8. Maintenance Service Logistics

Joachim Arts[1], Rob Basten[2], and Geert-Jan van Houtum[2]

Abstract Capital goods, such as manufacturing equipment, trains, and industrial printers, are used in the primary processes of their users. Their availability is of key importance. To achieve high availability, maintenance is required throughout their long life cycles. Many different resources such as spare parts, service engineers and tools, are necessary to perform maintenance. In some cases, e.g. for trains, also maintenance facilities are required. Maintenance service logistics encompasses all processes that ensure that the resources required for maintenance are at the right place at the right time. In a broader sense, it also includes maintenance planning and design-for-maintenance. We first discuss capital goods and the requirements that their users have, which leads us to basic maintenance principles and the structure of typical service supply chains. Next, various relevant decisions and supporting theories and models are discussed. Finally, we discuss the latest developments within maintenance service logistics.

1 Maintenance of capital goods

In this chapter we discuss (maintenance) service logistics for capital goods. Examples of capital goods are manufacturing equipment, trains, industrial printers, radar systems, MRI-scanners, and baggage handling systems. Their availability is of key importance for the users of capital goods, since they are used in the primary processes of the users. Downtime of a capital good may lead to delayed or even lost production, or it leads to inferior service levels and reduced satisfaction at customers. For example, downtime of the bottleneck machines in a semiconductor factory leads to reduced output when the factory produces at maximal capacity (24/7) and then the downtime costs are in the order of tens of thousands of euros per hour (ASML, 2013). There also exist situations where downtime or malfunction

University of Luxembourg, Luxembourg, Luxembourg · Eindhoven University of Technology, Eindhoven, The Netherlands

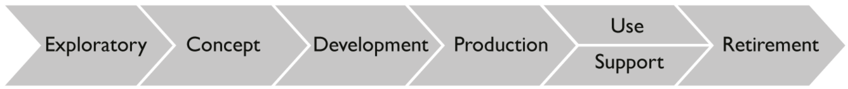

Fig. 1 Life cycle of a capital good (from International Organization for Standardization, 2008)

can lead to safety issues, for example in the case of radar systems on naval vessels or medical equipment in operating rooms.

Capital goods may have long life cycles of one to several decades. The life cycle typically follows the seven phases of the model of the International Organization for Standardization (2008), which is presented in Figure 1. In the first three phases, the needs of the users are identified, several ideas for solutions are explored and developed, and finally one is selected that is worked out to a complete design. Next, one or multiple units of the capital good are manufactured in the production phase. These units constitute the *installed base*. The use phase is the key phase, since the systems are then used for producing products or delivering services. The support phase, which encompasses the maintenance activities, is shown in parallel to this phase. Finally, the systems are removed from operation in the retirement phase.

The two key performance indicators of capital goods are its *availability* or overall equipment effectiveness on the one hand, and its *life cycle costs (LCC)* or *total costs of ownership (TCO)* on the other hand. The operational availability is the total time that a capital good is available to produce when required, divided by the total time that the capital good is required to produce. In case of 24/7 usage, this simplifies to the long run average number of available hours per day, divided by 24. When overall equipment effectiveness is used, which is especially popular in the process industry, it also incorporates that a system may not simply be available or not but may be producing at a lower rate: the overall equipment effectiveness is the achieved production capacity divided by the theoretical maximum production capacity. For instance, let us assume that a poultry processing line is used 16 hours a day and that it can process 10,000 chickens per hour when it is fully functioning. However, if currently it can only process 120,000 chickens a day, its overall equipment effectiveness is $120{,}000/(16 \times 10{,}000) = 75\%$.

The LCC are the total of all costs incurred during the seven phases of the life cycle for the whole installed base. This concept is typically used by original equipment manufacturers (OEMs). The TCO are the costs of a particular installed system as observed by the owner, and is closely related to LCC. Typically, TCO consists of acquisition costs at the end of the production phase, use and support costs, and costs (or benefits) related to the retirement of the system. However, more and more, equipment is not purchased but leased, and service contracts are performance based. In such cases, the total cost of ownership is built up differently. We discuss different service contract types in Section 4.2. Notice that a far-reaching service contract may also imply that the OEM (or system integrator) becomes fully responsible for achieving a certain availability of the capital good.

Both the LCC and the availability of a capital good are to a large extent determined by its design. As a result, we argue that service logistics in the *small sense* encompasses all processes that ensure that required resources for maintenance are at the right place at the right time. In the *broad sense*, it also includes the maintenance concept and design-for-maintenance, i.e., designing capital goods such that their maintenance will be easier and less costly. Let us denote the total downtime of a capital good in a year, say, by D^{t}. It consists of both planned and unplanned downtime, denoted by D^{p} and D^{u}, respectively. Planned downtime results from planned maintenance, which is scheduled P times per year, while unplanned downtime results from failures that happen F times per year on average (F is also called the failure frequency of the system). Unplanned downtime may consist of time to get a service engineer and possibly diagnosis tools at the system or to bring the system to a repair shop, diagnosis time to find the cause of a failure, waiting time for required spare parts and service tools, and the actual time to fix the problem. The diagnosis time and actual time to fix the problem is called the *active maintenance time*. The other time components constitute *waiting time*. Also planned downtime consists of actual maintenance time and waiting time. We denote the mean active maintenance time (mean waiting time) for planned and unplanned maintenance by M^{p} and M^{u} (W^{p} and W^{u}), respectively. We can then state:

$$D^{\mathrm{t}} = D^{\mathrm{p}} + D^{\mathrm{u}}, \tag{1}$$

where

$$\begin{aligned} D^{\mathrm{p}} &= P(W^{\mathrm{p}} + M^{\mathrm{p}}), \\ D^{\mathrm{u}} &= F(W^{\mathrm{u}} + M^{\mathrm{u}}). \end{aligned} \tag{2}$$

The failure frequency F is a result of the design (and production) and also the active maintenance times M^{p} and M^{u} are largely determined by the design of the capital good: some components are easily exchanged, while others are not. Design-for-maintenance can decrease F, M^{p}, and M^{u}. In order to further keep the active maintenance time low, typically failures are handled by a repair-by-replacement policy in which a failed part is replaced by a spare part. Planned maintenance reduces the failure rate F but increases P. One main advantage of replacing unplanned maintenance by planned maintenance is that it is typically executed much faster. One may be more efficient during the active maintenance time ($M^{\mathrm{p}} < M^{\mathrm{u}}$) and all required maintenance resources can be made available at the time that the planned maintenance starts ($W^{\mathrm{p}} \ll W^{\mathrm{u}}$). Another main advantage is that users of capital goods can adjust their production plans when maintenance is planned. Hence, having planned downtime is preferred over having unplanned downtime. Service logistics in the small sense can reduce the waiting times W^{p} and W^{u}.

Figure 2 gives an overview of maintenance strategies. Modificative maintenance concerns interchanging a component with a technologically superior component or a component with a lower failure rate (i.e., with longer times between failures). Often, this is not considered to be a maintenance strategy and we will also ignore it in the sequel. The other maintenance strategies are preventive and breakdown corrective maintenance. Under a breakdown corrective maintenance strategy, a part is

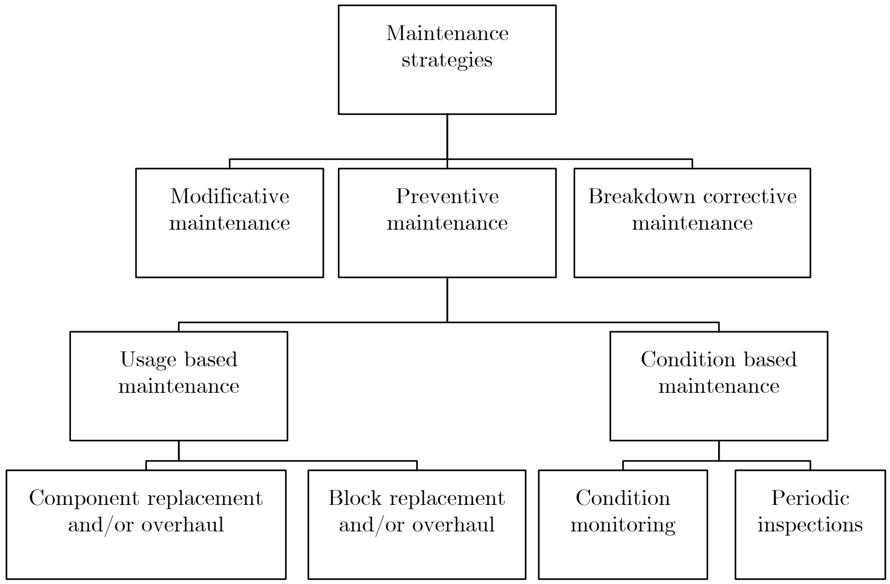

Fig. 2 Overview of maintenance strategies (from Arts, 2017)

replaced only after failure, while under a preventive maintenance strategy, the aim is to replace a part before failure. Of course, a part may still fail before it is replaced. Breakdown corrective maintenance is an attractive option for components that do not degrade or for which the degradation behavior is unclear, such as most electronic components. Such components generally have constant failure rates (also called hazard rates) and the distribution function of the time-to-failure can be modelled with an exponential distribution. This distribution is memoryless, which means that the probability of failure in the next period (a day, say), is independent of its age. This implies that replacing a functioning part by a new part does not bring any benefits.

For parts that do wear, it may be beneficial to follow a preventive maintenance strategy. These parts have a so-called increasing failure rate, meaning that over time, the probability of failure in the next period increases. The time-to-failure can then be modelled with a Weibull distribution with shape parameter $k > 1$, and it may then be worthwhile to replace a part by a new part after some amount of time or usage. There exist two preventive maintenance strategies: usage based maintenance and condition based maintenance (CBM; this type of maintenance is also called predictive maintenance). Under usage based maintenance, the total usage of a part is measured and maintenance is conducted when a certain threshold level has been reached. The usage of parts can be measured as time in the field (it is then also called age based maintenance), number of landings for an airplane, or number of sheets printed for an industrial printer, for example. Since the usage of the capital good is typically planned, also the moment that maintenance is to be performed can be planned. If there is a large set-up cost associated with maintenance, it may be beneficial to interchange several parts simultaneously, which leads to a block replacement and/or overhaul. Otherwise, maintenance can be performed on a single component. Barlow and Hunter (1960) developed the first usage based maintenance models, and new models are still being developed (see, e.g., Arts and Basten, 2016).

In CBM, the actual condition of a part is monitored, e.g., a vibration level, the number of metal particles in lubrication fluid, or a temperature. Depending on the nature of the capital good or the component, the condition can be measured either continuously through a sensor or periodically during inspections. The simplest form of CBM is to trigger maintenance when the measured condition exceeds a predetermined threshold, e.g., when the temperature rises above 90 degrees Celcius.

If there is a set-up cost associated with maintenance, for example because a service engineer needs to go to a customer or because the capital good needs to be shut down for maintenance, it may be beneficial to interchange several parts simultaneously (as in the block replacement policy discussed above).

2 Basics of maintenance service logistics (basic)

2.1 Service networks

Many different resources are required to perform maintenance, such as spare parts, service engineers, and tools. In some cases, e.g. for trains, also maintenance facilities are needed. To ensure that the required resources are at the right place at the right time, a service network or service supply chain is used, which consists of stocking locations for spare parts, repair shops, possibly maintenance facilities, and locations of service engineers. The involved costs are huge. All commercial airlines together are estimated to have over $40 billion worth of spare parts (Harrington, 2007). A single company such as ASML, which builds lithography equipment used in semiconductor manufacturing, owns spare parts worth tens of millions of euros (Kranenburg, 2006), and the US Coast Guard Aircraft owns inventories worth over $700 million (Deshpande et al, 2006). The demand intensity for spare parts changes during the life cycle of the installed base. Figure 3 shows this life cycle (not to be confused with the life cycle of one capital good, as shown in Figure 1).

We discuss the design of service networks in Section 3.1. We typically observe two types of service networks in practice, user networks and OEM networks. User networks are still encountered often, although maintenance activities are increasingly outsourced to OEMs. A good example is the military industry, where generally most maintenance is performed by military organizations themselves, despite the increased complexity of the equipment. This is partly due to the belief that a military organization should not be dependent on civilian support. Therefore, a typical example of a user network is the spare parts network of a military organization; the spare parts networks in the high-tech industry (e.g., computers, printing equip-

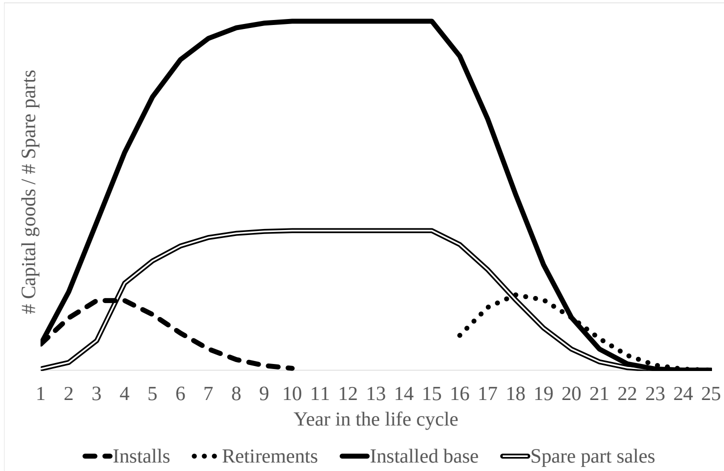

Fig. 3 Life cycle of the installed base

User networks (typical for military systems)	OEM networks (typical for high-tech systems)
Preventive maintenance dominates	Corrective maintenance dominates
Two echelon levels in one region	Global network with two echelon levels
No emergency option	Emergency option at highest echelon level
Mainly internal repair shops	External and internal repair shops
Relatively loose service targets	Strict/high service targets

Table 1 Characteristics of two network archetypes (cf. Basten and Van Houtum, 2014)

ment, medical systems) provide typical examples of OEM networks. Both types of networks often consist of two echelon levels, i.e., there are stocking locations at the local level and at a central level, but networks may consist of more echelon levels. We list the characteristics of both types of networks in Table 1 and show an archetypical example of a user network and an OEM network in Figure 4 and Figure 5, respectively. While these networks mainly represent the stocking locations for spare parts, the network is quite similar for expensive tools, and also for service engineers. Often, first line service engineers are located relatively close to the installed systems, while second line service engineers are located at the regional level. Third line service engineers are located at the main office or at the research and development department. (Sometimes, research and development itself is also involved in this third line).

In the rest of the chapter, we use the terminology that is common in an OEM network setting, and we focus our discussion on that setting. However, most of what we discuss is also applicable for user networks.

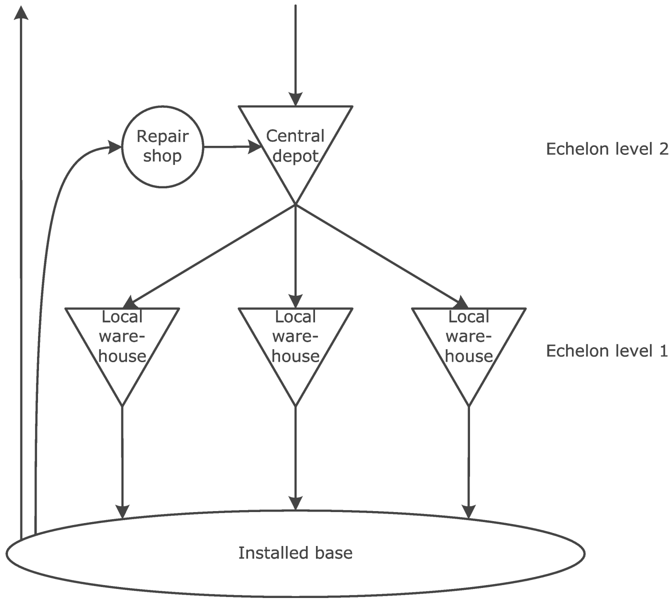

Fig. 4 Archetypical network of a user that maintains its own systems (cf. Basten and Van Houtum, 2014)

2.2 Case study: ASML

ASML is a Dutch equipment manufacturer in the semiconductor industry. ASML designs, supplies and integrates advanced technological equipment, called lithography machinery. The customers of ASML (such as Intel, AMD or Samsung) have billion dollar plants that house and exploit the equipment to produce semiconductors. The lithography step is one of many steps in the high tech production environment, but this step is crucial in terms of throughput and in terms of chip generation. In semiconductor plants, the ASML machines are the most capital intensive, which generally means that plants are developed such that ASML machines are the bottleneck of the process. Customers therefore generally strive for high occupation rates of the machines in their production processes. If an ASML machine breaks down, this can cause the whole production process at the customer to go down (cf. ESCF, 2012). As already stated above, downtime can result in costs in the order of tens of thousands of euros per hour.

ASML has service contracts with its customers in which an availability level of the machines is agreed upon. It is the responsibility of ASML to ensure that the machines sold are on average up and running for, for instance, 95% of the time. The downtime of the machines can be divided into downtime due to preventive maintenance (planned downtime) and downtime due to corrective maintenance (unplanned downtime). How the total yearly downtime is divided into planned and unplanned downtime can be decided by ASML. Internally at ASML, it has been decided to

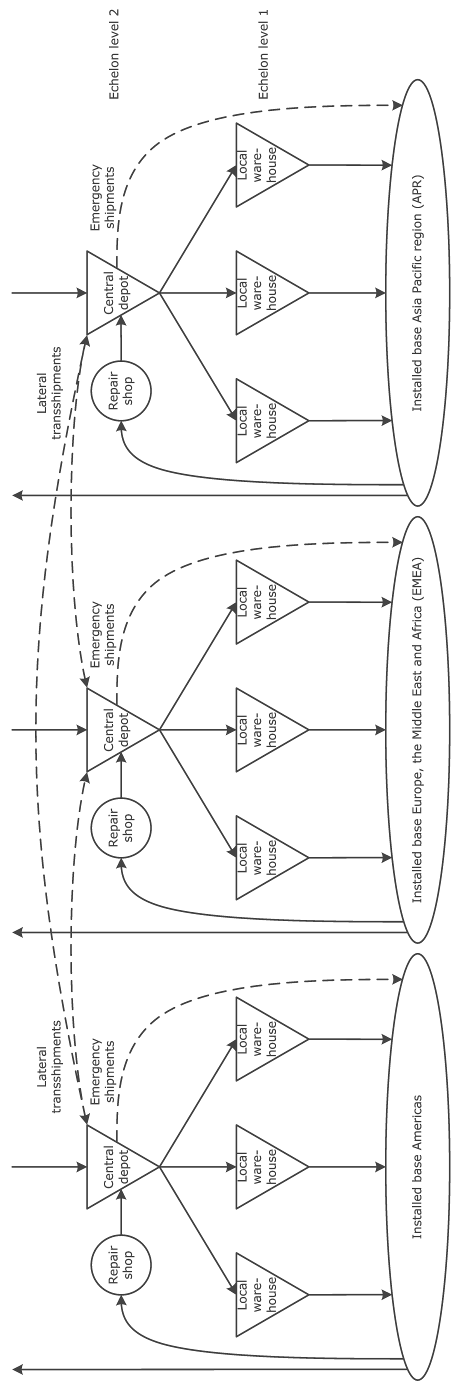

Fig. 5 Archetypical network of an OEM that maintains its sold systems (cf. Basten and Van Houtum, 2014)

divide the target for the total downtime into separate targets for planned downtime and unplanned downtime. The unplanned downtime is further divided into targets for the active mainenance time and waiting time.

Meeting the targets for unplanned downtime is most challenging. ASML has positioned spare parts, service engineers, and service tools in a global service network consisting of central and local customer service points. ASML has dozens of local support offices, including warehouses, around the world near its customers, and a few central warehouses (see Vliegen, 2009).

2.3 Decisions and decision levels

An OEM which provides maintenance services for systems installed at many different customers is faced with many different decisions at a strategic, tactical, and operational level. We describe the main decisions below, and an overview is displayed in Table 2. We discuss approaches and/or decision support models for different decisions and refer to sections within this chapter for more information where appropriate.

At the strategic level, the requirements regarding system availability should be taken into account in the product design (see also Section 4.2). For example, in order to meet very high system availability requirements, one can decide to install more reliable components, to build in redundant components, or to install extra sensors to enable predictive maintenance for certain components. The maintenance concept has to be decided immediately after, or, preferably, in parallel with the product design. This is also related to the portfolio of maintenance services that an OEM wants to offer, including the pricing of those services (Section 2.4). One needs clear standard service elements (regarding spare parts, service engineers, training of personnel of customers, technical documentation, and so on) with corresponding service processes of the OEM, and customers may be offered service contracts with varying combinations of service elements. An example of a standard service element is that the OEM guarantees spare parts deliveries within 1 week. Another service element may be that spare parts can be delivered within 24 hours in urgent situations, but one has to decide which customers can use that element. The OEM can decide that this is only possible for customers who have this element included in their service contract. Alternatively, the OEM may charge a higher price for this element for customers without a service contract. Another strategic decision concerns the locations of spare parts stocks and stocks of service tools and the division of the world in service areas with pools of service engineers (Section 3.1). The OEM also has to decide what procedures are in place when a part, tool or engineer is not available at the requested location. Procedures for emergency shipments of parts and tools are common in many service networks, but they need to be established before they can be used (to prevent ad hoc decisions as much as possible).

At the tactical level, the OEM has to decide on how many spare parts and service tools have to be kept on stock at each of the locations in the service network. When

Strategic	Tactical	Operational
Product design	Inventory levels spare parts	Allocation rules spare parts
Maintenance concept	Inventory levels service tools	Allocation rules service engineers
Portfolio maintenance services	Number of service engineers	Scheduling of service engineers
Design service process	Leadtimes repair loops	Priorities repair loops
Design service network	Batching decisions	Proactive movements of parts and tools
Control concept repair loops	Parameters of maintenance policies	Actions for specific customers

Table 2 Main decisions at the strategic, tactical, and operational decision level

demand levels are low, it is common to follow basestock policies, under which inventory positions are kept at constant levels. When demand levels are somewhat higher, it may be appropriate to include batching rules for replenishments of (new) parts at the central warehouse and possibly within the service network and for placing repair orders for failed parts. The OEM also has to decide on the number of service engineers per service region and the planned leadtimes for repairable spare parts that are repaired either at repair shops of the OEM itself or at external suppliers. While the maintenance concept has been decided at the strategic level, including the type of maintenance that is followed for the various components, the OEM or user can decide on the parameters of the maintenance policies at the tactical decision level.

At the operational level, an OEM first of all plans to follow the rules that have been determined at the tactical level. But, one may be faced with various types of deviations of what has been assumed there such as delayed deliveries of repaired spare parts by repair shops, exceptionally high demand rates for some parts, and long-term illness of multiple service engineers in the same service region. Further, at an operational level more detailed information is available which the OEM wants to take into account. Hence, allocation decisions have to be taken for spare parts and service tools, and, based on actual inventory levels, it has to be determined which parts should get priority in the repair loops. Further, scheduling of service engineers is arranged at that level. When condition information of critical components is monitored and one receives an indication of an upcoming failure of a specific component at a specific location in the world, it may be decided to move already a required spare part and possibly service tools to a place close to that specific location. If service contracts have been closed with many customers, for all of them, the OEM has to monitor continuously whether agreements specified in the contracts are still met. If, e.g. due to a temporary lack of resources, the OEM is not able to fulfill contract agreements in a timely fashion for a specific user, it may be decided to allocate additional resources close to that user's systems in the remaining contract period.

2.4 Service portfolio and service processes

Many different maintenance services are possible for capital goods. In service contracts, agreements may be specified regarding preventive, breakdown and modificative maintenance, and they can cover spare parts, service tools and service engineers. Contracts can also be closed for any subset of these elements. For each element, it has to be specified how quickly that element is provided and its price has to be determined. Different customers may ask for different speeds with which a service is delivered. An OEM has to carefully determine its service portfolio, in order to prevent too many different types of agreements with customers, which will lead to an unmanageable situation. For breakdown maintenance services for *large-scale computers*, service contracts are offered with response time targets of 2 hours, 4 hours, 8 hours, next business day, and second next business day. The response time starts when the customer calls the OEM to report a failure of the system. The call is received at a call center from where they will try to solve the problem remotely. If this is possible, the problem will be solved relatively quickly. If a failed part has to be replaced by a spare part, then that part will be delivered by a fast delivery from one of the local warehouses in the service network. The part has to arrive at the customer within the agreed response time. Some parts are replaced by the customer itself. Other parts have to be replaced by a service engineer. The customer agrees with the local service organisation when the service engineer will replace the required spare part. The customer has the right to demand the availability of a service engineer within the target response time, but he may also schedule the visit of the service engineer some time later. This depends on which part failed (for some parts, there is redundancy in the system, e.g. for hard disks) and on which processes are currently running.

In the above example, it becomes clear that response time is an appropriate *service measure* to base agreements on. Generally, which service measure is appropriate depends on the interaction between the customer and the OEM. A customer desires a smooth production or service process, and the OEM wants an efficient maintenance process. An important question is how strong the functioning of the system is related to the production or service process of the user. If failures of systems are strongly dependent on what production or service processes are executed on the system, on the utilization rate, or on the behavior of the operators, then this has to be taken account when defining agreements in service contracts. One may then decide to base agreements on response times (as in the above example) or on the total solution time per failure (consisting of waiting and repair time) rather than on system availability itself which also depends on the number of failures per time period (see (1)). We see different types of service measures in different industries. The use of response times and solution times is common for large-scale computers, printing systems and medical equipment. In the semiconductor industry, it is common to use the yearly or monthly downtime due to waiting for spare parts, which includes the effect of the number of failures.

3 Modeling and decision making (advanced)

3.1 Design of a service network

Consider an OEM with a service network of the type depicted in Figure 5. Suppose that the systems have to be maintained at the places where they are installed; this is common for e.g. manufacturing equipment. The OEM typically divides the world in regions that may coincide with one or more continents (e.g., Europe, Asia-Pacific, Americas). Per region, we have one central warehouse and multiple local warehouses for the spare parts and service tools (cf. Figure 5). Further, one has subregions (e.g., countries or parts of countries) with a pool of service engineers per subregion. These service engineers execute all first line maintenance services for the systems installed in their subregion. Service engineers may travel by car and may have a stock of spare parts in their car. In that case, these car stocks can be seen as local stockpoints as well. In several networks in practice, engineers have no car stocks, in which case a weaker coupling is obtained between the planning of spare parts and service tools on one hand and the planning of service engineers on the other hand.

For situations without car stocks and with tight targets for response or solution times, as in the large-scale computers example in Section 2.4, the choice of the locations for the central and local warehouses may be approached as a *facility location problem*. In this problem, customers are assumed to be distributed over a large area and they generate demands for spare parts and possibly service tools. There may be many candidate locations for the warehouses. Each customer has to be connected to a local warehouse that is at a sufficiently close distance to deliver parts/tools within x hours, where x is such that the highest level of service can be realized (we assume that this level has to be realized for many customers, but the set of customers with the highest service level may vary over time). Costs that are incorporated are fixed costs per local warehouse, the costs for the flows of the spare parts and the service tools, and inventory holding costs for spare parts and service tools. This leads to various types of mixed integer linear programming problems; see e.g. Candas and Kutanoglu (2007) and Rappold and Van Roo (2009).

Other related decisions that have to be taken are: (i) which components should be repaired and which should be discarded (repair or buy); (ii) where in the service network should parts be repaired; (iii) where to install manpower and equipment needed for these repairs? These decisions can be addressed by so-called *Level Of Repair Analysis (LORA) models*; see e.g. Basten et al (2015).

3.2 Forecasting

Planning service logistics inevitably starts with forecasting the demand for all resources needed for maintenance activities. Resources such as engineers and certain tools and parts are used often and then forecasting is not very different from traditional forecasting problems. However, many spare parts and also advanced service tools are used only sporadically. Then forecasting based on historic time series data (e.g., monthly demands) is difficult, because such data consists of many zeros. Such demand is also called *intermittent demand*. The usual forecasting techniques based on time series analysis such as exponential smoothing are ill-suited to deal with intermittent demand.

To deal with intermittent demand, time series forecasting methods have been developed that explicitly forecast the expected number of periods with zero demand. This forecast is then combined with a forecast for the demand quantity based on only periods with non-zero demand. These ideas have been pioneered by Croston (1972) and notable extensions have been made by Syntetos and Boylan (2005), who show how to reduce bias in this forecasting technique, and Teunter et al (2011), who show how to update the forecast for the time between positive demand in periods with zero demand. Machine learning algorithms such as neural networks have also been designed to forecast intermittent spare parts demand based on time series data as well as the more traditional bootstrapping and Box-Jenkins methods.

Another way to forecast intermittent demand for spare parts and service tools is to link the demand forecast to covariates such as planned maintenance and inspections. (Forecasting based on covariates is also called causal forecasting.) Rather than forecasting the demand directly, maintenance and inspections are forecasted. These forecasts are then linked with a forecast of the probability that such an inspection will lead to demand. These methods require knowledge on maintenance programs and planning, but yield more accurate forecasts. Wang and Syntetos (2011) and Romeijnders et al (2012) show how this idea works in different settings. For a situation with components that rarely fail in a small installed base, it is also useful to estimate the remaining useful lifetime of a component and construct a forecast based on that. Such remaining useful lifetime models are often technology specific, see e.g. Jardine et al (2006) for the prevalent technology of rotating machinery. Bian et al (2015) provide a Bayesian framework to update residual lifetime distributions based on measurement from sensors.

3.3 Inventory control of spare parts

As denoted in Figure 3, the demand of spare parts consists of three phases: a first, initial phase, a second phase with stationary demand, and a third, end-of-life phase. In all phases, the aim is to choose the inventory levels of spare parts such that agreements in service contracts can be met. These agreements are in terms of system availability, downtime or closely related measures. This leads to the formulation of multi-item inventory models in which relevant costs are minimized subject to constraints on or related to system availability. The relevant costs typically consist of inventory holding costs and transportation costs for replenishments, emergency shipments, and lateral transshipments. The system availability constraints may be

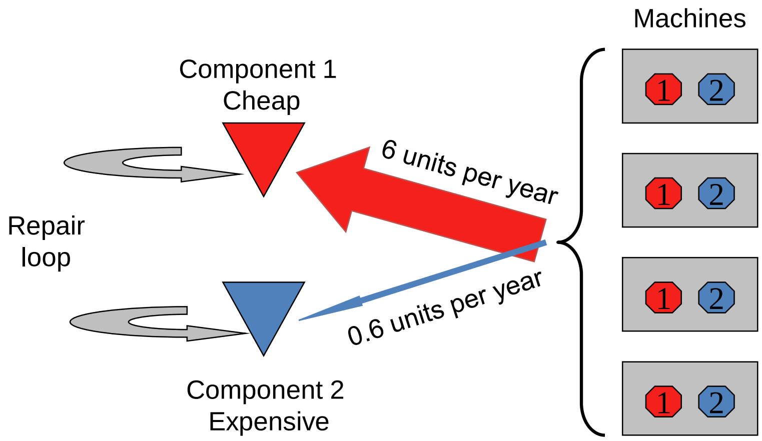

Fig. 6 Stocking of spare parts at a single production site

formulated for local warehouses and/or groups of installed machines. Solving these multi-item models requires a *system approach*, as introduced by Sherbrooke (1968).

Let us illustrate the main idea of the system approach by the following example. Consider a local stockpoint at a production site where multiple machines are installed. The machines have two critical components that are subject to failures. These components are called Component 1 and Component 2. For all machines together, Component 1 fails 6 times per year on average and Component 2 fails 0.6 times per year on average. These failures occur at random times (i.e., we assume that the underlying demand processes are Poisson). When a failure of a machine occurs, the component that has failed can be identified quickly. Next, the failed component is removed from the machine and replaced by a corresponding spare part. If this spare part is in stock, the replacement can be carried out immediately, within a very short time, and the downtime is negligible. In that case, the failed part is sent to a repair shop and will be returned to the local stock point as a ready-for-use part after 2 months (on average). If there is no spare part on stock, then the failed part is sent to the repair shop for an emergency repair, and the part will be returned as a ready-for-use part after 3 days (on average). In that case, the waiting time is 3 days, and we also have to pay an extra-high price for the repair at the repair shop. We assume that the machines are used for many years. Hence, we assume an infinite horizon and consider the performance in the long run, i.e., in steady state. Figure 6 shows a visualization of this example.

Let us assume that S_1 spare parts of Component 1 and S_2 spare parts of Component 2 are initially put in stock. Then, under the assumptions made above, we have a closed loop for the spare parts, and the stock on hand of Component i plus the number of parts in repair is always equal to S_i, $i = 1, 2$. The initial stock levels are our decision levels, and we seek to minimize the initial stock investment subject to an unplanned downtime constraint.

The price per new spare part of Component i is denoted by c_i, $i = 1,2$, where $c_1 = 100$ euros and $c_2 = 10,000$ euros. Let $\beta_i(S_i)$ be the probability that there is a ready-for-use part on stock when an arbitrary failure of a Component i occurs. The mean unplanned downtime per year (for all machines together) is then equal to (cf. (2))

$$D^{\mathrm{u}}(S_1,S_2) = 6\cdot(1-\beta_1(S_1))\cdot 3 + 0.6\cdot(1-\beta_2(S_2))\cdot 3 \quad \text{(in days)}.$$

Let us assume that this mean unplanned downtime per year may be at most 2.0 days. Then, our decision problem is as follows:

$$\begin{array}{lll} (\mathrm{P}) & \min & c_1S_1 + c_2S_2 \\ & \multicolumn{2}{l}{\text{subject to } D^{\mathrm{u}}(S_1,S_2) \le 2.0.} \end{array}$$

Notice that the downtime constraint ensures that the yearly number of fast repairs $(= 6\cdot(1-\beta_1(S_1)) + 0.6\cdot(1-\beta_2(S_2)))$ is limited and hence the extra costs for these repairs do not have to be included in the objective function. Further, the yearly mean number of repairs does not depend on the initial spare parts stocks and thus their costs are constant. As a result, the objective function only consists of the initial stock investment.

The performance of a given solution (S_1,S_2) is obtained by a steady-state analysis of the inventory system and the optimal solution can be obtained by enumeration for this small problem (see also Van Houtum and Kranenburg, 2015, Chapter 2). The optimal solution appears to be $(S_1^*,S_2^*) = (5,0)$ and the corresponding performance is:

$$\text{Initial stock investment} = 100\cdot 5 + 10,000\cdot 0 = 500 \text{ euros},$$
$$D^{\mathrm{u}}(S_1^*,S_2^*) = 6\cdot(1-\beta_1(S_1^*))\cdot 3 + 0.6\cdot(1-\beta_2(S_2^*))\cdot 3 = 1.86 \text{ days}.$$

So, this is the performance that we obtain by following a system approach.

Now suppose that, for some reason, we do not want to follow the system approach. In that case we model the unplanned downtime constraint at component level such that both initial stock levels can be optimized individually. Then the budget of 2.0 days for the unplanned downtime first has to be divided over the two components. We can do this proportionally to their demand rates, which gives a budget of $(6/6.6)\cdot 2.0 = 1.82$ days for Component 1 and a budget of $(0.6/6.6)\cdot 2.0 = 0.18$ days for Component 2. The problem to be solved for Component 1 is as follows:

$$\begin{array}{lll} (\mathrm{P}_1) & \min & c_1S_1 \\ & \multicolumn{2}{l}{\text{subject to } 6\cdot(1-\beta_1(S_1))\cdot 3 \le 1.82.} \end{array}$$

and a similar problem is obtained for Component 2. We then obtain $S_1^* = 3$ as optimal solution for Component 1 and $S_2^* = 1$ for Component 2. The resulting performance at machine level is then equal to:

$$\text{Initial stock investment} = 100 \cdot 3 + 10,000 \cdot 1 = 10,300 \text{ euros},$$
$$D^{\mathrm{u}}(S_1^*, S_2^*) = 6 \cdot (1 - \beta_1(S_1^*)) \cdot 3 + 0.6 \cdot (1 - \beta_2(S_2^*)) \cdot 3 = 1.62 \text{ days}.$$

This nicely meets the unplanned downtime constraint, but at an initial stock investment that is 20 times as high (10,300 vs. 500 euros)! This is purely due to the unnecessary decomposition of problem (P) into single-component problems. The system approach avoids such a decomposition and searches for a solution in a much larger solution space.

The theory on multi-item inventory models and the system approach are well developed. This theory is focussed on the second, stationary phase for spare parts, when generally reliable forecasts can be made for the demand rates and a steady-state analysis is appropriate. For inventory control in the first, initial phase and third, end-of-life phase, much less theory is available. In the initial phase, it may be difficult or even impossible to create reliable forecasts. Companies often use pragmatic approaches in that phase. Hardly any literature is available for this phase; for two recent contributions, see Martinetti et al (2017) and Van Wingerden et al (2017). In the third, end-of-life phase, also a separate approach is required. For many parts, it is not possible anymore to produce or buy new parts after some time point. At that moment, a last buy has to be made or one has to develop another sourcing possibility; see Pourakbar et al (2012) and Behfard et al (2015), and the references therein.

3.4 Repair loops

Repair-by-replacement significantly decreases the downtime of capital goods after a failure. Many replaced components are sufficiently expensive (and large) to warrant off-line repair. Such components are therefore called *repairables*. After being repaired, repairables are returned to stock as ready-for-use parts. Service logistics networks therefore consist of closed loops of repairable spare parts and repair shops where such parts are repaired or refurbished. Each repair shop in a service supply chain needs to be managed and this requires addressing the following three questions: (i) how many repair resources are needed; (ii) how should repairs of failed components be scheduled; (iii) how many lower level parts need to be stocked to facilitate the repair of components? We will discuss the main considerations and methods involved in making these decisions.

The sojourn time of a job in a repair shop consists mostly of waiting time for resources, such as technicians, tools, and lower level parts. The number of repair resources in a repair shop determines repair lead times that inventory control needs to account for. The number of resources such as technicians and tools in a repair shop can be dimensioned by using techniques from queueing theory. The fundamental machine-repair problem in queueing theory is a stylized model that was conceived for exactly this purpose. Taylor and Jackson (1954) designed this model for the engine repair shop of British Airways. Each repair shop will have specific characteristics and many variations to this queueing model have been developed for a variety of settings; see Koenigsberg (1982).

Scheduling in repair shops is done with the objective to avoid downtime of capital goods by preventing stockouts of repairable components. This is different from scheduling in production settings where the objective is to meet customer due dates or production efficiency targets. Repair shop scheduling rules should therefore take current stock levels of repairables into account, as well as shop characteristics and information. Hausman and Scudder (1982) test several scheduling rules in a simulation study and find that using smart scheduling rules yields a performance increase which allows a repairable item inventory reduction of 20% relative to a system with first come first service scheduling with equivalent performance.

One or more lower level parts are usually needed to complete the repair of a component. Stocking these lower level parts – which are also called Shop Replaceable Units (SRUs) – is therefore an important decision in repair shops. Muckstadt (1973) shows how inventory of SRUs and repairables can be controlled jointly when a component always breaks down due to exactly one SRU. In practice it is important to deal with the case where components need multiple SRUs to complete a repair. The inventory control of SRUs is then quite similar to the inventory control of assemble-to-order systems where it is also uncertain which parts will be needed to fulfill an order. Van Jaarsveld et al (2015) provide methods to minimize holding inventory of SRUs subject to the requirement that all parts needed for repair of an arbitrary component will be available with a high probability within a given time window.

3.5 Service tools

When a machine fails, one or multiple service tools may be needed to solve the failure. This includes diagnosis and calibration tools. If the number of different tools is low and their prices are low as well, then each service engineer may be given a kit with all relevant tools. However, service tools can also be expensive and/or there may be many different tools. In that case, one has a similar problem for service tools as for spare parts. One has to think of where to stock them and in which quantities. And when a machine fails, one aims to send the required service tools together with the spart part(s).

The planning of the service tools may be done independently from the planning of the spare parts. However, a machine failure can only be solved when both the required spare part(s) and service tool(s) are available. Hence, planning them in an integrated way is better, in particular when the same combinations of parts and tools are often needed.

Let us further elaborate on what is needed for an integrated planning. First of all, it requires that we have a very structured repair process. We need to have a list of all possible machine failures and a standard repair procedure per machine failure. The required part(s) and tool(s) per machine failure need to be known as well. Further, which machine failure occurred must be known already at the moment that the part(s) and tool(s) are sent to the failed machine. In addition, we need data on how often each machine failure occurs. If all these requirements are met, we can model coupled demands for parts and tools (as in assemble-to-order systems) in a multi-echelon network and this leads to complicated inventory problems. This integrated planning problem has been studied by Vliegen (2009). She studied this problem in a single-echelon, multi-location setting with lateral and emergency shipments (see Chapter 6 of her PhD thesis). She developed a heuristic solution procedure that takes the coupling in demands into account, but the computational complexity of this heuristic is relatively high. It is a great challenge to develop an efficient heuristic that works for instances of real-life size and that can be extended to multi-echelon networks with lateral and emergency shipments.

3.6 Service engineers

The maintenance of capital goods is executed by service engineers. Depending on the capital good, the service engineer travels to the capital good (e.g. manufacturing equipment) or the capital goods travel to the service engineer (e.g. aircrafts). In both cases, it is important to determine how many service engineers are needed so service all capital goods in a given area, and in the former situation, also the dispatching and routing of engineers to sites where capital goods need maintenance is important.

The number of service engineers that is available to respond to calls directly affects the waiting times W^{p} and W^{u} for maintenance in equation (1). Queueing theory can be used to determine the waiting time that results when a given number of service engineers need to handle all calls in a given area. Since this waiting time is also affected by the availability of spare parts, it is useful to consider stocking of spare parts and staffing of service engineers jointly; see Rahimi-Ghahroodi et al (2017). When engineers travel to capital goods, the travel times are also part of the waiting time. The waiting time due to travelling of service engineer can be affected by dispatching and routing decisions. However, dispatching and routing decisions also affect the effective available capacity of service engineers, because a service engineer is also "utilized" when traveling and the cumulative travel time that engineers experience is affected by routing and dispatch decisions. This interaction is studied in some more detail by Agnihothri and Karmarkar (1992).

Dispatching and routing decisions can also be affected by the availability of additional information through condition monitoring. The available condition information can be used to predict when and where a service engineer will be needed and this information in turn can be used advantageously in routing and dispatching; see Ichoua et al (2006). Dispatching and routing of emergency services such as ambulances have been leading in developing methods to support decision making in this area; see e.g. Maxwell et al (2010).

4 New developments in maintenance service logistics (state-of-the-art)

4.1 Technological developments

The developments in sensor technology enable that much more information on the degradation of components can be obtained against a low cost. In addition, the developments in communication technology enable that an OEM (or a third party) collects degradation data and other relevant data in one place in the world (or in the cloud). This gives a number of opportunities:

- Consider a component for which the prevalent failure mode is known and for which it is known how to measure the condition. By analyzing the data for many (failed) components, the OEM may be able to improve the modeling of the degradation process and hence the measurement of the condition. Or, the OEM may be able to improve the threshold that triggers maintenance.
- Consider a component for which the failure mode is not known or the failure mode is known in principle but it is not known how to measure the condition directly. Then data of many systems can be analyzed by regression or data mining techniques in order to construct failure predictors.
- For quite some components, it is known that their degradation process also depends on the environment or on the type of products that is being produced. This may be analyzed and can be incorporated in the failure prediction.

When generating failure predictions, it needs to be taken into account that predictions are only useful if there is a sufficient long time between the generation of the prediction and the time at which the failure will occur. Otherwise, one will be too late to replace a component preventively or take some other measure. For quite some components, a failure develops itself very quickly or even happens suddenly. In that case, CBM is not possible and one can stick to breakdown maintenance.

For some failure predictions, the OEM can be 100% sure that the failure will happen and there is a good prediction of the time at which the failure will occur. In that case preventive maintenance makes sense. If there is uncertainty about when the failure occurs, too much useful lifetime may be lost if a preventive maintenance action is applied to early. But the OEM can already ensure that the failure can be fixed quickly when the failure happens. E.g., if the failed component has to be replaced by a spare part, then the availability of a spare part can be assured (so that the waiting time is low; see (1)).

For failure predictions that are generated by regression analysis or data mining, the OEM has to make a tradeoff between false positives (a failure is predicted but will not occur) and false negatives (no prediction of a failure, but a failure does occur). Executing preventive maintenance based on such unreliable predictions may be unattractive. But, still such predictions can be used to initiate further investigations or to ensure that the failure can be fixed quickly when the failure does happen. This may be done by temporarily moving an extra spare part to a location. Topan et al

(2018) show that this can decrease costs significantly, even under a high percentage of false positives among the predictions. In these situations, high downtime costs due to waiting for spare parts are avoided by temporary placements of a required spare part in a local warehouse where normally that spare part is not kept on stock.

Another relevant technological development is additive manufacturing. Currently, many spare parts are kept on stock in local warehouse while demand rates are very low at those local levels. Nevertheless, the spare parts need to be kept on stock locally so that long waiting times are avoided when failures occur. It would be ideal if parts do not have to be kept on stock but if they can be quickly produced by additive manufacturing when a failure occurs. At this moment, we see multiple factors that limit the use of additive manufacturing: (i) only relatively simple components can be produced by additive manufacturing; (ii) before additive manufacturing can be applied, a costly design has to be made; (iii) the parts produced by additive manufacturing may have a lower reliability. When additive manufacturing develops further, these factors may become less limiting. In current research, it is investigated when additive manufacturing is a good option; see Knofius et al (2016), Song and Zhang (2017), and Westerweel et al (2016).

4.2 New business models for maintenance services

We already discussed maintenance services that can be offered by OEMs (or third parties) to users of capital goods. We see for many years already that more and more of the maintenance is taken over by OEMs. In the ultimate case, the OEM does not sell the system but offers the function of the system and the customers pay for the usage and the system availability. Rolls Royce does this already for a long time for engines of airplanes (Swartz, 2014). Rolls Royce offers so-called *power-by-the-hour contracts* under which Rolls Royce keeps the ownership of the engines and customers pay based on the number of hours that they use the engines. Other examples include ”pay per click” systems offered by professional copier manufacturers and cloud computing (e.g. IBM offers data center capacity). Obviously, also high system availabilities are guaranteed. This movement of offering the function instead of the function is known as *servitization*.

The movement to servitization for advanced capital goods has a number of advantages. The OEM (or a third party) will be responsible for the maintenance of many systems and will be much more efficient than maintenance departments of individual customers with respect to having well-trained service engineers, having specialists for the more complicated problems, and keeping spare parts and service tools on stock at a close distance of installed systems. Also they will be better able to develop advanced maintenance concepts, including CBM and the usage of data mining to generate failure predictions (this is only possible if you have data for many systems). Also, the OEM will be better able to realize a high system availability than an individual user. Further, under a power-by-the-hour contract, a customer knows directly what a system costs per unit of production or service that he produces. Another advantage is that an OEM is incentivized to design a more reliable product. When high system availabilities are required, the OEM may decide to use more reliable components or to build in extra sensors to facilitate CBM. An OEM may also think of tailored designs for different market segments with different system availability requirements. And, the OEM can use materials that can be re-used easily when installed systems retire.

Several of the advantages that one has under power-by-the-hour contracts hold also for *full service contracts* under which the customer still buys the system but the OEM is responsible for almost all maintenance activities and for realizing high system availabilities.

4.3 Control towers

Suppose that an OEM (or third party) closes power-by-the-hour and full service contracts with many customers and suppose that high system availabilities are required. Then using CBM or condition-based movements of spare parts will help to provide the required services in an efficient way. But this requires that decisions can be taken at the operational planning level based on the latest information. This may be realized by creating a *control tower*, like the control towers used in aviation, where the movements of all airplanes are followed and decisions are taken on the routes of airplanes, the order in which they can land at an airport, and so on. In a service logistics control tower, an OEM can bring multiple types of relevant information together: information on all service contracts, the performance for each contract until the current time point, the condition of critical components, the actual stocks of spare parts, possible problems with the supply of new spare parts, actual positions of service engineer, and so on. All this information needs be present in appropriate information systems (the current ERP systems will not suffice) and in a way that operational planners can quickly see where new problem arise. And, if so, the planner should be able to quickly decide what to do. They need simple rules or some decision support tool that enable them to decide quickly. The development of such control towers constitutes a great challenge, but will be key for success.

5 Further Reading

For various topics discussed in this chapter, we would like to suggest further readings:

- *Maintenance of capital goods*: For more literature on usage based maintenance models, we refer to any textbook on reliability engineering or maintenance optimization, e.g., Jardine and Tsang (2006), Pintelon and Van Puyvelde (2006), Ebeling (2001), and Arts (2017). For reviews on condition based maintenance, see Jardine et al (2006), Peng et al (2010), Prajapati et al (2012), and Olde Keizer

et al (2017). For a review on the topic of clustering of maintenance, see Nicolai and Dekker (2008); their review only includes one paper with CBM. Clustering of components that all use a CBM strategy is a recent topic in the literature, see, e.g., De Jonge et al (2016) or Olde Keizer et al (2016). Even more complicated is clustering of components that use different strategies, see e.g. Zhu (2015) and Arts and Basten (2016).
- *Service portfolio and service processes*: For different types of service contracts that are used in the industry, we refer to Cohen et al (2006) and Oliva and Kallenberg (2003).
- *Forecasting*: For forecasting methods for intermittent demand based on neural networks and several other machine learning and Box-Jenkins methods are studied by Lolli et al (2017), Moon (2013) and Pai and Lin (2006).
- *Inventory control of spare parts*: For textbooks on multi-item inventory models and the system approach, see Sherbrooke (2004), Muckstadt (2005), and Van Houtum and Kranenburg (2015). These books cover multi-echelon networks with a variety of properties that occur in practice, such as the use of emergency shipments, lateral transshipments, multiple customer classes, multiple machine types with commonality, and multiple indenture structures. Various models have been implemented in commercial software tools that are widely used in the industry, but there are also quite some companies with a tailor-made system-oriented solution (for more details, see Basten and Van Houtum, 2014, Section 8).
- *New business models for maintenance services*: For more information on servitization, see e.g. Avci et al (2015) and Agrawal and Bellos (2017), and the references therein.

Acknowledgements

The authors gratefully acknowledge the Netherlands Organisation for Scientific Research (NWO) and the Dutch Institute for Advanced Logistics (TKI Dinalog) for their support via the ProSeLoNext project and the NWO-TOP project on "Service Logistics for Advanced Capital Goods". Joachim Arts also gratefully acknowledges NWO for its support through his Veni grant.

References

Agnihothri S, Karmarkar U (1992) Performance evaluation of service territories. Operations Research 40(2):355–366

Agrawal VV, Bellos I (2017) The potential of servicizing as a green business model. Management Science 63(5):1545–1562

Arts J (2017) Maintenance modeling and optimization. BETA Working Paper 526

Arts JJ, Basten RJI (2016) Single-item models with minimal repair for multi-item maintenance optimization. BETA working paper 513

ASML (2013) Asml infographic about supply chain management. Video, `https://www.youtube.com/watch?v=dlKqO_L1rdc`

Avci B, Girotra K, Netessine S (2015) Electric vehicles with a battery switching station: Adoption and environmental impact. Management Science 61(4):772–794

Barlow R, Hunter L (1960) Optimum preventive maintenance policies. Operations Research 8(1):90–100

Basten RJI, Van Houtum GJ (2014) System-oriented inventory models for spare parts. Surveys in Operations Research and Management Science 19(1):34–55

Basten RJI, Van der Heijden MC, Schutten JMJ, Kutanoglu E (2015) An approximate approach for the joint problem of level of repair analysis and spare parts stocking. Annals of Operations Research 224(1):121–145

Behfard S, Van der Heijden MC, Al Hanbali A, Zijm WHM (2015) Last time buy and repair decisions for spare parts. European Journal of Operational Research 244(2):498–510

Bian L, Gebraeel N, Kharoufeh JaiIT (2015) Degradation modeling for real-time estimation of residual lifetimes in dynamic environments. IIE Transactions 47(5):471–486

Candas MF, Kutanoglu E (2007) Benefits of considering inventory in service parts logistics network design problems with time-based service constraints. IIE Transactions 39(2):159–176

Cohen MA, Agrawal N, Agrawal V (2006) Winning in the aftermarket. Harvard Business Review 84(5):129–138

Croston J (1972) Forecasting and stock control for intermittent demands. Operation 23(3):289–303

Deshpande V, Iyer AV, Cho R (2006) Efficient supply chain management at the U.S. Coast Guard using part-age dependent supply replenishment policies. Operations Research 54(6):1028–1040

Ebeling C (2001) Introduction to reliability and maintainablity engineering, 2nd edn. McGraw-Hill

ESCF (2012) Spare parts planning at asml. ESCF Operations Practices, `http://www.escf.nl/operation`

Harrington L (2007) From just in case to just in time. Air Transport World pp 77–80

Hausman W, Scudder G (1982) Priority scheduling rules for repairable inventory systems. Management Science 28(11):1215–1232

Ichoua S, Gendreau M, Potvin JY (2006) Exploiting knowledge about future demands for real-time vehicle dispatching. Transportation Science 40(2):211–225

International Organization for Standardization (2008) Systems and software engineering – System life cycle processes. ISO/IEC/IEEE 15288:2015

Jardine AKS, Tsang AHC (2006) Maintenance, Replacement, and Reliability. Theory and Applications. Dekker Mechanical Engineering, CRC Press, Boca Raton (FL)

Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing 20:1483–1510

De Jonge B, Klingenberg W, Teunter RH, Tinga T (2016) Reducing costs by clustering maintenance activities for multiple critical units. Reliability Engineering and System Safety 145:93–103

Knofius N, Van der Heijden M, Zijm W (2016) Selecting parts for additive manufacturing in service logistic. Journal of Manufacturing Technology Management 27:915–931

Koenigsberg E (1982) Twenty five years of cyclic queues and closed queue networks: A review. Journal of the Operational Research Society 33(7):605–619

Kranenburg B (2006) Spare parts inventory control under system availability constraints. PhD thesis, Eindhoven University of Technology

Lolli F, Gamberini R, Regattieri A, Balugani E, Gatos T, Gucci S (2017) Single-hidden layer neural networks forecasting intermittent demand. International Journal of Production Economics 183:116–128

Martinetti A, Braaksma AJJ, Ziggers J, Van Dongen LAM (2017) On the initial spare parts assortment for capital assets: A structured approach aiding initial spare parts assortment decision-making. In: Redding L, Roy R, Shaw A (eds) Advances in Through-Life Engineering Services, Springer, New York, pp 429–442

Maxwell MS, Restrepo M, Henderson SG, Topaloglu H (2010) Approximate dynamic programming for ambulance redeployment. INFORMS Journal on Computing 22(2):266–281

Moon S (2013) Predicting the performance of forecasting strategies for naval spare parts demand - a machine learning approach. Management Science and Financial Engineering 19(1):1–10

Muckstadt JA (1973) A model for a multi-item, multi-echelon, multi-indenture inventory system. Management Science 20(4):472–481

Muckstadt JA (2005) Analysis and Algorithms for Service Parts Supply Chains. Springer, New York

Nicolai RP, Dekker R (2008) Optimal maintenance of multi-component systems: A review. In: Kobbacy KAH, Murthy DNP (eds) Complex System Maintenance Handbook, Springer Series in Reliability Engineering, Springer, London (UK), pp 263–286

Olde Keizer MCA, Teunter RH, Veldman J (2016) Clustering condition-based maintenance for systems with redundancy and economic dependencies. European Journal of Operational Research 251:531–540

Olde Keizer MCA, Flapper SDP, Teunter RH (2017) Condition-based maintenance policies for systems with multiple dependent components: A review. European Journal of Operational Research 261(1):405–420

Oliva R, Kallenberg R (2003) Managing the transitions from products to services. International Journal of Service Industry Management 14:160–172

Pai P, Lin K (2006) Application of hybrid learning neural fuzzy systems in reliability prediction. Quality and Reliability Engineering International 22(2):199–211

Peng Y, Dong M, Zuo MJ (2010) Current status of machine prognostics in condition-based maintenance: a review. International Journal of Advanced Manufacturing Technology 50:297–313

Pintelon L, Van Puyvelde F (2006) Maintenance Decision Making. Acco, Leuven (Belgium)

Pourakbar M, Frenk JGB, Dekker R (2012) End-of-life inventory decisions for consumer electronics service parts. Production and Operation Management 21(5):889–906

Prajapati A, Bechtel J, Ganesan S (2012) Condition based maintenance: a survey. Journal of Quality in Maintenance Engineering 18(4):384–400

Rahimi-Ghahroodi S, Al Hanbali A, Zijm WHM, Van Ommeren JKW, Sleptchenko A (2017) Integrated planning of spare parts and service engineers with partial backlogging. OR Spectrum 39(3):711–748

Rappold JA, Van Roo BD (2009) Designing multi-echelon service parts networks with finite repair capacity. European Journal of Operational Research 199(3):781–792

Romeijnders W, Teunter R, van Jaarsveld W (2012) A two-step method for forecasting spare parts demand using information on component repairs. European Journal of Operational Research 220(2):386–393

Sherbrooke C (2004) Optimal Inventory Modeling of Systems: Multi-Echelon Techniques. Kluwer Academic, Boston/Dordtrecht/London

Sherbrooke CC (1968) Metric: A multi-echelon technique for recoverable item control. Operations Research 16(1):122–141

Song JS, Zhang Y (2017) Stock or print? impact of 3d printing on spare parts logistics. Working Paper

Swartz E (2014) The transformation of ge: From ”we bring good things to life” to industrial machines in the cloud. Blog, `http://www.metratech.com/blog/industrial-machines-in-the-cloud/` Last accessed: 7 May 2017.

Syntetos AA, Boylan JE (2005) The accuracy of intermittent demand estimates. International Journal of Forecasting 21(2):303–314

Taylor J, Jackson RRP (1954) An application ofthe the birth death process to the provision of spare machines. Operational Research Quarterly 5(4):95–108

Teunter R, Syntetos A, Babai M (2011) Intermittent demand: Linking forecasting to inventory obsolescence. European Journal of Operational Research 214(3):606–615

Topan E, Tan T, Van Houtum G, Dekker R (2018) Using imperfect advance demand information in lost-sales inventory systems. IISE Transactions

Van Houtum GJ, Kranenburg B (2015) Spare Parts Inventory Control under System Availability Constraints. International Series in OR & MS, Springer, New York

Van Jaarsveld W, Dollevoet T, Dekker R (2015) Improving spare parts inventory control at a repair shop. Omega 57:217–229

Van Wingerden E, Tan T, Van Houtum G (2017) Spare parts inventory control under uncertain demand rates. Working Paper

Vliegen I (2009) Integrated planning for service tools and spare for capital goods. PhD thesis, Eindhoven University of Technology

Wang W, Syntetos A (2011) Spare parts demand: Linking forecasting to equipment maintenance. Transportation Research Part E: Logistics and Transportation Review 47(6):1194–1209

Westerweel B, Basten R, Van Houtum G (2016) Traditional or additive manufacturing? assessing component design options through lifecycle cost analysis. BETA Working Paper 519

Zhu Q (2015) Maintenance optimization for multi-component systems under condition monitoring. PhD thesis, Eindhoven University of Technology