Operational Models and Methods for Risk Informed Nuclear Asset Management

Alexander Galenko, David Morton, Elmira Popova
Graduate Program in Operations Research
Department of Mechanical Engineering
The University of Texas at Austin
Austin, TX 78712, USA

Rick Grantom, Ernie Kee, Drew Richards, Alice Sun
South Texas Project Nuclear Operating Company
Risk Management Department
P.O. Box 289, Wadsworth, TX 77483, USA

Abstract: We optimize replacement time for a system that fails according to an increasing failure-rate distribution. The problem has a finite horizon, and our goal is to minimize expected total maintenance cost. The associated objective function is piece-wise convex but discontinuous, and we develop an algorithm to optimize replacement time. We illustrate our approach on a real problem from the South Texas Project Nuclear Operating Company in Bay City, Texas, USA.

1 Introduction

In this paper, we model and solve the problem of choosing a maintenance policy for a single item, from a particular class of policies, that minimizes total expected cost over a finite horizon. Our work is applied to a system at the South Texas Project (STP) Electric Generation Station.

STP is one of the newest and largest nuclear power plants in the US, and is an industry leader in safety, reliability and efficiency. STP has two nuclear reactors that together can produce 2,500 megawatts of electric power. The reactors went online in August 1988 and June 1989, and are the sixth and fourth youngest, respectively, of more than 100 such reactors operating nationwide. STP consistently leads all US nuclear plants in the amount of electricity its reactors produce.

The STP Nuclear Operating Company (STPNOC) manages the plant for its owners, who share its energy in proportion to their ownership interests, which as of July 2004 are: Austin Energy, The City of Austin, 16%, City Public Service of San Antonio, 40%, and Texas Genco LP, 44%. All decisions that the board of directors make are finite time since every nuclear reactor is given a license to operate. In the case of STP, 25 and 27 years remain on the licenses for the two reactors, respectively.

Equipment used to support production in long-lived (more than a few years) installations such as those at STP requires maintenance. While maintenance is being carried out, the associated equipment is typically out of service and the system operator may be required to reduce or completely curtail production. In general, maintenance costs include labor, parts and overhead. In some cases, there are safety concerns associated with certain types of plant disruptions. Overhead costs in-
clude hazardous environment monitoring and mitigation, disposal fees, license fees, indirect costs associated with production loss such as wasted feed stock, and so forth.

STPNOC works actively with EPRI to develop more robust methods to plan maintenance activities and prioritize maintenance options. Existing nuclear industry guidelines [2, 4, 5, 6] recommend estimating reliability and safety performance based on evaluating changes taken one at a time, using risk importance measures supplemented by heuristics to prioritize maintenance. In our view, the nuclear industry can do better. For example, Liming et al. [8] propose instead investigating “packages” of changes, and in their study of a typical set of changes at STPNOC, the projected plant reliability and nuclear safety estimates were found to be significantly different than when changes were evaluated one at a time. STPNOC is working with EPRI to improve its preventive maintenance reliability database with more accurate distributions in further support of quantifying the results of preventive maintenance. Here, we consider a single-item maintenance problem. That said, this paper’s model forms the basis for higher fidelity multi-item maintenance models that we are currently investigating. Furthermore, as we describe below, even though the model is mathematically a single-item model, in practice it is applied to multiple items within an equipment class.

Equipment can undergo either preventive maintenance (PM) or corrective maintenance (CM). PM can include condition-based, age-based, or calendar-based equipment replacement or major overhaul. Also included in some strategies is equipment redesign. In all these categories of PM, the equipment is assumed to be replaced or brought back to the as-purchased condition. CM is performed when equipment has failed unexpectedly in service at a more-or-less random time (i.e., the out-of-service time is not the operator’s choice as in PM).

There is an enormous literature on optimal maintenance policies for a single item that dates back to the early 1950s. The majority of the work covers maintenance optimization over an infinite horizon, see Valdez-Flores and Feldman [11] for an extensive review. The problem that we address in this paper is over a finite planning horizon, which comes from the fact that every nuclear power plant has a license to operate that expires in a finite predefined time. In addition the form of the policy is effectively predefined by the industry as a combination of preventive and corrective maintenance, as we describe below. Marquez and Heguedas [9] present an excellent review of the more recent research on maintenance policies and solve the problem of periodic replacement in the context of a semi-Markov decision processes methodology. Su and Chang [10] find the periodic maintenance policies that minimize the life cycle cost over a predefined finite horizon.

With respect to the form of the maintenance policy that we consider, commercial nuclear power industry practice establishes a set PM schedule for many major equipment classes. In an effort to avoid mistakenly removing from service sets of equipment that would cause production loss or safety concerns, effectively all equipment that can be safely maintained together are grouped and assigned a base calendar frequency. The grouping policy ensures PM (as well as most CM) will occur at the same time for equip-
ment within a group. The base frequency is typically either set by the refueling schedule (equipment for which at-power maintenance is either impossible or undesirable) or calendar-based. The current thinking for PM is typified by, for example, the INPO guidance [6], whereby a balance is sought between maintenance cost and production loss. By properly taking into account the probability of production loss, the cost of lost production, the CM cost and PM cost, the simple model described in this paper captures industry practice. Not accounted for in the model is the probability and cost of production loss due to PM during at-power operation as well as potential critical path extension for PM during scheduled outages (such as refueling outage). This is partly justified by the fact that PM is only performed during the equipment’s group assigned outage time (unlike CM when the equipment outage time is not predetermined) and because outage planning will (generally) assure PM is done off critical path. In practice, there is normally one or two major equipment activities (main turbine and generator, reactor vessel inspection, steam generator inspection, main condenser inspection) that, along with refueling, govern critical path during planned outages.

2 Maintenance Model

The model we develop has the following constructs. The problem has a finite horizon of length $L$ (e.g., 25 or 27 years). We consider the following (positive) costs: $C_{pm}$ – preventive maintenance cost, $C_{cm}$ – corrective maintenance cost, and $C_d$ – downtime cost, which includes all lost production costs due to a disruption of power generation. Let $N(t)$ be the counting process for the number of failures in the interval $(0, t)$. We assume that this process is “well behaved,” and we cannot, for instance, have an infinite number of failures in a finite time interval.

Let the random time to failure of the item from its as-new state be governed by distribution $F$ with density $f$. Further assume that each failure of the item causes a loss of production (i.e., a plant trip) with probability $p$ and in that case a downtime cost, $C_d > C_{cm}$, is instead incurred (this cost can include $C_{cm}$, if appropriate). The optimization model has a single decision variable $T$, which is the time between PMs, i.e., we assume constant interval lengths between PMs.

We consider the following form of a maintenance policy, which we denote $\langle P \rangle$:

$\langle P \rangle$: Bring the item to the “as-good-as-new” state every $T$ units of time (preventive maintenance) at a cost of $C_{pm}$. If it fails meanwhile then repair it to the “as-good-as-old” state (corrective maintenance) for a cost of $C_{cm}$ or $C_d$, depending on whether the failure induces a production loss.

The goal is to find $T \in A \subset [0, L]$ that minimizes the total expected cost, i.e.,

$$\min_{T \in A} z(T),$$

where

$$z(T) = C_{pm} \lceil L/T \rceil + [pC_d + (1-p)C_{cm}] \cdot \left\{ [L/T]E[N(T)] + E \left[ N \left( L - T \left\lfloor L/T \right\rfloor \right) \right] \right\}.$$
represent feasible lengths for the replacement interval.

The objective function, \( z(T) \), consists of two terms. The first term is the unit preventive maintenance cost, \( C_{pm} \), multiplied by the number of PMs incurred. We assume that at the end of the planning period, i.e., time \( L \), a PM must be performed and this is why the ceiling operator appears in this first term. The second term in the objective function consists of the expected number of times the item fails (i.e., the term in curly brackets) multiplied by \( c_d \) for the expected fraction of failures that cause a loss of production and \( C_{cm} \) for the remaining fraction. The second term in the curly bracket counts the number of failures in the time interval from the last PM until time \( L \).

In what follows, we will use \( \bar{C}_{cm} = pC_d + (1 - p)C_{cm} \) to denote the average maintenance and down-time cost we incur per failure event. Barlow and Hunter [1] show that if an item is repeatedly repaired to the “as-good-as-old” state then the failure-event process is nonhomogeneous Poisson in which the expected number of failures in the interval \((0, t)\) is: \[ \int_0^t q(u) \, du \] where \( q(u) = \frac{f(u)}{1 - F(u)} \) is the associated failure rate function.

We assume that the time to failure comes from a distribution with increasing failure rate (IFR), i.e., the failure rate function \( q(t) \) is increasing. For instance, if the time between failures is distributed as a Weibull random variable with parameters \( \alpha > 1 \) and \( \lambda > 0 \) with probability density function \( f(t) = \alpha \lambda^\alpha t^{\alpha-1} e^{-\lambda t} \), then the failure rate function is \( q(t) = \alpha \lambda^\alpha t^{\alpha-1} \), and \( E[N(T)] = (\lambda T)^\alpha \).

As the following proposition shows, under an IFR distribution, the objective function \( z(T) \) is a piecewise convex function that is increasing within each piece.

**Proposition 1** Assume that we follow maintenance policy \( (P) \), and that the time between failures is a random variable with an IFR distribution, i.e., the failure rate function \( q(t) \) is increasing. Then, the objective function \( z(T) \) is:

1. lower semicontinuous with discontinuities at \( T = L/n, n = 1, 2, \ldots, \) and
2. increasing and convex on each interval \((\frac{L}{n+1}, \frac{L}{n})\), \( n = 1, 2, \ldots, \)

**Proof**

We first show (ii). Assume that \( T \in (\frac{L}{n+1}, \frac{L}{n}) \) for some \( n \). Within this interval

\[
z'(T) = \frac{dz(T)}{dT} = n\bar{C}_{cm} \{ q(T) - q(L - nT) \}.
\]

Here, we have used \( q(t) = \frac{d}{dt} E[N(t)] \) and the fact that \( [L/T] = n \) is constant within this \( n \)-th interval. That \( z \) increases on this interval follows from the fact that \( z'(T) \) is positive since \( q \) is increasing and \( L - nT < T \) for \( T \in (\frac{L}{n+1}, \frac{L}{n}) \). To show convexity, we show that \( z''(T) \) is increasing. Let \( T + \Delta T \in (\frac{L}{n+1}, \frac{L}{n}) \). Then, \( q(T + \Delta T) \geq q(T) \) and \( q(L - n(T + \Delta T)) \leq q(L - nT) \), and hence \( z'(T + \Delta T) \geq z'(T) \). This completes the proof of (ii).

The convexity result for \( z(T) \) from (ii) shows that \( T = L, L/2, L/3, \ldots \) are the only possible points of discontinuity in \( z(T) \). The first term in \( z(T) \) (i.e., the PM term) is lower semicontinuous because of the ceiling operator. The second term (i.e., \( z(T) \) when \( C_{pm} = 0 \)) is continuous in \( T \). To show this it suffices to verify that

\[
[L/T] E[N(T)] + E \left[ N \left( L - T \left| \frac{L}{T} \right. \right) \right]
\]
has limit $nE[N(d)]$ for both $T \to d^-$ and $T \to d^+$, where $d = L/n$ for any $n = 1, 2, \ldots$. This follows in a straightforward manner using the bounded convergence theorem, e.g., [3, pp. 33]. Hence, we have (i).

Let

$$z^c(T) = C_{pm}(L/T) + \tilde{C}_{cm}(L/T)E[N(T)],$$

i.e., $z^c(T)$ is $z(T)$ with $\lfloor L/T \rfloor$ and $\lceil L/T \rceil$ replaced by $L/T$. The following proposition characterizes $z^c(T)$ and its relationship with $z(T)$.

**Proposition 2**

Let $D = \{d : d = L/n, n = 1, 2, \ldots \}$ denote the set of discontinuities of $z(T)$ (cf. Proposition 1). Then,

(i) $z^c(T)$ is quasiconvex on $[0, L]$.

Furthermore if $d \in D$ then

(ii) $z^c(d) = z(d)$, and

(iii) $\lim_{T \to d^-} [z(T) - z^c(T)] = C_{pm}$.

**Proof**

(i): We have

$$z^c(T) = \frac{C_{pm} + \tilde{C}_{cm}E[N(T)]}{T/L}. \tag{2}$$

The numerator of (2) is convex in $T$ because $E[N(T)]$ has increasing derivative $q(T)$. As a result, $z^c(T)$ has convex level sets and is hence quasiconvex.

(ii): The result is immediate by evaluating $z$ and $z^c$ at points of $D$.

(iii): Let $d \in D$. Then, there exists and positive integer $n$ with $d = L/n$. Note that due to the ceiling operator, we have the following result

$$\lim_{T \to d^-} \left\{ \left\lfloor \frac{L}{T} \right\rfloor - \frac{L}{T} \right\} = 1. \tag{3}$$

Based on (3) and the definitions of $z$ and $z^c$ it suffices to show

$$[L/T]E[N(T)] + E \left[ N \left( L - T \left\lceil \frac{L}{T} \right\rceil \right) \right]$$

has limit $nE[N(d)]$ as $T \to d^-$. This was established in the proof of part (i) in Proposition 1.

Proposition 2 shows that at its points of discontinuity, our objective function $z$ drops by magnitude $C_{pm}$ to agree with $z^c$. Proposition 2 also establishes a quasiconvexity property of $z^c$ useful in solving our original model (1). Before developing the associated algorithm, the following proposition shows how to solve a simpler variant of our model in which the set $A$ of feasible replacement intervals is replaced by $[0, L]$.

**Proposition 3**

Let

$$T^*_c \in \arg\min_{T \in [0, L]} z^c(T)$$

and let $n^*$ be the positive integer satisfying

$$\frac{L}{n^* + 1} \leq T^*_c \leq \frac{L}{n^*}.$$ Then,

$$T^* \in \arg\min_{T \in \left\{ \frac{L}{n^* + 1}, \frac{L}{n^*} \right\}} z^c(T)$$

solves $\min_{T \in [0, L]} z(T)$.

**Proof**

Proposition 1 says that $z(T)$ increases on each interval $(\frac{L}{n + 1}, \frac{L}{n})$, and is lower semicontinuous. As a result, $\min_{T \in [0, L]} z(T)$ is equivalent to $\min_{T \in D} z(T)$, where $D = \{d : d = L/n, n = 1, 2, \ldots \}$ is the set of discontinuities of $z$. This optimization model is, in turn,
equivalent to \( \min_{T \in D} z^c(T) \) since, part (ii) of Proposition 2 establishes that \( z(T) = z^c(T) \) for \( T \in D \). Finally, since \( z^c(T) \) is quasiconvex (see Proposition 2 part (i)), we know \( z^c(T) \) is nondecreasing on \([0, \hat{T}_c] \) and nonincreasing on \([\hat{T}_c, L] \). Proposition 3 follows.

\[ \square \]

3 Algorithm

Proposition 3 shows how to solve \( \min_{T \in [0, L]} z(T) \). The key idea is to first solve the simpler problem \( \min_{T \in [0, L]} z^c(T) \), which can be accomplished efficiently, e.g., via a Fibonacci search. Let \( \hat{T}_c \) denote the optimal solution to the latter problem. Then, the optimal solution to the former problem is either the first point in \( D \) to the right of \( \hat{T}_c \) or the first point in \( D \) to the left of \( \hat{T}_c \), whichever yields a smaller value of \( z(T) \) (or \( z^c(T) \) since they are equal on \( D \)).

As described above, our original model (1) instead has the optimization restricted over a finite grid of points \( A \subset [0, L] \). Unfortunately, it is not true that the optimal solution to (1) is given by either the first point in \( A \) to the right of \( \hat{T}_c \) or the first point in \( A \) to the left of \( \hat{T}_c \). However, the results of Proposition 1, part (ii) establish that within the set \( A \cap (\frac{L}{n+1}, \frac{L}{n}) \) we can restrict attention to the left-most point (or ignore the segment if the intersection is empty). To simplify the discussion that follows we will assume that \( A \) has been redefined using this restriction. Furthermore, given a feasible point \( T_A \in A \) with cost \( z(T_A) \) we can eliminate from consideration all points in \( A \) to the right of \( \min\{T \in D : T \geq \hat{T}_c, z^c(T) \geq z(T_A)\} \) and all points in \( A \) to the left of \( \max\{T \in D : T \leq \hat{T}_c, z^c(T) \geq z(T_A)\} \). This simplification follows from the fact that \( z \) increases on each of its intervals and is equal to the quasiconvex \( z^c \) on \( D \).

Our algorithm therefore consists of:

- **Step 0:** Let \( T^*_c \in \arg \min_{T \in [0, L]} z^c(T) \).

- **Step 1:** Let \( T^* \in A \) be the first point in \( A \) to the right of \( T^*_c \) and let \( z^* = z(T^*) \). Let \( T_A = T^* \).

- **Step 2:** Increment \( T_A \) to be the next point in \( A \) to the right. If \( z(T_A) < z^* \) then \( z^* = z(T_A) \) and \( T^* = T_A \). Let \( d \) be the next point to the right of \( T_A \) in \( D \). If \( z^c(d) \geq z^* \) then let \( T_A \) to be the first point in \( A \) to the left of \( T^*_c \) and proceed to the next step. Otherwise, repeat this step.

- **Step 3:** If \( z(T_A) < z^* \) then \( z^* = z(T_A) \) and \( T^* = T_A \). Let \( d \) be the next point to the left of \( T_A \) in \( D \). If \( z^c(d) \geq z^* \) then output \( T^* \) and \( z^* \) and stop. Otherwise, increment \( T_A \) to be the next point in \( A \) to the left and repeat this step.

Step 0 solves the simpler continuous problem for \( T^*_c \), as described above and Step 1 simply finds an initial candidate solution to (1). Then, Steps 2 and 3, respectively, increment to the right and to the left of \( T^*_c \), moving from the point of \( A \) in, say, \( A \cap (\frac{L}{n+1}, \frac{L}{n}) \) to the point (if any) in the next interval. Increments to the right and to the left stop when further points in that direction are provably suboptimal. The algorithm is ensured to terminate with an optimal solution to (1).
4 Implementation in a real production setting

For use at STPNOC, we created an Excel drop-down menu with two main tools called *Create Maintenance Policy Report* and *Evaluate Single Project*. The menu (an add-in) calls a C++ library where the optimization algorithm described in the previous section is implemented. The GUI is coded in Visual Basic for Excel.

The tool *Create Maintenance Policy Report* produces a detailed report for chosen units (as indicated by their unique labels from the collection of system labels) from a database consisting of all required parameters for the model. The items are grouped functionally by service description and equipment type. There are two types of equipment used in a hydraulic oil application: pressure dampening accumulators; and servo control valves. There are three groups that are high pressure, high temperature water valves: 18 inch shut-off gate valves; 16 inch, air-operated, plug-style flow control valves (continuous modulation service); and 4 inch air-operated, plug-style flow control valves (intermittent modulation service). The input parameters are the PM cost, CM cost, the shutdown cost, the shutdown probability, the length of the planning horizon, and the failure distribution. In this current implementation we assume a Weibull lifetime distribution. We anticipate adding the lognormal and other appropriate failure-time distributions in the near future.

The parameters of the Weibull distribution, α and λ, are estimated via a Bayesian model described in Yu et al. [12]. This model is implemented (as a Microsoft Access application) in the STPNOC production setting and is run on a monthly basis by the Risk Management Department. The resulting estimates populate a field in the database from which this tool reads.

The preventive maintenance cost, $C_{pm}$, and corrective maintenance cost, $C_{cm}$, are estimated from historical data by a factor model, see Yu et al. [13], which is a function of covariates specific for each functional group. Downtime (or lost production) cost is the cost of lost revenue minus the cost of fuel. For mid-cycle downtime (i.e., an unplanned outage), the cost of replacement power is typically the most significant cost. An unplanned outage downtime does not normally require a significant augmentation of staff augmentation, with associated costs (in fact, there is usually insufficient time to bring in extra contractors). We subtract the fuel cost because it is obviously saved when the plant is shut down. The planning horizon $L$ for a unit is the amount of time from the current date until the plant shutdown date associated with the plant’s federal license.

The probability of lost production $p$ is determined by the availability model of the plant. The STPNOC availability model simulates the frequency and impact of equipment failures. A fault tree logic representation with equipment failure rate, planned maintenance unavailability, and corrective maintenance repair time data is used to find the contribution of equipment to plant-level unavailability and event frequencies [7].

Figure 1 shows a snapshot indicating the structure of the database. Figure 2 shows the dialog window from which the user can select which units to include in the maintenance report that will be generated. Figure 3 is a
As the snapshot indicates the report is for a unit known as wsf5 from group ID3. The application has collected the associated model parameters from the database. The model was solved and the window shows the optimal solution as the suggested time of replacement as 23 (the time units are in 12-week periods).

The tool Evaluate Single Project allows the user to find the optimal replacement time for single unit, and differs from the above case in that here the user (rather than the database) enters all the parameters. Figure 4 shows that form.

The application we have developed has value at STPNOC. Maintenance engineers must set the times and types of preventive maintenance to be conducted for each piece of equipment. There are industry standards that require investigating economic preventive maintenance (the most widely-used is INPO AP-913, [6]). Although in many cases, the manufacturer may recommend maintenance schedules, manufacturer recommendations are in some cases “conservative”, that is, more extensive and higher frequency than necessary to maintain adequate performance in the plant application. In the absence of the types of analytical tools we have developed, the maintenance engineer must essentially “guess” at an “optimum” timing for the length of the maintenance period. The engineer may at times feel pressure to reduce
maintenance (to reduce short-term operating costs) but at the same time the engineer is expected to keep the equipment running well. This can be challenging and in some cases, the engineers may fail to “get it right” (optimum) because, from time to time, the frequency of preventive maintenance is reduced only to later be increased because, e.g., of poor equipment performance or plant shut downs. Typically, the feedback from the effect of a change to maintenance policy is not fast enough to make a fine-tuned correction. That is, at a time when cost overruns are being minimized, it is tempting to reduce maintenance on well-running equipment. By the time the of a reduction in maintenance is seen, the equipment performance may have significantly degraded and it is natural to overreact by increasing maintenance to a higher level than previous. Our analytical plug-in tool can in seconds generate a report for hundreds of units that can be used as a decision-support guide for engineers working on maintenance scheduling thus reducing pressure on them and reducing guessing in the process of forming the maintenance policy. By being careful to design the tool around the plant equipment numbering system, we can leverage the information available in the computerized maintenance management and computerized work management systems to minimize effort of the engineer in retrieving necessary maintenance and cost histories.

5 Acknowledgments

This work was supported by the National Science Foundation under grants DMI-0217927 and DMI-0457558, and by the South Texas Project Nuclear Operating Company under grant B02857.

References


