Estimating for Life

Three Transformative Enhancements for Life Cycle Cost Models

Fabian Eilingsfeld, Birgit Biehn, Christian Schindler

2024 SCAF/ICEAA International Training Symposium
Queen Elizabeth II Convention Centre, London

20 September 2024

Paper ITS05 – Handout

Industrieanlagen-Betriebsgesellschaft mbH
Einsteinstrasse 20
85521 Ottobrunn
Germany

Phone +49 89 6088 3748
eMail eilingsfeld@iabg.de • biehn@iabg.de • schindlerc@iabg.de

Estimating for Life: Three Transformative Enhancements for Life Cycle Cost Models

Fabian Eilingsfeld, Birgit Biehn, Christian Schindler
IABG mbH, Ottobrunn, Germany

Abstract. *Since the 1970s, software solutions for parametric cost estimation of products and services have been available, and the German Armed Forces (Bundeswehr) have long utilized these tools to project expenditures in advance. Traditionally, these solutions have focused more on estimating procurement costs, with less emphasis on calculating the cost of operation and support (O&S). However, with the Federal Ministry of Defence's initiative for comprehensive Life Cycle Cost Management (LCCM), the demand for more accurate cost estimates for the O&S phase has grown significantly. In response to this need, IABG mbH began developing new requirements in 2016 for an enhanced, LCCM-compliant cost estimation model. This paper explains how such an improved tool helps to model the top three "known unknowns" in the estimation of operation and support (O&S) costs:*

- *Increased failure rates due to early failures and ageing effects*
- *Additional costs for condition-based maintenance because no faults are found in repairable units*
- *Product obsolescence and its impact on replenishment spares cost*

The result is an improved accuracy in life cycle cost (LCC) estimation, which is intended to support optimal decision-making with a strong emphasis on economic efficiency.

0 Life cycle cost management in the Bundeswehr

The German Bundeswehr's Life Cycle Cost Management (LCCM) initiative is a strategic approach that covers the entire life cycle of defence systems.

> *"According to the Life Cycle Cost Management Directive (ZDv A-1510/1), LCCM is a business management method that involves the use of life cycle costs (LCC) for decisions in the cycle of products or services." (Bundeswehr, 2015, p. 4)*
>
> *LCCM serves to "reinforce the economic focus of decision-making within the Federal Ministry of Defence." (Bundeswehr, 2022b, p. 4)*

The aim of LCCM is to optimise the life cycle costs (LCC) of a system, from development and procurement through to operation and support (O&S).

0.1 Situation: Life cycle costs are the basis for decision-making

Life Cycle Cost Management (LCCM) starts with the design and planning of new systems. Comprehensive cost estimates and analyses should be conducted early in the process to ensure that long-term financial implications are considered during decision-making.

Before a solution is selected, life cycle costs for each proposed solution are planned (Bundeswehr, 2022b, p. 7). The wording is important here, because "planned" is not the same as "estimated". In principle, LCCM distinguishes between estimating, planning and collecting cost:

- Estimation: Forward-looking ("ex ante"), long-term time horizon (up to several decades), forecast of future expenditure up to the end of life, high to very high uncertainty
- Planning: Forward-looking ("ex ante"), short to medium-term time horizon (a few years), appropriation of line items for the federal budget, low uncertainty
- Collection: Rearward-looking ("ex post"), recording of all expenses already incurred over the life cycle, very low to no uncertainty

Estimating, planning, and collecting Life Cycle Costs (LCC) is a continuous process that gains increasing detail and accuracy as a project or programme progresses through its life cycle. Life Cycle Cost Management (LCCM) employs three distinct time horizons – short, medium, and long-term. Naturally, the further into the future an event is projected, the more challenging it becomes to accurately assign costs. The Bundeswehr's approach to improving the modelling of Operations and Support (O&S) costs directly addresses this inherent challenge (Bundeswehr, 2022a, p. 6).

- Situation: The operation and support (O&S) phase accounts for the largest share of life cycle costs; O&S cost should therefore be given appropriate consideration in the selection decision!
- Problem: Estimates of O&S cost are subject to higher uncertainty.
- Implication: How can the estimation accuracy of O&S cost be increased?
- Need: Implementation of all feasible measures to improve the modelling of O&S cost.

This causal relationship highlights the critical importance of accurately estimating operation and support (O&S) costs. They significantly impact the long-term budgetary resources needed and the overall cost-effectiveness of products and services. Consequently, the Bundeswehr must strive to develop realistic cost estimates for the O&S phase of any proposed solution ***before*** selecting the most viable alternative. These estimates should then be continuously refined throughout the implementation of the solution and across its entire life cycle.

0.2 Challenge: Estimating operation and support cost

Life cycle costs (LCC) as a selection criterion for proposed solutions are not new (Blanchard, 1978, p. 2). Quite a few textbooks on LCC have been published since the 1970s (e.g. Nowlan and Heap, 1978; Moubray, 1997). Some of the methods proposed therein have been implemented in various cost models for products and services (PRICE, 2001; PRICE, 2003). The Bundeswehr has long used commercial software to estimate the LCC of proposed solutions in advance. Solutions include products, services, projects and programmes in accordance with the process for Customer Product Management (CPM; Bundeswehr, 2018, p. 20).

The cost estimation software tools available on the market to date have traditionally focussed on procurement costs. These can be estimated with sufficient accuracy because the cost drivers in the development and production of systems are well understood. In contrast, the estimation of operation and support (O&S) costs in traditional cost models have been using a simplified approach:

- Units have constant failure rates, all failures are random failures
- Units are only repaired after they have failed (corrective maintenance)
- Spare parts remain unchanged over the entire service life and are technically identical to the units from the initial procurement

Operation and support costs are so difficult to estimate because in practice they are influenced by many variable factors that are difficult to predict. Costs can change over time due to market fluctuations, inflation, changes in technology, political decisions and other external influences. In addition, unexpected problems and adjustments may arise during operation that also affect cost. As the useful lives of Bundeswehr systems are very long, up to 50 years or longer, estimates of their O&S cost are subject to considerable uncertainty.

0.3 Solution: Modelling of "known unknown" cost drivers in operation and support

Several factors influencing the cost of a system's operation and support (O&S) are considered "known unknowns". This means that although their occurrence and impact on life cycle costs can be anticipated, their timing and extent remain uncertain. Despite long-standing awareness of these cost drivers, traditional cost models have largely ignored them.

The need for enhanced cost models for the O&S phase has grown since the Bundeswehr's initiative for comprehensive Life Cycle Cost Management (LCCM). In response, IABG mbH commissioned a study in 2016 (PRICE, 2016) to assess the requirements for a more advanced, LCCM-compliant cost estimation model.

As part of this study, the top three new functionalities were identified, targeting the most critical "known unknowns" affecting O&S costs:

1. Increased failure rates, primarily due to ageing effects
2. Additional costs for condition-based maintenance when no faults are found
3. Changes in spare parts costs due to product and technology obsolescence

The cost impacts of these three key factors will now be predicted using LCCM-capable software tools.

1 The cost of increased failure rates: The “bathtub curve”

The so-called “bathtub curve” (see Figure 1) visualises the failure behaviour of systems in use over time. The curve shows that the number of failures is relatively high at the beginning and end of the service life, while the probability of failure is lower in the middle of the service life. The curve resembles a bathtub in appearance, which explains its name.

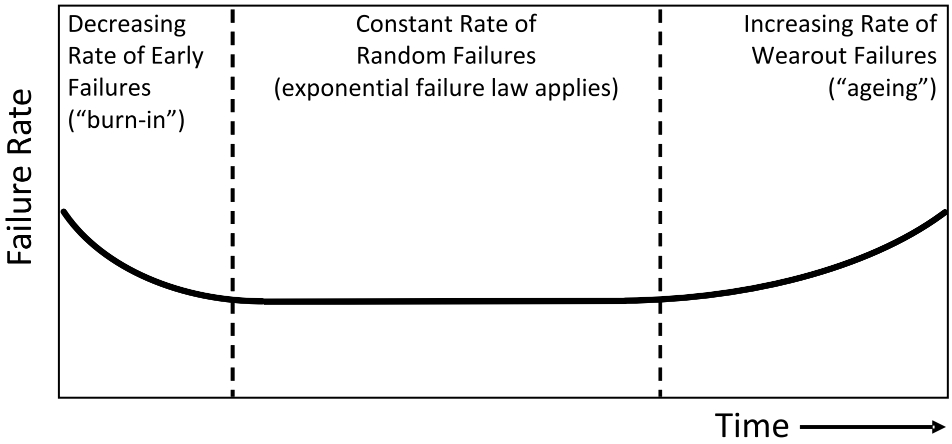


Figure 1 The bathtub curve (redrawn from Blanchard, 1978, p. 232)

The bathtub curve is problematic in terms of maintenance, as it highlights specific challenges and risks during different phases of a system’s service life:

- Decreasing early failures (start of the bathtub curve): Newly commissioned systems tend to fail more frequently. This may be due to manufacturing defects, a lack of quality assurance or installation problems (see Figure 1, left)
- Stable failure rates with random failures (centre of the bathtub curve): During this phase, systems tend to exhibit stable failure behaviour (see Figure 1, centre)
- Increasing age-related failures (end of the bathtub curve): In this phase, the probability of failures increases again as systems wear out and gradually reach the end of their service life (see Figure 1, right)

The random failure rate is calculated as the quotient of the total operating hours in the reporting period and the average operating time between two consecutive failures (Mean Time Between Failures (MTBF); see Eq. 1). Only random failures that obey the exponential failure law are considered when calculating the MTBF (Blanchard, 1978, p. 238).

Eq. 1 $$Random\ Failure\ Rate = \frac{Total\ operating\ hours\ in\ reporting\ period}{MTBF}$$

The challenge is to maintain the reliability and safety of systems even when systems and their components temporarily exhibit increased failure rates. This has an impact on operation and support (O&S) costs.

1.1 "Bathtub curve": the challenge

As far as the cost increase due to increased failure rates is concerned, the more important part of the bathtub curve is the "wearout" with typical age effects towards the end of the service life. Age effects have an impact on operation and support (O&S) costs and represent a challenge for decision-makers. In particular, age effects carry great weight when deciding whether and when it is cost-effective to replace an ageing system (Greenfield and Persselin, 2002, p. 19ff).

Since the 1970s, the useful life of many complex systems has been extended well beyond their originally planned service life (Pfingsten and Engelking, 2021). These extensions can be planned or unplanned. They usually occur because a replacement for a system is not available in time. The extension of the service life may mean that a higher proportion of systems will eventually be affected by wear and tear. Although random failures are still common, increasing wear and tear leads to an increase on the right-hand side of the bathtub curve in Figure 1.

There are numerous empirical studies on the bathtub curve, both on early failures (Crowley, 2004, p. 98; Pyles, 2003, pp. 64–65) and on age-related failures (Crowley, 2004, p. 285; Dixon, 2006, p. 27ff; Francis and Shaw, 2000, p. 4–14; Humphrey, 2002, p. 3; Kiley, 2001, p. 22; Stoll and Davies, 1993, p. 30; Trunkey, 2018, p. 1; Trunkey, 2019, p. 10–18; Valentin, 2003, p. 3ff).

The conclusion of numerous studies is: The bathtub curve reflects real conditions and has a significant impact on operation and support (O&S) cost.

It is essential to understand that the mechanisms underlying the left and right sides of the bathtub curve differ significantly. These distinctions are briefly outlined below.

Early failures

- These can extend over a period of up to ten years, with the typical period being two years
- The total failure rate at the start of operations can be up to four times the random failure rate; a good rule of thumb for the starting value is to double the random failure rate

Age-related failures

- These set in after ten to 15 years of use, often before the end of design life
- The observed annual growth rates range from 0.6 % to 14 %
- It is uncertain whether the annual growth will be linear or exponential!

The underlying factors driving these observable effects will be incorporated into an LCCM-compliant cost model.

1.2 "Bathtub curve": Implementation in the LCCM cost model

The bathtub curve is often explained in the scientific literature using Weibull curves (Sandborn, 2017, p. 260). Failure models based on Weibull curves are widely used in reliability and maintenance engineering (Sandborn, 2017, p. 296; Smith, 2011, p. 71ff). There, they are applied to analyse failure rates in retrospect ("ex post"), based on previously recorded operating data (Ben-Daya, 2009, pp. 61–63). However, they are far too detailed and therefore unsuitable for the predictive ("ex ante") estimation of operation and support (O&S) costs of new systems. Since cost estimators are generally not experts in reliability and failure behaviour, a simplified bathtub curve was configured for this purpose (see Figure 2). This has since been released as a software extension (PRICE, 2022, p. 21).

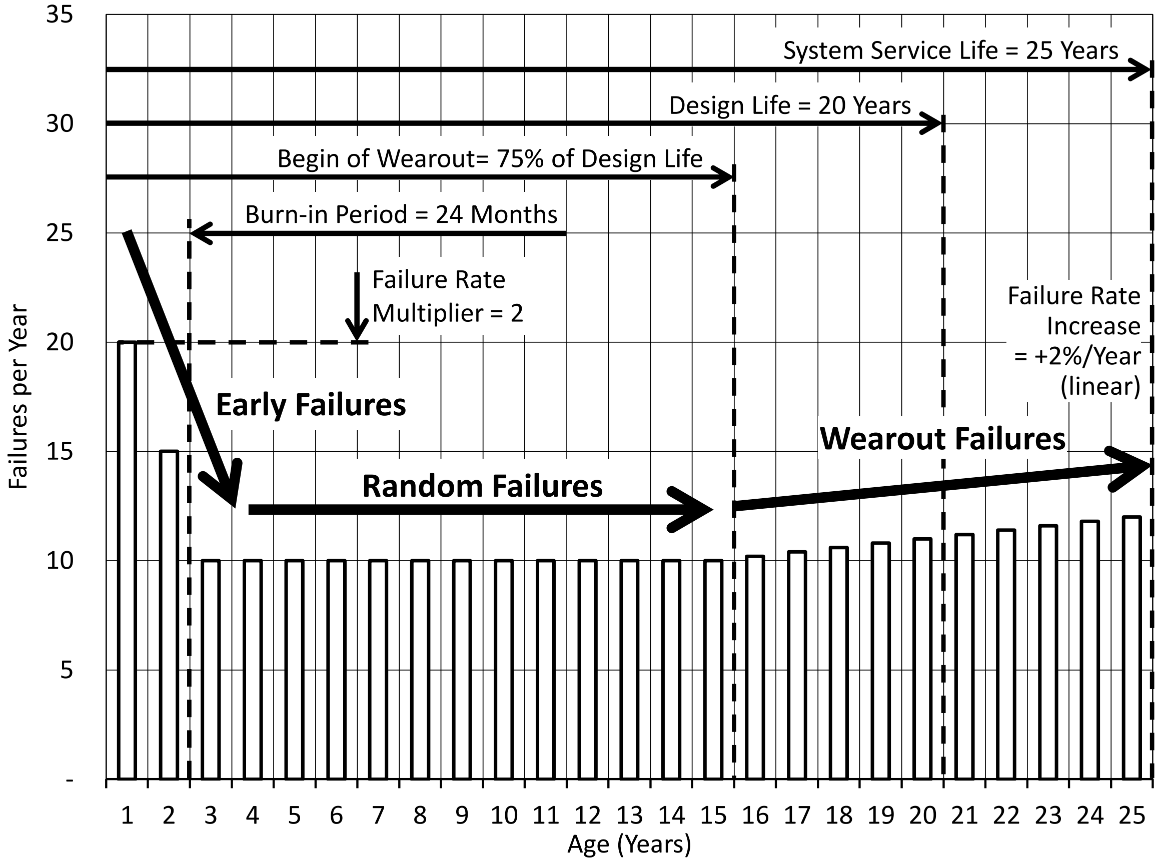


Figure 2: Default bathtub curve with annualised failure rate

The default settings of the bathtub curve are configured in such a way that it delivers an overall failure rate over a service life of 25 years that is 10% higher than the purely random failures. Please note: The time scale in Figure 2 does not refer to the absolute time in calendar years, but to the age of the system in question.

Early failures begin with a doubling of the random failures, i.e. the total failure rate at the beginning of operation is twice as high as the random failure rate. In practice, it is difficult to distinguish between early failures and random failures. What is visible from the outside is the initially increased and then decreasing total failure rate compared to the exponential failure distribution (Blanchard, 1978, p. 238). The time until the end of the early failures ("burn-in period") is given in months, the default setting is 24 months (two years).

The time between the end of early failures and the start of wearout failures is also called random failure plateau. Here, only random failures occur.

The onset of age-related wear and tear begins after 75 % of the planned design life. The latter is 20 years by default, meaning that the first signs of wear can be detected after 15 years. These manifest themselves in a failure rate that increases linearly by 2% per year.

All the above parameters can be changed by the user. If necessary, early failures and wearout failures can be switched on and off separately. By extending the useful life or adjusting the growth rate and characteristics (exponential instead of linear), the number of total failures can be changed from the default setting. The list of newly added input parameters for an LCCM-enabled cost model is shown in Table 1. The left-hand column shows the input parameters known from traditional cost models. The right-hand column contains the new inputs for LCCM.

Cost model, traditional	**Cost model, LCCM**
Total number of systems in use [number,	Number of systems deployed by age cohort [quantity, calendar year]
calendar year]	System service life [years]
Operating hours [h]	*Early failures:*
MTBF [h]	Burn-in period [months]
(models only random failures	Failure rate multiplier at the beginning
with a constant failure rate)	*Wearout failures:*
	System service life [years]
	Start of wearout [% of planned service life]
	Annual failure rate increase [%]
	Annual failure rate increase type [linear/exponential]

Table 1: Additional input parameters for modelling the bathtub curve

The most important new input parameter for LCCM is the number of systems deployed.

The key innovation for LCCM is the way the model handles fleet size by introducing age cohorts. Systems of different ages can show different failure rates within the same calendar year, unless all systems are simultaneously in the section of their bathtub curves with constant failures. This means that the number of systems in use and the total operating hours alone are not sufficient to calculate the total failures in a given calendar year.

To correctly determine the total failure rate of a fleet of systems over several years, a distinction must be made between the different system age cohorts. This is done by separating the entire fleet by the year of entry into service. All systems that are commissioned in the same (calendar) year are grouped into the same age cohort. If the total failure rate is calculated across several cohorts in a particular calendar year, each cohort has its own age on the time axis of the bathtub curve (in years, see Figure 2). The number of annual failures is a function of the age of the respective system. Very new and very old systems have higher failure rates. Age cohorts always have the same length (in years) but can comprise different numbers of systems independently of each other. The larger the cohort, the more total failures it experiences per calendar year.

To calculate the total failure rate, each age cohort is assigned its own bathtub curve, with its commissioning year as the starting year. The total failure rate over time is calculated by adding the bathtub curves of all age cohorts. The following simple example (see Figure 3) comprises five age cohorts of equal size from consecutive calendar years. The calculation of the total failure rate is shown graphically as a “stacking” of the total of five bathtub curves of all age cohorts.

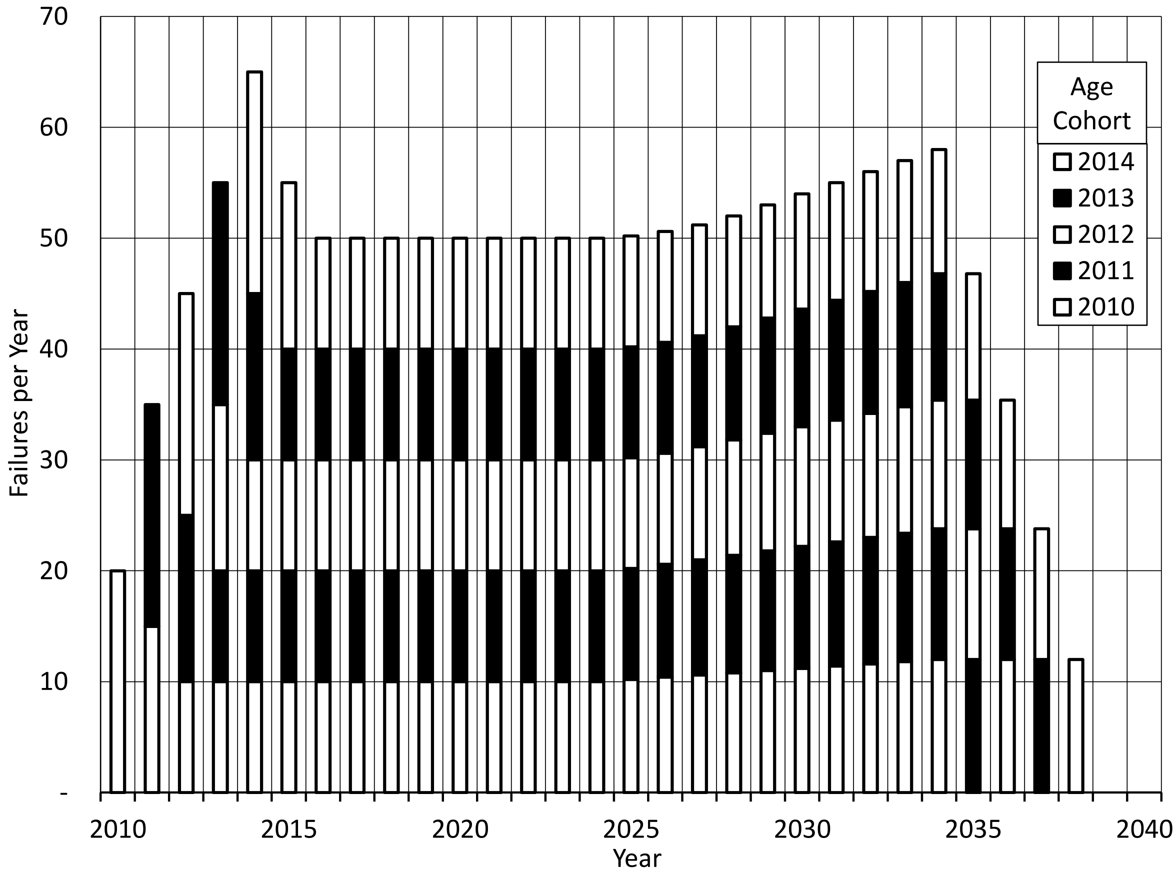


Figure 3: Calculation of the total annual failure rate by stacking bathtub curves of several age cohorts

By stacking the bathtub curves of several age cohorts, each separated by one year, the contour of the overall curve appears smoother than the simplified default curve in Figure 2. This becomes even more apparent when the stacked column chart from Figure 3 is converted into a line graph for the total failure rate, as shown in Figure 4.

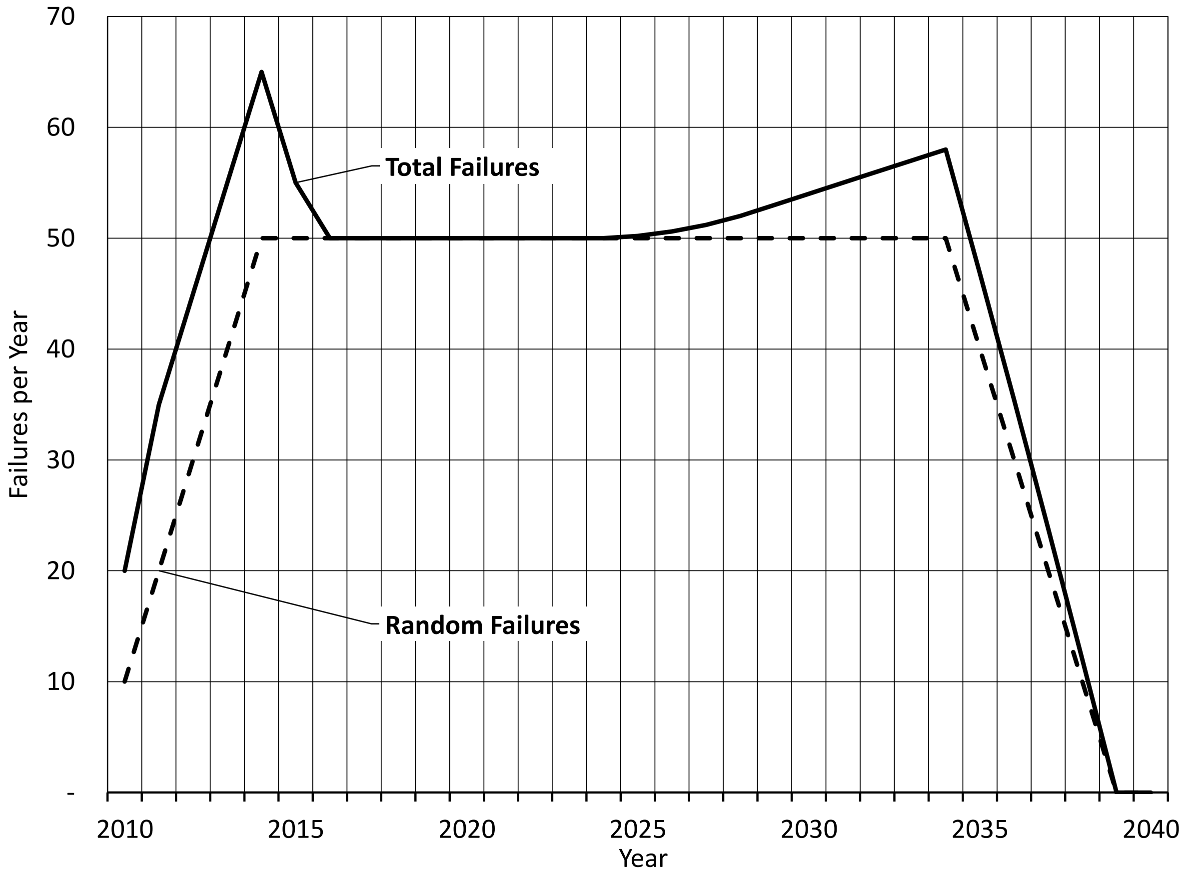


Figure 4: Total failure rate compared to random failures

The area below the solid line for total failures in Figure is 10 % larger than the area under the dashed line for random failures. This will apply as long as the default settings from the preconfigured bathtub curve as shown in Figure 2 are left unchanged. The value of 10 % additional failures over the entire service life should be seen as the lower limit. In the literature, much higher values are reported as typical, for example 70 % for military aircraft. If such an increase is ignored during planning, the budget for maintenance will later be too small by 70/(100 + 70) = 41 % (Crowley, 2004, p. 287).

1.3 Modelling increased failure rates: Discussion of benefits

In previous LCC analyses for the Bundeswehr, the costs of increased failure rates, particularly due to wearout, were identified as a problem. These are “known unknowns”. Traditional cost models only calculate with random failures and a constant failure rate based on the exponential failure law. By ignoring early and wearout failures, there is a significant risk of underestimating total maintenance costs.

Therefore, it is an LCCM-related improvement to plan for the bathtub curve from the beginning of a project. The probable increase in operation and support (O&S) costs and their distribution over time can already be estimated at this early stage. Existing cost models can be upgraded using a simple, preconfigured bathtub curve to model the effects of non-constant failure rates. The required inputs are simple to adjust and do away with Weibull distributions and similarly specialised methods. This means that cost estimators can look at the cost effects of early failures and wearout failures without having to be experts in reliability and failure probabilities.

As a result, the cost model provides more realistic estimates for maintenance costs. The example given (Crowley, 2004, p. 287) for the cost effects of the bathtub curve for military aircraft considers 70% additional costs (without inflation) over the entire useful life to be realistic.

2 The cost of condition-based maintenance: False alarms and "no fault found"

Many modern aerospace and defence systems employ condition-based maintenance with built-in tests (BIT) at system level. A BIT is run regularly, for example before every flight in the case of the Eurofighter multi-role combat aircraft. The purpose of a BIT at system level is to ensure that the system is in a safe and operational condition before going on a mission. The automatic self-tests help to detect potential failures at an early stage and ensure the safety of the system and its crew. The comparison between traditional (=corrective) and modern (=condition-based) maintenance is shown in Figure 5.

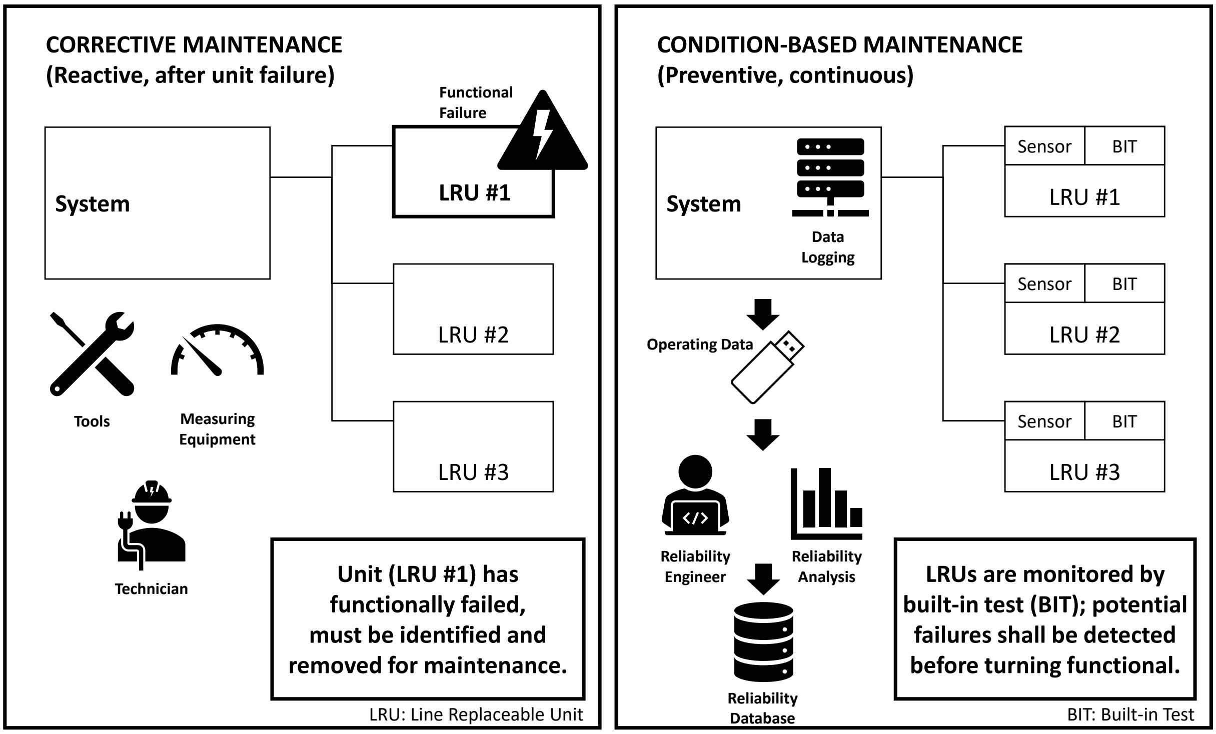


Figure 5: Comparison between corrective and condition-based maintenance

Overall, condition-based maintenance aims to improve the efficiency and reliability of systems by identifying potential failures at an early stage, before they can lead to functional failures. Corrective maintenance, on the other hand, aims to rectify failures only after they have occurred.

Corrective maintenance is reactive,
while condition-based maintenance is preventive.

2.1 Some history: The good old days of corrective maintenance

In traditional cost models for the operation and support (O&S) phase, which have been in development since the 1970s, the approach is straightforward (PRICE, 2001). These models focus exclusively on corrective maintenance, operating under the assumption that faulty units will manifest through dysfunctional behaviour. According to these models, a failure (F) is triggered in the cost calculations whenever the cumulative operating hours of a fleet reach an integer multiple of the Mean Time Between Failures (MTBF).

Maintenance and removal are considered only for units suspected of failing. Units that do not exhibit signs of failure are neither tested nor removed, adhering to the old maxim, "If it ain't broke, don't fix it."

The fact that there are more removals than failures is due to the incorrect assessment of units that are mistakenly identified as faulty during troubleshooting at system level, even though they are fully functional (False Failure (FF)). They are later declared false alarms after retest at the maintenance point ("shop") has confirmed that there is no fault. In traditional cost models the number of false alarms is calculated as a fixed percentage of failures. The sum of failures and false alarms gives the number of removals (R). A value of 20 % has long been established as the default setting for the so-called False Failure Fraction (FFF; PRICE, 2001, Chapter 10, p. 15).

Example FFF = 20 %; with 50 failures (F) → FF = FFF × F = 0.2 × 50 = 10, see Figure 6.

The challenge of condition-based maintenance can be better understood by comparing it with the old world of corrective maintenance. This is best done by analysing the workflow between the system and the maintenance point ("shop").

Figure 6 illustrates how the failure (F) of 50 units results in a total of 60 removals (R) and an equal number of shop visits (SV). It's important to note that the workflow to and from the maintenance point is depicted in a simplified manner in this diagram. In practical applications, cost models differentiate between various levels of maintenance (DIN, 2010, p. 35): Organisation (level 1), intermediate (level 2), depot (level 3).

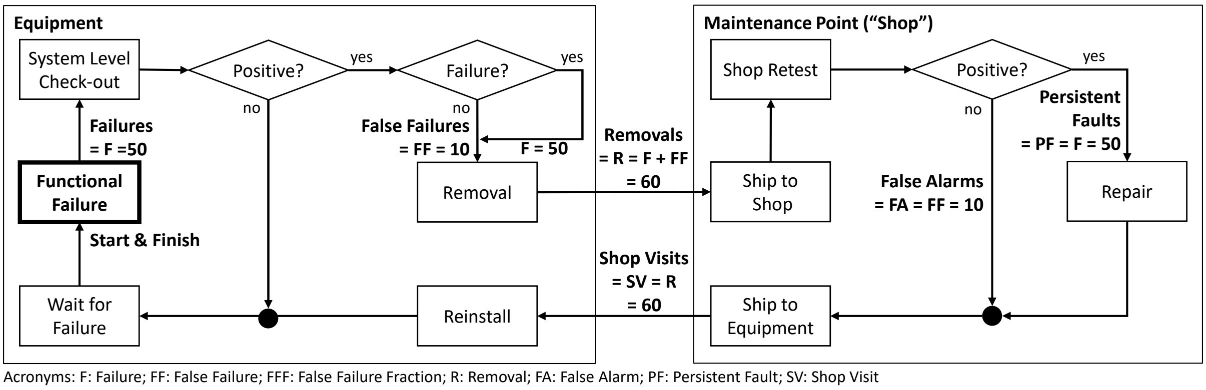


Figure 6: Workflow for corrective maintenance with 50 failures and 20% false failure fraction (FFF)

The cost drivers of corrective maintenance are shown in Figure 6:

- Number of failures (F): 50
- False failure fraction (FFF): 20 %
- Number of false failures (FF): 10, because: FF = FFF × F = 0.2 × 50 = 10
- Number of removals (R): 60, because: R = F + FF = 50 + 10 = 60
- All 60 removals end up at the maintenance point as a shop visit (SV)
- The 10 false failures (FF) retest negative at the shop,
  are flagged as false alarms (FA) and sent back unrepaired
- The 50 failed units (F) are confirmed as persistent faults (PF) and repaired

---

Result of corrective maintenance: 50 unplanned failures, a total of 60 removals, including 10 false alarms!

An important observation from the corrective maintenance workflow between equipment and the maintenance point in Figure 6 is that the False Failure Fraction (FFF) is a fixed percentage of (real) failures. So, in traditional cost models, the number of false failures (FF) is only driven by the number of failures.

This observation implies that the total number of systems in operation is not factored into the calculation of removals, rendering fleet size irrelevant to the overall cost of corrective maintenance.

Corrective maintenance only looks at units with functional failures,
only those will be removed and repaired!
Condition-based maintenance tests all units,
to find potential failures before they turn functional!

Condition-based maintenance is different. It is preventive and aims to identify potential failures in time, through testing, before they lead to functional failures and faults. This requires many repeated tests of **all systems in use**, which has an impact on operation and support (O&S) costs. This is a considerable challenge.

2.2 Condition-based maintenance means: Testing, testing, testing

The effectiveness and costs of condition-based maintenance are critically dependent on the accuracy of fault detection at the system level using built-in tests (BIT). Contrary to corrective maintenance, BIT should proactively identify potential failures before a unit fails and turns dysfunctional. The BIT must consistently analyse the operating parameters of units to identify any trends that might indicate a potential failure. These trends need to be accurately recognized. In practice, the effectiveness of this process varies, and test results can occasionally be wrong. Table 2 displays the potential outcomes of a BIT at the system level.

Table 2

	Genuine positive (unit is faulty)	**Genuine negative (unit is good)**
Test at system level (BIT): **Positive**	• True Positive (TP) • Potential failure detected • Unit no longer good • Unit removed, sent to shop	• False Positive (FP) • Potential failure detected • But: Unit still good • Unit removed, sent to shop
Test at system level (BIT): **Negative**	• False Negative (FN) • No potential failure detected • But: If no fault is found in later test, failure and unplanned maintenance will occur!	• True Negative (TN) • No potential failure detected • Unit still good • Do nothing!

Table 2: Possible results of a built-in test (BIT) at system level (PRICE, 2022, p. 91–95)

Those tasked with responding to BIT results can only observe the positive and negative outcomes of the tests, but not their accuracy. The validity of these results – true or false – only becomes evident during subsequent retesting at the maintenance point. Accurate BIT results lead to no wastage, ensuring efficiency. However, incorrect results can cause unnecessary additional work and significantly increase maintenance costs.

Removing good units for maintenance (false positives) is wasteful.
Depending on the maintenance concept, good units may be discarded prematurely.
At least they waste time and resources until they are confirmed as good and reinstalled.

Leaving faulty units on the system (false negatives) can lead to additional costs
due to unplanned failures, possible collateral damage and longer downtimes.

Therefore, any estimation of operation and support (O&S) costs must account for the prevalence of false test results. Given realistic conditions, it's reasonable to expect that BIT results will occasionally be incorrect. What implications does this have?

2.3 How accurate are tests at system level?

The quality of the test results depends on the probability of detection (P_{det}), also known as sensitivity. It is the most important parameter of tests at system level. In practical applications, it is an observable property of test tools and can be validated through a series of tests. P_{det} is the probability with which a test run at system level (BIT) positively detects a potential failure because the faulty unit to be tested does not pass the test. "One test run" means that each affected unit is only tested once (PRICE, 2022, p. 76). The sensitivity (P_{det}) of a test indicates the probability with which a faulty unit is correctly recognised (Eq. 2).

Eq. 2 $$\boldsymbol{Probability\ of\ Detection} = \boldsymbol{P_{det}} = \frac{Faulty\ units\ detected}{Faulty\ units\ tested} = Sensitivity$$

A sensitivity of a test to a fault of 80% means that if a sufficiently large number of tests were carried out and irrespective of the test preconditions, 80% of the units affected by the fault would be recognised and 20% of the units affected by the fault would not be recognised.

Therefore, 20% of the faulty units tested (and not 20% of all (!) tested units) would be false negatives. False negative units that are not detected as faulty in time will suffer unplanned failure during operation and require immediate maintenance.

Repeating tests at intervals within the observation period will increase the total fault detection (TFD), which is calculated from Eq. 3:

Eq. 3 $$\boldsymbol{Total\ Fault\ Detection} = \boldsymbol{TFD} = \boldsymbol{1-(1-P_{det})^{n}} = \boldsymbol{1-(1-P_{det})^{\frac{Observation\ Period}{Inspection\ Interval}}}$$

The exponent n denotes the number of tests at system level (BIT) within the observation period. Example: n = 2 means that a BIT is carried out twice within the observation period. Please note that the observation period in this simple example shall be significantly shorter than the MTBF, so that no new failures occur during observation. The numbers in Table 3 show how total fault detection is approaching 100% as n increases.

Number of tests *n*	**1**	**2**	**3**	**4**	**5**	**>5**
$TFD = 1 - (1 - P_{det})^n$	80%	96%	99.2%	99.84%	99.96%	≈100%

Table 3: Total fault detection (TFD) as a function of the number of tests during the observation period

Even with a low sensitivity of P_{det} = 80 %, total fault detection of a BIT can simply be increased to 96% by carrying out a second test pass. The process diagram in Figure 7 shows step by step how the numbers come about. From a total of 400 units, 50 of which are faulty, 48 faults are correctly identified over two test passes: 40 of which in the first pass (80% of 50) and eight in the second pass (80% of 10). All faulty units are removed for repair and replaced by "good" units without fault.

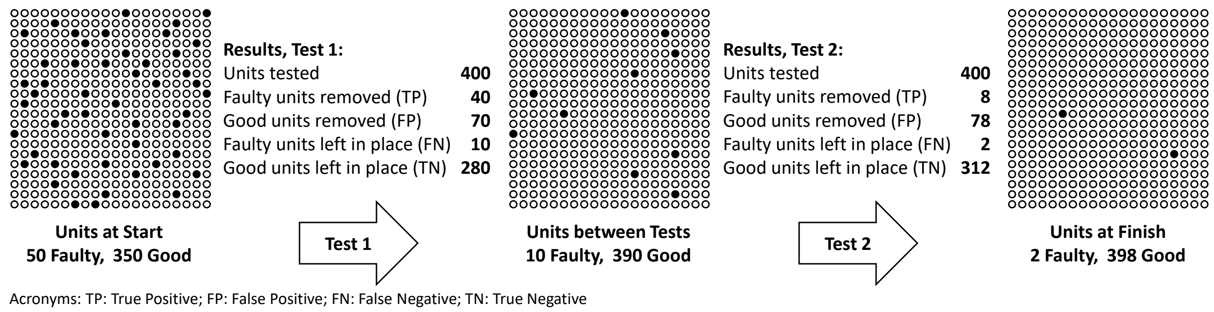


Figure 7: Total fault detection of a test with sensitivity 80% and two passes in the observation period

Below are the results of a built-in test (BIT) at system level with a Sensitivity P_{det} of 80%. After the second pass through the test, total fault detection equals 1 - (1 - 0.8)² = 1 - 0.04 = 96%.

- Units tested after two test passes: 400 + 400 = 800, of which are:
  - True positive (TP): 40 + 8 = 48
  - False positive (FP): 70 + 78 =148
  - True negative (TN): 280 + 312 =592
  - False negative (FN): 10 + 2 = 12 (after second pass, 2 FN remain in the system)

Either positive test results in the BIT or unplanned failure initiate the removal of units and their shipping to the maintenance point:

- Total units removed: 48 + 148 + 2 = 198, broken down into:
  - Genuine positives tested positive (TP): 48 (removed after test)
  - Genuine negatives tested positive (FP): 148
  - Genuine positives tested negative (FN): 2 (removed after unplanned failure)

The overall fault detection rate of 96% seems very high, as 48 out of 50 units are removed for preventive maintenance before they can cause any problems during operation. At second glance, however, the numbers in Figure 7 show the downside. A total of 148 good units (false positives) removed after failing the BIT is very high, at more than three times (!) the 48 genuine faults (true positives). This poses a big problem.

2.4 False positives: How false alarms occur

The good units removed in Figure 7 are false positives. Accordingly, the false positive rate (FPR) indicates the proportion of units incorrectly classified as positive while they are genuinely negative. In the example, a unit that is fully functional would be wrongly diagnosed as faulty. As shown in Eq. 4, FPR indicates the probability of removed genuine negatives causing a false alarm at a later shop retest.

Eq. 4 $False - Positive - Rate = FPR = (1 - P_{det})$; if P_{det} = 80%, then FPR = 20%.

Since the false positives are in fact fully functional, the false positive rate of 20% refers to the genuinely “good units”, see Eq. 5. As true positives (“faulty units”) are removed after each test, the number of genuine positives is reduced, and these are replaced by “good units”. Therefore, the number of genuine negatives (“good units”) increases from 350 to 390 after the first test (see Figure 7).

Eq. 5 $False\ Positives = FP = Genuine\ Negatives \times FPR = Good\ Units \times FPR$

According to Eq. 5, the number of false positives shall increase, from 70 in Test 1 to 78 in Test 2 (see Figure 7). Since all 70 + 78 = 148 false positives will be removed, this and their subsequent shipping to the maintenance point causes unnecessary effort and cost.

2.5 What happens during retest at the maintenance point?

After removal at system level, units are sent to the shop for maintenance. These comprise all units tested positive in the BIT (true positive and false positive) plus defective units previously tested as negative in the BIT (false negatives) that failed during operation before they could be tested again. These units are called Units Under Test (UUT; see Ungar, 2015, p. 1 and Eq. 6).

Eq. 6 $$Units\ Under\ Test = UUT = True\ Positives + False\ Positives + False\ Negatives$$

After arriving at the maintenance point ("shop"), a retest is carried out. Possible test outcomes are shown in Table 4. They determine what will happen next with the units.

	Genuine positive (unit is faulty)	**Genuine negative (unit is good)**
Shop Retest: **Positive**	**Persistent Fault (PF)** Unit fails retest, is confirmed as faulty Unit is repaired, sent back to system Unit is reinstalled, passes check-in at system	(Not applicable)
Shop Retest: **Negative**	**Intermittent Failure (IF)** Unit passes retest as good Unit is sent back to system unrepaired Unit is reinstalled, fails check-in at system Unit is removed, sent to shop again …	**False Alarm (FA)** Unit passes retest as good Unit is sent back to system unrepaired Unit is reinstalled, passes check-in at system

Table 4: Possible outcomes of a shop retest (Ungar, 2017, p. 9)

In a perfect world, all units removed at system level would test positive during retest at the maintenance point, be confirmed as faulty and repaired immediately. Accordingly, there is only one category of positive results.

Persistent Fault (PF): For units that have previously tested positive, the positive retest confirms the presence of a fault and retrospectively reveals the result of the BIT as a "True Positive" (TP). In the case of units that tested negative until recently and subsequently failed, these are retrospectively revealed as "false negatives" (FN). It is assumed that their functional failure will always lead to a "positive" result in the shop retest. Therefore, both, true positives (TP) and false negatives (FN, *after* functional failure!) make up the group of persistent faults (PF), see Eq. 7. These occur continuously or repeatedly and do not disappear spontaneously. All units identified as persistent faults (PF) are immediately forwarded for repair.

Eq. 7 $$Persistent\ Faults = PF = True\ Positives + Unplanned\ Failures\ (after\ BIT\ with\ FN)$$

However, the real world of condition-based maintenance is not perfect. In practical applications, there are many units that pass the shop retest as good ("negative"). These are divided into two categories:

Intermittent Failure (IF): Occasionally there are faulty units (true positives) that are not immediately recognised as positive during their first shop retest. These intermittent failures (IF) are genuine failures that only reveal themselves under certain conditions. Defective units that failed the BIT, i.e. are true positives (TP), pass the shop retest as good and are labelled as IF! Such units are prematurely categorised as not requiring repair and sent back to the system. There they are reinstalled, subjected to another BIT and ... fail the test. The units are therefore removed again and sent back to the shop. This is repeated until the fault is finally confirmed by a shop retest. Only then is the unit declared a persistent fault (PF). It is assumed that every intermittent failure (IF) will sooner or later become a PF. For intermittent failures (IF), an intermittent failure rate (IFR) must be specified (see Eq. 8). It describes the proportion of IFs in the total number of passes required by all true positives (TP) until they are ultimately confirmed as persistent failures (PF). A typical value is IFR = 20 %.

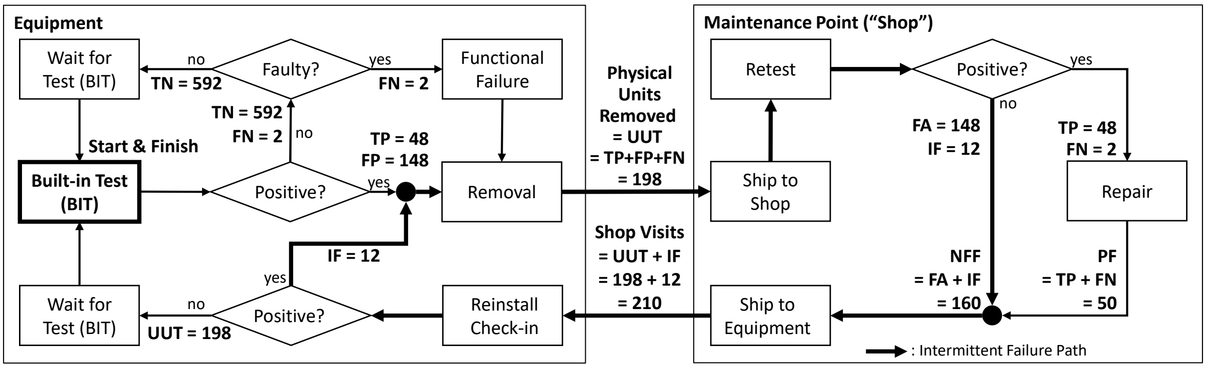
Eq. 8 $$Intermittent\ Failures = IF = True\ Positives \times \frac{IFR}{(1-IFR)}$$

False Alarm (FA): False alarm refers to units that have been removed for maintenance after a positive test at system level, although they are good (False Positive (FP)). These units pass the shop retest as good, since no fault can be detected. According to Eq. 9, this means that all (!) false positives (FP) coming from BIT will end up as false alarms (FA) in the shop retest (see also Eq. 4 and Eq. 5).

Eq. 9 $$False\ Alarms = FA = FP = False\ Positives$$

The workflow between the various stations in the maintenance process depends on the outcome of the two tests, system-level BIT and shop retest. Incorrect test results will lead to pointless movements of units.

The waste of resources in condition-based maintenance depends directly on how many units pass the shop retest with a negative result. The latter is named “No Fault Found (NFF)”. The quantitative effects can be illustrated using the previous example of corrective maintenance (see Figure 6) after it is switched to condition-based maintenance. The modified workflow is shown in Figure 8.



TP: True Positive; FN: False Negative; FP: False Positive; TN: True Negative; UUT: Units Under Test; PF: Persistent Fault; IF: Intermittent Failure; FA: False Alarm; NFF: No Fault Found; SV: Shop Visit. Note: To arrive at the cumulative total of 800 tests (BIT) after the second pass, FN = 10 from the first pass must be added!

Figure 8: Workflow for condition-based maintenance with 50 faulty units and two test passes at system level

The results at system level in Figure 8 are the same as in the example from Figure 7. The total number of units removed after a positive BIT or functional failure at system level is UUT = TP + FP + FN = 48 + 148 + 2 = 198.

All units removed from the system and sent to the maintenance point (“shop”) become units under test (UUT; see Eq. 6). After arriving at the shop, they undergo a retest to reproduce the result of the previous BIT. Accounting for intermittent failures (IF), it is assumed there are true positives (TP) that only fail under certain conditions. For them, an intermittent failure rate (IFR) of 20% is set. According to Eq. 8, the total number of intermittent failures is IF = TP × IFR / (1 - IFR) = 48 × 0.2 / (1 - 0.2) = 48 × 0.25 = 12.

Intermittent Failures (IF) are the reason why the number of shop visits is higher than the number of (physical) units under test (UUT)!

Eq. 10 $$Shop\ Visits = SV = Units\ Under\ Test + Intermittent\ Failures$$

According to Eq. 10, the total number of shop visits is SV = UUT + IF = 198 + 12 = 210.

So, following the workflow in Figure 8, it takes the removal of 198 (physical) units and 210 (!) shop visits to correctly identify and repair 50 faulty units.

The example for the maintenance point in Figure 8 shows that the total number of shop visits during which no fault is found is dramatically high. These are called No Fault Found (NFF).

Eq. 11 $$No\ Fault\ Found = NFF = False\ Alarms + Intermittent\ Failures$$

According to Eq. 11 the number of “No Fault Found (NFF)” is NFF = FA + IF = 148 + 12 = 160!

The figures show that the cause of most NFFs is false alarms (FA). The twelve intermittent failures (IF) are hardly significant compared to the 148 false alarms (FA).

Result of condition-based maintenance: 50 failures, only 2 of which were unplanned, 198 removals, 210 shop visits, 160 no faults found, 148 of which were false alarms!

The example for corrective maintenance described in ▶ Section 2.1 has the same number of failures (F = PF = 50) as the condition-based maintenance shown in Figure 8. However, the comparison in Table 5 shows that condition-based maintenance performs significantly worse than corrective maintenance in all metrics for logistical effort.

- The number of removals at system level is more than three times higher (UUT = 198 instead of 60).
- This mainly comes from the nearly 15-fold increase in the number of false alarms (FA = 148 instead of 10).
- Intermittent failures (IF) only occur in condition-based maintenance, but their overall impact is negligible (IF = 12 instead of 0)
- The number of No Fault Found (NFF = IF + FA) is 16 times higher (NFF = 160 instead of 10).
- The number of shop visits SV (SV = UUT + IF) is increased by a factor of 3.5 (SV = 210 instead of 60).
- The fraction of shop visits with “no fault found” (NFF% = NFF/SV) is increased by a factor of more than 4.5 (NFF% = 76.19 % instead of 16.67 %).

The metrics in Table 5 show that in over three quarters (160/210 = 76.19%) of shop visits (SV) the result is “NFF” (no fault found). This value may seem exaggerated, but it is not.

	F, PF	UUT	IF	FA	NFF	SV	NFF%
Corrective	50	60	-	10	10	60	**16.67%**
Condition-based	50	198	12	148	160	210	**76.19%**

Table 5: Comparison of false alarms (FA) and NFF (no fault found) between corrective and condition-based maintenance

Real-world examples show NFF values of up to 90% and more for condition-based maintenance, with the majority being due to false alarms (Ungar, 2017, p. 13). The costs of this issue are considerable. In 2012, the US Department of Defense announced that it was spending two billion dollars a year (!) on the removal of line replaceable units (LRUs) that were tested with NFF during the incoming inspection at the depot (Werner, 2015, p. 30). The costs arise from the removal, packaging, shipping and tracking of the many units tested as NFF. Since the majority of NFF are due to false alarms (FA) and there is no need to repair them, any attempt to repair them would be pointless. False alarms are simply sent from the system to the maintenance point and are returned unrepaired. Experts say that they will never be able to eliminate all NFF; however, they are confident the extent of the problem can be reduced (Werner, 2015, p. 31).

2.6 Why there are so many false positives

The workflow example in ▶ Sect. 2.1 showed how the number of false alarms in corrective maintenance is calculated as a fixed percentage of functional failures.

*In corrective maintenance, the false failure fraction (FFF) refers to the number of **faulty** units.*

The False Failure Fraction (FFF) determines how many of the units in use (as a percentage of faulty units) are removed for repair even though they are good units. Here, the false alarms depend solely on how many failures occur. An FFF of 20%, as used in Figure 6, leads to just 10 false alarms for 50 faulty units (see Table 5).

In contrast, calculation of the false-positive rate in ▶ Sect. 2.4 and the comparative example in ▶ Sect. 2.5 show that false alarms occur much more frequently with condition-based maintenance. They are a result of false positives (FP) in the test at system level (BIT).

The false positive rate (FPR) determines how many of the genuinely negative ("good") units tested at any point in time are incorrectly considered as potential failures and removed. Since all false positives (FP) later become false alarms (FA), their number depends on how many genuinely negative ("good") units go through the BIT per pass and how many passes are done. The FPR of 20 % shown (FPR = (1 - P_{det}) = 1 - 80 %) is the same figure as the FFF for corrective maintenance in Figure 6. Only here the 20% refers to the "good units", i.e. the genuine negatives from the two test passes in Figure 7. In the first pass these equal 400 - 50 = 350 units, and in the second pass 400 - 10 = 390 units. The total is 740, 20% of which equals the 148 false alarms observed after two test passes (see Table 5).

In condition-based maintenance, the false positive rate (FPR)
refers to the number of ***good*** *units!*

This highlights the disadvantage of frequent testing at system level, namely that faulty units are recognised and removed with each additional test, so that the number of true negatives increases with every pass. The false positive rate (see Eq. 4), on the other hand, remains constant. Therefore, the lower the sensitivity of the test at system level (P_{det}) and the shorter the inspection interval, the greater the number of false positives removed, see Eq. 12.

Eq. 12

$$False\ Positives = \left[\frac{Total\ operating\ hours}{Inspection\ interval} - \frac{1}{P_{det}} \times \left(1 - (1 - P_{det})^{\frac{MTBF}{Inspection\ interval}}\right) \times \frac{Total\ operating\ hours}{MTBF}\right] \times (1 - P_{det})$$

As more and more electronic units are monitored by a BIT (Built-in Test), there is a high probability that a fault detected at system level will initiate a removal, which will then lead to a false alarm in the shop (*Caution: The calculation scheme in Eq. 12 only works if inspection interval ≤ probability of fault detection × MTBF!).*

This situation is worrying because there are systems in which the ratio between MTBF and inspection interval becomes very large. Fighter aircraft, for example, have subsystems with a typical MTBF of 1000 hours and an average flight duration of 75 minutes or 1¼ hours. If testing is carried out between each flight, this means 800 tests within the MTBF (PRICE, 2022, p. 85).

The trend curves in Figure 9 illustrate where these extreme examples may lead. They show the ratio between removed and genuinely faulty units as a function of sensitivity (P_{det}) and MTBF/inspection interval ratio. The probability that a unit tested positive at system level and removed is genuinely faulty is becoming virtually zero as the inspection intervals become shorter. In contrast, the ratio of removed units to faulty units increases exponentially.

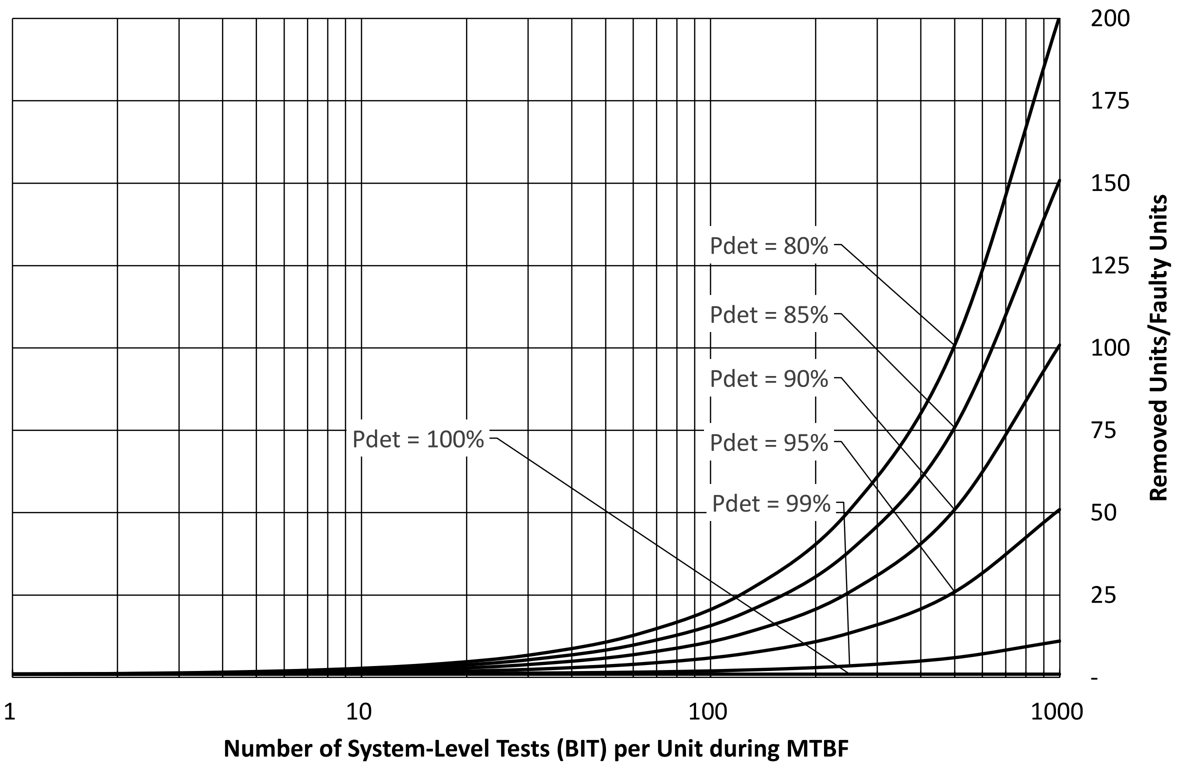


Figure 9: Ratio of removed to faulty units as a function of test frequency and sensitivity

For the sake of affordability alone, such extreme conditions will not be permitted in practice. It is more likely that maintenance will ensure that the sensitivity of tests at system level is close to 100%. Even an increase to 98% would be helpful. This value is a typical target, but one that is difficult to achieve for many systems (Ungar, 2015, p. 8).

Apart from the feasibility of better tests, it remains necessary to analyse the high rate of false alarms and NFFs observed in many places today, while keeping an eye on their cost impact.

2.7 False alarms and NFF: Implementation in the LCCM cost model

In a cost model optimized for Life Cycle Cost Management (LCCM), incorporating additional input parameters is essential for effectively modelling condition-based maintenance. The probability of detection (P_{det}) by the built-in test at the system level (BIT) emerges as a critical cost driver, largely influenced by technological MTBF and operational constraints like flight duration. Given its significance, sensitivity must be treated as a variable within the LCCM cost model. This adjustment enables a detailed analysis of learning effects in error detection and their subsequent impact on costs. The model requires the following inputs:

- An initial sensitivity value (P_{det}) at the start of the operation and support (O&S) phase (in %).
- A milestone defined by the cumulative operating time of the entire fleet (in hours).
- A target sensitivity value (P_{det}) to be achieved by the time of the milestone (in %).

For instance, at the onset of the O&S phase, the system-level BIT might have an initial detection probability (P_{det}) of 80%. After the fleet accumulates 100,000 operational hours, the target for P_{det} is set at 95%. This structured approach allows for a precise evaluation of improvements in fault detection capabilities over time.

These and the other new input parameters are listed in Table 6.

Cost models, traditional	**Cost model, LCCM**
Operating hours [h]	Probability of detection P_{det} at start [%]
MTBF (Mean Time Between Failures) [h]	Probability of detection P_{det} at goal [%]
MTTR (Mean Time to Repair)	Goal, in cumulative time of operation [hours]
(models only random failures	Inspection interval [hours]
with a constant failure rate)	Intermittent failure rate [%]
	Duration BIT at system level [hours]
	Continuous incoming inspection for maintenance [hours]
	Duration of function test after reinstallation [hours]
	Unplanned failure penalty in labour hours
	Unplanned failure penalty in material cost [€]

Table 6: Additional input parameters for modelling condition-based maintenance

Condition-based maintenance must aim to reach the highest possible sensitivity in built-in tests (BIT) at system level. This is the only way to operate a system in which unplanned failures can be prevented while keeping the number of NFFs and false alarms from becoming unbearably high.

2.8 Modelling condition-based maintenance: Discussion of benefits

Condition-based maintenance faces a frustrating and expensive issue: many units tested and flagged as faulty at the system level are sent for repair, only to be found fault-free upon retesting. These units are then labeled as “No Fault Found (NFF)” and returned without repair. Remarkably, it is not uncommon for over 70% of units sent in to be categorized as NFF (Werner, 2015, p. 28ff). The majority of these NFF cases are false alarms (FA), involving fully functional units that do not require any repair. The incidence of false alarms increases with lower sensitivity (P_{det}) of the built-in tests at the system level (BIT) and shorter testing intervals.

Historically, traditional cost models did not account for condition-based maintenance, focusing only on units removed after failure. The false alarm rate in corrective maintenance is typically assumed to be relatively low, around 20%.

To address these challenges, existing cost models have been enhanced with new functionalities to estimate the incidence of “no fault found” (NFF) and the proportion of false alarms (FA). New input parameters have been defined to facilitate these calculations, enabling detailed cost analysis and practical examples to be worked through. This allows planners to comprehensively analyze how NFF and false alarms affect maintenance costs.

Ultimately, the revised cost model provides more accurate estimates of the additional costs associated with preventive testing and handling of non-defective units. A typical finding shows that compared to corrective maintenance, these activities can result in a 300% increase in costs—accounting for removal, transport, testing, and reinstallation of units categorized as NFF, without factoring in inflation.

3 The cost of obsolete spare parts: Product and technology obsolescence

In the Bundeswehr environment, cost models are used to estimate the life cycle costs (LCC) of systems that typically have a long service life of 20, 30, 40 years or even longer. Over these long periods, the technology underlying the systems in use inevitably changes. This has an impact on the cost of spare parts. The Bundeswehr therefore demands that proposed solutions already take obsolescence into account and include statements on the subsequent financial requirements of obsolescence management (Bundeswehr, 2016, p. 33).

3.1 History: Technological change and cost

Since around 1975, commercially available cost estimation tools have used built-in models to calculate the impact of technology improvements on manufacturing costs (PRICE, 1981). They are based in part on the scientific model of innovation diffusion developed by Everett Rogers in 1962 (Rogers, 2003, p. 45). This assumes that an innovation in the form of a new technology, when it is first introduced, is the result of the work of one or more innovators. To be successful, the innovation must penetrate the market by being accepted by as many users as possible (Rogers, 2003, p. 161ff). New technologies are immature and therefore more expensive to develop and produce than mature ones. In the initial phase of a new technology, it is only used by the innovators themselves and early adopters. As the technology improves, its costs drop so that more and more potential buyers can afford products based on it. This in turn increases market demand, which encourages more manufacturers to join the supplier base. After reaching “critical mass”, the point at which further growth becomes self-sustaining, a technology continues its market penetration until the entire potential market is saturated, see cumulative market penetration (S-curve) in Figure 10. Its final percentage can remain below 100%, as not every buyer in a market is necessarily willing to accept an innovation (Rogers, 2003, p. 375)

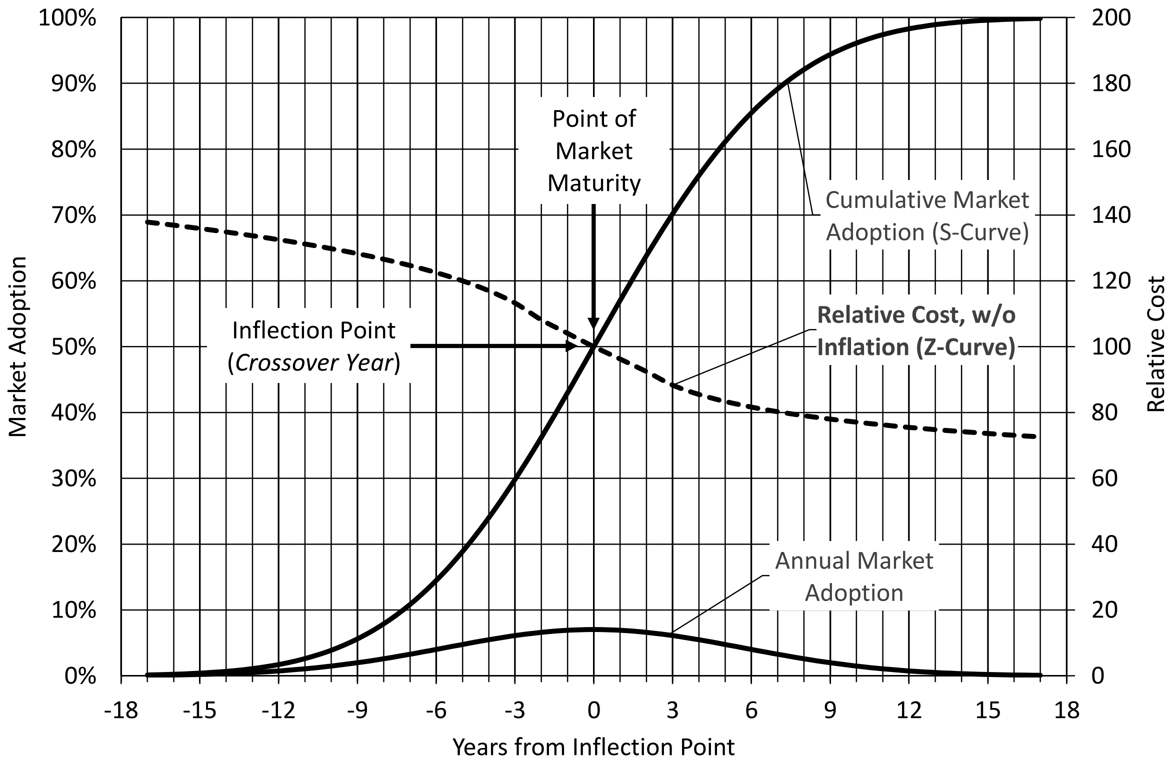


Figure 10: Relationship between market penetration rate, cumulative penetration and relative costs (based on: Rogers 2003, p. 478)

Rogers' original model from 1962 saw falling prices for innovations merely as a driving force for their diffusion through social systems (Rogers, 2003, p. 66ff). As a sociologist, Everett Rogers did not investigate the technical causes of cost reductions any further. The curve for relative cost as shown in Figure 10 therefore did not originate from Rogers himself. Instead, it was based on a later model that was first developed in the 1970s (PRICE, 1981, p. 1).

Gordon Moore first published what is now known as Moore's Law in 1965. It states that the number of transistors in an integrated circuit (IC) doubles approximately every two years (Moore, 1965, p. 114ff). Moore's Law is an observation and projection of a historical trend that is continuing over 50 years later. It is not a physical law, but an empirical relationship that can be derived from the continuous increase in experience in semiconductor production. The influence of this law was considerable and did not remain without consequences for the development of parametric cost estimation. When the first commercial cost estimation tool (PRICE H) came onto the market in 1975, it already had a sub-model that calculated the effects of technology improvement on the manufacturing costs of products. It was improved in 1981 and has remained virtually unchanged since then (PRICE, 1981, p. 4).

The effects of the interplay between increasing market penetration and technology improvement on product costs are shown in Figure 10. This shows how the cumulative percentage of market penetration (S-curve) for a product is linked to the relative cost reduction through technology improvement (Z-curve). The halfway point of market penetration, the year "0" in the graph, at which the turning point of the diffusion curve lies, is also defined as the point of full market maturity of the product. While cumulative market penetration follows the characteristic S-curve upwards, adjustment of relative product costs is directed downwards. It resembles an inverted S-curve, which is why it is also referred to as the "Z-curve". The striking symmetry of the three curves is intentional and may differ from the descriptions of other, more recent models (PRICE, 2003, Chapter 11, p. 3; Rogers, 2003, p. 383; Shermon, 2009, p. 119; PRICE, 2017, p. 331).

Over time, the three curves go through the following stages:

- The annual market penetration goes from zero to the maximum at the halfway point and then back to zero (symmetrical)
- The cumulative market penetration goes from zero to complete market saturation (S-curve; point-symmetrical)
- Relative product costs benefit from technology improvements, fall and follow an inverted S-curve (Z-curve; point-symmetrical)

In this model, the principle of product cost generation assumes that product cost only arises whenever the state on the Z-curve is frozen in a particular year. Only then can a new product be created using the latest technology. Note: The model described here was developed around 1980 and bears no relation to John Mankins' Technology Readiness Levels (TRLs) model, which was developed around ten years later (Shermon, 2009, p. 111ff).

3.2 Obsolescence of spare parts: The challenge

When cost adjustment curves (Z-curves) were normalised in 1981, 36 years of technology cycle, as shown in Fig. 1.9, still stretched into the distant future (PRICE, 1981, p. 1ff; PRICE, 2003, Chapter 14, p. 8ff). Ignoring inflation, the cost of a technology, once brought to market, could only go into one direction: downwards (as seen in Figure 10). Only around 2006, it was realised that an important aspect had been missing. It had to do with the question what happens when a technology becomes outdated and is replaced by a newer, higher technology? The answer is: obsolescence.

There are two types of obsolescence:
Product obsolescence and technology obsolescence.

Product obsolescence follows relatively short product cycles. Obsolescence is triggered by the supplier's decision to discontinue an existing product (End of Life; see DIN, 2008, p. 7). While reasons for discontinuation are not relevant, buyers who used the original product to operate and support a system must react (Bartels, 2012, p. 2ff; Galar, 2017, p. 351ff). Usually, product cycles are short and controlled by the supplier, five to seven years is a common value.

Technology obsolescence follows relatively long technology cycles. Products based on a more advanced technology than the original product enter the market. Demand for the original product line with older technology decreases. It therefore becomes increasingly unattractive for manufacturers and suppliers to maintain production capacity for the original product family. For a limited time, remaining stock can be utilised, or work outsourced to manufacture and support the product with older technology. But in the long term, fewer and fewer suppliers of older parts will survive. As the supplier base shrinks, availability decreases and experts with product-specific knowledge become fewer. Rising prices have the opposite effect of technology improvement. As soon as there are no more suppliers for older parts, technology obsolescence finally sets in. Technology cycles are long, typically 40 years (PRICE, 2022, p. 108–109), like in the timeline from Figure 10.

Before LCCM, a cost model would "freeze" the technology in a product no later than production start. The year of "technology freeze" thus becomes the technology year of the product and would remain the same for all production units, initial spares and subsequent replenishment spares! By default, spare parts cost is based on the average unit production cost of the original production quantity (see *Reference Quantity*; PRICE, 2001, Chapter 6, p. 25). This applies to the entire service life, even if it spans many decades. The graphic in Figure 11 uses a comparison with some real-world observations to show how inaccurate traditional cost models before LCCM represented reality.

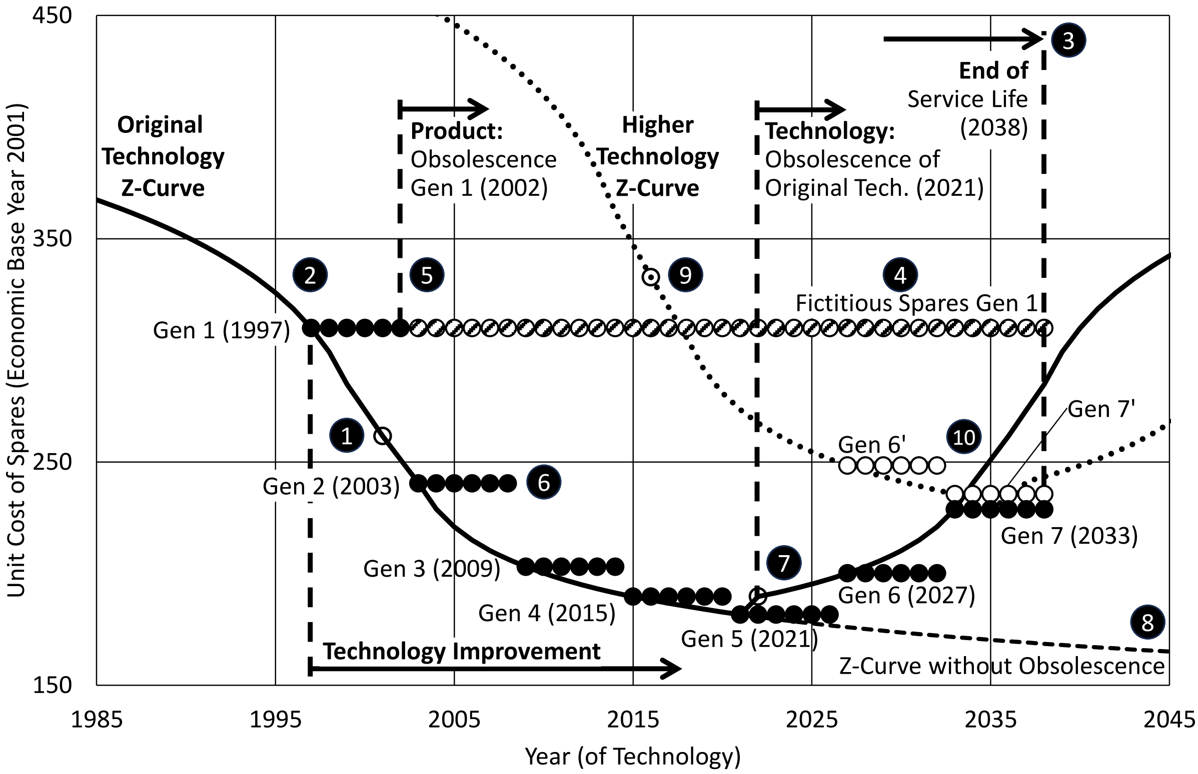


Figure 11: Correlation between product generation, technology improvement and obsolescence (based on PRICE 2022, p. 109–111)

Original old-school scenario, before LCCM:

- Unit cost of a product is based on a technology curve with an inflection point in 2001 (see ❶ in Figure 11)
- Production starts in 1997 as product generation 1 (Gen 1): technology is frozen, unit cost is plotted as black dots, there is no cost difference between production items and spares (see ❷)
- Service life ends in 2038, equalling a life cycle of 42 years (see ❸)
- Spare parts are technologically identical to the production units from 1997 and will be available until the end of operation in 2038 (“Fictitious spare parts Gen 1”, see ❹)
- **In this scenario, neither product nor technology obsolescence are considered!**

Realistic, practical scenario:

- The manufacturer regularly updates the product, one product cycle lasts 6 years; the first product generation is discontinued in 2002, marking product obsolescence of Gen 1 (see ❺)
- The second generation (Gen 2) takes over as successor of Gen 1 in 2003 (see ❻)
- Gen 2 up to Gen 5 benefit from decreasing unit costs due to technology improvements on the technology curve’s downward slope; this continues until 2022, when technology obsolescence sets in; with technology obsolescence progressing, unit costs of Gen 6 and Gen 7 will rise (see ❼)
- For comparison, see the “Z-Curve” (❽) without technology obsolescence
- From 2027, unit costs of Gen 6 and Gen 7, while still using the original technology, will approach a higher technology curve with an inflection point that is 15 years later, in 2016 (see ❾)
- After Gen 5, it will be necessary to decide if a Gen 6' and Gen 7' based on the higher, more recent technology would be a better alternative (unit cost plotted as hollow dots, see ❿); theoretically, a total of seven product generations can be realised by the end of service life in 2038
- **In this scenario, both, product and technology obsolescence are considered.**

The long horizontal row of hatched dots from 2003 to 2038 in Figure 11 represents the spare parts unit costs that a pre-LCCM model would have used to calculate the supply of obsolete units of the first product generation (Gen 1). The labelling “Fictitious spare parts Gen 1” shall point to the low likeliness of such a scenario. Traditional cost models simply lacked the input parameters needed to control the technology of spare parts independently from the original product.

The curve suggests that it would be more practical to substitute the original Gen 1 products with those from Gen 2 and subsequent generations. Cost models for LCCM shall achieve independent control of technology and obsolescence of replenishment spares, decoupled from production and initial spare parts supply. This leads to a significantly improved estimation of operation and support (O&S) costs if the real availability and any generational changes of spare parts can be mapped in the cost model.

3.3 Obsolescence of spare parts: The solution

To enable more realistic cost estimates for the operation and support (O&S) phase, mitigation strategies against obsolescence shall be considered in future. The longer the service life of a system, the more likely it is that obsolescence will affect the availability of spare parts. Fortunately, this problem can be mitigated in advance (Bundeswehr, 2016, p. 31ff). An LCCM-empowered model can support decision making by estimating the cost of alternative countermeasures. There are four options for mitigating obsolescence:

0. None (do nothing)
1. Last Time Buy (LTB). LTB avoids technical progress through final stockpiling of all required spare parts: freezing the existing technology, same technology curve, same technology year
2. Equivalent in form, fit and function (FFF). FFF utilises technical progress in the form of technology improvement, i.e.: further development of existing technology, same technology curve, later technology year)
3. Technology Refresh (TechRef). TechRef utilises technological progress in the form of technology growth, i.e. further development with newer technology, higher technology curve, later technology year

(Note: The additional option of having obsolete spare parts built-to-print, provided the user has the right to use and drawings, is not considered here. Simple reproduction of established products can be estimated using traditional cost models).

The rationales for alternative mitigation strategies against product obsolescence can be found in Table 7.

Product is discontinued	**Substitute with same technology available?**	**Substitute with higher technology available?**	**Mitigation**
No	-	-	None
Yes	No	No	Last Time Buy (LTB)
Yes	Yes	No	Equivalent in form, fit and function (FFF)
Yes	No	Yes	Technology refresh (TechRef)

Table 7: Rationale for alternative mitigation strategies against product obsolescence (PRICE 2022, p. 113)

Not taking any mitigation measures assumes that no product obsolescence occurs. This means that all spare parts are still available and remain the same throughout the entire operation phase. Such an option only makes sense for short service lives of a few years. This means that virtually all projects in the Bundeswehr environment – which often last several decades – will incur costs due to obsolescence (Bundeswehr, 2016, p. 7ff).

The three relevant countermeasures and their impact on spare parts costs are explained in the following subsections.

3.4 Last Time Buy (LTB)

Last Time Buy (LTB) as mitigation strategy after product obsolescence avoids technical progress. It is assumed that no equivalent alternative with the same technology is available for the obsolete product. With LTB, a sufficient quantity of the parts that are no longer produced or no longer available on the market is procured in advance and stored in order to cover the need for replacement parts over the remaining useful life of the system. This is why LTB is also known as *lifetime buy*.

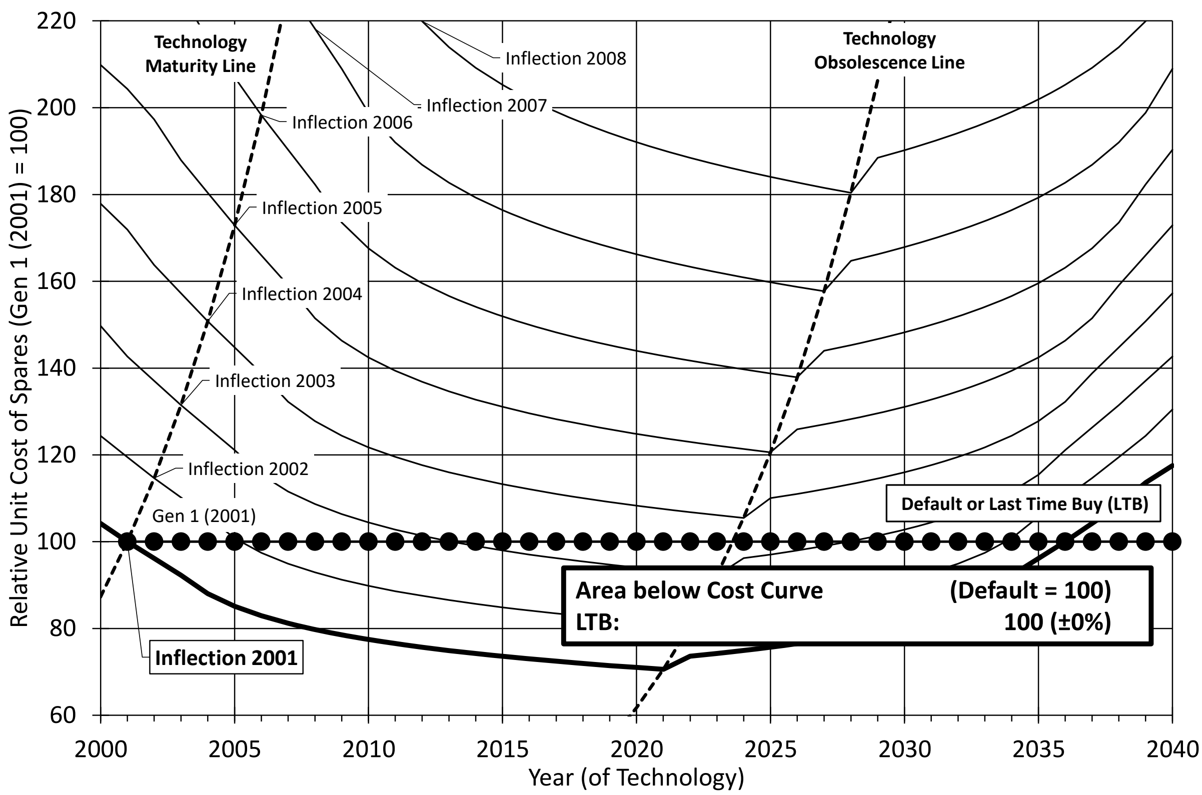


Figure 12: Relative spare part unit costs for obsolescence mitigation with "Last Time Buy" (based on PRICE 2022, p. 114)

The influence of a Last Time Buy (LTB) on spare parts costs is shown in Figure 12. The original product with its production units and spare parts is based on a technology with full market maturity in 2001; accordingly, the curve is called "Inflection 2001". Production also begins in 2001, i.e. the technology year is the same as the year of the inflection point, unlike in the example from Figure 10. Note: It is not mandatory that production begins at the inflection point of the technology curve; any other year is also possible.

The original product has a product life cycle of 5 years. Its production will be discontinued at the end of 2005. With the Last Time Buy (LTB), a conscious decision is made to not seek a successor for existing spare parts. Instead, it is ensured that sufficient spare parts are available until the end of the product's useful life. Accordingly, relative unit cost (before inflation) of spare parts remains constant over the entire service life. In addition, there will be costs for storage space and administration. They depend on the quantity and storage volume of the final stockpile but are not considered further here.

From the start of production, the constant unit costs of the selected technology year apply. Freezing technology at the start of production in 2001 leads to the horizontal unit cost line for all remaining years of service life (Figure 12). In this respect, LTB reflects the spare parts cost logic of traditional cost models. The only difference is that the outstanding spare parts requirement at the time of product obsolescence is summarised as a final stock and brought forward.

The following list shows which model inputs for an LTB act as cost drivers for spare parts:

- Product obsolescence: **Yes**; the original product is no longer produced
- Product successor: **No**; the existing product will not be replaced
- Product life cycle: **Not applicable**
- Technology improvement: **Not applicable**
- Technology obsolescence: **Not applicable**
- Technology growth: **Not applicable**
- Max. Number of technology refreshments: **Not applicable**
- Closure with Last Time Buy (yes/no): **Yes**; LTB is the sole mitigation here
- Time of Last Time Buy (year): **No later than product obsolescence occurs**

Global inputs, not shown in Figure 12:

- Costs for storage and management of the final stockpile: **Yes**, from the time of the LTB
- Costs for certification of substitute product(s): **No**, no substitute products required

When stocking spare parts, it should be noted that their permissible storage times may be limited. Furthermore, the probability of parts failing may increase due to ageing processes.

LTB can be combined with other strategies to finalise them. This makes sense, for example, as soon as production costs start to rise due to technology obsolescence.

3.5 Equivalent in form, fit and function (FFF)

Equivalent in Form, Fit, Function (FFF) employs technical progress in the form of technology improvement (further development of existing technology, same technology curve, later technology year). In this strategy, the costs of replacing the original product are determined by a successor product of a new generation based on the same technology (same Z-curve) but of a more recent date (later technology year). The successor product fulfils the same performance requirements but has the advantage of incorporating additional years of technology improvement, resulting in lower relative unit costs (before inflation). This applies for as long as the technology itself does not become obsolete.

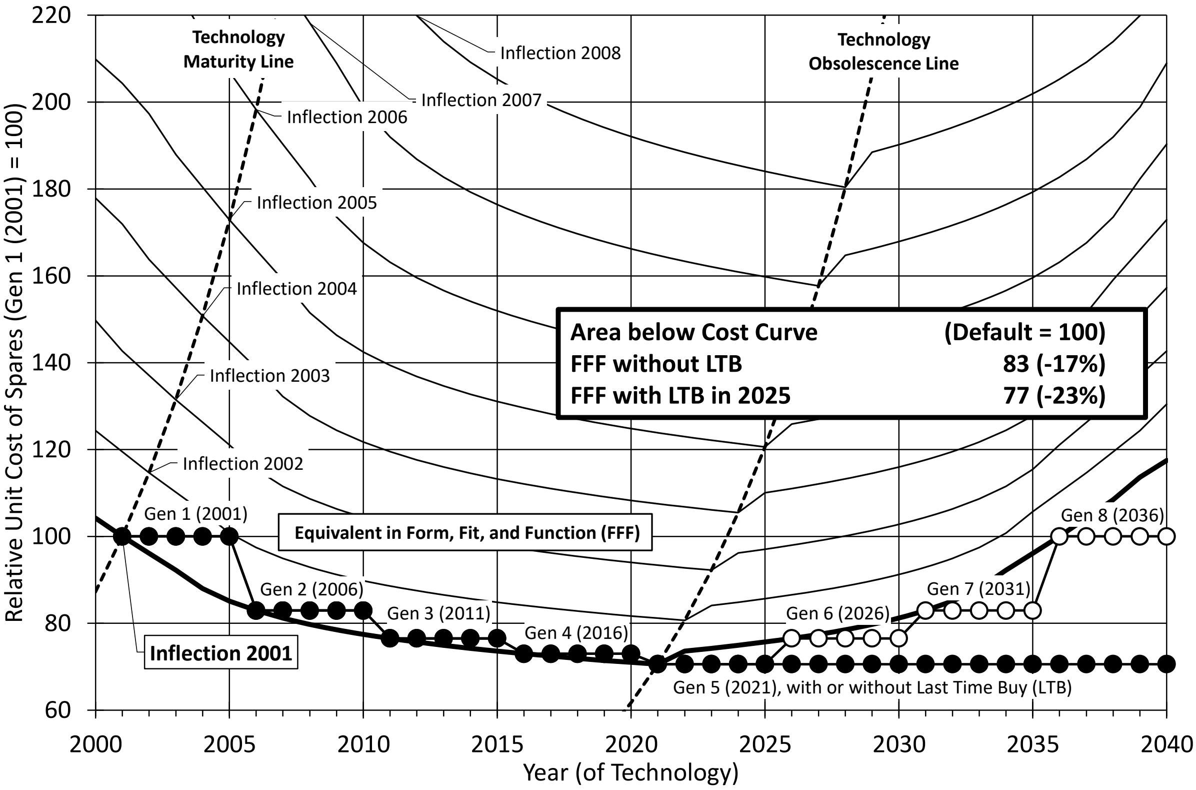


Figure 13: Relative spare part unit costs for obsolescence mitigation with "equivalent in FFF" (based on PRICE 2022, p. 114–115)

The influence of equivalent in FFF on the spare parts costs is shown in Figure 13. The initial product is identical to that shown in ▶ Sect. 3.4.

The original product of the first generation (Gen 1) has a product life cycle of five years. Its production will be discontinued at the end of 2005. The product family from which it originates has been developed further and offers a successor product (Gen 2). It offers the same form, fit and function (FFF) of its predecessor. It utilises the same technology with the same technology curve (Z-curve) but has a technology year five years later (2006). The relative unit costs are reduced in line with the technology improvement curve and frozen for the new generation (Gen 2). With a product life cycle of five years and a service life until 2040, this cycle can be repeated until eight product generations, from Gen 1 to Gen 8, have been completed.

However, it should be noted that in 2022, during the fifth generation (Gen 5), the technology used will itself begin to become obsolete! From this point onwards, the technology curve starts to rise again. Staying on the original curve for another three generations (Gen 6 to Gen 8) would therefore cause the relative unit costs to rise. In this case, it makes sense to carry out an LTB for the current generation (Gen 5) in 2025 before it is discontinued, and product obsolescence occurs.

The following list shows which model inputs for an equivalent in FFF act as cost drivers for spare parts:

- Product obsolescence: **Yes**; the original product is no longer produced
- Product successor: **Yes**; the existing product will be replaced
- Product life cycle: **Five years, max. eight product generations**
- Technology improvement: **Yes; each generation has a technology year five years later**
- Technology obsolescence: **Yes; starts in 2022, relevant from fifth generation (Gen 5) onwards**
- Technology growth: **Not applicable**
- Max. Number of technology refreshments: **Not applicable**
- Closure with Last Time Buy (yes/no): **Yes**; it is applicable
- Time of Last Time Buy (year): **Preferably at product obsolescence of fifth generation (Gen 5; 2025)**

Global inputs, not shown in Figure 13:

- Costs for storage and management of the final stockpile: **Yes**, if contract is concluded with LTB
- Costs for certification of substitute product(s): **Yes**, for each new product generation (Gen 2–8)

For equivalent in FFF, the area under the curve for relative spare part unit costs in Figure 13 is reduced compared to the default curve (constant value = 100). The reduction for equivalent in FFF without LTB is 17%. For equivalent in FFF with LTB in 2025, it is as much as 23%. Therefore, with forward-looking planning, costs (before inflation) for replenishment spares can be saved. This makes equivalent in FFF an attractive option against product obsolescence. This applies only as long as the primary aim is to minimise replenishment spares cost and manufacturers offer suitable successor products.

3.6 Technology Refresh (TechRef)

"Technology Refresh (TechRef)" utilises technological progress in the form of technology growth (further development with newer technology, higher technology curve, later technology year). This strategy is similar to the equivalent strategy in FFF, with the difference that the manufacturer replaces obsolete products with successors based on a later, higher technology curve. This means that the entire product family and the technology on which it is based are discontinued. The reason for using a higher technology may be increased performance, increased competition or increased customer expectations.

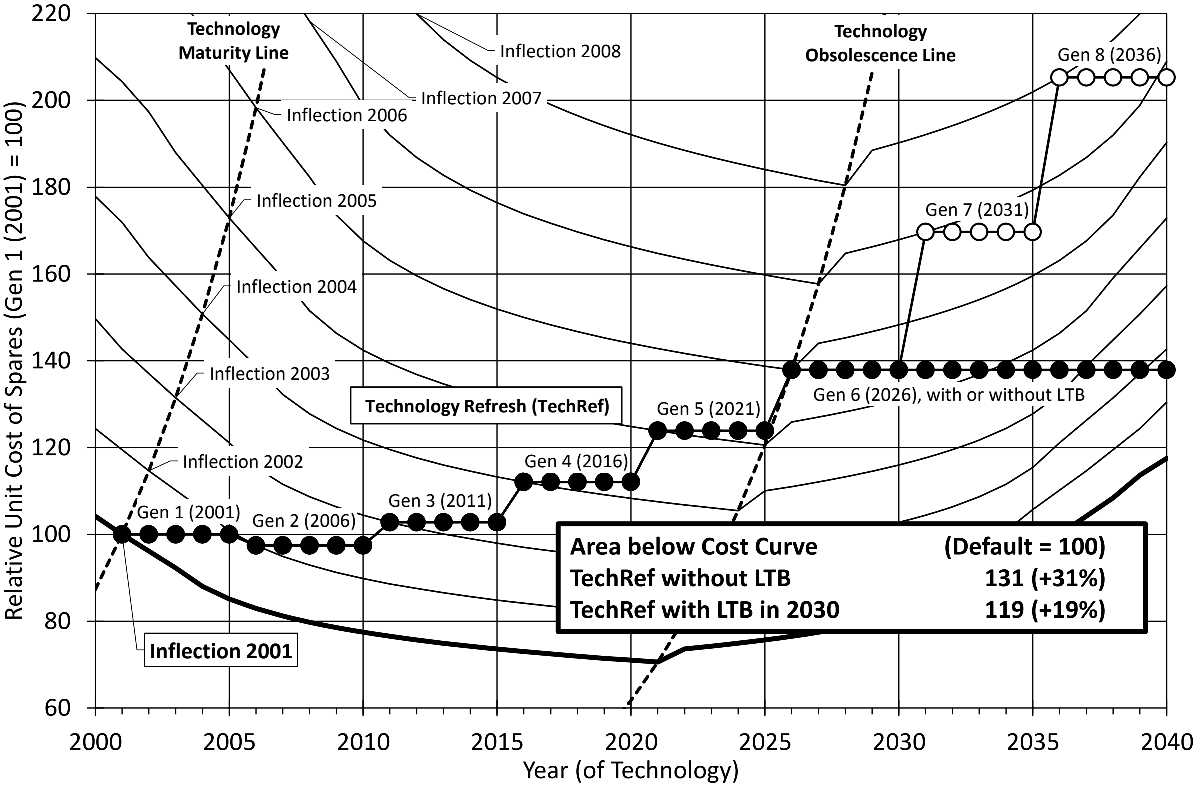


Figure 14: Relative spare part unit costs for obsolescence elimination with "technology refresh" (authors' illustration, based on PRICE 2022, p. 114–115)

The influence of technology refresh on spare parts costs is shown in Figure 14. The original product is identical to the one shown in ▶ Sect. 3.4.

The original product of the first generation (Gen 1) has a product life cycle of five years. Its production will be discontinued at the end of 2005. The product family from which it originates will no longer be developed further. Instead, there is a superior successor product. It utilises technology on a higher and later (inflection point 2002) technology curve. Relative unit costs are adjusted in line with the higher technology curve and frozen for the new generation (Gen 2) in the technology year 2006.

With a product life cycle of five years and a service life until 2040, this cycle is repeated until eight product generations, from Gen 1 to Gen 8, have been completed.

However, it should be noted that even the higher technology curves will experience the onset of technology obsolescence from Gen 7 onwards! In this case, it makes sense to carry out an LTB for the current Gen 6 by 2030 at the latest, before its product obsolescence sets in.

The following list shows which characteristic features of a technology refresh have an impact on the costs of replenishment spares:

The following list shows which model inputs for technology refresh act as cost drivers for spare parts:

- Product obsolescence: **Yes**; the original product is no longer produced
- Product successor: **Yes**; the original product is replaced by a successor of higher technology, with a later technology year
- Product life cycle: **Five years, max. eight product generations**
- Technology improvement: **Yes; each generation has a technology year five years later**
- Technology obsolescence: **Yes; starts in 2027, relevant from sixth generation (Gen 6) onwards**
- Technology growth: **Yes; each generation's technology curve has an inflection point one year later**
- Max. Number of technology refreshments: **Seven**
- Closure with Last Time Buy (yes/no): **Yes**; it is applicable
- Time of Last Time Buy (year): **Preferably at product obsolescence of sixth generation (Gen 6; 2030)**

Global inputs, not shown in Figure 14:

- Costs for storage and management of the final stockpile: **Yes**, if contract is concluded with LTB
- Costs for certification of substitute product(s): **Yes**, for each new product generation (Gen 2–8)

Technology Refresh (TechRef) is the costliest mitigation strategy for obsolescence. The area under the curve for relative spare part unit costs in Figure 14 is significantly larger compared to the default curve (constant value = 100). The increase for TechRef without LTB is 31%. For TechRef with LTB in 2030, this increase can be reduced to 19%.

3.7 Obsolescence mitigation of spare parts: Implementation in the LCCM cost model

Additional input parameters are required in a LCCM-enabled cost model to model product obsolescence and its influence on spare parts costs. These are listed in Table 8.

Cost model, traditional	Cost model, LCCM
Technology year [calendar year]	Desired mitigation [LTB, FFF, TechRef]
Factor for technology improvement	Start of product obsolescence [calendar year]
Factor for technology obsolescence	Product cycle [years]
(Product obsolescence is **not** modelled)	Technology growth with technology refresh [%]
	Planned number of technology refreshments
	Optional closure with Last Time Buy (yes/no)
	Time of last time buy [year]
	Storage space for final stockpile [m³]

Table 8: Additional input parameters for modelling various obsolescence mitigation strategies for spare parts

The most important innovation in the input parameters for LCCM concerns the option to choose between a total of four mitigation strategies against product obsolescence (including “None” as the default setting).

3.8 Modelling of obsolescence mitigation for spare parts: Discussion of benefits

The obsolescence of spare parts complicates the operation and support of military systems over their service life. If certain products are no longer available, alternative solutions must be found to keep the systems functioning. This has an impact on the cost of spare parts and therefore on the total cost of operation and support.

Traditional cost models have not yet been able to estimate the impact of product obsolescence. The technology available at the start of production of a system was frozen. This meant that every serviceable unit within a system and its spare parts remained technically unchanged over its entire service life. This was always the case, even for very long service lives of 40 years or more.

It doesn’t have to stay that way. Existing cost models can be upgraded with additional functions to model the technology improvements or technology leaps of spare parts independently of their parent system. Spare parts can go through several generations, either within the same technology or with a change to higher technologies. Spare part costs change accordingly. This allows planners to model alternative measures against product obsolescence and their costs in detail.

As a result, a cost model optimised for LCCM provides better cost estimates for the mitigation of product obsolescence in spare parts. The cost effects on replenishment parts shown here amount to 30% additional costs at worst in the case of technology refreshment (TechRef). In return, however, savings of over 20% can also be achieved if the product generations make extensive use of technology improvements in the case of equivalent in FFF (all figures without inflation).

4 Conclusion

The quality of operation and support cost estimates can be improved. It is sufficient to consider some influencing factors that have so far been ignored by traditional cost models. According to studies by IABG, the three most important are (1) non-constant failure rates due to teething troubles and ageing effects, (2) “No Fault Found” (NFF) in condition-based maintenance and (3) product obsolescence of spare parts.

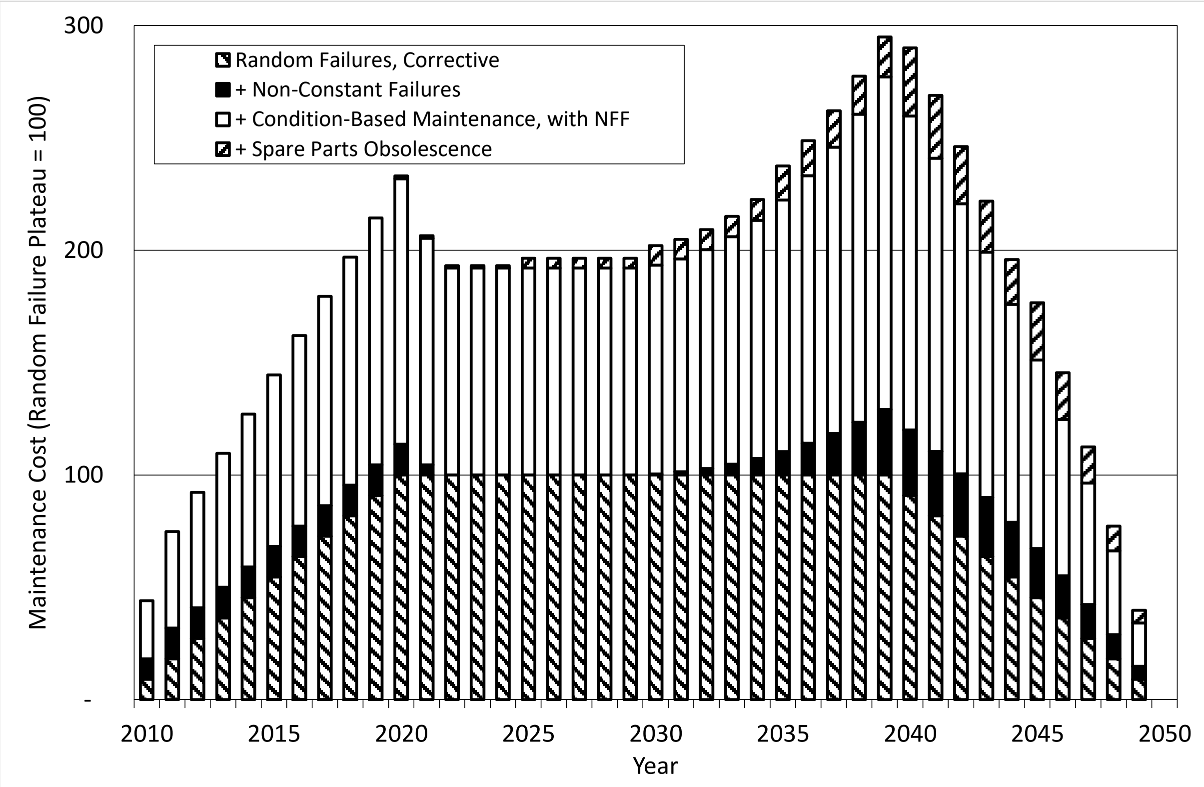


Figure 15: Enhanced life cycle cost estimating will lead to higher, more realistic O&S cost estimates

Considering the three influencing factors mentioned increases the estimated costs in various operation and support cost categories. Typical increases in the **cost of maintenance** are:

- plus 10–70% for additional effort due to higher failure rates,
- plus 100–300% for additional removal, transport, testing and reinstallation effort after NFF,
- plus 10–30% for mitigating product obsolescence of replenishment spares.

The stacked column chart in Figure 15 visualises the added costs stated above. The area under the stacked curve is 150% greater than the original baseline, the latter simply assuming corrective maintenance of units after random failures. All percentages stated are subject to uncertainty and their exact amount may depend on additional factors that have not yet been determined. This becomes even more valid the further an estimate looks into the future. Nevertheless, there is no reason to neglect the three big “known unknowns” presented in this paper, as they increase the estimating quality of operation and support cost.

When in doubt, the following life wisdom should always apply: “Completeness before accuracy.”

5 Bibliography

Bartels, B. et al: Strategies to the Prediction, Mitigation and Management of Product Obsolescence. Wiley, Hoboken (2012)

Ben-Daya, M.: Handbook of maintenance management and engineering. Springer, London (2009)

Blanchard, B.S.: Design and Manage to Life Cycle Cost. M/A Press, Portland (1978)

Bundeswehr: Life Cycle Cost Management - Methodological Documentation. Bundeswehr Division Regulation C1-1510/1-1, Bonn (2015)

Bundeswehr: Obsolescence management . Bundeswehr, Bonn (2016), Area Instruction C1-1500/3-7002

Bundeswehr: Customer Product Management (CPM). Central Service Regulation A-1500/3, Version 2, Bundeswehr, Bonn (2018)

Bundeswehr: Life Cycle Cost Management - Methodological Documentation. Division regulation C1-1510/1-1. version 2. Bundeswehr, Bonn (2022a)

Bundeswehr: Guidelines for life cycle cost management in the BAAINBw. Division regulation C1-1510/1-2. version 2. Bundeswehr, Bonn (2022b)

Crowley, C.: Meeting the ageing aircraft challenge. University of New South Wales, Sydney (2004)

DIN German Institute for Standardisation: Obsolescence management; EN 60242:2007. DIN, Berlin (2008)

DIN German Institute for Standardisation: Maintenance - Maintenance terminology; EN 13306:2010. DIN, Berlin (2010)

Dixon, M.: The maintenance costs of aging aircraft - insights from commercial aviation. RAND, Santa Monica (2006)

Francis, P. and Shaw, G.: Effect of aircraft age on maintenance costs. Centre for Naval Analyses, Alexandria (2000)

Galar, D. et al. Maintenance Costs and Lifecycle Cost Analysis. CRC, Boca Raton (2017)

Greenfield, V. and Persselin, D.: An economic framework for evaluating military aircraft replacement. RAND, Santa Monica (2002)

Humphrey, D. et al: Utilisation life of electronic systems - aging avionics usable life and wear-out issues. SAE, Warrendale (2002)

Kiley, G.: The effects of aging on the costs of operating and maintaining military equipment. Congressional Budget Office, Washington (2001)

Moore, G.E.: Cramming more components onto integrated circuits. Electron. **38.8**, 114–117 (1965)

Moubray, J. Reliability-centred maintenance. 2nd ed. Industrial Press, New York (1997)

Nowlan, F. and Heap, H.: Reliability-centred maintenance. Department of Defense, Washington (1978)

Pfingsten, S. and Engelking, N.: Extended lifetime: Tornado takes off again. https://www.bundeswehr.de/de/organisation/luftwaffe/aktuelles/verlaengerte-lebenszeit-tornado-hebt-wieder-ab-5029012 (2021). Accessed: 29 January 2024

PRICE Systems: Upgrading technology improvement in PRICE. Tech. Bull. 4, PRICE Systems, Moorestown (1981)

PRICE Systems: Your guide to PRICE HL: Estimating cost and schedule of hardware operation and support. PRICE Systems, Mount Laurel (2001)

PRICE Systems: Your guide to PRICE H: Estimating cost and schedule of hardware development and production. PRICE Systems, Mount Laurel (2003)

PRICE Systems: LCCM modelling Eurofighter. PRICE Systems, Rüsselsheim (2016)

PRICE Systems: Appendix C - Technology improvement in the hardware model, from: Description of the PRICE models: How the models work. Version 0.96. PRICE Systems, Mount Laurel (2017)

PRICE Systems: TruePlanning for Hardware Life Cycle Cost Management (True HLCCM) User Manual. Product Version 1.01. PRICE Systems, Rüsselsheim (2022)

Pyles, R.: Aging aircraft - USAF workload and material consumption life cycle patterns. RAND, Santa Monica (2003)

Rogers, E.M.: Diffusion of Innovations. 5th ed. New York, Free Press (2003)

Sandborn, P.: Cost Analysis of Electronic Systems. 2nd ed. Singapore, WSPC (2017)

Shermon, D. et al: Systems cost engineering: Programme affordability management and cost control. Gower, Farnham (2009).

Smith, D.: Reliability, maintainability and risk. 8th ed. Butterworth-Heinemann, Oxford (2011)

Stoll, L. and Davis, S.: Aircraft age impact on individual operating and support cost elements. Naval Aviation Maintenance Office, Patuxent River (1993)

Trunkey, D.: Operating costs of aging air force aircraft. Congressional Budget Office, Washington (2018)

Trunkey, D.: Operating Costs and Aging of Air Force Aircraft. NASEM - Air Force Studies Board, Washington, (2019)

Ungar, L. Causes and Costs of No Fault Found Events. https://www.researchgate.net/publication/273456994_Causes_and_Costs_of_No_Fault_Found_Events (2015). Accessed: 29 January 2024

Valentin, R. et al: Remaining life assessment of aging electronics in avionic applications. University of Maryland, College Park (2003)

Werner, D. A maddening, costly problem. Aerosp. Am. **53.2**, 28–33 (2015)

6 The authors

Fabian Eilingsfeld, Birgit Biehn and Christian Schindler work together in the Logistics & Cost Engineering department at IABG Industrieanlagen-Betriebsgesellschaft mbH in Ottobrunn, Germany. For many years, they have been supporting Bundeswehr projects with LCC data collection and the creation and updating of LCCM-compliant cost estimates. From 2019 to 2022, Fabian Eilingsfeld and Christian Schindler were involved in the development of True HLCCM, an extension of the PRICE TruePlanning cost estimating tool that has been specifically optimised for LCCM.

Contact: eilingsfeld@iabg.de ; biehn@iabg.de ; schindlerc@iabg.de