularity reporting frequencies by the maintenance depots and wagon types. Additionally, bar charts were used to display irregularity frequencies by wagon type. All the visualisations were created using Microsoft Power BI.

The analysis data indicate which maintenance workshop repaired the identified irregularity types on freight wagons, categorised by wagon type. The data set variables include the frequency of irregularity types, repair frequency of maintenance workshops and repair frequency of wagon types. In order to discover relationships between variables, appropriate statistical analyses were designed based on the characteristics of the data set. The statistical data type is count data, precluding the use of popular statistical analyses. Accordingly, the most frequently reported irregularities during the analysis period were identified, with a count defined as reporting a specific irregularity at least once. Therefore, repeated reports of the same irregularity for the same wagon have no impact.

After these determinations, tetrachoric correlations were conducted to identify relationships between irregularities by wagon type. Tetrachoric correlation is a maximum likelihood technique for estimating the correlation between count or binary variables [28]. This method, a particular case of polychoric correlation, is typically used when both associated variables are not continuous [29]. This correlation method is especially suitable for analysing count or binary data (e.g. presence/absence of irregularities as in the current study) as it estimates the correlation of an underlying average latent vector [30]. Additionally, the method's effectiveness depends on the quality of the data, with higher sample sizes required to achieve accurate results [31]. Given the characteristics of our variables, this method is entirely suitable for identifying correlations between irregularities. Tetrachoric correlation is calculated as follows to estimate the correlation between two variables [32]. This formula estimates the relationship between the data based on distribution assumptions [33].

$$\rho = \mathrm{COS}\left(\pi/(1+\sqrt{(ad/b/c)}\right) \quad (1)$$

In this formula, COS represents the cosine function where a, b, c and d represent numerical values of a 2x2 matrix for combinations of possible pairs of binary variables. To exemplify, d is the frequency of the instance that is irregularity -1 and 2 occurring simultaneously, a is the reverse, and variable b is the frequency of the instance that is irregularity -1 occurring while irregularity -2 is not, and c is vice versa.

Subsequently, regression models were developed to explain the irregularities encountered by each wagon type. Since the analysis of the data followed a Poisson distribution, irregularities were analysed using Poisson regression to prevent information loss. Poisson regression is preferred over logistic regression in data sets of this nature due to its more robust variance error for obtaining incidence rate ratios (IRR) [34]. This methodology is used when the dependent variable is count data and is suitable for estimating the number of events occurring per unit of time or within a certain interval. The general form of a Poisson regression model is expressed as a logarithmic function of the independent variables [35]. Poisson regression is also used for rate data, such as the number of events per capita over a given time period, and is generally more sensitive than linear regression methods [36]. In the general form of the Poisson regression, log-rate is modelled as a linear function of the explanatory variables as given below:

$$\log(\lambda i)= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_k x_{ik} \quad (2)$$

λi represents the expected rate of events for the i^{th} observation (e.g. the number of irregularities in a wagon type for the analysis period). β_0 is the constant term of the model. β_1 β_2 β_3 ... β_k are the coefficients of the independent variables in the model. x_{i1} x_{i2} x_{i3} ... x_{ik} represents the independent variables in the model (e.g. type of wagon, types of irregularities).

At this stage, stepwise regression was employed as an exploratory analysis technique to help identify significant relationships through a regression model form. Stepwise regression is a technique used to determine the most significant independent variables in the model. This method selects the variables that provide the strongest relationship between the variables in the model and excludes unnecessary variables. In this context, the analyses conducted in this step are defined as stepwise Poisson regression. This approach allows us to determine the strongest relationships between types of irregularities.

Previous analyses treated count data as binary variables, as the count was defined as whether or not an irregularity was reported for each wagon, regardless of how many times it was reported. Poisson regression analyses substitute dense relationships caused by frequency with transparent relationships. In the stepwise regression model, each explanatory variable is a significant predictor of the response variable count, and the likelihood ratio chi-square tests demonstrate that the model as a whole is significant.

The IRR is an essential indicator between response and explanatory variables in regression models. The incidence rate ratio for a binary predictor variable is the ratio of events in one category to the number of events in the other. The IRR coefficient is interpreted relative to 1. An IRR>1 indicates that the explanatory variable increases the likelihood of higher counts. An IRR<1 indicates that the explanatory variable increases the likelihood of lower counts. An IRR=1 indicates that the explanatory variable has no effect.

3. RESULTS

The data for the analysis is statistically characterised as count data. The data set comprises three variables and exhibits large-scale data attributes. These variables are the type of irregularity, the type of wagon and the maintenance workshop. The analyses were designed to uncover correlations and regression relationships among the types of irregularity. Before proceeding to statistical analyses, the relationships among the three variables in the data set were visualised using chord diagrams and heat maps. The insights provided by the data visualisation tools reveal the most frequently encountered types of irregularity, the most commonly used types of wagons and the locations where the maintenance and repairs of freight wagons were conducted during the analysis period in Turkey.

The visual depicted in *Figure 1* illustrates the wagon types that experience the most frequent irregularities. In the figure, wagon types are arranged from left to right based on the magnitude of irregularity frequency. Additionally, the length of each wagon type on the X-axis is proportional to the number of wagons in the respective wagon type. Accordingly, the wagon type that experiences the highest average number of malfunctions per wagon is FALS (665 0 331/2708). The three most commonly used wagon types in freight transportation in Turkey are as follows: FALS (665 0 331/2708), KS (330 1 001/2650), and SGSS (456 8 923/9772).

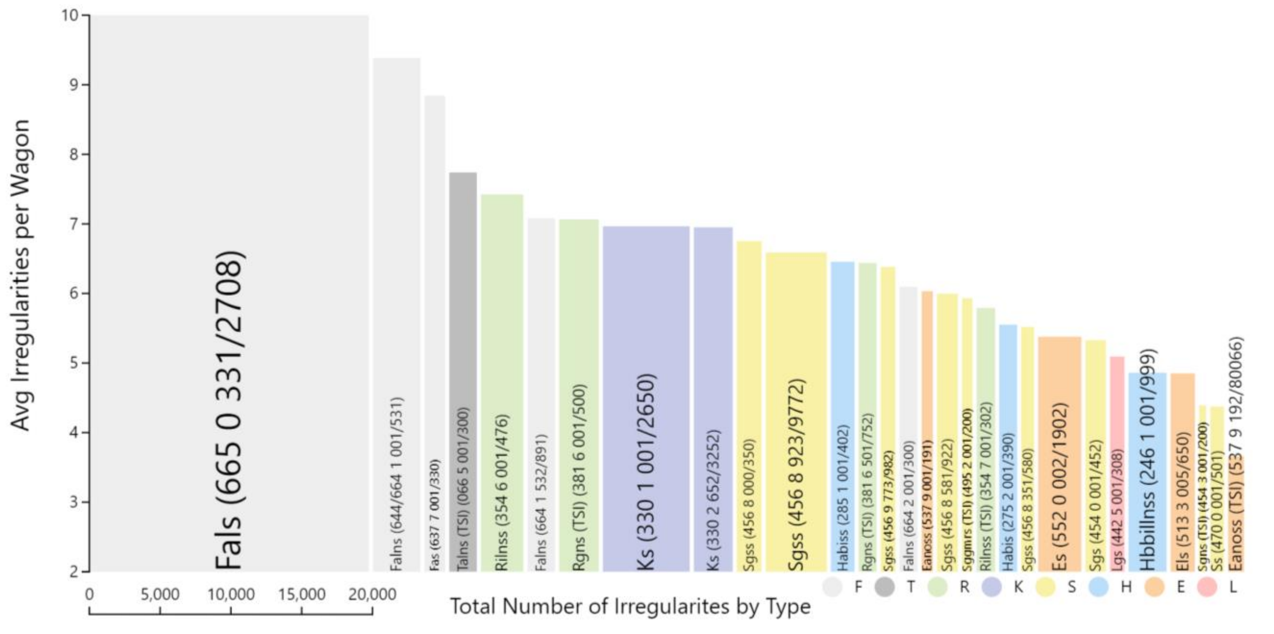

Figure 1 – Frequency bar chart for the irregularities by the wagon type

In *Figure 2*, the number of irregularities associated with freight wagon types and maintenance depots is visualised using a heat map. The X-axis of the heat map represents wagon types, while the Y-axis represents the locations where the freight wagons are repaired (maintenance depots). Each box where a row and column intersect is shaded according to the number of irregularities repaired for a particular freight wagon type at a specific location. The shading was applied based on the scale of repair counts provided next to the heat map.

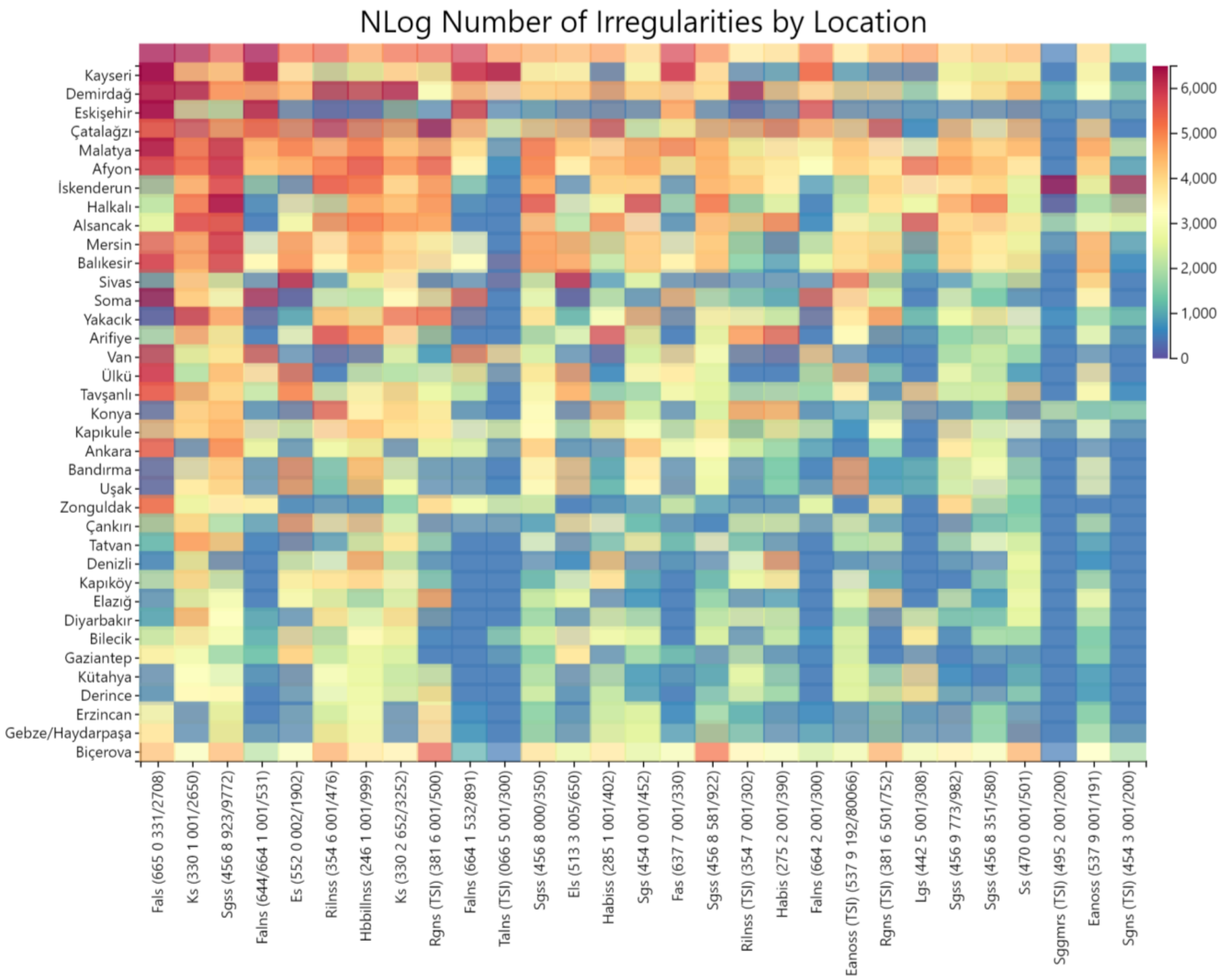

Figure 2 – Heat map for maintenance depots, wagon types and number of irregularities

The heat map visualises the distribution of 75,112 malfunctions across 29 types of freight wagons and 36 maintenance depots. The figure provides information on all irregularities concerning every freight wagon type and maintenance depots on the X and Y axes. For instance, FALS (665 0 331/2708) type freight wagons were most frequently repaired at workshops in Kayseri, Eskişehir, Malatya, Soma and Van. Conversely, workshops in Yakacık, Konya, Denizli, Kütahya and Derince had the lowest or zero repair frequencies for these freight wagons. Heat maps contribute to the identification of complex information that may be challenging to explore through traditional analyses with large data sets. For example, Iskenderun was identified as the most frequent location for repairing SGGMRS (TSI) (495 2 001/200) and SGNS (TSI) (454 3 001/200) type freight wagons.

Visualisation tools can be utilised to understand if there is any relationship between the types of irregularities identified in a particular type of freight wagon. To achieve this goal, chord diagrams showing whether there are statistical relationships between the identified types of malfunctions for each wagon type have been drawn. These diagrams subsequently guide the statistical analyses conducted. Diagrams were examined for each of the 29 freight wagons in the data set. For example, the chord diagram for FALS (665 0 331/2708) type freight wagons, which had the highest number of reported irregularities (attributed to the active number of wagons in use), is presented in *Figure 3*.

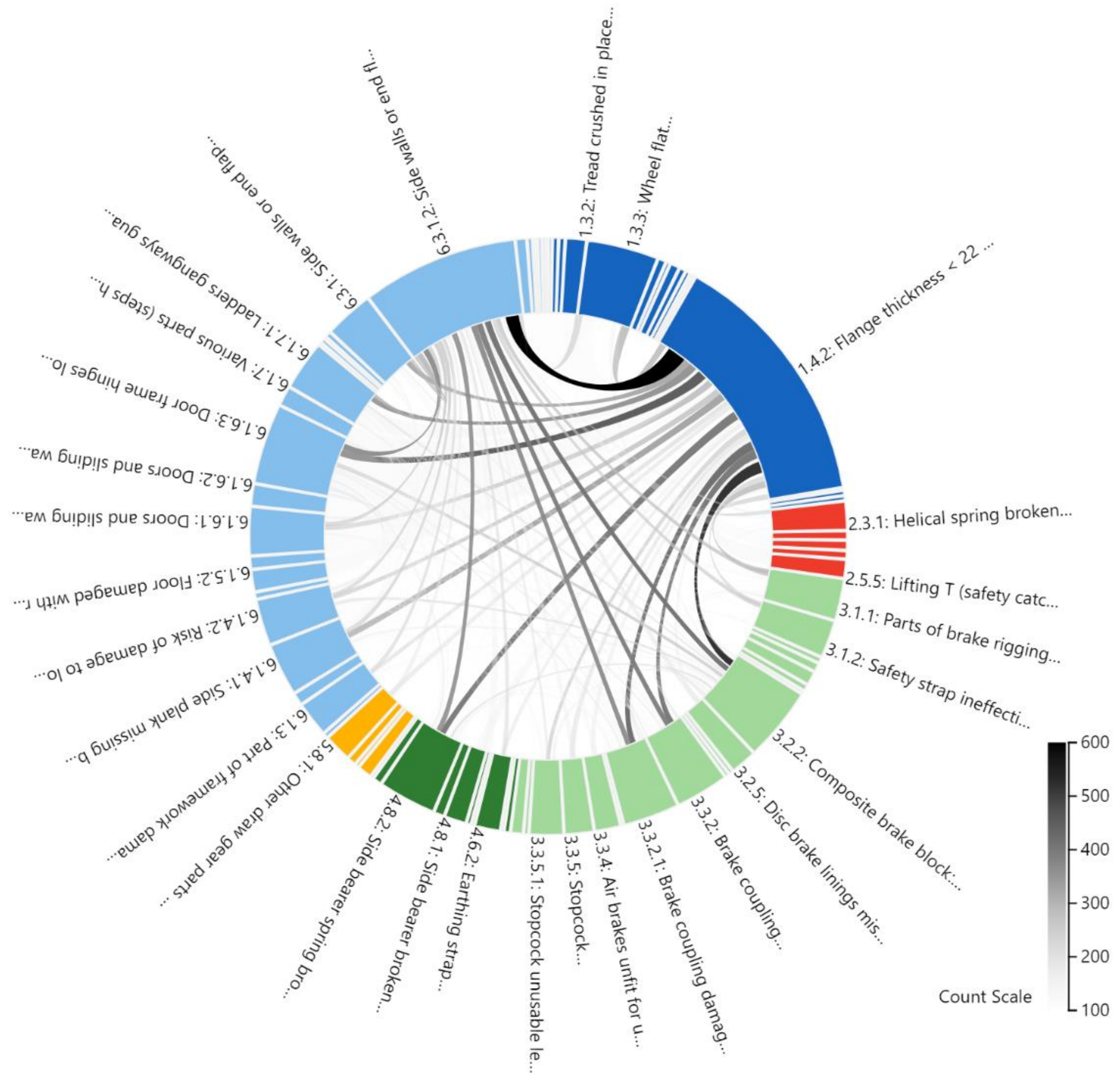

Figure 3 – Chord diagram of conjunct irregularities (paired irregularities)

Upon examining the chord diagram in *Figure 3*, it is evident that there are numerous relationships among irregularity types in FALS (665 0 331/2708) type freight wagons. Two pairs of irregularities appear to have relatively strong associations among these relationships. The first pair is between the irregularities ‘6.3.1.2 Side walls or end flaps damaged with the risk of losing load’ and ‘1.4.2 Flange thickness < 22 mm on wheel Ø > 840 mm or < 27.5 mm on wheel Ø 630 (330) – 840 mm or worn flange’. The second pair is between the irregularities ‘1.4.2 Flange thickness < 22 mm on wheel Ø > 840 mm or < 27.5 mm on wheel Ø 630 (330) – 840 mm or worn flange’ and ‘3.2.2 Composite brake block: missing, radial crack from friction surface through to plate edge (except at the designated expansion joint); the visible crumbling of the friction material over more than one-quarter of the block length, or metal inclusions – detached from back plate by more than 25 mm, cracked over more than 25mm in direction of wheel circumference, lowest thickness less than 10m.’

After examining the analysis data with visualisation tools, the most frequently reported irregularity pairs during the analysis period were identified without seeking causal relationships. Attention was focused on irregularity pairs with an observed conjunct reporting frequency of over 50. These examinations showed that numerous irregularity pairs were conjunctively reported in many types of freight wagons. Due to the size of the analysis data and the abundance of identified conjunct reporting numbers, the findings were presented with certain limitations in mind. Accordingly, irregularity pairs observed in each freight wagon type with a conjunct reporting rate exceeding 80% (which is very high) are presented in *Table 1*.

Table 1 – Irregularities with a high conjunct reporting rate

Freight Wagon Type	Irregularity I Code	Irregularity II Code	Obs. Count	Rate (%)
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm …	2.3.1 Helical spring broken	257	90.18
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm …	2.5.4 Spring cap(s) in contact with bogie frame	66	88.00
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm…	3.5.1 Hand brake clearly unfit for use	106	87.60
Fals 665 0 331 2708	6.3.1.2 Side walls or end flaps damaged	4.6.2 Earthing strap	223	86.10
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm…	4.8.1.1 Side bearer broken with no parts missing	79	84.95
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm…	2.5.5 Lifting T (safety catch) loose or missing	157	84.41
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm…	2.5.3 Damper ring(s) missing or broken…	71	82.56
Fals 665 0 331 2708	1.4.2 Flange thickness < 22 mm…	3.3.5.1 Stopcock unusable, leaking, warped…	268	80.48
Falns 644 664 1 001 531	1.3.3 Wheel flat	1.3.3.1 Wheel flat longer than 60 mm…	58	98.31
Falns 644 664 1 001 531	1.3.3 Wheel flat	1.3.2 Tread crushed in places of tire…	88	82.24
Falns 644 664 1 001 531	1.3.3 Wheel flat	6.3.1.2 Side walls or end flaps damaged…	76	80.85
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	6.2.2 Control gear, shutter retaining bracket…	53	96.36
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	3.2.2 Composite brake block…	73	94.81
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	6.3.1.2 Side walls or end flaps damaged…	146	90.68
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	3.3.2 Brake coupling	63	90.00
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	6.2.1 Ventilation flaps missing, damaged	158	89.77
Talns TSI 066 5 001 300	1.4.2 Flange thickness < 22 mm…	6.1.5.2 Floor damaged with risk of loss of load	73	89.02
Rilnss 354 6 001 476	6.6.1 Wagons with mechanical sheeting	4.6.2 Earthing strap	56	96.55
Rilnss 354 6 001 476	6.6.1 Wagons with mechanical sheeting	3.5.1 Hand brake clearly unfit for use	61	82.43
Rgns TSI 381 6 001 500	3.2.2 Composite brake block…	1.4.2 Flange thickness < 22 mm…	230	84.87
Ks 330 1 001 2650	6.4.3 Stanchion	6.4.4 Bolsters	306	94.15
Ks 330 1 001 2650	6.4.3 Stanchions	5.5.1 Buffer so slack…	64	90.14
Ks 330 1 001 2650	6.4.3 Stanchions	6.4.1 Drop sides	485	87.55
Ks 330 1 001 2650	6.4.1 Drop sides	6.4.4 Bolsters	274	84.31
Ks 330 1 001 2650	6.4.1 Drop sides	5.5.1 Buffer so slack…	58	81.69
Ks 330 1 001 2650	6.4.3 Stanchions	6.4.2 Hinges, pins, securing bolts missing…	156	81.68
Ks 330 2 652 3252	6.4.3 Stanchions	6.4.4 Bolsters	85	92.39
Ks 330 2 652 3252	6.4.3 Stanchions	6.04.2001 Drop sides	205	84.36

The findings presented in *Table 1* imply, for example, the following interpretation for row 1: For irregularities with code 1.4.2 reported in freight wagon type Fals 665 0 331 2708, the reporting rate of irregularities with code 2.3.1 is 90.18%, and the count of the conjunction is 257. Maintenance schedules for freight wagons can be planned based on the detected conjunct reporting rates of irregularity types. Additionally, maintenance or inspection procedures can be carried out for a freight wagon brought to a maintenance depot due to one of the irregularity pairs having high reporting rates.

After identifying irregularity pairs with high reporting rates, correlation relationships among irregularity types within the analysis data were investigated. Relationships above a certain correlation level were reported, considering the size of the analysis data. Specifically, correlation coefficients and levels were presented in *Table 2* within the context of tetrachoric correlation. In this context, moderate correlation corresponding to a 0.50 correlation coefficient and above and very high levels for tetrachoric correlations were deemed appropriate as the reporting threshold between irregularity pairs.

Table 2 – Correlation coefficients inferences

Correlation coefficients (Rho) lower limit	Correlation levels (Rho) upper limit	Correlation level inferences
0,90	1,00	very high
0,70	0,90	high
0,50	0,70	moderate
0,30	0,50	low
0,00	0,30	negligible

The entire data set was analysed to identify tetrachoric correlations between reported irregularities for each freight wagon. All correlation relationships above the level determined in *Table 3* (Rho>0.50) are reported.

Table 3 – Tetrachoric correlations between irregularities

Freight wagon type	Irregularity I code	Irregularity II code	Cor. coeff.	Sig. (p-)	Obs. count
Fals 665 0 331 2708	2.5.3 Damper ring(s) missing…	2.5.4 Spring cap(s) in contact…	0.693	0,000	29
Fals 665 0 331 2708	4.8.1.1 Side bearer broken…	4.8.2 Side bearer spring broken	0.552	0,000	67
Fals 665 0 331 2708	2.3.1 Helical spring broken	4.8.1.1 Side bearer broken…	0.525	0,000	47
Falns 644 664 1 001 531	1.3.3 Wheel flat	4.6.2 Earthing strap	0.634	0,000	42
Falns 644 664 1 001 531	1.3.2 Tread crushed in places…	1.3.3 Wheel flat	0.519	0,000	93
Falns 644 664 1 001 531	3.2.5 Disc brake linings missing…	6.1.6.3 Door frame, hinges…	0.508	0,000	33
Rilnss 354 6 001 476	4.6.2 Earthing strap	6.6.1 Wagons with mechanical…	0.614	0,000	56
Rgns TSI 381 6 001 500	1.4.2 Flange thickness < 22 mm…	3.2.2 Composite brake block…	0.515	0,000	230
Ks 330 1 001 2650	6.4.3 Stanchions…	6.4.4 Bolsters	0.681	0,000	306
Ks 330 1 001 2650	6.4.1 Drop sides	6.4.3 Stanchions…	0.660	0,000	485
Ks 330 1 001 2650	6.4.1 Drop sides	6.4.4 Bolsters	0.659	0,000	274
Ks 330 2 652 3252	6.4.1 Drop sides	6.4.3 Stanchions…	0.625	0,000	205
Ks 330 2 652 3252	6.4.3 Stanchions…	6.4.4 Bolsters	0.584	0,000	85
Hbbillnss 246 1 001 999	5.5.1 Buffer so slack…	6.1.6 Doors and sliding walls	0.578	0,000	23

When examining *Table 3*, it can be observed that moderate Tetrachoric Correlations were identified in seven freight wagon types across the entire analysis data set. The findings presented in the table imply the following interpretation, for instance, for row 1: The Tetrachoric Correlation Coefficient between '2.5.3 Damper ring(s) missing or broken, contact marks' and '2.5.4 Spring cap(s) in contact with bogie frame' is 0.6929 with a statistical significance level of 0.0000, and the count of the conjunction is 29. The identified correlations indicate statistically significant relationships between irregularity pairs. These findings can be utilised for maintenance and inspection purposes for a freight wagon brought to a repair workshop. Maintenance planning can be conducted based on these findings before any irregularities occur.

Significant relationships in tetrachoric correlation analysis may be attributed to a latent factor. In contrast to previous analyses, Poisson regression can explain causal relationships. These analyses trade-off concentrated relationships caused by frequency for crisp relationships. In this phase, stepwise regression is used as an exploratory analysis technique, which helps to define significant relationships through a regression model form. The R^2 value in regression analysis indicates the explanatory power of the model. The portion of the change in the response variable not explained by the model (1-R^2) is theoretically attributed to latent factors not included in the model. Numerous regression outputs were obtained from the analysis data. Regression models with high explanatory power were divided into two categories in the reporting phase. Regression models with an explanatory power of over 20% (R^2 > 0.20) and an observed count of response variables over 50 are reported in *Table 4*. In comparison, regression models with an explanatory power of over 17% (R^2 > 0.17) and an observed count of response variables over 45 are reported in *Table 5*.

The analysis data consists of a wide variety of random irregularities distributed across different wagon types. Many types of irregularities have been identified in the data set, classified according to the General Contract of Use for Wagons document. As previously mentioned, irregularities were analysed for all the irregularity combinations, and thus many regression analyses were conducted for these. Therefore, many of these are not presented in this study due to the threshold we set on R^2 for reporting. Considering the nature and characteristics of the analysis data, the R^2 thresholds established for the discovered regression models were deemed sufficient. R^2 values of Poisson regression models can often be relatively low due to the complexity and characteristics of the data, but this does not preclude the meaningful insights provided by the models [37]. Although the models resulting from Poisson regression analyses in the present study have relatively low R^2 values, they still provide significant insights.

Table 4 – Stepwise Poisson regression models (R2 > 0.2)

(a) Model 1 wagon type: **Fals 665 0 331 2708**	**Obs. count**: 57	**R^2:** 0.217
Response variable: 1.3.4.1 Metal build up over a length of > 60 mm…		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
1.3.2 Tread crushed in places…	2.028	0.000
1.3.4 Build-up of metal	3.093	0.000
1.3.4.2 Metal build up over a length…	5.557	0.000
3.2.2 Composite brake block…	2.098	0.000
1.2.2 Thermal overload due to braking…	3.281	0.003
4.8.1.1 Side bearer broken with no parts missing	2.282	0.004
Constant term	0.009	0.000

(b) Model 2 wagon type: **Falns 644 664 1 001 531**	**Obs. count:** 139	**R^2:** 0.251
Response variable: 1.4.2 Flange thickness < 22 mm on wheel Ø > 840 mm…		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
6.1.7.1 Ladders, gangways, guard rails…	1.972	0.001
3.5.1 Hand brake clearly unfit for use	2.164	0.000
3.3.2 Brake coupling	1.758	0.000
1.2.2 Thermal overload due to braking	3.846	0.000
6.1.4 Walls	2.989	0.000
2.5.5 Lifting T (safety catch) loose or missing	2.951	0.000
3.3.5.1 Stopcock unusable, leaking, warped…	4.933	0.000
2.4.2 Shackle, links displaced, missing, broken…	4.928	0.002
6.2.4.3 Opening roof control mechanism…	6.160	0.000
7.2.5 Direct or indirect fastenings…	21.413	0.000
1.3.8 Formation of grooves, hollows/furrows…	6.679	0.000
Constant term	0.159	0.000

(c) Model 3 wagon type: **Falns 644 664 1 001 531**	**Obs. count:** 53	**R^2:** 0.221
Response variable: 3.1.2 Safety strap ineffective		
Explanatory variables (predictors)	**IRR coefficient**	**Sig. (p-)**
4.6.2 Earthing strap	3.655	0.000
2.5.5 Lifting T (safety catch) loose or missing	3.347	0.001
6.3.1.2 Side walls or end flaps damaged…	1.891	0.000
6.3.1 Side walls or end flaps damaged	2.077	0.005
3.3.5 Stopcock	3.102	0.005
Constant term	0.049	0.000

(d) Model 4 wagon type: **Falns 664 1 532 891**	**Obs. count**: 504	**R^2:** 0.214
Response variable: 1.3.3 Wheel flat		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
3.2.5 Disc brake linings missing or cracked	1.395	0.000
6.1.7 Various parts...	1.828	0.000
3.1.5 Brake release pull broken or missing	2.927	0.009
1.3.1 Deviant width	0.219	0.000
6.1.6.1 Doors and sliding walls not fully closed	0.820	0.010
3.5.1 Hand brake clearly unfit for use	2.217	0.000
1.3.3.2 Wheel flat longer than 40 mm…	1.808	0.000
1.3.4 Build-up of metal	1.944	0.000
4.7.3 Bogie frame assembly, screw fastening…	0.307	0.000
6.1.6 Doors and sliding walls	2.754	0.000
3.3.1.1 Main brake pipe inoperative	3.728	0.000
2.5.5 Lifting T (safety catch) loose or missing	3.053	0.000
1.4.2 Flange thickness < 22 mm…	0.720	0.000
4.6.2 Earthing strap	1.817	0.000
6.1.1 Markings missing, illegible or incomplete	2.801	0.001
5.5.1 Buffer so slack that it can be de- pressed…	85.370	0.000
Constant term	1.379	0.000

(e) Model 5 wagon type: **Ks 330 1 001 2650**	**Obs. count:** 1026	**R^2:** 0.233
Response variable: 6.4.1 Drop sides		
Explanatory variables (predictors)	**IRR coefficient**	**Sig. (p-)**
6.4.3 Stanchions (detachable, pivoting, retractable)	1.300	0.000
6.4.3.1 Stanchions missing and necessary…	1.455	0.001
6.4.4 Bolsters	1.315	0.003
4.1.2 Solebar, headstock stressed by coupler…	1.363	0.000
3.3.1.1 Main brake pipe inoperative	2.795	0.001
Constant term	0.376	0.000

(f) Model 6 wagon type: **Ks 330 1 001 2650**	**Obs. count:** 1639	**R^2:** 0.231
Response variable: 6.4.3 Stanchions (detachable, pivoting, retractable)		
Explanatory variables (predictors)	**IRR coefficient**	**Sig. (p-)**
6.4.3.1 Stanchions missing and necessary…	0.640	0.001
6.4.4 Bolsters	1.245	0.006
7.1.1 Load visibly displaced…	0.258	0.001
5.6.1 Screw coupler non-operative	1.570	0.001
6.3.1 Side walls or end flaps damaged	1.436	0.001
6.4.1 Drop sides	1.470	0.000
Constant term	0.820	0.009

Table 5 – Stepwise Poisson regression models (0.20 > R2 > 0.17)

(a) Model 7 wagon type: **Fals 665 0 331 2708**	**Obs. count**: 129	**R^2:** 0.174
Response variable: 3.5.1 Hand brake clearly unfit for use …		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
6.1.4.2 Risk of damage to load due to humidity…	1.416	0.002
6.1.6.1 Doors and sliding walls not fully closed…	1.374	0.007
1.4.2 Flange thickness < 22 mm on wheel…	1.458	0.000
3.1.5 Brake release pull broken or missing	2.208	0.002
3.2.2 Composite brake block…	1.332	0.006
6.6.1 Wagons with mechanical sheeting…	1.974	0.003
6.1.5 Floor damaged	2.435	0.006
Constant term	0.015	0.000

(b) Model 8 wagon type: **Falns 644 664 1 001 531**	**Obs. count:** 69	**R^2:** 0.197
Response variable: 1.3.3.1 Wheel flat longer than 60 mm (wheel Ø > 840 mm)		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
2.3.1 Helical spring broken	1.962	0.000
1.3.3.2 Wheel flat longer than 40 mm…	3.261	0.000
6.1.7.1 Ladders, gangways, guard rails in poor…	2.063	0.010
1.3.2 Tread crushed in places…	1.510	0.000
6.1.6.2 Doors and sliding walls missing or derailed	1.639	0.008
Constant term	0.072	0.000

(c) Model 9 wagon type: **Falns 644 664 1 001 531**	**Obs. count:** 45	**R^2:** 0.261
Response variable: 2.3.1 Helical spring broken		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
1.3.3.1 Wheel flat longer than 60 mm…	3.721	0.000
1.3.4.2 Metal build up over a length of > 10 mm…	3.370	0.000
6.1.5.2 Floor damaged with risk of loss of load	5.859	0.000
6.1.4.1 Side plank missing, broken, split…	1.592	0.007
6.6.1 Wagons with mechanical sheeting…	26.843	0.001
Constant term	0.037	0.000

(d) Model 4 wagon type: **Falns 664 1 532 891**	**Obs. count**: 78	**R^2:** 0.178
Response variable: 3.3.4 Air brakes unfit for use but not labelled as such		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
1.3.2 Tread crushed in places of tyre…	1.943	0.000
5.7.1 Draw hook inoperative or in poor condition	3.658	0.006
3.1.1 Parts of brake rigging hanging down or broken	1.822	0.004
5.6.2 Hook for hanging screw coupler damaged…	3.699	0.001
3.3.1 Main brake pipe	3.047	0.000
6.4.4 Bolsters	3.622	0.000
4.6.2.1 One or more earthing straps ineffective…	6.501	0.010
Constant term	0.105	0.000

(e) Model 11 wagon type: **Falns 664 1 532 891**	**Obs. count:** 47	**R^2:** 0.245
Response variable: 3.1.2 Safety strap ineffective		
Explanatory variables (predictors)	**IRR coefficient**	**Sig.(p-)**
4.7.3 Bogie frame assembly, screw fastening…	3.885	0.000
1.2.2 Thermal overload due to braking…	7.382	0.001
1.3.3 Wheel flat	1.188	0.005
5.7.1 Draw hook inoperative or in poor condition	5.259	0.009
6.3.1.2 Side walls or end flaps damaged with risk…	1.404	0.010
6.3.1 Side walls or end flaps damaged	2.293	0.004
Constant term	0.032	0.000

(f) Model 12 wagon type: **Hbbillnss 246 1 001 999**	**Obs. count:** 45	**R^2:** 0.195
Response variable: 3.3.5 Stopcock		
Explanatory variables (predictors)	**IRR coefficient**	**Sig. (p-)**
6.2.2 Control gear, shutter retaining bracket…	5.023	0.000
3.3.1 Main brake pipe	5.286	0.006
1.8.3 Hot box	24.582	0.003
6.2.4.2 Opening roof derailed	7.546	0.001
Constant term	0.020	0.000

Table 4 presents the six regression models with the highest explanatory thresholds that are statistically significant. These models include explanatory variables that explain the response variable. The significance extracted from the regression model is the predictor of the portion of the change in the response variable explained by the model. For example, in Model 1, the explanatory variables explain 21.7% of the response variable (R^2: 0.217). Each explanatory variable is interpreted based on its IRR coefficient. For instance, the first explanatory variable in Model 1 is interpreted as follows: “The stepwise Poisson regression model explains ‘1.3.4.1 Metal build up over a length of > 60 mm...’ with ‘1.3.2 Tread crushed in places...’ and every reporting of ‘1.3.2 Tread crushed in places...’ changes the incidence rate (in this case, the number of irregularities) by 102.8% (IRR coefficient: 2.028).”

The models presented in *Table 5* in the second category also provide essential information about the causes of irregularities. The interpretation of these models is similar to those presented in the previous table. For example, in Model 7, the explanatory variables explain 17.4% of the response variable (R^2: 0.174). Each explanatory variable is interpreted based on its IRR coefficient. For instance, the first explanatory variable in Model 1 is interpreted as follows: “The stepwise Poisson regression model explains ‘3.5.1 Hand brake unfit for use’ with ‘6.1.4.2 Risk of damage to load due to humidity, risk of loss of load’ and every reporting of ‘6.1.4.2 Risk of damage to load due to humidity, risk of loss of load’ changes the incidence rate (in this case, the number of irregularities) by 41.6% (IRR coefficient: 1.416).” The explanatory variables in the model should be interpreted similarly. Other regression models presented in the tables should also be interpreted similarly when making maintenance plans.

4. CONCLUSION

This study conducted visualisation and statistical analyses of freight wagons actively used in Turkey. Visual tools provide information on which freight wagons encounter more irregularities and which maintenance workshops repair these irregularities. Additionally, with their statistical background, chord diagrams are powerful visual tools to show which irregularities are related to each other based on the type of freight wagons. Following the data visualisation phase, statistical analyses were conducted to identify the relationships between irregularities and to understand their causes. The high conjunct reporting rate irregularities and correlation relationships offer unique insights that can be utilised in freight wagons' maintenance and parts supply. In the final phase of the analysis, stepwise Poisson regression was employed to uncover causal relationships between irregularities. Numerous regression models were derived to explain the various irregularities encountered in freight wagons in Turkey. The reported regression outputs provide information on the most common faults in Turkish freight wagons. The models identify variables that explain the response variables. However, it should be noted that the models are limited by the R^2 value, which indicates how much of the variation in the response variable is explained by the model. The outcomes of this study, derived from a robust analytical framework, offer valuable insights for planning maintenance, parts supply and repair processes for freight wagons. Expert railway managers and engineers should carefully interpret and consider the results when developing these processes.

This study analysed irregularities according to the classification in the General Contract of Use for Wagons document. Standard codes allow for international comparisons and similar inferences for the same type of freight wagons in different countries. Similar correlation and regression relationships among irregularities in different data sets can be expected. Our study is geographically limited to Turkey, which might limit the generalisability of its findings. Additionally, the analysis does not account for operational variations between different railway companies or regions. Future studies can analyse irregularities encountered in freight wagons in different countries using this methodology. This study aims to explore a niche research area in the relevant literature. The causal and non-causal relationships between irregularities have been examined in the context of the characteristics of the analysis data. The research area opened by this study holds significant potential for future studies. Specifically, obtaining data sets containing information on the age of freight wagons or the kilometres travelled can help reduce the unexplained portion defined as latent factors in this study.

Although this study presents findings that can be considered a valuable input of information for long-term fleet optimisation, its primary focus is on demonstrating the positive impacts of short-term maintenance planning on the rapid and effective resolution of current irregularities. However, a short-term focus may have potential negative effects on long-term fleet optimisation. In order to optimise fleet performance in the long term, it is necessary to strategically plan maintenance processes and consider factors such as fleet aging, component wear and future cost monitoring. Future research should focus on better understanding these long-term factors and establishing a balance between short-term maintenance strategies and long-term fleet optimisation.

ACKNOWLEDGEMENTS

We would like to thank Erol Arıkan, deputy general manager of the State Railways Transportation Company of the Republic of Turkey, and Murat Durkan, head of the Vehicle Maintenance Department, for sharing the raw data set of this study with us.

REFERENCES

[1] Salonen A, Deleryd, M. Cost of poor maintenance: A concept for maintenance performance improvement. *Journal of Quality in Maintenance Engineering*. 2011;17(1):63-73. DOI: 10.1108/13552511111116259.

[2] Kour R, Tretten P, Karim R. eMaintenance solution through online data analysis for railway maintenance decision-making. *Journal of Quality in Maintenance Engineering.* 2014;20(3):262-275. DOI: 10.1108/JQME-05-2014-0026.

[3] Dinmohammadi F. A risk-based modelling approach to maintenance optimization of railway rolling stock: A case study of pantograph system. *Journal of Quality in Maintenance Engineering*. 2019;25(2) 272-293. DOI: 10.1108/JQME-11-2016-0070.

[4] Durazo-Cardenas I, et al. An autonomous system for maintenance scheduling data-rich complex infrastructure: Fusing the railways' condition, planning and cost. *Transportation Research Part C: Emerging Technologies*. 2018;89:234–253. DOI: 10.1016/j.trc.2018.02.010.

[5] Florian E, Sgarbossa F, Zennaro I. Machine learning-based predictive maintenance: A cost-oriented model for implementation. *International Journal of Production Economics*. 2021;236:108114. DOI: 10.1016/j.ijpe.2021.108114.

[6] Tusar MIH, Sarker BR. Maintenance cost minimization models for offshore wind farms: A systematic and critical review. *International Journal of Energy Research*. 2022; 46(4): 3739-3765. DOI: 10.1002/er.7425.

[7] Bevan A, et al. Optimisation of wheelset maintenance using whole-system cost modelling. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*. 2013;227(6):594-608. DOI: 10.1177/0954409713484712.

[8] Jaehn F, Rieder J, Wiehl A. Minimizing delays in a shunting yard. *Spectrum*. 2015;37:407–429. DOI: 10.1007/s00291-015-0391-1.

[9] Seliger G, Kross U, Buchholz A. Efficient maintenance approach by on-board monitoring of innovative freight wagon bogie. *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics 7-10 July 2003.* p. 407-411.

[10] Corsi T, Casavant K, Graciano T. A preliminary investigation of private railcars in north America author(s). *Journal of the Transportation Research Forum*. 2012;51:53-57. DOI: 10.5399/osu/jtrf.51.1.2811.

[11] Organisation for co-operation between railways committee. Agreement on use of freight wagons in international traffic (The PGV Agreement) 2022. https://en.osjd.org/api/media/resources/955?action=download [Accessed 20th Jan. 2024].

[12] Bigi F, et al. Long-term fleet management for freight trains: Assessing the impact of wagon maintenance through simulation of shunting policies. *Journal of Rail Transport Planning & Management*. 2024;29:100430. DOI: 10.1016/j.jrtpm.2023.100430.

[13] Muradian L, et al. Predictive model of risks in railroad transport when diagnosing axle boxes of freight wagons. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*. 2023;237(4):528-532. DOI: 10.1177/09544097221122043.

[14] Zavareh A, Fallahiarezoudar E, Ahmadipourroudposht M. Development of an optimized maintenance scheduling for emergency rescue railway wagons using a genetic algorithm: a case study of Iran railways company. *International Journal of Quality & Reliability Management*. 2023;40(6): 1540-1563. DOI: 10.1108/IJQRM-04-2022-0129.

[15] Xiao Q, et al. Wheel/rail-force–based maintenance interval extension of the C80 series wagon. *Railway Sciences*. 2023;2(4):514-524. DOI: 10.1108/RS-10-2023-0032.

[16] Zomer J, et al. The maintenance location choice problem for railway rolling stock. *Journal of Rail Transport Planning & Management*. 2021;20:100268. DOI: 10.1016/j.jrtpm.2021.100268.

[17] General Contract of Use for Wagons (GCU) Bureau. 2022. Technical report. https://gcubureau.org/wp-content/uploads/Contract/2022/20220101_GCU_EN_FULL.pdf [Accessed 20th Feb. 2024].

[18] Finnegan A, Sao SS, Huchko MJ. Using a chord diagram to visualize dynamics in contraceptive use: Bringing data into practice. *Global Health: Science and Practice*. 2019;7(4):598-605. DOI: 10.9745/GHSP-D-19-00205.

[19] Linial N, Nowik T. The expected genus of a random chord diagram. *Discrete & Computational Geometry*. 2011;45(1):161-180. DOI: 10.1007/s00454-010-9276-x.

[20] Jalali A. Supporting social network analysis using chord diagram in process mining. In Perspectives in Business Informatics Research: 15th International Conference, Prague, Czech Republic, September 15–16, 2016, Proceedings 15 (pp. 16-32). Springer International Publishing. DOI: 10.1007/978-3-319-45321-7_2.

[21] Koochaksaraei RH, Meneghini IR, Coelho VN. A new visualization method in many-objective optimization with chord diagram and angular mapping. *Knowledge-Based Systems*. 2017;138:134-154. DOI: 10.1016/j.knosys.2017.09.035.

[22] Lee YS, et al. Using chord diagrams to explore article themes in 100 top-cited articles citing Hirsch's h-index since 2005: a bibliometric analysis. *Medicine*. 2023;102(8):e33057. DOI: 10.1097/MD.0000000000033057.

[23] Babicki S, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic acids research*. 2016;44(W1):W147-W153. DOI: 10.1093/nar/gkw419.

[24] Zhao S, Guo Y, Sheng Q, Shyr Y. Advanced heat map and clustering analysis using heatmap3. *BioMed research international*. 2014. DOI: 10.1155/2014/986048.

[25] Netek R, Pour T, Slezakova R. Implementation of heat maps in geographical information system–exploratory study on traffic accident data. *Open Geosciences*. 2018;10(1):367-384. DOI: 10.1515/geo-2018-0029.

[26] Ryan MC, et al. Interactive clustered heat map builder: An easy web-based tool for creating sophisticated clustered heat maps. *F1000Research*. 2019;8. DOI: 10.12688/f1000research.20590.2.

[27] Bao W, et al. Mapping population distribution with high spatiotemporal resolution in Beijing using baidu heat map data. *Remote Sensing*. 2023;15(2):458. DOI: 10.3390/rs15020458.

[28] Debelak R, Tran US. Principal component analysis of smoothed tetrachoric correlation matrices as a measure of dimensionality. *Educational and Psychological Measurement*. 2013;73(1):63-77. DOI: 10.1177/0013164412457366.

[29] Maniu I, Costache C, Dumitraşcu DD. Adoption of green environmental practices in small and medium-sized enterprises: Entrepreneur and business policies patterns in Romania. *Sustainability*. 2021;13(9):49-68. DOI: 10.3390/su13094968.

[30] Grønneberg S, Jonas M, Njål F. Partial identification of latent correlations with binary data. *Psychometrika*. 2020; 85(4):1028-1051. DOI: 10.1007/s11336-020-09737-y.

[31] Schweizer K, Xuezhu R, Tengfei W. A comparison of confirmatory factor analysis of binary data on the basis of tetrachoric correlations and of probability-based covariances: A simulation study. *Quantitative Psychology Research: The 78th Annual Meeting of the Psychometric Society*. Springer International Publishing, 2015:273-292. DOI: 10.1007/978-3-319-07503-7_17.

[32] Bonett DG, Robert MP. Inferential methods for the tetrachoric correlation coefficient. Journal of Educational and Behavioral Statistics. 2005:30(2):213-225. DOI: 10.3102/10769986030002213.

[33] Digby PGN. Approximating the tetrachoric correlation coefficient. *Biometrics.* 1983:753-757. DOI: 10.2307/2531104

[34] Larsen LB, et al. Digital recruitment and acceptance of a stepwise model to prevent chronic disease in the Danish primary care sector: cross-sectional study. *Journal of Medical Internet Research*. 2019;21(1):e11658. DOI: 10.2196/11658.

[35] Warner P. Poisson regression. *Journal of Family Planning and Reproductive Health Care*. 2015;41(3):223-224. DOI: 10.1136/jfprhc-2015-101262.

[36] Cleophas TJ, Zwinderman AH, Cleophas TJ, Zwinderman AH. More on Poisson regressions: Poisson regressions with event outcomes per person or per population and per period of time. *Regression Analysis in Medical Research*. 2021:245-256. DOI: 10.1007/978-3-319-71937-5_14.

[37] Zeileis A, Christian K, Simon J. Regression models for count data in R. *Journal of Statistical Software*. 2008;27(8):1-25. DOI: 10.18637/JSS.V027.I08.

Ömür Akbayır, Serdar Benligiray, Ahmet Onay

Demiryolu Araçlarındaki Arıza Verilerinin Görsel ve İstatistiksel Analiz Yöntemleri: Türkiye'deki Yük Vagonları Üzerinde Bir Uygulama

Özet
Yük vagonlarının bakımı, demiryolu taşımacılığı şirketleri için önemli bir operasyonel süreç ve önemli bir maliyet faktörüdür. Yük vagonlarında tespit edilen arızaları analiz etmek, planlı bakım için farklı ve değerli içgörüler sunar. Bu çalışmanın temel amacı, bu arızaların özelliklerine ışık tutmak ve aralarındaki ilişkileri belirlemek için çeşitli teknikler sağlamaktır. Bu çalışma ayrıca, tespit edilen arızalarla ilgili olarak Türkiye'deki yük vagonlarının ve bakım depolarının genel özelliklerini ortaya koymaktadır. Bu çalışmada ısı haritaları ve akor diyagramları gibi yeni nesil görselleştirme araçları kullanılmıştır. Vagon tipine göre arıza çiftleri arasındaki ilişkileri belirlemek için, yüksek oranda birlikte görülme oranına sahip arızalar belirlenmiş ve tetrakorik korelasyon analizleri yapılmıştır. Son aşamada, her vagon tipi için arızaları açıklamak üzere Adımsal Poisson Regresyon Modelleri test edilmiştir. Bu çalışmada örneklenen analiz tekniklerinin, arızalar arasında birçok karşılıklı ilişkiyi ortaya çıkardığı kanıtlanmıştır. Bu çalışmada önerilen yöntemler, bakım planlaması, parça temini ve vagon onarım süreçleri için önemli bilgiler sağlama potansiyeline sahiptir. Ancak bunların pratikte uygulanabilmesi, uzman demir yolu yöneticileri ve mühendisleri tarafından dikkatli bir şekilde yorumlanmayı ve detaylı bir şekilde değerlendirilmeyi gerektirmektedir.

Anahtar Kelimeler
yük vagonları, arıza, bakım, bilgi görselleştirme, tetrachoric korelasyon, Poisson regresyon