University
of Southampton

Discussion Papers in
Management

To Train or To Repair? Training and Repair Policies for Stand-by Systems

Yeek-Hyun Kim

and

Lyn Thomas

Number M04-17
March 2004
ISSN 1356-3548

To Train or To Repair? Training and Repair policies for stand-by systems

by

Yeek-Hyun Kim, Lyn Thomas

Abstract

This research is concerned with developing repair and training strategies for standby equipment which maximise the time until the equipment is needed and it is unable to respond. Equipment can only be used if it is in an operable state and the users have had sufficient recent training on it. Thus it is necessary to decide when to maintain/repair the equipment and when to use the equipment for training. Both actions mean the equipment is not readily available for use in an emergency. We develop discrete time Markov decision process formulations of this problem in order to investigate the form of the optimal action policies which maximise the expected survival time until a catastrophic event when an emergency occurs and the equipment cannot respond. We also calculate the solution in a number of numerical examples.

1 Introduction

Stand-by equipment is only brought into operation when there is a vital need for it, for example, a hospital emergency power supply system, military equipment etc. We call the times when there is a vital need for a stand-by unit on initiating events and if the unit is not able to respond to an initiating event then it is deemed a catastrophic event.

This research is concerned with developing repair and training strategies which maximise the time until a catastrophic event for standby units in an uncertain environment. Equipment can only be used if it is in an operable state and if its users have had sufficient recent training with it. Thus as well as repairing and maintaining the equipment, it is necessary to train users. This is particularly clear in the military context where soldiers are constantly trained to operate the equipment satisfactorily under all conditions. However, a problem with training is that it increases the wear and tear of the stand-by unit even though it enhances the operator's ability to respond well to an initiating event. Another problem in the military context is that the training may be done away from where the equipment may be needed and so there is not time to move it between the training area and the front line say. In this research we look at the interaction between the need for training and the need to service the equipment. We develop discrete time Markov decision process formulations of the problem in order to investigate the form of the optimal action policies which maximise the expected survival time until a catastrophic event. The reason for focusing on the expected survival time rather than on cost is because we assume that the cost is immeasurably high if the system fails to respond when required.

The literature on maintenance, repair and replacement policies for deteriorating equipment started with the work of Barlow and Proschan[1965] and others, and as the surveys and bibliographies of McCall[1965], Pierskalla[1976], Sherif and Smith[1982], Monahan[1982], Thomas[1986], Valdes-Flores and Feldman[1989], Dekker[1996] and Wang[2002] indicate, it has continued apace to the present day. Almost all the literature concentrate on policies which minimise the average on discounted cost criteria. The idea of using a catastrophic event criterion to overcome the problem that failure will result in unquantifiably large cost was suggested first by Gaver, Jacobs and Thomas[1987], with other instances being considered by Kim and Thomas[2003a]. In all these cases the background environment and hence probability of initiating event is either fixed or follows a random Markovian process. Other authors such as Çinlar[1984], Çinlar and Özekici[1987], Çinlar et al[1989], Shaked and Shanthikumar[1989], Lefèvre and Milhaud[1990] and Özekici[1995][1996] have looked at maintenance in a random environment but in that case the unit is always in use so the changes in the environment age the equipment at different rates, but do not affect when it is needed. Wartman and Klutke[1994], Klutke et al[1996], , Yang et al[2000a][2000b] and Kissler et al[2002] study protective systems, such as circuit breakers, alarms, and protective relays, as well as standby systems, with non-self-announcing failures where the rate of deterioration is governed by random environment. Kim and Thomas[2003b], on the other hand, allow the deterioration of the equipment to be independent of the environment, but the environment affects the need for the equipment. Yeh[1995] studied an optimal maintenance model for standby systems focusing on availability and reliability as the criteria to optimise. None of these papers address the issue of how does the training of the operators affect the readiness of the unit which this paper considers.

We develop a Markov decision process model with random loss of learning in the training level in section 2. Numerical examples of these results are presented in section 3. In section 4 we examine a modified model where the effect of training does wear off, and look at when one should train, when one should repair as a function of the environmental situation, the training level and the state of the equipment. A numerical example of this situation and conclusions are given in section 5.

2 Training Model with Random Loss of Expertise

2.1 Introduction

In this model, there are several environmental situations which are graded from very dangerous to completely peaceful. Each environmental situation has its own probability of an initiating event occurring which increases as the situation gets more dangerous. There are three actions in this model in that the operator chooses among doing nothing, repairing, and training. At the end of this section we look at the special case in which we only consider do nothing and training. This corresponds to equipment which cannot be repaired though we do not consider the problem of when to replace such equipment.

2.2 Terminology

Possible Standby Unit Quality State, i Regular inspection of the standby unit gives information on the operation quality state of the units. This categorisation of the equipment into various states after inspection is fairly common in the military context. We assume the standby unit has N different unit quality states, i.e. $1, 2, \cdots, N$ where state 1 means that the standby unit is like new. The state $N-1$ means that it is in a poor but still operable state, while in state N, it is in a "down" condition which means that it will not work.

The Quality State Transition Probability Matrix(QSTPM), P_{ij} When the standby unit is in quality state i at the current stage, there is a probability, P_{ij} that it will be in state j at the next period where $i, j = 1, 2, \cdots, N$ and

$$\sum_{j=1}^{N} P_{ij} = 1, \text{ where } i = 1, 2, \cdots, N-1, N$$

We assume that the QSTPM satisfies a first order stochastic ordering condition so that $\sum_{j<k} P_{ij} \geq \sum_{j<k} P_{(i+1)j}$. We assume $P_{NN} = 1$ so once the standby unit reaches the "down" state N, it remains "down" until either it is repaired, or a catastrophic event occurs. There are several situations where equipment is classified as new, excellent condition, operable, failed and regular inspection of the equipment allows one to collect data to estimate the transition probabilities, P_{ij}.

Possible Environmental Situation, m We assume that there are M different environmental states, $1, 2, \cdots, M-1, M$. Environmental state 1 reflects the most peaceful environment in which there is the smallest probability, b_1 of an initiating event occurring. On the other hand, environmental state M is the most dangerous state with the highest probability, b_M of an initiating event occurring. We assume b_m is non-decreasing in the index of the environmental state m and $0 \leq b_m \leq 1$. These correspond to military states of readiness, such as the US DEFCON, or the UK, black/red/amber.

Environment Situation Transition Probability Matrix(ESTPM), $S_{mm(1)}$ The dynamics of the environmental situation is also described by a Markov chain with Environment Situation Transition Probability Matrix(ESTPM), $S_{mm(1)}$. If the environmental situation is m, $1 \leq m \leq M$ in the current stage, this changes to another environmental situation $m(1), 1 \leq m(1) \leq M$ with probability $S_{mm(1)}$ at the next stage, where

$$\sum_{m'=1}^{M} S_{mm(1)} = 1, \text{ with } m \text{ and } m(1) = 1, 2, \cdots, M-1, M$$

We assume the ESTPM also satisfies a first order stochastic ordering property so $\sum_{m(1)=1}^{k} S_{mm(1)} \geq \sum_{m(1)=1}^{k} S_{(m+1)m(1)}$ for any $m = 1, 2, \cdots, M-1$. The data for this can be obtained by historical analysis.

The Possible Actions There are three possible actions at each period, (1) do nothing, (2) repair and (3) training. The "do nothing" action means neither repair/maintenance nor training is undertaken. It is assumed the "repair" action which can be maintenance action, if the unit is still operable, but is a true repair in state N takes R unit time periods. This action is not perfect in that there is a probability R_r the unit will be in quality state r after the "repair" where $\sum_{r=1}^{N} R_r = 1$. If an initiating event occurs during repair period, the standby unit cannot respond to it, and so a catastrophic event occurs automatically.

Training Level, k The operator of the stand-by unit has L different training levels, i.e. $1, 2, \cdots, L$ where training level 1 is the best training level and training level L is the worst training level. If there is no training at the moment, the training level k goes to $k(1)$ at the next time stage with probability of $T_{kk(1)}$ which is the training level transition probability matrix(TLTPM). We assume there is no spontaneous improvement in training i.e. $T_{kk(1)} = 0$ if $k(1) < k$, and our model would allow the deterministic decrease in operator performance $T_{k,k+1} = 1$. Training may not be ideal and could be counter productive in that after a training exercise(which we assume to take one time period which defines the length of period in the model) the quality of training is p with probability w_p where $p = 1, 2, \cdots, L-1, L$. Training causes wear and tear on the equipment to a different extent than when it is not being used. Hence there is a transition probability matrix for the standby unit quality state variation caused by the training which is called wear and tear transition probability matrix(WTTPM), $\widetilde{P}_{ij}$. If $j < i$, $\widetilde{P}_{ij} = 0$. We assume that the WTTPM also satisfies a first order stochastic ordering condition so that $\sum_{j<k} \widetilde{P}_{ij} \geq \sum_{j<k} \widetilde{P}_{(i+1)j}$. We assume the wear and tear caused by training is more than the wear and tear caused by natural conditions and so require $\sum_{j\leq l} P_{ij} \geq \sum_{j\leq l} \widetilde{P}_{ij}$ where l is arbitrary quality state. If the unit is being repaired, no training is possible.

Catastrophic Event If an initiating event occurs either when the standby unit is down(in state N) or being repaired, a catastrophic event comes. To allow for the possibility that training could be aborted when an initiating event occurs, but only if the equipment is close to where it is needed, we say that if training is occurring there is a probability $(1-t)$ that the equipment can respond to the initiating event. Finally at an initiating event it is not enough for the equipment to be operating, but the training must be of a sufficient quality if there is not to be a catastrophic outcome. We assume that with training level k, one cannot successfully respond to an initiating event with probability K_k where K_k increases with k.

2.3 Model

The state space of this model S has three factors which are the unit quality state, training level, and environmental state, so

$$S = \{(i,k,m) \in S,\ i = 1,2,\cdots,\ N, k = 1,2,\cdots, L \text{ and } m = 1,2,\cdots,, M\}$$

where i, k and m mean the unit quality state, training level and the environmental situation respectively. When the unit is in quality state i, training level k and the environmental situation is in state m, $V(i,k,m)$ is the maximum expected number of periods until a catastrophic event occurs. Because we are looking for the best action policy, the optimality equation selects the best of the three actions.

$$V(i,k,m) \quad = \quad \max\{W_1(i,k,m), W_2(k,m), \delta_{iN} W_3(i,k,m)\} \tag{1}$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise}$$

where $W_1(i,k,m)$ is the expected period until a catastrophic event if nothing is done now, $W_2(k,m)$ is the expected period until a catastrophic event if a repair is performed now and $W_3(i,k,m)$ is the expected period until a catastrophic event if training is selected now. Hence $W_1(i,k,m), W_2(k,m), W_3(i,k,m)$ satisfy

$$\begin{aligned} & W_1(i,k,m) \\ = \; & (1 - b_m\phi_{iN})(1 + \sum_{j=1}^{N} P_{ij} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V(j,k(1),m(1))) \qquad (2) \\ & \text{where } \phi_{iN} = K_k \text{ if } i \neq N, \phi_{iN} = 1 \text{ if } i = N \end{aligned}$$

$$\begin{aligned} & W_2(k,m) \\ = \; & \sum_{j=1}^{R} \Pi_{t=0}^{j-1}(1-s^t) + \Pi_{t=0}^{R-1}(1-s^t) \sum_{k(1),\cdots,k(R)=1}^{M} \Pi_{t=1}^{R} T_{k(t-1)k(t)} \qquad (3) \\ & \cdot \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r V[r,k(R),m(R)] \\ & \text{where } s^0 = b_m, \;\; s^t = \sum_{m(t)=1}^{M} S_{m(t-1)m(t)} b_{m(t)}, \;\; m(0) = m \end{aligned}$$

$$\begin{aligned} & W_3(i,k,m) \\ = \; & \{1 - b_m \times [t + (1-t)K_k]\}(1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1))) \qquad (4) \end{aligned}$$

$(2), (3), (4)$ can be solved by value iteration where the nth iterate satisfies

$$V_n(i,k,m) = \max\{W_n^1(i,k,m), W_n^2(k,m), \delta_{iN} W_n^3(i,k,m)\} \tag{5}$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise, and}$$

$$\begin{aligned} & W_n^1(i,k,m) \\ = \ & (1 - b_m \phi_{iN})(1 + \sum_{j=1}^{N} P_{ij} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j, k(1), m(1))) \end{aligned} \tag{6}$$

where $\phi_{iN} = K_k$ if $i \neq N, \phi_{iN} = 1$ if $i = N$

$$\begin{aligned} & W_n^2(k,m) \\ = \ & \sum_{j=1}^{R} \Pi_{t=0}^{j-1}(1 - s^t) + \Pi_{t=0}^{R-1}(1 - s^t) \sum_{k(1),\cdots,k(R)=1} \Pi_{t=1}^{R} T_{k(R-1)k(R)} \\ & \cdot \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r V^{n-R}[r, k(R), m(R)] \end{aligned} \tag{7}$$

where $s^0 = b_m, \ s^t = \sum_{m(t)=1}^{M} S_{m(t-1)m(t)} b_{m(t)}, \ m(0) = m$

$$\begin{aligned} & W_n^3(i,k,m) \\ = \ & \{1 - b_m \times [t + (1-t)K_k]\}(1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j, p, m(1))) \end{aligned} \tag{8}$$

If we define the terminal value, $V_0(i,k,m) = 0$, $V_n(i,k,m)$ is a bounded increasing sequence of function and so converges to the limit $V(i,k,m)$. Standard results from Markov decision processes[Putterman,1994] show that the limit function satisfies the optimality equation $(1) \cdots (4)$.

Lemma 2.1 $V(i,k,m)$ is a

a) non-increasing function of i

b) non-increasing function of k

c) non-increasing function of m

where i is the quality state, and k and m are the arbitrary training level and environment situation state.

Proof The proofs use induction hypothesis on n in $V_n(i,k,m)$ and then the result[Putterman,1994] that $V(i,k,m)$ is the limit of the value iteration functions $V_n(i,k,m)$. Consider a) and defining $V_0(i,k,m) = 0$, then the property holds trivially for $n = 0$. So assume $V_{n-1}(i,k,m)$ is non-increasing in i. This together with the stochastic ordering condition of QSTPM and WTTPM implies

$$\begin{aligned}
& \sum_{j=1}^{N} P_{ij} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j,k(1),m(1)) \\
> & \sum_{j=1}^{N} P_{(i+1)j} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j,k(1),m(1))
\end{aligned}$$

$$\begin{aligned}
& \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j,p,m(1)) \\
> & \sum_{j=1}^{N} \widetilde{P}_{(i+1)j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V^{n-1}(j,p,m(1))
\end{aligned}$$

Hence, we can conclude that $W_n^1(i,k,m) \geq W_n^1(i+1,k,m)$ and $W_n^3(i,k,m) \geq W_n^3(i+1,k,m)$. Since $W_n^2(i,k,m) = W_n^2(i+1,k,m)$ from (7), it follows $V_n(i,k,m) \geq V_n(i+1,k,m)$. Hence the result holds for $V_n(i,k,m)$ and by convergence the results hold in the limit for $V(i,k,m)$. The proofs of b) and c) follow in a similar way.

Theorem 2.1 a) In state (i, k, m), the standby unit is repaired provided $i \geq i^*(k, m)$.

b) If the stand-by unit is down(in state N), it must be repaired immediately, so $N \geq i^*(k, m)$.

Proof The proof of a) follows because $V(i, k, m)$ is non-increasing function in i and using stochastic dominance of $P_{ij}, \widetilde{P}_{ij}$ ensures non-increasing property carries through to $\sum_j P_{ij} V(j, k, m)$. Hence $W_1(i, k, m)$ and $W_3(i, k, m)$ are non-increasing in i. Since $W_2(i, k, m)$ is independent of i, once $W_2(i, k, m) \geq W_3(i, k, m)$ and $W_2(i, k, m) \geq W_1(i, k, m)$, then the same inequalities must hold for larger i. So one repairs if $i \geq i^*(k, m)$.

In state $i = N$, the only allowed options are doing nothing or repair since training is not allowed in state $i = N$. $W_2(N, k, m) = W_2(k, m)$ since it does not depend on state N. For the proof of b), since the lengths of doing nothing(1 time period) and repair(R time periods) are different, we compare the expected survival period by doing nothing and repair at every single period during R periods. This is what Putterman[1994] calls *Uniformization*. We can let $W_{1(t)}(N, k(t), m(t))$,$W_{2(t)}(k(t), m(t))$ be the expected survival time if we do nothing or repair for the next period when we are t periods into the R-period repair. We also let $V_{(t)}(i, k(t), m(t)) = \max\{W_{1(t)}(i, k(t), m(t)), W_{2(t)}(i, (t), m(t)), \phi W_{3(t)}(i, (t), m(t))\}$ where $\phi = 0$ if $i = N$, $\phi = 1$ otherwise. Thus for $t = 0, \cdots, R-2$,

$$
\begin{aligned}
& W_{2(t)}(k(t), m(t)) \\
= \; & (1 - s^t) \sum_{k(t+1)=1}^{L} T_{k(t)k(t+1)} \sum_{m(t+1)=1}^{M} S_{m(t)m(t+1)} [1 + W_{2(t+1)}(k(t+1), m(t+1)]
\end{aligned}
$$

and for $t = R - 1$,

$$
\begin{aligned}
& W_{2(R-1)}(k(R-1), m(R-1)) \\
= \ & (1-s^{R-1}) \sum_{k(R)=1}^{L} T_{k(R-1)k(R)} \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r[1+V(r,k(R),m(R)]
\end{aligned}
$$

Firstly we compare $W_{1(R-1)}(N, k(R-1), m(R-1))$ and $W_{2(R-1)}(k(R-1), m(R-1))$ at $t = R-1$.

$$
\begin{aligned}
& W_{2(R-1)}(k(R-1), m(R-1)) - V_{(R-1)}(N, k(R-1), m(R-1) \\
\geq \ & W_{2(R-1)}(k(R-1), m(R-1)) - W_{1(R-1)}(N, k(R-1), m(R-1)) \\
= \ & (1-s^{R-1}) \sum_{k(R)=1}^{L} T_{k(R-1)k(R)} \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r[1+V(r,k(R),m(R)] \\
& -(1-s^{R-1}) \sum_{k(R)=1}^{L} T_{k(R-1)k(R)} \sum_{m(R)=1}^{M} S_{m(R-1)m(R)}[1+V(N,k(R),m(R)] \\
= \ & (1-s^{R-1}) \sum_{k(R)=1}^{L} T_{k(R-1)k(R)} \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r \\
& \cdot[V(r,k(R),m(R) - V(N,k(R),m(R)] \\
\geq \ & 0
\end{aligned}
$$

Hence, repair is better than doing nothing at $t = R-1$ since $V(i,k,m)$ is a non-increasing function in i. Using this result, we also find repair is the optimal action at $t = R-2$:

$$
\begin{aligned}
& W_{2(R-2)}(k(R-2), m(R-2)) - V_{(R-2)}(N, k(R-2), m(R-2)) \\
\geq \ & W_{2(R-2)}(k(R-2), m(R-2)) - W_{1(R-2)}(N, k(R-2), m(R-2)) \\
= \ & (1-s^{R-2}) \sum_{k(R-1)=1}^{L} T_{k(R-2)k(R-1)} \sum_{m(R-1)=1}^{M} S_{m(R-2)m(R-1)}
\end{aligned}
$$

$$\cdot [W_{2(R-1)}(k(R-1), m(R-1) - V_{(R-1)}(N, k(R-1), m(R-1)]$$
$$\geq \quad 0$$

By induction from $t = R-3$ to $t = 0$, we have same results. This means that , at quality state N, we should always repair.

Whereas when the equipment is in its worst state, Theorem 2.1 says one needs to repair it immediately, if the training levels are at their worst, it is not always the case that one should train. One needs to add same extra condition as Theorem 2.2 implies.

Theorem 2.2 If $K_L = 1$, $P_{ij} = \widetilde{P}_{ij}$, all i, j and $P_{1\bullet}$ stochastically dominates $R_\bullet$, this means that in some sense that repair is not as good as new, then one always trains in state $i = 1, k = L$.

Proof If training level is L with $K_L = 1$ and $i = 1$, (2), (3), (4) can be rewritten as

$$W_1(1, L, m) = (1 - b_m)(1 + \sum_{j=1}^{N} P_{1j} \sum_{m(1)=1}^{M} S_{mm(1)} V(j, L, m(1)))$$
$$W_2(1, L, m) = \sum_{j=1}^{R} \Pi_{t=0}^{j-1}(1 - s^t) + \Pi_{t=0}^{R-1}(1 - s^t) \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \sum_{r=1}^{N} R_r V[r, L, m(R)]$$
$$W_3(1, L, m) = (1 - b_m)(1 + \sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j, p, m(1)))$$

Comparing do nothing, W_1, and training, W_3, gives

$$W_3(1, L, m) - W_1(1, L, m)$$
$$= (1 - b_m)(1 + \sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j, p, m(1)))$$

$$
\begin{aligned}
& -(1-b_m)(1+\sum_{j=1}^{N} P_{1j} \sum_{m(1)=1}^{M} S_{mm(1)} V(j,L,m(1))) \\
= \;& (1-b_m)[\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1)) \\
& -\sum_{j=1}^{N} P_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,L,m(1))] \\
= \;& (1-b_m) \sum_{m(1)=1}^{M} S_{mm(1)} \sum_{p=1}^{L} w_p [\sum_{j=1}^{N} \widetilde{P}_{1j} V(j,p,m(1)) - \sum_{j=1}^{N} P_{1j} V^{n-1}(j,L,m(1))] \\
= \;& (1-b_m) \sum_{m(1)=1}^{M} S_{mm(1)} \sum_{p=1}^{L} w_p \sum_{j=1}^{N} P_{1j} [V(j,p,m(1)) - V(j,L,m(1))]
\end{aligned}
$$

Since $(1-b_m) \geq 0$ and $V(i,k,m)$ is a non-increasing function in k, $W_3(1,L,m) \geq W_1(1,L,m)$.

For the proof for $W_3(1,L,m) \geq W_2(L,m)$, suppose we can consider the case in which repair only takes 1 time period. It is obvious that the expected survival time if the repair takes 1 time period, ${}^1W_2(k,m)$, is longer than or equal to the expected survival time if the repair takes $R(R \geq 1)$ time period, $W_2(k,m)$. Hence, ${}^1W_2(k,m) \geq W_2(k,m)$. Now compare training with repair, if the latter only takes one time period, then

$$
\begin{aligned}
& W_3(1,L,m) - {}^1W_2(L,m) \\
= \;& (1-b_m)(1+\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1))) \\
& -(1-b_m)(1+\sum_{r=1}^{N} R_r \sum_{m(1)=1}^{M} S_{mm(1)} V(r,L,m(1))) \\
= \;& (1-b_m)[\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1))
\end{aligned}
$$

$$
\begin{aligned}
& - \sum_{r=1}^{N} R_r \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(r, L, m(1))] \\
= \;& (1 - b_m) \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} [\sum_{j=1}^{N} \widetilde{P}_{1j} V(j, p, m(1)) - \sum_{r=1}^{N} R_r V(r, L, m(1))]
\end{aligned}
$$

Since $(1 - b_m) \geq 0$, $V(i, k, m)$ is a non-increasing function in k and $\sum_{j=1}^{l} P_{1j} \geq \sum_{r=1}^{l} R_r$(stochastic ordering) where l is an arbitrary quality state, $W_3(1, L, m) \geq {}^1W_2(L, m)$. Hence,

$$
W_3(1, L, m) \geq \; {}^1W_2(L, m) \geq W_2(k, m)
$$

Therefore if the training level k is L with $K_L = 1$ and the quality state is 1 where $P_{ij} = \widetilde{P}_{ij}$ for all $i, j, \sum_{j=1}^{l} P_{1j} \geq \sum_{r=1}^{l} R_r$ for all l, then training is always optimal.

With slightly stronger conditions we can show we also repair or train in the worst training state.

Theorem 2.3 If $K_L = 1$ and $P_{ij} = \widetilde{P}_{ij}$ $\forall_{ij}$, then in state (i, L, m) one trains if $i < i^*(m)$ and repairs if $i \geq i^*(m)$.

Proof It is enough to show $W_1(i, L, m) \leq \max\{W_2(i, L, m), W_3(i, L, m)\}$ since then the non-increasingness in i of $W_3(i, L, m)$ and the fact $W_2(i, L, m)$ is independent of i gives the rest of the result. Since $P_{ij} = \widetilde{P}_{ij}$ and $\sum w_p V(j, p, m(1)) \geq V(j, L, m(1))$ by the non-increasing property of $V(j, k, m(1))$ in k, $W_3(i, L, m) \geq W_1(i, L, m)$ for all i and m.

If one considers equipment which is not repairable and so maintenance has no effect, then the only actions possible are training or doing nothing.

The optimality equation for $V(i, k, m)$ in this case satisfies

$$
V(i, k, m) \;=\; \max\{W_1(i, k, m), \delta_{iN} W_3(i, k, m)\} \tag{9}
$$

where

$$\begin{aligned} \delta_{iN} &= 0 \text{ if } i = N \\ \delta_{iN} &= 1 \text{ otherwise} \end{aligned}$$

and $W_1(i,k,m)$ and $W_3(i,k,m)$ are still defined by (2) and (4). In this case one can show that if one decides to train in state $(i,k^{'},m)$ one should train in all states $(i,k^{'},m), k^{'} \geq k$.

Theorem 2.4 In the non-repairable equipment special case, if one trains in state (i,k,m), one should train in all states $(i,k^{'},m), k^{'} \geq k$.

Proof

$$\begin{aligned} & V(i,k,m) \\ = \;& \max\{W_1(i,k,m), W_3(i,k,m)\} \\ = \;& \max\{(1-b_m K_k)(1+\sum_{j=1}^{N} P_{ij} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V(j,k(1),m(1))), \\ & [1-b_m \times (t+(1-t)K_k)](1+\sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1)))\} \end{aligned}$$

If training is optimal at training level k,

$$\begin{aligned} & W_3(i,k,m) \\ = \;& [1-b_m \times (t+(1-t)K_k)](1+\sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)} V(j,p,m(1))) \\ \geq \;& (1-b_m K_k)(1+\sum_{j=1}^{N} P_{ij} \sum_{k(1)=1}^{L} T_{kk(1)} \sum_{m(1)=1}^{M} S_{mm(1)} V(j,k(1),m(1))) \\ = \;& W_1(i,k,m) \end{aligned}$$

If we let $1-b_m\times(t+(1-t)K_k) = A(k), 1+\sum_{j=1}^{N}\widetilde{P}_{ij}\sum_{p=1}^{L} w_p \sum_{m(1)=1}^{M} S_{mm(1)}V(j,p,m(1))$ $= B$, $1-b_mK_k = C(k), 1+\sum_{j=1}^{N} P_{ij}\sum_{k(1)=1}^{L} T_{kk(1)}\sum_{m(1)=1}^{M} S_{mm(1)}V(j,k(1),m(1))$ $= D(k)$, the above relation can be rewritten by

$$\begin{aligned} A(k)B &\geq C(k)D(k) \\ B &\geq \frac{C(k)}{A(k)}D(k) \end{aligned}$$

If we think of $\frac{C(k)}{A(k)}$,

$$\frac{C(k)}{A(k)} = \frac{1-b_mK_k}{1-b_m\times(t+(1-t)K_k)}$$

If $b_m, t > 0$, $\frac{C(k)}{A(k)}$ is a decreasing function of k from $\frac{1}{1-b_mt}$ at $k = 1$(perfect training level : $K_k = 0$) to 1 at $k = L$(worst training level $K_k = 1$).

Because of this and $D(k)$ being a decreasing function in k, hence

$$B \geq \frac{C(k)}{A(k)}D(k) \geq \frac{C(k+1)}{A(k+1)}D(k) \geq \frac{C(k+1)}{A(k+1)}D(k+1)$$

Hence,

$$A(k+1)B \geq C(k+1)D(k+1)$$

This means that $W_3(i,k+1,m) \geq W_1(i,k,m)$ and so if training is optimal at (i,k,m), it is also optimal for (i,k',m), $k' > k$.

3 Examples

Consider a problem with five environment situations, 1(most peaceful environment), 2, 3, 4, 5 (most dangerous environment), two training levels $1, 2$, 10 unit quality states, 1(new), $2, \cdots, 9$, 10 (down) and with the Quality State Transition Probability Matrix(QSTPM) and Wear and Tear Transition Probability Matrix(WTTPM) given by Table 1 and 2. We also assume that repair is not perfect but given by R_r in Table 4 and takes 1 time period($R = 1$). The probability of an initiating event, b_m is $\{0.1, 0.2, 0.4, 0.6, 0.7\}$ from environmental situation 1 to 5. The value of K_k, w_p is $K_k = (0, 1), w_p = (1, 0)$ for training level 1,2. The probability that training can not respond to an initiating event, t is 0.7. We use different versions of this problem with different training transitions. For example 1 $T_{kk'} = \begin{bmatrix} 0.6 & 0.4 \\ 0 & 1 \end{bmatrix}$, and for example 2, $T_{kk'} = \begin{bmatrix} 0.01 & 0.99 \\ 0 & 1 \end{bmatrix}$. So training has a positive effect for $\frac{1}{0.4} = 2.5$ periods in example 1 and $\frac{1}{0.99} = 1$ period in example 2.

The results for example 1 and 2 are shown in Figure 1 to Figure 4. We know that the expected survival period is a non-increasing function in quality state i, training level k and environment situation m. Because the transition probability from training level 1 to training level 2 in example 2 is bigger than in example 1, the expected survival period for example 1 is longer than that for example 2. Since the effect of training lasts for a shorter time in example 2 than in example 1, training occurs even when the training level is high in example 2 as we can see in Figure 3, but does not occur in that state in example 1. However one trains more in the poor training level state in example 1 than example 2. In Figure 2, training is optimal when $k = L$(i.e. level 2 in example 1) and $i \neq N$. However, this is not always true as example 2 shows. From Figure 4, repair is optimal even though the training level is $k = 2$(worst training level) and the quality state is in working condition($i = 9$).

Example 3 is the case of $P_{ij} = \widetilde{P}_{ij}$. In reality this means that training does not cause any more wear and tear than doing nothing. In this example, $P_{i,i+1} = \widetilde{P}_{i,i+1} = 1$ where $i = 1, \cdots, 5$.Otherwise, $P_{iN} = \widetilde{P}_{iN} = 1$ where $i = 6, - - -, N$. We use the $T_{kk'}$ used in example 2. The other conditions are the same as in previous examples. The results for example 3 are shown in Figures 5 and 6. Training is always optimal when

$k = 2$(worst training level) and $i = 1$. For the worst training level($k = 2$), the quality state increases, the optimal action will change from training to repair provided $K_2 = 1$ and $P_{ij} = \widetilde{P}_{ij}$.

Example 4 looks only at the do nothing or training problem where repair is not possible. We assume 5 training levels for this example and Table 5 gives the training level transition probability whereas the environmental state and quality state transitions stay as in example 2. The values of K_k and w_p are $K_k = (0.0, 0.2, 0.5, 0.7, 1.0)$, $w_p = (0.6, 0.2, 0.1, 0.05, 0.05)$ for training level $k = 1, 2, 3, 4, 5$ respectively. Figures 7 and 8 show the results of example 4. In this case once training is optimal at training level k, training is optimal for all $k^{'} \geq k$.

Table 1. Quality State TPM, P_{ij}

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	0.2	0.2	0.2	0.1	0.08	0.05	0.05	0.05	0.05	0.02
2	0	0.2	0.2	0.2	0.1	0.1	0.08	0.05	0.04	0.03
3	0	0	0.2	0.2	0.2	0.1	0.1	0.1	0.05	0.05
4	0	0	0	0.2	0.2	0.2	0.15	0.1	0.1	0.05
5	0	0	0	0	0.2	0.3	0.2	0.1	0.1	0.1
6	0	0	0	0	0	0.2	0.3	0.2	0.2	0.1
7	0	0	0	0	0	0	0.2	0.3	0.3	0.2
8	0	0	0	0	0	0	0	0.3	0.4	0.3
9	0	0	0	0	0	0	0	0	0.4	0.6
10	0	0	0	0	0	0	0	0	0	1

Table 2. Wear and Tear TPM, $\widetilde{P}_{ij}$

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.05	0.05
2	0	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1
3	0	0	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1
4	0	0	0	0.1	0.1	0.2	0.2	0.2	0.1	0.1
5	0	0	0	0	0.1	0.1	0.2	0.2	0.2	0.2
6	0	0	0	0	0	0.1	0.2	0.3	0.2	0.2
7	0	0	0	0	0	0	0.1	0.3	0.3	0.3
8	0	0	0	0	0	0	0	0.2	0.4	0.4
9	0	0	0	0	0	0	0	0	0.3	0.7
10	0	0	0	0	0	0	0	0	0	1

Table 3. Environmental Situation TPM, $S_{mm'}$

$m \setminus m'$	1	2	3	4	5
1	0.4	0.3	0.2	0.05	0.05
2	0.2	0.4	0.23	0.1	0.07
3	0.1	0.2	0.4	0.2	0.1
4	0.05	0.15	0.2	0.3	0.3
5	0.05	0.1	0.15	0.2	0.5

Table 4. Repair TPM, R_r

r	1	2	3	4	5	6	7	8	9	10
R_r	0.2	0.2	0.1	0.1	0.1	0.1	0.08	0.05	0.05	0.02

Table 5. Training Level TPM for example 4, $T_{kk'}$

$k\backslash k'$	1	2	3	4	5
1	0.3	0.3	0.2	0.1	0.1
2	0	0.3	0.3	0.2	0.2
3	0	0	0.4	0.3	0.3
4	0	0	0	0.4	0.6
5	0	0	0	0	1

Figure 1. Result of example 1, training level 1

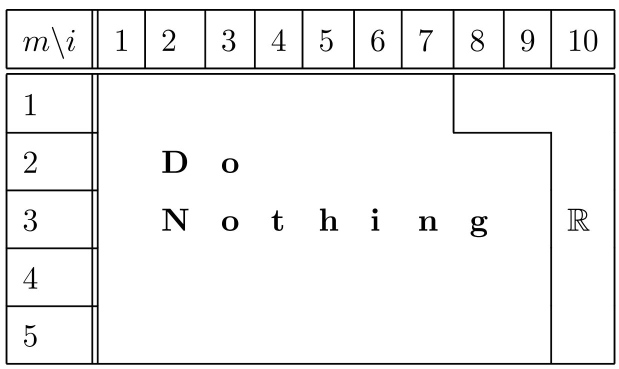


Figure 2. Result of example 1, training level 2

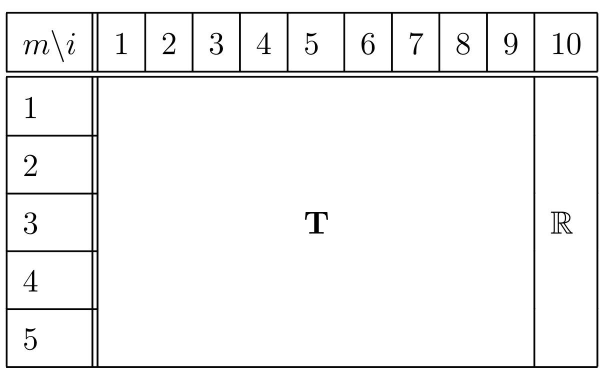


Figure 3. Result of example 2, training level 1

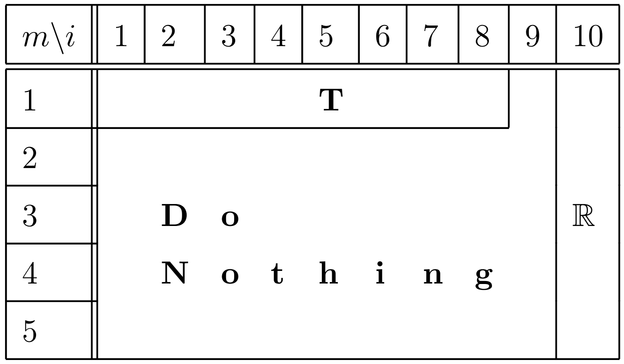


Figure 4. Result of example 2, training level 2

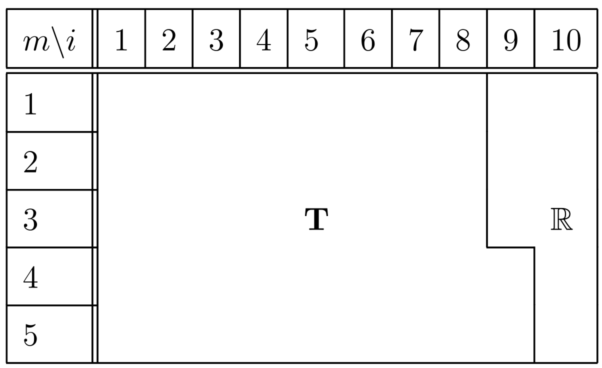


Figure 5. Result of example 3, training level 1

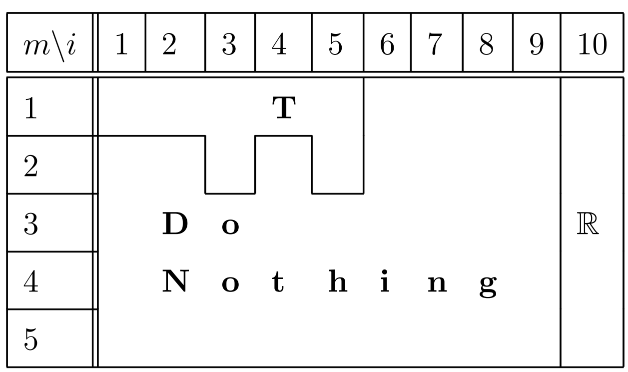


Figure 6. Result of example 3, training level 2

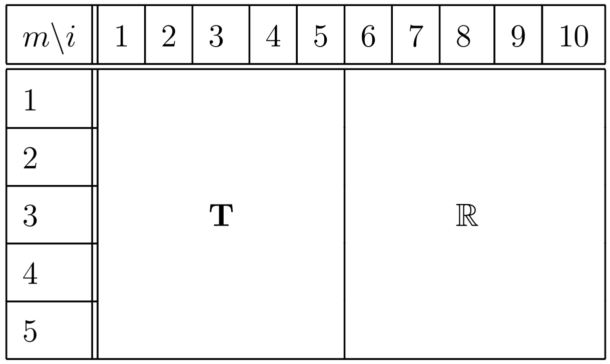


Figure 7. Result of example 4, environment 1

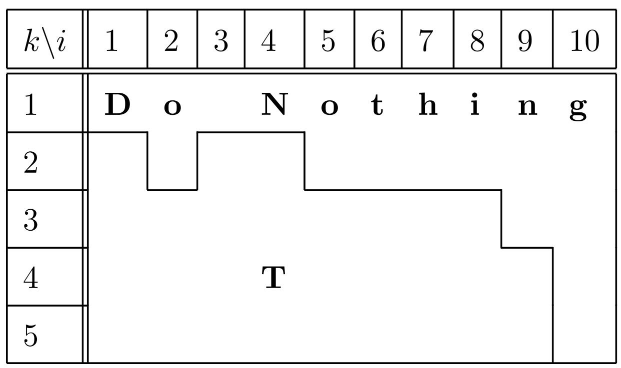
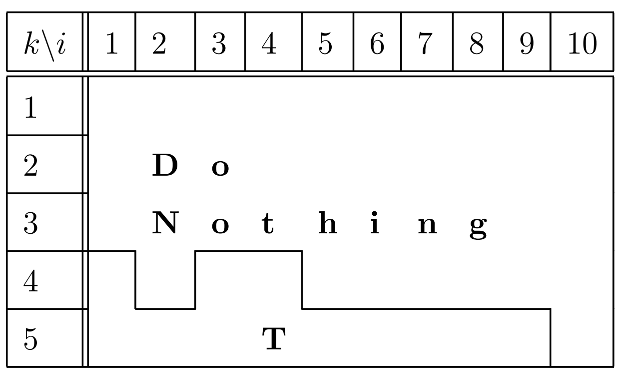


Figure 8. Result of example 4, environment 5



4 Training Model with Continuous Loss of Expertise

In this section we consider the problem where the expertise obtained by training on the equipment is gradually lost over time rather than subject to random changes as in section 2. To do this we define an expertise index which shows how well trained the operator of the stand-by unit has been. An operator with higher values in this index is likely to perform better. Apart from the modification of an expertise index instead of a training level, the other conditions are the same as in section 2.

To define the expertise index, we assume that if it is at level $\mathbb{T}, \mathbb{T} \geq 0$, then

a) when no training occurs at the next period, it moves to $\alpha\mathbb{T}$

b) when training occurs at the next period, it moves to $\alpha\mathbb{T} + 1$

So in a sense, expertise obtained through training dissipates geometrically(the equivalent of exponentially in discrete time) and each period of training of training adds 1 unit to the expertise level whatever it is. Thus training all the time gives us a level of $1 + \alpha + \alpha^2 + \alpha^3 + \cdots = (1 - \alpha)^{-1}$, while no training gives an expertise of 0. In order to make the index easy to understand, we multiply the above index by $(1 - \alpha)$ to arrive at one where all values are between 0 and 1, and if we let $T = (1 - \alpha)\mathbb{T}$, training changes T into $\alpha T + (1 - \alpha)$, while no training changes T into αT. If the expertise index is T, the probability the operator can not respond to satisfactorily to an initiating event is f_T where $0 \leq f_T \leq 1$, $f_{T'} < f_T$ if $T < T'$.

The state space of this model S has three factors which are the unit quality state, training level, and environmental state, so

$$S = \{(i, T, m) \in S,\ i = 1, 2, ---,\ 0 \leq T \leq 1 \text{ and } m = 1, 2, ---, , M\}$$

where i, T and m mean the unit quality state, expertise index and the environmental situation respectively. When the unit is in quality state i, expertise index T and the environmental situation is in state m, $V(i, T, m)$ is the maximum expected number of periods until a catastrophic event occurs.

$$V(i, T, m) \quad = \quad \max\{W_1(i, T, m), W_2(T, m), \delta_{iN} W_3(i, T, m)\} \tag{10}$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise}$$

where

$$W_1(i,T,m) = (1 - b_m\delta_{iN})(1 + \sum_{j=1}^{N} P_{ij} \sum_{m(1)=1}^{M} S_{mm(1)} V(j, \alpha T, m(1))) \quad (11)$$

where $\phi_{iN} = f_T$ if $i \neq N, \phi_{iN} = 1$ if $i = N$

$$\begin{aligned} & W_2(T,m) \\ = & \sum_{j=1}^{R} \Pi_{t=0}^{j-1}(1 - s^t) + \Pi_{t=0}^{R-1}(1 - s^t) \sum_{m(R)=1}^{M} S_{m(R-1)m(R)} \\ & \cdot \sum_{r=1}^{N} R_r V[r, \alpha^R T, m(R)] \end{aligned} \quad (12)$$

where $s^0 = b_m,\ s^t = \sum_{m(t)=1}^{M} S_{m(t-1)m(t)} b_{m(t)},\ m(0) = m$

$$\begin{aligned} & W_3(i,T,m) \\ = & \{1 - b_m \times [t + (1-t) f_T]\}(1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{m(1)=1}^{M} S_{mm(1)} V(j, \alpha T + (1-\alpha), m(1))) \end{aligned} \quad (13)$$

$(9), (10), (11), (12)$ can be solved using value iteration. The results of the previous section extend to this model and the proofs follow by induction on value iteration. The value iteration scheme satisfies equation $(9), (10), (11), (12)$ with V_n, W_n on the L.H.S. and V_{n-1} on the R.H.S.

Lemma 4.1 $V(i, T, m)$ is a

a) non-increasing function of i

b) non-increasing function of T

c) non-increasing function of m

where i is the quality state, T and m are the arbitrary expertise index and environment state.

Proof As in lemma 2.1, the proofs use induction hypothesis on n in $V_n(i, T, m)$ and then the result that $V(i, T, m)$ is limit of the value iteration functions $V_n(i, T, m)$. If we consider a) and define $V_0(i, T, m) = 0$, then the property holds trivially for $n = 0$. So assume $V_{n-1}(i, T, m)$ is non-increasing in i. This together with the stochastic ordering condition of P_{ij} and $\widetilde{P}_{ij}$ implies

$$W_n^1(i, T, m) \geq W_n^1(i + 1, T, m)$$

$$W_n^3(i, T, m) \geq W_n^3(i + 1, T, m)$$

Since $W_n^2(i, T, m) = W_n^2(i + 1, T, m)$ from (11), it follows $V_n(i, T, m) \geq V_n(i + 1, T, m)$. By convergence the results hold in the limit for $V(i, T, m)$. The proofs of b) and c) follow in a similar way.

Theorem 4.1 In state N(the down), one should always repair.

Proof The proof follows as b) in Theorem 2.1.

5 Example

In the example of this model, there are also 5 different environmental situation states and 10 different unit quality states. The probability of an initiating event, b_m, where

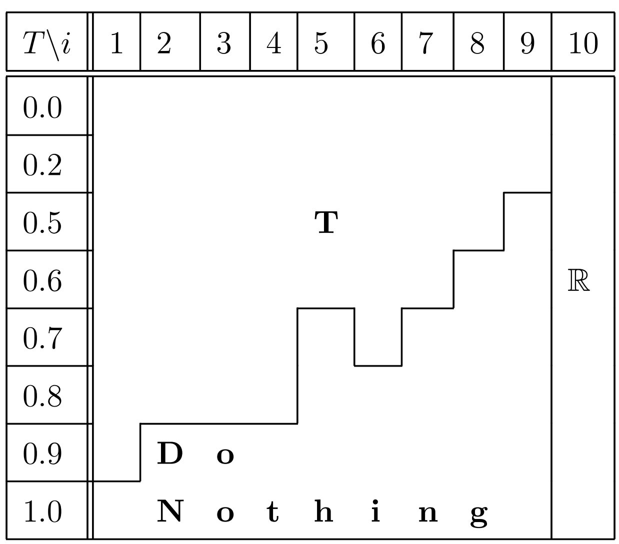
$1 \leq m \leq 5$ is $(0.1, 0.2, 0.4, 0.6, 0.7)$. Repair takes 1 time period($R = 1$) in this example. The transition matrices for quality state, environment situation are the same as in previous examples. The probability which training can not respond an initiating event, t is 0.7. The discount factor for the expertness index, α is 0.6. The effect of the repairs has the following distributions in Table 6.

Table 6. Repair TPM, R_r

r	1	2	3	4	5	6	7	8	9	10
R_r	0.1	0.1	0.2	0.2	0.1	0.1	0.08	0.05	0.05	0.02

The results in Figure 9 shows that the expected survival period is also non-increasing function in quality state and environment situation and non-decreasing in expertness index. The results in general show the pattern in that as the quality state increases(gets worse) one initially trains, then does nothing, and then repairs. However this pattern is violated in the case of $T = 0.7$.

Figure 9. Result of example, environment 3



6 Conclusions

The models presented in this paper show that there is a strong interaction between the quality state of the stand-by unit, the general environment state, the training level of the operator and the decision on whether to repair or train. In the training model with random loss of expertise, the expected survival time until a catastrophic situation decreases as the unit quality states, the training levels and environment situations worsen. One always repairs when the unit is down. Also, once repair is optimal for a given quality state, repair is optimal for worse quality states. When there is no difference between QSTPM and WTTPM, if the training level is at its lowest which means that it can not respond to an initiating event, one always trains or repairs. One trains if the quality state is good and there is same quality level below which one repairs. If one adds the extra condition $\sum_{j<l} P_{1j} > \sum_{r<l} R_r$ to the above condition, one can show one always trains when the quality state is new. If the repair action is not available and the quality state is in working condition, once training is optimal at a certain training level, one always trains at worse training levels.

In the training model with continuous loss of expertise, we find that one always repairs when the unit is down. If the unit is operating then as the expertise index increases, one moves from training to doing nothing, but this need not always be the case.

References.

Barlow, R. E., Proschan, F., 1965. Mathematical theory of reliability, John Wiley, New York

Çinlar, E., 1984. Markov and semimarkov models of deterioration. In Reliability Theory and Models(M. Abdel-Hamid, E. Çinlar, J. Quinn, eds), Academic Press, New York, 3-41.

Çinlar, E. and Özekici, S., 1987. Reliability of complex devices in random environments. Prob. Engrg. Inform. Sci. 3, 97-115.

Çinlar, E., Shaked, M. and Shanthikumar, J.G., 1989. On lifetimes influenced by a common environment. Stoch. Proc. Appl. 33, 347-359.

Dekker, R., 1996. Applications of maintenance optimization models: A review and analysis. Reliability Engineering and System Safety 51(3), 229-240.

Kiessler, P.C., Klutke, G.-A. and Yang, Y., 2002. Availability of periodically inspected systems subject to Markovian degradation. J. Appl. Prob. 39, 700-711.

Kim, Y.H., Thomas L.C., 2003a. Repair strategies in an uncertain environment: Markov Decision Processes Approach. Working Paper M03-11, University of Southampton, UK

Kim, Y.H., Thomas L.C., 2003b. Repair strategies in an uncertain environment: Stochastic Game Approach. Working Paper M03-14, University of Southampton, UK

Klutke, G.-A., Wortman, M. and Ayhan, H., 1996. The availability of inspected systems subject to random deterioration. Prob. Engrg. Inform. Sci. 10, 109-118.

Lefèvre, C. and Milhaud, X. 1990. On the association of the lifelengths of components subjected to a stochastic environment. Adv. App. Prob. 22, 961-964.

McCall, J.J., 1965. Maintenance policies for stochastically failing equipment: A survey. Management Science 11(5), 493-524.

Mertens, J. F., Neyman, A., 1981. Stochastic Game. Int. Journal of game Theory 10(2), 53-66.

Monahan, G.E., 1982. A survey of partially observable Markov decision processes: theory, models and algorithms. Management Science 28, 1-16.

Özekici, S., 1995. Optimal maintenance policies in random environments. Eur. J. Oper. res. 82, 283-294.

Özekici, S., 1996. Complex systems in random environments. In Reliability and Maintenance of Complex Systems(S. Özekici, ed), Springer Verlag, New York, 137-157.

Pierskalla, W.P., Voelker, J.A., 1976. A survey of maintenance models: the control and surveillance of deteriorating systems. Naval Research Logistics Quarterly 23, 353-388

Putterman, M. L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons

Shaked, M. and Shanthikumar, J. G., 1989. Some replacement policies in a random environment.. Prob. Engrg. Inform. Sci. 3, 117-134.

Sherif, Y.S., Smith, M.L., 1981. Optimal maintenance models for systems subject to failure-A review. Naval Research Logistics Quarterly 28(1), 47-74

Thomas, L.C., 1986. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. Reliability Engineering and System Safety 16 (4), 297-309

Thomas, L.C., Jacobs, P.A., Gaver, D.P.,1987. Optimal Inspection Policies for Standby Systems. Communication in Statistics-Stochastic Models 3(2), 259-273

Valdez-Flores, C. Feldman, R.M., 1989. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. Naval Research Logistics 36, 419-446

Wang, H., 2002. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 139, 469-489

Wartman, M. and Klutke, G.-A. 1994. On maintained systems operating in a random environment. J. Appl. Prob. 31, 589-594.

Yang, Y. and Klutke, G.-A., 2000a. Improved inspection schemes for deteriorating equipment. Prob. Engrg. Inform. Sci. 14, 445-460.

Yang, Y. and Klutke, G.-A., 2000b. Lifetime characteristics and inspection schemes for Lèvy degradation processes. IEEE Trans. on Rel. 49, 337-382.

Yeh, L. 1995. An optimal inspection-repair-replacement policy for standby systems. J. Appl. Prob. 32, 212-223.