A MARKOV DECISION PROCESS WITH UNOBSERVABLE STRENGTH UNDER MARKOVIAN ENVIRONMENT

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

PINAR ALTINKESER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

JANUARY 2020

Approval of the thesis:

A MARKOV DECISION PROCESS WITH UNOBSERVABLE STRENGTH UNDER MARKOVIAN ENVIRONMENT

submitted by **PINAR ALTINKESER** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Yasemin Serin
Head of Department, **Industrial Engineering** ________________

Prof. Dr. Yasemin Serin
Supervisor, **Industrial Engineering, METU** ________________

Examining Committee Members:

Assoc. Prof. Dr. Z. Pelin Bayındır
Industrial Eng., METU ________________

Prof. Dr. Yasemin Serin
Industrial Engineering, METU ________________

Assoc. Prof. Dr. İsmail S. Bakal
Industrial Eng., METU ________________

Assoc. Prof. Dr. Seçil Savaşaneril
Industrial Eng., METU ________________

Assist. Prof. Dr. Benhür Satır
Industrial Eng., Çankaya University ________________

Date: 28.01.2020

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Pınar Altınkeser

Signature:

ABSTRACT

A MARKOV DECISION PROCESS WITH UNOBSERVABLE STRENGTH UNDER MARKOVIAN ENVIRONMENT

Altınkeser, Pınar
Master of Science, Industrial Engineering
Supervisor: Prof. Dr. Yasemin Serin

January 2020, 57 pages

This thesis focuses on a system which survives or fails depending on the stress it is exposed to and strength occurrences. Stress and strength are random quantities generated by two factors, an unobservable Markovian environment and the actions applied. The distributions are known but only the resulting level of realized stress can be observed. The objective is to find an optimal policy that only uses available information to minimize the long run average cost of running this system. The base decision model is modified to handle lack of information, and a new model is built. Value of information is measured by comparing the optimal objective values of two models. Conflicting performance measures are also evaluated.

Keywords: Stress Strength Reliability Models, Markov Decision Process, Markov Modulated Environment

ÖZ

MARKOV ÖZELLİKLİ ORTAMDA GÖZLEMLENEMEYEN DAYANIKLILIK ALTINDA BİR MARKOV KARAR SÜRECİ

Altınkeser, Pınar
Yüksek Lisans, Endüstri Mühendisliği
Tez Danışmanı: Prof. Dr. Yasemin Serin

Ocak 2020, 57 sayfa

Bu tez, yaşam döngüsü uygulanan stres seviyesine ve sahip olduğu dayanıklılığa bağlı olan bir sisteme odaklanmaktadır. Bu sistemin ömrü stres ve dayanıklılık ilişkisine bağlı olarak devam etmekte veya bitmektedir. Stres ve dayanıklılık değerleri iki faktörden etkilenen rassal büyüklüklerdir. Bu iki faktör gözlemlenemeyen Markov özellikli bir çevresel etki ve alınan aksiyonlardır. Bu faktörlerin sonucu olan dağılımlar bilinmektedir, ancak sadece sonuçlanan stres seviyesi gözlenebilmektedir. Amaç mevcut bilgi ile uzun vadedeki ortalama maliyeti en düşük olan politikaları bulmaktır. Temel karar modeli, gözlemlenemeyen değişkenler nedeniyle güncellenmiş, yeni bir model oluşturulmuştur. Bu iki model kıyaslanarak ulaşılamayan bilginin değeri ölçülmüştür. Çatışan amaç fonksiyonları ile çözülen modellerin sonuçları karşılaştırılmıştır.

.

Anahtar Kelimeler: Stres Dayanıklılık Güvenilirlik Analizi, Makrov Karar Süreci, Markov Özellikli Ortam

To my beloved family…

ACKNOWLEDGEMENTS

I would like to express that it was a great opportunity for me to study with Prof. Dr. Yasemin Serin during my thesis study. I would like to thank for her patient guidance, enthusiastic encouragement and useful critiques in this process. She consistently allowed this thesis to be my own work but steered me in the right direction whenever she thought I needed it. It was a great privilege to work and study under her guidance.

I would like to thank my company ASELSAN INC. for letting me have time for this study. I felt continuous support of my managers/directors Bülent Bilgin, Tolga İpek and Ümit Bülbül during my study. I am very lucky to have the opportunity to know them and work with them.

I would like to express my gratitude to my dear friends, especially Asst. Prof. Dr. Ayşe Selin Kocaman who encouraged and motivated me a lot during my study.

I would like to thank my beloved family members, my mother Seher Altınkeser, my father Ahmet Altınkeser, my brother Çınar Altınkeser, my sister Semiha Altınkeser and my little princess niece Ezgi Altınkeser for their love and guidance in whatever I pursue. Last but not least, I wish to thank my loving and supportive fiance, İzzet Serbest, who provide unending inspiration and continuous encouragement throughout my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE SURVEY

1.1. Introduction

The stress–strength models have been used for reliability assessment of some systems especially during the design phase. Stress can be defined as any factor that can cause failure of a component or system reaching a certain level. The concept of stress is mostly used in electronic systems or components where stress can be medium temperature, humidity, effective voltage, current, etc. Strength of a system is the maximum level of stress it can sustain. Therefore, stress-strength analysis defines reliability as the probability that the strength being larger than the stress.

We consider a system with a random strength that is affected by an unobservable environmental condition. We have to apply a random stress to this system; we can only partially control the probability distribution of the stress because unobservable environment has an effect on that distribution also. We cannot observe the external environmental condition; however, we know that it can be modelled by a Markov chain and we know how it governs probability distributions of the stress and the strength. Applying stress may mean using the system in a certain fashion as well as adjusting its working conditions. As a result, stress controls have different costs. As the environmental condition evolves, we need to select the stress controls so that this system runs in a desirable way.

We construct a Markov decision process (MDP) model to solve this problem. Its states have observable as well as unobservable variables; but it carries sufficient information to compute the transition probabilities together with the selected actions. A policy of a Markov decision process describes an action for every possible state; hence assumes observable state information. To overcome that difficulty, we force the process to take actions without taking unobservable part of the state into consideration. Although there are other methods in literature to handle MDP’s with unobservable states their solutions may not be very practical.

In literature, there are studies that take either stress or strength as time dependent processes. Rather, we assume that they are time independent, but the strength is affected by the Markov modulated external conditions and stress is probabilistically controllable by the decision maker. In other words, actions available to the decision maker change the distribution of the random stress on the system whereas the Markov modulated environment changes the distribution of the random strength of the system. We derive optimal policies with respect to a cost minimization objective. We also evaluate other performance measures such as probability of failure, mean time to failure. We investigate how the optimal policy and performance measures change with respect to some parameters.

The organization of the next chapters is as follows. In Chapter 1, the main idea and the purpose of the study is given. A brief literature review related to the topic and the contributions of this study are also presented in Chapter 1. Chapter 2 includes stress-strength reliability definition and main assumptions of the study conducted. In Chapter 3, the decision process of the problem is discussed. This chapter also includes the models built in the scope of the study. The results of the computational experiments are demonstrated in Chapter 4, which also includes parametric analysis and variations of the models. In the last chapter of the thesis, concluding remarks and possible future research directions are discussed.

1.2. Literature Survey

In reliability-maintenance interface, there are many studies in the literature emphasizing one or the other more. The topic is great interest of electronic, mechanical and civil design fields. Studies by Johnson (1998) and Kotz et al. (2003) can be seen for extensive review of the topic. Reliability is studied related to a lifetime distribution or as stress-strength collusion. The lifetime distribution is mostly modeled by a discrete probability distribution so that it can be handled with Markov chains or in case of a control problem, with Markov decision processes. In stress-strength reliability studies, one or both can be taken as time dependent processes. Expected failure time computation together with some statistical estimation problems are often studied. Optimization comes into picture in terms of some controllable parameters such as length of the inspection, inventory size of the spare parts, etc. to minimize a cost function; these decisions mostly are time/state independent.

An earlier study by Dhillon (1980) offers models for different stress-strength inference scenarios. A time-dependent stress-strength model with random strength and constant stress which are repeatedly applied at random cycle times was researched by Siju and Kumar (2016). Xie and Shen (1991) focused on reliability growth by changing strength distribution parameters. In the present study, the effects of changing stress mean, and variance are calculated. Markov decision process is used to model the problem, and actions are based on the mean and variance of the normally distributed stress. Many studies have addressed estimation problems for stress-strength reliability. A good example is the study by Bhuyan and Dewanji (2017) which focuses on the cases where both stress and strength are time dependent. Stress depends on random damages due to the shocks arriving at random points, strength has a deterministic decreasing curve. The study evaluates two sampling plans to make the estimation stronger.

In most of the studies, systems are assumed to have two possible states, working or failed. There have been also studies which involves intermediate states to the model, such as the study by Eryilmaz (2011). Eryilmaz (2011) defined multiple states based on the difference between stress and strength levels. Qin. et. al. (2017), on the other hand, also defined more than two states, where both stress and strength are exponentially distributed random variables and used the ratio of the strength and stress values instead of the difference between them.

Godoy et al. (2013) proposed a decision technique for spare parts ordering, using stress-strength interference. Probability of stress being less than strength is defined as Condition-based service level (CBSL) in the study. The paper deals with the situations where both stress and strength are stochastic, and one of them is constant while the other is stochastic. In our model, we assume both stress and strength are normally distributed random variables, where the parameters of the distribution are determined by the actions and an unobservable environmental condition.

Markov modulated environment is one of the main assumptions of this study. Markov modulated environment indicates an unobservable environmental factor which is modelled by a Markov chain. This factor impacts the main state that we observe the behavior of. Although there are many studies about Markov modulated models on inventory problems or other fields, there has been no study about Markov modulated stress strength reliability models. The concept of Markov modulated environment is widely used for Poisson Processes in the literature. Markov Modulated Poisson Process (MMPP) is a poisson process in which arrival rate varies according to a Markov process. Ramesh (1995) proposed a MMPP model and its extensions. There are many applications of MMPP study available in the literature. A good example is the study by Scott and Smyth (2003) in which MMPP is used for web surfing behavior analysis. In this study, a special case of MMPP is used which satisfies certain rules and referred as The Markov-Poisson cascade (MPC). By help of this specific case, click rate for computer users browsing through web is modelled.

Another study by Andronov and Gertsbakh (2014), which is also related to failure time of a system, focuses on a system working in an environment in which failure rates of subsystems are jointly modulated by a continuous time Markov chain. System failure distribution function is derived using order statistics for subsystem lifetimes. This study basically shows the impact of random environment on the distribution function of the lifetime of a system.

MDPs in which the state of the system at each decision stage is not precisely known are called *Partially observable Markov decision processes* (POMDPs). Drake (1962) introduced the first POMDP model where in addition to the uncertain dynamics of an MDP, the state of the world is only partially observed through a noisy channel. In POMDP models, the relation between true state of the system and the observation is denoted by a relationship matrix. Corotis et al. (2005) proposed and advanced POMDP model, which does not have the assumption that the relationship between the observation and the true state of the process depends only on the prior control action. Policies produced by POMDP models are not very practical to implement. In the present study, unobservability is dealt with adding some constraints to the original MDP rather than POMDP methodology. Serin and Kulkarni (2005) added a special case of unobservability and proposed a model to handle it with constraints. We adopt that model in our study as in the study by Satır (2010). They also solved the problem with deterministic policy restrictions. Linear programming models for MDP's provide means of handling constraints to the problem. Such models with different properties as well as under additional constraints can be found in the book by Altman (1999). Randomized policies under constraints are tried to be avoided in some studies. For example, Ross (1989) discusses the problem of finding equivalent nonrandomized, but nonstationary, policies for single constraint case.

Maintenance decisions that are modeled as Markov decision processes mostly take a deteriorating condition as the state variable. Using the structure of the transitions, they obtain structured optimal policies. Reliability can now be defined in terms of these conditions that the process follows.

The present study proposes a dynamic decision-making process in which the states are partially governed by an unobservable process. (i) Maintenance/repair actions result in explicit distributions of the stress (may be of the strength) (ii) State of the process carries information resulting from the collusion of a random strength and a random stress. Optimal decisions with respect to a minimum cost objective are computed using Markov decision process methodology.

Unobservability brings a realistic dimension to maintenance problems. It is often the case that the repair/maintenance action changes as you learn more about the system.
(iii) We model unobservability in maintenance/repair problem as constraints to MDP. That makes the MDP problem harder to solve. The problem under consideration is new with properties (i), (ii) and (iii) to our best knowledge.

CHAPTER 2

STRESS-STRENGTH RELIABILITY PROBLEM

The concern of this study is to model stress-strength reliability of a system and obtain a decision policy that takes actions based on observations while optimizing its performance. As stated in the first section, the strength of the system is defined as the highest level of stress the system can sustain. Failure occurs when stress, X, exceeds strength, Y. Therefore, probability of survival, namely the reliability, R, is defined as P(X<Y). If the strength is fixed and known, and if the stress is a random variable with a pdf $f_x(x)$, then the reliability is

$$R = 1 - \int_{Y}^{\infty} f_X(x)dx \quad (2.1)$$

as demonstrated in Figure 2.1.

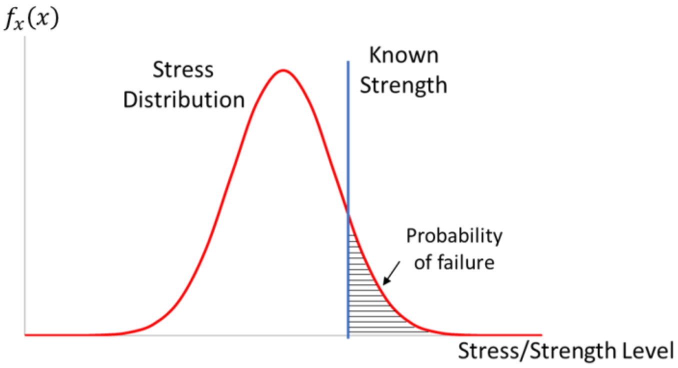

Figure 2.1: Normally distributed stress and fixed strength

If both stress and strength are normally distributed random variables, as in Figure 2.2, the interference area is an indication of the probability of failure, 1- P(X<Y).

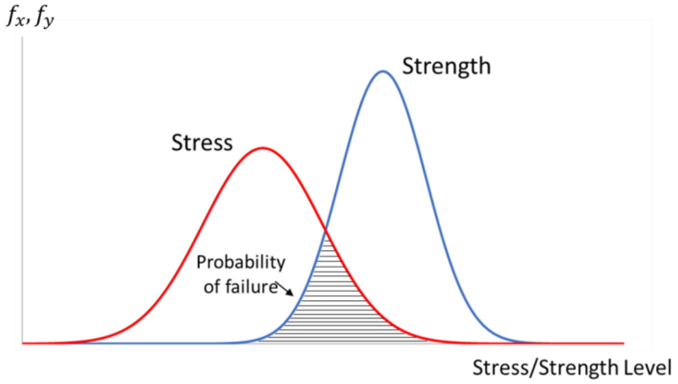

Figure 2.2: Normally distributed stress and strength

If X and Y both have independent normal distributions, then the failure probability is

$$P(X \geq Y) = 1 - R = 1 - P(X < Y) = 1 - \int_0^{\infty} f_X dx \int_x^{\infty} f_Y dy \qquad (2.2)$$

In the present study, we assume the strength of the system is normally distributed with mean μ_y and variance σ_y^2, as usually assumed in literature. Stress, X, independent of Y, is also assumed to be normally distributed and distribution parameters are assumed to depend on the actions applied. Each action leads to a different set of parameters of the stress distribution; however, it cannot change the strength of the system. In other words, it is possible to lower the mean of the stress distribution by changing action, which is desirable for low failure probability. However, each action has different costs. The actions which lead to lower distribution mean for stress have higher costs. Mean or variance of the stress that an item is exposed to can be changed in various ways. For example, the mean of a current can be decreased by adding parallel units, the variance of the temperature can be decreased by using an air conditioner to control extreme temperature values, etc.

Besides, we have the following assumptions: We cannot measure or observe strength of the system; this is a realistic assumption as the strength of the system is defined as the highest level of stress the system can sustain. However, we assume that stress level can be measured that is a realization of the stress random variable. Stress measurement at a time point will be referred as *the observation* at that time point. We aim to be able to select actions at a time based on the observation at that time. The model constructed here aims to find the best policy that minimizes the long-term average cost.

CHAPTER 3

MARKOV MODULATED RELIABILITY DECISION MODEL

3.1. Markov Modulated Nature

Markov modulated environment indicates an external environmental condition "*Nature*", which has an impact on the state of the system and can be modeled by a Markov Chain, called "*Nature Process*", and cannot be observed directly. Only the stress of the system can be measured/observed, which is partially a result of Nature. We assume that Nature Markov Chain has state space SN= {Good (G), Bad (B)} and can be defined by the transition probability matrix, Q. Nature here represents any factor that cannot be observed easily but has an effect on the state of the system, such as outside temperature for a device in a box, network voltage for a converter, working state of another device for an electrical unit using same current, etc.

An example unit is a silicon carbide-based photovoltaic converter, which is affected by environmental factors. External Markovian factor (Nature) is assumed to be sunload, namely the intensity of sunlight on the unit. Sunload can be measured by using special sensors which frequently require maintenance or repair. Users usually prefer to measure resulting factors, such as stress, instead of using these sensors. Stress factor is voltage that the unit is exposed to. The unit fails at a certain level of voltage and this level is dependent on other environmental factors.

The voltage generated by solar panel increases by increasing sunload, therefore the voltage the unit is exposed to, namely stress increases. Increasing sunload also increases the temperature in the power converter box and consequently the breakdown voltage of the unit increases. Higher temperature enables all silicon carbide-based semiconductor products to sustain higher levels of voltage. There are many actions to affect voltage level of the unit. We may use smart sensors and regulators to control extreme stress levels, which decrease stress variance. These sensors are costly, it is beneficial to evaluate if it is worth to buy sensors or not. Another action might be changing solar panel position or angle in order to decrease the voltage caused by sunload. We may apply a comprehensive maintenance to ensure the lowest stress level that the unit can have, which corresponds to “replace” action in our model.

Another example unit might be a high voltage high power supply unit which is in a case but still affected by external factors. Nature factor is air temperature, stress factor is humidity in the case. When the air temperature gets higher, the humidity increases because of vaporization. With increasing temperature, due to the physical changes of the components, the humidity level the unit can sustains decreases. A certain level of humidity causes failure of the unit. There are many mean changing actions related to air conditioning. Cost of the action depends on how much energy is spent for air conditioning. Variance changing actions might be covering the case by a material to control extremely high humidity levels. After failure, the air in the case should be purified by a special air-purifier to make humidity level as low as possible. This action corresponds to “replace” action in our model and computations.

3.2. The Decision Process

There are decisions/actions available to us to control the system stress in a “desirable” fashion determined by our objective. We now explain the decision process in a timely order.

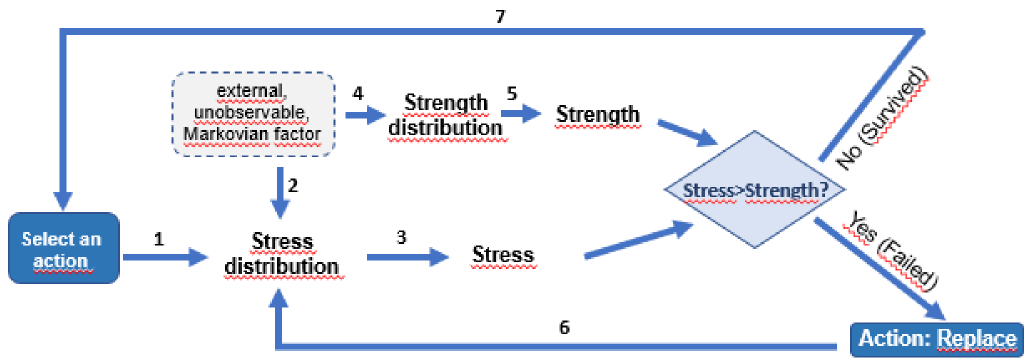

Figure 3.1: Decision Process Flow

The flow of the process can be seen in Figure 3.1. We assumed that there are three actions available from action set A= {1,2,3, replace}, each producing different mean or variance for the stress of the system. "Replace" action is reserved for a failed system only; we have to replace a failed system (arrow 6 in Figure 3.1). At each period, the unobserved Nature evolves with a Markovian transition. *i)* It produces a random strength (arrow 4 and arrow 5 in Figure 3.1) that is also unobservable to us from a known distribution depending on the state of the Nature. *ii)* An action is applied to the system at every period (arrow 7 in Figure 3.1). Together with the unobserved state of the Nature, the action results in a known probability distribution for the stress of the system (arrow 1 and arrow 2 in Figure 3.1). *iii)* Then we observe/measure a stress realization from that distribution (arrow 3 in Figure 3.1).

We first describe the details and computation of the parameters of the Markov decision process (MDP). Then we define MDP under consideration. Finally, we construct the mathematical model to find the optimal policy of that MDP.

The problem that this study deals with is determining the policy based on the state variables that should be measured/observed. There are available actions and an action is defined as a control on the stress of the process with a cost. That is, each action together with the current Nature leads to a stress distribution for the next period and we observe the resulting stress. On the other hand, the distribution of the unobservable strength is determined by the Nature only. We assume that the strength is Y_k when

Nature is k and is distributed with N (μ_{yk}, σ_{yk}^2) and the stress is X_{ka} when Nature is k and action is "a" and is distributed with N (μ_{xka}, σ_{xka}^2) where k ϵ SN and a ϵ A, A ={1, 2, 3, replace}, the set of available actions. Actions in our model only effect the mean and the variance of the resulting stress. It is possible to model actions effecting the strength also. Distribution of X corresponding to each state of Nature under each action can be seen in Table 3.1.

Table 3.1. *Distributions of the stress X corresponding to state of Nature under each action*

		Nature	
		Good	**Bad**
Action	**1**	X_{g1}~$N(\mu_{xg1}, \sigma_{xg1}^2)$	X_{b1}~$N(\mu_{xb1}, \sigma_{xb1}^2)$
	2	X_{g2}~$N(\mu_{xg2}, \sigma_{xg2}^2)$	X_{b2}~$N(\mu_{xb2}, \sigma_{xb2}^2)$
	3	X_{g3}~$N(\mu_{xg3}, \sigma_{xg3}^2)$	X_{b3}~$N(\mu_{xb3}, \sigma_{xb3}^2)$

The natural state definition of such a process would be (X, Y). There are some difficulties with that state definition:

i) The strength Y is unobservable; we cannot base decisions on an unobservable process.

ii) Even if we observe Y, it would not help to construct the transition probabilities because rather than the particular realization of Y, the distribution generated by the nature determines the strength in the next period. Since Nature is also is unobservable, the actions should not be based on that process either.

iii) The stress X is observable with a continuous set of values, but as with the strength, the stress distribution generated by the nature and the selected action determines the value in the next period rather than the particular realization of X.

So, for eliminating these difficulties, we construct the state of the MDP with the properties:

i) The strength Y comes in our state definition only through the information if a failure occurs ($X \geq Y$) or not ($X < Y$). If failure occurs, the system is immediately replaced.
ii) Since the Nature is unobservable, the same action is used whatever the unobserved state of Nature is.
iii) We discretize the stress into a finite number of classes as low, medium and high in order to obtain a discrete state space. Clearly, a more detailed stress classification results in more realistic representation of the system.

Therefore, the state keeps the information that if the system is up or failed and the level of the observed stress together with the Nature state. To discretize stress, we select two threshold levels T1 and T2 and defined three stress classes as summarized in Table 3.2. At any time point, what we can observe is a (discrete) *observation process O(n),* taking values form *the observation set* SO = {Low-Up (L), Medium-Up (M), High-Up (H), Failed (F)}, where L, M and H are as given in Table 3.2 and "Up" indicates that X<Y so that the system is working and "Failure" indicates that X ≥Y. Possible observations can be seen in Figure 3.2 in (X, Y) plane.

Table 3.2. *Possible observations*

Definition	Observation	Represantation
X<T1, X<Y	Low-Up	L
T1≤X <T2, X<Y	Medium-Up	M
T2≤X, X<Y	High-Up	H
X≥Y	Failed	F

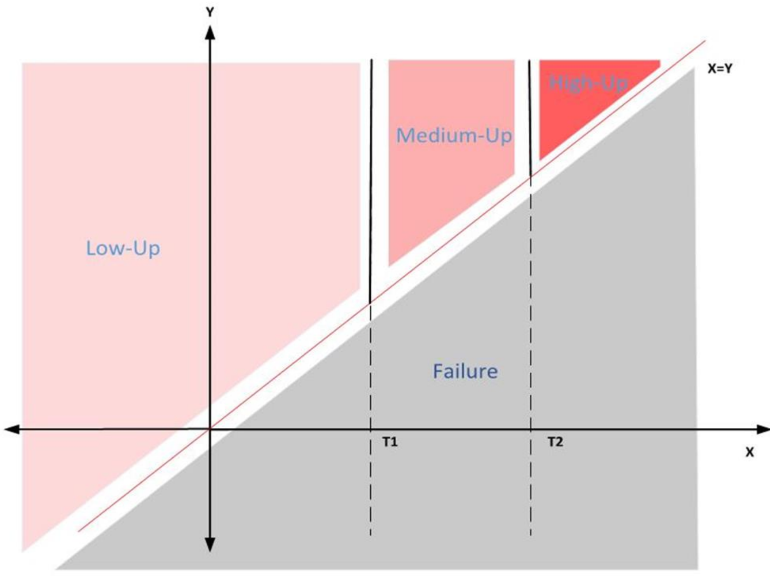

Figure 3.2: Observations

The observation process O(n) is not Markovian under the actions in A. Now we construct the state of the MDP under consideration at time n as (N(n), O(n)) where N(n) is the Nature, O(n) is the observation and A(n) is the action taken at time n. Therefore, the state space is

S = {(G, L), (G, M), (G, H), (B, L), (B, M), (B, H), Failure (F)}

= {(k, O): k ∈ SN, O ∈ SO}

where G means "Good" and B means "Bad".

3.3. Transition Probabilities

For each action $a \in$ A, we need to compute the transition probabilities

$p_{(k1,O1),(k2,O2)}(a) =$

$P\{N(n+1) = k2,\ O(n+1) = O2 |\ N(n) = k1,\ O(n) = O1,\ A(n) = a\}$

and construct the transition matrix

$P(a) = [p_{(\text{k1,O1}),(\text{k2,O2})}(a)]$ =

$$\begin{bmatrix} P_{(G,L)(G,L)}(a) & P_{(G,L)(G,M)}(a) & P_{(G,L)(G,H)}(a) & P_{(G,L)(B,L)}(a) & P_{(G,L)(B,M)}(a) & P_{(G,L)(B,H)}(a) & P_{(G,L)F}(a) \\ P_{(G,M)(G,L)}(a) & P_{(G,M)(G,M)}(a) & P_{(G,M)(G,H)}(a) & P_{(G,M)(B,L)}(a) & P_{(G,M)(B,M)}(a) & P_{(G,M)(B,H)}(a) & P_{(G,M)F}(a) \\ P_{(G,H)(G,L)}(a) & P_{(G,H)(G,M)}(a) & P_{(G,H)(G,H)}(a) & P_{(G,H)(B,L)}(a) & P_{(G,H)(B,M)}(a) & P_{(G,H)(B,H)}(a) & P_{(G,H)F}(a) \\ .. & .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. & .. \\ P_{F(G,L)}(a) & P_{F(G,M)}(a) & .. & .. & .. & .. & \mathrm{P}_{FF}(a) \end{bmatrix}$$

Take the transition from (k1, O1) to (k2, L), that is, observing low stress (and the system is up) and Nature =k2 after taking action a in Nature=k1 can be defined as

$$P_{(k1,O1)(k2,L)}(a) \;=\; P\{N(n+1) = k2,\; O(n+1) = L \mid N(n) = k1,\; O(n) = O1,\; A(n) = a\}$$

$$=\; Q_{k1k2}\, P\{X_{k2a} <\; Y_{k2},\; X_{k2a} < \mathrm{T1}\}$$

where (k1, O1), (k2, L) ∈ S, a ∈ A. Note that the stress at time n, $O(n) = O1$, does not affect the stress or Nature at time n+1. All transition probabilities are independent from the stress of the previous state, they are the same for the starting states (N(n), O(n)) =(G,L), (G, M), (G, H) and they are the same for the starting states (N(n), O(n)) =(B,L), (B, M), (B, H); that is,

$$P_{(G,L)(k2,O2)}(a) =\; P_{(G,M)(k2,O2)}(a) =\; P_{(G,H)(k2,O2)}(a)$$

and

$$P_{(B,L)(k2,O2)}(a) =\; P_{(B,M)(k2,O2)}(a) =\; P_{(B,H)(k2,O2)}(a)$$

for a ∈ A. This results in transition probability matrices having repeated rows for two groups of states.

Similarly,

$$P_{(k1,O)(k2,M)}(a) =\; P\{N(n+1) = k2,\; O(n+1) = M \mid N(n) = k1,\; O(n),\; A(n) = a\}$$

$$=\; Q_{k1k2} P\{X_{k2a} <\; Y_{k2},\; \mathrm{T1} < X_{k2a} < \mathrm{T2}\}$$

$P_{(k1,O)(k2,H)}(a) = P\{N(n+1) = k2,\ O(n+1) = H \mid N(n) = k1,\ O(n),\ A(n) = a\}$

$$= Q_{k1k2}P\{X_{k2a} < Y_{k2},\ X_{k2a} > \mathrm{T2}\}$$

For example, probability that, we will observe a low stress and the unobservable Nature is good in the next period if we observed a low stress and the unobservable Nature was good in the current period under action 1 is

$P_{(G,L)(G,L)}(1) = P\{N(n+1) = G,\ O(n+1) = L \mid N(n) = G,\ O(n),\ A(n) = 1\}$

$$= Q_{11}P\{X_{g1} < Y_g,\ X_{g1} < \mathrm{T1}\}$$

where $X_{g1} \sim N(\mu_{xg1}, \sigma_{xg1}^2), Y \sim N(\mu_{yg}, \sigma_{yg}^2)$

We take the second term on the right-hand side:

$$P\{X_{g1} < Y_g, \quad X_{g1} < \mathrm{T1}\} = \int_{-\infty}^{\mathrm{T1}} f_X dx \int_{x}^{\infty} f_Y dy$$

Since both variables are normally distributed, we use numerical approximation by discretizing integral as follows:

$$\int_{x}^{\infty} f_Y dy = 1 - F_Y(x)$$

where F_y is the cdf of Y that can be computed in MATLAB.

$P\{\mu_{\mathrm{xg1}} - 6\sigma_{\mathrm{xg1}} < X_{g1} < \mu_{\mathrm{xg1}} + 6\sigma_{\mathrm{xg1}}\} > 0.99$ since X is normally distributed,

$$P\{X_{g1} < Y_g, \quad X_{g1} < \mathrm{T1}\} \cong \int_{\mu_{\mathrm{xg1}} - 6\sigma_{\mathrm{xg1}}}^{\mathrm{T1}} f_X(x)\left(1 - F_y(x)\right) dx$$

$$\cong \sum_{k=1}^{n} f_X(\mu_{xg1} - 6\sigma_{xg1} + k\Delta x)\left(1 - F_Y(\mu_{xg1} - 6\sigma_{xg1} + k\Delta x)\right)\Delta x$$

where $n = \frac{T1-(\mu_{xg1}-6\sigma_{xg1})}{\Delta x}$ and Δx is sufficiently small.

Similarly,

$P_{(G,L)(G,M)}(1) = Q_{11}P\{X_{g1} < Y_g,\ T1 < X_{g1} < T2\}$

where $X_{g1} \sim N(\mu_{xg1}, \sigma^2_{xg1}), Y \sim N(\mu_{yg}, \sigma^2_{yg})$

$$P\{X_{g1} < Y_g, \quad T1 < X_{g1} < T2\} \cong \int_{T1}^{T2} f_X(x)(1 - F_Y(x))dx$$

$$\cong \sum_{k=1}^{n} f_X(T1 + k\Delta x)\left(1 - F_Y(\mu_{xg1} - 6\sigma_{xg1})\right)\Delta x$$

where $n = \frac{T2-T1}{\Delta x}$ and Δx is sufficiently small.

$P_{(G,L)(G,H)}(1) = Q_{11}P\{X_{g1} < Y_g,\ X_{g1} > T2\}$

where $X_{g1} \sim N(\mu_{xg1}, \sigma^2_{xg1}), Y \sim N(\mu_{yg}, \sigma^2_{yg})$

Also,

$$P\{X_{g1} < Y_g,\ X_{g1} > T2\} \cong \int_{T2}^{\mu xg1+6\sigma xg1} f_X(x)(1 - F_Y(x))dx$$

$$\cong \sum_{k=1}^{n} f_X(T2 + k\Delta x)\left(1 - F_Y(\mu_{xg1} - 6\sigma_{xg1} + k\Delta x)\right)\Delta x$$

where $n = \frac{\mu_{xg1}+6\sigma_{xg1}-T1}{\Delta x}$ Δx is sufficiently small.

Transition probabilities of other states and actions are calculated in the same manner.

When the process is in Failure (F) state, the system is replaced. If Nature is in state k, replacement leads X to be normally distributed with mean μ_{newk} and variance σ_{newk}^2 where k ∈ SN. Distribution parameters of newly replaced unit are dependent on the Nature condition and can be seen in Table 3.3.

Table 3.3. *Distributions of X after replacement*

Nature	**Good**	$X_{Rg} \sim N(\mu_{newg}, \sigma_{newg}^2)$
	Bad	$X_{Rb} \sim N(\mu_{newb}, \sigma_{newb}^2)$

Steady state probability of Nature being Good, π_G and being Bad, $\pi_B = 1 - \pi_G$, are used to calculate transition probabilities from state F as

$$P_{F(G,L)}(replace) = \pi_G P\{X_{Rg} < Y_g,\ X_{Rg} < \text{T1}\}$$

$$P_{F(G,M)}(replace) = \pi_G P\{X_{Rg} < Y_g,\ \text{T1} < X_{Rg} < \text{T2}\}$$

$$P_{F(G,H)}(replace) = \pi_G P\{X_{Rg} < Y_g,\ T2 < X_{Rg}\}$$

$$P_{F(B,L)}(replace) = \pi_B P\{X_{Rb} < Y_b,\ X_{Rg} < \text{T1}\}$$

$$P_{F(B,M)}(replace) = \pi_B P\{X_{Rb} < Y_b,\ \text{T1} < X_{Rb} < \text{T2}\}$$

$$P_{F(B,H)}(replace) = \pi_B P\{X_{Rb} < Y_b,\ T2 < X_{Rb}\}$$

where $X_{Rg} \sim N(\mu_{newg}, \sigma_{newg}^2), X_{Rb} \sim N(\mu_{newb}, \sigma_{newb}^2),$

$Y_g \sim N(\mu_{yg}, \sigma_{yg}^2)$ *and* $Y_b \sim N(\mu_{yb}, \sigma_{yb}^2)$

Cost of each action depends on the current stress and Nature conditions. Immediate (one period) cost of operating such a system usually depends on the current state, the action taken and possibly the state transited that can be unconditioned. If the random cost of being in state (N(n), O(n)) and taking action A(n) is

$C\left(\left(N(n), O(n)\right), A(n)\right)$ then the expected costs of state-action pairs $(k, O) \in$ S and $a \in$ A

$$C_{(k,O)a} = E\left(C\left(\left(N(n), O(n)\right), A(n)\right) \mid N(n) = k, O(n) = O, A(n) = a\right)$$

are computed for the present process.

After defining and computing all related parameters we have a well-defined Markov Decision Process {(N(n), O(n), A(n)): n=0, 1, …} where the state (N(n), O(n)) ∈S, A(n)∈A, with transition matrices P(a), for each a∈A, and the expected cost of taking action a and being in state (k, O) is $C_{(k,O)a}$. We can use one of the available methods, linear programming model, policy iteration algorithm or value iteration algorithm to find the optimal policy to minimize long-run average cost. We prefer the linear programming model because observability restrictions can be handled by linear programming relatively easily. We want to find policies to minimize the average cost over an infinite horizon.

3.4. Linear Programming Model for the MDP

The Markov chain defined above is unichain under any policy. So, we construct the following linear programming model, called Stress Strength Reliability model (SSR), for the MDP problem with parameters defined in the previous section.

$$Min \sum_{a \in A} \sum_{(k,O) \in S} C_{(k,O)a} \pi_{(k,O)a} \tag{3.1}$$

subject to

$$\sum_{a \in A} \pi_{(k2,O2)a} = \sum_{a \in A} \sum_{i \in S} \pi_{(k1,O1)a} P_{(\mathrm{k1,O1}),(\mathrm{k2,O2})}(a) \;\; \forall\, (k2, O2) \in S \tag{3.2}$$

$$\sum_{a \in A} \sum_{(k,O) \in S} \pi_{(k,O)a} = 1 \tag{3.3}$$

$$\pi_{(k,O)a} \geq 0 \qquad \forall\ (k,O) \in S,\ a \in A \tag{3.4}$$

where $\pi_{(k,O)a}$ is the steady state probability of taking action a and being in state (k,O).

The objective function (3.1) includes average long-term cost of the policy. Constraints (3.2) - (3.3) ensure $\pi_{(k,O)a}$ values being steady state probabilities. SSR model gives a policy to minimize average long run cost for each state. From the solution of this linear program, optimal policy will be obtained. $\pi_{(k,O)a}$ values, which are steady state probabilities, indicate the optimal policy. If $\pi_{(k,O)a}$ >0, this means that the policy uses action a when state is i. However, state (k,O) includes unobservable information about Nature conditions. For this reason, the "feasible" policy for our problem should give actions according to the observations only. For instance, assigning different decisions for (G, L) and (B, L) is not "feasible", due to the lack of information about Nature condition. In other words, (G, L) and (B, L) states are not distinguishable for the decision maker. That is why, the model should not allow different decisions for the states belong to same observation, such as (G, L) and (B, L). Hence, we introduce constraints (3.5) - (3.7) to ensure same decisions for states (G, L) and (B, L), (G, M) and (B, M), (G, H) and (B, H) as unobservability is handled by Serin (2005).

$$\frac{\pi_{(G,L)a}}{\sum_{l\in A}\pi_{(G,L)l}} = \frac{\pi_{(B,L)a}}{\sum_{l\in A}\pi_{(B,L)l}} \qquad \forall\ a \in A = 1,2,3 \tag{3.5}$$

$$\frac{\pi_{(G,M)a}}{\sum_{l\in A}\pi_{(G,M)l}} = \frac{\pi_{(B,M)a}}{\sum_{l\in A}\pi_{(B,M)l}} \qquad \forall\ a \in A = 1,2,3 \tag{3.6}$$

$$\frac{\pi_{(G,H)a}}{\sum_{l\in A}\pi_{(G,H)l}} = \frac{\pi_{(B,H)a}}{\sum_{l\in A}\pi_{(B,H)l}} \qquad \forall\ a \in A = 1,2,3 \tag{3.7}$$

SSR Model with these constraints is now called Markov Modulated Stress Strength Reliability (MMSSR) model.

3.5. The Main Characteristics of the Model

It is well-known that there is a deterministic optimal policy for model SSR. A proof of this result uses two facts:

i) The states are positive recurrent

ii) In any basis of this linear program, there can be at most $|S|$ positive variables.

Hence, the extreme points of the linear program correspond to the deterministic policies. So, it can easily be proved that if L linear constraints are added to the model there can be at most L randomization as shown in Ross (1989). For example, in case of a single constraint, the optimal policy randomizes between two actions in one state; that is for the state (k, O), $\pi_{(k,O)a1} > 0$, $\pi_{(k,O)a2} > 0$ and $\pi_{(k,O)a} = 0$ for $a \neq a1$, $a \neq a2$. Adding more constraints may increase number of randomizations. MMSSR model is obtained by adding non-linear constraints to SSR. So, we expect randomization in the optimal policy, but number of randomizations is difficult to estimate.

Action structure is one of the main characteristics of our model. The actions do not have fixed transition probabilities for state pairs. Each action leads normal distribution of stress with a different parameter set. So, the distributions related to the actions are independent from the current stress, namely the first parameter of the states. The distributions produced by the actions depend only on the second parameter of the current state, which is nature condition. This leads us to have a special kind of transition matrix for each action. The probabilities are the same for different current stress level categories, but different for different nature conditions. Using this special structure, actions as “do nothing” or “reinforce one level” are not relevant for our model.

CHAPTER 4

COMPUTATIONAL RESULTS

4.1. A Representative Example

We compute the transition probabilities for SSR and MMSSR in MATLAB, since the transition probabilities of the model need to be evaluated using numerical integration. We used GAMS Minos solver to solve SSR (linear) and MMSSR (nonlinear) models.

We first give the parameters of the MDP process. We used two Nature states, three actions and three stress level categories in most of our computations and parametric analysis, however, we also give an example with high number of actions and stress level categories to see the performance of the model better.

The transition probability matrix, Q, for the Nature process with state space SN={Good, Bad} is assumed to be

$$Q = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

Distribution parameters corresponding to actions for the representative example can be seen in Table 4.1.

Table 4.1. *Distributions parameters of the stress X corresponding to state of Nature under each action*

		Nature	
		Good	**Bad**
Action	**1**	X_{g1}~N(5, 10)	X_{b1}~ N(5, 10)
	2	X_{g2}~N(7, 10)	X_{b2}~ N(7, 10)
	3	X_{g3}~N(7, 5)	X_{b3}~N(7, 5)

Distribution parameters used for newly replaced units stress are shown in Table 4.2.

Table 4.2. *Distribution parameters of X after replacement*

Nature	**Good**	X_{Rg}~N(4, 2)
	Bad	X_{Rb}~ N(4, 2)

Strength distribution parameters differs according to the nature conditions. Bad nature causes lower strength. Assumed parameters can be seen in Table 4.3.

Table 4.3. *Distribution parameters of Y*

Nature	**Good**	Y~N(10, 4)
	Bad	Y~N(8, 4)

Threshold values to map the stress level to intervals are T1=6, T2=11.

Cost of actions are given in Table 4.4. Actions that decrease the stress level that the unit is exposed to, may correspond to adding parallel units or increasing tolerance of the unit, namely investment. Therefore, these actions are relatively expensive. The most expensive action we use is action 2, which leads to stress with the lowest mean. Action 3 decreases the stress variance, for example using a regulator to control extreme stress levels, also requires investment, and has higher cost than action 1, but not as expensive as mean stress decreasing actions.

Table 4.4. *Cost of actions for each state*

		Actions		
		1	**2**	**3**
States	**(G, L)**	10	2	5
	(G, M)	300	20	80
	(G, H)	1000	60	100
	(B, L)	10	2	5
	(B, M)	300	20	80
	(B, H)	1000	60	100
	F	2000	2000	2000

The state F has a predetermined action, replace. Replacement cost is so high that the model tends not to get into state F. When SSR is run with these parameters, optimal long run average cost is 326.6 and $\pi_{(k,O)a}$ values are given in Table 4.5. The optimal policy shown in Table 4.6.

Table 4.5. *Positive $\pi_{(k,O)a}$ values of the optimal policy for SSR of the Representative Example*

		Actions		
		1	**2**	**3**
States	**(G, L)**	0.34642		
	(G, M)		0.15852	
	(G, H)			0.00345
	(B, L)	0.27035		
	(B, M)	0.07345		
	(B, H)		0.00037	
	F	0.14745		

Table 4.6. *Optimal Policy for SSR of the Representative Example*

State	**Action**
(G, L)	1
(G, M)	2
(G, H)	3
(B, L)	1
(B, M)	1
(B, H)	2

This policy offers action 2 for state (G, M), and action 1 for state (B, M). However, Nature conditions are not visible to us, therefore, (G, M) and (B, M) states are not distinguishable. That is why, MMSSR is run to have a “feasible” policy. Because of adding new constraints, optimal long run average cost is increased to 327.4 and optimal policy becomes independent of Nature conditions. Optimal $\pi_{(k,O)a}$ values and the optimal policy are given in Table 4.7 and Table 4.8 respectively. Note that optimal policy for SSR in Table 4.6 is not feasible for MMSSR.

Table 4.7. *Positive $\pi_{(k,O)a}$ values of the optimal policy for MMSSR of the Representative Example*

		Actions		
		1	**2**	**3**
States	**(G, L)**	0.37270		
	(G, M)	0.14812		
	(G, H)			0.00239
	(B, L)	0.27775		
	(B, M)	0.07137		
	(B, H)			0.00029
	F	0.12738		

Table 4.8. *Optimal Policy for MMSSR for the Representative Example*

State	**Action**
(G, L)	1
(G, M)	1
(G, H)	3
(B, L)	1
(B, M)	1
(B, H)	3

As can be seen, same actions are offered for states (G, L) and (B, L), (G, M) and (B, M), (G, H) and (B, H).

One important observation about the optimal policy for MMSSR is that it is not randomized as can be seen in Table 4.7 and Table 4.8. The constraints (3.5) - (3.7) cut the optimal solution for SSR but the optimal policy is a deterministic policy, again an extreme point of SSR. We encounter this result frequently in the computations below. We analyze this situation in the following sections.

4.2. Parametric Analysis

The parameters of the problem are the distribution parameters of the stress and the strength and the cost coefficients. In this section, we present the analysis about the effects of the parameter changes on the optimal policy and the long run average cost. We did not find optimal policies minimizing the probability of failure π_F (since replacement is compulsory in failure state) we keep the record of failure probability to see its trade off with the cost in this analysis.

4.2.1. Replacement Cost

Optimal policy and long run average cost are calculated for increasing replacement cost while other parameters are kept constant. Same input parameters of the representative example in 4.1. are used other than costs of actions. The cost parameters of the actions can be seen in Table 4.9. As can be seen, cost of each action changes depending on not only the observation level but also the nature condition.

Table 4.9. *Cost of actions for replacement cost analysis*

State	**1**	**2**	**3**
(G, L)	10	2	5
(G, M)	300	20	40
(G, H)	1000	60	90
(B, L)	10	2	5
(B, M)	600	50	80
(B, H)	1400	100	150
F	1000	1000	1000

The replacement cost is increased from 1000 to 12000 incrementally. The optimal policies encountered in this analysis are summarized in Table 4.10.

The optimal long run average costs, optimal policies as well as the failure probabilities π_F for SSR and MMSSR can be seen in Figure 4.1 and Figure 4.2 respectively. Policy changing points can also be seen in the figures.

Table 4.10. *Actions of the optimal policies encountered in parametric analyses*

	Policies											
State	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**
(G, L)	1	1	1	1	1	1	1	1	1	1	1	1
(G, M)	3	3	3	1	1	1	1	2	3	2	1	3
(G, H)	2	3	3	3	3	1	1	2	3	3	3	2
(B, L)	1	1	1	1	1	1	1	1	1	1	1	1
(B, M)	2	2	3	3	1	1	1	2	3	1	1	3
(B, H)	2	2	2	3	3	3	1	2	3	2	1	2

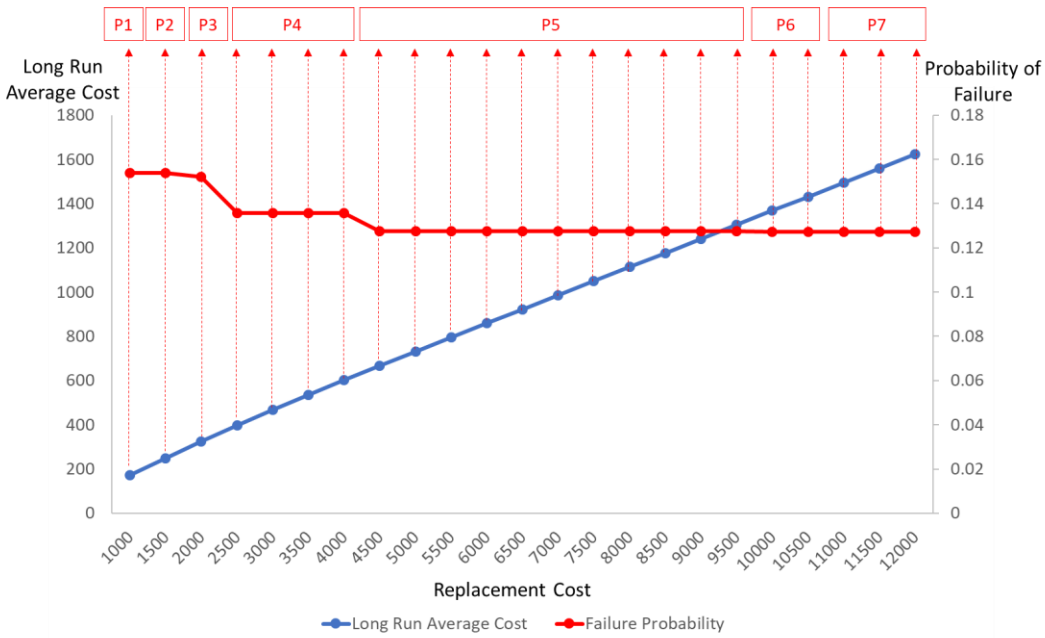

Figure 4.1: Optimal cost, failure probability π_F and optimal policy change against replacement cost for SSR

As can be seen from Figure 4.1, replacement cost between 1000 and 12000 give 7 different optimal policies assuming we do an exhaustive evaluation. Figure 4.2 shows the same information for MMSSR.

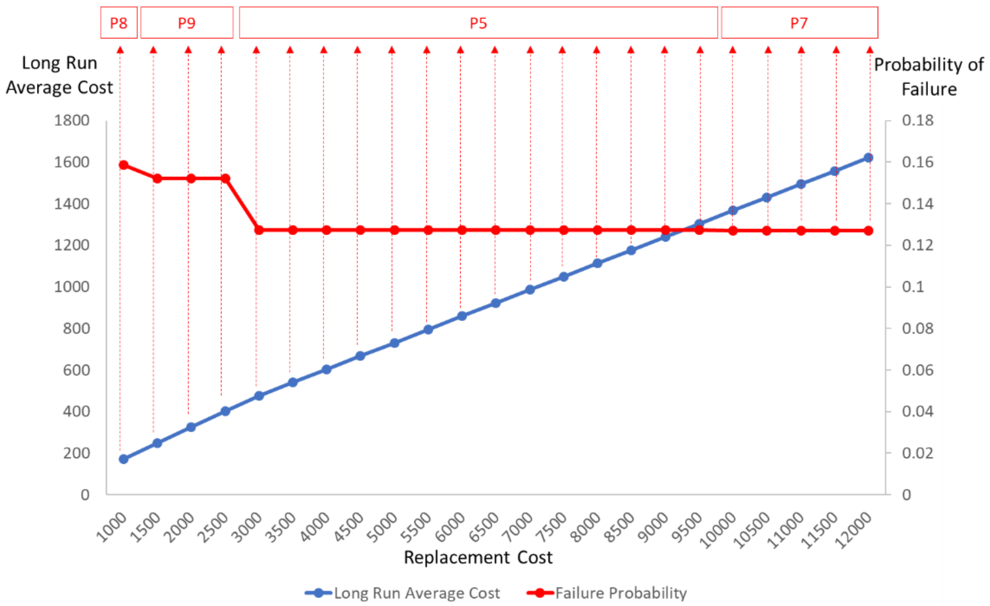

Figure 4.2: Optimal cost, failure probability π_F and optimal policy change against replacement cost for MMSSR

As can be seen, MMSSR proposes 4 different policies for the same range of replacement cost values. Both figures show that replacement cost is dominantly effective on the long run average cost. Probability of failure is a performance measure conflicting with the cost objective. Since the objective of the model is minimizing the cost, replacement cost is directly a penalty for the failure state. When the replacement cost increases while all other parameters are constant, the models prefer to be in Failure state less frequently. Consequently, failure probability decreases while the objective value increases.

We observe that the optimal costs are almost identical for SSR and MMSSR. This can be interpreted as MMSSR catches almost the same optimal cost by changing the optimal policy “slightly” to a feasible policy. Although the optimal costs are close to each other, the steady state distributions, hence the failure probabilities are quite different since the optimal policies are different in those two models.

4.2.2. New Unit Quality

A new unit is used after replacement. Keeping the cost of the new unit constant, we want to see how the stress level of the new unit affects the long run average cost and probability of failure. The model enables different new unit mean stress values for different nature conditions, however we used same values, called μ_{new}, for both nature conditions. Same input parameters of the representative example in 4.1. are used other than costs of actions and new unit mean stress level. The cost parameters of the actions used can be seen in Table 4.9. We change new unit mean stress (μ_{new}) from 0.5 to 9 incrementally. The optimal long run average costs, optimal policies as well as the failure probabilities π_F for SSR and MMSSR can be seen in Figure 4.3 and Figure 4.4 respectively.

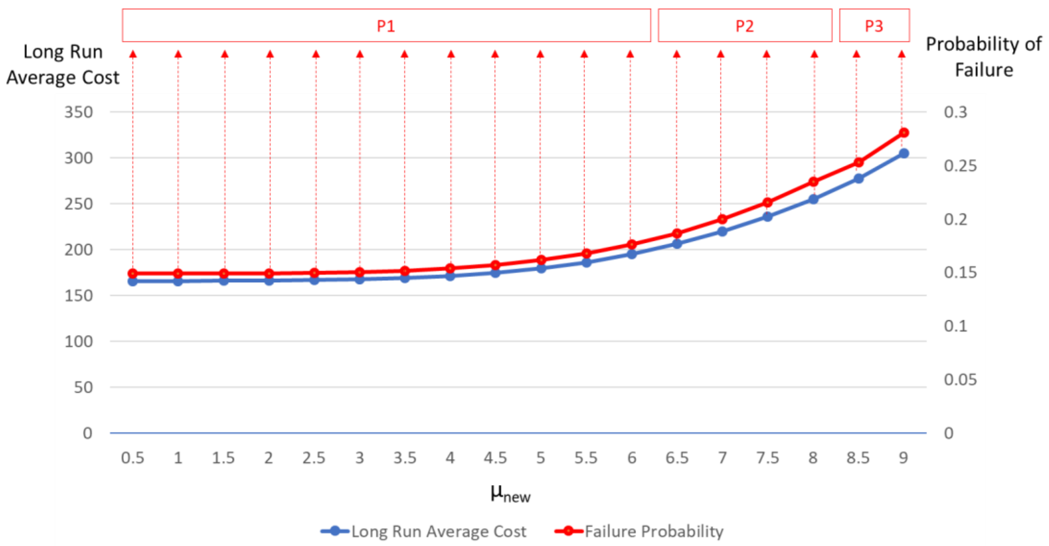

Figure 4.3: Optimal cost, failure probability π_F and policy change against μ_{new} for SSR

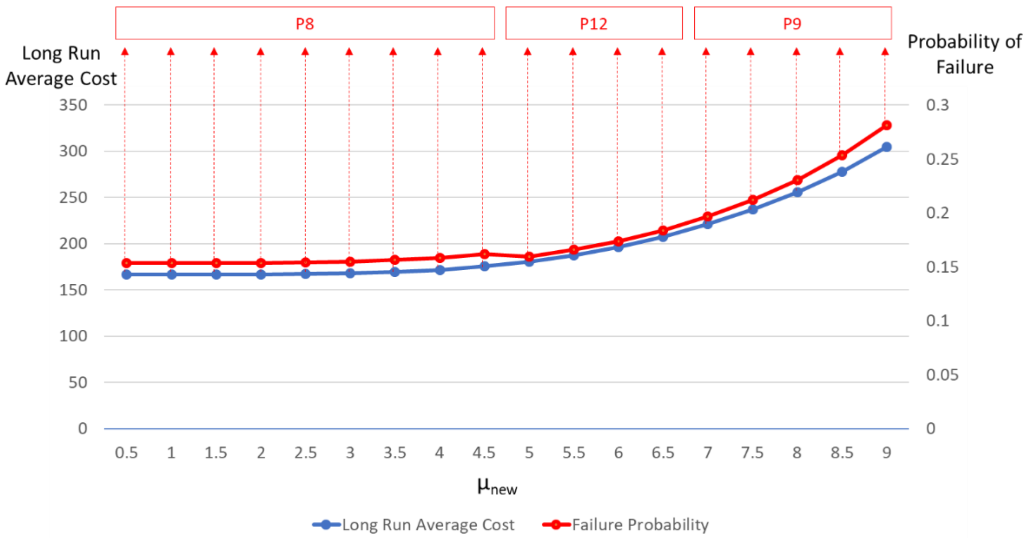

Figure 4.4: Optimal cost, failure probability π_F and policy change against μ_{new} for MMSSR

SSR and MMSSR both offer 3 different optimal policies in given μ_{new} range. Since P1, P2 and P3 do not satisfy MMSSR constraints, the resulting policies are different for both models. The graphs show that, as the mean stress of new unit increases, namely as the quality of the new unit gets worse while replacement cost is same, the failure probability increases, increasing also the cost of failure.

4.2.3. New Unit Quality and Cost

Selection of new unit is a potential issue for designers. We face with higher cost if we want to have more reliable products. Replacement cost increases when μ_{newg} and μ_{newb} decreases. As we assume μ_{newg} and μ_{newb} to be equal in our calculations, we name this value μ_{new}. Replacement cost and μ_{new} relation is assumed as shown in Figure 4.5.

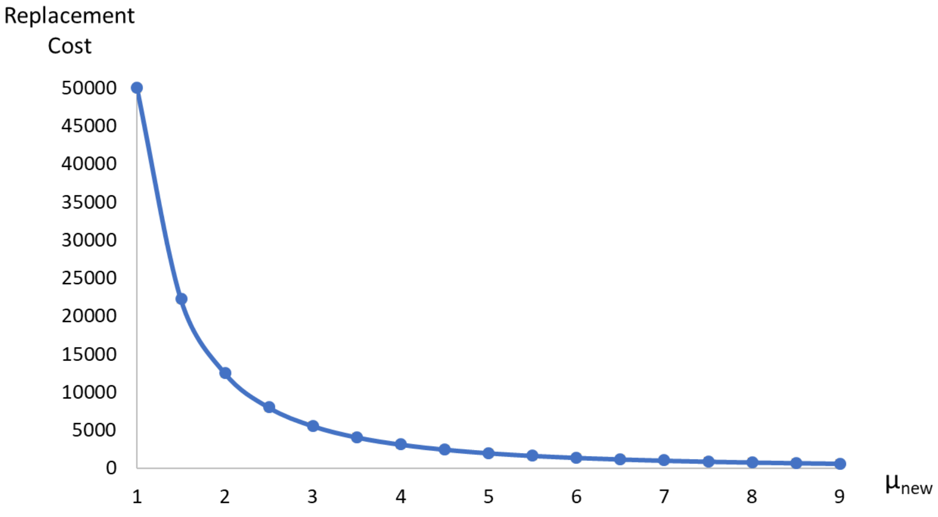

Figure 4.5: Replacement cost against the stress mean of new unit (μ_{new})

$$\text{Replacement Cost}(\mu_{\text{new}}) = 50000\ \mu_{\text{new}}^{-2} \tag{4.1}$$

Figure 4.6 and Figure 4.7 show both long run average cost, the optimal policy and the failure probability π_F change for different quality, namely mean stress values using SSR and MMSSR models respectively.

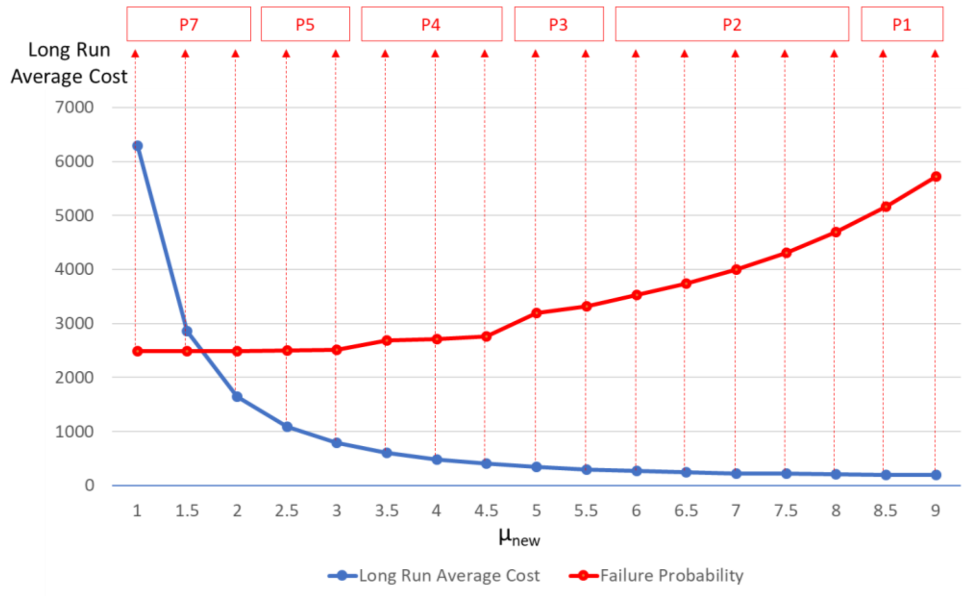

Figure 4.6: Optimal cost, failure probability π_F and optimal policy change against μ_{new} for SSR using equation (4.1)

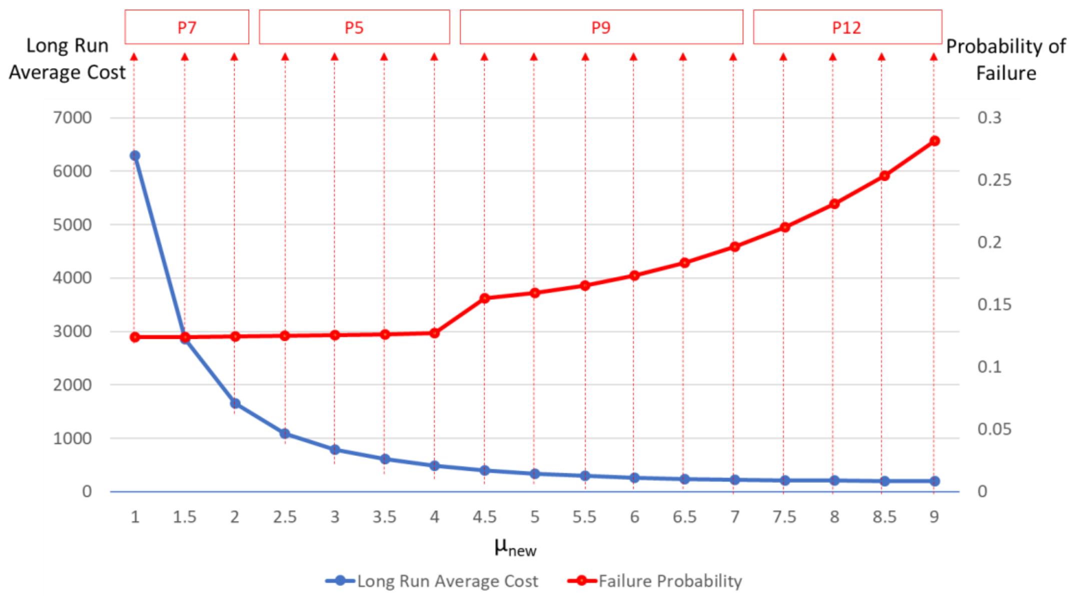

Figure 4.7: Optimal cost, failure probability π_F and optimal policy change against μ_{new} for MMSSR using equation (4.1)

As can be seen from the figures above, long run average failure probability is a conflicting measurement with long run average cost for our problem. The cost increases with decreasing mean stress level, while average failure probability decreases. This means, new units should be chosen considering vital importance of the unit. If number of failures is not a concern, then there is no reason to decrease mean stress level. If the unit is a safety-critical unit, then the best unit, namely the unit with the highest cost will be chosen with the related optimal policy.

4.3. Value of Information

SSR Model uses exact state information to find the optimal policy while MMSSR gives optimal policies that use only partial information about the state. In terms of mathematical models, MMSSR is SSR under some more constraints. Optimal objective in MMSSR ≥ Optimal objective in SSR. Failure probability comparison is difficult to make under current cost structure. The resulting difference of the objective values of SSR and MMSSR can be called *value of exact state information*. This would be interpreted as the maximum amount that the decision maker should be willing to pay to reveal the unobservable state. We define

$$\%\ value\ of\ information =$$

$$\frac{optimal\ objective\ value\ of\ MMSSR - optimal\ objective\ value\ of\ SSR}{optimal\ objective\ value\ of\ SSR}$$

For our models, value of exact state information at different replacement costs can be seen in Figure 4.8.

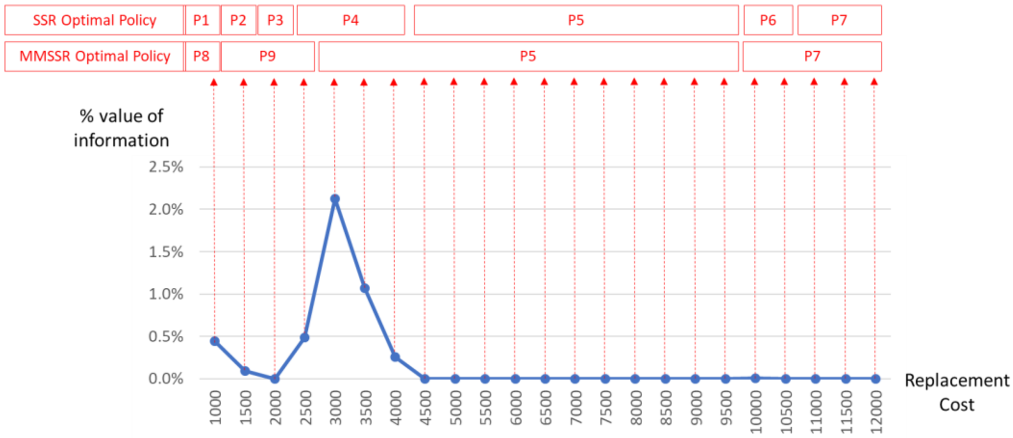

Figure 4.8: Value of information against replacement costs keeping mean stress fixed

Figure 4.8 is produced using optimal costs of Replacement Cost Analysis in 4.2.1. and shows how the additional constraints of MMSSR impacts long run average cost. The value of information is zero at points for which the optimal policies are the same.

For our models, value of exact state information at different mean of the stress of the new unit, μ_{new} can be seen Figure 4.9 that is produced using optimal costs of New Unit Mean Stress Analysis in 4.2.2. Here, we also define

$$\%\ gain\ in\ failure\ probability = \frac{\pi_F\ from\ MMSSR - \pi_F\ from\ SSR}{\pi_F\ from\ SSR}$$

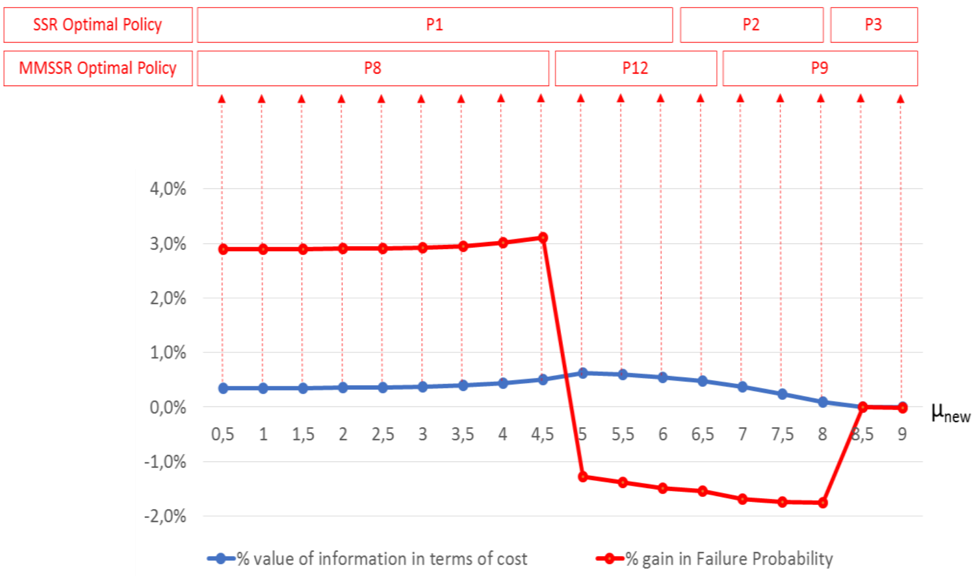

Figure 4.9: Value of information against mean stress keeping the replacement cost fixed

As can be seen in the figure, for the mean stress levels between 0.5 and 4.5, value of information in terms of cost and % gain in failure probability are positive. This means that, because of not knowing the nature condition, we must pay more and fail more frequently. However, for the mean levels higher than 4.5, MMSSR offers optimal policies with higher cost but less failure probabilities. The large change between μ_{new}= 4.5 and μ_{new}=5 is due the MMSSR optimal policy change from P8 to P12.

As stated before, the optimal costs we observe are very close for SSR and MMSSR. This is due to the input data we used in the models. Although our models allow differentiation of all parameters for different nature conditions such as resulting mean stress of actions, mean strength, new unit mean stress, cost of actions, we assumed same parameter values for "Good" and "Bad" nature conditions, to be able to observe effect of one parameter at a time. The main factor creating the value of information is the difference between cost of actions in different nature conditions. Since the cost of actions in different conditions are close in our computations, all the optimal policies have close long run average costs and the value of information can be considered as low.

The representative example in Section 5.1 the nature conditions do not change the cost of actions, as can be seen from Figure 4.4. To check the above argument, we run this representative example under the scenario that the costs are highly dependent on the nature condition. SSR and MMSSR models are solved under the same parameters but the costs in Table 4.11.

Table 4.11. *Alternative cost of actions for the representative example*

State	**1**	**2**	**3**
(G, L)	20	1	5
(G, M)	30	5	12
(G, H)	50	30	55
(B, L)	1000	50	250
(B, M)	1100	150	300
(B, H)	2000	600	450
F	2500	2500	2500

Under the cost structure given in Table 4.11, the optimal long run average cost of SSR and MMSSR are 465.7 and 587.8 respectively, which makes the % value of information %26.2.

4.4. Optimizing Other Performance Measures

4.4.1. Probability of Failure and First Passage Time to Failure

Probability of failure, π_F, another performance measure, can be the objective function to minimize for the model. MMSSR with failure probability minimization model is as follows.

$$Min\ \pi_{F1}+\pi_{F2}+\pi_{F3} \tag{3.8}$$

subject to (3.2) - (3.7)

Optimal policy of the model is observed to select the action with the lowest mean stress, which is the most expensive action as expected. This is mostly due to the selected means and the variances of the random strength and random stress in this study. The result may change with different selection of the parameters.

Another objective is maximizing the expected first passage time to failure state which gives the same results with minimizing failure probability. Expected first passage time (EFPM) from state i to j for any Markov Chain is defined as follows:

$$EFPM_{ij} = 1 + \sum_{t \neq j} P_{it} EFPM_{tj}$$

where P_{it} is transition probability from state i to state t and $EFPM_{ij}$ is expected first passage time from state i to state j.

If the steady state distribution of a policy R is $(\pi_{(k1,O1)a}, (k1, O1) \in S, a \in A)$ then the transition probabilities under that policy can be written as

$$P_{(\mathrm{k1,O1}),(\mathrm{k2,O2})}(R) = \sum_{a \in A} P_{(\mathrm{k1,O1}),(\mathrm{k2,O2})}(a) \frac{\pi_{(k1,O1)a}}{\sum_{l \in A} \pi_{(k1,O1)l}}$$

We use the following model for maximizing expected first passage time from state (G, L) to F.

$$Max\ EFPM_{(G,L)F} \qquad (3.9)$$

subject to

(3.2) - (3.7)

$$EFPM_{(k1,O1)F} = 1 + \sum_{(k1,O1)\neq(k2,O2)} \sum_{a\in A} P_{(k1,O1),(k2,O2)}(a) \frac{\pi_{(k1,O1)a}}{\sum_{l\in A} \pi_{(k1,O1)l}} EFPM_{(k2,O2)F}$$

$$\forall\ (k1, O1), (k2, O2) \in S \qquad (3.10)$$

We analyze expected first passage time to failure state and probability of failure as side performance measures when minimizing long run average cost. Figure 4.10 and Figure 4.11 present first passage time and failure probability levels of New Unit Mean Stress and Cost Analysis in 4.2.3.

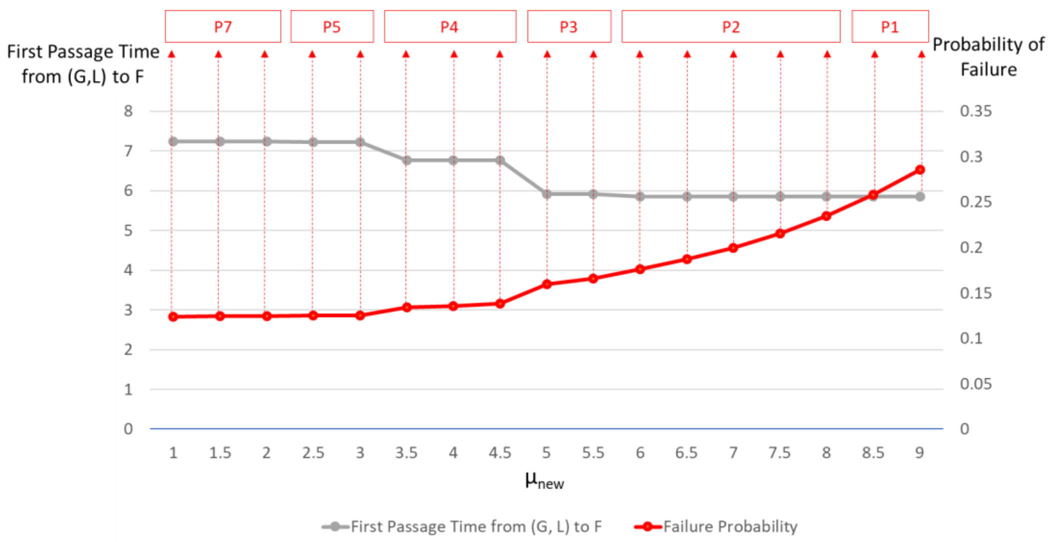

Figure 4.10: Failure Probability and First Passage Time change against Stress Mean for SSR

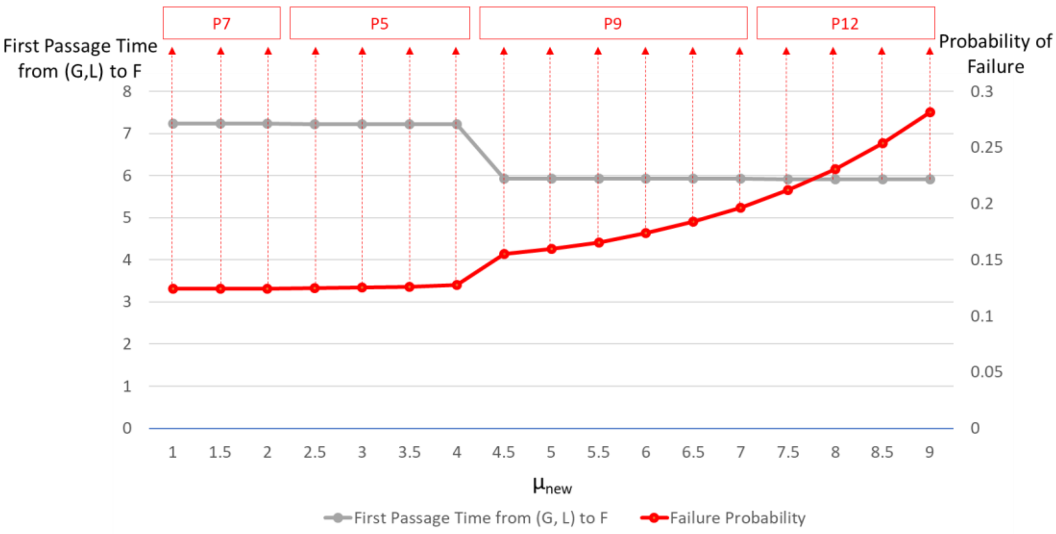

Figure 4.11: Failure Probability and First Passage Time change against Stress Mean for MMSSR

Figures show that, expected first passage time is decreasing in failure probability, as expected. Expected first passage time from other states to failure state also have the same behavior, which can be seen in Appendix A.

4.4.2. Example for Randomized Optimal Policies

For the original MDP problem with no additional constraints, we use a linear programming model, SSR. We know that there is a deterministic optimal policy. However, possibility of obtaining randomized solutions arises when we add constraints, especially nonlinear constraints to obtain MMSSR. With the present cost structure and parameter value, we have not obtained any randomized optimal policies. The optimal policy of SSR is cut by constraints (3.5) - (3.7) since they are not feasible for MMSSR most of the time, but the latter optimal policies are also deterministic with very close optimal objective values. Although desirable, this result makes us to ask the question if these constraints always lead to deterministic optima. So, we make some more trials to see if it is the original feasible set (original transition probabilities) or the constraints or the objective function that create this result. We observe that with some other objective functions, the optimal policies can be highly randomized.

We present some arbitrary objective functions with randomized optimal policies for MMSSR. An example case is taken from New Unit Mean Stress Level analysis. Input parameters can be found in Appendix B. MMSSR is run with the following objective functions:

min $2\pi_{(G,H)1} + \pi_{(B,H)3}$

min $\pi_{(G,M)3} + \pi_{(B,M)1}$

Resulting steady state probabilities are given in Table 4.12 and Table 4.13 respectively. As can be seen, randomized optimal policies come into the picture.

Table 4.12. *Optimal $\pi_{(k,O)a}$ values for MMSSR by minimizing $2\pi_{(G,H)1} + \pi_{(B,H)3}$*

		Actions		
		1	**2**	**3**
States	**(G, L)**	0.10572		0.13996
	(G, M)	0.23827		
	(G, H)		0.00347	
	(B, L)	0.07082		0.09376
	(B, M)	0.11601		
	(B, H)		0.00048	
	F	0.23151		

Table 4.13. *Optimal $\pi_{(k,O)a}$ values for MMSSR by minimizing $\pi_{(G,M)3} + \pi_{(B,M)1}$*

		Actions		
		1	**2**	**3**
States	**(G, L)**	0.19979	0.01511	
	(G, M)		0.24193	
	(G, H)	0.00552		
	(B, L)	0.13607	0.01029	
	(B, M)		0.11638	
	(B, H)	0.00071		
	F	0.27420		

4.5. Effect of Increasing Number of States and Number of Actions

Keeping the state space of Nature process and its transition probability matrix the same, we define a system with 6 stress level intervals and 10 actions to get a more precise solution and see the behavior of the model in detail Distributions corresponding to actions can be seen in Table 4.14.

Table 4.14. *Distributions of X in 10-action scenario*

		Nature	
		Good	**Bad**
Action	**1**	X_{g1}~N(2, 4)	X_{b1}~N(2, 4)
	2	X_{g2}~N(3, 4)	X_{b2}~N(3, 4)
	3	X_{g3}~N(4, 4)	X_{b3}~N(4, 4)
	4	X_{g4}~N(5, 4)	X_{b4}~N(5, 4)
	5	X_{g5}~N(6, 4)	X_{b5}~N(6, 4)
	6	X_{g6}~N(7, 4)	X_{b6}~N(7, 4)
	7	X_{g7}~N(8, 4)	X_{b7}~N(8, 4)
	8	X_{g8}~N(9, 4)	X_{b8}~N(9, 4)
	9	X_{g9}~N(10, 4)	X_{b9}~N(10, 4)
	10	X_{g10}~N(11, 4)	X_{b10}~N(11, 4)

Distribution parameters used for newly replaced units stress are shown in Table 4.15.

Table 4.15. *Distributions of X after replacement in 10-action scenario*

Nature	**Good**	X_{Rg}~N(3, 2)
	Bad	X_{Rb}~ N(2, 2)

Strength distribution parameters differs according to the nature conditions. Bad nature causes lower strength. Strength parameters can be seen in Table 4.16.

Table 4.16. *Distributions of Y in 10-action scenario*

Nature	**Good**	Y~N(10, 4)
	Bad	Y~N(7, 4)

Five threshold values are used to discretize stress level and map the values to intervals. Threshold levels and intervals are given in Table 4.17. Possible observations can be seen in Figure 4.12 in (X, Y) plane.

Table 4.17. *Possible observations for 10-action scenario*

Definition	**Observation**	**Representation**
X<2, X<Y	Level 1-Up	L1
2≤X <5, X<Y	Level 2-Up	L2
5≤X <8, X<Y	Level 3-Up	L3
8≤X <10, X<Y	Level 4-Up	L4
10≤X <13, X<Y	Level 5-Up	L5
13≤X, X<Y	Level 6-Up	L6
X≥Y	Failed	F

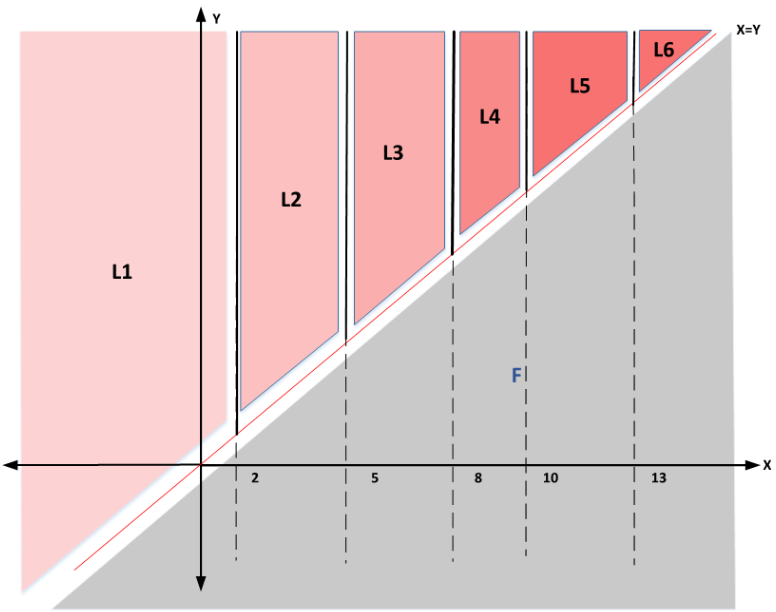

Figure 4.12: Observations for 10-action scenario

Cost of actions assumed in this scenario is given in Table 4.18.

Table 4.18. *Cost of actions for each state in 10-action scenario*

	Actions									
States	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(G, L1)	0	10	10	10	10	10	10	10	10	10
(G, L2)	500	385	280	205	155	115	80	50	30	20
(G, L3)	700	560	420	300	230	170	120	75	45	30
(G, L4)	1050	840	630	450	345	255	180	110	65	45
(G, L5)	1575	1260	945	675	520	375	270	165	100	68
(G, L6)	2350	1900	1400	1010	780	550	405	250	150	102
(B, L1)	0	10	10	10	10	10	10	10	10	10
(B, L2)	10000	7700	5600	4100	3100	2300	1600	1000	600	400
(B, L3)	14000	11200	8400	6000	4600	3400	2400	1500	900	600
(B, L4)	21000	16800	12600	9000	6900	5100	3600	2200	1300	900
(B, L5)	31500	25200	18900	13500	10400	7500	5400	3300	2000	1360
(B, L6)	47000	38000	28000	20200	15600	11000	8100	5000	3000	2040
F	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000

State 13 is the failure state, which has a predetermined action = replace, same as previous scenarios. When the first model, SSR is run with these parameters, long run average cost is 1606.5 and $\pi_{(k,O)a}$ values are given in Table 4.19.

Table 4.19. *Optimal $\pi_{(k,O)a}$ values for SSR of the 10-action scenario*

		Actions									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	(G, L1)	0.236									
	(G, L2)	0.237									
	(G, L3)	0.068									
	(G, L4)	0.016									
	(G, L5)	0.003									
	(G, L6)	>0									
States	**(B, L1)**	0.138									
	(B, L2)						0.150				
	(B, L3)								0.046		
	(B, L4)									0.006	
	(B, L5)									>0	
	(B, L6)										>0
	F	0.100									

MMSSR is run to determine action without nature condition information. Because of adding new constraints, the objective cost is increased to 2367.6 and optimal policy becomes independent of Nature conditions. $\Pi_{(k,O)a}$ values. The optimal $\pi_{(k,O)a}$ values can be seen in Table 4.20.

Table 4.20. *Optimal $\pi_{(k,O)a}$ values for MMSSR of the 10-action scenario*

		Actions									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
States	**(G, L1)**	0.129									
	(G, L2)			0.285							
	(G, L3)				0.141						
	(G, L4)	0.01									
	(G, L5)		>0								
	(G, L6)	>0									
	(B, L1)	0.096									
	(B, L2)			0.201							
	(B, L3)				0.069						
	(B, L4)	0.002									
	(B, L5)		>0								
	(B, L6)	>0									
	F	0.065									

Value of information is %47 for this example.

As shown in the previous examples, deterministic optimal policies of MMSSR is thanks to the objective function. We present three other objective functions with randomized optimal policies for MMSSR. We have observed that most of the possible objective functions, other than long run average cost, bring randomized optimal policies. Three arbitrary examples are as follows.

min $\pi_{(B,L1)1}$

min $\pi_{(B,L2)3} + \pi_{(B,L2)4} + \pi_{(B,L2)7}$

min $\pi_{(G,L4)4} + 3\pi_{(G,L4)5} + 7\pi_{(G,L5)4} + 2\pi_{(G,L5)5}$

Resulting steady state probabilities of MMSSR by minimizing $\pi_{(B,L1)1}$ is given in

Table 4.21. Optimal $\pi_{(k,O)a}$ values of MMSSR by minimizing the other two functions above can be seen in Appendix C.

Table 4.21. *Optimal $\pi_{(k,O)a}$ values for MMSSR by minimizing* $\pi_{(B,L1)1}$

		Actions									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
States	**(G, L1)**			0.046							
	(G, L2)				0.036			0.146			
	(G, L3)						0.192				0.006
	(G, L4)			0.068							
	(G, L5)				0.012						
	(G, L6)				>0						
	(B, L1)			0.039							
	(B, L2)				0.027			0.110			
	(B, L3)						0.084				0.003
	(B, L4)			0.012							
	(B, L5)				0.001						
	(B, L6)				>0						
	F	0.216									

CHAPTER 5

CONCLUSION AND FUTURE WORK

In this thesis, we focus on a system that survives a period depending on its strength and stress occurrences at that period. These two are random quantities generated by an unobservable Markovian environment as well as the actions applied at that period. Their distributions are known but only the level of realized stress information is available. The objective is to find an optimal policy that only uses available information to minimize the long run average cost of running this system. Although we expect randomized policies due to unobservability of some state information, the optimal policies are deterministic with the present cost function. We checked if this is a structural property of the transition matrices and we concluded that it is not. Randomized optimal policies exist with different objective functions. We also evaluate other performance measures such as probability of failure, mean time to failure. We investigate how the optimal policy and performance measures change by changing different parameters. We also measure the value of observing the unobservable variables by comparing the optimal objective values. The cost and failure probability trade-off is also measured.

The analysis made in this study can be extended in many future directions. Firstly, different objective functions can be tried. For instance, failure probability and long run average cost trade-off can further be analyzed using multiple objective methodology.

Expected first passage time can be written as a nonlinear function of model variables. So, we tried maximizing the expected time to failure objective. We observed that this gives the policy that minimizes the failure probability in our limited number of runs. Other functions of time to failure can also be considered.

The stress and the strength distributions can be taken in a stochastic order so that the transition matrices may have a structure such as increasing failure rate (IFR). In this case, a somewhat structured optimal policies may exist.

The stress and strength can also be taken as dependent random variables coming from a bivariate normal distribution. In such a model, we can also assume actions affecting both of them.

Another valuable study might be using actions to resolve unobservability sequentially as inspections with changing details, recovering some more information after every action although we used “maintenance” in a very broad sense, as changing distributions of the stress and the strength.

REFERENCES

[1] Jonhson, R. A. (1998). “Stress-Strength Models for Reliability”, *Handbook of Statistic*, 7:27-54.

[2] Kotz, S, Lumelskii, Y., Pensky M. (2003). “The Stress-Strength Model and Its Generalizations: Theory and Applications”, World Scientific.

[3] Dhillon, B. (1998). “Stress-Strength Reliability Models”, *Microelectronics Reliability*, 20(4):513–516.

[4] Siju, K. C., Kumar, M. (2016). “Reliability Analysis of Time Dependent Stress-Strength Model with Random Cycle Times”, *Perspectives in Science*, 8:654-657.

[5] Xie, M., Shen, K. (1991). “Some New Aspects of Stress-Strength Modelling”, *Relialbilitv Engineering and System Safety*, 33(1):131-140.

[6] Bhuyan, P., Dewanji, A. (2017). “Estimation of reliability with cumulative stress and strength degradation”, *Statistics*, 51 (4):766-781.

[7] Eryilmaz, S. (2011). “A new perspective to stress–strength models”, *Annals of the Institute of Statistical Mathematics*. 63(1):101-115.

[8] Qin, H., Jana, N., Kumar, S., Chatterjee, K. (2016). “Stress-strength Models with More than Two States under Exponential Distribution”, *Communications in Statistics - Theory and Methods*. 46(1):120-132.

[9] Godoy, D., Pascual, R., Knights, P. (2013). “Critical spare parts ordering decisions using conditional reliability and stochastic lead time”, *Reliability Engineering & System Safety*. 119:199-206.

[10] Ramesh, N. I. (1995). “Statistical analysis on markov-modulated poisson processes”, *Environmetrics.* 6:165-179.

[11] Scott, S., Smyth P (2003). “The Markov Modulated Poisson Process and Markov Poisson Cascade with Applications to Web Traffic Modelling”, *Bayesian Statistics*. 7:1-10.

[12] Andronov, A., Gertsbakh, I. (2014). “Signatures in Markov-Modulated Processes”, *Stochastic Models*. 30(1).1-15

[13] Drake, A. W. (1962). “Observation of a markov process through a noisy channel,” Ph.D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology.

[14] Corotis, R., Ellis, H. J., Jiang M. (2005). “Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision processes”, *Structure and Infrastructure Engineering*, 1(1):75-84.

[15] Serin, Y., Kulkarni V. G. (2005). “Markov decision processes under observability constraints”, *Mathematical Methods of Operations Research*, 61(2):311-32.

[16] Satır, B. (2010). “An analysis of benefits of inventory and service pooling and information sharing in spare parts management systems”, Ph.D. Thesis, Department of Industrial Engineering, Middle East Technical University.

[17] Altman, E. (1999). “Constrained Markov Decision Processes”, London: Chapman & Hall/CRC, 199

[18] Ross, K. W. (1989). “Randomized and Past-Dependent Policies for Markov Decision Processes with Multiple Constraints”, *Operations Research*, 37(3):474-477

APPENDIX A

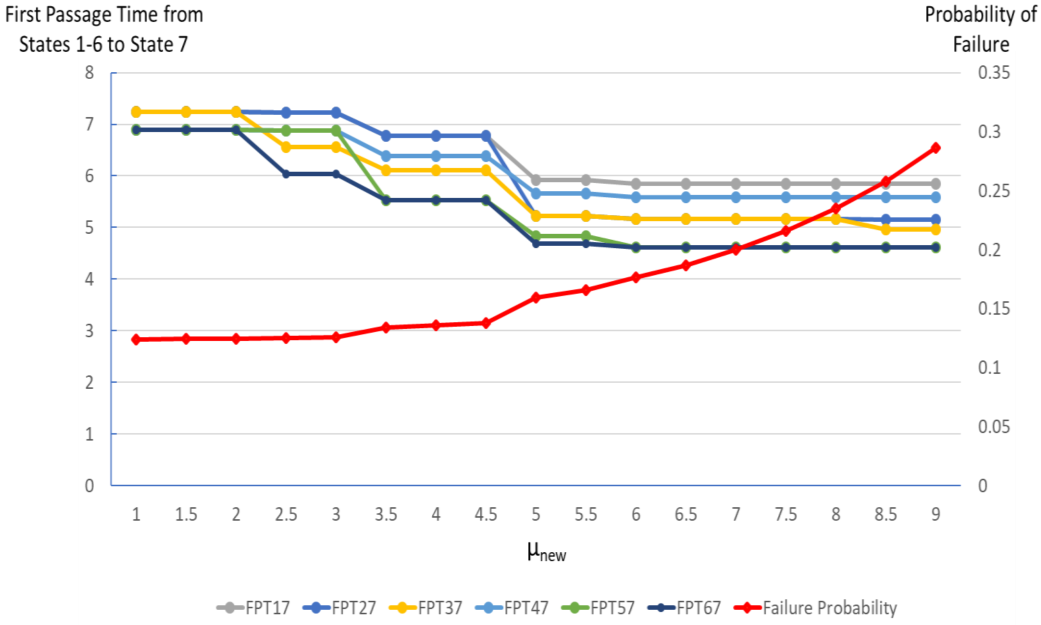

Figure A.1: First Passage Time from all 6 states to State F in SSR

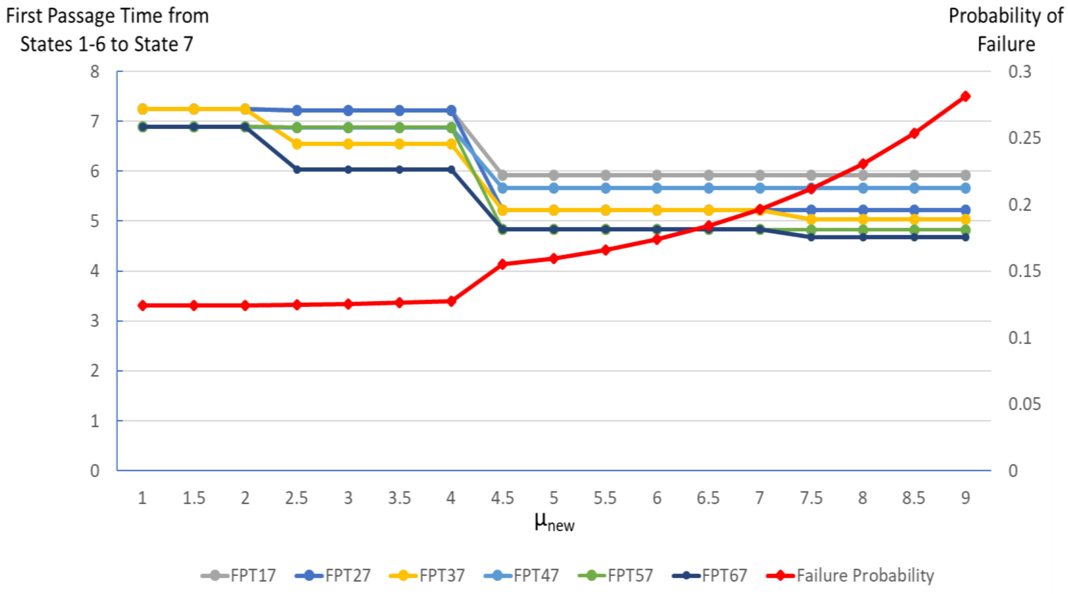

Figure A.2: First Passage Time from all 6 states to State F in MMSSR

APPENDIX B

Table B.1. *Distributions parameters of the stress X corresponding to state of Nature under each action for the example case*

		Nature	
		Good	**Bad**
Action	**1**	X_{g1}~N(5, 10)	X_{b1}~ N(5, 10)
	2	X_{g2}~N(8.5, 10)	X_{b2}~ N(8.5, 10)
	3	X_{g3}~N(8.5, 5)	X_{b3}~N(8.5, 5)

Table B.2. *Distribution parameters of X after replacement for the example case*

Nature	**Good**	X_{Rg}~N(4, 2)
	Bad	X_{Rb}~ N(4, 2)

Table B.3. *Distribution parameters of Y for the example case*

Nature	**Good**	Y~N(10, 4)
	Bad	Y~N(8, 4)

Table B.4. *Cost of actions for each state for the example case*

		Actions		
		1	**2**	**3**
States	**(G, L)**	10	10	8
	(G, M)	5	0.25	0.3
	(G, H)	10	0.5	0.75
	(B, L)	0.05	10	13
	(B, M)	0.7	5	0.7
	(B, H)	12.5	1.2	1.5
	F	2.5	2.5	2.5

APPENDIX C

Table C.1. *Optimal* $\pi_{(k,O)a}$ *values for MMSSR by minimizing* $\pi_{(B,L2)3} + \pi_{(B,L2)4} + \pi_{(B,L2)7}$

		Actions									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
States	**(G, L1)**							0.031			
	(G, L2)					0.175					
	(G, L3)					0.056	0.189				
	(G, L4)				0.006	0.058					
	(G, L5)				0.001	0.006					
	(G, L6)					>0					
	(B, L1)							0.029			
	(B, L2)					0.132					
	(B, L3)					0.024	0.082				
	(B, L4)				0.001	0.010					
	(B, L5)				>0	>0					
	(B, L6)					>0					
	F	0.198									

Table C.2. *Optimal* $\pi_{(k,O)a}$ *values for MMSSR by minimizing* $\pi_{(G,L4)4} + 3\pi_{(G,L4)5} + 7\pi_{(G,L5)4} + 2\pi_{(G,L5)5}$

		Actions									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
States	**(G, L1)**			0.043							
	(G, L2)							0.175			
	(G, L3)				0.088	0.044	0.049				0.017
	(G, L4)						0.070				
	(G, L5)	0.004					0.010				
	(G, L6)				>0						
	(B, L1)			0.038							
	(B, L2)							0.136			
	(B, L3)				0.038	0.019	0.021				0.007
	(B, L4)						0.013				
	(B, L5)	>0					0.001				
	(B, L6)				>0						
	F	0.229									