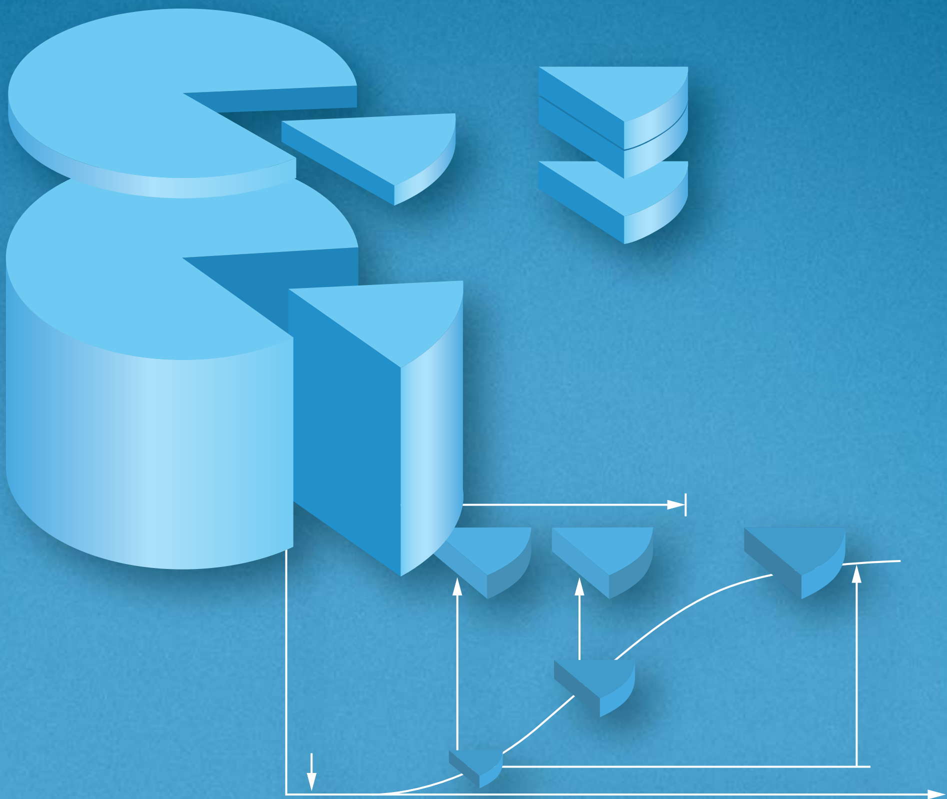
2021-05-05 - SOCOS



Quality Management in the Bosch Group | Technical Statistics

# 6. Evaluation of **Field Data**

BOSCH
Invented for life

# Quality Management in the Bosch Group
# Technical Statistics

## No. 6

## Evaluation of Field Data

### Edition 3.2021

**Edition 3.2021**

# Table of Contents

# 1 Introduction

High product quality is one of our company's top priorities. Assuring product reliability, as an integral part of our quality effort, requires long-term planning. Quality targets, including reliability targets, are defined at the very beginning of the design process for new products and are often subject to binding agreements with automotive manufacturers.

The durability testing of our products, which is carried out in parallel with the design and manufacturing process, is defined in accordance with these targets. However, final proof that the quality targets have been achieved requires the observation of product quality in the field. This includes the collection and analysis of warranty data. Automotive manufacturers and suppliers use statistics derived from this as control parameters in their overall quality control loop.

This publication describes procedures that yield suitable statistics for assessing product reliability in the field, which can be used for quality control and for comparative supplier assessments.

> *NOTE: For the sake of simplicity and in accordance with common usage, the word "time" is used in this text to represent all possible lifetime characteristics which are specific for a damage mechanism. Typically, this involves information such as time / operating time (hours, months, years), covered distance, mileage, number of load changes, actuations, switching operations, work cycles, revolutions. So it can be measurable or countable characteristics.*

In this context, field data refers to the totality of all data that is generated in the field in connection with the use of a product. In a narrower sense, this includes all data associated with errors, faults, defects and failures that lead to customer complaints.

The first edition of this booklet used the term "field data" in this narrower sense almost exclusively to refer to complaint and failure data, i.e., the more negative aspects of the topic for all stakeholders. This continues to be the focus of the Chapters 2 to 7.

In a broader sense, however, this also includes information on usage such as operating hours, driving times, consumption, load collectives or customer feedback in the form of evaluations, experiences, wishes and suggestions for improvement. The focus here is therefore on a more future-oriented basic attitude and preventive aspects. This topic is addressed in particular in Chapter 9.

It is not easy to distinguish between these positions and points of view. Ultimately, all the methodological approaches and activities presented should serve the customer benefit and contribute to customer satisfaction. Nevertheless, the following focal points can be identified in this booklet.

- Chapters 2 to 7: Basics, visualization and simple statistical approaches
- Chapters 8 and 9: Systematic collection of field data with respect to product usage
- Chapters 10 to 12: Evaluations based on the Weibull distribution

## 1.1 Goals of the Field Data Evaluation

The standards ISO 9001 and IATF 16949 contain numerous requirements that directly or indirectly involve customer complaints and field failures, particularly in the chapters on:

- Risk analyses
- Customer communication
- Problem solving
- Customer complaints and field failure test analysis
- Nonconformity and corrective action

Independently of this, however, the field observation and analysis of field data is also in the company's own interest, e.g. for the purpose of

- Early detection of problems due to design defects, manufacturing defects, defective supplier parts (early warning system)
- Estimation of costs from field complaints; reporting on external defect costs
- Statistical modelling and prognosis of the further failure behavior and corresponding costs development (necessary accruals)
- Estimation of the product reliability and derivation of design changes
- Investigation of the correlation of test and field data to optimize the reliability assurance of future products
- Derivation of strategies/recommendations regarding warranty and maintenance periods

As a rule, there are only a few response options:

- Software update (over-the-air update for IoT-enabled products)
- Customer service action
- Recall

*NOTE: A documented procedure for the processing of customer complaints according to ISO 9001 and IATF 16949 does not cover any product monitoring obligations in the context of product liability.*

*LITERATURE NOTE: [VDA Rekl], [VDA OtA]*

## 1.2 Failure Analysis and Problem Solving

The failure analysis (diagnosis, part analysis) should provide an unambiguous statement about the returned product's technical functionality. The result should also provide information as to whether a complaint can be acknowledged or not. [VDA Field] assumes that the product complained about can be examined for this purpose as part of a standard test or load test. However, there are also cases in which the input of electrical, mechanical or hydraulic energy leads to complete internal destruction of the product.

First ascertainable indications on the product are e.g.

- the location: Where on the part is the fault located?
- the failure mode: What on the part is affected?

From a technical point of view, in the broadest sense, it is about determining why a part has failed. The analysis is a kind of detective work to trace what happened and to explain what cause(s) finally led to the failure.

In this context, terms are often used such as failure or damage mechanism, damage or failure process and failure model. What is meant in all cases is a physical, chemical, elektrochemical or other process, which leads or led to a failure.

Examples are usually time-dependent processes such as material fatigue, crack formation, migration, or corrosion. Spontaneous failures due to overload is once apart.

[VDA 3.2] defines damage mechanisms as “processes that lead to a gradual change of a unit's properties due to stresses”.

The goal of analysis is the systematic study of a fact and its causes. Corresponding methodical approaches are described in [Booklet 16]. The focus here is in particular on cause-effect relationships of products, components and production processes. The booklet also contains a procedure for problem solbing in case of product problems.

Where appropriate, the methods described in this Booklet No. 6 can support the collection of facts and help answer the W-questions: what, where, when, who, how much/many?

## 1.3 Problems with the Statistical Evaluation

The high expectations for the possibilities of field data evaluation presented in Section 1.1 are often contrasted by severely limiting framework conditions.

If abnormalities in the field behavior of a product are detected early, there are only few data available. At the same time, all interest groups understandably expect forecasts on the further course of the failure behavior that are as reliable as possible, even at this early stage. However, it should be clear to everyone that large uncertainties are inevitable in statistical modeling and analyses based on small amounts of data.

From a statistical point of view, this means in particular that confidence intervals will then turn out to be very large. A possibly necessary division of the few existing data, e.g. because of competing failure mechanisms or product variants, aggravates the situation even more.

From a statistical point of view, this means in particular that confidence intervals will then turn out to be very large.



The opposite case, i.e. a rapidly increasing number of failures, at best pleases the statistician, as the large data basis makes his work easier. For the manufacturer, however, this means having to take care to quickly identify the root causes, isolate the affected products, and apply effective problem solving mehods.

In diagrams which represent a time axis as an abscissa (horizontal axis), the data points furthest to the right correspond to the maximum lifetime reached by a product of the data set. Extrapolations beyond this time to derive statements about the behavior of the products to even larger times should be regarded with skepticism. This also applies in particular to products that have not been in the field for long.

In general, field data are limited regarding operating time or mileage of the products. Even with generous warranty promises from vehicle manufacturers, there is usually a limit to a maximum mileage. In addition, the warranty for certain vehicle components may be reduced in time or excluded. As a rule, the manufacturer does not receive any information about failures that occur after the warranty or guarantee period, unless he has his own breakdown service.

However, such services usually require fee-based warranty extensions or breakdown coverages. Different warranty periods and country-specific laws thus complicate the evaluation of field data.

Provided that no failures occurred on the products of a given production period until a time t after commisioning, only a statement about the minimum reliability based on the success-run principle is possible [Booklet 13].

## 1.4 Data Basis

For the evaluation of field data, the following information is required, for example, depending on the question:

- Date (PD in unencoded form, dd.mm.yyyy) month or quarter of production
- Number of pieces produced this month, production quantity (products, vehicles)
- Purchase date of the product or registration date of the vehicle
- Failure date / complaint date
- Time in service, operating time / mileage until failure (in time units or km/mls)
- If applicable, mileage distribution of the considered population

It is usually advantageous, to order the data corresponding to the production date.

[VDA Field] also lists some information which can be relevant for the error analysis and to narrow down affected products, e.g.

- Vehicle data, like engine and transmission variants, special equipment, e.g. trailer hitch
- Operating conditions, e.g. any specific conditions in the country in question and climatic particularities, fuel quality



## 1.5 Data Quality

Experience has shown that data analysis must be preceded by a review and adjustment of the data basis. In evaluations of data related to passenger cars, the following observations were made:

- The data set contains 0-km failures
- the date of sale is before the date of production, the “storage time” is negative
- poduction date and failure date are identical; operating time is zero
- in a few cases unusually high mileages, e.g. more than 169,000°km/year
- single implausible data, e.g. 2 km in 7 months for a registered car
- also possible: very long storage times before commissioning, e.g. more than one year
- products of differend product classes are mixed (e.g. failure data of 2-, 4- and 6-cylinder engines in the same data set.)
- different measurement units, e.g. in 1,000 km instead of km
- Input errors, transmission errors, typing errors, missing data

*NOTE 1: 0-km oder 0-h failures correspond to lifetime zero. They cannot be described in the context of a Weibull or Lognormal distribution ($log(0) \rightarrow -\infty$).*

*NOTE 2: It cannot be completely ruled out that data sets contain statistical outliers because products are used outside the intended use or are misused, e.g.*

- *Use of vehicles on race tracks or in racing operations on public roads*
- *Use of the brake pedal for bodybuilding when the vehicle is stationary*
- *Continuous operation of do-it-yourselfer tools in the professional area*
- *Overloading of electrical equipment due to installation of incorrect/unsuitable tools*

# 2 Basics

When dealing with field data and possible risks in the field, it is important to mention very basic concepts and terms to prevent misinterpretation and misunderstanding. Some of everyday language has to be sharpened by defining the exact meaning in this context. This will be the topic of this chapter.

## 2.1 Field Data and Field Failures

As already described in the Chapter "Introduction", field data comes in two flavours: field loads and field failures.

In short, field loads are about datasets where the loads have been collected, that a specific fleet has undergone over time in the field, but independent on any failures that occurred.

These field loads are of special importance when it comes to mathematically model the failure behavior over time in the field to make forecasts. A deeper insight into the process of systematically collect such information is provided in Chapter 9.

Field failure data, however, is always in the scope of a concrete issue, where the quality does not meet the expectations in field. This data is an important source for quality evaluation, but also legal regulations request the monitoring of field data. Therefore, it is a matter of course for the BOSCH Group to have the appropriate processes installed to deal with field data.

To understand field data, its quality, and its coverage, it is necessary to understand the steps that are in between a failure in the field and the knowledge of it for the BOSCH Group.

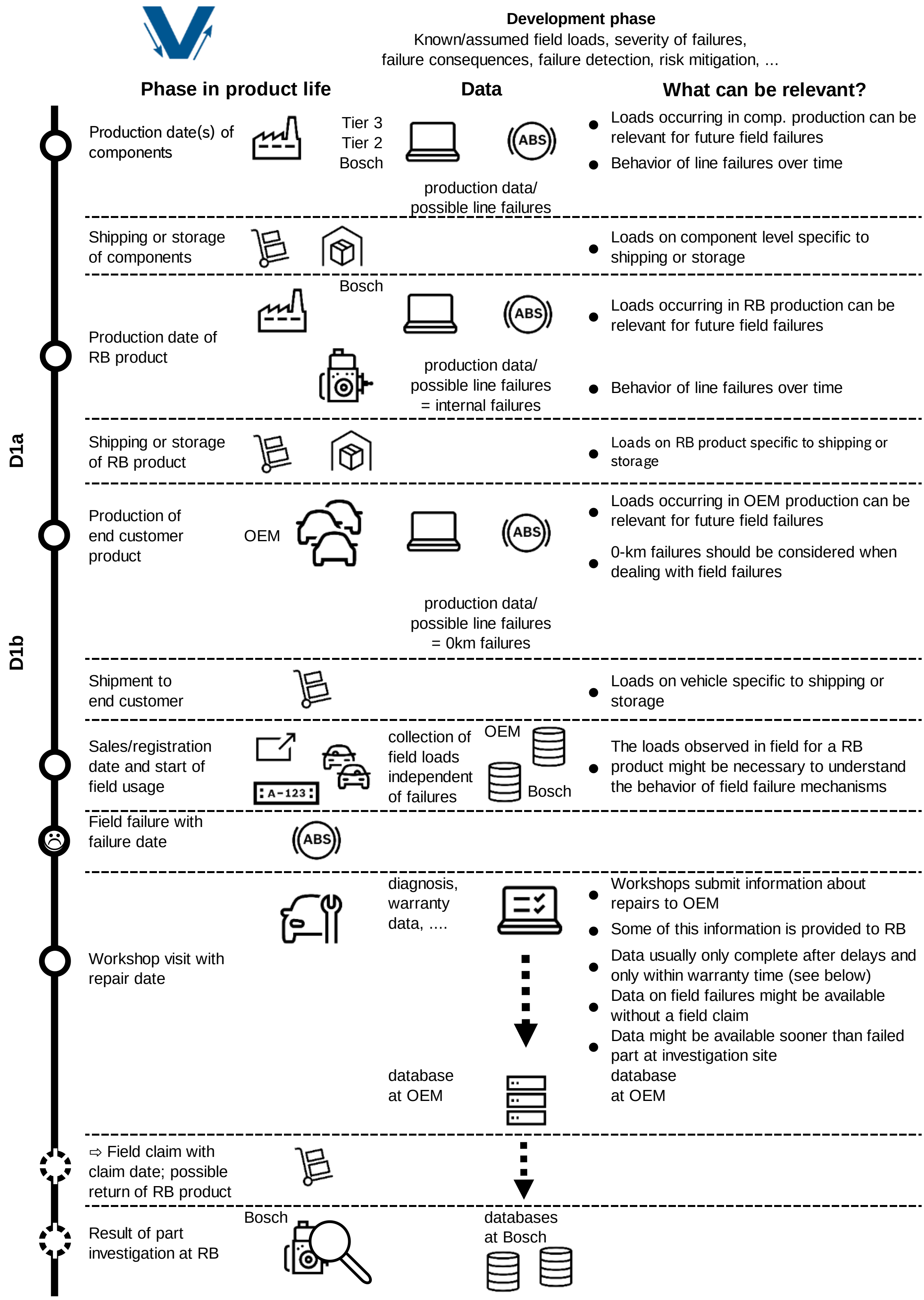
At first, the customer takes the car to a workshop due to failure. Dependent on the failure and the urgency for a fix, this will already obscure the real failure time. Two things have to come together there: the workshop has to find the failure correctly and it has to report it in the OEMs database.

Second, for the part to return to the OEM and/or the BOSCH Group, the workshop has to take part in the OEM's process to return failed parts. For many customers of the BOSCH Group, special agreements are in place that only parts from specific markets are passed further on for investigation to the BOSCH Group. After such an investigation has been finished, a root cause can be assigned to the specific failure, if this was not possible before by any other means like vehicle diagnostic data (OBD data).

Therefore, it might make a big difference what "kind" of field failure data is used and one has to understand their shortcomings: For many automotive customers of the BOSCH Group, it is possible to evaluate directly failure data from customers' warranty or other field databases. However, here, the root cause is often unclear and even Bosch responsibility for the failure might not necessarily be clear and confirmed.

## 2.2 Lifecycle of a Product in the Context of Field Data

The typical lifecycle of a Bosch product is depicted schematically and exemplarily for the Automotive area in the following Figure 1 to clarify some of the stages and basic terms that are relevant, if a failure occurs.



A product's relevant life does not start with its final assembly before it is delivered to the endcustomer. The history of a product begins with the production of its components. It is relevant under which conditions a product is produced or to which stresses the product or its components have been exposed. The Bosch product might be the final end customer product or a component of a bigger system itself. Therefore, dependent on the field failure, it is necessary to consider these different levels.

There are different production dates of the components, Bosch product, or end product. Furthermore, in case Bosch acts as a Tier 2 supplier, there might even be other intermediate products. Closely connected with this is the production timeperiod, production batch, delivery batch, which might be important to determine the affected volume of parts.

# 3 Partial Market Factors

Under the assumption that the failure behavior of products in a partial market can be transferred to other countries/markets, the complaint or failure quote can be determined using so-called partial market factors.

As a rule, the partial market corresponds to one or more countries or to a region.

The approach to projection via partial market factors assumes that all complaints from a partial market are reported and the offending products from that market are available for analyses.

From a vehicle producer's point of view, the partial market factor is determined, as a first approximation, by the ratio of the production/registration numbers in the partial market and the total production in the period under consideration.

Otherwise, the following ratios are possible as partial market factors:

- Ratio of the number of failed parts returned to the producer and the total number of failed parts in the field
- Ratio of the number of parts failed in a certain region and the total number of parts failed worldwide
- Ratio of the number of parts produced in a certain region and the total number of parts produced worldwide

However, the applicability of this rather simple method must be questioned if, for example, the reporting and returning behavior of the customer is unreliable or if there are market or model-specific differences in the equipment share of vehicles.

# 4 Preparation and Visualization of Field Data

## 4.1 Stair-Step Table

Experience shows that zero-mileage failures and field failures of products can often be attributed to specific failure causes if only products are affected that were produced during a particular production period. One would therefore expect to find a systematic relationship between zero-mileage or field quality and the production date. As a rule, there is no systematic relationship between zero-mileage or field quality and the purchase date. Hence, the purchase date is usually not a relevant point of reference.

The stair-step table is a representation of product failures ordererd according to reporting period (month, quarter) and production period (production date, month, quarter). To generate this table, all the failures reported from the beginning of the production month to the end of the reporting month (complaints by end customers that have been accepted as warranted by RB) are summed up and given in absolute terms and/or as a proportion (in ppm) of the total number produced during the relevant period.

| | ⟵ Reporting Quarter (RQ) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3Q.2017 | 2Q.2017 | 1Q.2017 | 4Q.2016 | 3Q.2016 | 2Q.2016 | 1Q.2016 | 4Q.2015 | 3Q.2015 | 2Q.2015 | 1Q.2015 | 4Q.2014 | 3Q.2014 |
| **PQ** | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm | ppm |
| 3Q.2014 | 3,193 | 3,181 | 3,079 | 2,972 | 2,747 | 2,559 | 2,223 | 1,805 | 1,314 | 831 | 413 | 147 | 12 |
| 4Q.2014 | 3,353 | 3,324 | 3,258 | 3,113 | 2,890 | 2,662 | 2,095 | 1,403 | 861 | 480 | 99 | 4 | |
| 1Q.2015 | 2,381 | 2,319 | 2,228 | 2,110 | 1,929 | 1,698 | 1,145 | 750 | 406 | 147 | 17 | | |
| 2Q.2015 | 2,574 | 2,558 | 2,360 | 2,178 | 1,766 | 1,342 | 765 | 457 | 132 | 17 | | | |
| 3Q.2015 | 3,331 | 3,233 | 2,913 | 2,541 | 1,891 | 1,250 | 492 | 209 | 33 | | | | |
| 4Q.2015 | 3,190 | 3,021 | 2,650 | 2,123 | 1,484 | 817 | 237 | 27 | | | | | |
| 1Q.2016 | 3,739 | 3,500 | 2,862 | 2,409 | 1,632 | 518 | 0 | | | | | | |
| 2Q.2016 | 2,684 | 2,451 | 1,680 | 1,011 | 451 | 36 | | | | | | | |
| 3Q.2016 | 1,601 | 1,230 | 493 | 110 | 12 | | | | | | | | |
| 4Q.2016 | 768 | 621 | 197 | 18 | | | | | | | | | |
| 1Q.2017 | 332 | 173 | 14 | | | | | | | | | | |
| 2Q.2017 | 38 | 15 | | | | | | | | | | | |
| 3Q.2017 | 0 | | | | | | | | | | | | |

Table 1: Stair-step table; status: end 12.2017

## 4.2 Nevada Chart

The Nevada Chart is a standard form for collecting data used for a Weibull Analysis. For each time period, when parts have been put into service, there is a volume and a certain number of field returns by failure date.

The amount of parts in service minus returns are the still operating parts in the field. In the Weibull analysis, the parts still in operation are treated as "suspensions“.

In the example below, 1,750 parts have been put into service in March 2016. In April 2016, 3 parts failed and have been returned. The remaining 1,747 parts are still operating in the field and possibly fail in the subsequent months.

At the end of the analysis period, all parts which were put into service and did not fail are considered as “suspensions”.

This chart needs to be created for every failure mode and evaluated separately.

| No. of parts | In Service Date | Number of field returns | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Feb. 16 | Mar. 16 | Apr. 16 | May 16 | Jun. 16 | Jul.16 | Aug. 16 | Sep. 16 | Oct. 16 | Nov. 16 | Dec. 16 |
| 1,500 | Jan. 2016 | 0 | 3 | 1 | 0 | 2 | 3 | 3 | 1 | 3 | 4 | 1 |
| 1,750 | Feb. 2016 | | 1 | 4 | 1 | 2 | 2 | 3 | 2 | 0 | 2 | 2 |
| 1,750 | Mar. 2016 | | | 3 | 4 | 2 | 5 | 3 | 4 | 1 | 1 | 2 |
| 2,000 | Apr. 2016 | | | | 4 | 1 | 2 | 2 | 1 | 2 | 1 | 1 |
| 2,500 | May 2016 | | | | | 0 | 3 | 2 | 1 | 2 | 1 | 1 |
| 3,000 | Jun. 2016 | | | | | | 1 | 1 | 2 | 2 | 4 | 3 |
| 2,500 | Jul. 2016 | | | | | | | 5 | 2 | 2 | 4 | 3 |
| 1,500 | Aug. 2016 | | | | | | | | 1 | 3 | 0 | 2 |
| 1,500 | Sept. 2016 | | | | | | | | | 2 | 4 | 1 |
| 2,500 | Oct. 2016 | | | | | | | | | | 0 | 1 |
| 3,500 | Nov. 2016 | | | | | | | | | | | 1 |
| 3,500 | Dec. 2016 | | | | | | | | | | | |

Table 2: Nevada Chart

*NOTE: Creating the Nevada chart by hand is a tedious and error-prone process, especially when non-trivial delay times must be considered.*

## 4.3 Isochrones Chart

The failure proportions of products of equal age (same time gap between production and failure report) can be taken from this table and used to plot a so-called isochrones chart (contour plot). Isochrones are lines of equal product age. Failures are not shown against the reporting month, but against the production month.

Fig. 2 shows isochrones for product ages of 0, 3, 6, 9, 12, 24 Month In Service (MIS). The 0-months isochrone represents product failures at ages of less than one month. In this case, the production month and reporting month are identical, corresponding to the entries along the jagged lower edge of the stair-step table.

By following an isochrone to the right, changes in product quality over time are immediately apparent. It's very useful to make a note of any quality improvement actions, such as design changes or changes in manufacturing and assembly processes, along the time axis.

The effects of such measures will then, after a corresponding delay, be apparent from the subsequent development of the isochrones. Likewise, the behavior of the lowest isochrones enables early recognition of any critical developments concerning the product quality for a specific production quarter.

Since the failures of products from a particular production month are captured as cumulative totals at the end of each reporting month, the isochrones can never cross. The most that can happen is that no further failures of products from a given production month are observed. In this case, the failure proportion of this production month does not increase any further, and subsequent isochrones all run through the same point.

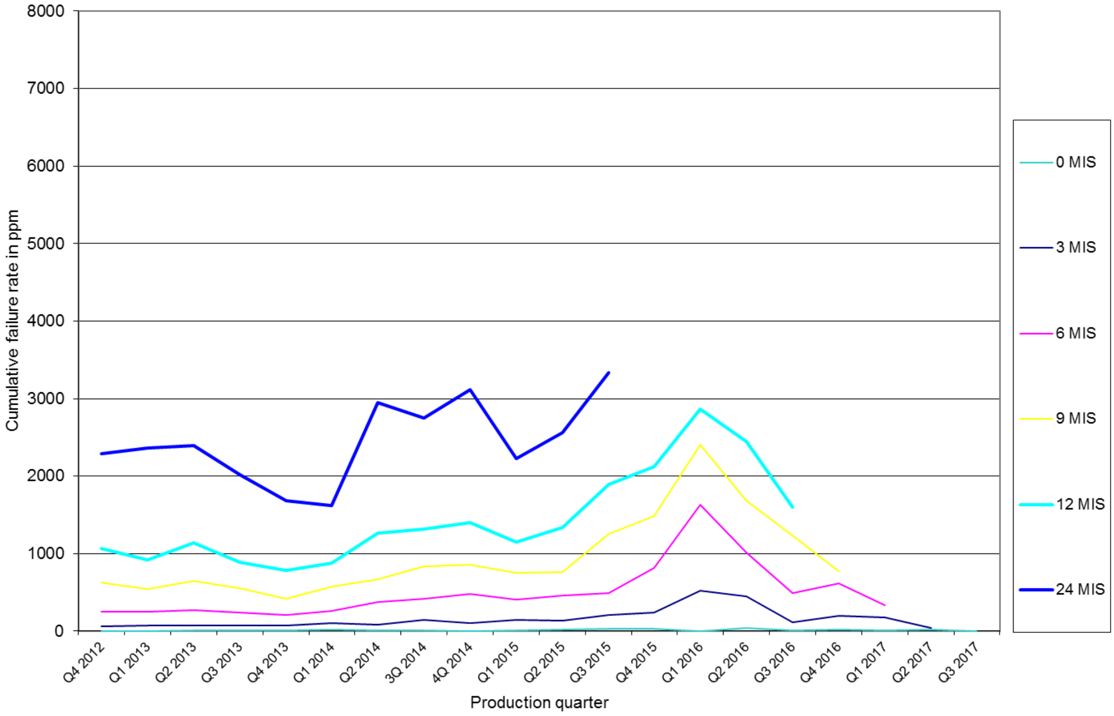



Fig. 2: Isochrones chart; status: end 12.2017

The isochrone values for a particular production month (read from the bottom up) correspond to the values in the row for that production month in the stair-step table (read from right to left). A given isochrone corresponds to one of the 'diagonals' in the stair-step table.

The stair-step table and the isochrone chart reflect the number of actual, reported and recognized failures. This figure may not be identical to the number of failures that have actually occurred, due to causes such as the following:

- The failure only occurs under certain operating conditions
- The driver does not notice the failure
- The driver does not feel the need to register a complaint, since the failure is felt to be insignificant
- The dealer workshop is unable to find the real cause of failure or just puts the vehicle owner off.



Of course, the vehicle manufacturer's service behavior plays a role as well. The manufacturer may, for example:

- Initiate a recall campaign for all the affected vehicles
- Instruct workshops to look out for this type of fault whenever a vehicle is serviced
- Instruct workshops to ask vehicle owners whether they have ever observed this specific type of fault.

Whenever some of the products manufactured over a certain period of time are shipped to other countries or regions where data collection is less than complete, the reported proportion of failures will be smaller than the actual proportion by a certain factor. This so-called partial market factor must be taken into account when reimbursing the OEM customers for warranty costs.

Determining the necessary warranty budget requires a projection for the likely number of failure reports received by the end of the warranty period (and beyond, if the company pursues a goodwill repair policy). However, due to the problems described above, such projections are subject to a fairly large degree of uncertainty.

*NOTE: In the literature there ace occasionally isochrone charts where the isochrones have a slightly different labeling (i.e. <1, <2, ..., <n) from the one used in our example, even though their meaning is the same. Another slight difference can be that the appropriate production quantity is noted at the top for each production quarter.*

*Obviously, the isochrones chart can be modified to suit individual practical requirements.*

## 4.4 Bar Chart: Months In Service (MIS)

The chart shows claims per million of sold components (cpm) based on vehicle manufacturing years.

Data sources are systems containing customer claim data e.g. Global Warranty Analyzer (GWA).

In it all claims are taken into account (Bosch responsibility, customer responsibility, in specification and open/under investigation).

The analysis performed with 3 MIS data (claims occurred within the first 3 month after vehicle registration), focusses on launch performance of a product. An 24 MIS analysis focusses on product robustness.

Due to the fact that the analyzed data is based on current customer data, it is assumed that all field claims are covered and therefor no partial market factor is applied. Projections are not part of this chart.

Depending on the data analyzed, the chart gives a more general overview like in the example below or can be done on a more detailed level.




Fig. 3: Bar charts: Months in Service

## 4.5 Scatter Diagram: Part Time in Service (PTiS)

The PTiS cloud represents product failures by product age against production date.

PTiS in month is the period of time from the date when the part left the production plant until it failed, this includes storage, shipping, transportation until the part is installed to a vehicle and the vehicle is put into service. PTiS is indicated for each production date.

The line separating past/future shows the maximum possible time in service for each production date at the time the diagram was created.

Claims known line is “line separating past/future” minus average period of time from repair/claim date until the claim is known in the Bosch Warranty Data Base (IQIS) and has been analyzed. The area between “Line separating past/future” and “Claims known” is a period of time where claims may already occur but are not know in the Warranty Data Base and/or are not analyzed.

Time D1 until registration of vehicle is the average period of time from part production date until the vehicle was put into service/registered.

The PTiS cloud visualizes failure behavior related to production date and time in service. As can also be seen from Figure 5 more recent production dates will always show a positive trend, the PTiS Cloud explains this trend by visualization of delay times and maximum possible time in service for each production date.

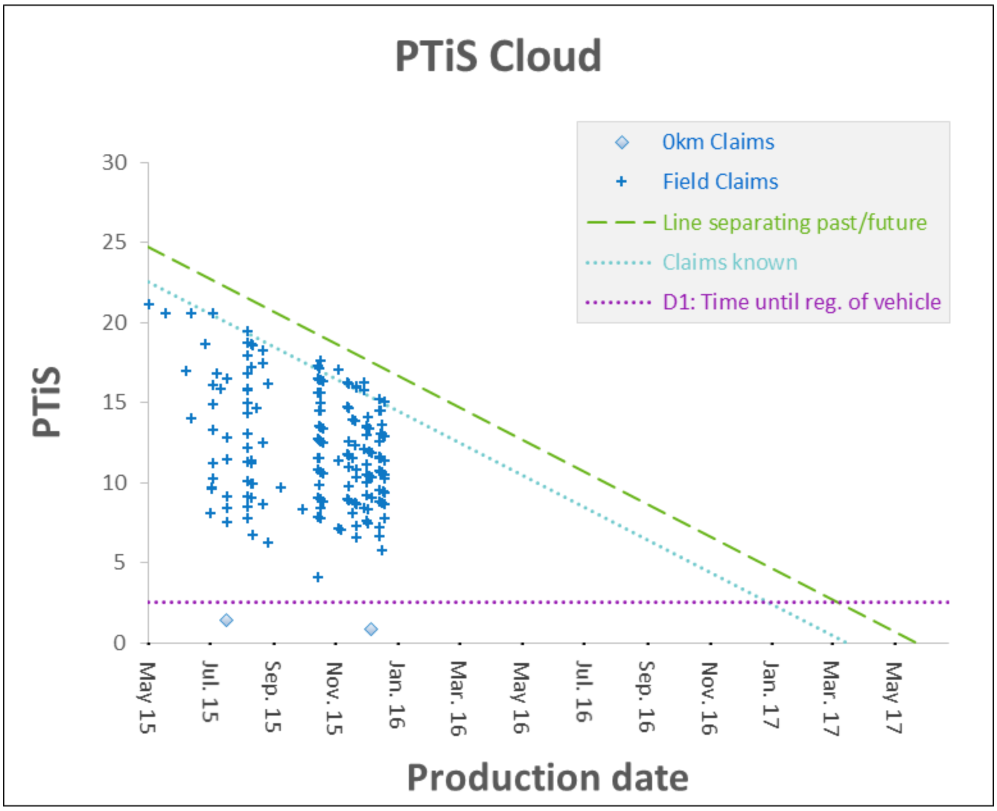



Fig. 4: Scatter plot

# 5 Projection Factors

## 5.1 Purpose of Projection Factors

When observing product performance in the field, there is a need to establish as soon as possible what the final complaint rate is likely to be, based on the product's present failure behavior in the field.

Such projections are needed both for business management of warranty cases (including production planning and planning of final stock levels at the time manufacture of this product is discontinued) and for reporting.

If the current (actual) complaint rates for various production dates are analyzed at a particular point in time, this will always show a positive (downward) trend, at least for more recent production dates (see Fig. 5). In other words, the more recent the production date, the smaller the complaint rate.



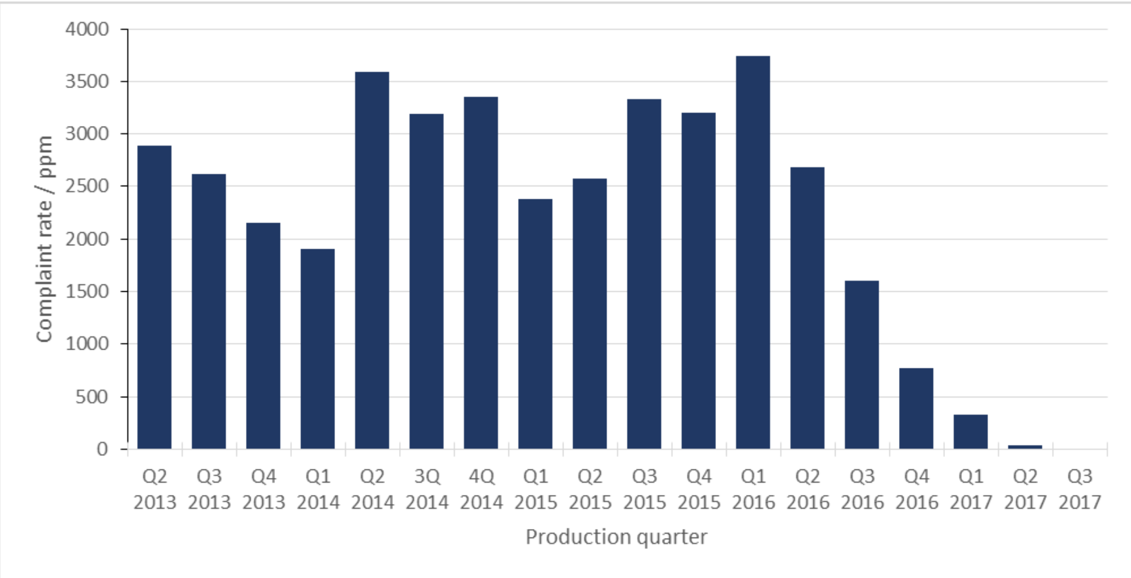



Fig. 5: Chart of current complaint rates for various production periods.

As discussed in Section 2.2, the cause of this effect is the statistical distribution of time intervals between the products' manufacture and the reporting of failures.

It's only after a certain amount of time, which considerably exceeds the warranty period (plus a limited goodwill period, where appropriate), that the cumulative total of complaints will asymptotically approach its final value. In the case of a one-year warranty period, the final value is typically only reached after three to four years (from production date).

The aim of the projection is to estimate the likely final complaint rate based on the current complaint rate. To determine this estimated final value, the current complaint rate must be multiplied by the projection factor.

## 5.2 Determining Projection Factors

Projection factors can be determined on the basis of the complaint rates known for a product (or comparable product) at a particular point in time. Monitoring the development of the complaint rates against time elapsed since the production date (product age) shows that the various production periods approach different final values (Fig. 6).

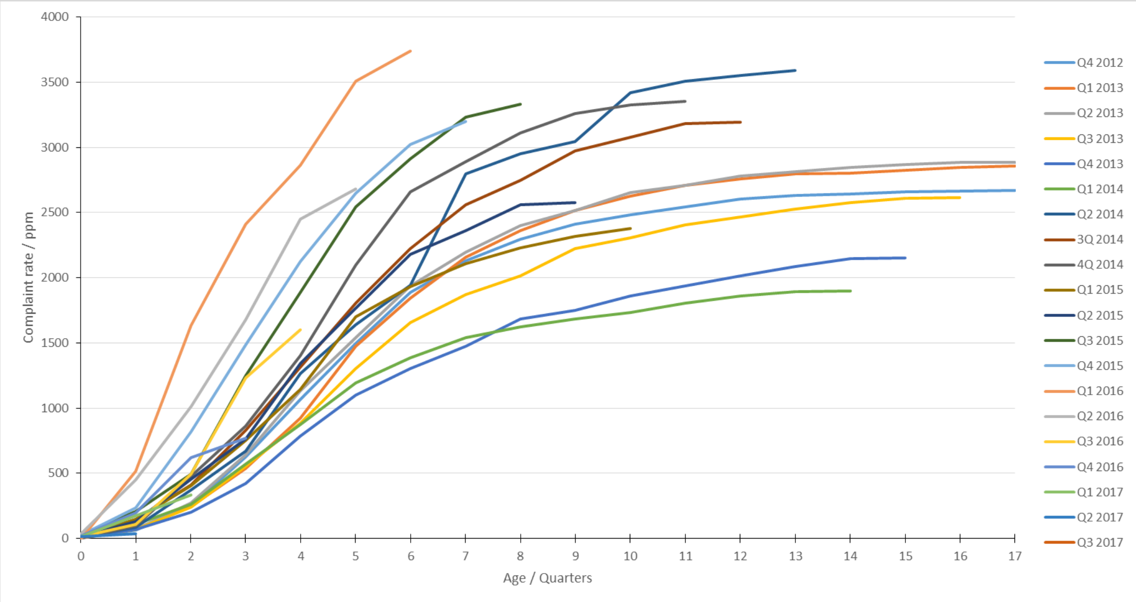



Fig. 6: Development of complaint rates over time (in ppm, relative to the quantity pro-duced)

Dividing the ppm values represented by the points on a curve (in Fig. 6) by the final value that the curve is approaching and plotting the percentages calculated in this manner against time yields a standardized plot as shown in Fig. 7.

If the curves all have a similar shape, i.e. if the standardized complaint rate behavior over time is independent of the production date, then an averaging process can be used to arrive at an average curve which may be assumed to be representative of this product's behavior. The projection factors then correspond to the reciprocals of the standardization factors.

The curve for the production quarter (PQ) 2Q.15 reaches 2,360 ppm after seven quarters (Fig. 6). After seven quarters, the complaint rates of the comparable quarters have on average reached about 76 % of their final value (Fig. 7).

If the time pattern of the ppm numbers for the production quarter under consideration is similar to the pattern observed in the other quarters, then the complaint rate will reach a final value of approximately $\frac{2{,}360\ ppm}{0.76} \approx 3{,}105\ ppm$.

The projection factor thus has a value here of $\frac{1}{0.76} \approx 1.32$.

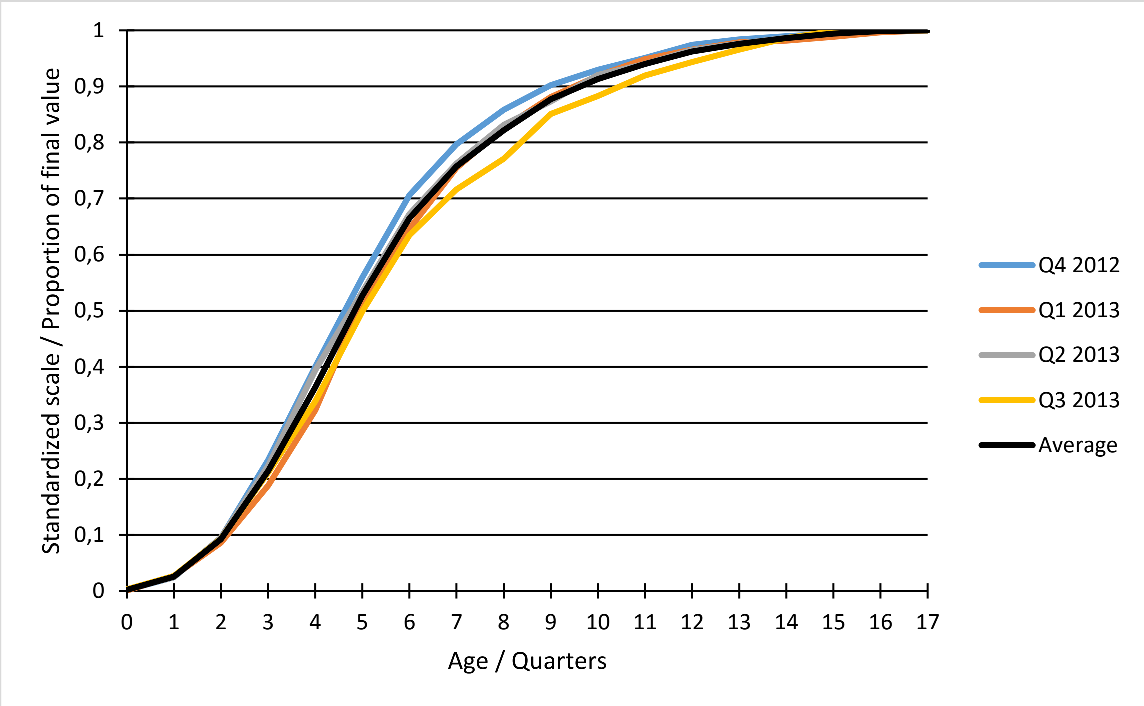



Fig. 7: Development of the standardized (expressed as proportion of the final value) complaint rate over time (for four production quarters).

### 5.2.1 Simple Calculation Method Based on the Stair-step Table

Below follows a description of a simple method for calculating projection factors based on the data contained in a stair-step table (see Section 4.1).

We are looking for the expected final complaint rate for a specific production quarter. The complaint rate value for this production quarter, as known at the time of the analysis, can be found in the column for the most recent reporting quarter (RQ, second column from the left). It represents a certain product age. We now look for a comparable value relating to a production quarter that's very much "older" and has a stable end value. Production quarters with an "age" of more than 14 quarters can be assumed to have settled down to a stable end value.

Data (production quarters) of equal age can be found in fields lying on a diagonal running from bottom left to top right in the stair-step table.

The final value for the older production quarter is again found in the column for the most recent reporting quarter. Dividing this final value by the comparable value yields a projection factor for this age, based on the data for the older production quarter. This can now be multiplied by the current value of the production quarter that interests us, to arrive at the expected end value for this production quarter (cf. the example on the following page).

This method assumes that the complaint rate for the quarter under consideration will develop in the same way as was the case for the earlier quarter some years ago that served as the basis for the projection. In other words, there is a tacit assumption that the general state of affairs concerning the factors summarized in Section 5.2.4 (e.g. manufacturing and development quality, reporting behavior) has not changed in any significant way.

Table 3

| | Reporting Quarter (RQ) | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PQ** | 3Q.2017 | 2Q.2017 | 1Q.2017 | 4Q.2016 | 3Q.2016 | 2Q.2016 | 1Q.2016 | 4Q.2015 | 3Q.2015 | 2Q.2015 | 1Q.2015 | 4Q.2014 | 3Q.2014 | 2Q.2014 | 1Q.2014 | 4Q.2013 | 3Q.2013 | 2Q.2013 |
| 2Q.2013 | 2,884 | 2,884 | 2,871 | 2,845 | 2,816 | 2,783 | 2,712 | 2,655 | 2,518 | 2,398 | 2,199 | 1,939 | 1,539 | 1,136 | 645 | 275 | 70 | 7 |
| 3Q.2013 | 2,615 | 2,611 | 2,579 | 2,525 | 2,467 | 2,403 | 2,308 | 2,226 | 2,016 | 1,872 | 1,659 | 1,304 | 885 | 554 | 243 | 69 | 7 | |
| 4Q.2013 | 2,153 | 2,146 | 2,085 | 2,017 | 1,938 | 1,858 | 1,749 | 1,683 | 1,474 | 1,302 | 1,039 | 784 | 422 | 204 | 68 | 11 | | |
| 1Q.2014 | 1,899 | 1,893 | 1,862 | 1,804 | 1,736 | 1,684 | 1,625 | 1,542 | 1,388 | 1,194 | 874 | 572 | 262 | 108 | 15 | | | |
| 2Q.2014 | 3,589 | 3,551 | 3,506 | 3,422 | 3,044 | 2,953 | 2,798 | 1,946 | 1,633 | 1,264 | 670 | 371 | 87 | 8 | | | | |
| 3Q.2014 | 3,193 | 3,181 | 3,079 | 2,972 | 2,747 | 2,559 | 2,223 | 1,805 | 1,314 | 831 | 413 | 147 | 12 | | | | | |
| 4Q.2014 | 3,353 | 3,324 | 3,258 | 3,113 | 2,830 | 2,662 | 2,085 | 1,403 | 861 | 430 | 93 | 4 | | | | | | |
| 1Q.2015 | 2,381 | 2,319 | 2,228 | 2,110 | 1,929 | 1,638 | 1,145 | 750 | 406 | 147 | 11 | | | | | | | |
| 2Q.2015 | 2,574 | 2,558 | 2,360 | 2,178 | 1,766 | 1,342 | 765 | 457 | 132 | 17 | | | | | | | | |
| 3Q.2015 | 3,331 | 3,233 | 2,913 | 2,541 | 1,891 | 1,250 | 492 | 209 | 33 | | | | | | | | | |
| 4Q.2015 | 3,199 | 3,021 | 2,650 | 2,123 | 1,484 | 817 | 237 | 27 | | | | | | | | | | |
| 1Q.2016 | 3,739 | 3,509 | 2,862 | 2,409 | 1,632 | 518 | 0 | | | | | | | | | | | |
| 2Q.2016 | 2,684 | 2,451 | 1,680 | 1,011 | 451 | 36 | | | | | | | | | | | | |
| 3Q.2016 | 1,601 | 1,230 | 493 | 110 | 12 | | | | | | | | | | | | | |
| 4Q.2016 | 768 | 621 | 197 | 18 | | | | | | | | | | | | | | |
| 1Q.2017 | 332 | 173 | 14 | | | | | | | | | | | | | | | |
| 2Q.2017 | 38 | 15 | | | | | | | | | | | | | | | | |
| 3Q.2017 | 0 | | | | | | | | | | | | | | | | | |

**Example:**

Sought: Final value for PQ 4Q.2015

Current value: 3,199 ppm

Age: 7 quarters (3Q.2017 – 4Q.2015)

Comparable value: 1,474 ppm (intersection PQ 4Q.13/RQ 3Q.15)

Final value (PQ 4.13): 2,153 ppm

Projection factor: 2,153 ppm / 1,474 ppm = 1.46

Projected final value for PQ 4Q.15 : 3,199 ppm * 1.46 = 4,671 ppm

### 5.2.2 Graphical Method

Following the basic approach outlined in Section 5.2.1, final complaint rates can also be determined using contour lines (Fig. 8). To do this, one follows an isochrone (line of equal age) to find a comparative value in the past (relating to an earlier production quarter). According to the data available today, this value has led to a certain final value (uppermost isochrone for the corresponding production quarter). This yields the prediction factor to be applied to the quarter under consideration. The drawbacks of this method are as described above (Section 5.2.1).

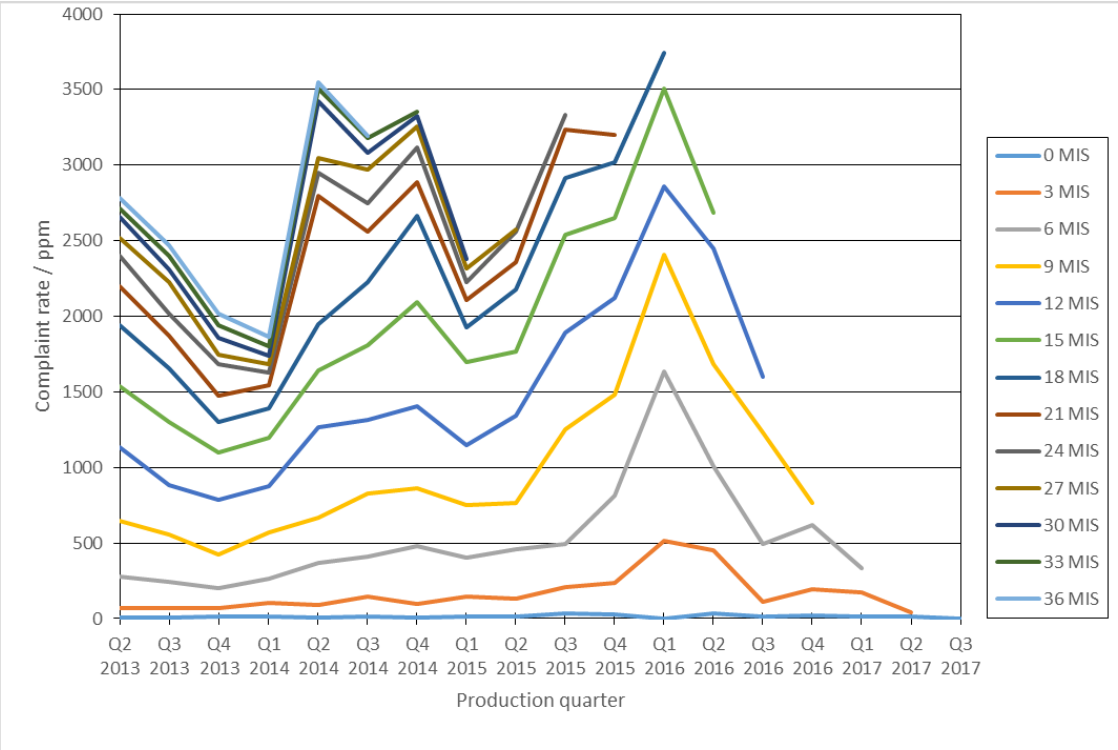



Fig. 8: Isochrones chart based on the data in Table 3

### 5.2.3 Computational Method

The following method is recommended in order to ensure that more recent data are also included in the determination of the prediction factors, thus putting the calculation on a broader footing.

Using a suitable software program (e.g. Excel), the stair-step table (Table 4) is converted to the format shown in Table 5. This shows the cumulative complaint rates for each production quarter against product age, measured in quarters elapsed since the production date. It can be seen that the values listed from right to left in each row (specific production date) of the stair-step table (Table 4) now appear in the corresponding column (relating to that same production date) of Table 5. The values found in these columns, read from top to bottom, correspond to the points of intersection of the isochrones (0, 1, 2, ...) with the vertical line representing the relevant PD in the isochrone chart (Fig. 8).

In each column (relating to a specific PQ) of Table 4, we divide each value by the one above it and enter the result in the lower of the two corresponding cells in Table 5. Each of these numbers indicates the factor by which the complaint rate has increased in comparison with the value of the preceding quarter. Of course, this only yields a useful result where the denominator is > 0.

The following example illustrates the procedure:

Of the products manufactured in production quarter 2Q.13, 645 ppm had given rise to a complaint by the time the first two quarters had elapsed (age = 2 quarters). The cumulative complaint rate reached 5,092 ppm by the end of the following quarter (age = 3 quarters). In other words, it rose by a factor of $\frac{645}{275} = 2.35$. This number thus corresponds to an "instantaneous projection factor" and is entered in the appropriate cell of Table 3.4.3.2 (column: 2Q.2013, row: age = 3).

The "instantaneous projection factors" are then averaged for each age across all of the PQs (horizontally), and the averages are recorded in the average column on the right. Note that cells that contain the ∞ symbol are not taken into account in this.

The projection factors sought (PF column) are then calculated by multiplying together all the averages from that row downwards.

Thus the number 3.74 (in the PF column) results from multiplying all the averages together: $1.57 \cdot 1.34 \cdot 1.19 \cdot 1.14 \cdot \ldots \cdot 1.00 = 3.74$. This is the projection factor for the fourth quarter following the production date. Likewise, the projection factor for the fifth quarter is found by multiplying all of the averages from that row (age = 5 quarters) onwards: $1.34 \cdot 1.19 \cdot 1.14 \cdot \ldots \cdot 1.00 = 2.38$.

This method ensures that all the available data from the stair-step table are used to determine the projection factors.

One might think that it would make sense to use age-dependent weighting when calculating the averages. However, corresponding studies have shown that a broader base yields better results.

The values in the row that corresponds to an age of one quarter, as well as some of the values in the row for an age of two quarters, exhibit a great deal of variation. This applies both to the initial rate of complaints reported and the "instantaneous projection factors" derived from them. This is due to the various factors covered in Section 5.2.4.

The first few age quarters in particular often yield projection factors that are significantly greater than 10.

Due to the evident uncertainty involved, it is recommended that practitioners should not normally try to calculate complaint rate projections on the basis of the two or three earliest quarters.

| Age | Production Quarter (PQ) 2Q.2013 | 3Q.2013 | 4Q.2013 | 1Q.2014 | 2Q.2014 | 3Q.2014 | 4Q.2014 | 1Q.2015 | 2Q.2015 | 3Q.2015 | 4Q.2015 | 1Q.2016 | 2Q.2016 | 3Q.2016 | 4Q.2016 | 1Q.2017 | 2Q.2017 | 3Q.2017 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 7 | 7 | 11 | 15 | 8 | 12 | 4 | 11 | 17 | 33 | 27 | 0 | 36 | 12 | 18 | 14 | 15 | 0 |
| 1 | 70 | 69 | 68 | 108 | 87 | 147 | 93 | 147 | 132 | 209 | 237 | 518 | 451 | 110 | 197 | 173 | 38 | |
| 2 | 275 | 243 | 204 | 262 | 371 | 413 | 430 | 406 | 457 | 492 | 817 | 1,632 | 1,011 | 493 | 621 | 332 | | |
| 3 | 645 | 554 | 422 | 572 | 670 | 831 | 861 | 750 | 765 | 1,250 | 1,484 | 2,409 | 1,680 | 1,230 | 768 | | | |
| 4 | 1,136 | 885 | 784 | 874 | 1,264 | 1,314 | 1,403 | 1,145 | 1,342 | 1,891 | 2,123 | 2,862 | 2,451 | 1,601 | | | | |
| 5 | 1,539 | 1,304 | 1,099 | 1,194 | 1,633 | 1,805 | 2,085 | 1,638 | 1,766 | 2,541 | 2,650 | 3,509 | 2,684 | | | | | |
| 6 | 1,939 | 1,659 | 1,302 | 1,388 | 1,946 | 2,223 | 2,662 | 1,929 | 2,178 | 2,913 | 3,021 | 3,739 | | | | | | |
| 7 | 2,199 | 1,872 | 1,474 | 1,542 | 2,798 | 2,559 | 2,830 | 2,110 | 2,360 | 3,233 | 3,199 | | | | | | | |
| 8 | 2,398 | 2,016 | 1,683 | 1,625 | 2,953 | 2,747 | 3,113 | 2,228 | 2,558 | 3,331 | | | | | | | | |
| 9 | 2,518 | 2,226 | 1,749 | 1,684 | 3,044 | 2,972 | 3,258 | 2,319 | 2,574 | | | | | | | | | |
| 10 | 2,655 | 2,308 | 1,858 | 1,736 | 3,422 | 3,079 | 3,324 | 2,381 | | | | | | | | | | |
| 11 | 2,712 | 2,403 | 1,938 | 1,804 | 3,506 | 3,181 | 3,353 | | | | | | | | | | | |
| 12 | 2,783 | 2,467 | 2,017 | 1,862 | 3,551 | 3,193 | | | | | | | | | | | | |
| 13 | 2,816 | 2,525 | 2,085 | 1,893 | 3,589 | | | | | | | | | | | | | |
| 14 | 2,845 | 2,579 | 2,146 | 1,899 | | | | | | | | | | | | | | |
| 15 | 2,871 | 2,611 | 2,153 | | | | | | | | | | | | | | | |
| 16 | 2,884 | 2,615 | | | | | | | | | | | | | | | | |
| 17 | 2,884 | | | | | | | | | | | | | | | | | |

| Age | Production Quarter (PQ) 2Q.2013 | 3Q.2013 | 4Q.2013 | 1Q.2014 | 2Q.2014 | 3Q.2014 | 4Q.2014 | 1Q.2015 | 2Q.2015 | 3Q.2015 | 4Q.2015 | 1Q.2016 | 2Q.2016 | 3Q.2016 | 4Q.2016 | 1Q.2017 | 2Q.2017 | 3Q.2017 | Average | PF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |
| 1 | 10.00 | 9.86 | 6.18 | 7.20 | 10.88 | 12.25 | 23.25 | 13.36 | 7.76 | 6.33 | 8.78 | 1.00 | 12.53 | 9.17 | 10.94 | 12.36 | 2.53 | | 9.67 | 9.7 |
| 2 | 3.93 | 3.52 | 3.00 | 2.43 | 4.26 | 2.81 | 4.62 | 2.76 | 3.46 | 2.35 | 3.45 | 3.15 | 2.24 | 4.48 | 3.15 | 1.92 | | | 3.22 | 31.2 |
| 3 | 2.35 | 2.28 | 2.07 | 2.18 | 1.81 | 2.01 | 2.00 | 1.85 | 1.67 | 2.54 | 1.82 | 1.48 | 1.66 | 2.49 | 1.24 | | | | 1.96 | 61.1 |
| 4 | 1.76 | 1.60 | 1.86 | 1.53 | 1.89 | 1.58 | 1.63 | 1.53 | 1.75 | 1.51 | 1.43 | 1.19 | 1.46 | 1.30 | | | | | 1.57 | 96.2 |
| 5 | 1.35 | 1.47 | 1.40 | 1.37 | 1.29 | 1.37 | 1.49 | 1.43 | 1.32 | 1.34 | 1.25 | 1.23 | 1.10 | | | | | | 1.34 | 128.8 |
| 6 | 1.26 | 1.27 | 1.18 | 1.16 | 1.19 | 1.23 | 1.28 | 1.18 | 1.23 | 1.15 | 1.14 | 1.07 | | | | | | | 1.20 | 153.9 |
| 7 | 1.13 | 1.13 | 1.13 | 1.11 | 1.44 | 1.15 | 1.06 | 1.09 | 1.08 | 1.11 | 1.06 | | | | | | | | 1.14 | 174.9 |
| 8 | 1.09 | 1.08 | 1.14 | 1.05 | 1.06 | 1.07 | 1.10 | 1.06 | 1.08 | 1.03 | | | | | | | | | 1.08 | 188.3 |
| 9 | 1.05 | 1.10 | 1.04 | 1.04 | 1.03 | 1.08 | 1.05 | 1.04 | 1.01 | | | | | | | | | | 1.05 | 197.4 |
| 10 | 1.05 | 1.04 | 1.06 | 1.03 | 1.12 | 1.04 | 1.02 | 1.03 | | | | | | | | | | | 1.05 | 207.0 |
| 11 | 1.02 | 1.04 | 1.04 | 1.04 | 1.02 | 1.03 | 1.01 | | | | | | | | | | | | 1.03 | 213.3 |
| 12 | 1.03 | 1.03 | 1.04 | 1.03 | 1.01 | 1.00 | | | | | | | | | | | | | 1.02 | 218.3 |
| 13 | 1.01 | 1.02 | 1.03 | 1.02 | 1.01 | | | | | | | | | | | | | | 1.02 | 222.6 |
| 14 | 1.01 | 1.02 | 1.03 | 1.00 | | | | | | | | | | | | | | | 1.02 | 226.1 |
| 15 | 1.01 | 1.01 | 1.00 | | | | | | | | | | | | | | | | 1.01 | 228.0 |
| 16 | 1.00 | 1.00 | | | | | | | | | | | | | | | | | 1.00 | 228.7 |
| 17 | 1.00 | | | | | | | | | | | | | | | | | | 1.00 | 228.7 |

Tables 4 (top) and 5 (bottom). For explanation see text.

### 5.2.4 Limitations of the Method

The “quality” of the projection factors determined and thus the projection itself is limited by various factors. These include:

- Product-specific reporting behavior
- Sporadic batch reports made by OEM customers
- Long storage periods at RB or the customer’s premises
- Failure behavior of the products concerned
- Customer usage patterns
- Various warranty/goodwill periods.

The influence of these noise factors must be assessed in each individual case.

Experience shows that projections should only be made for quarters that go back at least two to three quarters (from the time of the analysis), since otherwise there won’t be a sufficient database for a reliable prediction.

Failure Forecast

This chapter explains how future failures can be predicted based on field data. More specifically a forecast consists of how many parts will have failed at a given date in the future. Typically, this information is given in terms of expected value and lower and upper bound.

### 5.2.5 Collecting Information

The main input information for a field forecast are the currently known claims as well as the monthly production quantities. Notice that depending on the specific case at hand, a higher resolution may be needed, i.e. weekly or daily production volumes.

### 5.2.6 Production Quantities

When defining the relevant production quantities, it is imperative that only products which will fail due to the root cause at hand are selected. Any major mistakes will void the analysis altogether. This leads directly to two questions:

1. During which time period were parts with a deviation manufactured?
2. What fraction of parts in said production period carries the deviation?

This first question can only be answered by Problem Solving. The relevant production period coincides with the TRC being active. Hence, it is bounded by when the technical root cause started and the corrective measures.

Unfortunately, many times this information is not available at the time the risk analysis is being prepared. As an intermediate solution for the first question, the relevant time period can be defined from the earliest manufacturing date of the claims until the present. This is quite contrary to inituition as it does not include the largest number of parts possible. However, at a second glance, in the absence of certainty regarding the time frame, it is the most conservative thing to do as it assumes minimum field experience and thereby maximizes the forecasted failure rate.

If for said time period, monthly production volumes are not yet available and only the total volume is known, it is best to evenly distribute the total volume across the production period. Bear in mind, however, that if within said production period an event occurred that influenced the production, it is necessary to split the production period. These influences may pertain to the deviation as well as its propagation. Such events may include but are not limited to maintenance, containment, process changes.

Depending on the situation, the volume may also need to be split according to lines, machines, nests. The second question is subject of section 11.1.

### 5.2.7 Information about Logistics

Of equal importance to the analysis is the proper consideration of logistic information. Especially, if the time in field is relatively short at the time the analysis is prepared, the delay times have considerable influence on the outcome. The long term forecast critically depends on the warranty time, as it limits the time during which a part can be claimed.

Additionally, it should not be neglected that only a subset of all failures is actually claimed. This can be modelled with a so-called field factor, that encodes the probability that a failure will actually be claimed.

#### Delay Times

There are four kinds that need to be considered:

- D1a — The time it takes for a part to reach the OEM plant, i.e. time spent between Bosch production and OEM production.
- D1b — The time between OEM production and the registration of the vehicle.
- D2-0km — The time it takes for a claim to reach Bosch from the OEM, i.e. the claim is analyzed and assigned to the case at hand.
- D2-Field — The time it takes for a claim to reach Bosch from the field, i.e. the claim is analyzed and assigned to the case at hand.

It is clear, that these delay times limit the amount of field experience. Therefore, properly accounting for them is crucial.

They can be derived from the general claim database. This exploits the fact the delay time is independent of the root cause of failure. For most cases, it is sufficient to model the respective delays as a single number, e.g. the part takes 2 months to reach OEM production (D1a=2 months). However, it may also occur that the additional precision is required, because the claims exhibit large variations in delay times. In cases like this, literature suggests to use statistical distributions to reflect at which point in time a given number of parts is in the field.

#### Warranty time and lifetime

For the sake of brevity, only the warranty time is elaborated here. The ideas and concepts remain the same for inclusion of the lifetime into the analysis. The warranty influences the analysis at two points.

First of all, it limits the field experience in the context of fitting a Weibull distribution. Suppose some parts at risk have been in the field for 5 years. That may lead to the conclusion that the field experience stretches over 5 years as well. However, because the warranty time is typically 3 years, the field experience is capped at 3 years.

Furthermore, the warranty time needs to be taken into account as part of the forecast. Once the Weibull parameters are known and the forecast is made, it is crucial to bear in mind that parts older than the warranty time cannot be claimed.

### 5.2.8 Claim Information

The claim database should contain the following fields:

1. Production date @ Bosch
2. Vehicle manufacturing date @ OEM
3. Registration date of the vehicle
4. Failure date
5. Reporting date, i.e. the date when the claim appears in the case database

From this information the time in field can be deduced. Furthermore, the delay times can be deduced given a sufficient number of claims.

For most cases the above is sufficient. If, however, the load activating the failure depends, say, on mileage, additional information is needed. Other than mileage, loads may be modelled in terms of power on time, power on cycle, seasonal effects to name a few.



### 5.2.9 Determine the volume with deviation

In the past section, the information that is needed for a risk analysis was elaborated on. We briefly touched the fact that the number of parts at risk is crucial for a meaningful analysis. In this section, some methods are introduced to aid this process.

If every part is affected, i.e. carries the deviation, then this step is not relevant.

In most cases statistical methods will be used to estimate the volume with deviation.

- If all delivered products carry the deviation, skip this step (e.g. software problem)
- Estimate, upper and lower bound of the volume with deviation;
  Standard methods are:

  - Control sampling

  - Test gate method

  - Distribution and experience based estimators

### 5.2.10 Estimating the Number of Failures Until a Given Date

- The failure occurrence over time is described in most cases by a Weibull model
- For delivered products with deviation determine their total exposed load (e.g. lifetime) until today
  - MIS: For the failed products = failure date registration date
  - { MIS: For the products survived until today = last update registration date D2 = last update production date Bosch product D2 D1
- From the propagation of the deviation it should be derived whether the third parameter of the Weibull model has to be used.
- Minimal lifetime corresponds to a positive third parameter
- Preaging in production corresponds to a negative third parameter

## 5.3 Risk analysis for the 8D Process

### 5.3.1 Introduction

In the context of a risk analysis we distinguish between four different phases. Within these phases the target and even the methodological approach change.

The technical part of the complaint process is the 8D process.

### 5.3.2 Phase 0: Predetection Phase

Obviously the risk of field failures starts as soon as a vehicle with a specific deviation is used in the field. The deviation refers to a problem as defined in [Booklet 16]. Typically the root cause is not yet known to Bosch or misjudged as not relevant for the field or treated as background noise. Here we consider deviation which propagates from a Bosch product only. A deviation of a unknown technical root cause is typically detected within the complaint process. The technical part of process is called 8D process. As soon as Bosch became the problem aware and a 8D process is started for this deviation the prephase ends.



### 5.3.3 Phase 1: Early Phase of Risk Analysis

The early phase starts whith the 8D process to the failed product. Within the 8D process we want to understand and fix the technical root cause. At the same time we would like to know the expected additional field failures from the production volume already in field. If many failures already have occurred, customer escalation has happened or even safety critical failure modes are possible this question cannot be postponed. Depending on the outcome of the answer decisions with strong economical impact are possible. In these cases no answer is not an option.

The answer to this question of potential failure quantity depends heavily on the as yet unknown technical root cause. To manage the conflict arising from the time required to analyze the problem and develop a solution on the one hand, and the need for a failure prognosis on the other, it is necessary to calculate/create an appropriate prediction using assumptions and expert knowledge.

The typical situation in this phase is that

1. we need assumption about the technical root cause

2. not all possible failure activation mechanism are known

3. quantities with deviation are not completely known

One must be aware of the following: It is very important to start the risk analysis immediately. In this phase the risk analysis is based on assumptions and is called preliminary. This preliminary risk analysis is required as soon as possible and has to be reworked continuously.

### 5.3.4 Phase 2: Establishing the forecast (D4 until D6)

As soon as the technical root cause is understood (closure of D4) and all possible failure activation mechanisms are known (propagation of the deviation understood) the volumes with deviation can be estimated from TRC and tests. The failure forecast is derived from physical understanding of the root causes, field data and tests. Used assumptions must be plausible and agreed with the customer (internal/external).

### 5.3.5 Phase 3: Field observation (starts with D7)

As long the field behavior (0 km, field) is consistent with the failure forecast (number of failures is below the upper bound of the failure forecast) the forecast needs not to be reworked, i.e. the risk analysis is up-to-date. In IQIS the Q8 collects all claims with the same TRC (in practice there might be temporarily several Q8s).



### 5.3.6 Summary

- The risk analysis team has to be aware in which phase of the risk analysis it is and act accordingly. This makes the handling of a quality issues more effective and efficient. Even unproductive escalations can be avoided or reduced.
- Based on current knowledge, it may be necessary to return to earlier phases and go through them again.

# 6 Several identical components in the product

Often, several of the same components are installed in vehicles, e.g. spark plugs, ignition coils, injectors, ABS sensors, or window motors. In the event of problems on a four-cylinder engine, for instance, abnormalities will usually already be apparent in operation and indicated by the motor indicator light, even if only one of the cylinders is affected.

The following calculation can only lead to a meaningful result if the number of defective parts is correctly determined and documented, and further limiting conditions are met (e.g. narrowly limited production period, occurrence of the defect is independent of the mileage or operating time, independence of the individual component failures from each other).

The probability that, for example, exactly $k$ of four identical components are defective can be calculated using the probability function of the binomial distribution.

$$P(n,k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Here are the results $P(4,k)$ for $k = 0, 1, 2, 3, 4$ and $p = 0{,}02$:

| $k$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(4,k)$ | 0.9224 | 0.0753 | 0.0023 | 3.14E-05 | 1.60E-07 |

When evaluationg field data, the proportion $p$ is not known. However, one can try to estimate it in the following way. The ratio of the probabilities for $k = 2$ and $k = 1$ is $\frac{0.0023}{0.0753} = 0.0306$. The remaining ratios can be calculated in the same way. In the case of $p = 0.1 = 10\ \%$, we can expect that about 16.7 % of the repairs will require the replacement of two components instead of just one. The probabilities for $k > 2$ are negligible when $p$ is small.

| p | 2 / 1 | 3 / 1 | 4 / 1 |
|---|---|---|---|
| 0.01 | 0.0152 | 0.0001 | 2.58E-07 |
| 0.02 | 0.0306 | 0.0004 | 2.12E-06 |
| 0.05 | 0.0789 | 0.0028 | 3.64E-05 |
| 0.10 | 0.1667 | 0.0123 | 3.43E-04 |

Table 6

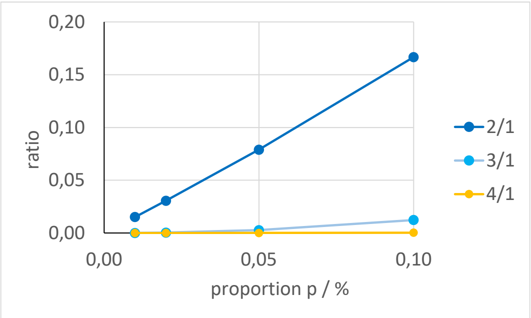



Fig. 9

Thus, if the numbers $N_1$ and $N_2$ of repairs in which one or two components were exchanged are known, this ratio $r = \frac{N_2}{N_1}$ can be used to infer the proportion $p$ of the population: $p = \frac{2 \cdot r}{3 + 2 \cdot r}$.

*NOTE: The described situation corresponds to a bowl model with a proportion p of defective units in the population, s. [Booklet 2]. It assumes homogeneous mixing of defect-free and defective components.*

*The calculation can only lead to a meaningful result if the number of defective parts is correctly determined and documented.*

# 7 Establishing the Annual Mileage Distribution

Where possible, the annual mileage distribution should be determined on the basis of independent representative data (e.g. manufacturer's data). From a statistical point of view, an estimation based on sample data would only be legitimate if a random(representative) sample had been defined prior to the first use of the vehicles, i.e. at a time when it was not yet known how many of the vehicles were going to fail.

The failures included in warranty data tend to be based on a large variety of failure mechanisms. Hence it seems justified to estimate the annual mileage distribution on the basis of the available data. However, the result might be biased, due to the fact that the exclusive consideration of warranty cases constitutes a negative selection principle, and should therefore be checked for plausibility.

For each individual case, it is possible to calculate an average annual mileage based on the difference between the registration date and the failure date and the vehicle's mileage at the time of failure.

$$annual\ mileage = \frac{mileage\ at\ time\ of\ failure}{failure\ date - registration\ date} \cdot 365\ days$$ (production date known) or

$$annual\ mileage = \frac{mileage\ at\ time\ of\ failure}{failure\ month - registration\ month} \cdot 12\ months$$ (production month known)

Experience has shown that the lognormal distribution represents a suitable model for this distribution (see e.g. [VDA 3.14] and [Pauli]). Plotting the data on lognormal probability paper or performing a statistical goodness-of-fit test will show whether this assumption is justified in this particular case. Fig. 10 below shows an example of a good fit.

A random variable is said to follow the lognormal distribution if the logarithm of this variable follows the normal distribution. The statistics $\mu$ and $\sigma$ of the underlying normal distribution can be calculated as the average and standard deviation of the logarithms of the data (s. [Booklet 3]).

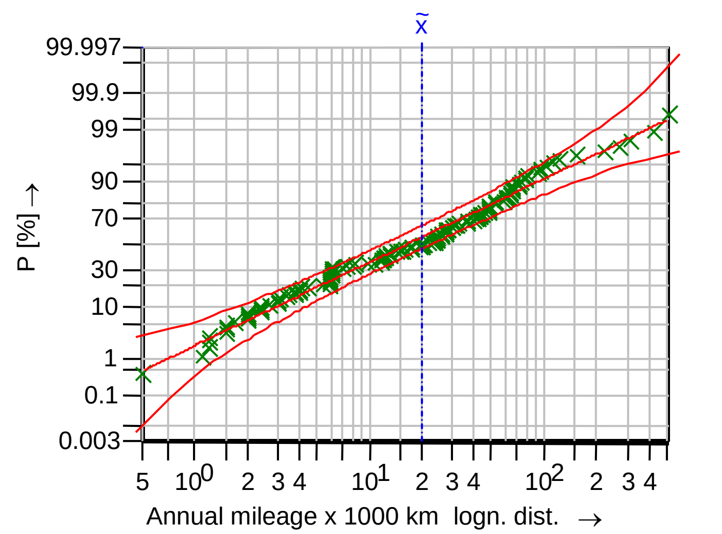



Fig. 10: Lognormal plot of annual mileage. All of the values lie within the 90% confidence zone.

*NOTE: The annual mileages (mileage distributions) may vary by country, for instance.*

# 8 Systematic Collection of Field Data

For the development of a new product, it is important to realistically assess the loads that occur in the field in order to be able to design the product for the planned lifetime in an operationally reliable yet cost-effective manner, as well as to avoid critical operating conditions and failures.

Systematic collection of field data is about obtaining functional and loading data from products that are in use by the enduser.

Modern possibilities of sensor technology and signal processing allow such data to be generated, collected, condensed, stored and transmitted. For this, however, current products must already have the necessary hardware. The corresponding components must therefore already be present and usable in the current product in order to be able to obtain the desired information for a new product of the next generation or even the one after that.

In a sense, one encounters the chicken-and-egg problem here. The development must already know which data are needed to create the conditions for data recording in form of hardware und software. In this context, also the term “Design to Diagnose” is used. The products should be designed in such a way that they can be (at least) sufficiently well diagnosed and analyzed.

The development must already know which data are needed to create the conditions for data recording in form of hardware und software.

For example, the use of an additional sensor may initially be perceived as an additional cost factor that makes product development and manufacturing more expensive. At second glance, however, there may be a positive cost/benefit effect in relation to the service life of the product, e.g. through lower service and warranty costs due to improved diagnostic possibilities.

## 8.1 Data Security

When collecting field data, requirements in the area of data security and confidentiality enjoy the highest priority. The collection of personal data is only permitted if there is a legal basis and this fulfills a purpose covered in the basis. Examples of possible legal bases include the legitimate interest of the manufacturer (e.g., in the area of product quality and safety) or the explicit consent to data collection by the so-called "data subject", i.e., the user of the monitored product.

If it can be ensured that no person-related data is collected during the data collection process, this does not release the user from the duty of care in handling the data.

Therefore, when designing a data collection solution, the involvement of the legal service with a focus on data protection is required, including to evaluate whether a personal reference is present and, if so, what legal basis can be used for collection.

## 8.2 Data Analysis in the Context of Field Tests

One goal of field data evaluation is to derive practice-oriented requirements and test profiles suitable for this purpose. A good possibility to collect such field data are field tests with a limited number of users. If the end user is not able to commission and test devices directly in his area, there is in some cases the possibility to invite him to Bosch and have the devices tested there.

A critical aspect here is the selection of a representative group of users which can have a significant effect on the test results.

Data collected from the user are often time series with a high sampling rate. After several hours of operation, large amounts of data in the range of several giga-bytes can accumulate. In the interest of comparability of such data, data reduction is necessary. Classification methods are often used for this purpose.



### 8.2.1 Classification Methods

One distinguishes counting methods, 1-parametric and 2-parametric classification methods.

Examples for counting methods:

- Event counter (e.g. starts of a device, switch operations)
- Time recording (e.g. operating time of a motor)

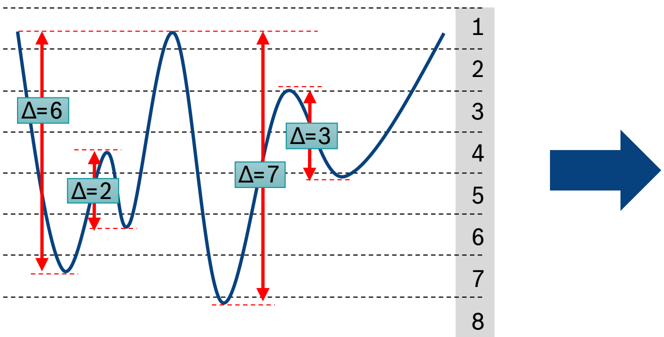
1-parametric

- Time-at-level counting
- Range-pair counting
- 1-parametric rainflow counting

2-parametric counting

- 2-parametric rainflow counting
- 2-parametric time-at-level counting

In many cases, the operating dynamics at the user has a greatly influences the load on the components. The operating dynamics can be recorded using the range-pair method or rainflow counting. In the following example the range-pair method is used to determine the load cycles endured by a device in the field test. For this purpose, the frequency of the cycles is evaluated depending on their level (range “delta”).




| delta | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Frequency | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |

Fig. 11: Range-pair counting

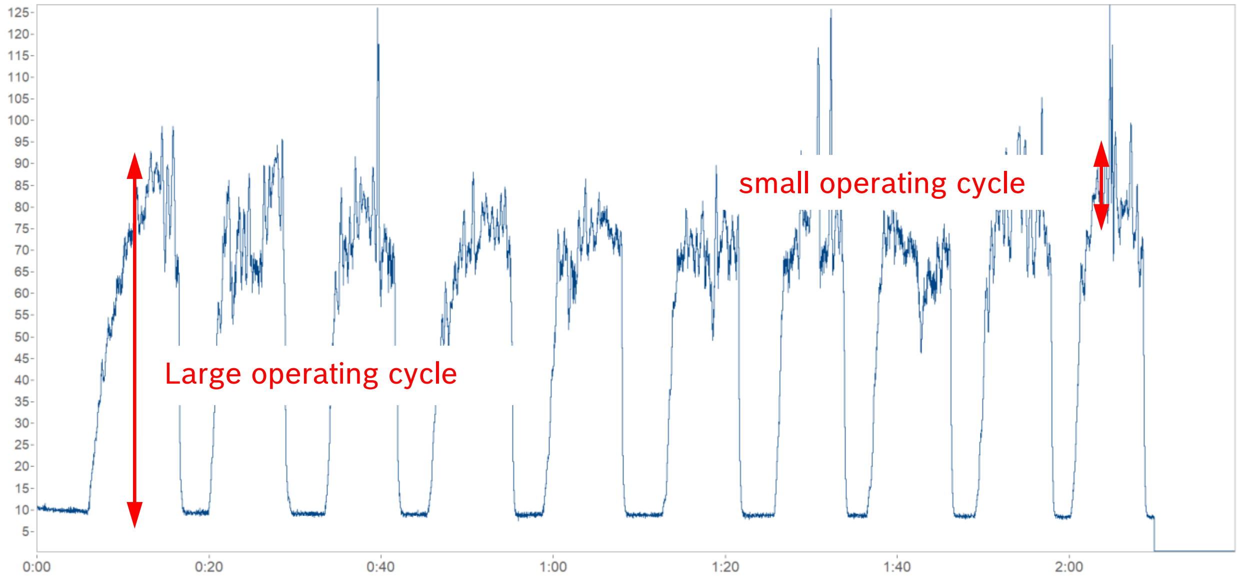



Fig. 12: Typical measured time series.

By range-pair counting the time series were converted into the following bar chart.

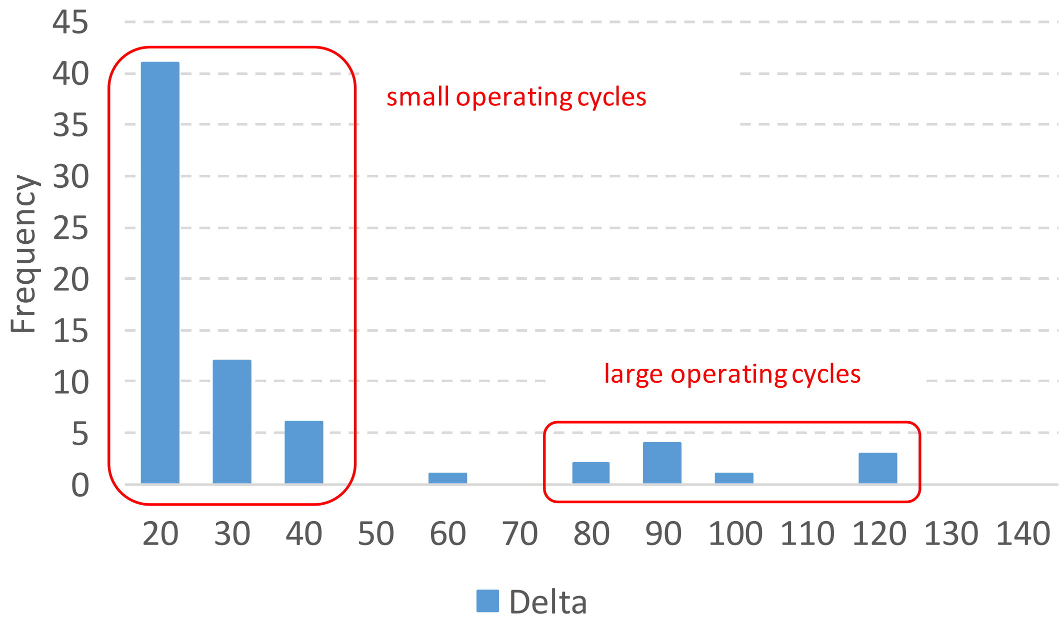



Fig. 13: Frequencies of small and large operating cycles rsulting from range-pair counting

In total, 10 great operating cycles with ranges between 80 and 120 units were counted and 60 smaller ones with load cycles between 30 and 40. These values can be used to distinguish between different users and load profiles. However, for a comparison it is necessary to normalize to the same reference quantities (e.g. frequency per hour, day, km).

Further methods and explanations of load collectives and counting methods can be found e.g. in [Haibach], [Köhler] and [DIN 45667].

### 8.2.2 Gaining Information from Load Collectives

Different information can be derived from load collectives depending on how they are viewed.

**Individual consideration:** The load collective of an individual product (over a period of time) can be checked against the specification of the components of the product, e.g. to derive statements about the remaining life of the product (at constand load).

**Grouped consideration:** A product-specific load collective, when compared to the individual or grouped load collectives of products of the same type and application environment, can provide indications of under- or overloading of the product relative to the average (detection of outliers)

The sum of all load collectives can be used to determine a specific quantile, in relation to a load-determining characteristic, e.g., the 95% quantile: 95 % of all considered products have a lower load.



### 8.2.3 Data Quality, Data Adjustment

Sufficiently good data quality is a basic prerequisite for any data analysis. Otherwise, the results obtained from the data can lead to false conclusions.

The respective use case determines what is understood by “sufficiently good data quality”. If, for example, a data set is incomplete for a large part of the measured variables, but the data for a characteristic that is essential for gaining information is complete, then there is nothing to prevent the targeted analysis of this characteristic.

Criteria that can be used in assessing data quality include:

- Number of recorded data points vs. number of expected data points (e.g. well applicable at a fixed sampling frequency),
- Number of data gaps (e.g. with scanning per minute: number of gaps > 5 min),
- Maximum time period without data recording within the data set.

Scaled to longer time periods:

- Number of days with data gaps
- Maximum number of days without data per month/year

Depending on the use case and its sensitivity to missing data, appropriate thresholds must be defined, beyond which the data should not be used or should be used only in a restricted manner.

What advantages does the customer get from data collection and Condition Monitoring?

- No more maintenance than necessary, failure-free operation, low service costs
- Short term repair, because the service technician already has prior information about the problem
- Software updates online possible
- Recommendation for suitable device (device class, device program)

What advantages does the manufacturer of the product get?

- Correct constructive design of the product (no under-/over engineering)
- Chances for new service offers and business models?

# 9 Weibull Distribution

At the beginning of the 1950's, under very general conditions, the Swedish engineer W. Weibull [Weibull] developed the universal probability function, from which he could show that in many cases, the latter represented a good description of the properties of real objects:

- Yield strength of a steel
- Size distribution of fly ash
- Fiber strength of Indian cotton
- Fatigue life of a steel

Weibull wrote in [Weibull] that he "has never been of the opinion that this function is always valid". However, the Weibull distribution function has proven itself in many cases as a model and is inseparably connected with the evaluation of lifetime investi-gations today.

Initial basis for a Weibull evaluation are mostly failure times of test objects, subjected to stress up to failure on respectively one or several test benches.

If one subdivides the time axis into intervals of equal size and plots the number of failures in each interval versus the number of products still intact at the beginning of the latter (instantaneous failure rate), then the result will be a representation as in Fig. 14.

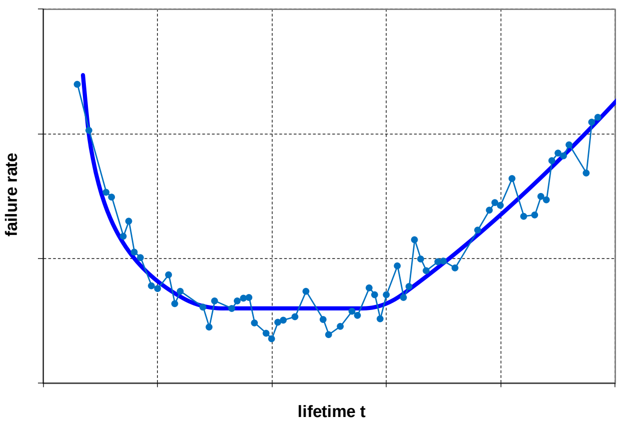



Fig. 14: „Bathtub curce" (schematic)

Due to the similarity with a longitudinal cross-section of a bathtub, the graph of $\lambda(t)$ has the name "bathtub curve".

The empirically determined failure quote (dots) is an estimation of the the-oretical failure rate $\lambda(t)$, illustrated by the continuous curve.

The bathtub curve in general results through superposition of three typical failure modes and is hence subdivided into the following sections:

*NOTE 1: When projected to the life of a human being, the failure modes mentioned above correspond somehow to life phases of childhood (still birth, sudden infant death, death through children sicknesses), adulthood (acci-dents in profession or leisure time, fatal diseases, operations) and retirement (decrepitude, increased danger of accident).*

*NOTE 2: By the Weibull distribution, the three areas can only be described separately. In the literature, therefore, model approaches repeatedly emerge with tree or more parameters which are able to describe the bathtub curve with only one distribution function, e.g. the IDB distribution, proposed by Hjorth (Increasing, Decreasing, Bathtub-shaped).*

In this chapter we give a short introduction into the Weibull distribution and its usage inside risk analysis. Be aware that the Weibull alone is not a risk analysis as long as the key elements are missing which were treated in the prior chapters.

## 9.1 Two-Parameter Weibull Distribution

Parameters of the two-parameter Weibull distribution are the characteristic life $T$ (also denoted as scale parameter) and the shape parameter $b$ . By plotting measured failure times $t_i$ of products, components or assemblies on Weibull paper, estimate values of these two parameters can be determined. Methods for this are given in Section 9.6.

| | |
|---|---|
| $t$ | Lifetime |
| $b$ | The shape parameter or Weibull exponent $b$ determines the slope of the straight lines in the Weibull plot, (this clarifies the German designation “Ausfallsteilheit” and the term “Weibull slope”) and is characteristic for the failure mode. |
| $T$ | The characteristic life $T$ gives the time until 63 % of the products of a population have failed. By substituting $T$ in the distribution function, $F(t)$ , one attains $F(T) = 1 - \frac{1}{e} =$ 63.2 %. |
| $t_0$ | Failure-free time |

**Formulas and Definitions**

| Function | Definition with respect to a population | Definition with respect to a unit |
|---|---|---|
| Probability density function<br>$f(t) = \frac{b}{T} \cdot \left(\frac{t}{T}\right)^{b-1} \cdot e^{-\left(\frac{t}{T}\right)^b}$ | Derivative of the cumulative distribution function.<br>$f(t) \cdot dt$: Population fraction failing within the following time interval $dt$ | $f(t) \cdot dt$: probability that a unit which has survived age $t$ will fail within the following time interval $dt$ |
| Cumulative distribution function (cdf)<br>$F(t) = 1 - e^{-\left(\frac{t}{T}\right)^b}$ | Population fraction failing by age $t$<br>(from the beginning of stress) | Probability that a unit fails by age $t$ |
| Reliability function<br>$R(t) = 1 - F(t)$ | Population fraction surviving a pre-given age $t$ | Survival probability (survivorship) function: probability that a unit survives beyond age $t$ |
| Failure rate<br>$\lambda(t) = \frac{f(t)}{R(t)} = \frac{b}{T} \cdot \left(\frac{t}{T}\right)^{b-1}$ | $\lambda(t) \cdot dt$: fraction of the remainders which fails within the following time intervall $dt$<br><br>(empirically: instantaneous failure rate or failure quota). | $\lambda(t) \cdot dt$: probability that a unit of the remainders which have survived age $t$ will fail within the following time interval $dt$ |

## More Relations

The expectation value of t, i.e. the mean time to failure, is

$$E(t) = MTTF \ = T \cdot \Gamma\left(1 + \frac{1}{b}\right)$$

$t_q$ is the $q$-quantile of the Weibull distribution. This is the time until which the population's proportion $q$ has failed:

$$t_q = MTTF \cdot \frac{\left[-ln\left(1-\frac{q}{100}\right)\right]^{\frac{1}{b}}}{\Gamma\left(1+\frac{1}{b}\right)} \quad \text{or} \quad t_q = T \cdot \left[-ln\left(1 - \frac{q}{100}\right)\right]^{\frac{1}{b}}$$

$t_{0,5}$ is the median of the distribution, also denoted as $B_{50}$: $t_{0,5} = T \cdot \left(ln(2)\right)^{\frac{1}{b}}$



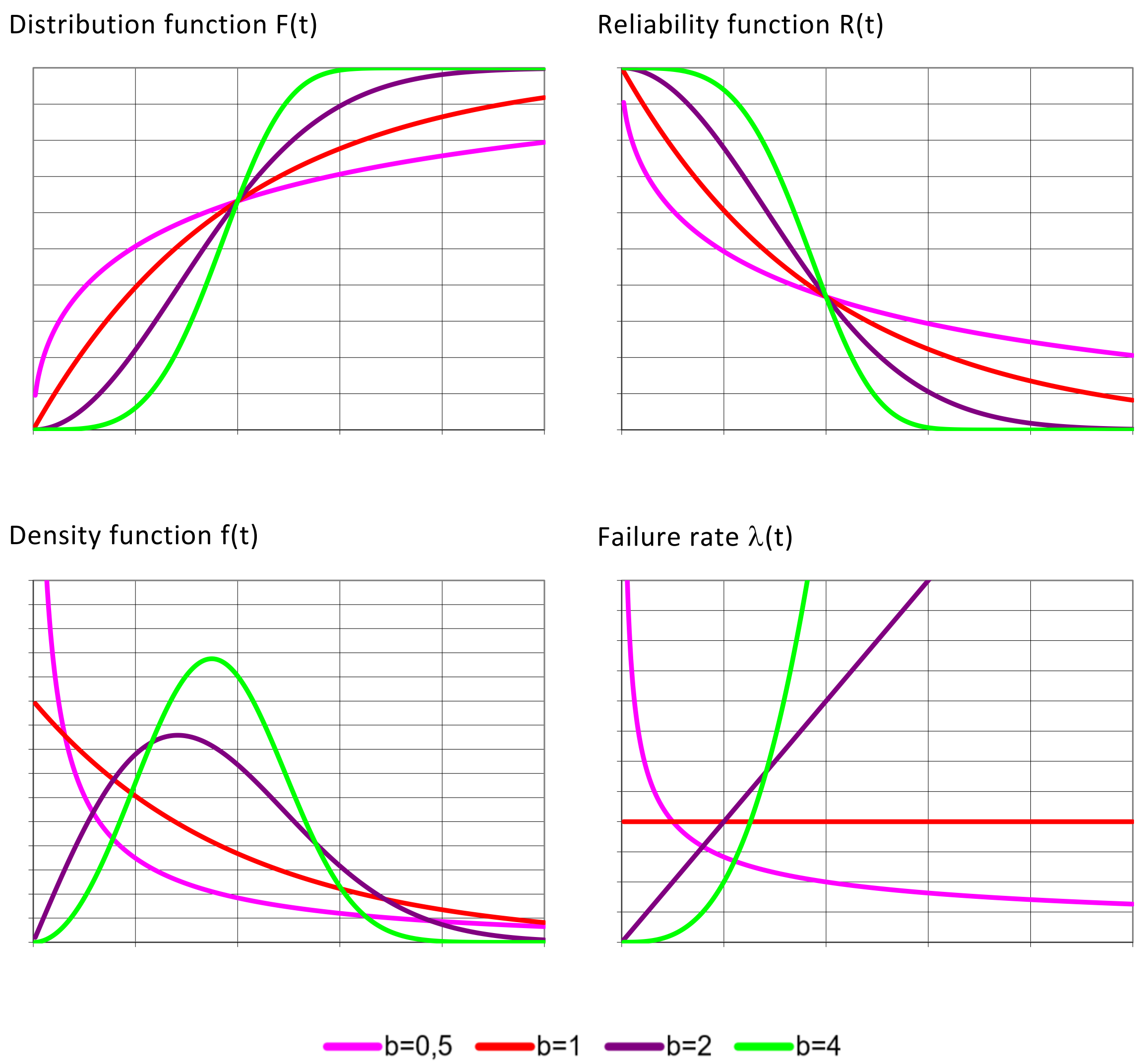



Fig. 15: Courses of F(t), R(t), f(t) und $\lambda$(t) for T=1 and several values of b

## 9.2 Weibull Plots

As the following equation shows, the transformation $ln(-ln(1-F(t))$ transforms Weibull distributed data $(t_i;\ F(t_i))$ into a linear relationship:

$$y_i = ln(-ln(1-F(t_i)) = b \cdot ln(t_i) - b \cdot ln(T)$$

It corresponds to the equation of a straight line $y = b \cdot x + a$

with slope $b$ and intersect $a = -b \cdot ln(T)$.

Therefore, plotting the points $y_i = ln(-ln(1-F(t_i))$ against $ln(t_i)$ results in a representation that approximately corresponds to a straight line with slope b.

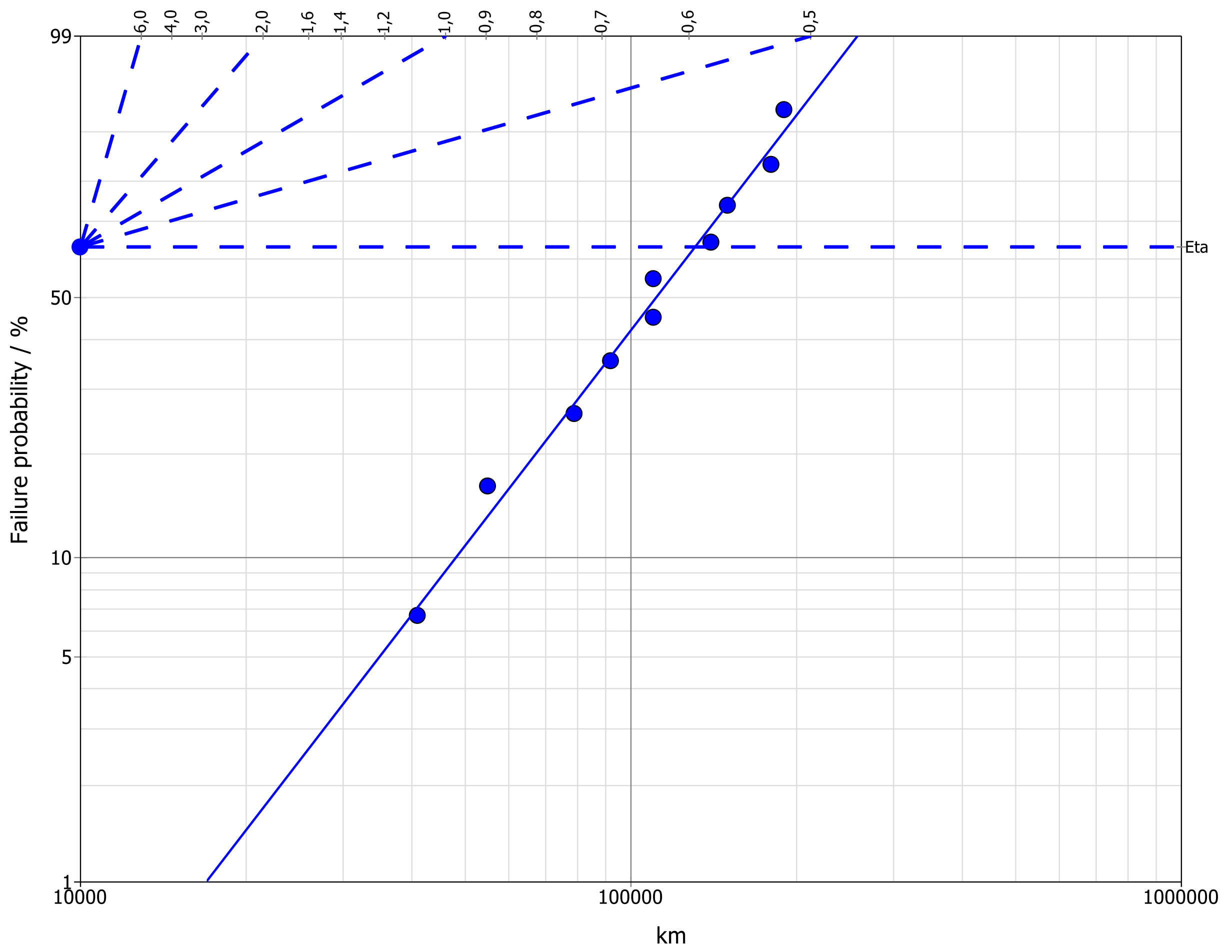



Fig. 16: Weibull plot with ten data points; $b = 2.25$ and $T = 130{,}784\ km$

## 9.3 Interpretation of Slope b

The bathtub curve (Chapter 11) is generally obtained by superimposing three failure types and is therefore divided into the following sections:

- Decreasing failure rate: early failures ($b < 1$)
- Constant failure rate: random failures ($b = 1$)
- Inreasing failure rate: failures due to wear and tear ($b > 1$)

The Weibull distribution allows all three failure modes to be described separately.

When projected to the life of a human being, the failure modes mentioned above correspond somehow to life phases of childhood (still birth, sudden infant death, death through children sicknesses), adulthood (accidents in profession or leisure time, fatal diseases, operations) and retirement (decrepitude, increased danger of accident).



| | |
|---|---|
| $b = 1$ | Constant falilure rate, e.g., the failure behavior doesn't change with time. Example: failures of electronic components. A Weibull distribution with $b = 1$ is identical to the exponential distribution.<br><br>In the area of "random failures", the failure probability is independent of the prehistory of the product. The same proportion of the products still intact at the beginning of each time interval always fails within equivalent time intervals. |
| $b < 1$ | Decreasing failure rate, e.g. in the run-in phase of a product. This behavior with $b < 1$ means that the failure probability is high at the beginning and decreases with time ("teething problems" of a product). The frequently used term "early failures" can be misleading and should not be used in this context. |
| $b > 1$ | Increasing failure rate, e.g. failures due to wear, material fatigue<br><br>In the area of wear-out failures, the failure probability increases constantly. In other words it is increasingly probable that a product still intact up to that time will fail within the subsequent time interval. |

Special cases:

| | |
|---|---|
| $b = 2$ | Linearly increasing failure rate<br>For $b = 2$ the Weibull distribution is identical to the Rayleigh distribution (which is similar to a lognormal distribution). It describes events with mechanical wear like in a gear. |
| $b \approx 3.6$ | For $b \approx 3.60235$ the skewness of the Weibull distribution becomes infinitesimally small and its density function is similar to that of a normal distribution (bell-shaped curve) but is defined only for $x \geq 0$. |

## 9.4 Three-Parameter Weibull Distribution

### 9.4.1 Technical Meaning of the Failure-free Time

Several technical products are described to have an “integrated lifetime”, because a part or several parts of the product become mechanically worn out, e.g. automobile tyres, brake pads and discs, clutches, potentiometers, brushes in electric motors and alternators. In such cases it is possible to de-termine the lifetime from abrasion, which is measured after a predetermined test time or after a certain km-number, and the known “material reserve”. This prediction nonetheless assumes that the law according to which the abrasion occurs (e.g. linear), is at least approximately known.

In case of several failure causes, a certain time $t_0$ must be completed before a failure can occur due to this cause.

In addition to the examples of abrasive wear listed, the following should also be mentioned here.

- Corrosion due to electrochemical processes
- Crack formation and material fatigue fractures under cyclic bending.
- Disruption of solder joints due to temperature changes with the effect of a loose contact or complete failure of an electronic component. This creates fine cracks that can eventually lead to the fracture of the solder joint.

Naturally, one tries to avoid failures due to such known mechanisms by constructive design or suitable maintenance measures. Therefore, they appear rather rarely in the evaluation of field data.

### 9.4.2 Mathematical Consideration of the Failure-free Time

The three-parameter Weibull distribution serves as a model for the statistical description of the phenomena listed in 11.3.1 as examples.

| | |
|---|---|
| $F(t) = 1 - e^{-\left(\frac{t-t_0}{T-t_0}\right)^b}$ | Probability function |
| $f(t) = \frac{b}{T-t_0} \cdot \left(\frac{t-t_0}{T-t_0}\right)^{b-1} \cdot e^{-\left(\frac{t-t_0}{T-t_0}\right)^b}$ | Density function |
| $\lambda(t) = \frac{f(t)}{R(t)} = \frac{b}{T-t_0} \cdot \left(\frac{t-t_0}{T-t_0}\right)^{b-1}$ | Failure rate |

In case $t_0 > 0$ the Weibull plot show an approximative line which is curved towards the X-axis. If the model assumption is correct, a transformation of the form $t^* = t - t_0$ leads to a sequence of points which can be approximated by a straight line. The parameters of interest can be determined on the basis of this straight line.

### 9.4.3 Notes on Parameter b when taking into account $t_0$

Due to the transformation mentioned in Section 9.2, Weibull distributed data is represented as a straight line in the Weibull plot.

The addition of a failure-free time $t_0$ would shift the line to the right by the amount $t_0$ if the time axis was scaled linearly. Because of the logarithmic scaling, however, data points are shifted less to the right the further they already lie to the right (large values of t). Therefore the transformation $t^* = t + t_0$ causes the curvature of the straight line to a convex curve.

When evaluating data with failure-free time $t_0$, we have the opposite situation. Treating the data with the 2-parameter Weibull distribution results in a line that only starts above (to the right of) $t_0$. If the evaluation software fits a straight line to this curve in a purely formal way, it generally has a large slope corresponding to an $b$-value of 6, 7 or more.

Only by subtracting the failure-free time $t_0$ does the curvature disappear and a much smaller value $b$ is usually determined. What does it mean now if $b$ then lies in the range of about 1 or below? Obviously, in this case it does not make sense to speak of random failures and even less of early failures.



## 9.5 Determination of the Distribution Parameters

### 9.5.1 Method of Least Squares

From the function $F(t, b, T)$ follows by twofold logarithmizing

$$ln(-ln(1 - F(t_i)) = b * ln(t_i) - b \cdot ln(T)$$

and plotting the quantity $ln(-ln(1 - F(t_i))$ against $ln(t_i)$ a straight line

$$y = g(x) = a + b \cdot x$$

with slope $b$ and intercept $a = -b \cdot ln(T)$.

Using the least squares method, the slope $b$ and intercept $a$ can be calculated relatively easily. There are two approaches for this (s. [Sachs]).

**Regression from Y on X**

In the regression of Y on X, the sum of squared perpendicular distances is minimized.

Then the slope $b_{yx}$ is:

$$b_{yx} = \frac{\sum_{i=1}^{n} x_i \cdot y_i - \frac{(\sum_{i=1}^{n} x_i) \cdot (\sum_{i=1}^{n} y_i)}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}} \quad \text{with} \quad y_i = ln(-ln(1 - F(t_i)) \quad \text{and} \quad x_i = ln(t_i)$$

and the intercept $a_{yx}$ is:

$$a_{yx} = \frac{\sum_{i=1}^{n} y_i - b_{yx} \cdot \sum_{i=1}^{n} x_i}{n} \quad \text{and due to} \quad a_{yx} = -ln(T) \quad \text{finally} \quad T = T_{yx} = exp(-a_{yx}).$$

**Regression from X on Y**

In the regression of X on Y, the sum of squared horizontal distances is minimized.

Then the slope $b_{xy}$ is:

$$b_{xy} = \frac{\sum_{i=1}^{n} x_i \cdot y_i - \frac{(\sum_{i=1}^{n} x_i)\cdot(\sum_{i=1}^{n} y_i)}{n}}{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}} \quad \text{with} \quad y_i = ln(-ln(1-F(t_i)) \quad \text{and} \quad x_i = ln(t_i)$$

and the intercept $a_{xy}$ is:

$$a_{xy} = \frac{\sum_{i=1}^{n} x_i - b_{xy} \cdot \sum_{i=1}^{n} y_i}{n}$$

By rearranging the above equation $y = a + b \cdot x$ we get

$$x = h(y) = \frac{1}{b} \cdot y - \frac{1}{b} \cdot a = b^* \cdot y - \frac{1}{b} \cdot \left(-b \cdot ln(T)\right) = b^* \cdot y + ln(T).$$

Therefore, the slope $b^* = \frac{1}{b_{xy}}$ and $T^* = T_{xy} = exp(a_{xy})$.

The straight line $g(x)$ tends to be flatter than $h(x)$, i.e. $b \leq b^*$.

### 9.5.2 Median Ranks Approach

In a Weibull plot the quantity $Y = ln(-ln(1-F(t_i))$ is plotted against $X = ln(t_i)$. While the failure times $t_i$ and thus the positions on the X axis are known, the question is how to determine the associated Y values.

The formula for Y contains the failure probabilities $F(t_i)$ of the i-th failure. The best estimations for the corresponding cumulative frequencies $H_i$, which can be allocated to the first, second, ..., n-th failure, are the so-called “median ranks”. They can be calculated using the binomial distribution, by solving the equation:

$$\sum_{i=1}^{n} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} = 50\,\% \quad \text{with} \quad i = 1, 2, 3, \dots, n\,.$$

In general, it is sufficient to use Benard’s approximation: $H_i = \frac{i - 0{,}3}{n + 0{,}4}$

The median rank approach can be explained as follows.

We consider a hypothetical, large population of e.g. 1000 parts and the corresponding failure times $t$. If we put the times in an ascending order, we can number them from 1 to 1000.

Thus, the time $t_i$ corresponds to position $i$ in the sequence. However, we can just as well specify that this position corresponds to the proportion $\frac{i}{1000} \cdot 100\%$ of the population. For example, 20 % of the values $t_i$ are then smaller than $t_{200}$.

We can illustrate the situation by writing the percentages 0.1 %, 0.2 %, …, 100 % on individual balls and putting them into a lottery drum.

If we now draw samples of size 10 very often and order each of them according to size, we obtain a distribution of percentages with a median value at 6.7 % for the first value. The distribution of the greatest value has a median of 93.3 %. So, in this way, for every rank $i$, you get a percentage.

The procedure assumes that after each draw the 10 balls are put back into the lottery drum.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $H_i$ | 6.7 % | 16.2 % | 25.9 % | 35.5 % | 45.2 % | 54.8 % | 64.5 % | 74.1 % | 83.8 % | 93.3 % |

The probability that percentage $x$ appears at position $b$ in the drawing result corresponds to the combined probability that

- there are numbers smaller than $x$ at the positions $1, 2, \dots, i-1$,
- $x$ appears at position $i$, and
- there are numbers greater than $x$ at the positions $i+1, i+2, \dots, 10$.

It is given by the binomial distribution. By summing up, we finally get the probability that there is a number $\leq x$ at the position $i$. This corresponds to the formula given above, with $H_i = p$.

*NOTE 1: The above example explains the mathematical background for the case $n = 10$, but is transferable to any n.*

*NOTE 2: In the same way, you can calculate the best estimators for the first, second, ..., sixth lotto number. These are the numbers 5, 13, 21, 28, 36, and 44. However, it is not possible to use it to increase the chances of winning in lotto. The number 5 simply means that the smallest number in a draw with equal probability of 50 % is less than or greater than 5.*

*NOTE 3: The designation “median ranks” becomes understandable only if one interprets the percentages as “ranks”. The median (50 %-value) then refers in each case to the distribution of these ranks.*

### 9.5.3 Rank Regression

In classical statistics, the linear regression, where the sum of the squared vertical distances is minimized, is called regression from Y on X, as it is described in the upper part of Section 9.6.1.

This assumes that X and Y are independent variables and that the relationship between an explanatory, non-random characteristic X and a target characteristic Y is describable by a function of the form $Y = a + b \cdot X + \epsilon$. In particular, for scheduled tests, X is set to the desired values $x_i$ and the corresponding $y_i$ are measured.

Weibull++ also distinguishes with respect to the rank regression the two approaches described in Section 10.5.1 and recommends the regression in the direction of the axis on which the "uncertain" values are represented.

Generally, a test on a test bench runs until all $n$ tested parts have failed. If the test were repeated several times with $n$ new parts, different values would always be obtained for the time until the first failure. However, the same cumulative frequency $H_1$ is always assigned to the first failure with rank 1. The same applies to the other ranks $i$ and associated $H_i$.The "uncertainty" therefore consists in the values on the X-axis, i.e., X is a random variable (a variate). Weibull++ recommends minimizing the horizontal distances of the points from the straight line in this case, but calls this a rank regression on X (RRX).

Also the software Minitab considers lifetime data plotted on the X-axis as random results and minimizes the horizontal distances of the points from the straight line, when „Least Squares Estimation: failure time(X) on rank(Y)" is selected.

*NOTE: It is unclear why the notation in the Software Weibull++ differs from the usual notation of classical statistics. The results obtained with Weibull++ and Minitab are identical.*

As an advantage of rank regression, it should be mentioned that it can give a measure of how well the fitted line describes the point sequnece in the form of the correlation coefiicient, i.e., how well the model fits the observed data.

When evaluating field data, one usually has to deal with comparatively few failures but a large number of still intact products. For the evaluation of such data, the Maximum-Likelihood method is to be given preference. (s. MLE in Section 9.6.5).

### 9.5.4 Determination of the Weibull Parameters in EXCEL

According to Section 11.4.1, plotting $y_i = ln(-ln(1 - F(t_i))$ against $x_i = ln(t_i)$ results in a straight line with slope $b$ and intercept $a = -b \cdot ln(T)$.

Slope and intercept can be easily calculated in EXCEL in the following way (given here for the regression Y on X):

$b = SLOPE(Y\_values;\ X\_values)$

$T = EXP(-INTERCEPT(Y\_values;\ X\_values)/b)$

### 9.5.5 Maximum Likelihood Method

The Maximum Likelihood Estimation (MLE) is a numerical method for determining estimate values for the parameters of a presumed distribution.

By this method, $b$ and $T$ are determined so that under the assumption of a Weibull distribution, the probability will be at the maximum just to observe the measured values $t_i$ as found.

This probability is expressed by the so-called likelihood function $L$. The best possible estimate for $b$ and $T$ are those parameter values for which $L$ attains a maximum value. The best possible approximating straight line to the points in the Weibull plot corresponds to these values.

Details can be found in [Booklet 13] and [VDA 3.2], for example.

In addition, with the help of this method, one can determine a 95%-confidence interval. The limiting lines of this confidence interval are presented as curved lines above or below the approximating straight line in a Weibull plot.



*NOTE 1: In contrast to the regression method, this approach does not require failure probabilities and is therefore independent of the rank distribution.*

*NOTE 2: The equation* $\frac{\partial ln(L)}{\partial b} = 0$ *cannot be solved analytically. However, the solution b can be approximated, e.g. using Newton's method. With known* $b, T$ *then results from the following expression:* $T = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} t_i^b\right)^{1/b}$ *(see also [DIN 61649])*

*NOTE 3: The ML estimations for b and T react sensitively to outliers.*

## 9.6 Confidence Intervals

The term confidence interval denotes an interval calculated from sample values, which covers the true but unknown parameter of a distribution with a predetermined probability, the confidence level. For example, 95 % is chosen as the confidence level. This probability means that in case of a frequently repeated and permissible application of this method the calculated confidence intervals roughly cover the parameter in 95 % of the cases and does not cover it in only 5 % of the cases.

In connection with the Weibull plot, a confidence interval is a two dimensional area that is constrained above and below the calculated best fitting straight line by curved lines, the confidence limits. All imaginable straight lines in this interval correspond to Weibull distributions, which can produce the actually observed points as random results. An example is shown in Fig. 18.

In [VDA 3.2], methods or formulas are given, with whose help confidence intervals for the parameter $b$ and $T$ of the Weibull distribution, and the $B_{10}$-life (generally $B_q$-life) can be calculated (s. also [Sachs]).

Confidence limits in the Weibull plot can also be calculated with the help of the likelihood function $L$ (Section 9.6.5).

## 9.7 Exponential Distribution

The exponential distribution is a model for describing random failures, which can often be applied when dealing with electronic components, for example.

This distribution results as a special case of the Weibull distribution with form parameter $b = 1$. Its distribution function is $F(t) = 1 - e^{-\lambda \cdot t}$. Therein, $\lambda$ denotes the (constant) failure rate.

The exponential distribution has proven itself effectively for expressing radioactive decays. The validity of this law has always been reconfirmed in numerous physical experiments.

In connection with lifetime investigations, the exponential distribution can conversely only be valid for a limited time period ("bottom of the bathtub curve") since every technical product will finally fail due to wear, aging and fatigue after a sufficiently long time (the failure rate $\lambda$ will then be time-dependent).



When evaluating a data set in a Weibull plot, the value $b = 1$ will practically never be exactly attained. For practical applications, according to [VDA 3.2], in the case of $0.5 \leq b \leq 1.2$ one can presume "random failures".

# 10 Evaluations Based on the Weibull Distribution

## 10.1 Example of a Completed Test

Let us consider a completed test in which 30 vehicles were running till they failed due to the same failure mechanism. The mileage of each reached are recorded.

After 28 failures the test is stopped. In this case our task is not to predict when the next car will stop but we want to get a statistical model about the lifetime of the cars regarding this failure mechanism.

Here are the data collected as multiples of 1,000 km, in ascending order. The numbers are to be read column by column from top to bottom and from left to right.

| | | | | | | |
|---|---|---|---|---|---|---|
| 14 | 17 | 23 | 27 | 30 | 35 | 45 |
| 15 | 22 | 25 | 30 | 31 | 38 | 48 |
| 16 | 23 | 27 | 30 | 34 | 38 | 52 |
| 16 | 23 | 27 | 30 | 35 | 41 | 53 |

The following approximation formula can be used to calculate the associated relative cumulative frequency $H_i$ for each i-th failure.

| i | km | $H_i$ / % |
|---|---|---|
| 1 | 14,000 | 2.5 |
| 2 | 15,000 | 6.0 |
| 3 | 16,000 | 9.5 |
| ... | ... | ... |

$$H_i = \frac{i - 0.3}{n + 0.4} \cdot 100\ \% \qquad \text{with} \qquad n = 28$$

Alternatively, other approximation formulas for $H_i$ are also common.

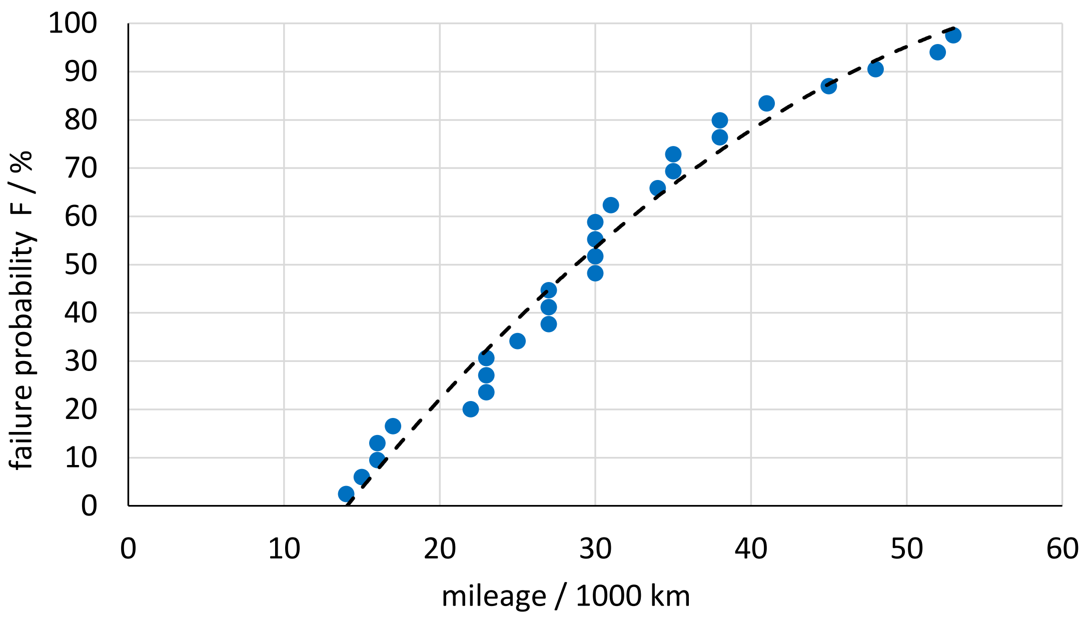



Fig. 17: Determination of an appropriate probability distribution F(t) for the given failure data

Since both axes are linearly scaled, the point sequence in Fig. 17 shows a curved course. According to Section 9.2, we find a straight line with slope $b$ and intercept $a = -b \cdot ln(T)$, when plotting $ln(-ln(1-F(t_i))$ against $ln(t_i)$.

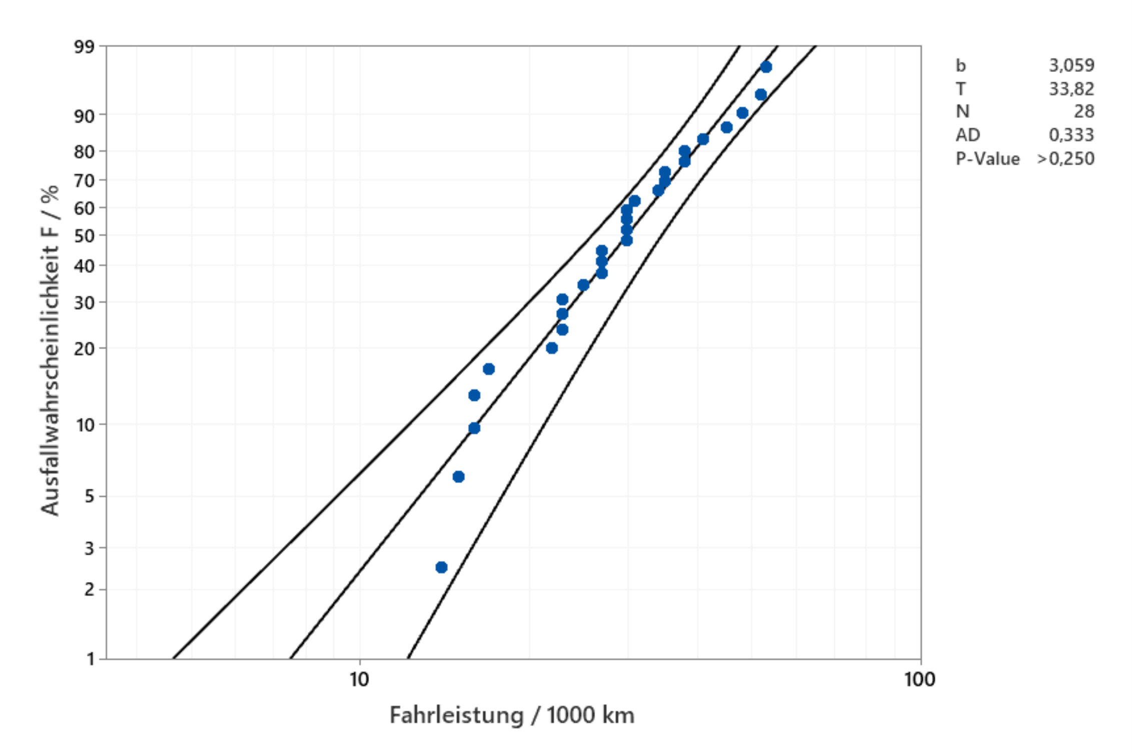



Fig. 18: Weibull plot of the sample data. Here the failure probability $F(t)$ is plotted over the logarithmized lifetime $ln(t)$. In the present case the lifetime characteristic is the mileage until failure.

The slope $b$ and characteristic lifetime $T$ are displayed in the diagram when using the relevant statistics programs. However, they can also be easily calculated using the formulas in Section 9.6.1.

In the present case we get $b = 3.1$ and $T = 34{,}000\ km$.

## 10.2 Data Sources

The example in Section 10.1 is about a data set for a completed test. The goal here is essentially to evaluate the test result technically. In contrast, the goal of field data evaluation is in particular to make a prediction of the further product behavior at an early stage, when failures are observed. So we are dealing with failure data at a point in time when only a portion of the population has failed by then.

Thus, the number of the last failure is not identical with $n$ in Section 10.1. It is instead obvious to choose the number of produced parts for $n$. But even in this you would make a mistake, because not all produced parts are also sold and are in the field.

So, in general, there is information about the parts that have failed, but there is no information about the parts that are still in operation.



For a sound modeling of the field events, at least

- the monthly production numbers,
- the distribution of mileages (or operation hours) of the population
- and an estimation of delay times (delay of sales, reporting delay)

are usually required for vehicles. See also Section 2.4.

The reference quantity $n$ significantly influences the evaluation result and its correct determination is not an easy task.

Further aspects may have to be taken into account in the calculation:

- Only a part of the population has gained high operation times at the time of observation.
- Does a partial market factor play a role?
- Is the restriction to a specific period of manufacturing always reasonable?

In case of vehicles a type related consideration is recommended. This also applies if we consider several equal parts per vehicle. Only the first failure of each vehicle is evaluated.
Since the data acquisition is frequently done on a monthly basis also the Weibull analysis should be updated monthly. It is not useful to do several up-dates in the same month.

## 10.3 Sudden-death Method for Field Failures

Sudden-death testing is a time-saving method used for investigating failure behavior in bench tests (cf. [VDA 3.1], [Booklet 13]). The total number of test items is divided into a number of subgroups. Each subgroup is then tested until one item fails in each subgroup.

As outlined in [VDA 3.1], this method can also be applied to field failures. To facilitate this, the failures reported are viewed as the first failures occurring within artificial subgroups of "test items". In forming these artificial subgroups, the $n_f$ failures reported are distributed evenly across the total number of items, $n$, produced during the reference period. Each subgroup comprises approximately $k = \frac{n - n_f}{n_f + 1} + 1$ products. The $k - 1$ items within a subgroup that did not fail are treated as suspended units.



**Example:**

PD: 10/98 Number of failures: $n_f = 24$ Production total: $n = 7{,}902$

Subgroup size $k = \frac{n - n_f}{n_f + 1} + 1 = \frac{7{,}902 - 24}{24 + 1} + 1 \approx 316$

| Failure No. | km until failure | Failure No. | km until failure | Failure No. | km until failure | Failure No. | km until failure |
|---|---|---|---|---|---|---|---|
| 1 | 500 | 7 | 1,800 | 13 | 4,100 | 19 | 9,800 |
| 2 | 600 | 8 | 2,900 | 14 | 4,500 | 20 | 10,400 |
| 3 | 900 | 9 | 3,000 | 15 | 5,000 | 21 | 11,500 |
| 4 | 930 | 10 | 3,050 | 16 | 5,100 | 22 | 12,000 |
| 5 | 1,000 | 11 | 3,300 | 17 | 6,000 | 23 | 14,000 |
| 6 | 1,500 | 12 | 4,000 | 18 | 6,900 | 24 | 20,000 |

So, in the present case, for each failed unit, there are 315 units that have not failed up to the corresponding number of km.

However, some software programs for life data analysis allow users to enter both the failed and non-failed items and enable direct analysis of such subgrouped data by means of the maximum likelihood method (see [VDA 3.1], [Booklet 13]).

Fig. 11.3 shows the result of such an evaluation.

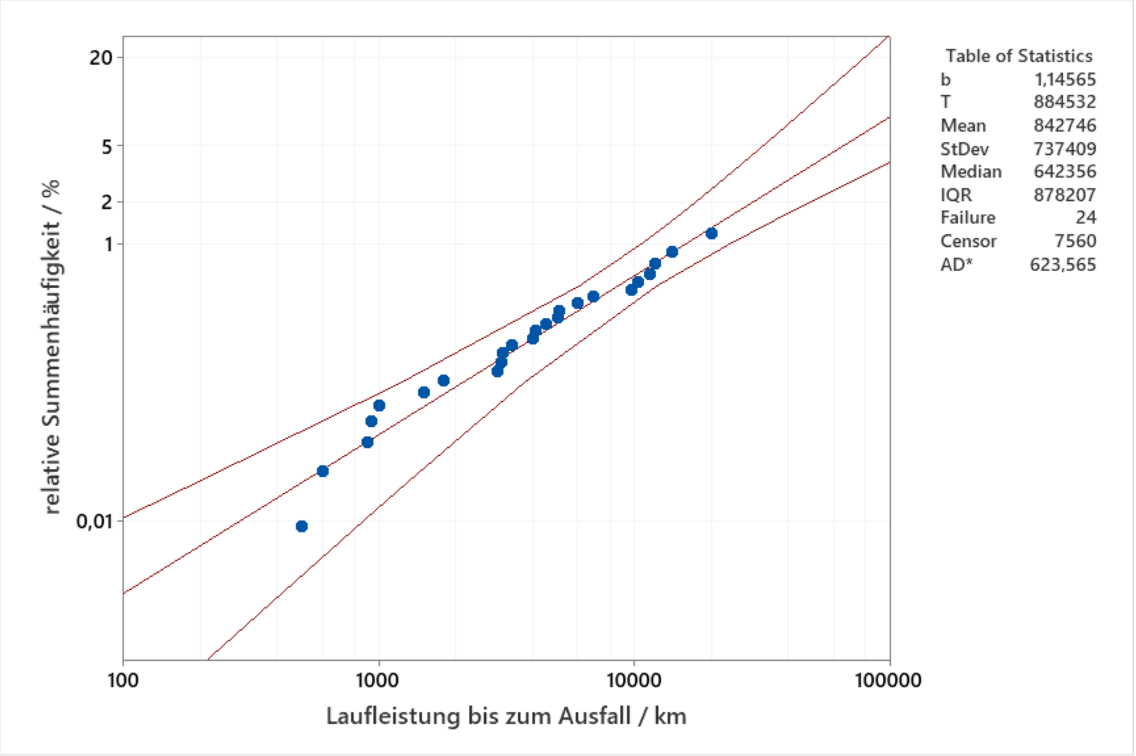



Fig. 19: Result of an evaluation with the Maximum Likelihood Method, incl. confidence limits on the 95 % level.

Rounded results:

| b | T | $B_1$ | $B_5$ | $B_{10}$ |
|---|---|---|---|---|
| 1.15 | 885,000 km | 26,000 km | 66,000 km | 124,000 km |

*NOTE: This method assumes that all the "suspended" units of a group have exactly the same mileage as the unit that has failed. Since all groups have the same size (here, k = 316), an assumption is made that the number of parts that have performed a mileage of 4,000 km is the same as the number of parts that ran 20,000 km. In other words, the annual mileage distribution (cf. Chapter 8) is not taken into account.*

*Mixed distributions do not yield a straight line with constant b, but yield instead a curve. The graphical procedure described in [VDA 3.1] then no longer applies.*

As the chart shows, the relative cumulative frequencies are in the area below 1 %. An extrapolation, to determine T, for instance, corresponds to an extension of the straight line far beyond this area. Regardless of whether this determination is made graphically or numerically, there is a great deal of uncertainty associated with it.

[VDA 3.1] describes a graphical method for analyzing such data on Weibull paper. After plotting the points corresponding to the failures, a line of best fit is drawn which represents the “first failures”. Parallel shifting of this “line of first failures” to make it pass through a calculated point then yields the required line for the Weibull distribution of the population.

To this straight line belongs the value $B_{50} = T \cdot (ln(2))^{\frac{1}{b}}$.

The first failure of a group of $k = 316$ units is assigned the cumulative frequency

$H_1 = \frac{1 - 0.3}{316 + 0.4} \cdot 100\ \% \ \approx 0.221\ \%$

on the Y-axis and the median $B_{50}$ of the first failures on the X-axis.

By a parallel shift of this “line of first failures” through the point $(H_1;\ B_{50})$ we finally find the searched straight line to the Weibull distribution of the population.

The charakteristic life can be calculated using the formula for $t_q$ from Section 9.1:

$T = t_q \cdot \left[-ln\left(1 - \frac{q}{100}\right)\right]^{-\frac{1}{b}}$ with $q = H_1$

## 10.4 Possible Analysis for Constant Failure Rates

The preceding example with data from 10/98 yields a value of $b = 1.1$ for the shape parameter. Since the confidence range for b includes $b = 1$, the following deliberations assume that this case corresponds to the random failure scenario ($b = 1$) meaning that the failures are independent of the mileage. Similar analyses performed for other production months have shown this assumption to be justified.

The difference between the purchase date and the complaint date corresponds to the vehicle's "operating time" (of course, this does not, in this case, imply round-the-clock operation). Using the mileage shown on the vehicle's odometer, it is then possible to calculate the average annual mileage of this vehicle. This can yield very high annual mileage projections (far in excess of 100,000 km/60,000 miles) which do not appear realistic. Since the median is much less sensitive to such outliers than the arithmetic mean, use of the median, rather than the mean, is recommended.

The median for the data set used in Section 10.3 is approximately 20,100 km per year. Division by twelve yields an estimate for the average mileage per month, in this case approximately 1,675 km/month.

For each production date from July 1997, the age in months as of September 1999 was determined, e.g. 9/99 – 7/97 = 26 months. By 9/99, a vehicle with a production date of 7/97 will thus have an expected mileage of approximately 26 · 1,675 km = 43,550 km. Multiplying this mileage figure by the production total for this month yields the total mileage covered by all the vehicles manufactured in 7/97, as of 9/99.

This calculation was performed for each production date, to calculate the total mileage of all the manufactured parts (cf. table below).

Remembering that there is a time gap of approximately 1 month between the production and purchase dates, the true operating time of the parts will actually be one month less. Hence the calculation was repeated, with each age reduced by one month.

Calculation for each PD and summation of total mileage for all the manufactured units, e.g.:

| PD | Prod. Total | Age (months) | $PM \cdot Age \cdot 1{,}675$ | $PM \cdot (Age - 1) \cdot 1{,}675$ |
|---|---|---|---|---|
| ... | ... | ... | ... | ... |
| 1/99 | 13.190 | 8 | $1.8 \cdot 10^8\ km$ | $1.5 \cdot 10^8\ km$ |
| 2/99 | 12.903 | 7 | $1.5 \cdot 10^8\ km$ | $1.3 \cdot 10^8\ km$ |
| 3/99 | 21.576 | 6 | $2.2 \cdot 10^8\ km$ | $1.8 \cdot 10^8\ km$ |
| ... | ... | ... | ... | ... |
| 8/99 | 15.343 | 1 | $2.8 \cdot 10^7\ km$ | 0 km |
| | | | $\Sigma \quad 1.94 \cdot 10^9\ km$ | $\Sigma \quad 1.61 \cdot 10^9\ km$ |

Result: $1.94 \cdot 10^9\ km$ Since 5/98 there were 102 failures.

Failure rate: $\lambda = \frac{102\ failures}{1.94 \cdot 10^9\ km} \approx \frac{53\ ppm}{1{,}000\ km}$

Taking into account the storage delay, the operating time will be reduced by one month, i.e. the age of each unit will be one month less. Failure rate: $\lambda = \frac{102\ failures}{1.61 \cdot 10^9\ km} = \frac{63\ ppm}{1{,}000\ km}$

## 10.5 Analysis Using the Mileage Distribution

Prerequisites:

- All failures are known, no partial market.
- The period under analysis should be so long ago that all vehicles have a fairly similar operating time.
- The mileage distribution is known (probability plot or histogram).

The first step consists in defining appropriate classes for the mileage data.

The reported failures are then allocated to the various classes. In addition, the mileage distribution is used to determine how many of the non-failed parts are likely to belong to each of the various classes. This means that we will have a figure for both the number of failed parts and the number of non-failed parts in each mileage class.

[VDA 3.1] presents an example with a median-rank analysis based on the Johnson method. Using appropriate software, an analysis according to the maximum likelihood method is equally possible.

In analyses performed during the warranty period, the prerequisite of the cars having fairly similar operating times is not fulfilled. Section 11.7 describes some alternatives for this case.

## 10.6 Data Summary

The procedure shown in Section 10.3 assumes that the vehicles have been in operation for approximately the same amount of time. An evaluation is therefore only possible for a specific month in which the vehicles were manufactured (PD) or registered (RD).

However, since the mostly few failures usually occur on vehicles of different ages, the data basis for an evaluation is very small if only a single month is considered.

In order to evaluate all failures of vehicles from several production or registration months together, it is necessary to determine the distribution of the mileages of all vehicles of the considered population up to the time of evaluation.

For the following consideration we assume that the yearly mileage $L_{12}(x)$ is known (see Section 7). If $L_{12}(x)$ is a lognormal distribution with geometric mean $\bar{x}_g$ and shape parameter $\varepsilon$ the logarithmized mileages $z = ln(x)$ are normally distributed. We denote the associated distribution by $F_{12}(\bar{z}_{12};\ s_{12})$. So it has the mean $\bar{z}_{12}$ and the standard deviation $s_{12}$.

The mean value of the distribution $F_j(s)$ after an operating time of j months then is:

$$\mu_j = \mu_{12} + ln\left(\frac{j}{12}\right).$$

Thus, the mileage distribution after 6 months has the mean value

$$\mu_6 = \mu_{12} + ln\left(\frac{6}{12}\right) = \mu_{12} - 0.693.$$

The mean value therefore shifts to the left by the value $-0.693$.

The standard deviation is $s_j = s_{12}$.

The perhaps somewhat surprising fact $s_j = s_{12}$ can be explained as follows.

The equation $ln(a \cdot x_i) = ln(a) + ln(x_i)$ implies

$$\bar{z} = \overline{ln(a \cdot x_i)} = \frac{1}{n} \cdot \sum_{i=1}^{n} ln(a \cdot x_i) = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(ln(a) + ln(x_i)\right) = \bar{z} + ln(a)$$

$$s(z) = s\left(ln(a \cdot x)\right) = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(ln(a \cdot x_i) - \bar{z}\left(ln(a \cdot x)\right)\right)^2}$$

$$s(z) = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(ln(a) + ln(x_i) - \bar{z} - ln(a)\right)^2}$$

$$s(z) = s\left(ln(a \cdot x)\right) = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(ln(a \cdot x_i) - \bar{z}\left(ln(a \cdot x)\right)\right)^2}$$

$$s(z) = s\left(ln(a \cdot x)\right) = s\left(ln(x)\right)$$

The standard deviation is the same as for the yearly mileage.

Because of this one-to-one mapping the number of vehicles in the individual km classes can be determined via the normal distribution.

## 10.7 Projections During the Warranty Period

[Reichmann] describes a method that can be used for an approximate calculation of short-term prognoses, provided that the Weibull parameter $b$ is known for the product in question. The method takes both the storage time and the reporting delay into account. This paper presents two possible ways to estimate the failure rate to be expected by the end of the warranty period, by weighting the known failure totals or by weighting the production quantities. Creating suitable EXCEL spreadsheets makes the practical implementation of these methods very much easier.

A simple projection model which also takes the annual mileage distribution into account is described in [Pauli]. If the distribution $L(s)$ of the annual mileage is known (cf. 6.2), then it is possible to calculate, for each mileage s at which a failure occurred, the proportion $1 - L(s)$ of vehicles from the relevant population which have not yet reached this mileage and can still fail. The reciprocal value $\frac{1}{1-L(s)}$ of this proportion is the projection factor by which the reported number of failures at mileage s has to be multiplied in order to obtain the corrected (projected) figure.



## 10.8 Weibull Evaluation of a Stair-Step Table

| Beginning of quarter | End of quarter | Production Volume: | 1,163 | 1,357 | 1,439 | 2,046 | 1,357 | 1,112 | 1,943 | 1,429 | 11,846 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Production Quarter: | 1Q.2017 | 2Q.2017 | 3Q.2017 | 4Q.2017 | 1Q.2018 | 2Q.2018 | 3Q.2018 | 4Q.2018 | Σ |
| 0 | 1 | Reporting quarter 1 | 9 | 14 | 13 | 19 | 1 | 7 | 15 | 3 | 81 |
| 1 | 2 | Reporting quarter 2 | 21 | 42 | 47 | 43 | 28 | 16 | 37 | | 234 |
| 2 | 3 | Reporting quarter 3 | 15 | 20 | 46 | 23 | 25 | 22 | | | 151 |
| 3 | 4 | Reporting quarter 4 | 18 | 20 | 17 | 21 | 18 | | | | 94 |
| 4 | 5 | Reporting quarter 5 | 15 | 15 | 10 | 33 | | | | | 73 |
| 5 | 6 | Reporting quarter 6 | 7 | 6 | 22 | | | | | | 35 |
| 6 | 7 | Reporting quarter 7 | 9 | 13 | | | | | | | 22 |
| 7 | 8 | Reporting quarter 8 | 6 | | | | | | | | 6 |
| | | | | | | | | | | Σ | 696 |

Table 7: Stair-step table; status: end 2017

Based on the failures added up in the last column and the total production quantity, an evaluation is possible within the framework of the Weibull model.

So all the 81 parts that failed in the 1st quarter were in operation for a maximum of 1 quarter.

Using the approximation formula $F(t_i) = \frac{i - 0.3}{n + 0.4}$ with i = 1, 2, 3, …, 696 and n = 11,846

the corresponding function values of the Weibull distribution for the failure times $t_i$ can be determined.

The times are identical line by line, i.e.

$t_i = 1$ for i = 1, 2, 3, ..., 81

$t_i = 2$ for i = 82, 83, 84, ..., 315

$t_i = 3$ for i = 316, 317, 318, ..., 466

...

$t_i = 8$ for i = 691, 692, 693, ..., 696

In the Weibull plot, the markings for the failures in each quarter therefore lie vertically one above the other.

The parameters of the Weibull distribution can be determined using the least squares method (Section 9.6.1).

No distinction is made between different failure types. Also different operating hours or mileages are disregarded here.

A constructive change in one production quarter, e.g. to eliminate a failure cause, is not taken into account in the above evaluation. It therefore makes sense to evaluate together only those production quarters that correspond to a certain status of the product.

## 10.9 Weibull Evaluation to Isochrones Charts

If a line perpendicular to the time axis is drawn in the isochrones chart over a production quarter, the intersection points of the perpendicular lines with the isochrones correspond to the cumulative frequencies $F(t_i)$ for the different operating times $t_i$ (MIS).

Thus, just that information is given which is needed to create a Weibull plot. However, because the isochrones to the larger MIS values end earlier and earlier, the number of support points $[t_i;\ F(t_i)]$ for the younger production quarters is smaller the further to the right they lie on the time axis. [VDA 3.3]

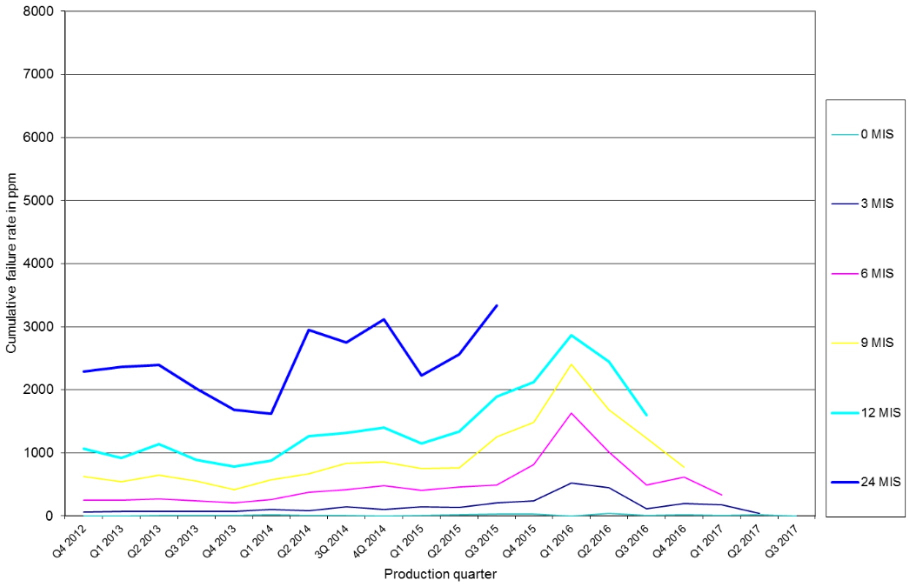



Fig. 20

## 10.10 Mixture Distribution

A mixture distribution of failure times occurs when products in a population, which are subject to different failure modes, are put together (e.g. due to different manufacturing conditions or material charges). This means, for instance, that the 1st group only fails due to the cause A and the 2nd group only due to the cause B. In Weibull's model, these two failure modes correspond to different slopes $b_1$ and $b_2$ and/or different characteristic lives $T_1$ and $T_2$.

If data from such a mixture distribution are plotted on Weibull paper, one does not get a straight line but mostly gets a curve with a curvature bending away from the time axis. So long as the cause of failure can be determined for each part, it is possible to perform separate Weibull analyses after the respective separation of the data.

If one mixes two populations of size $n_1$ or $n_2$ whose distribution functions are $F_1(t)$ and $F_2(t)$, then one gets a distribution function of the mixture:

$$F(t) = \frac{n_1}{n} \cdot F_1(t) + \frac{n_2}{n} \cdot F_2(t) \quad \text{with} \quad n = n_1 + n_2.$$

In general: $F(t) = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i \cdot F_i(t)$ with $n = \sum_{i=1}^{k} n_i$.

Based on the density functions (relative frequencies) it is analogous to:

$$f(t) = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i \cdot f_i(t).$$

Another situation lies on hand when an individual product can fail due to different reasons. In this case, one speaks of competing failure modes.

The reliability function $R(t)$ of a part is then equal to the product of the reliability functions based on individual $k$ failure modes:

$$R(t) = R_1(t) \cdot R_2(t) \cdot \ldots \cdot R_k(t).$$

Weibull analysis of a data set based on non-distinguishable, competing failure modes (which as such are not distinguishable through investigation of the parts), is only possible with computer aid.



## 10.11 Separation of Different Failure Mechanisms

The following Weibull plot illustrates failures on a vehicle component without considering the causes (failure mechanisms). It can be seen that the sequence of points deviates significantly from the best-fit line.

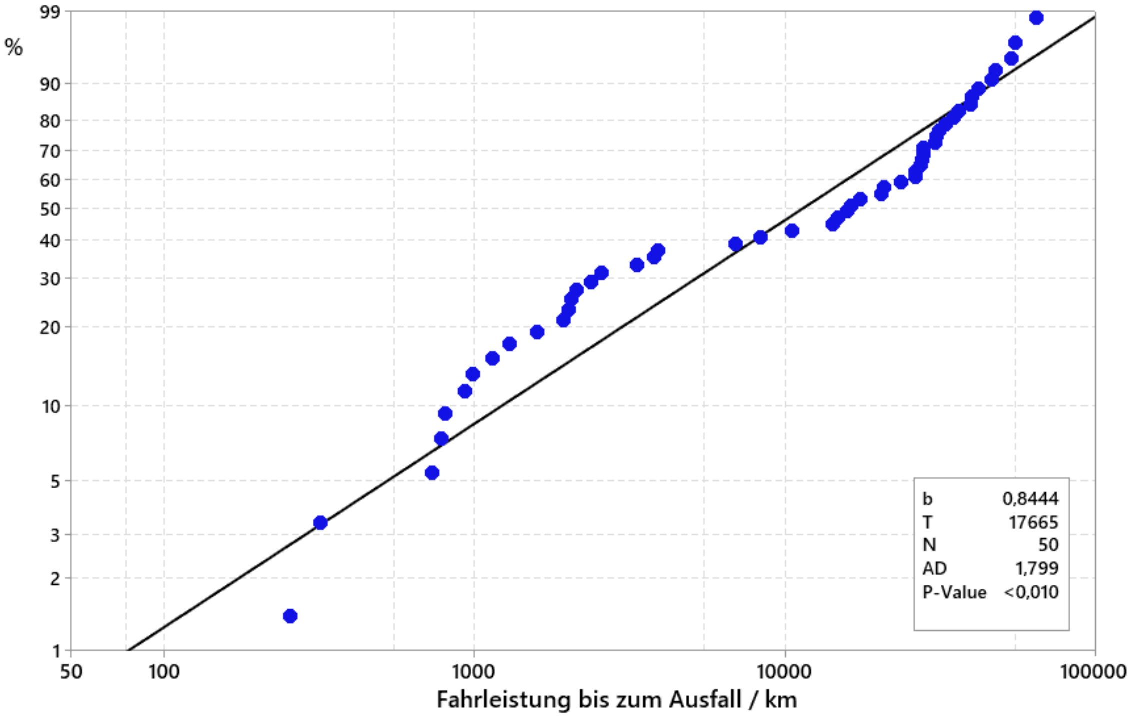



Fig. 21 Weibull diagram of field data without considering the failure cause

As expected, the points appear to follow a piecewise convex or concave curve, which is different from a straight line.

This is a mixed distribution (s. Section 10.10) of the form:

$$F(t) = \frac{20}{50} \cdot \left(1 - e^{-\left(\frac{t}{T_1}\right)^{b_1}}\right) + \frac{30}{50} \cdot \left(1 - e^{-\left(\frac{t}{T_2}\right)^{b_2}}\right)$$

A careful investigation of the parts showed that they failed because of two completely different failure mechanisms. The distribution of lifetimes of a component or product is following the Weibull distribution only if there is only one mechanism which determines the lifetime. Therefore, it is necessary to evaluate each failure mechanism by separate Weibull diagram. The result is displayed in the following figure.

This results in widely separated straight lines with clearly different slopes and characteristic lifetimes.

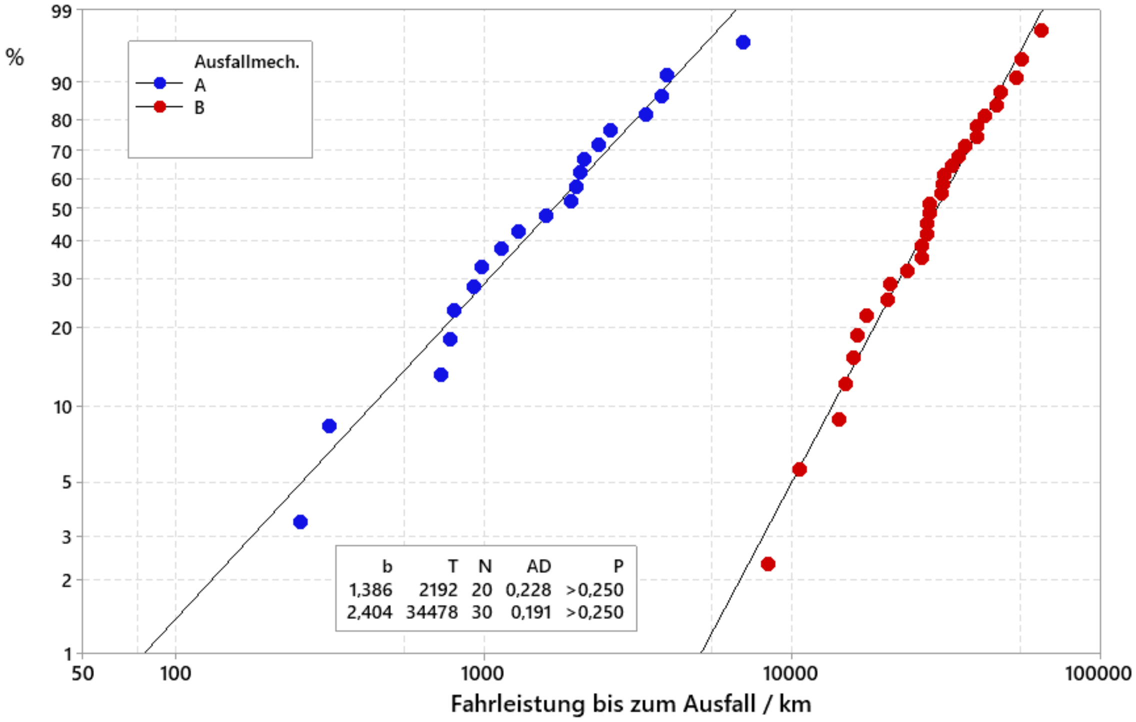



Fig. 22: Weibull diagram of the same field data after separating regarding failure mechanisms A and B.

# 11 Error Sources

## 11.1 Affected Proportion q

So far it has been tacitly assumed that finally all parts will fail due to the considered cause. However, in practical problems irregularities occur in a specific time period at a specific location or production line. The the error will refer only to a certain affected proportion $q$.

The cumulative frequency $F(t)$ then does not reach the limit value 100 %, but approaches the value $q$ according to the product $q \cdot F(t)$.

## 11.2 Uncertainty in the choice of model

The future is not predictable. In particular, for the time without failure and the affected proportion $q$ one has to accept the following: The observed data in an early stage allow different interpretations. The task is to narrow down the range for the model by using of additional context knowledge.

Every data input is a potential error source. Assuming that failure reports are complete and correct the uncertainty has two main sources:

- insufficient knowledge of the processes in the field (partial market factor, distribution of mileages, delay times) and
- wrong selection of the model ($t_0$, $q$) due to due to misconceptions about the failure mechanism.

# 12 Simulation of Field Data

If we interpret the uniformly distributed random numbers generated with a random number generator in an interval [0, 1] as the values of the cumulative relative frequency $F(t)$ and insert them in the following equation

$$t = T \cdot \left[ln\left(\frac{1}{1-F(t)}\right)\right]^{\frac{1}{b}} + t_0$$

then we obtain the Weibull distributed failure times $t$, with $b$, $T$ and $t_0$ as parameters of the Weibull distribution.

*NOTE: If b is in cell A4 and T in cell A2, the EXCEL formula is: =A$2*POWER(LN(1/(1-RAND()));1/A$4)*

If one has estimated the values of the parameters $b$ and $T$ (if applicable, also $t_0$) on the basis of available real data, one can examine with the help of such a simulation whether the observed field data and representations can be reproduced in the Weibull plot.

In addition, the mileage distribution, the reporting delay and the affected proportion, for example, must be taken into account.

# 13 Terms and Abbreviations

## 13.1 Symbols and Acronyms

| | |
|---|---|
| b | Parameter of the Weibull distribution |
| CS | Control Sampling |
| FLRA | First Level Risk Analysis |
| LoRA | Levels of Risk Analysis |
| LP | Linear Prediction |
| LT | Lifetime (Test) |
| LT-Tool | Tool to calculate step 3 from a lifetime test |
| | |
| GDPR | General Data Protection Regulation |
| GWA | Global Warranty Analyzer |
| MIS | Months in Service |
| PD | Production Date |
| PF | Projection Faktor |
| ppb | parts per billion |
| ppm | parts per million |
| PQ | Production Quarter (quarter of production) |
| PTiS | Part Time in Service |
| PV | Production Volume (produced quantity, number of pieces) |
| RQ | Reporting Quarter |



## 13.2 Terms and Definitions

**0-km-/0-h failure**

0-km failures are in general failures of products, which occur before delivery to the end customer (e.g. at the original equipment manufacturer; also inspection goods and stock failures). This also applies analogously to products where the operating time is of interest (0-h failures). The exact definition and delimitation against field failures can vary depending on the area and product.

**8D method**

8D is a problem solving approach by a team working on eight necessary steps or disciplines. At the beginning it is important to describe the problem, i.e. the deviation from the desired state, in detail. In the following steps, the aim is to find a short-term remedy and carefully determine the causes. Finally, longer-term effective and preventive measures are taken to ensure that the same problem cannot occur again in the future.

**activation mechanism**

Physical, chemical, mechanical, situative processes or prerequisites that are necessary for a failure to occur. Typically, each step during the propagation of the deviation has a specific activation mechanism.

**aging failures, wear-out failures**

Here, aging in the broadest sense means a change in structure, composition, or properties due to a natural time-dependent process. Aging failures in the scope of Weibull's theory are characterised by a shape parameter $b > 1$. Examples for aging processes:

- Transportation processes: diffusion, vaporisation, whisker formation, formation of deposits, material movement under the influence of electric fields (electromigration)
- Chemical reactions: oxidation, sulphide formation, polymerisation, cor-rosion
- Physical and structural changes: crack growth, fracture, plastic deformation, material fatigue, creep, adhesive joints becoming loose, welding, recrystallisation
- Wear: mechanical abrasion

**failure**

In general, failure means that a unit (product) considered free of fault at the beginning, loses specified properties during a time interval of a certain stress. Failure can mean a complete breakdown (no function; result e.g. a vehicle not able to start; so-called hard failure) or an impermissible deviation (due to timely drift) from a specified nominal value or range of values (also characteristic line; e.g. defective lambda-control without perceptible power loss; soft failure).

**affected parts / volume**

All parts that are affected by the technical root cause. They have to be distinguished, for example, from the parts that were produced in the considered time-period.

**B-life**

Age at which a given percentage of items have failed

NOTE: "$B_{10}$"-life is the age at which 10 % of items (e.g. bearings) have failed. Sometimes it is denoted by the L (life) value. B lives may be read directly from the Weibull plot or determined more accurately from the Weibull equation. The age at which 50 % of the items fail, the $B_{50}$ life , is the median time to failure. [DIN 61649]

**clean date**

Production date from which a defect has been eliminated and production is free of defects again

**claim delay**

Period between repair and notification of the OEM or supplier. [VDA 3.3]

**complaint**

Expression of dissatisfaction made to an organization related to its product or service, or the complaints-handling process itself, where a response or resolution is explicitely or implicitely expected [ISO 9000]

**damage mechanism**

By a damage mechanism one understands processes that lead to a gradual change of a unit's properties due to stresses. [VDA 3.2]

**dependability**

Ability to perform as and when required [ISO 9000]

Dependability is used as a collective term for the time-related quality characteristics of an item. Dependability includes availability, reliability, recoverability, maintainability, and maintenance support performance. Cf. [IEC 60050-191]

Reliability is the product property to perform a required function over a specified service life. The service life is divided into operating and downtime. [VDA 3.2]

Characteristics (nature) of an item with regard to its suitability to meet the reliability requirement during or after specified periods of time under specified application conditions. [DIN 40041]

**expectation**

Integral of the product of $x$ and probability density $f(x)$ above the number line

NOTE: The expectation of a continuous probability distribution is denoted by $E(X)$ and calculated as follows: $E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx$

The expectation of a discrete probability distribution is denoted by $E(X)$ and calculated as follows: $E(X) = \sum_{i=1}^{n} x_i \cdot p(x_i)$

In accordance with [DIN ISO 3534-1]

**failure mode**

[AIAG FMEA] also uses the terms failure or failure of a function. Manner in which an item could fail to meet or deliver the intended function. Failure modes are derived from the functions.

Examples: loss of function, degradation of function, intermittent function, partial function, unintended function, exceeding function, delayed function, persisting on a certain value, wrong direction

In accordance with [AIAG FMEA]

**failure quota**

Number of failures in a time interval related to the number of survivors at the beginning of that interval and the interval size

**field**

The entirety of the natural environment, situations and conditions during the use phase at the end customer. In particular, this includes the actual application, operating and environmental conditions to which the products under consideration are subject. The term "field" in a narrower sense can also be applied to the totality of all users. The conditions "in the field" are not repeatable in contrast to artificially created laboratory conditions.

> *NOTE: The term "field", which is difficult to define, was probably taken from the field of empirical social sciences. The collection of sociological data is done, for example, through questioning, observation or experimental studies. Terms such as field observation and field research are derived from this.*

**field action**

Field action is the generic term for recall and service/customer service action.

**field complaint**

Complaint after transfer of the complaint product to the dealer organization / end cus-tomer (unless agreed differently with the customer)

**field data**

In this context, field data refers to the totality of all data that is generated in the field in connection with the use of a product. In a narrower sense, this includes all data associated with errors, faults, defects and failures that lead to customer complaints. In a broader sense, however, this also includes information on usage such as operating hours, driving times, consumption, load collectives or customer feedback in the form of evaluations, experiences, wishes and suggestions for improvement.

**field failure**

Effect that is observable or experienced by the end customer and does not reflect the intended, expected, or wanted behavior of the product.

**Fix-as-fail**

Fix-as-fail means that a product is only repaired or replaced when it no longer performs its function. The opposite strategy would be of a preventive nature, i.e. a part is regularly serviced or replaced before it is likely to fail: "change before fail".

**goodwill**

Material or financial compensation given by Bosch without legal obligation (contract, leg-islation)

**guarantee**

This is the manufacturer's independent promise to the relevant owner of the product not just to the direct contract partner as in the case of the warranty to take action in the event of faults. The scope of the guarantee is determined by the manufacturer himself. The owner, not only the manufacturer's contract partner, therefore also has a direct claim with respect to the manufacturer.

**GWA**

RB abbr.: Global Warranty Analyzer

GWA is a web-based tool for evaluating, displaying and reporting customer warranty data.

**IoT, Internet of Things**

A global infrastructure for the information society, enabling advanced services by inter-connecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

**IQIS**

RB abbr.: Integrated Quality Improvement System

IQIS is a SAP based IT system mainly used in the BBM for the processing of external (and also internal) complaints with the 8D method.

**lifetime**

The lifetime is the time during which a unit can be exposed to a damage mechanism uniterruptedly until failure [VDA 3.2]

> *EDITORIAL NOTE: In the metrological sense, service life is a quantity that has a value that can be expressed by a number and a reference. Instead of the quantity time (duration of use, operating hours, months, years), other quantities are also used in practice, such as distance travelled (mileage indicated in km or mls, number of load cycles, bending stresses, actuations, switching operations, operating cycles, revolutions.*
>
> *It is common practice to speak of lifetimes in the sense of times even when using one of these alternatives, which is practically unavoidable in connection with the fixed terms "characteristic lifetime" or $B_{10}$-life, for example.*

**load**

Sum of the mechanically, chemically, thermally and electromagnetically induced loads that are applied externally on the product.

**load collective, load spectrum**

A load spectrum (load collective) is a classified totality of measured data or counted values that indicates the absolute or relative frequencies of loads to which a product is subjected or subjected during operation. A load spectrum is usually determined empirically by using suitable classification methods to determine how frequently stresses of a certain magnitude are contained in the stress-time function.

Ultimately, such a load spectrum for a characteristic can thus be represented as a frequency diagram or histogram. A load spectrum therefore no longer contains any information about the sequence of the loads. However, the sequence can have an influence on the damage progression.

While it is never repeated in an identical manner in real operation, it is repeatable under largely reproducible conditions in the context of test bench endurance runs.

[Booklet 13] contains a chapter on "Quantitative Reliability Assessment" and remarks on load spectra.

**load cycle**

A load cycle describes the load time curve of a dynamic load. In case of a harmonic (sinusoidal) oscillation it corresponds to a complete oscillation period. For non-harmonic oscillations (in the area of fatigue strength) half oscillation cycles "from peak to peak" are determined and combined to complete oscillation cycles.

**MTTF, mean lifetime**

$MTTF$ designates the mean time (empirical mean, expectation value of the lifetime distribution) between the initial operation and the time of failure.

**OEM**

Abk. engl.: Original Equipment Manufacturer

Mostly used synonymously with automotive manufacturer

**operating time, stress duration**

Operating time is the time during which a damage mechanism takes effect. It generally consists of proportions of the operating and downtime, as the operating time does not need to match the stress duration. For example, certain electronic components are still energized or subject to corrosion even with the vehicle taken of the road. [VDA 3.2]

**operating time**

The operating time is the time throughout which a unit works according to its intended purpose. In contrast, the unit is not operated during downtime. [VDA 3.2]

Time interval for which the item is in an operating state

NOTE: "Operating time" is generic, and should be expressed in units appropriate to the item concerned, e.g. calendar time, operating cycles, distance run, etc. and the units should always be clearly stated.

[DIN 61649], [IEC 61649]

> *EDITORIAL NOTE: It can be seen that this definition gives a rather general definition of operating time. While the calendar time probably corresponds approximately to the stress duration according to [VDA 3.2], the number of operating cycles and the distance run are operating times according to the understanding of [VDA 3.2].*



**ordinal characteristic**

Qualitative characteristic for whose characteristic values an ordinal relationship exists.

[DIN 55350-12]

> *EDITORIAL NOTE: An ordinal characteristic can be assigned values on an ordinal scale.*
>
> *EXAMPLES:*
>
> - *The characteristic "clothing size" with the values XXS, XS, S, L, XL, XXL, ..., 7XL,*
> - *The characteristic "(European) shoe size" with the values 14, 15, 16, ..., 48, 49.*
>
> *The values of these characteristics can be classified and differentiated, e.g. using the relationships "smaller than", "is equal" or "greater than".*

**partial market**

A partial market is a market segment (e.g. a country, a region) from which all products subject to complaint are submitted by the end users via the dealer network of the original equipment manufacturers.

**partial market factor**

The partial market factor is a ratio smaller than one, the reciprocal of which serves as an extrapolation factor to infer the total market from an "observed" submarket.

**probability plot**

The probability plot is a distribution-specific diagram in which the y-axis is selected in such a way that a straight line is formed when data from a distribution underlying the plot is displayed. Thus a given data set can be checked graphically for compatibility with the selected distribution model. [Booklet 3]

## range

Difference between two adjacent extreme values of an oscillation; vertical distance peak-to-peak according to DIN 50100; difference between maximum and minimum deflection

### rank

The values of an ordinal characteristic are subject to a natural sequence. Each value takes a fixed position in comparison to the others.

In the above example of clothing sizes, the value "XL" is in the fifth place, it has the rank 5. One also says that the rank number is five. The corresponding value XL is called order statistic.



### recall

Any measure aimed at actively (e.g. by means of a letter to the end consumer or public information) recalling a product already supplied to the end consumer.

> *NOTE: Recall can also be the complete removal of the product from the market, the exchange, correction, repair or sorting of it. A recall is frequently undertaken for safety or legal reasons and with the involvement of the competent authorities.*

### risk

Combination of the probability of occurrence of harm and the severity of that harm

[DIN EN ISO 12100]

### risk analysis

Combination of the specification of the limits of the machine, hazard identification and risk estimation [DIN EN ISO 12100]

### risk evaluation

Judgement, on the basis of risk analysis, of whether the risk reduction objectives have been achieved [DIN EN ISO 12100]

### risk assessment

Overall process comprising a risk analysis and a risk evaluation [DIN EN ISO 12100]

### Rayleigh distribution

Special case of the Weibull distribution for $b = 2$

**service / customer service action**

Any measure whereby a product already delivered to a customer is examined or changed without a recall/withdrawal

*NOTE: Service / customer service action can relate to the entire population affected (e.g. repair in connection with regular maintenance) or be limited to the portion that the complaint refers to (so-called "fix as fail"). A service / customer service action is frequently undertaken for customer satisfaction reasons.*

**service life**

The service life is devided into operating and downtime. The operating time is the time throughout which a unit works according to its intended purpose. In contrast, the unit is not operated during downtime. [VDA 3.2]

**stress**

Total or part of the impacts to which the unit is or may be exposed [DIN 40041]

Local effect of the load on the design element with respect to the considered damage/failure mechanism (e.g. induced voltage, temperature distribution or mass conversion during a chemical reaction)

**time in service**

Time interval (since commissioning) in which the product is in use at the end customer. In the automotive sector, the quantity MIS, Months in Service is typically used.

**TRC, Technical Root Cause**

Reasons for admitting the interaction of causing conditions for the problem/fundamental problem, which are proven by logical (why?) and functional (how?) relations.

**warranty**

The contractual or legal duty of the manufacturer and seller to assume responsibility for product defects.

# 14 Literature


[CDQ 0907] CDQ 0907 Complaint Process
(Central Directive Quality, only available to RB internally)

[AIAG FMEA] AIAG & VDA FMEA Handbook, 2019

[DIN 40041] DIN 40041 Zuverlässigkeit — Begriffe (dependability concepts; German only)

[DIN 45667] DIN 45667 Klassierverfahren
(Classification methods for evaluation of random vibrations; German only)

[DIN 61649] DIN EN 61649 Weibull-Analyse (engl. version: IEC 61649 Weibull analysis)

[ISO 12100] DIN EN ISO, Safety of machinery General principles for design Risk assessment and risk reduction

[IATF 16949] Automotive Quality Management System Standard —Quality management system requirements for automotive production and relevant service parts organizations, IATF 16949

[ISO 9001] DIN EN ISO 9001:2015, Quality management systems — Requirements (ISO 9001:2015)

Verband der Automobilindustrie VDA (publisher: German Association of the Automotive Industry) — Volume 3: Reliability Assurance of Car Manufacturers and Suppliers

[VDA 3.1] Part 1: Reliability Management

[VDA 3.2] Part 2: Reliability methods and tools

[VDA 3.3] Part 3: Case Studies in the Reliability control loop

[VDA Field] VDA Volume: Joint Quality Management in the Supply Chain — Marketing and Service — Field failure analysis

[VDA OtA] VDA Volume: Recall Management Using Over-the-Air Updates

[VDA Rekl] VDA Volume: Standardized Process for Handling Customer Complaints

Booklet Series “Quality Management in the Bosch Group,Technical Statistics”,

[Booklet 2] Booklet No. 2, Discrete Characteristics

[Booklet 3] Booklet No. 3, Evaluation of Measurement Series

[Booklet 12] Booklet No. 12, Terms of Quality Engineering

[Booklet 13] Booklet No. 13, Methods of design for reliability, verification and validation

[Booklet 16] Booklet No. 16, Problem Solving

[Haibach] E. Haibach, Betriebsfestigkeit, Springer-Verlag, Berlin

[Hartung] J. Hartung, Statistik, Oldenbourg-Verlag

[Köhler] M. Köhler et. al., Zählverfahren und Lastannahme in der Betriebsfestigkeit, Springer-Verlag, Berlin

[Stockinger] K. Stockinger: Datenfluß aus dem Feld, in: Masing (Hrsg.), Handbuch Qualitätsmanagement, Hanser-Verlag München, 1994, S. 681ff

[Sachs] J. Hedderich, L. Sachs: Angewandte Statistik, Springer-Verlag

[Pfeifer] T. Pfeifer: Qualitätsmanagement, Hanser-Verlag, München, 1993, S. 294ff

[Pauli] B. Pauli: Eine neue Methode zur Bestimmung der kilomerterabhängigen Lebensdauerverteilung von Kfz-Komponenten. ATZ Automobiltechnische Zeitschrift 100 (1999) Heft 4

[Reichmann] A. Reichmann: Praktische Verfahren zur Zuverlässigkeitsbewertung von Kfz-Komponenten, Diplomarbeit, Gesamthochschule Wuppertal, Fachbereich Sicherheitstechnik, 2000

[Weibull] W. Weibull: Fatigue testing and the analysis of results, Pergamon Press, 1961

[Weibull] A Statistical Distribution Function of Wide Applicability, J. Appl. Mech. 18, 1951, 293-297

# 15 Index

2021-05-05 - SOCOS

**Robert Bosch GmbH**
C/QMM Tilsch
Wiener Strasse 42-46
70469 Stuttgart
Germany
Phone: +49 711 811 - 0
**www.bosch.com**

2021-05-05 - SOCOS

**Robert Bosch GmbH**
C/QMM Tilsch
Wiener Strasse 42 - 46
70469 Stuttgart
Germany
Phone +49 711 811 - 0
**www.bosch.com**