

COMPONENT REDESIGNS AND THE IMPACT OF THEIR IMPLEMENTATION POLICY

Steffi Neefs
Douniel Lamghari-Idrissi

ASML Netherlands B.V.
De Run 6501
Veldhoven, 5504DR, THE NETHERLANDS

Rob Basten
Geert-Jan van Houtum

School of Industrial Engineering
Eindhoven University of Technology
P.O. Box 513
Eindhoven, 5600MB, THE NETHERLANDS

ABSTRACT

An OEM who maintains a fleet of complex systems strives for high system availability for its customers. Frequently failing components lead to system unavailability and high maintenance costs. Consequently, the OEM might decide to upgrade components. We develop a model that quantifies the impact of the introduction of an upgraded component on the OEM's costs and number of failures to define the best implementation strategy. Using a Markov process, we evaluate four policies differing in the *roll-out strategy* of new parts, either immediate or corrective, and the *phase-out strategy* of old parts, either rework or salvage. The model is used in a case study at ASML. We conclude that, in the case study, reworking is preferred over salvaging as the phase-out strategy and corrective replacements are generally preferred over immediate replacements for the roll-out strategy.

1 INTRODUCTION

Lithography systems are high-tech systems and the most expensive assets used in the chip manufacturing process. Modern front-end wafer fabs are designed around them, making them the bottleneck station (Kopp and Mönch 2020). As a result, the average downtime costs of these systems are 72,000 euros per hour (ASML Holding N.V. 2014) and it can take up to 27 days for a chip manufacturer to recover from 24 hours of downtime (Lamghari-Idrissi et al. 2022). Due to the high impact of downtime on their operations, chip manufacturers often prefer to outsource the maintenance of lithography systems to the OEM. This provides an incentive to the OEM to redesign components with a high failure rate in order to improve their reliability (Schaefers et al. 2021), which is common practice for lithography systems (e.g., Yang et al. 2019; Van Schoot et al. 2020; Niimi et al. 2020; Mizoguchi et al. 2021). Inspired by the practice of ASML, an OEM selling lithography systems, our work focuses on the implementation strategy of these redesigned components. Such an implementation strategy consists of two parts: the roll-out strategy and the phase-out strategy.

The *roll-out strategy* concerns the timing of the upgrades. An OEM's roll-out strategy can be either corrective or immediate. With a corrective roll-out, the old parts are replaced as soon as they fail. With an immediate roll-out, the OEM upgrades the systems in the field immediately after the new part is available, i.e., preventively before the old part has failed. Although performance-based contracts are often used for complex systems (Kim et al. 2007), OEMs typically do not have authority over the maintenance strategy of the assets. Therefore, the OEM depends on the customers' willingness to replace components preventively.

The *phase-out strategy* concerns the phasing-out of the old components. The OEM may decide to either salvage or rework the old parts. When the part is salvaged, it is either scrapped or harvested. When the part is reworked, it is upgraded to the latest design.

In the literature, technological change and component obsolescence are the main reasons for studying the problem of part replacements. Most researchers study the decision of whether or not to replace the

old component in the system (e.g., Hartman 2000; Childress and Durango-Cohen 2005; Forootani et al. 2023). When the decision is made to replace the old component, the subsequent question is how the roll-out strategy and phase-out strategy affect the implementation of the new component. The literature on this topic is scarce. Mercier and Labeau (2004) and Mercier (2008) are one of the first to study the combination of a corrective and preventive roll-out strategy. Mercier and Labeau (2004) introduce the so-called K strategy, meaning that old units are correctively replaced by new ones until K old units have been replaced. The remaining old units are then preventively replaced. Mercier (2008) extends the work by Mercier and Labeau (2004) by incorporating general failure rates in order to better model degradation. However, phase-out strategies or customers’ willingness to adopt the new part are not incorporated in their models.

To the best of our knowledge, Öner et al. (2015) and Driessen (2018) are the only ones who take into account both the roll-out and phase-out strategy (to some extent). Öner et al. (2015) study two strategies to implement a new component: (1) Upgrade all systems preventively just after redesign, or (2) upgrade systems one-by-one correctively. They assume zero stock of old parts, and failed old parts are immediately salvaged. Driessen (2018) considers three different strategies to implement a new component: (1) Replace all old parts preventively, (2) produce all new parts at once before the start of the horizon and replace correctively, or (3) produce new parts in two batches and replace old parts correctively. He assumes the stock of old parts to be used only when there is a stock-out of new parts, and all failed old parts are salvaged. Both studies assume that when the choice is made for a preventive (immediate) roll-out strategy, all parts are replaced preventively at the beginning of the time horizon. They thus do not take into account the customers’ willingness to implement the new component in their installed base. In addition, both studies assume that all failed old parts are salvaged. They do not regard the potential to rework parts such that they can be put back on stock as an as-good-as-new new part.

In the literature, no methodologies are defined that help OEMs with making decisions on both the roll-out and phase-out strategy. To fill this gap, we define four distinct policies that differ in (1) roll-out strategy and (2) phase-out strategy (see Table 1). The first letter of the policy indicates the roll-out strategy, P(reventive) or C(orrective), and the second letter indicates the phase-out strategy, R(ework) or S(alvage).

Table 1: Four policies considered in this research.

Policy	PR	PS	CR	CS
Roll-out strategy	Immediate	Immediate	Corrective	Corrective
Phase-out strategy	Rework	Salvage	Rework	Salvage

The OEM’s choice for these strategies impacts its cost of service and reduces the number of (unscheduled) failures due to known failure modes compared to when the new part would not have been introduced. This can have major effects on the users’ utilization, throughput, cycle time, availability, and WIP, and it highly impacts their total cost of ownership. We develop a model to support the economic comparison of the strategies and we derive insights about the effect of various relevant parameters on this outcome.

The main contribution of this research is that we create a model to determine the impact of upgrading parts to improve reliability on the OEM’s cost of service and failures. We are the first to take into account the impact of customers’ acceptance of preventive replacements in this field of research. In addition, we include the possibility to rework old parts to the latest design. Sustainability has become more and more a crucial concern in the semiconductor industry (Shu et al. 2021) and many other industries. This research could aid with making chip manufacturing more sustainable by increasing system uptime and reusing materials instead of salvaging them.

Our paper is organized as follows. In Section 2, we introduce the model and assumptions. In Section 3, we formulate a finite horizon discrete-time Markov process. Section 4 presents a case study performed at ASML and provides the key managerial insights. We conclude in Section 5.

2 MODEL FORMULATION

In Table 1, we presented the four policies considered in this research. We now formulate the model that enables us to evaluate these policies. Table 2 introduces our notation and Figure 1 presents the graphical representation of our model.

Table 2: Notation.

$N \in \mathbb{N}$	Size of the installed base	$n \in \mathbb{N}$	Number of old parts operating at $t = 1$
$x \in \mathbb{N}$	Number of old parts on stock before $t = 1$	$s \in \mathbb{N}$	Base stock level
$p_r^o \in [0, 1]$	Prob. that old part can be reworked	$p_r^n \in [0, 1]$	Prob. that new part can be repaired
$\lambda^o \in \mathbb{R}^+$	Failure prob. of old part	$\lambda^n \in \mathbb{R}^+$	Failure prob. of new part
$\mu^o \in \mathbb{R}^+$	Rework prob. of old part	$\mu^n \in \mathbb{R}^+$	Rework prob. of new part
$h \in \mathbb{R}^+$	Holding costs of new parts	$b \in \mathbb{R}^+$	Penalty costs for emergency repair
$d \in \mathbb{R}^+$	Failure costs for one system	$f \in \mathbb{R}^+$	Costs to preventively replace a part
$z^o \in \mathbb{R}^+$	Costs to rework an old part	$z^n \in \mathbb{R}^+$	Costs to repair a new part
$w^o \in \mathbb{R}^-$	Salvage costs of old parts	$w^n \in \mathbb{R}^-$	Salvage costs of new parts
$p^n \in \mathbb{R}^+$	Cost price of new parts	$T \in \mathbb{N}$	Modeling horizon

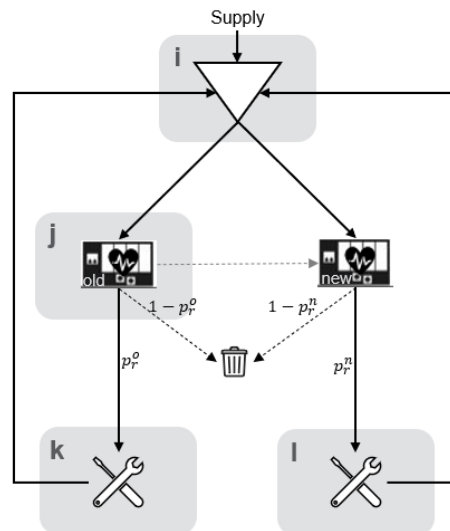


Figure 1: Graphical representation of the model.

We consider an installed base of N systems. Each system has a single unit of the part, meaning there is no redundancy. We model over a finite horizon of T periods, which we refer to as the modeling horizon. The modeling periods are numbered $t \in \{1, \dots, T\}$. Considering the long lifetimes of capital intensive systems, we assume that the installed base has a remaining lifetime that is equally long or even longer than the time horizon of T periods. In the case of a corrective roll-out strategy, all N systems contain an old part at the start of the horizon. In the case of an immediate roll-out strategy, we assume that $N - n$ customers accept the preventive replacement before the start of the model horizon. The remaining n old operating parts will be replaced correctly. To preventively replace a part, a cost f is incurred.

We assume that all old parts on stock are reworked or salvaged before the start of the horizon. For the policies with an immediate roll-out strategy, the preventively replaced parts will also be reworked or salvaged before the start of the horizon. This is a reasonable assumption when the failure rate of the part is low, which is often the case for capital intensive systems (van Houtum and Kranenburg 2015). This leaves us with zero stock of old parts at the start of the horizon.

All parts fail independently, and the failure time for each old part is geometrically distributed with failure probability λ^o per period. The failure time of each new part is also geometrically distributed, but with failure probability λ^n per period. We assume that $\lambda^n \leq \lambda^o$. When an old or new part fails, a failure cost d is incurred. Then, as soon as an old/new part fails in the field, it is replaced by a new part from stock.

If the failed part is a new part, the part is inspected after its failure. With probability $0 \leq p_r^n \leq 1$, the part can be repaired and is sent to the repair center for repair. With probability $(1 - p_r^n)$ this part cannot be repaired and is salvaged against cost w^n . The repair lead time of a new part is geometrically distributed with repair probability μ^n per period. Once a part is repaired in the repair center, the part is sent back to the local warehouse. To repair the new part, repair costs z^n are incurred, which includes the costs for shipment.

If the failed part is an old part, it is reworked or salvaged, depending on the phase-out strategy. If the phase-out strategy is to rework the part, with probability $0 \leq p_r^o \leq 1$, the part can be reworked and is sent to the repair center for rework. With probability $(1 - p_r^o)$ the part cannot be reworked and will be salvaged against cost w^o . The rework lead time of an old part is geometrically distributed with rework probability μ^o per period. As soon as a part is reworked in the repair center, the part is sent back to the stock point. To rework the old part, rework costs z^o are incurred, which includes the costs for shipment. All reworked and repaired parts are considered as-good-as-new. If the phase-out strategy is to salvage the old parts, that will be done so immediately against cost w^o . Note that this results in a simplified version of the model where $p_r^o = 0$, i.e., the probability that an old part can be reworked is zero.

As soon as an old or new part fails, it is replaced by a ready-for-use part from stock. The stock point follows a continuous review $(s - 1, s)$ inventory policy. We assume that at the beginning of the horizon, there are s ready-for-use parts on stock. Depending on the policy, this consists of new parts on stock, either bought new (policies PS and CS) or reworked old parts of which there are x on stock before the start of the horizon (policies PR and CR). As soon as the inventory position is equal to level $s - 1$, the stock is replenished with one item to level s . This inventory policy is often used for expensive, slow-moving spare parts (van Houtum and Kranenburg 2015). Every time an order for a new part is placed, i.e., if the inventory position is $s - 1$, a price p^n per part is paid. Note that in our model, the inventory position equals the on-hand inventory at the stock point plus the amount of parts in the repair center (either for repair or for rework). For simplicity, we assume that the replenishment lead time for new parts is 0. If the on-hand inventory at the stock point is 0, due to many parts being in repair or rework, an emergency repair is requested when a part fails. For every emergency repair a penalty b is incurred on top of the normal rework and repair costs.

3 EVALUATION APPROACH

Costs are incurred for the actions occurring prior to the model horizon. We outline these costs in Section 3.1. For each of the four policies, we calculate the total expected costs over T periods, using a discount factor $0 \leq \alpha \leq 1$. We assume that at most one event occurs in a period t : either the failure or the repair or rework of a part. This assumption is reasonable if the length of a period is sufficiently short (say a day), and if the failure rate of parts is low, which is often the case in industries using capital intensive systems (Driessen 2018). We can model the problem as a discounted, finite horizon, discrete-time Markov process. We present the Markov process that describes the operation of the system in Section 3.2 and the evaluation of the total expected failures and costs in Sections 3.3 and 3.4, respectively.

3.1 Pre-horizon Costs

For all $N - n$ preventive replacements a cost f is incurred. We assume that these $N - n$ parts are then either reworked (policy PR and CR) or salvaged, thereby leading to value (policy PS and CS). To calculate how many new parts we need to buy at the start of the horizon, we need to know how many old parts we

have on stock, which we denote as x . If $x + N - n \leq s$, all $x + N - n$ parts are reworked or salvaged, and $(s - (x + N - n))$ parts are procured at price p^n . Note that $N = n$ for the policies following a corrective roll-out strategy. When the phase-out strategy is to rework old parts (policy PR) and the current stock of old parts plus the preventively replaced parts is already higher than the intended basestock level, i.e., $x + N - n > s$, then s parts are reworked and the remainder is salvaged. In this case, no parts need to be procured. Table 3 presents an overview of the pre-horizon costs V_0 for all policies.

Table 3: Pre-horizon costs of the model.

Policy	Pre-horizon costs
Policy PR	$V_0 = (N - n) \cdot f + (x + N - n) \cdot z^0 + (s - x - N + n) \cdot p^n$, if $x + N - n \leq s$ $V_0 = (N - n) \cdot f + s \cdot z^0 - (x + N - n - s) \cdot w^0$, if $x + N - n > s$
Policy PS	$V_0 = (N - n) \cdot f - (x + N - n) \cdot w^0 + s \cdot p^n$
Policy CR	$V_0 = x \cdot z^0 + (s - x) \cdot p^n$, if $x \leq s$ $V_0 = s \cdot z^0 - (x - s) \cdot w^0$, if $x > s$
Policy CS	$V_0 = s \cdot p^n - x \cdot w^0$

3.2 Discrete-Time Markov Process

The sequence of events in a period t is given as follows.

1. The current state is observed.
2. Holding costs h are paid for all parts that are currently on stock.
3. The inventory position is observed. If the inventory position is $s - 1$, then a replenishment order is placed. A new part is ordered for price p^n per part. The ordered part will arrive at the beginning of the next period.
4. An event may occur, i.e., at most one by assumption:
 - (a) If a failure occurs, the OEM incurs a failure cost d and replaces the failed part by a new part from stock. If there is no stock on-hand to replace the failed part (all parts are at rework or repair), the OEM will perform an emergency repair against penalty cost b . In this case, a failed old (new) part at the repair center is urgently reworked (repaired) to replace the failed part. With probability $1 - p_r^0$ ($1 - p_r^n$), the failed part cannot be reworked (repaired) and will be salvaged. The OEM then incurs the expected discounted salvage costs w^0 (w^n) for this. With probability p_r^0 (p_r^n) the failed old (new) part can be reworked (repaired), and is subsequently sent to the repair center.
 - (b) If the rework (repair) of an old (new) part finishes, the OEM incurs rework (repair) costs z^0 (z^n). The part is then sent to the stock point. This ready-for-use part can only be used to replace a failed part in the next period.
 - (c) If no failure, repair or rework occurs, no additional costs are incurred.
5. The period ends and we transition to the next state.

The state space of the model is four-dimensional: (1) the number of operating old parts N^0 (then the number of operating new parts is $N^n = N - N^0$), (2) the number of old parts being reworked X^0 , (3) the number of new parts in repair X^n , and (4) the number of ready-for-use parts on stock S^n . Policies PS and CS consider a phase-out strategy where old parts are salvaged. For these policies, the probability that parts are reworked after failure is equal to $p_r^0 = 0$, implying $X^0 = 0$.

The number of operating old parts N^0 can never be higher than n . Consequently, the number of parts in rework X^0 can never be higher than n . Parts that finished repair and rework are sent to the inventory. To avoid explosion of the state space, we bound the number of parts in the inventory S^n to be at most equal to the size of the installed base N . This assumption is reasonable as the failure and repair rates are such that the optimum will always be less, which can be checked after optimization. In

addition, we assume emergency repairs. Therefore, the number of parts in inventory is at least equal to 0. Similar reasoning can be applied for the parts in repair X^n . The state space is defined by $\mathcal{S} = \{(S^n, N^o, X^o, X^n) : 0 \leq S^n \leq N, 0 \leq N^o \leq n, 0 \leq X^o \leq n, 0 \leq X^n \leq N\}$. We refer to a state $\mathfrak{s} = (i, j, k, l) \in \mathcal{S}$, where i, j, k, l correspond to S^n, N^o, X^o, X^n respectively. $(s, n, 0, 0)$ Note that the inventory position $(S^n + X^o + X^n)$ cannot become lower than $s - 1$, because a new part is purchased if the inventory position has dropped to $s - 1$. It can become higher than s , e.g. if there are more than s parts in repair or rework. At the beginning of the modeling horizon, there are s parts in inventory, n old parts operating, and 0 parts in rework or repair. Therefore, we define the starting state as $\mathfrak{s}_{\text{start}} = (s, n, 0, 0)$.

We consider the one step transition probability from state $\mathfrak{s} = (i, j, k, l) \in \mathcal{S}$ to state $\mathfrak{s}' \in \mathcal{S}$. Let $p(\mathfrak{s}' | \mathfrak{s})$ denote this one step transition probability. In Table 4, we present these transition probabilities for the events that they are triggered by. We present the entire version of the one step transition probability $p(\mathfrak{s}' | \mathfrak{s})$ for each $\mathfrak{s}, \mathfrak{s}' \in \mathcal{S}$ in Table 9 in Appendix A. Using value iteration, we can find the mean values for the number of failures and costs, as outlined in Sections 3.3 and 3.4, respectively.

Table 4: One step transition probabilities.

$\mathbf{p}(\mathfrak{s}' \mathfrak{s})$	Event
$j\lambda^o p_r^o$	An old part fails and can be reworked
$j\lambda^o(1 - p_r^o)$	An old part fails and cannot be reworked
$(N - j)\lambda^n p_r^n$	A new part fails and can be repaired
$(N - j)\lambda^n(1 - p_r^n)$	A new part fails and cannot be repaired
$k\mu^o$	An old part is reworked
$l\mu^n$	A new part is reworked
$1 - \sum_{\hat{\mathfrak{s}} \in \mathcal{S}: \hat{\mathfrak{s}} \neq \mathfrak{s}'} p(\hat{\mathfrak{s}} \mathfrak{s})$	No failure, rework or repair occurs

Note that the transition probabilities depend on the state at the moment of the transition. For example, the inventory position in state \mathfrak{s} determines whether a new part will arrive in on-hand inventory in state \mathfrak{s}' . Moreover, the on-hand inventory in state \mathfrak{s} determines whether a failed part can be replaced by a new part from inventory or through an emergency shipment.

3.3 Failure Calculations

The expected number of failures in state $\mathfrak{s} \in \mathcal{S}$ is $Y(\mathfrak{s}) = j\lambda^o + (N - j)\lambda^n$. Notice that we have a finite horizon problem, meaning that we can easily calculate the recursive formulas. Let $Q_t(\mathfrak{s})$ be the total expected number of failures over periods $t, t + 1, \dots, T$, given that we are in state $\mathfrak{s} \in \mathcal{S}$ at the start of period t . The expected number of failures over the full time horizon for the OEM can then be expressed as $Q = Q_1(\mathfrak{s}_{\text{start}})$. The recursion for $Q_t(\mathfrak{s})$ for all periods $1 \leq t \leq T$ is given by

$$Q_t(\mathfrak{s}) = Y(\mathfrak{s}) + \sum_{\mathfrak{s}' \in \mathcal{S}} p(\mathfrak{s}' | \mathfrak{s}) Q_{t+1}(\mathfrak{s}') \quad \forall \mathfrak{s} \in \mathcal{S} \quad (1)$$

3.4 Cost Calculations

The expected cost $\sigma(\mathfrak{s})$ incurred in one period while in state $\mathfrak{s} \in \mathcal{S}$ consists of inventory holding costs $H(\mathfrak{s})$, procurement costs $P(\mathfrak{s})$, repair, rework and failure costs $F(\mathfrak{s})$, emergency repair penalty costs $E(\mathfrak{s})$, and the salvage costs $S(\mathfrak{s})$:

$$\sigma(\mathfrak{s}) = H(\mathfrak{s}) + P(\mathfrak{s}) + F(\mathfrak{s}) + E(\mathfrak{s}) + S(\mathfrak{s}).$$

We next explain each of these cost elements.

In each state $\mathfrak{s} \in \mathcal{S}$, inventory holding costs $H(\mathfrak{s}) = hi$ have to be paid.

If in state $\mathfrak{s} \in \mathcal{S}$ the inventory position is equal to $s - 1$, a new part needs to be procured, incurring procurement costs p^n . Therefore,

$$P(\mathfrak{s}) = \begin{cases} p^n & \text{if } i+k+l = s-1 \\ 0 & \text{otherwise.} \end{cases}$$

Rework (Repair) costs z^o (z^n) are paid when a part is successfully reworked (repaired) and failure costs d are paid when a failure occurs. The probabilities that we observe a rework, repair or failure of an old part, or a failure of a new part are $k\mu^o$, $l\mu^n$, $j\lambda^o$, and $(N-j)\lambda^n$, respectively. Therefore, the expected repair, rework and failure costs in state $\mathfrak{s} \in \mathcal{S}$ are

$$F(\mathfrak{s}) = z^o k\mu^o + z^n l\mu^n + d(j\lambda^o + (N-j)\lambda^n)$$

Emergency repair costs b are paid when there is a failure during stockout. We can therefore express the expected emergency repair penalty costs when being in state $\mathfrak{s} \in \mathcal{S}$ as

$$E(\mathfrak{s}) = \begin{cases} b(j\lambda^o + (N-j)\lambda^n) & \text{if } i = 0 \\ 0 & \text{otherwise.} \end{cases}$$

A salvage cost w^o (w^n) is incurred when the state transition is triggered by the failure of an old or new part that cannot be repaired. The probability of this transition is $j\lambda^o(1-p_r^o)$ and $(N-j)\lambda^n(1-p_r^n)$ respectively. The expected salvage costs when being in state $\mathfrak{s} \in \mathcal{S}$ is thus

$$S(\mathfrak{s}) = w^o(j\lambda^o(1-p_r^o)) + w^n((N-j)\lambda^n(1-p_r^n)) \quad \forall \mathfrak{s} \in \mathcal{S}$$

Let $V_t(\mathfrak{s})$ be the total expected discounted costs over periods $t, t+1, \dots, T$, given that we are in state $\mathfrak{s} \in \mathcal{S}$ in the beginning period t . The total cost for the OEM can then be expressed as $V = V_0 + V_1(\mathfrak{s}_{\text{start}})$. The recursion for $V_t(\mathfrak{s})$ for all periods $1 \leq t \leq T$ is given by

$$V_t(\mathfrak{s}) = \sigma(\mathfrak{s}) + \alpha \sum_{\mathfrak{s}' \in \mathcal{S}} p(\mathfrak{s}' | \mathfrak{s}) V_{t+1}(\mathfrak{s}') \quad \forall \mathfrak{s} \in \mathcal{S} \quad (2)$$

We assume $V_{T+1}(\mathfrak{s}) = 0$ for all $\mathfrak{s} \in \mathcal{S}$.

4 RESULTS

Some years ago, ASML introduced an upgraded component in the field. We use the parameters in this case as the base case in our case study. Next, we perform a sensitivity analysis where, each time, we vary one parameter compared to the base case. The *normalized* parameters of the base case and the sensitivity analysis are outlined in Section 4.1. Sections 4.2 and 4.3 elaborate on the results of the base case and the sensitivity analysis, respectively.

4.1 Base Case and Sensitivity Analysis Parameters

In Table 5, we summarize the values of all parameters. ASML had a worldwide installed base of 4,500 systems in 2016 (Oomen 2016). Due to computation time constraints, we set the installed base size equal to $N = 100$, and assume that every customer operates one system. On average, 14 out of 100 customers accepted a preventive replacement. This means that for policies with an immediate roll-out strategy, $N - n = 14$ parts are replaced preventively before the start of the horizon and we have $n = 86$ old operating parts at the start of the horizon. The failure cost is $d = 3,839$, which is a cost based on the mean time to get the system up again after a failure of this part. A preventive replacement costs approximately 70% of this failure cost. The cost price of the old and new parts are $p^o = 8,825$ and $p^n = 10,000$, respectively. We assume that a holding cost of 17% of the cost price is paid per part per year, i.e. $h = 4.66$ per part per day. Parts are salvaged at 35% of the cost price, i.e. $w^o = -3,089$ for the old part and $w^n = -3,500$ for the new part. The base stock level is equal to $s = 2$ for both the old and new part. The repair yield of the part is $p_r^n = 0.5$. In the real case, old parts are not reworked into the new version, but we assume that the rework yield is $p_r^o = 0.5$ as well. The rework and repair rates are $\mu^o = \mu^n = 3/365$. The rework and repair costs

are 50% of the new part’s cost price, i.e. $z^o = z^n = 5,000$. Lastly, the emergency shipment penalty costs is equal to $b = 350$, which is a value based on the average penalty paid for the part’s emergency shipments in 2020 and 2021. Note again that all the aforementioned values are normalized. We present the results of the base case in Section 4.2.

The choices of the parameters other than their values in the base case are given in Table 6. We study whether an overestimation of the failure rate improvement leads to a wrong choice of policy by altering λ^n . Moreover, we investigate the impact that the number of preventively replaced parts before the start of the horizon and the size of the installed base have on the choice of policy, by altering $N - n$ and N , respectively. Lastly, we present the case where the costs of reworking a part equal the cost price, i.e. $z^o = p^n$. We present the results of the sensitivity analysis in Section 4.3.

Table 5: Values of factors/parameters for the base case (normalized).

N	n	T (years)	λ^o	λ^n	μ^o, μ^n	p_r^o, p_r^n	s	α
100	86	10	0.16/365	0.02/365	3/365	0.5	2	0.9995
x	p^n	h	w^o	w^n	z^o, z^n	d	f	b
	(€/part)	(€/part/day)	(€/part)	(€/part)	(€/part)	(€/failure)	(€/replacement)	(€/shipment)
2	10,000	4.66	-3,089	-3,500	5,000	3,839	2,687	350

Table 6: Alternative values of factors/parameters for the sensitivity analysis (normalized).

λ^n	$N - n$	N	z^o
0.05/365, 0.1/365	50, 100	50, 75, 100	10,000

4.2 Base Case Results

Table 7 presents the results of the base case (which have been verified with simulation). It shows that policy CR is preferred for its lowest costs, while policy PR is preferred when the number of failures is the main objective. Policies with a phase-out strategy to rework old parts (policies PR and CR) result in 6 – 8% lower costs compared to policies where old parts are salvaged. It is thus cheaper for the OEM to rework parts than to salvage them and buy new ones. Moreover, policies with an immediate roll-out strategy (policies PR and PS) result in 16 – 23% higher costs for the OEM, but 11% fewer failures than the policies with a corrective roll-out strategy (policies CR and CS). The higher costs can be explained by the costs incurred at the start of the horizon when preventively replacing parts. The lower number of failures are subsequently caused by the average installed base having a higher reliability as there are more new parts (with a higher reliability) installed at the start of the horizon.

Besides the evaluation of the four policies, we need to examine whether the development of the new component is worth it for the OEM. Therefore, for each policy, we calculate the return on investment (ROI) of the implementation of the new component for the OEM as the difference between the costs for the OEM when keeping the old component and the costs of implementing the new component according to that policy (see Table 7). Note that this value of the ROI excludes the costs for the development of the new component, i.e., it resembles the maximum cost that the development may incur in order for the new component to be successful. The higher the ROI, the better. Keeping the old component would generate 160 failures over a time horizon of $T = 10$ years, resulting in total expected costs for the OEM of €677,996. Table 7 shows that the ROI thus varies between €68,794 for policy PS and €183,366 for policy CR.

If the OEM would not have performance based contracts with its customers, it would not be responsible for the downtime at the customer. Considering its ROI, a corrective roll-out strategy combined with rework (policy CR) would then be favored. This may be at the expense of the customer, as the expected number of unscheduled failures is lower for immediate replacement policies than for corrective policies. Downtime of

Table 7: Results of the base case (normalized).

Policy	OEM cost	Unscheduled failures	OEM's ROI	Downstream supply chain cost	OEM's CpFR	Δ CpFR
PR	573,661	80	104,335	4,894,492	7,171	-1.46%
PS	609,202	80	68,794	4,929,660	7,615	-7.21%
CR	494,630	90	183,366	5,375,474	7,066	-
CS	535,225	90	142,771	5,422,463	7,646	-7.59%
Keep old component	677,996	160				

lithography systems due to unscheduled failures is very costly for the additional 100,000 euros to calculate the *downstream supply chain costs*. We therefore calculate the *downstream supply chain costs* by increasing the failure cost d with an additional 100,000 euros to account for this downtime at the customer, which is a conservative increase. From Table 7, we observe that when the OEM would choose to introduce the new components according to policy CR, this would indeed be at the expense of the customer. In fact, policy CR is 10% more expensive for the downstream supply chain than policy PR.

In practice, management would have to decide on the best compromise for both parties. To this end, we calculate the OEM's cost per failure reduction (CpFR) as

$$CpFR = \frac{OEM\ cost}{(\# failures\ keep\ old\ component - \# failures\ policy)}$$

This metric provides insights into what policy has the greatest impact on the number of unscheduled failures at the lowest costs for the OEM. We find that policy CR remains the preferred policy. Moreover, we calculate

$$\Delta CpFR = \frac{CpFR\ best\ policy - CpFR\ policy}{CpFR\ policy}$$

to determine to what extent this policy is preferred over the other policies. We find that the benefit of choosing policy CR over policy PR is only 1.46%, whereas it is more than 7% for the policies where the phase-out strategy is to salvage old parts. This highlights the importance of choosing the right policy to introduce the upgraded component to the field.

4.3 Sensitivity Analysis

In our sensitivity analysis, we vary one parameter at a time compared to the base case. We vary the values of λ^n , $N - n$, N , and z^o . We have tested more parameters, but we focus on the most interesting results here.

We derive the following managerial insights. Policies with an immediate roll-out strategy again result in the lowest number of unscheduled failures. Policy CR results in the lowest costs and highest ROI for the OEM, and policy PR results in the lowest downstream supply chain costs. Only when it becomes (almost) equally expensive to rework a part as to buy one, policies CS and PS are preferred, respectively. This result is not surprising as it then becomes cheaper to buy the part and salvage it, than to rework it.

Similar to the base case, we thus find that the OEM and its customers may have conflicting preferences for the choice of implementation policy. The corrective policies result in the lowest OEM costs and highest ROI, thereby favored by the OEM. The immediate replacement policies result in the lowest downstream supply chain costs and lowest number of unscheduled failures, thereby favored by the OEM's customers. Kopp et al. (2020) show in their research that lithography systems get the lowest proportion of preventive maintenance compared to other stations in the wafer fabrication process. Our results show that the immediate implementation of an upgraded component results in the largest improvement for the customer. As a result, customers of lithography systems would benefit from increasing the share of preventive maintenance despite the fact that it is the bottleneck workstation in many modern front-end wafer fabs.

We again examine the CpFR to find the best compromise between the OEM's and its customers' interests. Table 8 presents the values for the aforementioned parameters and the resulting CpFR for the best and second best policies. In addition, we present the delta between the two.

Table 8: Sensitivity analysis results (normalized).

Variables	Variable values	OEM cost best policy	CpFR best policy	CpFR 2nd best policy	Δ CpFR
Base Case		Policy CR 494,630	Policy CR 7,066	Policy PR 7,171	-1.46%
λ^n	$\lambda^n = 0.05/365$	Policy CR 546,485	Policy CR 9,936	Policy PR 10,113	-1.75%
	$\lambda^n = 0.1/365$	Policy CR 632,920	Policy CR 21,097	Policy PR 21,848	-3.44%
$N - n$	$N - n = 50$	Policy CR 494,630	Policy CR 7,066	Policy PR 7,399	-4.50%
	$N - n = 100$	Policy CR 494,630	Policy CR 7,066	Policy PR 7,569	-6.86%
N	$N = 50$	Policy CR 252,170	Policy CR 7,205	Policy PR 7,362	-2.13%
	$N = 75$	Policy CR 373,373	Policy CR 7,045	Policy PR 7,297	-3.46%
	$N = 125$	Policy CR 615,911	Policy CR 6,999	Policy PR 7,164	-2.31%
z^o	$z^o = 10,000$	Policy CS 383,746	Policy CS 5,482	Policy PS 5,987	-8.43%

The policies where the phase-out strategy is to rework old parts (policies CR and PR) are considered the best two policies when regarding the CpFR. Moreover, a corrective roll-out is preferred for all cases that were tested. Nevertheless, the delta between the two is relatively small. Only when it becomes (almost) equally expensive to rework a part as to buy one, policies CS and PS are preferred.

When examining the Δ CpFR in Table 8, we observe that the smaller the failure rate improvement of the new component, the more a corrective roll-out strategy is preferred over an immediate roll-out strategy. The relative difference between the expected number of failures becomes smaller for a smaller failure rate improvement. On the one hand, this implies that the benefit of choosing a policy with an immediate roll-out strategy is bigger when the reliability improvement is bigger. This result is intuitive, given that the immediate replacements increase the overall reliability of the installed base at the start of the horizon through the component’s increased reliability. On the other hand, it implies that these policies are more sensitive to a wrong estimation of the failure rate improvement.

Moreover, as expected, we observe a linear relationship between the size of the installed base and the expected costs and failures. It does not affect the choice for the best policy. A similar relationship is found for the customers’ acceptance of the preventive replacements. Every additional preventive replacement before the start of the horizon costs €5,650, but also prevents 0.7 failures in the following $T = 10$ years. Therefore, the difference in expected costs and failures between the policies having a corrective or immediate roll-out strategy becomes larger as more parts are preventively replaced before the start of the horizon. This results in an increase of the preference for a corrective roll-out strategy when more customers accept the preventive replacement before the start of the horizon.

5 CONCLUSION

OEMs who maintain a fleet of complex systems strive for the best performance of their systems, including high system availability for their customers. They want few failures at the lowest costs of service, as it benefits both themselves and their customers. We have developed a model that quantifies the impact of the introduction of an upgraded component on the OEM’s costs of service and the number of failures to help OEMs select the best implementation strategy of the upgraded component. Using a Markov process, we evaluate four policies that differ in *roll-out strategy* (immediate, corrective) and *phase-out strategy* (rework, salvage). The model is used in a case study at ASML. Our model enables OEMs to consciously select

the policy that best meets their needs, but also identifies what is the best compromise for the OEM and its customers. With the developed model, we are the first to take into account the impact of customers' acceptance of preventive replacements. Moreover, we include the possibility to rework old parts. The results show that reworking old parts is preferred over salvaging, unless the costs to rework become (almost) equally high as the cost price of the new part. Policies with a corrective roll-out are preferred for the OEM and policies with an immediate roll-out are preferred for the OEM's customers. When considering the number of unscheduled failures saved with the implementation of the new component, a corrective roll-out is generally found to be the best compromise between the two. Future studies may consider preventive roll-out strategies that assume that parts are replaced before the machine is likely to fail or non-linear production time losses to find better compromises between the OEM and the customer. Making the right decision for a certain implementation strategy secures an extended lifetime of the machines as well as the re-use of materials.

A ONE STEP TRANSITION PROBABILITY

Table 9: One step transition probability.

$j\lambda^o p_r^o$	if $s' = (i-1, j-1, k+1, l), i > 0, 0 < j \leq n-k, s \leq i+k+l \leq N$
$j\lambda^o p_r^o$	if $s' = (i, j-1, k+1, l), i > 0, 0 < j \leq n-k, i+k+l = s-1$
$j\lambda^o p_r^o$	if $s' = (i, j-1, k, l), i = 0, 0 < j \leq n-k, k > 0, s \leq k+l \leq N$
$j\lambda^o p_r^o$	if $s' = (i+1, j-1, k, l), i = 0, 0 < j \leq n-k, k > 0, k+l = s-1$
$j\lambda^o p_r^o$	if $s' = (i, j-1, k+1, l-1), i = 0, j > 0, k = 0, s \leq l \leq N$
$j\lambda^o p_r^o$	if $s' = (i+1, j-1, k+1, l-1), i = 0, j > 0, k = 0, l = s-1$
$j\lambda^o(1-p_r^o)$	if $s' = (i-1, j-1, k, l), i > 0, 0 < j \leq n-k, s \leq i+k+l \leq N$
$j\lambda^o(1-p_r^o)$	if $s' = (i, j-1, k, l), i > 0, 0 < j \leq n-k, i+k+l = s-1$
$j\lambda^o(1-p_r^o)$	if $s' = (i, j-1, k-1, l), i = 0, 0 < j \leq n-k, k > 0, s \leq k+l \leq N$
$j\lambda^o(1-p_r^o)$	if $s' = (i+1, j-1, k-1, l), i = 0, 0 < j \leq n-k, k > 0, k+l = s-1$
$j\lambda^o(1-p_r^o)$	if $s' = (i, j-1, k, l-1), i = 0, j > 0, k = 0, s \leq l \leq N$
$j\lambda^o(1-p_r^o)$	if $s' = (i+1, j-1, k, l-1), i = 0, j > 0, k = 0, l = s-1$
$(N-j)\lambda^n p_r^n$	if $s' = (i-1, j, k, l+1), i > 0, j \leq n-k, s \leq i+k+l \leq N$
$(N-j)\lambda^n p_r^n$	if $s' = (i, j, k, l+1), i > 0, j \leq n-k, i+k+l = s-1$
$(N-j)\lambda^n p_r^n$	if $s' = (i, j, k-1, l+1), i = 0, j \leq n-k, s \leq k+l = 0$
$(N-j)\lambda^n p_r^n$	if $s' = (i+1, j, k-1, l+1), i = 0, j \leq n-k, k = s-1l = 0$
$(N-j)\lambda^n(1-p_r^n)$	if $s' = (i-1, j, k, l), i > 0, j \leq n-k, s \leq i+k+l \leq N$
$(N-j)\lambda^n(1-p_r^n)$	if $s' = (i, j, k, l), i > 0, j \leq n-k, i+k+l = s-1$
$(N-j)\lambda^n(1-p_r^n)$	if $s' = (i, j, k, l-1), i = 0, j \leq n-k, l > 0, s \leq k+l \leq N$
$(N-j)\lambda^n(1-p_r^n)$	if $s' = (i+1, j, k, l-1), i = 0, j \leq n-k, l > 0, k+l = s-1$
$(N-j)\lambda^n(1-p_r^n)$	if $s' = (i, j, k-1, l), i = 0, j \leq n-k, s \leq k+l = 0$
$(N-j)\lambda^n(1-p_r^n) + k\mu^o$	if $s' = (i+1, j, k-1, l), i = 0, j \leq n-k, k = s-1l = 0$
$k\mu^o$	if $s' = (i+1, j, k-1, l), j \leq n-k, k > 0, s \leq i+k+l \leq N$
$k\mu^o$	if $s' = (i+2, j, k-1, l), j \leq n-k, k > 0, i+k+l = s-1$
$l\mu^n$	if $s' = (i+1, j, k, l-1), j \leq n-k, l > 0, s \leq i+k+l \leq N$
$l\mu^n$	if $s' = (i+2, j, k, l-1), j \leq n-k, l > 0, i+k+l = s-1$
$1 - \sum_{\hat{s} \in \mathcal{S}: \hat{s} \neq s'} p(\hat{s} s)$	if $s' = (i, j, k, l), s \leq i+k+l \leq N$
$1 - \sum_{\hat{s} \in \mathcal{S}: \hat{s} \neq s'} p(\hat{s} s)$	if $s' = (i+1, j, k, l), i+k+l = s-1$
0	otherwise.

REFERENCES

- ASML Holding N.V. 2014, 04. "ASML - Supply Chain Management". https://www.youtube.com/watch?v=7JwSF_6sdyo. Accessed 12th April 2023.
- Childress, S., and P. Durango-Cohen. 2005. "On Parallel Machine Replacement Problems with General Replacement Cost Functions and Stochastic Deterioration". *Naval Research Logistics (NRL)* 52(5):409–419.
- Driessen, J. 2018. *Life Cycle Costs Optimization for Capital Goods*. Ph.D. thesis, Eindhoven University of Technology.
- Forootani, A., M. G. Zarch, M. Tipaldi, and R. Iervolino. 2023. "A Stochastic Dynamic Programming Approach for the Machine Replacement Problem". *Engineering Applications of Artificial Intelligence* 118:105638.

- Hartman, J. 2000. "The Parallel Replacement Problem with Demand and Capital Budgeting Constraints". *Naval Research Logistics* 47(1):40–56.
- Kim, S.-H., M. Cohen, and S. Netessine. 2007. "Performance Contracting in After-Sales Service Supply Chains". *Management Science* 53(12):1843–1858.
- Kopp, D., M. Hassoun, A. Kalir, and L. Mönch. 2020. "SMT2020—A Semiconductor Manufacturing Testbed". *IEEE Transactions on Semiconductor Manufacturing* 33(4):522–531.
- Kopp, D., and L. Mönch. 2020. "Matheuristics for Qualification Management Decisions in Wafer Fabs". *IEEE Transactions on Semiconductor Manufacturing* 33(4):511–521.
- Lamghari-Idrissi, D., R. van Hugten, G.-J. van Houtum, and R. Basten. 2022. "Increasing Chip Availability Through a New After-Sales Service Supply Concept at ASML". *INFORMS Journal on Applied Analytics* 52(5):460–470.
- Mercier, S. 2008. "Optimal Replacement Policy for Obsolete Components With General Failure Rates". *Applied Stochastic Models in Business and Industry* 24(3):221–235.
- Mercier, S., and P.-E. Labeau. 2004. "Optimal Replacement Policy for a Series System With Obsolescence". *Applied Stochastic Models in Business and Industry* 20(1):73–91.
- Mizoguchi, H., T. Oga, K. Kakizaki, J. Fujimoto, and B. Fechner. 2021. "Excimer Lasers for Lithography and Annealing". In *Handbook of Laser Micro-and Nano-Engineering*, 367–427. Springer.
- Niimi, G., S. Nagai, T. Hori, Y. Ueno, T. Yanagida, K. Miyao, H. Hayashi, Y. Watanabe, T. Abe, H. Nakarai et al. 2020. "Update of Development Progress of the High Power LPP-EUV Light Source Using a Magnetic Field". In *Extreme Ultraviolet (EUV) Lithography XI*, Volume 11323, 492–505. SPIE.
- Oomen, P. 2016, December. "The Service Logistics Control Tower.". <http://www.servicelogisticsforum.nl/over-het-slf/activiteiten-archief/2016-6-themamiddagasml>. Accessed 12th April 2023.
- Schaefers, T., S. Ruffer, and E. Böhm. 2021. "Outcome-Based Contracting from the Customers' Perspective: A Means-End Chain Analytical Exploration". *Industrial Marketing Management* 93:466–481.
- Shu, Y., B. Porter, E. Soh, N. Farmakidis, S. Lim, Y. Lu, J. Warner, and H. Bhaskaran. 2021. "Nanoscale Bilayer Mechanical Lithography Using Water as Developer". *Nano Letters* 21(9):3827–3834.
- van Houtum, G.-J., and B. Kranenburg. 2015. *Spare Parts Inventory Control under System Availability Constraints*. New York: Springer.
- Van Schoot, J., E. van Setten, K. Troost, S. Lok, J. Stoeldraijer, R. Peeters, J. Benschop, J. Zimmerman, P. Graeupner, L. Wischmeier et al. 2020. "High-NA EUV lithography Exposure Tool: Program Progress". In *Extreme Ultraviolet (EUV) Lithography XI*, Volume 11323, 15–24. SPIE.
- Yang, N., K. Aslam, R. Schiffelers, L. Lensink, D. Hendriks, L. Cleophas, and A. Serebrenik. 2019. "Improving Model Inference in Industry by Combining Active and Passive Learning". In *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 253–263.
- Öner, K., G. Kiesmüller, and G.-J. van Houtum. 2015. "On the Upgrading Policy After the Redesign of a Component for Reliability Improvement". *European Journal of Operational Research* 244(3):867–880.

AUTHOR BIOGRAPHIES

STEFFI NEEFS was a Master's student at the school of Industrial Engineering and Innovation Sciences at the Eindhoven University of Technology. Her Master thesis project, conducted at ASML, formed the basis for this paper. She graduated in October 2022 and started her career in the industry. Her email address is steffi.neefs@asml.com.

DOUNIEL LAMGHARI-IDRISSI Douniel Lamghari-Idrissi worked in emergency response services and in the pharmaceutical industry. He joined ASML Netherlands B.V. in 2013, and works in customer support. He is also an assistant professor at Eindhoven University of Technology. His current research topic is the design of after-sales services that lead to more stable operations for the asset users. His email address is douniel.lamghari-idrissi@asml.com

ROB BASTEN is an associate professor at Eindhoven University of Technology. He works on maintenance and other after sales services. He is especially focused on using new technologies to improve such services, and in human behavior in the collaboration between decision makers and decision support systems. His email address is r.j.i.basten@tue.nl.

GEERT-JAN VAN HOUTUM is full professor of maintenance and reliability at Eindhoven University of Technology. He works on design and control of spare parts networks, predictive maintenance concepts, and product design choices that have a strong effect on system availability and TCO. His e-mail address is g.j.v.houtum@tue.nl.