Optimal Maintenance and Repair Policies under Nonlinear Preferences

Stephen D. Boyles[1]; Zhanmin Zhang[2]; and S. Travis Waller[3]

Abstract: This paper is concerned with infrastructure maintenance and repair policies when the managing agency does not value expenses linearly. Such nonlinearity naturally arises in the stochastic world of reliability engineering, manifesting, for instance, as risk aversion: the ideal maintenance and repair policy has a low cost variance, as well as a low average cost. Another type of nonlinear behavior arises when one tries to match future expenditures with an externally determined budget. Dynamic programing techniques are applied to create two algorithms (FindPolicy and EvalPolicy) which are of use in this problem: FindPolicy determines an optimal maintenance policy, while EvalPolicy allows a previously determined policy to be evaluated according to a broad class of measures of effectiveness. These algorithms are applied to a hypothetical bridge facility and to determine and evaluate short-term and long-term maintenance policies. In this example, large reductions in solution variance are attainable with only slight increases in average cost.

DOI: 10.1061/(ASCE)1076-0342(2010)16:1(11)

CE Database subject headings: Rehabilitation; Infrastructure; Maintenance; Costs.

Author keywords: Rehabilitation; Infrastructure; Maintenance; Costs.

Introduction

It is difficult to overstate how important transportation infrastructure is to modern society due to the vast mobility and economic benefits that it provides. In this light, proper maintenance of this infrastructure is of vital importance due to the large capital expenditure and construction time needed to construct new facilities. At the same time, maintenance activities have nontrivial costs, and agencies responsible for maintenance have limited budgets, making it necessary to determine what maintenance actions to perform, and when, in order to ensure a well-functioning system with a reasonable cost. Because it is impossible to forecast the future deterioration and maintenance needs of a facility with complete accuracy, maintenance *policies* are developed, which specify many different possible maintenance and rehabilitation (M&R) schedules, along with a set of rules specifying which schedule is to be used under which realization.

Broadly speaking, two separate problems need to be solved: the problem of expenditure planning, in which an agency must decide how much money is needed to achieve a long-term maintenance goal, and the problem of budget allocation, addressing the short-term decisions of which repair actions to perform once the actual budget and facility status are known with more certainty.

[1]Assistant Professor, Univ. of Wyoming, Dept. 3295, 1000 E. University Ave., Laramie, WY 82071 (corresponding author). E-mail: sboyles@uwyo.edu

[2]Associate Professor, The Univ. of Texas at Austin, 1 University Station C1761, Austin, TX 78712. E-mail: z.zhang@mail.utexas.edu

[3]Associate Professor, The Univ. of Texas at Austin, 1 University Station C1761, Austin, TX 78712. E-mail: stw@mail.utexas.edu

Note. This manuscript was submitted on June 24, 2008; approved on May 11, 2009; published online on February 12, 2010. Discussion period open until August 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Infrastructure Systems***, Vol. 16, No. 1, March 1, 2010. ©ASCE, ISSN 1076-0342/2010/1-11–20/$25.00.

In the literature, several models have been constructed to address these problems; for instance, Mori and Ellingwood (1994) constructed a nonlinear program to minimize inspection, repair, and expected failure costs, and Madanat and Ben-Akiva (1994) developed a dynamic programing method for finding an optimal maintenance and inspection policy in the presence of inspection error. Smilowitz and Madanat (2000) extended this to the network level, considering the best policies for a number of facilities joined by a common budget constraint, as do Gao and Zhang (2006). Kong and Frangopol (2001) presented a software program to find an optimal maintenance schedule using a number of parameters to represent facility deterioration. Robelin (2006) and Robelin and Madanat (2007) incorporated history dependence into a dynamic programing model, allowing limited information on past maintenance history to influence the current decision. Robelin and Madanat (2008) extended this analysis to a system of facilities.

All of the facility-level models mentioned share the assumption that the objective is to minimize long-term expected costs. While reasonable in some cases, other concerns are also important in defining a truly "optimal" policy—it is also desirable for policies to be robust to future conditions, performing well under a variety of future scenarios. This phenomenon is well known in the finance literature, leading, for example, to the use of objectives minimizing some combination of mean portfolio performance and its variance (Markowitz 1952), value-at-risk (Duffie and Pan 1997), or conditional value-at-risk (Rockafellar and Uryasev 2000). Essentially, the presence of uncertainty leads to risk aversion in decision making, financially, and otherwise. This effect is also seen in other areas of transportation engineering, such as mode choice (Pinjari and Bhat 2006) or route choice (Liu et al. 2004; Small et al. 2005; Liu et al. 2007), in which travel reliability is seen to play a large role in trip-making behavior.

Several authors have suggested alternate methods of finding policies which reduce uncertainty: Kuhn and Madanat (2006), Gao and Zhang (2008), and Marfitano (2008) applied techniques of robust optimization which solve "minimax" problems intended

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

to optimize worst-case results. While robust optimization methods carry the advantage of not requiring probability distributions to be defined for uncertain parameters, the solutions they generate are often overconservative (Bertsimas and Sim 2004); more information on these methods can be found in Ben-Tal and Nemirovski (2002) and elsewhere. Alternately, Kuhn and Madanat (2005) and Madanat et al. (2006) developed methods of addressing uncertainty in the deterioration and repair model, and Li and Puyan (2006) used stochastic programing methods to address budget uncertainty.

However, even these techniques cannot capture the full range of possible objectives: for instance, in the short run, minimizing expected cost may not even be desirable: given a fixed budget, one usually opts to use as much of it as possible to improve facility reliability. Even in such cases, uncertainty about the future means it is desirable to find a maintenance policy in which deviation from the budget is minimized (overexpenditure being problematic for the obvious reason, and underexpenditure problematic because it indicates inefficient allocation of resources and missed opportunities to improve reliability). This "target seeking" behavior cannot be captured simply by trying to minimize expected expenditures.

Both of these are examples of nonlinear preferences on the part of the managing agency. To the writer's knowledge, the problem of finding an optimal M&R policy for such preferences has not been previously studied in the literature. Thus, this paper develops two efficient, dynamic programing-based algorithms (FindPolicy and EvalPolicy) to solve it for a broad class of user behaviors. These algorithms can address both the expenditure planning and budget allocation problems when applied at different time scales. The Problem Statement section describes this class of preferences and outlines the structure of the problem, presenting suitable notation and defining basic concepts. The dynamic programming algorithms are given in the Solution Methods section along with their derivations. The Numerical Example section demonstrates this algorithm using a fictitious bridge. Finally, the Conclusion section briefly summarizes the paper, emphasizing the key conclusions and pointing to future extensions of this work.

Problem Statement

This paper describes exact algorithms for finding an optimal M&R policy under nonlinear preferences and evaluating it under various measures of effectiveness, a problem which is defined in this section. In turn, this section describes the context of this problem, the deterioration assumptions, how a utility function can be used to represent different preferences, and how risk attitudes can be incorporated in this framework.

Problem Context and Deterioration Model

Assume that an agency is responsible for maintaining a single deteriorating facility. For ease, we discretize this problem in several ways. First, we assume that decisions about whether to perform any action on the facility are made on a periodic basis, indexed by the variable t, and that the specific time at which maintenance is actually performed during each period is irrelevant. Second, we assume that the condition of the facility can be represented by a discrete scalar $\beta \in B$ (called the *reliability index*). These assumptions are common; Robelin and Madanat (2007) define β in relation to the instantaneous probability of failure, with the latter probability equal to $\Phi(-\beta)$, with Φ as the cumulative distribution function for the standard normal distribution. Integer values of β between 1 and 15 are used for the reliability index in their work, and we adopt the same definition in this paper. Third, we assume a finite set of *actions* U and their associated costs $c_u(\beta,t) \in \mathbb{Z}^+$, at most one of which can be performed in any given period; these represent different types of maintenance activities, including complete replacement. For convenience, we include the "do nothing" action in this set with a cost of zero. The set U is assumed to be constant over time; however, the costs c_u are allowed to vary over time (perhaps representing inflation or a discount factor) or according to the condition β of the facility when the action is performed.

The goal is to find an appropriate policy π to describe how M&R should be performed on this facility, within a finite time horizon $\bar{T} \in \mathbb{Z}^+$ representing the number of discrete time periods in the analysis period; the length of each period, as well as the time horizon $\bar{T}$ itself depend on whether a short-term or long-term analysis is sought. As it turns out, the algorithms developed in this paper function equally well for either case; the difference between these time scales lies primarily in the measures of effectiveness used. As future conditions are unknown, a policy specifies which action (if any) is to be taken for each possible *state* the facility can exist in. For this problem, we define a state x as a five tuple (t,β,m,τ,e), indicating the current time t, reliability level β, the last maintenance action m performed, the time τ since this action was performed, and the amount of money e spent to date. Let X represent the set of all possible states for this facility. Then a policy π is a function mapping each state X to an element of U.

In the algorithms that follow, it is desirable for the state space to be as small as possible. In particular, expenditures should be measured in units allowing the costs c_u to be represented as small integers. For example, Tao et al. (1995) considered three possible bridge maintenance actions with per-girder costs of \$1,170, \$3,800, and \$8,400. In our model, one might opt to approximate these actions' costs as multiples of \$1,200, respectively, giving costs of 1, 3, and 7 units. Defining the set E_t as the set of possible values of the expenditure e at time t, this choice of costs yields sets E_t of size $7t$. A better approximation could be obtained with a smaller unit of measurement (say, \$200) with a corresponding increase in the size of sets E_t. Define $c_{\max}=\max_{\beta,t,u}\{c_u(\beta,t)\}$; this constant is useful for establishing the complexity results in the next section.

Regarding deterioration of the facility, we take the Markovian assumption that the probability p_{ij} of moving from one state i to another j does not depend on any additional information; the current state suffices for knowing the transition probabilities. We also write $p_{i,j|u}$ to express the transition probabilities conditional on performing action u in state i. Under the assumption that these probabilities are stationary and do not depend on the to-date expenditures (except as captured by the current reliability level and the last maintenance action), we can specify the transition probabilities using a smaller set of states, namely, $Y=B\times U\times T$, with $T=\{1,2,\ldots,\bar{T}\}$ as the set of periods within the time horizon. We assume that these probabilities are known *a priori*; these might be based directly on historical experience, or taken from simulation—Robelin and Madanat (2007) described an algorithm to estimate these probabilities from a physical deterioration model using Monte Carlo sampling.

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

Utility Function

To model different preferences, define a utility function $f(\beta, e)$ representing agency satisfaction for having a facility in condition β at the end of the time horizon $\bar{T}$, having spent e during that time. (In general, f could depend on the last maintenance action m which was performed τ periods ago; however, this may be a less common preference, and these indices are therefore suppressed for brevity.) We seek a policy π such that $E[f]$ is maximized. f can account for concepts such as salvage value, achieving M&R goals (e.g., meeting or exceeding a long-term reliability standard), total budget spent, and so on; the generality of the algorithms presented in this paper is based in the flexibility enabled by different definitions of f, possibly for different time scales. For this reason, we discuss several possible choices, and the goals they represent.

For the long-term budget planning problem, if the goal is to minimize expected expenditure while satisfying a minimum reliability level β_0, one might define $f(\beta, e)=-e$ for $\beta \geqslant \beta_0$ and $f(\beta, e)=-e-M$ otherwise, where $M>0$ is a suitably large constant to penalize failure to meet the minimum reliability standard; the expected cost of the resulting policy can then be used in forming budget requests and forecasts. Another possible goal is to simply maximize the probability that the facility's reliability index is at least β_1, in which case $f(\beta, e)=1$ for $\beta \geqslant \beta_1$ and 0 otherwise, suffices. These options can also be combined: $f(\beta, e)=-e+N$ for $\beta \geqslant \beta_1$, $-e$ for $\beta_0 \leqslant \beta < \beta_1$, and $-e-M$ for $\beta < \beta_0$ attempt to minimize expected cost, while penalizing failure to meet a minimum standard β_0 and rewarding achievement of a desired goal β_1, with M and N suitably defined.

Other options are available for finding short-term budget allocation policies. Although budget constraints are not always considered for single-facility problems, in the short term, they are highly relevant in practice. If the cost of a recommended action exceeds currently available funds, a suboptimal action may need to be performed instead, or maintenance deferred until the next budgeting period. Depending on the agency, money may be borrowed or emergency funding requested, but these actions are costly (the former directly, due to interest; the latter often indirectly, from a loss of trust or public image), and these costs should be reflected in the utility function.

If the budget b available until time $\bar{T}$ is known, adding a large penalty constant to f whenever $e>b$ discourages policies that exceed the budget. Other options are to include a hard budget constraint by fixing $f(\beta, e)=-\infty$ for $e>b$ (effectively prohibiting any overspending), or to progressively penalize greater violations of the budget, such as by subtracting an additional term $(e-b)^2$ from f whenever $e>b$. Or, if budget rules allow unspent funds to be "rolled over" to the next time period, $f(\beta, e)=-e$ for $\beta \geqslant \beta_0$, $e \leqslant b$, $f(\beta, e)=-e-M$ for $\beta < \beta_0$, and $f(\beta, e)=-e-(e-b)^2$ for $e>b$ attempts to minimize present expenditures while meeting a minimum reliability level β_0 and satisfying the current budget, allowing greater future flexibility. The flexibility of the utility function f allows this type of short-term tactical spending decisions to be studied, along with the long-term strategic planning that deterioration models usually consider.

Risk Preferences

Economists have suggested that risk preferences are of great importance in stochastic environments. Many proposed risk measurse can be incorporated within the utility function f. Perhaps the simplest option is to attempt to minimize the variance of utility. Since f must be completely specified *a priori*, one cannot directly minimize utility variance, which requires that the expected solution utility be known before the values of f can be calculated (since $\mathrm{Var}[f]=E[(f-E[f])^2]$). However, an appropriate choice of utility function can ensure a reduction in utility variance, as compared to the policy π which simply maximizes $E[f]$.

Proposition 1. *Let* θ *be a policy maximizing the expected value of the utility function* $-(f-f^*)^2$, *where* $f^*=E_\pi[f]$. *Then* $\mathrm{Var}_\theta[f] \leqslant \mathrm{Var}_\pi[f]$.

Proof

$$\begin{aligned}
\mathrm{Var}_\theta[f] &= E_\theta[f^2]-(E_\theta[f])^2 \\
&= E_\theta[f^2]-(E_\theta[f])^2+[2f^*E_\theta[f]+(f^*)^2]-[2f^*E_\theta[f] \\
&\quad +(f^*)^2] \\
&= E_\theta[(f-f^*)^2]-[E_\theta[f]-(f^*)]^2 \\
&\leqslant E_\theta[(f-f^*)^2] \leqslant E_\pi[(f-f^*)^2] \\
&\quad \{\text{since } \theta \text{ minimizes } E[(f-f^*)^2]\} = \mathrm{Var}_\pi[f]
\end{aligned}$$

□

Essentially, this approach minimizes the deviation from a specified "target" utility value f^*. By choosing f^* to be the expected utility from a utility-maximizing policy, Proposition 1 ensures a reduction in utility variance, while also avoiding the problem that directly minimizing variance may result in a very poor expected utility. The mean-variance technique, developed by Markowitz (1952) in the finance domain, proposes an alternate way of avoiding this problem by maximizing a weighted difference of expected utility and utility variance. The same approach can be applied here by choosing the utility function $f-\lambda(f-f^*)^2$, with λ an appropriately chosen weighting constant.

Asymmetric preferences are often observed. For instance, overspending the budget by a given amount is much worse than underspending by the same amount, or missing a reliability target for a facility is worse than exceeding the target. This can be represented directly in the f function as described above, or through a utility function of the form $[f-f^*]^+-\mu[f^*-f]^+$, with $[\cdot]^+=\max\{\cdot, 0\}$, and μ a weighting constant describing how falling short of the target utility f^* should be penalized, relative to exceeding it.

Prospect theory provides yet another approach for modeling preferences under uncertainty. Proposed by Kahneman and Tversky (1979), this approach attempts to describe behavior through a two-phase process: first, possible outcomes are ranked, and a reference point is chosen. The desirability of each outcome is determined by an S-shaped value function based on its desirability relative to the reference point. This function is usually chosen to represent loss aversion, where outcomes worse than the reference point are penalized more heavily. Next, an action is chosen maximizing the expected value of the value function. Given a reference utility f^* and a value function $v(\beta, e, f^*)$, this approach can be modeled in our framework by choosing our utility function to coincide with the value function for all possible outcomes.

Thus, the function f allows a very flexible representation of agency preferences under uncertainty. One should note that previous Markovian policy models do not include the current expenditures e in the state definition; while this is not necessary when preferences are linear, it is important to include it for nonlinear preferences. An intuitive reason for this is given by considering spending behavior with a budget constraint: one should be allowed to differentiate between the cases when one has already

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

spent most of the budget, and the case where one has only spent a small portion of it. Since such distinctions are possible in reality, and relevant with nonlinear objectives, current expenditure must be included in the state definition.

Solution Methods

In this section, we describe exact algorithms to find a policy π minimizing $E[f]$ (FindPolicy) and to evaluate a given policy according to a metric defined by another utility function g (EvalPolicy). Both FindPolicy and EvalPolicy apply dynamic programing techniques originated by Bellman (1957). Dynamic programing can be used to solve discrete problems satisfying certain properties, namely, optimal substructure and overlapping subproblems. Optimal substructure requires that components of an optimal solution should also optimize smaller subproblems; for instance, for the problem of finding an optimal M&R policy, consider some optimal policy for times $1,2,\ldots,\bar{T}$. If we only consider the portions of this policy for, say, $T\geqslant 2$, this should itself be an optimal policy for times $2,3,\ldots,\bar{T}$: if it were not, we could update some part of both this and the original policy as well, obtaining better solutions for both.

The overlapping subproblem criterion essentially requires that the problem be solvable sequentially, often by working backward. When finding an optimal M&R policy, one can start at the last time period $\bar{T}$ and consider the optimal policy for all the states corresponding to the previous period $\bar{T}-1$. Since there is only one time period left, knowing the value of the facility in each possible final state, it is relatively easy to find the optimal actions at period $\bar{T}-1$. Knowing this, one proceeds to find the optimal actions at $\bar{T}-2$, only having to look at $\bar{T}-1$, because the final period has already been accounted for.

Algorithm FindPolicy

FindPolicy determines the optimal action to take for every possible state. This is done by working backward from the final time interval, so at every point one can assume that the future actions and outcomes have been determined for every possible sequence of states the facility will follow as it deteriorates and is repaired.

To assist in this process, we associate two labels L and π with each state $x=(t,\beta,m,\tau,e)\in X$, where L denotes the expected value of the objective function, assuming that the policy developed thus far will be implemented until the time horizon, and π stores the optimal action to take at state x. Taken together, the labels $\pi(x)$ define the policy we seek.

FindPolicy is initialized by setting final utility levels $L(x)=f(\beta,e)$ for all states $x=(\bar{T},\beta,m,\tau,e)$ corresponding to the final time interval. Given these final utility values and the deterioration model, the optimal actions during time interval $\bar{T}-1$ can be found: for each state, calculate the expected utility that would be obtained from applying each possible action; the action maximizing that expected utility is optimal and assigned to that state.

Now, consider an arbitrary time interval t, and assume that these labels have been found for all time intervals $t+1,\ldots,\bar{T}$, and we want to set the labels for the states associated with time t. For each possible action u, we can calculate temporary labels e^u denoting the expected objective function value if the policy were to use action u in this state. After doing this for all actions, choose

Algorithm 1. FindPolicy

1: {Initialization}
2: **for all** $t\in\{1,2,\ldots,\bar{T}\}$, $y=(\beta,m,\tau)\in Y$, $e\in E_t$ **do**
3: **if** $t=\bar{T}$ **then**
4: $L(\bar{T},y,e)\leftarrow f(\beta,e)$
5: **else**
6: $L(\bar{T},y,e)\leftarrow -\infty$
7: **end if**
8: **end for**
9: {Iteration}
10: **for all** $t=\bar{T}-1,\ \bar{T}-2,\ldots,1$ **do**
11: **for all** $y\in Y$, $e\in E_t$ **do**
12: **for all** $u\in U$ **do**
13: $L^u\leftarrow \Sigma_{z\in Y}p_{y,z|u}L(t+1,z,e+c_u(\beta,t))$
14: **if** $L^u(t,y,e)<L(t,y,e)$ **then**
15: $L(t,y,e)\leftarrow L^u(t,y,e)$
16: $\pi(t,y,e)\leftarrow u$
17: **end if**
18: **end for**
19: **end for**
20: **end for**
21: **return** π

the action u^* maximizing L^u, set the label L equal to L^{u^*}, and set μ to u^*. Denoting (β,m,τ) by y, the temporary label can be calculated as

$$L^u=\sum_{z\in Y}p_{y,z|u}L[t+1,z,e+c_u(\beta,t)] \tag{1}$$

This formula accomplishes the calculation needed for the backward recursion. The algorithm thus considers time intervals $\bar{T}-1$, $\bar{T}-2$, and so on, in turn before terminating at the first time period with the optimal policy fully specified. This is presented formally below, as Algorithm 1.

From a computational standpoint, one does not need to consider *all* states $z\in Y$ when evaluating the sum in Eq. (1) since many of the transition probabilities $p_{y,z|u}$ must be zero. For instance, if one is performing action u at time t, the succeeding state will certainly have $m=u$ for the last maintenance action, and $\tau=1$ for the time since that action was performed: all transition probabilities to states with different m and τ values must be zero. Taking this into account, the time and space complexity of FindPolicy are as follows.

Proposition 2. *Algorithm FindPolicy can be implemented to require* $O(\bar{T}^3|B||U|c_{\max})$ *space and run in* $O(\bar{T}^3|B|^2|U|^2c_{\max})$ *time.*

Proof. FindPolicy stores two variables for each state x. As the five dimensions of X are of size $\bar{T}$, $|B|$, $|U|$, $O(\bar{T})$, and $\bar{T}c_{\max}$, respectively, storing these labels requires $O(\bar{T}^3|B||U|c_{\max})$ space.

The loops in lines 10–12 of Algorithm 1 require the summation in line 13 to be performed $O(\bar{T}^3|B||U|^2c_{\max})$ times. As depicted, each summation requires adding $O(\bar{T}|B||U|)$ terms; however, since the only unknown component of z is the facility reliability in the next stage, this term can be calculated by only adding $|B|$ terms. The overall complexity immediately follows. □

Algorithm EvalPolicy

The second algorithm, EvalPolicy, allows a specified policy π to be evaluated according to a given measure of effectiveness (not

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

Algorithm 2. EvalPolicy

```
1:  {Initialization}
2:  for all x ∈ X do
3:    if x=x_0 then
4:      ρ(x) ← 1
5:    else
6:      ρ(x) ← 0
7:    end if
8:  end for
9:  {Iteration}
10: for all t=1, ..., T̄−2, T̄−1 do
11:   for all y=(β,m,τ) ∈ Y, e ∈ E_t do
12:     for all z ∈ Y do
13:       ρ(t+l, z, e+c_{π(t,y,e)}(β)) ← ρ(t+1, z, e+c_{π(t,y,e)}(β)) + ρ(t,y,e) p_{y,z|π(t,y,e)}
14:     end for
15:   end for
16: end for
17: return Σ_{y∈Y,e} ρ(T̄,y,e) g(y,e)
```

necessarily the same one which π optimizes). For instances, although π may have been chosen to minimize expected cost, an agency may be interested in knowing the probability that the actual expenditures under this policy exceed some threshold value, or facility condition exceeds a certain target level, or in calculating the standard deviation in expenditures. Even if none of these are the primary goal, EvalPolicy allows policies to be compared using these (or other) metrics.

EvalPolicy uses a forward recursion: given an initial (current) state for the facility and the values of the desired measure of effectiveness for all states at the final time $\bar{T}$, the probability of each final state is calculated, allowing direct calculation of the chosen metric's expected value. We denote this metric by $g(\beta,m,\tau,e)$. Although serving a similar purpose to f in FindPolicy, we distinguish the notation for two reasons: first, we intend to clarify that the policy π need not maximize g; second, it is quite plausible that this metric may be a function of the state elements m and τ at the end of the time horizon, in contrast to FindPolicy, where we suggest the opposite.

For each state $x \in X$, we store a label $\rho(x)$ denoting the probability that state x occurs if policy π is followed; clearly we have $\Sigma_{\beta,m,\tau,e}\rho(t,\beta,m,\tau,e)$ for all $t \in 1,2,\ldots,\bar{T}$. Thus, for a state $x=(t,\beta,m,\tau,e)$ we calculate

$$\rho(x) = \sum_{y \in Y} \sum_{e \in E_{t-1}} \rho(t-1,y,e) p_{y,(\beta,m,\tau)|\pi(t-1,y,e)}$$

if we know the probabilities of each state occuring during the preceding time interval. Thus, beginning with $\rho(x_0)=1$ for the initial state x_0, we proceed in increasing order of time to calculate these probabilities for all states. Upon completion, calculation of the expected value of g is straightforward. This is formally presented as Algorithm 2.

Proposition 3. *Algorithm EvalPolicy can be implemented to require $O(\bar{T}^2|B||U|c_{\max})$ space and run in $O(\bar{T}^3|B|^2|U|c_{\max})$ time.*

Proof. As shown in Algorithm 2, EvalPolicy stores a label for each state x. Although $|X|$ is $O(\bar{T}^3|B||U|c_{\max})$ as shown in the proof of Proposition 2 at any point the algorithm only uses labels from the current time interval and its immediate predecessor, allowing reuse of storage space and reducing the space complexity by a factor of $\bar{T}$.

The loops in lines 10 and 11 require lines 12–14 to be repeated $O(\bar{T}^3|B||U|c_{\max})$ times. Using the same logic as in the proof of Proposition 2, only $|B|$ states need to be considered in the innermost loop, providing the time complexity. □

While the Markovian deterioration model is useful for finding optimal policies, one may wish to evaluate policies using a more accurate model. This may occur if the Markov transition probabilities are taken from a simplified version of another deterioration model (for instance, estimated from repeated simulations). A Monte Carlo approach can then be used to simulate the facility's condition as the maintenance policy is applied and estimate the desired measures of effectiveness. If the Markov probabilities are only approximate, this approach may yield more accurate estimates; however, this comes at a cost in computation time since it lacks the efficient forward recursion and exact computation of expected value found in EvalPolicy.

Numerical Example

In this section, we demonstrate how FindPolicy and EvalPolicy can be applied in determining M&R policies, both in the long-term expenditure planning and in the short-term budget allocation contexts. As an example, we consider a bridge facility which exists in one of 15 reliability levels or conditions ($B=\{1,2,\ldots,15\}$). Maintenance decisions are scheduled according to 6-month intervals, and budgets are determined biennially. The minimum allowable reliability index is three; we control for this in the model by identifying all of the states corresponding to a lower condition and setting their f label to $-\infty$. The deterioration model and parameters are taken from Frangopol et al. (2001), and the transition probabilities are estimated using the procedure in Robelin and Madanat (2007).

In any given year, three possible actions can be performed (in addition to the do nothing option). Minor repairs, representing routine maintenance, costing between 1 and 3 units depending on the reliability level ($c_1=2-\lfloor\beta/5\rfloor$), increase the reliability level by γ_1, which is drawn from a lognormal distribution with a mean and standard deviation of two reliability units. Major repairs represent a more involved repair action: the cost is three times that of minor repairs, and the reliability is increased by γ_2, which is drawn from a lognormal distribution with mean of 7 and standard deviation of 3. The third action, replacement, has a cost of 21 and restores the facility to new condition ($\beta=15$). This choice of actions, along with their costs and effects, follows Tao et al. (1995).

We first consider the long-term planning decision, assuming a 20-year time horizon. The long-term goal is to have this facility have a reliability index equal to 12 or greater at the end of the time horizon, and the agency wishes to know the minimum expected cost needed to satisfy this condition. However, recognizing that this cost may be prohibitively high depending on the initial state of the bridge, this long-term goal is not enforced as a hard constraint. Rather, the agency decides that achieving this goal is "worth" five units of cost; that is, it is willing to spend up to five additional units on maintenance actions if doing so would result in $\beta \geq 12$. Thus, the appropriate choice of utility function is $f_1(\beta,e)=-e$ for $\beta<12$ and $f(\beta,e)=5-e$ for $\beta \geq 12$. With this choice of time horizon and model parameters, the state space X consists of 3,969,000 elements. Implemented on a 2.6 GHz Pentium III machine using Windows XP with 4 Gbit random access

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

Table 1. Comparison of Utility-Maximizing Long-Term Policies

β_0	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$\Pr(\beta \geq 12)$	$E[f]$	$\sigma[f]$
3	13.868	9.454	10.869	4.868	0.614	−16.000	19.467
4	9.985	8.819	9.439	4.807	0.484	−13.962	20.093
5	6.498	6.192	8.770	4.432	0.428	−11.094	16.534
6	5.121	4.276	9.096	4.296	0.473	−5.597	11.916
7	4.990	3.079	11.009	3.797	0.697	−3.928	8.401
8	4.571	2.359	12.463	2.661	0.866	−2.600	5.403
9	3.699	2.023	13.131	1.716	0.941	−1.222	3.701
10	2.895	1.662	13.353	1.336	0.962	−0.158	2.915
11	2.287	1.461	13.398	1.277	0.965	1.537	2.036
12	1.806	1.342	13.401	1.277	0.965	2.020	1.950
13	1.301	1.225	13.398	1.277	0.965	2.525	1.872
14	0.776	1.090	13.392	1.278	0.965	3.049	1.789
15	0.671	1.040	13.391	1.278	0.965	3.614	1.545

memory, FindPolicy requires 12.7 s of computation time, on average, and EvalPolicy requires 0.19 s.

Algorithm FindPolicy is then applied to determine the optimal policy, and EvalPolicy is used multiple times to evaluate this policy according to several metrics: the expected value and standard deviation of expenditures, the expected value and standard deviation of the reliability index at the end of the time horizon, the probability of achieving the reliability goal, and the expected value and variance of utility. These results are shown in Table 1

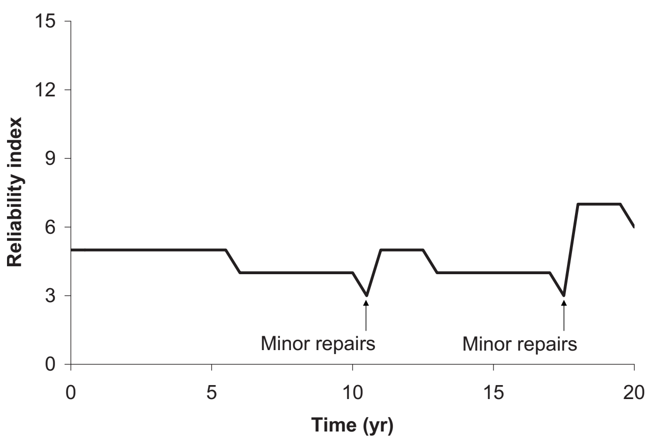

Fig. 1. One realization of a long-term utility-maximizing policy

for different possible initial values for bridge reliability (denoted β_0). (In all cases, it is assumed that minor repairs were performed in the year prior to analysis.) Fig. 1 plots the reliability over time for one particular realization of this policy, indicating when maintenance is performed; since the deterioration model is stochastic, naturally this plot only represents one possible outcome.

As an alternative, the agency is interested in a policy which is "robust" to future uncertainty in some sense. To accomplish this, FindPolicy is used a second time, with $f_2(\beta,e)=-[f_1(\beta,e)-f_1^*]^2$, with f_1 as before and f_1^* equal to the expected value of f_1 under the first policy. This choice of utility function is appropriate because it penalizes deviation from the average outcome and thus encourages a policy with smaller variance in utility. Following this, EvalPolicy is then applied using the same measures of effectiveness; these results are shown in Table 2. Note that, for each initial reliability value, the robust policies have a slightly lower expected utility ($E[f]$) but substantially lower standard deviation of utility ($\sigma[f]$), indicating that policies with more predictable outcomes are readily available and are not significantly worse in terms of expected utility. (Average expenditures are much higher, but, according to the specified utility function, this is nearly balanced by a corresponding increase in facility condition.) Also note that the expected reliability and probability of attaining the goal are higher for the robust policies along with expected expenditures; an intuitive interpretation of this result is spending more to improve facility condition is less affected by future uncertainty

Table 2. Comparison of "Robust" Long-Term Policies

β_0	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$\Pr(\beta \geq 12)$	$E[f]$	$\sigma[f]$
3	23.014	0.566	14.294	2.901	0.952	−18.250	1.793
4	20.026	1.457	14.832	1.009	0.971	−15.170	2.002
5	16.32	2.943	14.271	2.171	0.892	−11.861	4.376
6	9.391	3.738	11.461	4.147	0.555	−6.615	5.012
7	7.272	3.097	11.533	4.297	0.584	−4.354	3.887
8	7.168	2.065	13.229	3.352	0.78	−3.269	2.796
9	5.951	1.188	14.293	2.089	0.935	−1.276	2.160
10	4.830	1.289	14.445	1.516	0.938	−0.141	1.612
11	3.356	0.736	14.576	0.773	1.000	1.642	1.058
12	3.190	0.559	14.617	0.757	0.999	1.807	0.954
13	2.667	0.705	14.626	0.733	1.000	2.331	0.772
14	2.185	0.549	14.575	0.783	0.999	2.813	0.490
15	1.456	0.818	14.354	0.891	0.999	3.540	0.178

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

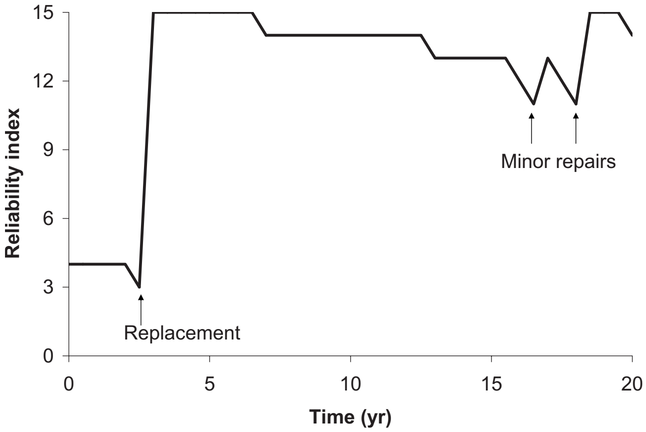

Fig. 2. One realization of a long-term robust policy

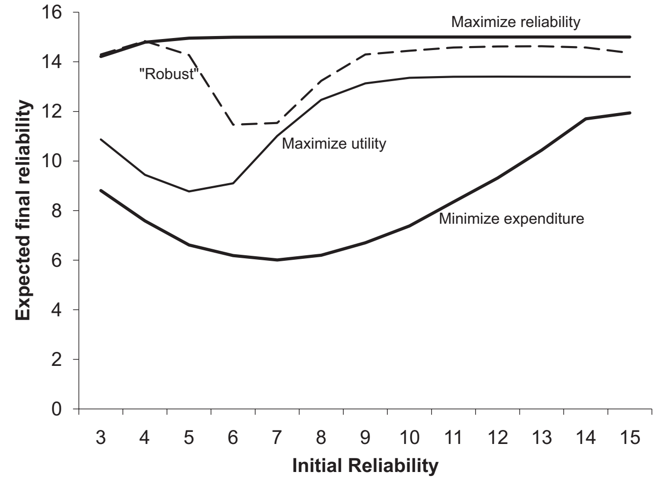

Fig. 4. Expected final reliability for four policies

than the first policy, which solely sought to maximize expected utility. (Of course, this result could vary with a different utility function or problem parameters.)

Fig. 2 plots reliability over time for one realization of a robust policy, providing graphical demonstration of how higher average expenditures for this policy translate into less variation in utility. By completely replacing the facility when reaching a degraded condition and then performing routine maintenance, less variation is seen in the final outcome, as compared to the utility-maximizing policy, where the replacement action is much rarer. Intuitively, the robust policy plans for a replacement expenditure at some point in the time horizon, while the utility-maximizing policy does not; the result is higher average expenditure under the former, but less unpredictability in both expenditure and final facility condition.

These results are typical of the utility-maximizing and robust policies. Figs. 3 and 4 plot the expected expenditure and final reliability for these policies, showing how they vary according to the initial reliability. For comparison, the cost-minimizing and reliability-maximizing policies are also plotted. One can observe that these policies adopt a compromise between these two objectives, achieving higher reliability than the cost-minimizing policy would, and requiring lower expenditure than the reliability-maximizing policy.

As a final long-term planning consideration, recall that EvalPolicy uses the Markovian deterioration model to evaluate policies. Since the Markovian transition probabilities were estimated from a deterioration model from Frangopol et al. (2001), it is worthwhile to compare EvalPolicy's results with that of simulation using the latter model. Table 3 shows the expected utility for the utility-maximizing and robust policies, both as calculated by EvalPolicy, and as estimated through simulation based on 5,000 Monte Carlo trials for each initial reliability level. The results are mostly consistent, although EvalPolicy seems to have greater difficulty matching the simulated results for very low initial reliability in the case of the utility-maximizing policy. Excluding these, the root-mean-square difference between expected utilities is slightly greater than 1. Given the utility specification, this is approximately equivalent to the cost of one minor repairs. Considering the twenty-year time horizon, this does not seem unreasonable.

After considering these long-term policies, a budget recommendation is made. Now, assume that the actual amount of money allocated for the first two years is five units, which is higher than expected. The relevant short-term question is how to make the best use of this funding while not exceeding the assigned budget. Three possible utility functions achieving this purpose are $f_3(\beta,e)=\beta$ for $e\leqslant 5$ and $\beta-10$ otherwise (assigning a

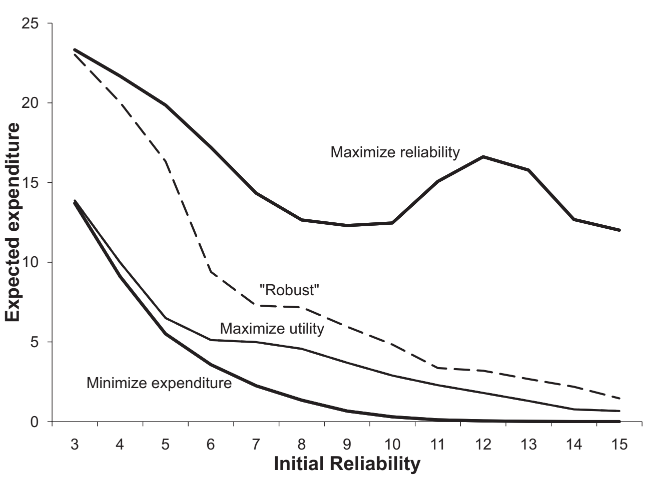

Fig. 3. Expected expenditure for four policies

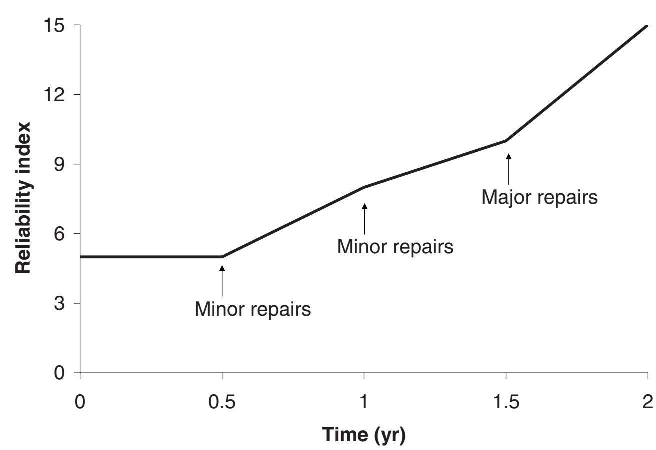

Fig. 5. One realization of a short-term policy with a constant penalty for exceeding the budget (f_3)

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

Table 3. Comparison of EvalPolicy and Simulated Deterioration

Initial reliability	Utility maximizing policy			"Robust" policy		
	EvalPolicy	Simulation	Diff	EvalPolicy	Simulation	Diff
3	−16.000	−11.834	−4.166	−18.250	−18.472	0.222
4	−13.962	−8.442	−5.520	−15.170	−16.290	1.120
5	−11.094	−5.735	−5.359	−11.861	−12.759	0.898
6	−5.597	−3.950	−1.647	−6.615	−8.005	1.390
7	−3.928	−2.061	−1.867	−4.354	−5.207	0.853
8	−2.600	−1.121	−1.479	−3.269	−3.990	0.721
9	−1.222	−0.649	−0.573	−1.276	−2.285	1.009
10	−0.158	−0.053	−0.105	−0.141	−1.169	1.028
11	1.537	0.498	1.039	1.642	0.272	1.370
12	2.020	1.245	0.775	1.807	1.223	0.584
13	2.525	1.983	0.542	2.331	1.524	0.807
14	3.049	2.782	0.267	2.813	1.198	1.615
15	3.614	3.037	0.577	3.540	1.756	1.784

penalty of ten reliability units for exceeding the budget), $f_4(\beta,e)=\beta$ for $e \leqslant 5$ and $-\infty$ otherwise (effectively eliminating the possibility of overspending), and $f_5(\beta,e)=\beta$ for $e \leqslant 5$ and $\beta -2(e-5)^2$ otherwise (penalizing greater violation of the budget by progressively larger amounts). For each of these, FindPolicy is applied (using a two-year time horizon) to determine the optimal policy for each objective, and EvalPolicy is used to determine the expected value and standard deviation of expenditures and reliability at the end of this budget period; these are shown in Table 4, and possible realizations for these policies are plotted in Figs. 5–7. This choice of a time horizon results in a state space consisting of 37,800 elements, and the algorithms execute in less than a tenth of a second.

Note that the budget is frequently exceeded if the initial reliability index is very low, even under objective f_4 which supposedly prevents exceeding the budget in an absolute sense; the implication is that it is impossible to satisfy this requirement while also ensuring that the bridge does not fall below the minimum reliability index standard. In this demonstration, this conflict was resolved by relaxing the budget constraint, rather than the minimum reliability constraint—in terms of implementation, the $-\infty$ quantities are replaced by large negative numbers, and the $-\infty$ corresponding to the budget constraint was chosen to be smaller in magnitude than the $-\infty$ corresponding to the reliability constraint. Also note that, for policy f_3, the expected final reliability indices are higher for $\beta_0 \in \{3,4\}$ than for $\beta_0 \in \{5,6,7\}$. This occurs because, for utility function f_3, violation of the budget by any amount incurs the same penalty; thus, once more than five cost units have been spent, there is no penalty for greater spending, indicating that this utility function may not be the best way to account for the budget.

Although this example is clearly a great simplification of the budget planning and allocation process, it nevertheless demonstrates how FindPolicy and EvalPolicy can be used to support agencies in making such decisions, providing quantitative analysis of a variety of objectives.

Conclusion

In this paper, we present techniques for finding and evaluating M&R policies which account for certain types of nonlinear agency preferences regarding maintenance expenditures, including behaviors such as risk aversion, goal setting, and target seeking, both in the long-term (budget planning) and in the short-term

Table 4. Comparison of Short-Term Policies

Initial reliability	f_3				f_4				f_5			
	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$
3	19.44	8.115	13.571	2.837	6.73	7.436	6.967	4.230	5.790	0.535	6.685	2.504
4	14.766	9.496	12.042	3.507	4.236	2.440	6.777	2.986	5.514	0.500	7.707	2.516
5	6.710	5.437	9.365	2.921	4.820	0.572	8.434	2.613	5.090	0.286	8.624	2.461
6	4.295	0.883	9.872	2.753	4.267	0.455	9.859	2.752	6.000	0	12.169	2.151
7	4.417	0.519	11.114	2.694	4.417	0.519	11.114	2.694	5.137	0.768	11.872	2.284
8	4.634	0.537	12.551	2.387	4.634	0.537	12.551	2.387	4.867	0.636	12.799	2.131
9	4.816	0.504	13.901	1.707	4.816	0.504	13.901	1.707	4.85	0.516	13.938	1.630
10	4.853	0.513	14.752	0.813	4.853	0.513	14.752	0.813	4.856	0.513	14.755	0.800
11	5.000	0	14.998	0.051	5.000	0	14.998	0.051	5.000	0	14.998	0.051
12	3.790	1.169	15	0.017	3.790	1.169	15	0.017	3.790	1.169	15	0.017
13	3.182	1.076	15	0.004	3.182	1.076	15	0.004	3.182	1.076	15	0.004
14	3.023	0.496	15	0	3.023	0.496	15	0	3.023	0.496	15	0
15	1.216	0.626	15	0	1.216	0.626	15	0	1.216	0.626	15	0

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

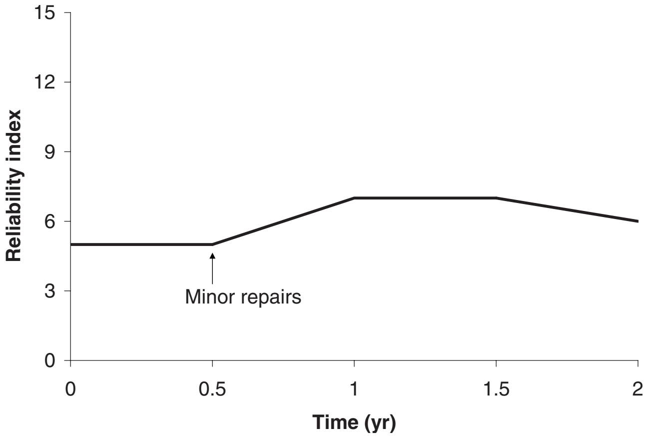

Fig. 6. One realization of a short-term policy with a hard budget constraint (f_4)

(budget allocation) horizons. Using the framework of a history-dependent Markovian deterioration model, two dynamic programing-based algorithms are developed to find optimal policies with respect to these preferences and to evaluate such policies according to a wide variety of measures of effectiveness. A hypothetical bridge was used as an example, indicating how both short-term and long-term policies can be derived and compared using these algorithms.

Still, additional field data should be applied to see if these conclusions drawn from a fictitious example can be extended to general practice. Another useful extension of this work is to consider the macroscopic, network-level optimization problem, where multiple facilities must be maintained, under a common budget or other bundling constraint.

Acknowledgments

The writers would like to thank Aristeidis Pantelias and three anonymous referees for their valuable comments and suggestions. Of course, the writers remain solely responsible for the contents of the paper. The first writer would also like to thank Samer Madanat, whose excellent presentation at the University of Texas inspired the research presented here.

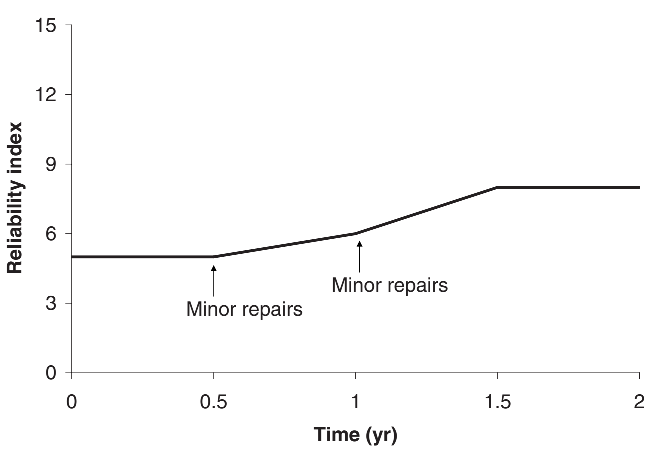

Fig. 7. One realization of a short-term policy with a variable penalty for exceeding the budget (f_5)

Notation

The following symbols are used in this paper:

B = set of possible reliability indices;
$c_{\max}$ = cost of most expensive action, in appropriate units;
c_u = cost of a maintenance action;
E_t = set of possible expenditures at time t;
e = expenditures to-date;
f = utility function representing agency preferences;
g = utility function used to evaluate a policy;
L = labels used in algorithm FindPolicy;
m = last maintenance action performed;
$p_{i,j|u}$ = probability of progressing from state i to state j if maintenance action u is performed;
T = set of time indices;
$\bar{T}$ = time horizon;
t = current time;
U = set of possible maintenance actions;
X = set of all possible facility states;
x = facility state (including t, β, m, τ, and e);
Y = smaller set of states used to represent transition probabilities;
$\mathbb{Z}^+$ = set of nonnegative integers;
β = facility reliability index;
π = maintenance and rehabilitation policy (formally, a function mapping X to U); and
τ = time since last maintenance action was performed.

References

Bellman, R. (1957). *Dynamic programming*, Princeton University Press, Princeton, N.J.

Ben-Tal, A., and Nemirovski, A. (2002). "Robust optimization—Methodology and applications." *Math. Program. Ser. B*, 92, 453–480.

Bertsimas, D., and Sim, M. (2004). "The price of robustness." *Oper. Res.*, 52(1), 35–53.

Duffie, D., and Pan, J. (1997). "An overview of value-at-risk." *Journal of Derivatives*, 4, 7–49.

Frangopol, D. M., Kong, J. S., and Gharaibeh, E. S. (2001). "Reliability-based management of highway bridges." *J. Comput. Civ. Eng.*, 15(1), 27–34.

Gao, L., and Zhang, Z. (2006). "Robust optimization for managing pavement maintenance and rehabilitation." *Proc., 85th Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, D.C.

Gao, L., and Zhang, Z. (2008). "Robust optimization for managing pavement maintenance and rehabilitation." *Proc., 87th Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, D.C.

Kahneman, D., and Tversky, A. (1979). "Prospect theory: an analysis of decision under risk." *Econometrica*, 47, 263–291.

Kong, J. S., and Frangopol, D. M. (2001). "Life-cycle reliability-based maintenance cost optimization of deteriorating structures with emphasis on bridges." *J. Struct. Eng.*, 129(6), 818–828.

Kuhn, K. D., and Madanat, S. M. (2005). "Model uncertainty and the management of a system of infrastructure facilities." *Transp. Res., Part C: Emerg. Technol.*, 13, 391–404.

Kuhn, K. D., and Madanat, S. M. (2006). "Robust maintenance policies for Markovian systems under model uncertainty." *Comput. Aided Civ. Infrastruct. Eng.*, 21, 171–178.

Li, Z., and Puyan, M. (2006). "A stochastic optimization model for high-

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

way project selection and programming under budget uncertainty." *Proc., 9th Int. Conf. on Applications of Advanced Technology in Transportation*, ASCE, Reston, Va.

Liu, H. X., He, X., and Recker, W. (2007). "Estimation of the time-dependency of values of travel time and its reliability from loop detector data." *Transp. Res., Part B: Methodol.*, 41(4), 448–461.

Liu, H. X., Recker, W., and Chen, A. (2004). "Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data." *Transp. Res., Part A: Policy Pract.*, 38(6), 435–453.

Madanat, S., and Ben-Akiva, M. (1994). "Optimal inspection and repair policies for infrastructure facilities." *Transp. Sci.*, 28(1), 55–62.

Madanat, S., Park, S., and Kuhn, K. (2006). "Adaptive optimization and systematic probing of infrastructure system maintenance policies under model uncertainty." *J. Infrastruct. Syst.*, 12(3), 192–198.

Marfitano, S. C. (2008). "Robust optimization for budget allocation at network level using clustering techniques." MS thesis, The Univ. of Texas at Austin, Austin, Tex.

Markowitz, H. M. (1952). "Portfolio selection." *Coll. Math. J.*, 7(1), 77–91.

Mori, Y., and Ellingwood, B. R. (1994). "Maintaining reliability of concrete structures. II: Optimum inspection/repair." *J. Struct. Eng.*, 120(3), 846–862.

Pinjari, A., and Bhat, C. (2006). "On the nonlinearity of response to level of service variables in travel mode choice models." *Transp. Res. Rec.*, 1977, 67–74.

Robelin, C.-A. (2006). "Facility-level and system-level stochastic optimization of bridge maintenance and replacement decisions using history-dependent models." Ph.D. thesis, Univ. of California, Berkeley, Calif.

Robelin, C.-A., and Madanat, S. M. (2007). "History-dependent bridge deck maintenance and replacement optimization with markov decision processes." *J. Infrastruct. Syst.*, 13(3), 195–201.

Robelin, C.-A., and Madanat, S. M. (2008). "Reliability-based system-level optimization of bridge maintenance and replacement decisions." *Transp. Sci.*, 42(4), 508–513.

Rockafellar, R. T., and Uryasev, S. (2000). "Optimization of conditional value-at-risk." *J. Risk*, 2, 21–41.

Small, K. A., Winston, C., and Yan, J. (2005). "Uncovering the distribution of motorists' preferences for travel time and reliability." *Econometrica*, 73(4), 1367–1382.

Smilowitz, K., and Madanat, S. (2000). "Optimal inspection and maintenance policies for infrastructure networks." *Comput. Aided Civ. Infrastruct. Eng.*, 15, 5–13.

Tao, Z., Corotis, R. B., and Ellis, J. H. (1995). "Reliabiltiy-based structural design with Markov decision processes." *J. Struct. Eng.*, 121(6), 971–980.

Downloaded 07 Jun 2011 to 129.132.202.35. Redistribution subject to ASCE license or copyright. Visit http://www.ascelibrary.org

Optimal Maintenance and Repair Policies under Nonlinear Preferences

Stephen D. Boyles * Zhanmin Zhang † S. Travis Waller ‡

*Assistant Professor, University of Wyoming, sboyles@mail.utexas.edu, Dept. 3295, 1000 E. University Ave., Laramie, WY 82071

†Associate Professor, The University of Texas at Austin, z.zhang@mail.utexas.edu, 1 University Station C1761, Austin, TX 78712

‡Associate Professor, The University of Texas at Austin, stw@mail.utexas.edu, 1 University Station C1761, Austin, TX 78712

Abstract

This paper is concerned with infrastructure maintenance and repair policies, when the managing agency does not value expenses linearly. Such nonlinearity naturally arises in the stochastic world of reliability engineering, manifesting, for instance, as risk aversion: the ideal maintenance and repair policy has a low cost variance, as well as a low average cost. Another type of nonlinear behavior arises when one tries to match future expenditures with an externally-determined budget. Dynamic programming techniques are applied to create two algorithms (FindPolicy and EvalPolicy) which are of use in this problem: FindPolicy determines an optimal maintenance policy, while EvalPolicy allows a previously-determined policy to be evaluated according to a broad class of measures of effectiveness. These algorithms are applied to a hypothetical bridge facility, to determine and evaluate short-term and long-term maintenance policies. In this example, large reductions in solution variance are attainable with only slight increases in average cost.

1 Introduction

It is difficult to overstate how important transportation infrastructure is to modern society, due to the vast mobility and economic benefits that it provides. In this light, proper maintenance of this infrastructure is of vital importance, due to the large capital expenditure and construction time needed to construct new facilities. At the same time, maintenance activities have nontrivial costs, and agencies responsible for maintenance have limited budgets, making it necessary to determine what maintenance actions to perform, and when, in order to ensure a well-functioning system with a reasonable cost. Because it is impossible to forecast the future deterioration and maintenance needs of a facility with complete accuracy, maintenance *policies* are developed, which specify many different possible maintenance and rehabilitation (M&R) schedules, along with a set of rules specifying which schedule is to be used under which realization.

Broadly speaking, two separate problems need to be solved: the problem of expenditure planning, in which an agency must decide how much money is needed to achieve a long-term maintenance goal, and the problem of budget allocation, addressing the short-term decisions of which repair actions to perform once the actual budget and facility status are known with more certainty.

In the literature, several models have been constructed to address these problems; for instance, Mori and Ellingwood (1994) construct a nonlinear program to minimize inspection, repair, and expected failure costs, and Madanat and Ben-Akiva (1994) develop a dynamic programming method for finding an optimal maintenance and inspection policy, in the presence of inspection error. Smilowitz and Madanat (2000) extend this to the network level, considering the best policies for a number of facilities joined by a common budget constraint, as do Gao and Zhang (2006). Kong and Frangopol (2001) present a software program to find an optimal maintenance schedule, using a number of parameters to represent facility deterioration. Robelin (2006) and Robelin and Madanat (2007) incorporate history dependence into a dynamic programming model, allowing limited information on past maintenance history to influence the current decision. Robelin and Madanat (2008) extend this analysis to a system of facilities.

All of the facility-level models mentioned share the assumption that the objective is to minimize long-term expected costs. While reasonable in some cases, other concerns are also important in defining a truly "optimal" policy — it is also desirable for policies to be robust to future conditions, performing well under a variety of future scenarios. This phenomenon is well-known in the finance literature, leading, for example, to the use of objectives minimizing some combination of mean portfolio performance and its variance (Markowitz, 1952), value-at-risk (Duffie and Pan, 1997), or conditional value-at-risk (Rockafellar and Uryasev, 2000). Essentially, the presence of uncertainty leads to risk aversion in decision making, financially and otherwise. This effect is also seen in other areas of transportation engineering, such as mode choice (Pinjari and Bhat, 2006) or route choice (Liu et al., 2004; Small et al., 2005; Liu et al., 2007), in which travel reliability is seen to play a large role in trip-making behavior.

Several authors have suggested alternate methods of finding policies which reduce uncertainty: Kuhn and Madanat (2006), Gao and Zhang (2008), and Marfitano (2008) apply techniques of robust optimization which solve "minimax" problems intended to optimize worst-case results. While robust optimization methods carry the advantage of not requiring probability distributions to be defined for uncertain parameters, the solutions they generate are often overconservative (Bertsimas and Sim, 2004); more information on these methods can be found in Ben-Tal and Nemirovski (2002) and elsewhere. Alternately, Kuhn and Madanat (2005) and Madanat et al. (2006) develop methods of addressing uncertainty in the deterioration and repair model, and Li and Puyan (2006) use stochastic programming methods to address budget uncertainty.

However, even these techniques cannot capture the full range of possible objectives: for instance, in the short run, minimizing expected cost may not even be desirable: given a fixed budget, one usually opts to use as much of it as possible to improve facility reliability. Even in such cases, uncertainty about the future means it is desirable to find a maintenance policy in which deviation from the budget is minimized (overexpenditure being problematic for the obvious reason, and underexpenditure problematic because it indicates inefficient allocation of resources, and missed opportunities to improve reliability). This "target seeking" behavior cannot be captured simply by trying to minimize expected expenditures.

Both of these are examples of nonlinear preferences on the part of the managing agency. To the author's knowledge, the problem of finding an optimal M&R policy for such preferences has not been previously studied in the literature. Thus, this paper develops two efficient, dynamic programming-based algorithms (FindPolicy and EvalPolicy) to solve it for a broad class of user behaviors. These algorithms can address both the expenditure planning and budget allocation problems, when applied at different time scales. Section 2 describes this class of preferences and outlines the structure of the problem, presenting suitable notation and defining basic concepts. The dynamic programming algorithms are given in Section 3, along with their derivations. Section 4 demonstrates this algorithm using a fictitious bridge. Finally, Section 5 briefly summarizes the paper, emphasizing the key conclusions and pointing to future extensions of this work.

2 Problem Statement

This paper describes exact algorithms for finding an optimal M&R policy under nonlinear preferences and evaluating it under various measures of effectiveness, a problem which is defined in this section. In turn, this section describes the context of this problem, the deterioration assumptions, how a utility function can be used to represent different preferences, and how risk attitudes can be incorporated in this framework.

2.1 Problem Context and Deterioration Model

Assume that an agency is responsible for maintaining a single, deteriorating facility. For ease, we discretize this problem in several ways. First, we assume that decisions about whether to perform any action on the facility are made on a periodic basis, indexed by the variable t, and that the specific time at which maintenance is actually performed during each period is irrelevant. Second, we assume that the condition of the facility can be represented by a discrete scalar $\beta \in B$ (called the *reliability index.*) This assumptions are common; Robelin and Madanat (2007) define β in relation to the instantaneous probability of failure, with the latter probability equal to $\Phi(-\beta)$, with Φ the cumulative distribution function for the standard normal distribution. Integer values of β between 1 and 15 are used for the reliability index in their work, and we adopt the same definition in this paper. Third, we assume a finite set of *actions* U and their associated costs $c_u(\beta, t) \in \mathbb{Z}^+$, at most one of which can be performed in any given period; these represent different types of maintenance activities, including complete replacement. For convenience, we include the "do nothing" action in this set, with a cost of zero. The set U is assumed to be constant over time; however, the costs c_u are allowed to vary over time (perhaps representing inflation or a discount factor) or according to the condition β of the facility when the action is performed.

The goal is to find an appropriate policy π to describe how M&R should be performed on this facility, within a finite time horizon $\overline{T} \in \mathbb{Z}^+$ representing the number of discrete time periods in the analysis period; the length of each period, as well as the time horizon $\overline{T}$ itself depend on whether a short-term or long-term analysis is sought. As it turns out, the algorithms developed in this paper function equally well for either case; the difference between these time scales lies primarily in the measures of effectiveness used. As future conditions are unknown, a policy specifies which action (if any) is to be taken for each possible *state* the facility can exist in. For this problem, we define a state x as a 5-tuple (t, β, m, τ, e) indicating the current time t, reliability level β, the last maintenance action m performed, the time τ since this action was performed, and the amount of money e spent to date. Let X represent the set of all possible states for this facility. Then a policy π is a function mapping each state X to an element of U.

In the algorithms that follow, it is desirable for the state space to be as small as possible. In particular, expenditures should be measured in units allowing the costs c_u to be represented as small integers. For example, Tao et al. (1995) consider three possible bridge maintenance actions with per-girder costs of \$1,170, \$3,800, and \$8,400. In our model, one might opt to approximate these actions' costs as multiples of \$1,200, respectively giving costs of 1, 3, and 7 units. Defining the set E_t as the set of possible values of the expenditure e at time t, this choice of costs yields sets E_t of size $7t$. A better approximation could be obtained with a smaller unit of measurement (say, \$200), with a corresponding increase in the size of sets E_t. Define $c_{max} = \max_{\beta,t,u}\{c_u(\beta, t)\}$; this constant is useful for establishing the complexity results in the next section.

Regarding deterioration of the facility, we take the Markovian assumption, that the probability p_{ij} of moving from one state i to another j does not depend on any additional information; the current state suffices for knowing the transition probabilities. We also write $p_{i,j|u}$ to express the transition probabilities conditional on performing action u in state i. Under the assumption that these probabilities are stationary and do not depend on the to-date expenditures (except as captured by the current reliability level and the last maintenance action), we can specify the transition probabilities using a smaller set of states, namely $Y = B \times U \times T$, with $T = \{1, 2, \ldots, \overline{T}\}$ the set of periods within the time horizon. We assume that these probabilities are known *a priori*; these might be based directly on historical experience, or taken from simulation — Robelin and Madanat (2007) describes an algorithm to estimate these probabilities from a physical deterioration model, using Monte Carlo sampling.

2.2 Utility Function

To model different preferences, define a utility function $f(\beta, e)$ representing agency satisfaction for having a facility in condition β at the end of the time horizon $\overline{T}$, having spent e during that time. (In general, f could depend on the last maintenance action m which was performed τ periods ago; however, this may be a less common preference, and these indices are therefore suppressed for brevity.) We seek a policy π such that $E[f]$ is maximized. f can account for concepts such as salvage value, achieving M&R goals (e.g., meeting or exceeding a long-term reliability standard), total budget spent, and so on; the generality of the algorithms presented in this paper is based in the flexibility enabled by different definitions of f, possibly for different time scales. For this reason, we discuss several possible choices, and the goals they represent.

For the long-term budget planning problem, if the goal is to minimize expected expenditure while satisfying a minimum reliability level β_0, one might define $f(\beta, e) = -e$ for $\beta \geq \beta_0$ and $f(\beta, e) = -e - M$ otherwise, where $M > 0$ is a suitably large constant to penalize failure to meet the minimum reliability standard; the expected cost of the resulting policy can then be used in forming budget requests and forecasts. Another possible goal is to simply maximize the probability that the facility's reliability index is at least β_1, in which case $f(\beta, e) = 1$ for $\beta \geq \beta_1$, and 0 otherwise, suffices. These options can also be combined: $f(\beta, e) = -e + N$ for $\beta \geq \beta_1$, $-e$ for $\beta_0 \leq \beta < \beta_1$, and $-e - M$ for $\beta < \beta_0$, attempts to minimize expected cost, while penalizing failure to meet a minimum standard β_0 and rewarding achievement of a desired goal β_1, with M and N suitably defined.

Other options are available for finding short-term budget allocation policies. Although budget constraints are not always considered for single-facility problems, in the short term, they are highly relevant in practice. If the cost of a recommended action exceeds currently available funds, a suboptimal action may need to be performed instead, or maintenance deferred until the next budgeting period. Depending on the agency, money may be borrowed or emergency funding requested, but these actions are costly (the former directly, due to interest; the latter often indirectly, from a loss of trust or public image), and these costs should be reflected in the utility function.

If the budget b available until time $\overline{T}$ is known, adding a large penalty constant to f whenever $e > b$ discourages policies that exceed the budget.

Other options are to include a hard budget constraint by fixing $f(\beta, e) = -\infty$ for $e > b$ (effectively prohibiting any overspending), or to progressively penalize greater violations of the budget, such as by subtracting an additional term $(e-b)^2$ from f whenever $e > b$. Or, if budget rules allow unspent funds to be "rolled over" to the next time period, $f(\beta, e) = -e$ for $\beta \geq \beta_0$, $e \leq b$, $f(\beta, e) = -e - M$ for $\beta < \beta_0$, and $f(\beta, e) = -e - (e-b)^2$ for $e > b$ attempts to minimize present expenditures while meeting a minimum reliability level β_0 and satisfying the current budget, allowing greater future flexibility. The flexibility of the utility function f allows this type of short-term tactical spending decisions to be studied, along with the long-term, strategic planning that deterioration models usually consider.

2.3 Risk Preferences

Economists have suggested that risk preferences are of great importance in stochastic environments. Many proposed risk measurse can be incorporated within the utility function f. Perhaps the simplest option is to attempt to minimize the variance of utility. Since f must be completely specified *a priori*, one cannot directly minimize utility variance, which requires that the expected solution utility be known before the values of f can be calculated (since $Var[f] = E[(f - E[f])]$). However, an appropriate choice of utility function can ensure a reduction in utility variance, as compared to the policy π which simply maximizes $E[f]$:

Proposition 1. *Let θ be a policy maximizing the expected value of the utility function $-(f-f^*)^2$, where $f^* = E_\pi[f]$. Then $Var_\theta[f] \leq Var_\pi[f]$*

Proof.

$$\begin{aligned}
Var_\theta[f] &= E_\theta[f^2] - (E_\theta[f])^2 \\
&= E_\theta[f^2] - (E_\theta[f])^2 + (2f^* E_\theta[f] + (f^*)^2) - (2f^* E_\theta[f] + (f^*)^2) \\
&= E_\theta[(f-f^*)^2] - (E_\theta[f] - (f^*))^2 \\
&\leq E_\theta[(f-f^*)^2] \\
&\leq E_\pi[(f-f^*)^2] \text{ (since } \theta \text{ minimizes } E[(f-f^*)^2] \text{)} \\
&= Var_\pi[f]
\end{aligned}$$

□

Essentially, this approach minimizes the deviation from a specified "target" utility value f^*. By choosing f^* to be the expected utility from a utility-maximizing policy, Proposition 1 ensures a reduction in utility variance, while also avoiding the problem that directly minimizing variance may result in a very poor expected utility. The mean-variance technique, developed by Markowitz (1952) in the finance domain, proposes an alternate way of avoiding this problem, by maximizing a weighted difference of expected utility and utility variance. The same approach can be applied here, by choosing the utility function $f - \lambda(f-f^*)^2$, with λ an appropriately-chosen weighting constant.

Asymmetric preferences are often observed. For instance, overspending the budget by a given amount is much worse than underspending by the same amount; or missing a reliability target for a facility is worse than exceeding the target. This can be represented directly in the f function as described above, or through a utility function of the form $[f - f^*]^+ - \mu[f^* - f]^+$, with $[\cdot]^+ = \max\{\cdot, 0\}$, and μ a weighting constant describing how falling short of the target utility f^* should be penalized, relative to exceeding it.

Prospect theory provides yet another approach for modeling preferences under uncertainty. Proposed by Kahneman and Tversky (1979), this approach attempts to describe behavior through a two-phase process: first, possible outcomes are ranked, and a reference point is chosen. The desirability of each outcome is determined by an S-shaped value function, based on its desirability relative to the reference point. This function is usually chosen to represent loss aversion, where outcomes worse than the reference point are penalized more heavily. Next, an action is chosen maximizing the expected value of the value function. Given a reference utility f^* and a value function $v(\beta, e, f^*)$, this approach can be modeled in our framework by choosing our utility function to coincide with the value function for all possible outcomes.

Thus, the function f allows a very flexible representation of agency preferences under uncertainty. One should note that previous Markovian policy models do not include the current expenditures e in the state definition; while this is not necessary when preferences are linear, it is important to include it for nonlinear preferences. An intuitive reason for this is given by considering spending behavior with a budget constraint: one should be allowed to differentiate between the cases when one has already spent most of the budget, and the case where one has only spent a small portion of it. Since such distinctions are possible in reality, and relevant with nonlinear objectives, current expenditure must be included in the state definition.

3 Solution Methods

In this section, we describe exact algorithms to find a policy π minimizing $E[f]$ (FindPolicy) and to evaluate a given policy according to a metric defined by another utility function g (EvalPolicy). Both FindPolicy and EvalPolicy apply dynamic programming techniques originated by Bellman (1957). Dynamic programming can be used to solve discrete problems satisfying certain properties, namely, optimal substructure and overlapping subproblems. Optimal substructure requires that components of an optimal solution should also optimize smaller subproblems; for instance, for the problem of finding an optimal M&R policy, consider some optimal policy for times $1, 2, \ldots, \overline{T}$. If we only consider the portions of this policy for, say, $T \geq 2$, this should itself be an optimal policy for times $2, 3, \ldots, \overline{T}$: if it were not, we could update some part of both this and the original policy as well, obtaining better solutions for both.

The overlapping subproblem criterion essentially requires that the problem be solvable sequentially, often by working backwards. When finding an optimal M&R policy, one can start at the last time period $\overline{T}$, and consider the optimal policy for all the states corresponding to the previous period $\overline{T} - 1$. Since there is only one time period left, knowing the value of the facility in each possible final state, it is relatively easy to find the optimal actions at period $\overline{T} - 1$. Knowing this, one proceeds to find the optimal actions at $\overline{T} - 2$, only having to look at $\overline{T} - 1$, because the final period has already been accounted for.

3.1 Algorithm FindPolicy

FindPolicy determines the optimal action to take for every possible state. This is done by working backward from the final time interval, so at every point one can assume that the future actions and outcomes have been determined for every possible sequence of states the facility will follow as it deteriorates and is repaired.

To assist in this process, we associate two labels L and π with each state $x = (t, \beta, m, \tau, e) \in X$, where L denotes the expected value of the objective function, assuming that the policy developed thus far will be implemented until the time horizon, and π stores the optimal action to take at state x. Taken together, the labels $\pi(x)$ define the policy we seek.

FindPolicy is initialized by setting final utility levels $L(x) = f(\beta, e)$ for all states $x = (\overline{T}, \beta, m, \tau, e)$ corresponding to the final time interval. Given these final utility values and the deterioration model, the optimal actions during time interval $\overline{T} - 1$ can be found: for each state, calculate the expected utility that would be obtained from applying each possible action; the action maximizing that expected utility is optimal, and assigned to that state.

Now, consider an arbitrary time interval t, and assume that these labels have been found for all time intervals $t+1, \ldots, \overline{T}$, and we want to set the labels for the states associated with time t. For each possible action u, we can calculate temporary labels e^u denoting the expected objective function value if the policy were to use action u in this state. After doing this for all actions, choose the action u^* maximizing L^u, set the label L equal to L^{u^*}, and set μ to u^*. Denoting (β, m, τ) by y, the temporary label can be calculated as:

$$L^u = \sum_{z \in Y} p_{y,z|u} L(t+1, z, e + c_u(\beta, t)) \tag{1}$$

This formula accomplishes the calculation needed for the backward recursion. The algorithm thus considers time intervals $\overline{T}-1, \overline{T}-2$, and so on, in turn before terminating at the first time period with the optimal policy fully specified. This is presented formally below, as Algorithm 1.

From a computational standpoint, one does not need to consider *all* states $z \in Y$ when evaluating the sum in (1) , since many of the transition probabilities $p_{y,z|u}$ must be zero. For instance, if one is performing action u at time t, the succeeding state will certainly have $m = u$ for the last maintenance action, and $\tau = 1$ for the time since that action was performed: all transition probabilities to states with different m and τ values must be zero. Taking

Algorithm 1 FindPolicy

```
{Initialization}
for all t ∈ {1, 2, ..., T̄}, y = (β, m, τ) ∈ Y, e ∈ E_t do
    if t = T̄ then
        L(T̄, y, e) ← f(β, e)
    else
        L(T̄, y, e) ← −∞
    end if
end for
{Iteration}
for all t = T̄ − 1, T̄ − 2, ..., 1 do
    for all y ∈ Y, e ∈ E_t do
        for all u ∈ U do
            L^u ← Σ_{z∈Y} p_{y,z|u} L(t + 1, z, e + c_u(β, t))
            if L^u(t, y, e) > L(t, y, e) then
                L(t, y, e) ← L^u(t, y, e)
                π(t, y, e) ← u
            end if
        end for
    end for
end for
return π
```

this into account, the time and space complexity of FindPolicy are as follows:

Proposition 2. *Algorithm FindPolicy can be implemented to require* $O(\overline{T}^3|B||U|c_{max})$ *space and run in* $O(\overline{T}^3|B|^2|U|^2c_{max})$ *time.*

Proof. FindPolicy stores two variables for each state x. As the five dimensions of X are of size $\overline{T}$, $|B|$, $|U|$, $O(\overline{T})$, and $\overline{T}c_{max}$, respectively, storing these labels requires $O(\overline{T}^3|B||U|c_{max})$ space.

The loops in lines 10–12 of Algorithm 1 require the summation in line 13 to be performed $O(\overline{T}^3|B||U|^2c_{max})$ times. As depicted, each summation requires adding $O(\overline{T}|B||U|)$ terms; however, since the only unknown component of z is the facility reliability in the next stage, this term can be calculated by only adding $|B|$ terms. The overall complexity immediately follows. $\square$

3.2 Algorithm EvalPolicy

The second algorithm, EvalPolicy, allows a specified policy π to be evaluated according to a given measure of effectiveness (not necessarily the same one which π optimizes). For instances, although π may have been chosen to minimize expected cost, an agency may be interested in knowing the probability that the actual expenditures under this policy exceed some threshold value, or facility condition exceeds a certain target level; or in calculating the standard deviation in expenditures. Even if none of these are the primary goal, EvalPolicy allows policies to be compared using these (or other) metrics.

EvalPolicy uses a forward recursion: given an initial (current) state for the facility, and the values of the desired measure of effectiveness for all states at the final time $\overline{T}$, the probability of each final state is calculated, allowing direct calculation of the chosen metric's expected value. We denote this metric by $g(\beta, m, \tau, e)$. Although serving a similar purpose to f in FindPolicy, we distinguish the notation for two reasons: first, we intend to clarify that the policy π need not maximize g; second, it is quite plausible that this metric may be a function of the state elements m and τ at the end of the time horizon, in contrast to FindPolicy, where we suggest the opposite.

For each state $x \in X$, we store a label $\rho(x)$ denoting the probability that state x occurs if policy π is followed; clearly we have $\sum_{\beta,m,\tau,e} \rho(t, \beta, m, \tau, e)$ for all $t \in 1, 2, \ldots, \overline{T}$. Thus, for a state $x = (t, \beta, m, \tau, e)$ we calculate

$$\rho(x) = \sum_{y \in Y} \sum_{e \in E_{t-1}} \rho(t-1, y, e) p_{y,(\beta,m,\tau)|\pi(t-1,y,e)}$$

if we know the probabilities of each state occuring during the preceding time interval. Thus, beginning with $\rho(x_0) = 1$ for the initial state x_0, we proceed in increasing order of time to calculate these probabilities for all states. Upon completion, calculation of the expected value of g is straightforward. This is formally presented as Algorithm 2.

Proposition 3. *Algorithm EvalPolicy can be implemented to require* $O(\overline{T}^2|B||U|c_{max})$ *space and run in* $O(\overline{T}^3|B|^2|U|c_{max})$ *time.*

Proof. As shown in Algorithm 2, EvalPolicy stores a label for each state x. Although $|X|$ is $O(\overline{T}^3|B||U|c_{max})$ as shown in the proof of Proposition 2, at

Algorithm 2 EvalPolicy

```
1: {Initialization}
2: for all $x \in X$ do
3:    if $x = x_0$ then
4:       $\rho(x) \leftarrow 1$
5:    else
6:       $\rho(x) \leftarrow 0$
7:    end if
8: end for
9: {Iteration}
10: for all $t = 1, \ldots, \overline{T} - 2, \overline{T} - 1$ do
11:    for all $y = (\beta, m, \tau) \in Y$, $e \in E_t$ do
12:       for all $z \in Y$ do
13:          $\rho(t+1, z, e + c_{\pi(t,y,e)}(\beta)) \leftarrow \rho(t+1, z, e + c_{\pi(t,y,e)}(\beta)) + \rho(t, y, e) p_{y,z|\pi(t,y,e)}$
14:       end for
15:    end for
16: end for
17: return $\sum_{y \in Y, e} \rho(\overline{T}, y, e) g(y, e)$
```

any point the algorithm only uses labels from the current time interval and its immediate predecessor, allowing reuse of storage space and reducing the space complexity by a factor of $\overline{T}$.

The loops in lines 10 and 11 require lines 12–14 to be repeated $O(\overline{T}^3 |B||U| c_{max})$ times. Using the same logic as in the proof of Proposition 2, only $|B|$ states need to be considered in the innermost loop, providing the time complexity. $\square$

While the Markovian deterioration model is useful for finding optimal policies, one may wish to evaluate policies using a more accurate model. This may occur if the Markov transition probabilities are taken from a simplified version of another deterioration model (for instance, estimated from repeated simulations). A Monte Carlo approach can then be used to simulate the facility's condition as the maintenance policy is applied, and estimate the desired measures of effectiveness. If the Markov probabilities are only approximate, this approach may yield more accurate estimates; however, this comes at a cost in computation time, since it lacks the efficient forward recursion and exact computation of expected value found in EvalPolicy.

4 Numerical Example

In this section, we demonstrate how FindPolicy and EvalPolicy can be applied in determining M&R policies, both in the long-term expenditure planning and short-term budget allocation contexts. As an example, we consider a bridge facility which exists in one of fifteen reliability levels or conditions ($B = \{1, 2, \ldots, 15\}$). Maintenance decisions are scheduled according to six-month intervals, and budgets are determined biennially. The minimum allowable reliability index is three; we control for this in the model by identifying all of the states corresponding to a lower condition, and setting their f label to $-\infty$. The deterioration model and parameters are taken from Frangopol et al. (2001), and the transition probabilities are estimated using the procedure in Robelin and Madanat (2007).

In any given year, three possible actions can be performed (in addition to the "do nothing" option). **Minor repairs**, representing routine maintenance, costing between 1 and 3 units depending on the reliability level

($c_1 = 2 - \lfloor \beta/5 \rfloor$), increase the reliability level by γ_1, which is drawn from a lognormal distribution with a mean and standard deviation of two reliability units. **Major repairs** represent a more involved repair action: the cost is three times that of minor repairs, and the reliability is increased by γ_2, which is drawn from a lognormal distribution with mean seven, and standard deviation of three. The third action, **replacement**, has a cost of 21, and restores the facility to new condition ($\beta = 15$). This choice of actions, along with their costs and effects, follows Tao et al. (1995).

We first consider the long-term planning decision, assuming a 20-year time horizon. The long-term goal is to have this facility have a reliability index equal to 12 or greater at the end of the time horizon, and the agency wishes to know the minimum expected cost needed to satisfy this condition. However, recognizing that this cost may be prohibitively high depending on the initial state of the bridge, this long-term goal is not enforced as a hard constraint. Rather, the agency decides that achieving this goal is "worth" five units of cost; that is, it is willing to spend up to five additional units on maintenance actions if doing so would result in $\beta \geq 12$. Thus, the appropriate choice of utility function is $f_1(\beta, e) = -e$ for $\beta < 12$ and $f(\beta, e) = 5 - e$ for $\beta \geq 12$. With this choice of time horizon and model parameters, the state space X consists of 3,969,000 elements. Implemented on a 2.6 GHz Pentium III machine using Windows XP with 4 GB RAM, FindPolicy requires 12.7 seconds of computation time, on average, and EvalPolicy requires 0.19 seconds.

Algorithm FindPolicy is then applied to determine the optimal policy, and EvalPolicy is used multiple times to evaluate this policy according to several metrics: the expected value and standard deviation of expenditures, the expected value and standard deviation of the reliability index at the end of the time horizon, the probability of achieving the reliability goal, and the expected value and variance of utility. These results are shown in Table 1 for different possible initial values for bridge reliability (denoted β_0). (In all cases, it is assumed that minor repairs were performed in the year prior to analysis.) Figure 1 plots the reliability over time for one particular realization of this policy, indicating when maintenance is performed; since the deterioration model is stochastic, naturally this plot only represents one possible outcome.

As an alternative, the agency is interested in a policy which is "robust" to future uncertainty in some sense. To accomplish this, FindPolicy is used a second time, with $f_2(\beta, e) = -(f_1(\beta, e) - f_1^*)^2$, with f_1 as before and f_1^* equal to the expected value of f_1 under the first policy. This choice of utility function is appropriate because it penalizes deviation from the average outcome, and thus encourages a policy with smaller variance in utility. Following this, EvalPolicy is then applied using the same measures of effectiveness; these results are shown in Table 2. Note that, for each initial reliability value, the "robust" policies have a slightly lower expected utility ($E[f]$), but substantially lower standard deviation of utility ($\sigma[f]$), indicating that policies with more predictable outcomes are readily available and are not significantly worse, in terms of expected utility. (Average expenditures are much higher; but, according to the specified utility function, this is nearly balanced by a corresponding increase in facility condition.) Also note that the expected reliability and probability of attaining the goal are higher for the "robust" policies, along with expected expenditures; an intuitive interpretation of this result is spending more to improve facility condition is less affected by future uncertainty than the first policy, which solely sought to maximize expected utility. (Of course, this result could vary with a different utility function or problem parameters.)

Figure 2 plots reliability over time for one realization of a "robust" policy, providing graphical demonstration of how higher average expenditures for this policy translate into less variation in utility. By completely replacing the facility when reaching a degraded condition and then performing routine maintenance, less variation is seen in the final outcome, as compared to the utility-maximizing policy, where the replacement action is much rarer. Intuitively, the "robust" policy plans for a replacement expenditure at some point in the time horizon, while the utility-maximizing policy does not; the result is higher average expenditure under the former, but less unpredictability in both expenditure and final facility condition.

These results are typical of the utility-maximizing and "robust" policies. Figures 3 and 4 plot the expected expenditure and final reliability for these policies, showing how they vary according to the initial reliability. For comparison, the cost-minimizing and reliability-maximizing policies are also plotted. One can observe that these policies adopt a compromise between these two objectives, achieving higher reliability than the cost-minimizing policy would, and requiring lower expenditure than the reliability-maximizing policy.

As a final long-term planning consideration, recall that EvalPolicy uses the Markovian deterioration model to evaluate policies. Since the Markovian transition probabilities were estimated from a deterioration model from Frangopol et al. (2001), it is worthwhile to compare EvalPolicy's results with that of simulation using the latter model. Table 3 shows the expected utility for the utility-maximizing and "robust" policies, both as calculated by EvalPolicy, and as estimated through simulation based on five thousand Monte Carlo trials for each initial reliability level. The results are mostly consistent, although EvalPolicy seems to have greater difficulty matching the simulated results for very low initial reliability in the case of the utility-maximizing policy. Excluding these, the root-mean-square difference between expected utilities is slightly greater than 1. Given the utility specification, this is approximately equivalent to the cost of one minor repairs. Considering the twenty-year time horizon, this does not seem unreasonable.

After considering these long-term policies, a budget recommendation is made. Now, assume that the actual amount of money allocated for the first two years is five units, which is higher than expected. The relevant short-term question is how to make the best use of this funding while not exceeding the assigned budget. Three possible utility functions achieving this purpose are $f_3(\beta, e) = \beta$ for $e \leq 5$ and $\beta - 10$ otherwise (assigning a penalty of ten reliability units for exceeding the budget); $f_4(\beta, e) = \beta$ for $e \leq 5$ and $-\infty$ otherwise (effectively eliminating the possibility of overspending); and $f_5(\beta, e) = \beta$ for $e \leq 5$ and $\beta - 2(e-5)^2$ otherwise (penalizing greater violation of the budget by progressively larger amounts). For each of these, FindPolicy is applied (using a two-year time horizon) to determine the optimal policy for each objective, and EvalPolicy is used to determine the expected value and standard deviation of expenditures and reliability at the end of this budget period; these are shown in Table 4, and possible realizations for these policies are plotted in Figures 5 to 7. This choice of a time horizon results in a state space consisting of 37,800 elements, and the algorithms execute in less than a tenth of a second.

Note that the budget is frequently exceeded if the initial reliability index is very low, even under objective f_4 which supposedly prevents exceeding the budget in an absolute sense; the implication is that it is impossible to satisfy this requirement while also ensuring that the bridge does not fall below the minimum reliability index standard. In this demonstration, this conflict was resolved by relaxing the budget constraint, rather than the minimum reliability constraint — in terms of implementation, the $-\infty$ quantities are replaced by large negative numbers, and the $-\infty$ corresponding to the budget constraint was chosen to be smaller in magnitude than the $-\infty$ corresponding to the reliability constraint. Also note that, for policy f_3, the expected final reliability indices are higher for $\beta_0 \in \{3, 4\}$ than for $\beta_0 \in \{5, 6, 7\}$. This occurs because, for utility function f_3, violation of the budget by any amount incurs the same penalty; thus, once more than five cost units have been spent, there is no penalty for greater spending, indicating that this utility function may not be the best way to account for the budget.

Although this example is clearly a great simplification of the budget planning and allocation process, it nevertheless demonstrates how FindPolicy and EvalPolicy can be used to support agencies in making such decisions, providing quantitative analysis of a variety of objectives.

5 Conclusion

In this paper, we present techniques for finding and evaluating M&R policies which account for certain types of nonlinear agency preferences regarding maintenance expenditures, including behaviors such as risk aversion, goal setting, and target seeking, both in the long-term (budget planning) and short-term (budget allocation) horizons. Using the framework of a history-dependent, Markovian deterioration model, two dynamic programming-based algorithms are developed to find optimal policies with respect to these preferences, and to evaluate such policies according to a wide variety of measures of effectiveness. A hypothetical bridge was used as an example, indicating how both short-term and long-term policies can be derived and compared using these algorithms.

Still, additional field data should be applied to see if these conclusions drawn from a fictitious example can be extended to general practice. Another useful extension of this work is to consider the macroscopic, network-level optimization problem, where multiple facilities must be maintained, under a common budget or other bundling constraint.

Acknowledgements

The authors would like to thank three anonymous referees for their valuable comments and suggestions. Of course, the authors remain solely responsible for the contents of the paper.

Notation

c_{max}	Cost of most expensive action, in appropriate units
c_u	Cost of a maintenance action
e	Expenditures to-date
f	Utility function representing agency preferences
g	Utility function used to evaluate a policy
m	Last maintenance action performed
$p_{i,j\|u}$	Probability of progressing from state i to state j if maintenance action u is performed
t	Current time
x	A facility state (including t, β, m, τ, and e)
E_t	Set of possible expenditures at time t
L	Labels used in algorithm FindPolicy
T	Set of time indices
$\overline{T}$	Time horizon
U	Set of possible maintenance actions
VOR	Value of reliability parameter used to generate properly efficient policies
Y	Smaller set of states used to represent transition probabilities
X	Set of all possible facility states
β	Facility reliability index
π	A maintenance and rehabilitation policy (formally, a function mapping X to U)
τ	Time since last maintenance action was performed
B	Set of possible reliability indices
$\mathbb{Z}^+$	Set of non-negative integers

References

Bellman, R. (1957). *Dynamic Programming.* Princeton, NJ: Princeton University Press.

Ben-Tal, A. and A. Nemirovski (2002). Robust optimization — methodology and applications. *Mathematical Programming, Series B 92*, 453–480.

Bertsimas, D. and M. Sim (2004). The price of robustness. *Operations Research 52*(1), 35–53.

Duffie, D. and J. Pan (1997). An overview of value-at-risk. *Journal of Derivatives 4*, 7–49.

Frangopol, D. M., J. S. Kong, and E. S. Gharaibeh (2001). Reliability-based management of highway bridges. *Journal of Computing in Civil Engineering 15*(1), 27–34.

Gao, L. and Z. Zhang (2006). Robust optimization for managing pavement maintenance and rehabilitation. In *Proceedings of the 85th Annual Meeting of the Transportation Research Board*, Washington, DC.

Gao, L. and Z. Zhang (2008). Robust optimization for managing pavement maintenance and rehabilitation. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, DC.

Kahneman, D. and A. Tversky (1979). Prospect theory: an analysis of decision under risk. *Econometrica 47*, 263–291.

Kong, J. S. and D. M. Frangopol (2001). Life-cycle reliability-based maintenance cost optimization of deteriorating structures with emphasis on bridges. *Journal of Structural Engineering 129*(6), 818–828.

Kuhn, K. D. and S. M. Madanat (2005). Model uncertainty and the management of a system of infrastructure facilities. *Transportation Research Part C 13*, 391–404.

Kuhn, K. D. and S. M. Madanat (2006). Robust maintenance policies for Markovian systems under model uncertainty. *Computer-Aided Civil and Infrastructure Engineering 21*, 171–178.

Li, Z. and M. Puyan (2006). A stochastic optimization model for highway project selection and programming under budget uncertainty. In *Proceedings of the 9th International Conference on Applications of Advanced Technology in Transportation*, Chicago, IL. American Society of Civil Engineers.

Liu, H. X., X. He, and W. Recker (2007). Estimation of the time-dependency of values of travel time and its reliability from loop detector data. *Transportation Research Part B 41*(4), 448–461.

Liu, H. X., W. Recker, and A. Chen (2004). Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation Research Part A 38*(6), 435–453.

Madanat, S. and M. Ben-Akiva (1994). Optimal inspection and repair policies for infrastructure facilities. *Transportation Science 28*(1), 55–62.

Madanat, S., S. Park, and K. Kuhn (2006). Adaptive optimization and systematic probing of infrastructure system maintenance policies under model uncertainty. *Journal of Infrastructure Systems 12*(3), 192–198.

Marfitano, S. C. (2008). Robust optimization for budget allocation at network level using clustering techniques. Master's thesis, The University of Texas at Austin.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance 7*(1), 77–91.

Mori, Y. and B. R. Ellingwood (1994). Maintaining reliability of concrete structures. II: Optimum inspection/repair. *Journal of Structural Engineering 120*(3), 846–862.

Pinjari, A. and C. Bhat (2006). On the nonlinearity of response to level of service variables in travel mode choice models. *Transportation Research Record 1977*, 67–74.

Robelin, C.-A. (2006). *Facility-level and system-level stochastic optimization of bridge maintenance and replacement decisions using history-dependent models*. Ph. D. thesis, University of California, Berkeley.

Robelin, C.-A. and S. M. Madanat (2007). History-dependent bridge deck maintenance and replacement optimization with markov decision processes. *Journal of Infrastructure Systems 13*(3), 195–201.

Robelin, C.-A. and S. M. Madanat (2008). Reliability-based system-level optimization of bridge maintenance and replacement decisions. *Transportation Science 42*(4), XXX–XXX[1]

Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–41.

Small, K. A., C. Winston, and J. Yan (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica 73*(4), 1367–1382.

Smilowitz, K. and S. Madanat (2000). Optimal inspection and maintenance policies for infrastructure networks. *Computer-Aided Civil and Infrastructure Engineering 15*, 5–13.

Tao, Z., R. B. Corotis, and J. H. Ellis (1995). Reliabiltiy-based structural design with Markov decision processes. *Journal of Structural Engineering 121*(6), 971–980.

[1]This paper will appear in the November 2008 issue of Transportation Science; page numbers not currently available.

List of Tables and Figures

Table 1: Comparison of utility-maximizing long-term policies

β_0	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$\Pr(\beta \geq 12)$	$E[f]$	$\sigma[f]$
3	13.868	9.454	10.869	4.868	0.614	-16.000	19.467
4	9.985	8.819	9.439	4.807	0.484	-13.962	20.093
5	6.498	6.192	8.770	4.432	0.428	-11.094	16.534
6	5.121	4.276	9.096	4.296	0.473	-5.597	11.916
7	4.990	3.079	11.009	3.797	0.697	-3.928	8.401
8	4.571	2.359	12.463	2.661	0.866	-2.600	5.403
9	3.699	2.023	13.131	1.716	0.941	-1.222	3.701
10	2.895	1.662	13.353	1.336	0.962	-0.158	2.915
11	2.287	1.461	13.398	1.277	0.965	1.537	2.036
12	1.806	1.342	13.401	1.277	0.965	2.020	1.950
13	1.301	1.225	13.398	1.277	0.965	2.525	1.872
14	0.776	1.090	13.392	1.278	0.965	3.049	1.789
15	0.671	1.040	13.391	1.278	0.965	3.614	1.545

Table 2: Comparison of "robust" long-term policies

β_0	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$\Pr(\beta \geq 12)$	$E[f]$	$\sigma[f]$
3	23.014	0.566	14.294	2.901	0.952	-18.250	1.793
4	20.026	1.457	14.832	1.009	0.971	-15.170	2.002
5	16.32	2.943	14.271	2.171	0.892	-11.861	4.376
6	9.391	3.738	11.461	4.147	0.555	-6.615	5.012
7	7.272	3.097	11.533	4.297	0.584	-4.354	3.887
8	7.168	2.065	13.229	3.352	0.78	-3.269	2.796
9	5.951	1.188	14.293	2.089	0.935	-1.276	2.160
10	4.830	1.289	14.445	1.516	0.938	-0.141	1.612
11	3.356	0.736	14.576	0.773	1.000	1.642	1.058
12	3.190	0.559	14.617	0.757	0.999	1.807	0.954
13	2.667	0.705	14.626	0.733	1.000	2.331	0.772
14	2.185	0.549	14.575	0.783	0.999	2.813	0.490
15	1.456	0.818	14.354	0.891	0.999	3.540	0.178

Table 3: Comparison of EvalPolicy and simulated deterioration

Initial	Utility maximizing policy			"Robust" policy		
Reliability	EvalPolicy	Simulation	Diff	EvalPolicy	Simulation	Diff
3	−16.000	−11.834	−4.166	−18.250	−18.472	0.222
4	−13.962	−8.442	−5.520	−15.170	−16.290	1.120
5	−11.094	−5.735	−5.359	−11.861	−12.759	0.898
6	−5.597	−3.950	−1.647	−6.615	−8.005	1.390
7	−3.928	−2.061	−1.867	−4.354	−5.207	0.853
8	−2.600	−1.121	−1.479	−3.269	−3.990	0.721
9	−1.222	−0.649	−0.573	−1.276	−2.285	1.009
10	−0.158	−0.053	−0.105	−0.141	−1.169	1.028
11	1.537	0.498	1.039	1.642	0.272	1.370
12	2.020	1.245	0.775	1.807	1.223	0.584
13	2.525	1.983	0.542	2.331	1.524	0.807
14	3.049	2.782	0.267	2.813	1.198	1.615
15	3.614	3.037	0.577	3.540	1.756	1.784

Table 4: Comparison of short-term policies

Initial Reliability	f_3				f_4				f_5			
	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$	$E[e]$	$\sigma[e]$	$E[\beta]$	$\sigma[\beta]$
3	19.44	8.115	13.571	2.837	6.73	7.436	6.967	4.230	5.790	0.535	6.685	2.504
4	14.766	9.496	12.042	3.507	4.236	2.440	6.777	2.986	5.514	0.500	7.707	2.516
5	6.710	5.437	9.365	2.921	4.820	0.572	8.434	2.613	5.090	0.286	8.624	2.461
6	4.295	0.883	9.872	2.753	4.267	0.455	9.859	2.752	6.000	0	12.169	2.151
7	4.417	0.519	11.114	2.694	4.417	0.519	11.114	2.694	5.137	0.768	11.872	2.284
8	4.634	0.537	12.551	2.387	4.634	0.537	12.551	2.387	4.867	0.636	12.799	2.131
9	4.816	0.504	13.901	1.707	4.816	0.504	13.901	1.707	4.85	0.516	13.938	1.630
10	4.853	0.513	14.752	0.813	4.853	0.513	14.752	0.813	4.856	0.513	14.755	0.800
11	5.000	0	14.998	0.051	5.000	0	14.998	0.051	5.000	0	14.998	0.051
12	3.790	1.169	15	0.017	3.790	1.169	15	0.017	3.790	1.169	15	0.017
13	3.182	1.076	15	0.004	3.182	1.076	15	0.004	3.182	1.076	15	0.004
14	3.023	0.496	15	0	3.023	0.496	15	0	3.023	0.496	15	0
15	1.216	0.626	15	0	1.216	0.626	15	0	1.216	0.626	15	0

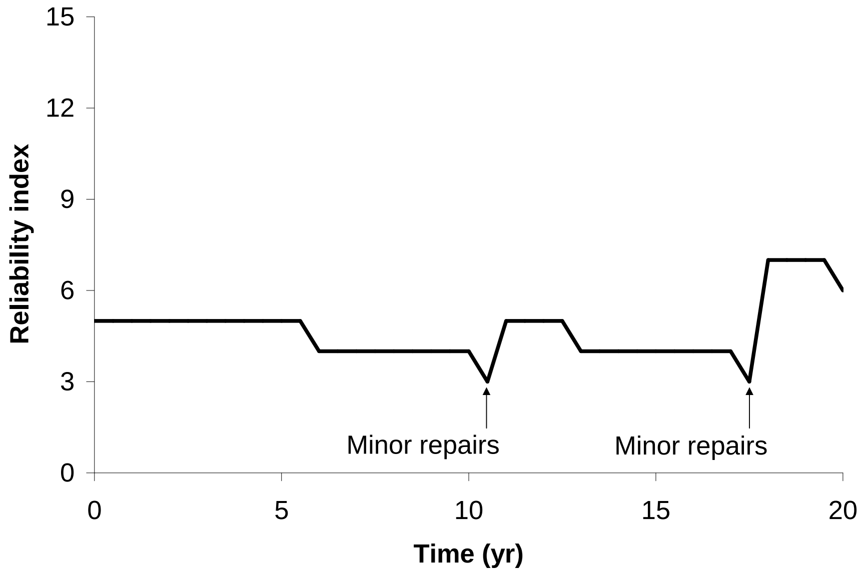

Figure 1: One realization of a long-term, utility-maximizing policy

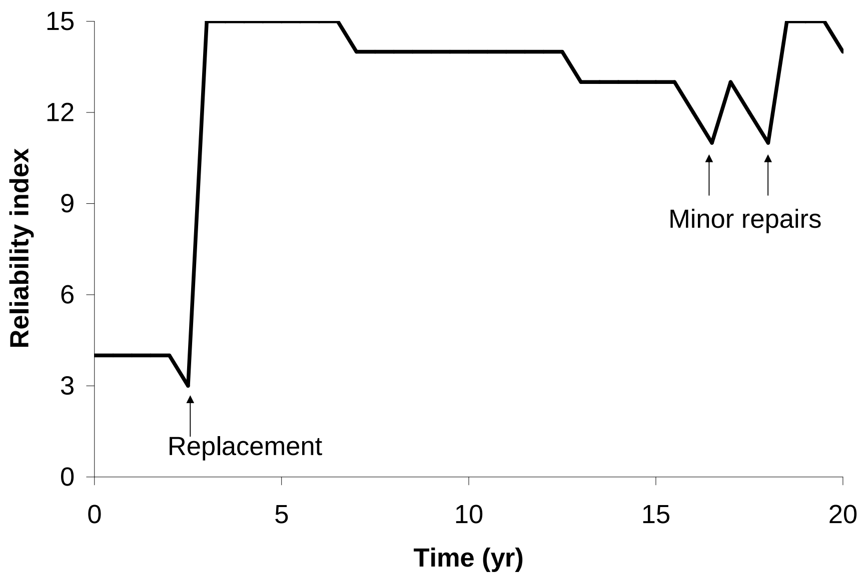

Figure 2: One realization of a long-term, “robust” policy

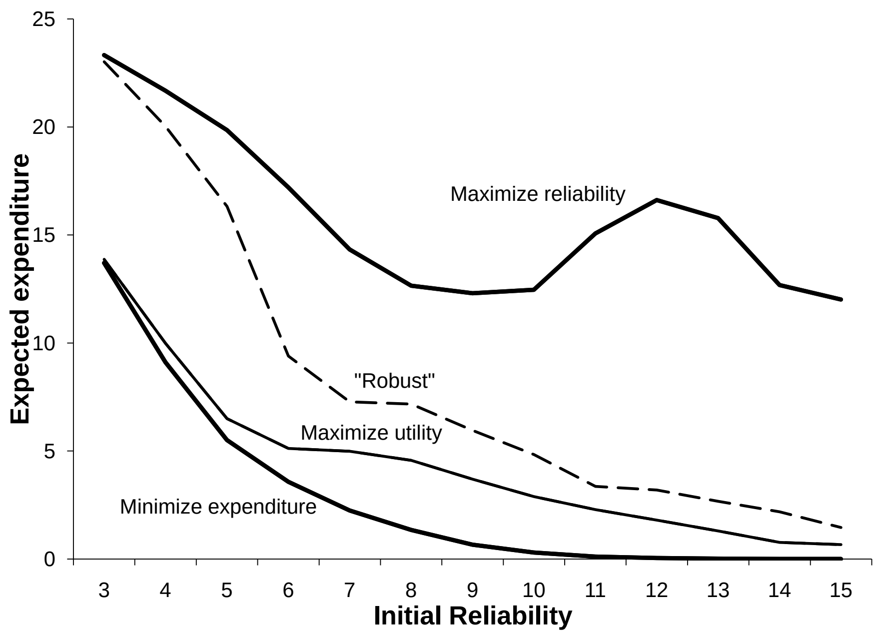

Figure 3: Expected expenditure for four policies

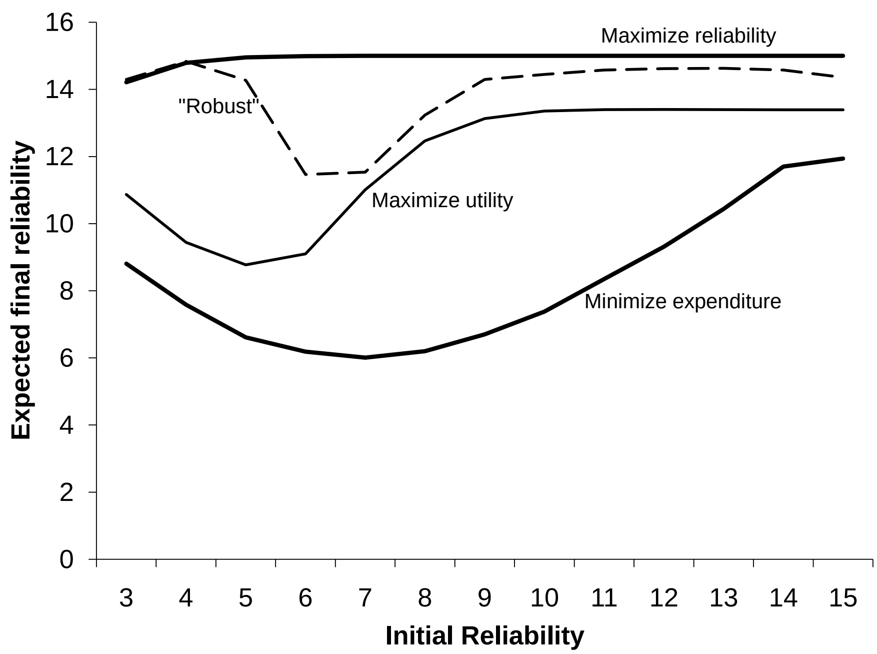

Figure 4: Expected final reliability for four policies

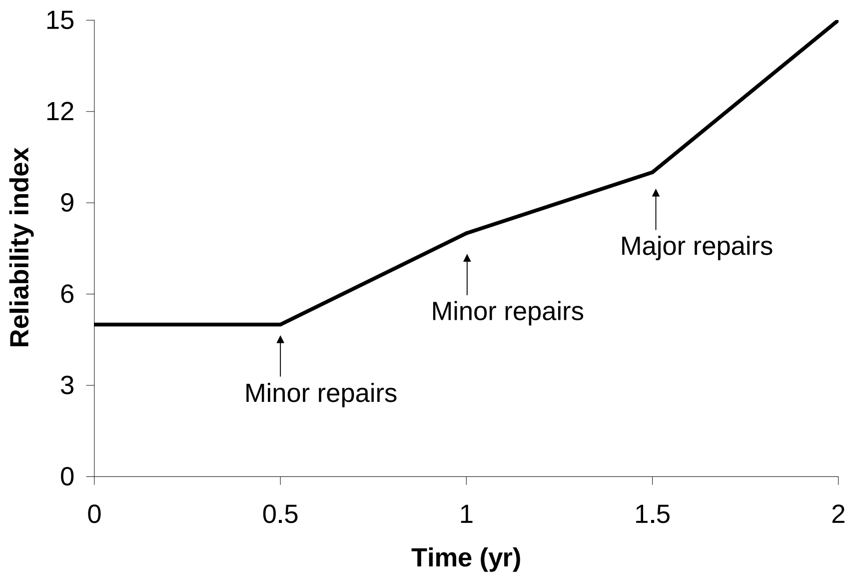

Figure 5: One realization of a short-term policy with a constant penalty for exceeding the budget (f_3)

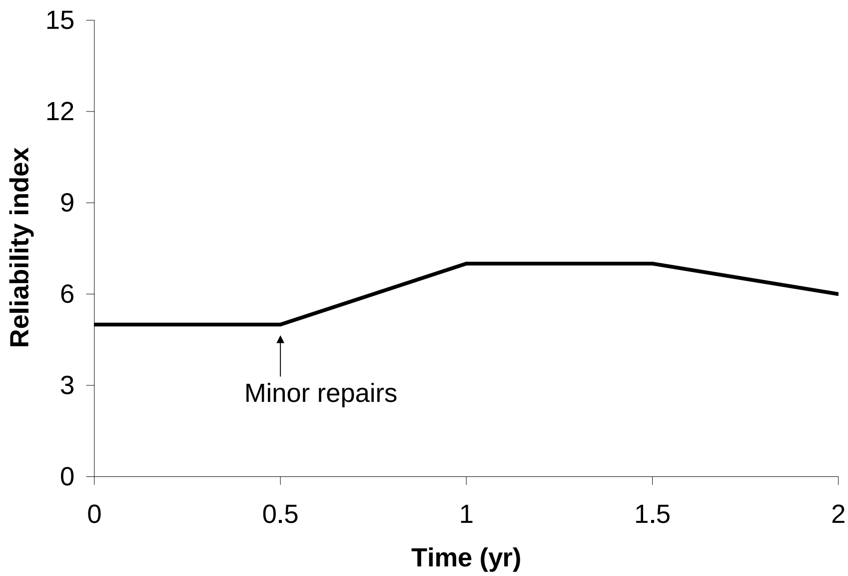

Figure 6: One realization of a short-term policy with a hard budget constraint (f_4)

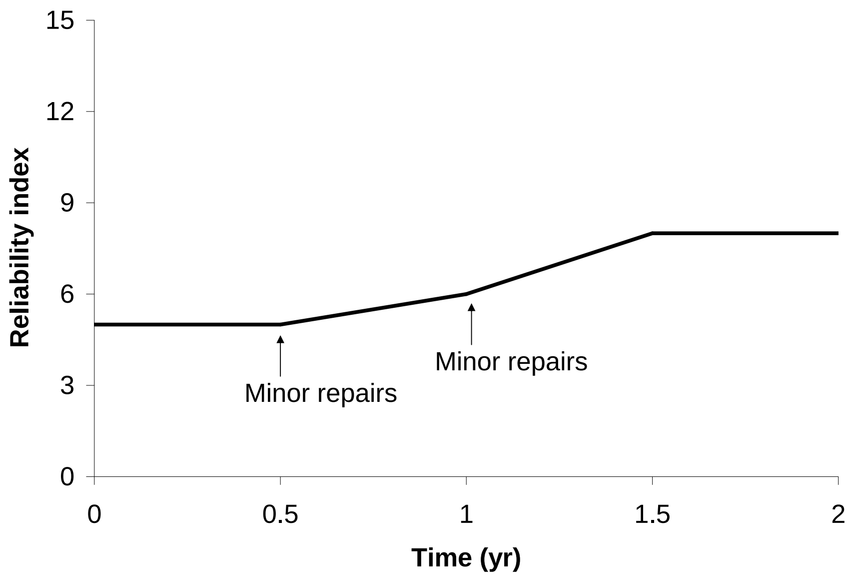

Figure 7: One realization of a short-term policy with a variable penalty for exceeding the budget (f_5)