Chapter 2
System Degradation and Maintenance

2.1 Introduction

Every system (product, plant or infrastructure) is unreliable in the sense that it degrades and eventually fails. Maintenance is needed to compensate for this unreliability. Any decision-making with respect to maintenance requires a proper understanding of the degradation processes over time and the actions of maintenance from a system life cycle perspective. The life cycle from the manufacturer (or builder) perspective is different from that of the customer (or owner). This chapter looks at these issues using concepts from reliability theory and the characterisation of maintenance actions. It forms the basis for the modelling of the degradation process and maintenance actions which is the focus of the next chapter.

The outline of the chapter is as follows. Section 2.2 deals with system life cycle from both customer (owner) and manufacturer (builder) perspectives. Section 2.3 looks the characterisation of system (product, plant and infrastructure) performance. Section 2.4 deals with product reliability and looks at various issues such as linking component reliability to product reliability and different notions of reliability from a product life cycle perspective. Section 2.5 looks at maintenance and the characterisation of different types of maintenance actions appropriate for products and plants and Sect. 2.6 looks at maintenance of infrastructures.

2.2 System Life Cycle

The life cycle of a system is basically the period of time during which it is in existence, either conceptually or physically, and may be defined in various ways. The life cycle for products differs somewhat from that for plants or infrastructures.

D. N. P. Murthy and N. Jack, *Extended Warranties, Maintenance Service and Lease Contracts*, Springer Series in Reliability Engineering,
DOI: 10.1007/978-1-4471-6440-1_2, © Springer-Verlag London 2014

2.2.1 Products

The life cycle for a product (consumer, commercial or industrial) is commonly referred to as the product life cycle (PLC).

2.2.1.1 Manufacturer Perspective

The PLC for a standard consumer durable or an industrial product, from the point of view of the manufacturer, is the time from initial concept of the product to withdrawal of the product from the marketplace. The life cycle involves six stages, as indicated in Fig. 2.1.[1]

The process begins with the idea of building a product to meet some customer requirements, such as performance targets. This is usually based on a study of the market and the potential demand for the product being planned. The next step is to carry out a feasibility study. This involves determining if it is possible to achieve the targets within specified cost limits. This analysis is done in the front-end stage (Stage 1) of Fig. 2.1.[2]

If the analysis indicates that the project is feasible, an initial product design is undertaken. A prototype is then built and tested. It is not unusual at this stage to find that achieved performance levels of the prototype product are below the target values. In this case, further product development is undertaken to overcome the problem. These define the Stages 2 (Design) and 3 (Development) of the PLC as shown in Fig. 2.1. Once these are achieved, the next step is to carry out trials to determine performance of the product in the field and to start a pre-production run. This is required because the manufacturing process must be fine-tuned and quality control procedures established to ensure that items produced have the same performance characteristics as those of the final prototype.

After this, the production and marketing efforts begin. These constitute Stages 4 (Production) and 5 (Marketing) of the PLC shown in Fig. 2.1. The items are produced and sold. Production continues until the product is removed from the market because of obsolescence and/or the launch of a new product. Post-sale support of the product continues at least until expiration of the warranty on the last item sold but can continue beyond this point in terms of spare parts, service contracts, etc. This defines Stage 6 (post-sale) of the PLC.

2.2.1.2 Customer Perspective

From the consumer's viewpoint, the PLC is the time from the purchase of an item to its discarding when it reaches the end of its useful life or is replaced earlier due

[1] The number of stages in the PLC can vary. For more on this, see Murthy et al. (2008).

[2] The Front End stage is also often referred to as the Feasibility stage.

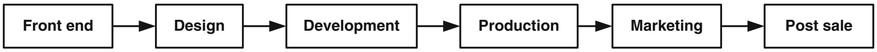


Fig. 2.1 Product life cycle (manufacturer perspective)

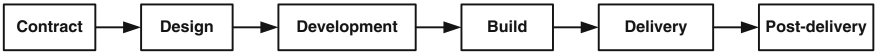


Fig. 2.2 Product life cycle for plants and infrastructures (manufacturer perspective)

to either technological obsolescence or the item being no longer of any use. The life cycle involves the following three phases:

- Purchase.
- Operation and maintenance.
- Discarding (leading to replacement by a new one).

2.2.2 Plants and Infrastructures

2.2.2.1 Builder Perspective

The life cycle for a custom built system (product, plant or infrastructure) is slightly different and is as shown in Fig. 2.2. Here, the initial requirements of the plant or infrastructure are specified by the owner and then jointly agreed through discussions leading to a contract that specifies the final agreed requirements. The builder then builds the plant or infrastructure to the specifications stated in the negotiated contract. The process then follows basically the same steps as those for products.

2.2.2.2 Owner Perspective

From the consumer's viewpoint, the life cycle is the time from the initiation of the process and to discarding or upgrading the plant or infrastructure. As such the life cycle involves all the phases (except post delivery) shown in Fig. 2.2 and the following additional phases after the delivery phase:

- Operation and maintenance.
- Discarding or major upgrade (leading to a new life cycle).

2.2.3 Salvage Value

The salvage value of a system is the value of the system at the end of its economical or useful life. It is used in accounting to determine depreciation amounts and to determine deductions for taxation purposes. The value can be a best guess of the end value (or determined by a regulatory body such as the Taxation Department). It depends on the state of the system and is influenced by factors such as usage, maintenance and technological obsolescence. It is also referred to as the *residual value*.

2.3 Characterisation of System Performance

Every system (product, plant and infrastructure) is designed for some specified performance as illustrated by the following example:

- Electric bulb (Product): To produce light.
- Engine (Product): To operate to some specified efficiency.
- Power station (Plant): To produce specified output with cost/unit below some specified value.
- Rail system (Infrastructure): To provide passenger service to some specified schedule (frequency and punctuality) at a cost below some specified value.

The performance of a system is a function of the *condition* or *state* of the system. The state of a system, in turn, depends on the state of its elements. We first look at the characterisation of component state and then the characterisation of the state of products, plants and infrastructures.

2.3.1 Characterisation of Component State

The condition of a component (of a product) degrades with time (and usage) and can be characterised through a variable $X(t)$ which represents the *state* of the component. Note that $t = 0$ corresponds to the instant a new component is put into use for the first time. We have three different characterisations with increasing degrees of detail.

2.3.1.1 Characterisation 1 (Binary)

Here, $X(t)$ is binary valued with

- $X(t) = 1$ corresponding to the component being in the *working state* (performance satisfactory or acceptable), and
- $X(t) = 0$ corresponding to component being in the *failed state* (performance is unsatisfactory or unacceptable).

The component starts in the working state and changes to the failed state after a period $\tilde{T}$ as shown in Fig. 2.3. $\tilde{T}$ is the time to failure (or lifetime of the

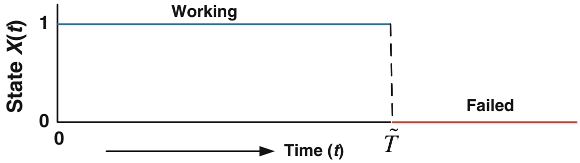


Fig. 2.3 Time to failure (binary characterisation)

component). This is a random variable[3] as the time instant of change from working to failed is uncertain.

A typical example where this characterisation is appropriate is an electric bulb where the state changes from working to failed in a very short time so that it can be viewed as being instantaneous.

2.3.1.2 Characterisation 2 (Finite Number of Levels)

Here, $X(t)$ can assume values from the set $\{1, 2, \ldots, K\}$ with

- $X(t) = 1$ corresponding to component performance being fully acceptable (component is in the good *working state*),
- $X(t) = i,\ 1 < i < K,$ corresponding to component performance being partially acceptable (component is in a working state with a higher value of i implying a higher level of degradation) and,
- $X(t) = K$ corresponding to component performance being unacceptable (component is in the *failed state*).[4]

The time to failure of the component is given by $\tilde{T} = \inf\{t : X(t) = K\}$ as shown in Fig. 2.4. Let $\tilde{T}_i$ denote the duration for the time the component state is $i,\ 1 \le i \le K - 1$. This is a random variable, and as result the time to failure is the sum of $(K - 1)$ random variables.

A typical example where this characterisation is appropriate is the wear in a tire where no wear corresponds to state 1 and complete wear corresponds to state K.

2.3.1.3 Characterisation 3 (Infinite Number of Levels)

This is an extension of the above case with $K = \infty$. $X(t)$ is now a non-decreasing continuous time stochastic process as shown in Fig. 2.5. Here, a higher value of $X(t)$ implies greater degradation, and the component failure time is given by $\tilde{T} = \inf\{t : X(t) = x*\}$.[5]

---

[3] See Appendix A for a definition of a random variable and an introduction to probability theory.

[4] The numbering of states is arbitrary. One can easily reverse the order so that the lower the state the greater the degradation.

[5] In some cases $X(t)$ could be non-increasing with lower values corresponding to greater degradation. In this situation the curve in Fig. 2.5 would be downward sloping.

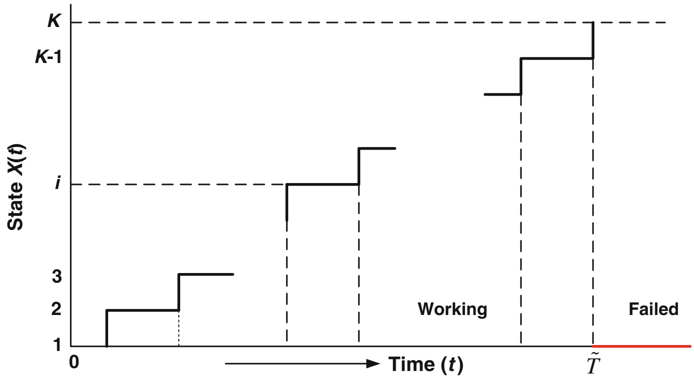
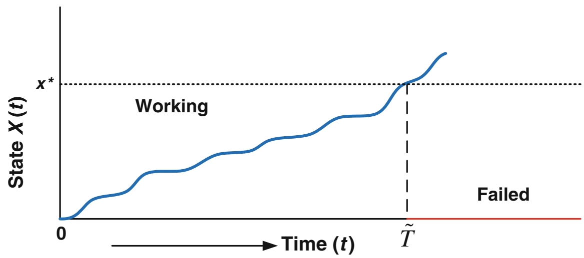


Fig. 2.4 Time to failure (multistate characterisation with finite states)

Fig. 2.5 Time to failure (multistate characterisation with infinite states)



A typical example, where this characterisation is appropriate, is failure due to crack growth in a pipe. The state depends on the crack length, and failure occurs when the crack length reaches some critical value.

Comment: The rate of deterioration of the state depends on factors that impact on the stress (thermal, mechanical, etc.) on the component. The stresses are, in turn, influenced by the load or throughput of the system.

2.3.2 Characterisation of Product (Plant) State

At the product (plant) level, the characterisation of the state is more complex and two approaches can be used. In the first approach, the product (plant) is viewed as a black box and the characterisation is done in a manner similar to the previous subsection with $X(t)$ denoting the state [defined by the output of the product (plant)]. The second approach views the product (plant) in terms of its components. Each component is characterised using Characterisation 1 discussed earlier, and fault tree analysis (FTA) is used to link the product-(plant) level characterisation to the component-level characterisation.

2.3.2.1 Fault Tree Analysis

A fault tree is a logic diagram that displays the relationship between a potential event affecting product (plant) performance and the reasons or underlying causes for this event. The reason may be failures (primary or secondary) of one or more components of the system, environmental conditions, human errors and other factors.

A fault tree illustrates the state of the system denoted the TOP event (binary characterisation—working/failed) in terms of the component states (binary characterisation—working/failed) denoted basic events. The connections are done using gates, where the output from a gate is determined by the inputs to it. A special set of symbols (for gates and basic events) is used for this purpose.[6]

2.3.2.2 Multiunit Plants and Service Facilities

Many industrial plants and service facilities have multiunits—for example a power plant having three units (each with output capacity 50, 100 and 200 MW, respectively), a bus operator having a fleet of K buses. In this case, the system state (from a business-level perspective) is best done in terms of the different levels of output. In the case of the power plant, the different levels correspond to 350 MW (all three units working), 0 MW (all three failed) and six different levels (50, 100, 150, 200, 250 and 300 MW) depending on the number of units (1 or 2) in failed state.

Comment: The rate of deterioration of the state depends on several factors such as the production (throughput) rate, environmental factors and maintenance actions.

2.3.2.3 Fleet

A fleet refers to multiple units of an asset (such as machines, automobiles, ships aircraft, computers, etc.). A fleet can be viewed as a multiunit system, where each unit operates independently and a failure of a unit does not result in the failure of the system but can affect the overall performance (e.g. production capacity) of the system.

---

[6] The extension of FTA to the case where the performance is based on Characterisation 2 is more complex. For further details see, Blischke and Murthy (2000) or Rausand and Høyland (2004).

2.3.3 Characterisation of Infrastructure State

The characterisation of the state of an infrastructure is still more complex for the following reasons:

- There are several types of infrastructure—road, rail, utilities (gas, water sewerage, etc.), concrete structures (dams, buildings, bridges, etc.) and communication networks, etc., and each is different.
- Most infrastructures involve two types of elements—(i) discrete or lumped (similar to a product or plant), and (ii) distributed (with a spatial dimension). The characterisation of the state for the discrete elements is similar to that for products and plants discussed earlier and hence will not be discussed here.
- If one focuses on the distributed elements, the term *quality* is often used to indicate the state or condition. This in turn is defined through terms such as *damage*, *defects*, etc. Also, the characterisation of failure is not so clear. Often failure is defined to occur when the quality falls below some specified norm.
- Often there are several parties involved each with a different objective.
- The performance characterisations for each party are different and each involves a multitude of variables.
- The degradation of the state of the infrastructure is influenced by several factors such as weather, state of the system, usage intensity and output of (or throughput through) the system.
- Each party's performance of interest is different, but they are all functions of the state of the infrastructure.
- Safety also plays a role as poor condition of the asset can lead to dramatic consequences, e.g. in the case of tracks, roads, etc.

We confine our discussion to road infrastructure.

2.3.3.1 Road Infrastructure[7]

Road infrastructure consists of pavements (or "roads") and other items such as traffic signals, signs, etc. There are two types of pavements—rigid and flexible. Rigid pavements consist of a thick concrete top surface. Flexible pavements have a flexible layer on top of the surface.

When a road is built, the surface is dugout down to the designed depth of the intended road. Preparation is carried out on the ground now exposed below (such

[7] The material for the remainder of this section is based on Worm and van Harten (1996). For other issues relating to road maintenance, can be found in Dekker et al. (1998) and Rose and Bennett (1992).

as compaction). The road itself will then be built up above, usually consisting of several layers. The two bottom layers are as follows:

- Subgrade: The ground that is exposed once the ground has been dugout ready to build the road. The top level of this is termed the formation
- Capping: This is a layer added above the subgrade to protect it in new constructions (and often constitutes the formation).

This is followed by four more layers (in ascending order from bottom to top) are

- Sub-base,
- Base,
- Binder Course,
- Surface Course.

The nearer the surface, the profile needs to be more flatter as an uneven surface will be uncomfortable for vehicle occupants and will wear more quickly (as each time a vehicle hits a bump the hammering effects impacts on the surface). These factors are the main reasons for the layered construction of the road. Weight on any unbound material will compact it down with time, as material is forced down and fills gaps. For this reason, during construction of each layer compaction is carried out.

The most commonly used material for use in sub-base and base is an unbound material made from crushed rock, crushed slag, crushed-concrete and recycled aggregates. The binder course helps distribute the load of traffic above onto the base course, which is usually a weaker material. Materials used include open-graded macadam,[8] dense-coated macadam and rolled asphalt. Surface courses are laid in a wide range of bituminous materials, ranging in thickness from 20 to 40 mm. The material selected is dependent on the anticipated traffic intensity. Asphalt pavement is known for its durability and resilience.

The deterioration of a road depends on the materials used in the construction of the road and several other factors. In the case of asphalt pavement, the deterioration is because the materials that make up asphalt begin to break down over time and are affected by elements such as rain, sunlight and chemicals that come into contact with the pavement surface. The liquid asphalt binder that is the “glue” of the pavement begins to lose its natural resistance to water, allowing it to penetrate into and underneath the pavement. Once this happens, the surface can quickly fall prey to a number of different types of deterioration. The premature deterioration of asphalt pavement is usually due to failures in construction and/or human error and includes the following factors:

[8] Compacted broken stone usually bound with tar or asphalt (also referred to as bitumen).

Table 2.1 Groups of damage

Groups of damage	Features of damage
Texture	Ravelling[a], skidding resistance[b]
Evenness	Transverse and longitudinal evenness[c], irregularities[d], roughness[e]
Soundness	Transverse and longitudinal cracks[f], crazing[g], potholes[h], marginal strip, edge damage[i], kerb
Miscellaneous	Water run-off, verge[j]

[a] *Ravelling* loss of aggregate (used in road construction) due to (i) cohesive failure of the bituminous mortar, or (ii) adhesive failure in the adhesive zone

[b] *Skid resistance* characterises the cumulative effects of snow, ice, water, loose material and the road surface on the traction produced by the wheels of a vehicle

[c] *Longitudinal* (*transverse*) *evenness* measurement of longitudinal (transverse) profiles for determination of rutting. A rut is sunken track or groove made by the passage of vehicles within pavement layers that accumulates over time

[d] *Irregularity* something irregular, such as a bump in a smooth surface

[e] *Roughness* deviations of surface from true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads and drainage. It is defined using the International Roughness Index (IRI)—deviations (in metres) per kilometre. Roughness is a function of age, strength, traffic loading, potholes, cracking, ravelling, rutting, environment, etc

[f] *Longitudinal* (*transverse*) *cracking* cracks that run along (perpendicular to) the road

[g] *Craze* a fine crack in a surface of the road

[h] *Pothole* an open cavity in road surface with at least 150 mm diameter and at least 25 mm depth

[i] *Edge Damage* loss of bituminous surface material (and possibly base materials) from the edge of the pavement, expressed in square metres per km

[j] *Verge* a strip of grass or other vegetation beside a road

- Insufficient or improperly compacted base below the asphalt.
- Over or under compaction of the asphalt.
- Improper temperature of the asphalt when applied.
- Poor drainage.

Cracks, potholes, edge defects, depressions and corrugations are the significant road defects observed in the field. Traffic, age, road geometry, weather, drainage, construction quality as well construction material and maintenance policy are the major factors that affect the deterioration of a road. The state of a road surface can be described by a state vector where each component corresponds to one of the groups of damage, and these (and the features of damage for each group) are indicated in Table 2.1.

The quality of a road is often described through a function involving one or more of these features. Figure 2.6 illustrates the quality deterioration over time (as defined through some feature such as ravelling) with no maintenance actions.

The *quality standards* (also referred to as *norms*) for a road are derived from the lowest acceptable value for these features. They can be (i) local—for segments (for example 100 m in length) of a road (or lane) or (ii) global—the whole length of the

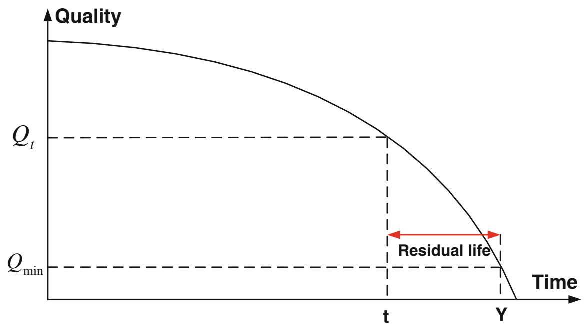


Fig. 2.6 Deterioration of road quality over time

road. In Fig. 2.6 at time t (since the construction of the road), the quality is Q_t and Y denotes the time when the quality reaches the minimum acceptable level $Q_{\min}$ at which instant CM action is needed. The interval $(Y - t)$ provides a window over which PM action can be initiated to avoid the need for CM action.

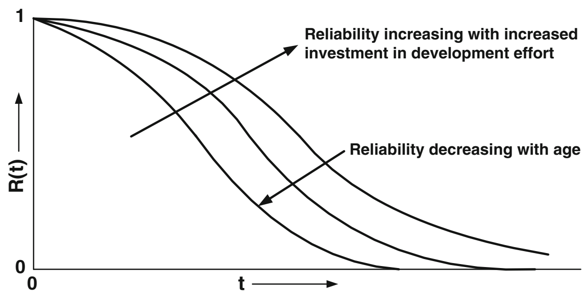
The performance of road transport is a complex function of the quality of the road. It is a vector that characterises the flow rate (number of cars passing per unit time) which would depend on the number of lanes open for traffic and the quality of ride. These depend on the speed of travel, which in turn, depends on the condition of the road (potholes, roughness to ensure grip) and weather conditions (rain, snow, etc.). From the public perspective, the quality of ride and safety are the important performance measures. The latter is also of importance to the regulators. From a road owner's perspective, cost of maintenance (to ensure the minimum standards for safety) and profits (in the case of toll roads operated by private business enterprises) are two important performance measures.

2.4 Reliability

The reliability of a product (component or some intermediate element) conveys the concept of dependability, successful operation or performance and the absence of failures. It is an external property of great interest to both manufacturer and consumer. Unreliability (or lack of reliability) conveys the opposite. More technical definitions of reliability are the following:

> The ability of an item to perform a required function, under given environmental and operational conditions and for a stated period of time (ISO 8402 1986).
>
> The reliability of a system is the probability that the product (system) will perform its intended function for a specified time period when operating under normal (or stated) environmental conditions (Blischke and Murthy 2000).

Fig. 2.7 Plots of reliability functions



The reliability is given by a function $R(t;\,\theta)$ with the following properties[9]:

- $R(t)$ is a non-increasing function of t, $0 \leq t < \infty$.
- $R(0) = 1$ and $R(\infty) = 0$.

Typical plots of $R(t)$ are shown in Fig. 2.7.

2.4.1 Linking Product and Component Reliabilities

The linking of component reliabilities to product reliability is done through the structure function. This function can be obtained using either FTA (discussed in the previous subsection) or a reliability block diagram (RBD).

2.4.1.1 Reliability Block Diagram

In a RBD, each component is represented by a block with two end points. When the component is in its working state, there is a connection between the two end points. This connection is broken when the component is in a failed state. A product (system) can be represented as a network of such blocks, each with two end points. The product (system) is in the working state if there is a connected path between the two end points. If no such path exists, then the system is in a failed state.

2.4.1.2 Structure Function

A product contains n components, and $X_i(t)$, $1 \leq i \leq n$, denotes the state of component i at time t, with

[9] θ is the set of parameters for the reliability function. Often we will suppress this and use $R(t)$ instead of $R(t;\theta)$ for notational ease. $F(t) = 1 - R(t)$ is called the failure distribution function and characterises the time to first failure (a random variable).

$$X_i(t) = \begin{cases} 1 & \text{if component } i \text{ is working at time } t \\ 0 & \text{if component } i \text{ is failed at time } t \end{cases} \quad (2.1)$$

Let $X(t) = (X_1(t), X_2(t), \ldots, X_n(t))$ denote the state of the n components at time t, and $X_S(t)$ (a binary variable) denote the state of the system at time t. Then, from FTA or the RBD, one can derive an expression of the form

$$X_S(t) = \phi\left(\underset{\sim}{X}(t)\right), \quad (2.2)$$

which links the component states to the system state. $\phi(\cdot)$ is called the *structure function.*[10]

Let $\underset{\sim}{R}(t) = (R_1(t), R_2(t), \ldots, R_n(t))$ denote the set of reliability functions of the n components of the product and $R_S(t)$ the reliability function for the system. If the component failures are independent, then

$$R_S(t) = \phi\left(\underset{\sim}{R}(t)\right) \quad (2.3)$$

so the system reliability can be expressed in terms of the component reliabilities. When failures are not independent, deriving the expression for the structure function is more complicated.

2.4.2 *PLC Perspective: Different Notions of Reliability*

From a product life cycle perspective, there are several different notions of reliability. Figure 2.8 (Murthy et al. 2008) shows how these are sequentially linked and the factors that affect them. We briefly discuss four reliability concepts.

2.4.2.1 Design Reliability

At the design stage, the desired product reliability is determined through a trade-off between the cost of building in reliability and the consequences of failures. This trade-off is discussed in detail in Murthy et al. (2008). From this, one derives the reliability specification at the component level. One then evaluates the design reliability.

[10] The details can be found in many books on reliability; see, for example, Blischke and Murthy (2000) and Rausand and Høyland (2004).

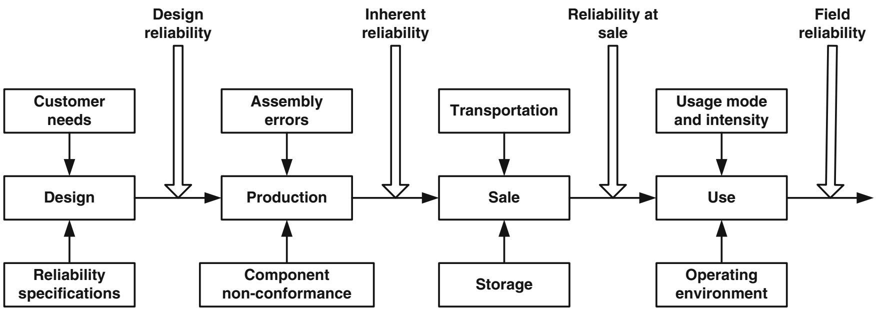


Fig. 2.8 Different notions of reliability (standard product)

2.4.2.2 Inherent Reliability

For standard products produced in volume, the reliability of the produced item can differ from the design reliability because of assembly errors and component non-conformance. The reliability of produced items is the "inherent reliability" of the product.

2.4.2.3 Reliability at Sale

After production, the product must be transported to the market and is often stored for some time before it is sold. The reliability of an item at sale depends on the mechanical load (resulting from vibrations during transport) and impact load (resulting from mishandling) to which it has been subjected, the duration of storage and the storage environment (temperature, humidity, etc.). As a result, the reliability at sale can differ from the inherent reliability. Once an item is sold, it may either be stored for an additional time (if the item has been purchased for later use or is used as a spare), or it may be put into operation immediately. The additional storage time may again affect its reliability.

2.4.2.4 Field Reliability

The reliability performance of an item in operation depends on the length and environment of prior storage and on operational factors such as the usage intensity (which determines the load—electrical, mechanical, thermal and chemical—on the item), usage mode (whether used continuously or intermittently) and operating environment (temperature, humidity, vibration, pollution, etc.) and, in some instances, on the human operator. The reliability performance of an item in operation is often referred to as "field reliability".

2.5 Maintenance of Products and Plants

Maintenance consists of the different functions (or activities) necessary to keep a system in, or restoring it to, an acceptable state (or operating condition). Maintenance involves one or more of the following actions:

- Servicing
- Testing/Inspection
- Removal/Replacement
- Repair/Overhaul
- Modification.

Comment: In the literature, the term "MRO" is used extensively. It is acronym for the following actions:

M Maintenance (minor PM actions)
R Repair (CM actions)
O Overhaul (major PM action).

2.5.1 Corrective Maintenance

The failure of a system is due to the failure of one or more of its components. CM actions are actions to restore a failed system to operational state by rectification actions (repair or replace) on the failed components.

2.5.1.1 Classification of CM Actions

Let x denote the time of first failure. The different types of CM actions and their impact on system reliability are as follows:

Back to New: This involves the replacement of a failed item by a new one. As such the system reliability at time t is given by $R_1(t) = R_0(t - x)$ for $t > x$ where $R_0(\cdot)$ is the reliability of a new system. Note that this is appropriate for maintenance actions at the component level.

Minimal Repair: Here, the reliability of the item is unaffected by the maintenance action. As such, the reliability after repair is the same as that just before failure. This is an appropriate characterisation at the system level if the failure is due to one or few components and either repairing or replacing them has very little impact on the overall system reliability. In this case, the system reliability at time t is given by $R_2(t) = R_0(t)/R_0(x)$ for $t > x$.[11]

[11] This follows from simple argument based on conditional probability (see Appendix A).

Imperfect Repair: Here, the reliability of the item is affected by the repair action. One can define two types of imperfect repairs—case (i) and case (ii). In the former case, the reliability after repair is better than what was just before failure. This characterises the situation where the failure is a major failure requiring the replacement of several components by new ones so that the overall reliability improves. In this case, the system reliability at time t, $t > x$, is given by $R_3(t)$ with $R_2(t) < R_3(t) < R_1(t)$. In the latter case, the reliability after repair is lower than that just before failure. This usually is the effect of poor quality of repair that degrades the reliability so that the system reliability at time t, $t > x$, is given by $R_4(t)$with $R_4(t) < R_2(t)$.

Figure 2.9 shows the impact of the different types of CM actions on the reliability of the system after a failure.

2.5.1.2 Repair Time

In general, the time to carry out a CM action is uncertain and needs to be characterised as a random variable. If the variability in the repair time relative to the mean time to repair is small, then one can treat it as a deterministic quantity (the mean time to repair).[12]

2.5.1.3 Repair versus Replace

When a repairable item fails, there is an option to either repair or replace it by a new (or used) item. The optimal decision is usually based on cost considerations and the impact of the actions on future failures of the item involved.

2.5.2 Preventive Maintenance

PM actions are actions to control system degradation and reduce the likelihood of failure

2.5.2.1 Classification of PM Actions

The different types of PM actions are as follows:

Clock-based maintenance: Here, PM actions are carried out at set times.

[12] Typically, the time taken to repair or replace a failed item is often very much smaller than the time between failures (in a statistical sense) so that one can ignore repair times and treat the repairs as being instantaneous for the purpose of modelling of failures over time. This is discussed in Chap. 3.

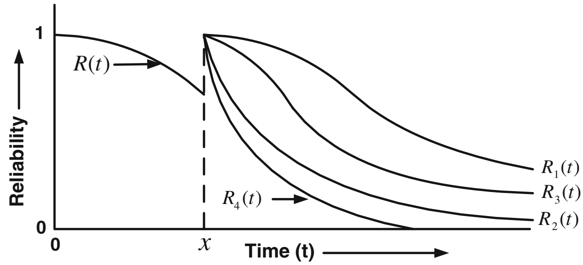


Fig. 2.9 Impact of different types of CM action on system reliability

Age-based maintenance (*ABM*): Here, PM actions for an item (component or higher level element) are based on the age of the item.[13]

Usage-based maintenance (*UBM*): Here, PM actions are based on total output (or usage) of the item since the last PM action.

Condition-based maintenance (*CBM*): Here, PM actions for an item are based on the condition of the item being maintained. This involves monitoring of one or more variables characterising the wear process (e.g. crack growth in a mechanical component).

Opportunity-based maintenance (*OBM*): This is applicable for multicomponent items, where maintenance actions (PM or CM) for a component provide an opportunity to carry out PM actions on one or more of the remaining components contained in the item.

Design-out maintenance (*DOM*): This involves carrying out modifications through re-design of one or more components so that the new components have better reliability characteristics.

Imperfect PM Actions: Here, the reliability characteristics improve after a PM action but not to as-good-as new and are similar to imperfect CM actions.

Overhaul (*Shutdown Maintenance*): In the case of complex products and plants, major overhaul involves dismantling the whole system and replacing components that have deteriorated significantly. The reliability characteristics improve significantly after an overhaul. However, the reliability of the system after overhaul decreases with the number of overhauls as indicated in Fig. 2.10.

Let $R_j(t)$ denote the system reliability after the jth overhaul with t denoting the time subsequent to the overhaul (one is using a local clock that is reset to zero after each overhaul).[14] Note that $j = 0$ corresponds to a new system. Then, we have the following:

- $R_j(t)$ is a decreasing function of t (the effect of degradation)
- $R_{j+1}(t) < R_j(t),\ j \geq 0$, implying that the reliability of a system subjected to $(j + 1)$ overhauls is inferior to that subjected to j overhauls.

[13] The first two types of PM are also referred to as Time-based maintenance (TBM).

[14] Here the subscripts refer to the number of times an item has been subjected to overhauls and should not be confused with the notation in Sect. 2.4.1 where it refers to reliability of different components.

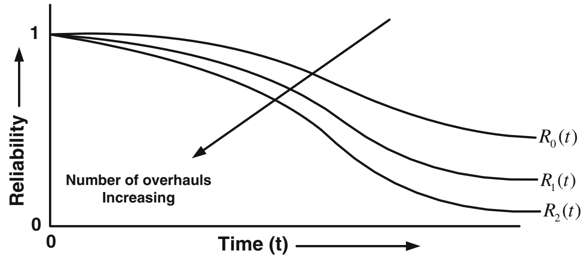


Fig. 2.10 Effect of overhauls on system reliability

2.5.2.2 Time for PM Action

In general, the time to carry out a PM action is predictable and as such it can be characterised as a deterministic quantity.

2.5.3 Maintenance Costs

Maintenance costs can be divided into two categories.

2.5.3.1 Direct Costs

The direct costs (which are viewed as part of the maintenance budget) are as follows:

- cost of manpower
- cost of material and spares
- cost of tools and equipment needed for carrying out maintenance actions
- overhead costs
- etc.

2.5.3.2 Indirect Costs

In addition, many other costs are affected either directly or indirectly by maintenance (or, more precisely, by lack of an effective maintenance policy). The costs involved depend on the nature of the business. In the case of a manufacturing operation, some of these costs are as follows:

- Equipment-related
  - accelerated wear because of poor maintenance
  - excessive spare parts inventory
  - unnecessary equipment redundancy
  - excessive energy consumption

- Production-related
  - rework
  - excessive scrap and material losses
  - idle operators due to breakdowns
  - delays in fulfilling orders
- Product-related
  - quality and reliability issues
- Customer-related
  - Customer dissatisfaction
  - Negative word-of-mouth effects.

2.5.4 Some Maintenance Policies

2.5.4.1 Maintenance of Products

Several different maintenance policies for products have been proposed and studied.[15] Examples of some of these policies are the following:

Policy 1 (*Age Policy*): Replace an item (PM action) by a new item when it reaches age v after being put into use or on failure (CM action) should the item fail earlier. This policy is characterised by the decision variable set $\Upsilon \equiv \{v\}$.

Policy 2 (*Block Policy*): Replace an item (PM action) by a new item at set times $t_k = kv$, $k = 1, 2, \ldots$. Failures between PM actions are rectified (CM action) by replacing the failed item by a new one. This policy is also characterised by the decision variable set $\Upsilon \equiv \{v\}$.

Policy 3 (*Periodic Policy*): Replace an item (PM action) by a new one at set times $t_k = kv$, $k = 1, 2, \ldots$. Failures between PM actions are minimally repaired (CM action). This policy is also characterised by the decision variable set $\Upsilon \equiv \{v\}$.

2.5.4.2 Maintenance of Plants

Policies for the maintenance of plants policies involve imperfect PM and overhauls with the above three policies used at the product or component levels. Two such policies are the following.

Policy 4 (*Imperfect PM*): The item is subjected to K imperfect PM actions before it is replaced by a new item. The time instants at which these actions are carried out are given by $\{t_k, 1 \le k \le K\}$ with $t_i < t_j$ for $i < j$. The reduction in the

[15] Nakagawa (2005) deals with the modelling and analysis of several maintenance policies.

failure hazard function during the kth PM action is δ_k. The item is replaced at time t_{K+1}. All failures between PM actions are rectified through minimal repair. This policy is characterised by the decision variable set $\Upsilon \equiv \{K; (t_k, \delta_k), 1 \le k \le K; t_{K+1}\}$.

Policy 5 (*Overhaul*): The item is subjected to the first overhaul after it has been in operation for a period t_0 and then is subsequently subjected to a sequence of overhauls. After the kth overhaul, the system is kept in operation for a period t_k after which it is subjected to an overhaul if $k < K$ or else is replaced by a new unit after being in operation for t_K after the last overhaul. All failures in between overhauls are repaired minimally. This policy is characterised by the decision variable set $\Upsilon \equiv \left\{K; t_j, 0 \le j \le K; \delta_j, 0 \le j \le K-1\right\}$.

2.5.5 Fleet Maintenance

There are several issues that need to be taken into account in the context of fleet maintenance.[16] Some of these are as follows:

- The age and the condition of units in a fleet can vary significantly so that the units are not statistically similar. The main reasons for this include (a) the units are purchased at different time points, (b) the usage of each unit can be quite different and hence the degradation levels of the units with the same age can be quite different, and (c) the ages of constituent components of a unit can be quite different due to the maintenance history. This raises an issue—how to control the "health level" of the fleet by appropriate maintenance and replacement decisions.
- Fleet maintenance needs to coordinate with production (or service) requirements and needs to take into account resource constraints.
- The failure consequence of a unit strongly depends on the configuration of a fleet and the functional requirements assigned to units within the fleet. This implies that the fleet maintenance needs to consider the priority of each unit and devise appropriate maintenance policies.
- The technological evolution of the unit makes maintenance options multidimensional—repair or replacement; if replacement, whether a particular unit should be replaced with a unit with same technology unit or one with more advanced technology. This implies that one needs to take into account technological evolution in the decision-making process when retiring (or replacing) old or degraded units.

Because of the multiunit nature, group and opportunistic maintenance are appropriate for fleet maintenance. Many different policies have been proposed and we discuss a few of them.

[16] For more information, see Cassady et al. (1998).

2.5.5.1 Group Maintenance Policies

One can define three types of group maintenance policies for a fleet.

Type I Policies: Here, the maintenance actions are based on age of the fleet. A group replacement is performed when the fleet reaches an age T.

Type II Policies: Here, the maintenance is based on the number of failed units. If the fleet is monitored continuously, then maintenance actions are initiated when the number of failed units reaches m. At this instant, all failed units are replaced with new ones (CM action) and all functioning units are serviced (PM action) so that they are restored to good-as-new. When the monitoring is not continuous, then the fleet is inspected at discrete time instants and maintenance actions are initiated only if the number of failed units is equal to or greater than m.

Policy 6 (Assaf and Shanthikumar 1987): The system is inspected at discrete time instants. Upon an inspection, the failed units are repaired if the number of failed units is greater than or equal to m; otherwise, they are left idle (failed state). The time to the next inspection is decided based on the number of failed units. The decision variables are m and the state-dependent inspection time instants.

Type III Policies: Here, maintenance action is based on both age and number of failed units.[17] The maintenance actions are initiated (for the continuous monitoring case) when the fleet reaches an age T or at the time instant when the number of failed units reaches m, whichever comes first. All failed units are replaced with new ones (CM action), and all functioning units are serviced (PM action) so that they become good-as-new.

Policy 7 (Park and Yoo 1993): The fleet consists of a group of identical units. Each unit is replaced on failure during the interval $(0, T)$. Beyond this interval, failed units are left idle until the number of failed units reaches a specified number m, when a block replacement is performed. The decision variables of the policy are T and m.

2.5.5.2 Opportunistic Maintenance Policies

Ritchken and Wilson (1990) deal with a fleet of machines in a production line and propose two types of opportunistic maintenance (Type I and II, respectively). In Type I opportunistic maintenance, CM action on a failed unit needs to be performed without any delay and PM actions on non-failed units can be advanced if appropriate and possible. In Type II opportunistic maintenance, failed units can be kept idle (failed state) for some amount of time so that one can postpone CM action to coincide with the first PM opportunity subsequent to the failure. A Type II policy involving only CM actions and two types of failures (minor and major or catastrophic) is the following:

[17] Type I and Type II policies are special cases of Type III policies.

Policy 8 (Sheu and Jhang 1997): The policy involves two intervals—[0, $T-w$] and [$T-w$, T]. Minor failures are rectified by minimal repairs at any time, and major failures are rectified immediately through replacements in the first interval and are not rectified in the second interval so that the failed units remain idle. Group maintenance is conducted at time T or when the number of failed units reaches $m(\leq n)$, where n is the total number of units whichever comes first. The decision variables of the policy are w, T and m.

Comment: In some cases, one or more units are cannibalised to provide spares for the other units.

2.6 Maintenance of Infrastructures

Maintenance of infrastructures include services such as clearing (snow, any object hindering the operation, etc. in the case of road and rail tracks) and cleaning (routine cleaning of buildings, rolling stock, vegetation growth on the sides of roads and rail tracks, etc.) and fixing (damaged road signs) for safe operations. These are referred to as service/operations and are different from PM and CM actions relating infrastructure per se. PM and CM are infrastructure specific, and we discuss these for road infrastructure. PM actions include inspection to monitor and assess the condition of the infrastructure. Based on the inspection results, e.g. the severity of the fault to traffic and the availability of resources, the decision is made to rectify the fault immediately or it is planned for a later stage considering all the risks to traffic and business, etc. Also, since failure is not so well defined, there is a blurring of PM and CM actions. In general, PM actions are those tasks that can be carried out in a short time period without too much interruption to the normal operation of the infrastructure. In contrast, CM actions take a longer time to complete (possibly running into months) and affect normal operations in a significant manner and are costly.

The main purpose of maintenance actions (PM and CM) is to control infrastructure degradation due to age, usage, load carried and other environmental factors, etc. and restore it to normal operating condition in the case of failure or other faults. In some industry sectors, OPEX (Operating Expenditure) denotes the expenditure associated with service/operations and PM actions and CAPEX (Capital Expenditure) the expenditure associated with CM actions and upgrades, etc. Major maintenance and investment involve a great deal of expenditure but have a direct influence on the financial and operational performance of the infrastructure.[18]

[18] In a 2009 report released by the American Association of State Highway and Transportation Officials (USA) about 50 % of the roads in the USA are in bad condition with urban areas worse.

Table 2.2 Work types for road maintenance

Work class	Work type	Work activity/operation
Routine maintenance	Routine pavement	Patching, edge-repair, crack sealing, spot re-gravelling, shoulders repair, etc.
	Drainage	Culvert repairs, clearing side drains
	Routine miscellaneous	Vegetation control, markings, signs
Periodic maintenance	Preventive treatment	Fog seal, rejuvenation
	Resurfacing	Surface dressing, slurry seal, cape seal, re-gravelling
	Rehabilitation	Overlay, mill and replace, inlay
	Reconstruction	Partial reconstruction, full pavement reconstruction
Special	Emergency	Clearing debris, repairing washout/ subsidence, traffic accident removal, etc.
	Winter	Snow removal, salting, gritting, etc.
Improvement	Widening	Partial widening, lane addition
	Realignment	Horizontal and vertical geometric improvements, junction improvement
	Off-carriageway	Shoulders addition, shoulders upgrading, side drain improvement, etc.
Construction	Upgrading	Upgrading by changing the surface class
	New section	Expanding of an existing section (with more lanes), new section (link)

2.6.1 Road Infrastructure

Pavements are designed for an expected service (design) life that can vary from 10 to 60 years, and for asphalt pavements, the typical life is 40 years. On each lane sector (for a multilane road), the initiating event for maintenance can be of two kinds:

1. End of the technical lifetime of the asphalt,
2. Economic depreciation of the road surface before its technical lifetime is over.

Maintenance is considered in the whole life cost of the road with CM actions at 10, 20 and 30 year milestones, and there is considerable freedom for maintenance planning with 15–20 possible actions per lane sector, from which the best choice has to be made. Maintenance (PM and CM) include many activities, and these are listed Table 2.2 [adapted from Archondo-Callao (2008)].

References

Archondo-Callao R (2008) Bituminous road deterioration, highway development and management (HDM-4). The World Bank, Washington

Assaf D, Shanthikumar JG (1987) Optimal group maintenance policies with continuous and periodic inspections. Manage Sci 33:1440–1452

Blischke WR, Murthy DNP (2000) Reliability: modeling, prediction and optimization. Wiley, New York

Cassady CR, Nachlas JA, Murdock WP (1998) Comprehensive fleet maintenance management. In: Proceedings of the IEEE international conference on systems, man and cybernetics, vol 5. IEEE, New York, pp 4665–4669

Dekker R, Plasmeijer R, Swart J (1998) Evaluation of a new maintenance concept for the preservation of highways. IMA J Math Appl Bus Ind 9:109–156

ISO 8402 (1986) Quality vocabulary. ISO, Geneva

Murthy DNP, Rausand M, Osteras T (2008) Product reliability: performance and specifications. Springer, New York

Nakagawa T (2005) Maintenance theory of reliability. Springer, London

Park KS, Yoo YK (1993) (τ, k) block replacement policy with idle count. IEEE Trans Reliab 42:561–565

Rausand M, Høyland A (2004) System reliability theory: models, statistical methods, and applications, 2nd edn. Wiley, Hoboken

Ritchken P, Wilson JG (1990) (m, T) group maintenance policies. Manage Sci 36:632–639

Rose G, Bennett D (1992) Locating and sizing road maintenance depots. Eur J Oper Res 63:151–163

Sheu SH, Jhang J (1997) A generalized group maintenance policy. Eur J Oper Res 96:232–247

Worm JM, van Harten A (1996) Model based decision support for planning of road maintenance. Reliability Eng Syst Safety 51:305–316

Springer

http://www.springer.com/978-1-4471-6439-5

Extended Warranties, Maintenance Service and Lease Contracts
Modeling and Analysis for Decision-Making
Murthy, D.N.P.; Jack, N.
2014, XVIII, 399 p. 101 illus., Hardcover
ISBN: 978-1-4471-6439-5