Research Article

Paweł Dziedziak*, Tomasz Szczepański, Andrzej Niewczas, and Marcin Śleżak

Car reliability analysis based on periodic technical tests

https://doi.org/10.1515/eng-2021-0061
received November 17, 2020; accepted March 10, 2021

Abstract: This article presents the results of the analysis of data from 500 vehicle control stations concerning periodic tests of passenger cars in the number of 2 million. Based on the collected data, a database of vehicles has been prepared to cover technical information, such as model, mass, engine capacity and maximum power. This information has been summarized with the exploitation data such as age of the car, distance traveled, the area in which the vehicle was most frequently used, the use for business or private purposes, and intensity of operation that was measured by the ratio of kilometers traveled to the age of the car. This information constituted a set of information has been summarized with the exploitation data such as age of the car, distance traveled, the area in which the vehicle was most frequently used, the use for business or private purposes, and intensity of operation that was measured by the ratio of kilometers traveled to the age of the car. This information constituted a set of factors affecting the reliability of the car. Information about reliability itself was the results of control tests, including occurring defects and their significance. The occurring defects were codified, which allowed their allocation to particular categories describing car systems, such as steering, braking and driving. While the significance of the fault resulted from the test result: positive (minor defects) or negative (important defects). This article presents a factor and the correlation analysis of the database, which assesses which construction and operating factors have a significant impact on damages occurring in individual car systems.

Keywords: vehicles reliability, car exploitation, periodic vehicles tests

1 Introduction

Vehicle defects are a common problem related to, among others, road safety, logistics and transport economics. Hence, the possibility of predicting faults, even in statistical terms, can be very helpful in the rational management of car operation. This topic is often taken up in scientific papers on the reliability of technical objects [1–4].

However, there are a number of problems here, including difficulty in selecting a research sample that is representative enough to draw conclusions of a general nature. It is worth noting that the reliability of such complex objects, which are vehicles for road traffic, depends on a relatively large number of factors – not only from the design features of the vehicles themselves but also from the road conditions in which they operate, as well as from the manner of use imposed by the driver. This necessitates the multifactor analysis, which makes the selection of a research sample difficult. Consequently, the available research samples may be too small. On their basis, the methodology of the statistical analysis can be tested, but the conclusions drawn from such analysis are significantly limited [5–8].

One of the biggest challenges in the reliability analysis is forecasting the occurrence of faults in vehicles. All estimated calculations in this respect not only lead to increased road safety but also lead to improved traffic organization and minimized costs related to road transport [2,8,9]. The literature on the subject has many statistical and probabilistic tools that allow for an estimated prognosis of the occurrence of defects. To maximize the accuracy of these tools, a relatively large statistical sample is needed, which faces the difficulties discussed earlier [5,6,15]. It is worth following at least a few new studies into how to predict vehicle faults and faults in general.

Reference [10] presents a model of the car operation system in the form of a stationary stochastic process. It allows to determine probabilistic indicators of the current and border readiness of the system. The semi-Markov process was used to model the vehicle operation system. On the example of a sample of police cars, experimental
tests of the intensity of the use of the fleet and the vehicle uptime were carried out, and it was shown that the tested transport system is characterized by a satisfactory, stationary readiness factor.

Reference [9] compares the operational efficiency of city buses at different stages of their useful life. Based on the research, the operational efficiency characteristics were determined. The new division of the time of use proposed in the work made it possible to determine the most rational limit of vehicle use until the end of its operation.

Reference [11] presents methods and applications of stochastic modeling for simple damage streams of vehicles, the model of which is the Poisson process. Evaluation of the parameters of the damage processes of individual parts of vehicles is performed based on the operational database of a homogeneous fleet of vehicles operated for 5 years. The operational database is dynamically updated with new events over time.

The comparisons of readiness based on the Markov model with continuous time carried out in the previous study [12] showed differences between the model with and without the failure of occurrence interval. Furthermore, when considering the threshold values, it is important to balance the maintenance costs and lifetime of multistate components.

In ref. [13], the method of dynamic evaluation of fuzzy operational readiness (availability) of a multi-state system operating in the mode of minor damage and repairs was investigated. In this study, the theory of fuzzy sets was used. The fuzzy Markov model with discrete time and fuzzy transition probability matrix was used to analyze the fuzzy probability of the state of each element at any discrete time.

The publication [7] focuses on the assessment of the dynamics of the rotor production line, which is part of the electric motor production system. It is assumed that degradation is a heterogeneous Markov process with continuous time. Application research was carried out based on the life cycle data of seven stations. In addition, operating parameters were developed for both individual workstations and the whole system.

Reference [14] presents an analysis of the costs of work and maintenance of fleets. An optimization model was proposed that combines the traditional life cycle cost analysis and the Monte Carlo simulation model.

Research of this type is carried out in the form of theoretical considerations or based on a database covering a selected fleet of vehicles. As a result, such a database is limited in size.

Recently, however, there is the possibility of using a database of unprecedented coverage, storing information about the technical condition of cars in the country. This information is obtained from vehicle control stations. New research material available can be used to predict reliability features. Importantly, the number of entries in the database corresponds to the number of periodic vehicle tests, which in Poland reaches 40 million a year. This is the solution to the problem of insufficient data set size. The presented database operates within the CEPIK 2.0 system, which has been working since the beginning of 2018. So far, the material collected in this database has not been used for scientific analysis. The research presented here will propose an algorithm analyzing data from periodic vehicle diagnostic tests, which was used to predict defects in cars.

The following information is available in the CEPIK 2.0 database:
- Selected vehicle technical data, such as curb weight, axle loads, engine type, fuel.
- Selected vehicle registration data, including place and date of registration and place and date of first registration.
- Date and place of the diagnostic test.
- Information on the detected defects, including information about the successful or unsuccessful result of diagnostic tests, information about the observed major malfunctions, remarks regarding minor malfunctions.
- Other information such as road course, manufacturer and model of the vehicle.

The aforementioned data constitute comprehensive information not only about the current state of the vehicle but also about its history and conditions of use. For example, a combination of first place and current registration information shows whether cars were imported into the country as a used vehicle. The place of registration allows to conclude about the geographical area in which the car is most often used. The test location will allow conclusions about the specifics of driving, climate or the character of the local road infrastructure. The date of the first registration combined with the road course of the car allows to draw conclusions regarding the intensity of use. Some assumptions can also be made on the basis of engine capacity and power because the dynamic driving capabilities of the vehicle can be reflected in the nature of its actual operation. The manufacturer and the model as well as selected technical parameters can inform about design features that affect reliability. Some of the aforementioned information may have a relatively small
correlation, for example, high engine power does not always mean a dynamic way of driving. However, in a sufficiently large statistical sample, it will be relatively easy to verify this type of relationship.

This article will present the results of preliminary tests showing the methodology of data analysis and preliminary results, covering a research sample with 10,000 diagnostic tests.

2 Analyzed material and test method

Centralna Ewidencja Pojazdów i Kierowców (The Central Register of Vehicles and Drivers), shortcut: CEPIK, is a database in which information about vehicles registered in the country as well as drivers and candidates for drivers is stored. The second, extended version of this database, which entered into force in 2019, also includes information on periodic technical inspections of vehicles. Therefore, it is possible to use this information to analyze the reliability of vehicles.

Technical inspection data are sent to CEPIK by vehicle inspection stations. In this dissertation, the data on periodic tests coming directly from the control stations were used. These are the same data that are sent to the central database. This direct way of obtaining the results of experimental research in this study was the most advantageous because CEPIK, as a newly introduced IT system, still contains some errors. Information provided directly by vehicle inspection stations was divided into the so-called data packages, which facilitated their processing and possible elimination of errors at the IT level.

The amount of information on periodic technical inspections available for research purposes includes about 1,000,000 tests each year, which come from nearly 1,000 vehicle inspection stations. They contain data on all diagnostic tests performed in the country. For practical reasons, it was decided to limit this number due to the technical possibilities of data processing.

Data from vehicle control stations were collected in a database. Each entry in the database represented one diagnostic test, described by the following features:
- Order number
- Test date
- Initial characters of vehicle registration number
- Vehicle producer
- Vehicle type and model
- Vehicle identification number (VIN)
- Date of first registration in the country
- Date of first registration in general
- Vehicle type
- Vehicle category
- Type of fuel
- Type of test performed (periodic or additional)
- Test result (positive or negative)
- Date of next test
- Defects found (in the form of a list of fault codes)
- Vehicle road course

Based on this information, the derived information has been determined, such as
- The age of the car expressed in years (based on the difference between the date of the test and the date of first registration)
- Information about the fact of importing the vehicle from abroad (based on the discrepancy between the date of the first registration and the date of the first registration in the country)
- Information on the significance of the faults (based on the code describing the faults found)
- Intensity of use (as the quotient of road course expressed in kilometers and the age of the car expressed in years, thanks to which the average annual mileage is determined)

The database prepared in this way was equipped with a filtering tool. For discontinuous changing features (e.g., the type of vehicle), the filter allows to choose the interesting value. For constantly changing features, the filter allows you to select a range of interesting values. In this way, it was possible to create histograms and their analysis in the form of searching for correlations between selected vehicle features and the frequency of the identified defects during diagnostic tests. The results of this analysis are presented in Section 3.

The analyzed research sample included a total of 10,3016 vehicles, divided into the following types:
- 9,157 passenger cars, which is 88.8% of the total
- 1,010 trucks, which is 9.8% of the total
- 94 motorcycles, which is 0.9% of the total
- 19 special cars, which is 0.2% of the total
- 18 mopeds, which is 0.2% of the total
- 7 other motor vehicles, which is 0.1% of the total

Among the vehicles mentioned, 62.7% are vehicles powered by gasoline and 37.3% powered by diesel oil. Forty-nine percent of all vehicles were purchased domestically, and 51% were purchased abroad (the so-called individual imports).
In the analyzed sample, 2019 defects were found. The majority (1,758) were significant defects, causing a negative result of the diagnostic test. The remaining 261 faults were marked as small, without causing a negative result.

The analysis of the results consists of three steps:
- Analysis of the road mileage of vehicles
- Analysis of the relative age of vehicles
- Analysis of the intensity of vehicle use

The following elements were used in the first two stages:
- Histogram for the selected field of analysis
- Failure frequency for all analyzed vehicles
- Failure frequency for vehicles divided into vehicle fuel type
- Failure frequency for vehicles divided by their origin.

The last stage was limited to:
- Histogram for the vehicle operation intensity
- Failure frequency for all analyzed vehicles

In this case, a more detailed analysis of the intensity of exploitation was abandoned because the preliminary analysis showed a relatively small correlation between the frequency of faults and the intensity of exploitation. Therefore, this area of analysis is considered less important for the presented research.

The calculations were made in the MS Excel. The described database was copied to the program sheet in a tabular form. For the calculations, both formulas built into the program package as well as own formulas were used.

The next lines of the sheet corresponded to the cars from the database. Conversely, the following columns of the sheet contained the appropriate vehicle parameters:
- Entry number
- Vehicle brand (e.g., FIAT)
- Vehicle type and model (e.g., Punto)
- Type of vehicle (e.g., passenger)
- Type of fuel
- Test result (positive or negative)
- Fault number
- Defect number decoding (major faults, minor faults or no faults)
- Vehicle mileage
- Vehicle age in the country
- Overall age of the vehicle
- Information on the origin (imported or purchased in the country)
- Intensity of use

The fault codes were, for example, as follows: UD: 4.2.1.c, where the second character indicated the significance of the fault. Formula (1) categorizes the defects into major or minor defects:

$$= \text{IF} \left( \text{MID} (H11; 2; 1) < "\"; \text{MID} (H11; 2; 1); "B" \right).$$ (1)

The information about the origin of the vehicle used the following formula:

$$= \text{IF} \left( K8 = L8; "N"; "S" \right).$$ (2)

The intensity of use was calculated as the ratio of the vehicle mileage to its age.

Above the following columns:
- Vehicle type
- Type of fuel
- Test result
- Fault decoding
- The origin of the vehicle

there was a place to enter the desired values against which the database was filtered.

Also above the following columns:
- Vehicle mileage
- Overall age of the vehicle
- Intensity of use

there was a place to enter the desired value ranges against which the database was filtered.

To perform the filtering operation, the database table contained one more column containing information whether a given vehicle met the filtering criteria. If the criteria are met, the value 1 appears in the cell, otherwise the value 0.

The following formula is used to verify that the filtering criteria are met:

$$= \text{IF} \left( \text{AND} \left( \text{OR} \left( E3 = "\"; E3 = E7 \right); \text{OR} \left( F3 = "\"; F3 = F7 \right); \text{OR} \left( G3 = "\"; G3 = G7 \right); \text{AND} \left( J7 \geq J8; J7 \leq J8 \right); \text{AND} \left( L7 \geq L8; L7 \leq L8 \right); \text{OR} \left( M3 = "\"; M3 = M7 \right); \text{OR} \left( I3 = "\"; I3 = I7 \right); 1; 0 \right) \right).$$ (3)

Above the column with the assessment of meeting the filtering criteria, there were cells that calculate

- the sum of the values in the column (which was the sum of the vehicles meeting the filtering criteria) and
- percentage of the number of eligible vehicles out of all vehicles.

The filtering tool prepared in this way allowed to find the following information from the database:

- The number of vehicles in a given age range, mileage or intensity of use (information for creating histograms)
– The number of vehicles with defects in a given age range, mileage or intensity of use
– Combinations of the aforementioned filter criteria with the fuel type or vehicle origin criterion applied simultaneously.

Data for the dependencies presented in the charts were prepared using the aforementioned tool. The assumed values of age or mileage ranges and additional assumptions for filtering the results were recorded in a tabular form, and then, these data were entered into the filtering tool and the result was recorded in the created table. Based on this, it was possible to prepare charts.

The tools built into MS Excel were used to approximate the dependencies such as:
– Setting a trend line
– Displaying the trendline equation
– Calculation of the coefficient of determination $R^2$

### 3 Research results

Figure 1 shows the histogram of a road vehicle course. It shows that the majority of vehicles subjected to a diagnostic test at a vehicle control station was within the range of 100,000–250,000 km.

For each road course from Figure 1, the frequency of faults was determined as the share of the number of faults relative to the number of cars in a given range. The frequency of faults in the function of the road course is shown in Figure 2.

As shown in graphs in Figures 1 and 2, the share of defects in the vehicles tested increases, almost in the whole range, along with the road course. This is in line with expectations because this share corresponds to the probability of occurrence of faults, which increases with the generalized operating time of the vehicle (in this case, the generalized operating time is the road course). At the same time, the uncertainty of the result increases with the value of the road course (deviation of results relative to the regression line increases). This phenomenon is associated with a decrease in the number of vehicles in subsequent ranges of the road course.

Figure 3 shows the relationship between the share of vehicles with faults among vehicles in a given road range: for vehicles fueled with gasoline (red line) and diesel (black line). In both cases, the percentage value of the share reaches 25%, while for gasoline-powered cars, the function increases slightly earlier (with lower road course values) and then decreases (with high road course values and small numbers of vehicles taken into account in the research). Meanwhile, for diesel cars, the
share of faults remains relatively low (around 5%) up to a road course value of 200,000 km, followed by a sharp jump in the share of faults to the value of 20–25%. For the subsequent values of the waveform, this level remains at an almost constant value.

Figure 4 presents the dependence of the share of vehicles with faults among vehicles in a given road course range: for preused vehicles (red line) and purchased in the country (black line). It should be noted that the share of defects in the tested vehicles reaches higher values for vehicles purchased domestically. It reaches 30% for routes over 200,000 km and 40% over 400,000 km. Meanwhile, vehicles purchased abroad rarely have a fault share above 20%. The exceptions are vehicles with low mileage, below 50,000 km, for which the share of defects in imported vehicles clearly exceeds that of vehicles purchased domestically.

Figure 5 shows the histogram of the age of the vehicles analyzed. Its shape is similar to the histogram of the waveform (from Figure 1), which seems understandable. Figure 6 presents the dependence of the share of vehicles with faults among vehicles in a given age range.

It is worth noting that vehicles older than 10 years have a relatively constant value of the share of defects at the level of 15%. Compared with the analogous relationship, but in the field of road course (Figure 2), it can be stated that the probability of faults is less correlated with the age of the vehicle than with its road range.

Figure 7 shows the relationship between the share of vehicles with faults among vehicles in a given age range for vehicles powered by gasoline (red line) and diesel oil (black line). Both relationships show a relatively high agreement, in both value and their distribution in the field of care age.

Figure 8 shows the dependence of the share of vehicles with faults among vehicles in a given age range for vehicles purchased domestically (black line) and imported (red line). The differences in the values and
the shape of these functions are similar to those of ana-
logous functions in the road course domain (Figure 4),
although in this case, they are less visible. Similarly, cars
bought domestically have higher levels of faults.

An additional area has also been prepared in which
the occurrence of defects is worth considering. It is the
intensity of vehicle use, expressed in the average annual
road course. Figure 9 shows the histogram of vehicle use
intensity. The histogram shows that the most common
average annual road course is between 5,000 and
20,000 km. In turn, the intensity of exploitation over
25,000 km per year is relatively rare.

Figure 10 shows the dependence of the frequency of
faults on the intensity of operation. This function does
not seem to be significantly correlated with the intensity
of operation because the function visible on the graph
increases initially and decreases after exceeding the
value of 20,000 km. This means that the probability of
occurrence of defects will be influenced by the road
course of the vehicle and its age rather than by the inten-
sity of use.

An analysis of the most common defects was also
carried out. The following is a list of faults whose share
among the total number of faults exceeded 1%:
- Excessive wear of the pivot, pivot bearings or rocker
  pins
- Excessive looseness in steering rod joints (joints)
- Damage to the light source illuminating the rear regis-
  tration plate
- The brake does not work on one side or the vehicle
  pulls excessively to the side
- The shock absorber is damaged and shows large leaks
  or damage
- Excessive leakage of any fluid that may threaten the
  environment or the safety of other road users
- No clear border of light and shadow of the passing
  beam
- Damage to the “stop” light source
- Excessive corrosion affecting the rigidity of the chassis
  structure or the frame and components attached thereto
- The SRS self-diagnostic indicator indicates any type of
  failure in the airbag system
- The tire tread depth not in accordance with the
  requirements of the regulation on technical conditions
- Damage to the light source of the front, rear and side
  position lamps or marker lamps
- The emission of gaseous pollutants exceeds the max-
  imum values specified in the regulation on technical
  conditions petrol engine
- No continuous increase in braking force (sticking) in
  the service brake system
- Efficiency less than the minimum values contained in
  the regulation on technical conditions in the service
  brake system

Figure 8: Dependence of the share of vehicles with faults among
vehicles in a given age range for vehicles purchased domestically
(black line) and imported (red line).

Figure 9: Histogram of vehicle use intensity.

Figure 10: Dependence of the share of vehicles with faults among
vehicles in a given range of intensity of exploitation.
The light beam does not hit the license plate surface
- Excessive delay in the operation of the brakes on any wheel in the service brake system
- Lack of required braking performance in the parking braking system
- The OBD system displays emission related fault codes (gasoline engine)
- Insufficient braking force on at least one wheel of the auxiliary (emergency) braking system (if present as a separate system)
- Damage or cracking of the suspension spring
- Damage to the light source illuminating the registration plate
- The condition of the brake pipes may cause breakdown or breakage
- Clearly leaking exhaust system (in the context of noise)
- Insufficient braking force on at least one wheel in the service brake system
- Insecure attachment of a front, side, rear position lamp or end-outline marker lamp
- A warning indicator indicates a defective antilock braking system (ABS)
- No light, source of light or damage to it (applies to driving beam or passing beam)
- Efficiency less than the minimum values contained in the regulation on technical conditions in the auxiliary (emergency) braking system (if present as a separate system)
- Exceeding the value of gaseous emissions, referred to in the Regulation on technical conditions (compression-ignition engine)
- Uncertain attachment or poor condition of the floor in the vehicle
- Damage to the reversing light source
- Excessive wear of brake pads or linings
- The brakes do not work on one side in the parking braking system
- Excessive wear of steering joints connection
- The convergence value of wheel alignment exceeds the permissible limits for a given type of vehicle
- No document issued by the transport technical inspection and a data plate on the tank or cylinder, confirming its efficiency
- Cracks or discoloration of the glass or transparent pane or plastic (if allowed)
- Defective front, side, rear position lamp or end-outline lamp
- No emission control device installed, modification of the device or clearly incorrect operation (gasoline engine)
- Leaks affecting the measurement of exhaust emissions (gasoline engine)

The aforementioned faults are listed in order from most to least frequent. As it can be seen, they can be divided into several categories, referring, for example, to individual vehicle systems, such as braking system, steering system or vehicle lighting. Defects corresponding to legal irregularities were also registered, such as the lack of a document issued by the transport technical inspection.

4 Summary

The presented test results are compatible with expectations and indicate that the frequency of defects increases with the age of the car and its road course. The degree of correlation of this relationship increases significantly with the size of the test sample, expressed in the number of vehicles subjected to the diagnostic test. This confirms the credibility of the test results.

The frequency of faults among the vehicles tested in the considered groups is more strongly correlated with the mileage of vehicles than with their age.

In the field of the intensity of operation, the occurrence of faults in vehicles reaches the highest values for average intensity of 15,000–20,000 km/year.

Comparison of the occurrence of faults in petrol- and diesel-fueled vehicles indicates a higher failure rate of gasoline vehicles in the field of the vehicle road course, but this phenomenon is not observed in the field of vehicle age.

Comparison of the occurrence of defects in vehicles purchased domestically and purchased abroad and imported indicates a higher failure rate of vehicles purchased domestically in the case of vehicles with a mileage of over 100,000 km.

These observations show that the operating time is not the only significant factor conditioning the occurrence of vehicle defects. Both construction and functional factors are also important. The impact of these factors on faults will be analyzed in further research planned by the authors.

Conflict of interest: Authors state no conflict of interest.

References

[1] Dziedziak P. Problematyka merytoryczno-prawna związana z badaniami technicznymi w kontekście funkcjonowania Biura Patronatu ITS nad SKP – wybrane zagadnienia (Substantive and legal issues related to technical research in the context of
the functioning of the MTI Patronage Office over SKP – selected issues). Wydawnictwo Konferencyjne Badania Techniczne w Świecie Obowiązujących Przepisów; 2013.


