SEMI-MARKOV AND DELAY TIME MODELS OF MAINTENANCE

KWANG-MOOK CHOI

T.I.M.E. Research Institute
Department of Computer and Mathematical Science
University of Salford, Salford, UK

Submitted in Partial Fulfilment for the Degree of
Doctor of Philosophy, January 1997

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

Figure | Page

Figure		Page

LIST OF TABLES

Table		Page

ACKNOWLEDGEMENTS

I would like to express my great thanks to my supervisor, Professor A. H. Christer, for his excellent supervision, untiring interest, support, encouragement, and remarkable help for this research.

To the staff in the Centre of Operational Research and Applied Statistics, visitors, and friends, thank you for your very kind and untiring help and assistant over the period of this study. I particularly thank Dr. W. Wang for his continued interest and Professor F. A. Van der Duyn Shouten for his valued input comments early in the study.

Thanks are also due to my sponsor, R. O. K. Army, who supported me during the research period.

Finally, I am greatly indebted to my wife, Kap-Sook, and the children, Seon-Jin and Jae-Woo, for their patience, understanding, and support during the past four years.

ABSTRACT

This thesis is concerned with modelling inspection policies of facilities which gradually deteriorate in time. The context of inspection policies lends itself readily to probabilistic modelling. Indeed, many of the published theoretical models to be found in the literature adopt a Markov approach, where states are usually 'operating', 'operating but fault present', and 'failed'. However, most of these models fail to discuss the 'fit' of the model to data, and virtually no examples of actual applications or case-studies are to be found.

In a series of recent papers dating from 1984, a robust approach to solve these problems has been introduced and developed as the Delay Time Model (DTM). The central concept for this model is the delay time, h , of a fault which is the time lapse from when a fault could first be noticed until the time when its repair can be delayed no longer because of unacceptable consequences. The bottle neck in delay time modelling is how to estimate the delay time distribution parameters. Two methods for estimating these parameters have been developed, namely the subjective method and the objective method.

Markov models have the advantage of an extensive body of theory. There are, however, difficulties of definition, measurement, and calculation when applying Markov models to real-world situations within a maintenance context. Indeed, this problem has motivated the current research which aims to explore the two modelling methodologies in cases where comparison is valid, and also to gain an insight as to how robust Markov inspection models can be as decision-aids where Markovian properties are not strictly satisfied. It will be seen that a class of inspection problems could be solved by a semi-Markov model using the delay time concept. In this thesis, a typical semi-Markov inspection model based upon the delay time concept is presented for a complex repairable system that may fail during the course of its service lifetime and the results are compared. Finally, a case study of the semi-Markov inspection model and the delay time model is discussed.

Chapter 1

INTRODUCTION

This thesis is concerned with modelling inspection policies of facilities such as machines, vehicles, or buildings which gradually deteriorate in time. In general, most items eventually wear out and fail. The time of failure, however, is not known in advance, and so it is both possible and probable that items will fail when in operation. These failures can be quite expensive not only due to repairing or replacing the item, but also because of the disruption and delay involved to the purpose of the operation of the system. Thus such items are often subject to a maintenance policy.

The primary function of maintenance is to control the condition of facilities. It involves actions to be carried out in order to inspect, repair, replace or modify a component or group of components of a system. In the past, maintenance was regarded as a necessary evil and unavoidable activity. This perception of maintenance has slowly changed over the past four decades and now maintenance has in many cases become not only an important part of an organisation, but may also be considered as a profit making activity. In an organisation, an objective of a maintenance function may be to maximize the availability of an operating facility in a safe condition, or to minimize the cost while ensuring a certain level or percentage of facility availability. The fulfilment of this objective, however, would require not only adequate engineering and technological skills but also management skills to effectively plan, organise, direct and control maintenance activities and resources.

Some of the problems associated with maintenance include determination of inspection frequencies. The basic purpose behind an inspection is to determine the state of equipment. Once indicators, such as bearing wear, gauge readings, quality of product, which may be used to describe the state, have been specified, and the inspection made to determine the values of these indicators, then some further maintenance action may be taken, depending on the state. The point when an inspection should take place ought to be influenced by the costs and benefits of the inspection, which will in turn be related to the indicators used to describe the state of the equipment and the benefits of the inspection, such as the detection and correction of minor defects before major breakdown occurs.

At an inspection, defects are presumed identified and repaired so that equipment is restored to a specified condition, often regarded as new. Between inspections, defects may arise and be obvious, in which case appropriate action is taken. Alternatively, a defect may lie dormant for a period until it either matures in severity to become obvious, or is identified at an inspection. Sometimes, and for equipment such as a computer or piece of software, very specific checks are required to determine if equipment is working, whilst at other times equipment can only be checked by destruction. Safety and defence equipment are typical examples here. In such a case, statements as to the likely availability of equipment in an operating state is based upon statistical evidence derived from trials and samples, and is well documented in the literature on reliability.

In some of the early work, Barlow and Proschan [1965] devised optimal inspection schedules which are subject to two-states, namely a good state and a failed state. Though simple in form, the schedules proved less than simple to compute, and a number of authors presented approximations to the optimal inspection schedules which were easier to calculate (see Munford and Shahani [1972] and Nakagawa and Yassui [1979]). In recent times, as pointed out in Christer [1984], many of the published theoretical models of industrial inspection problems adopt a Markov approach where the states are "operating", "operating but fault present", and "failed". Each state is associated with a cost or downtime in the broad sense of the term with transitions between them occurring according to probabilistic laws, the occurrence of inspections associated with maintenance actions, and repair upon failure.

Though the models provide ideas for possible model-building blocks along with some qualitative insight as to how a system might behave if only it would oblige the model's assumptions, a major interest is in the solution procedures. Much insight has been gained in the task of investigating and solving various types of theoretical models, but the task of building an inspection model for an identified plant appears to be relatively unexplored. Also, many papers assume that the working condition of the system can be expressed as a discrete-time Markov chain with a new state, degraded states, and a failed state, and that the state transition probabilities can be determined. It is, however, difficult in general to define the degraded states for a deteriorating system. Even when it is possible to sensibly define the states of the system, it is often more difficult to determine the state transition probabilities. We presume that this is the reason that so few papers presenting Markov models make any mention of the 'fit' of the model to data, on present examples of actual applications or case-studies.

In a series of recent papers dating from 1984, what has proved to be a robust approach to solve these problem has been introduced and developed as the delay time concept and model. In 1982, Christer exploits the ideas of a "delay time" for a fault, which arose originally as a side issue in modelling building maintenance. Fundamental to most engineers' experience is the idea that defects do not just appear as failures, but are present for a while before becoming sufficiently obvious to be noticed and declared as failures. The time lapse from when a defect could first be reasonably expected to be identified at an inspection to the consequential failure repair if no corrective action is taken has been termed the delay time h of the fault. The bottle neck in delay time modelling is how to estimate the delay time distribution parameters. Two main methods for estimating these parameters have been developed, namely the subjective method and the objective method.

Markov models have the advantage of an extensive body of theory. There are, however, difficulties of definition, measurement, and calculation when applying Markov models to real-world situations within a maintenance context. Indeed, this problem has motivated the current research which aims to explore the two modelling methodologies in cases where comparison is valid, and also to gain an insight as to how robust Markov inspection models can be as decision-aids where Markovian properties are not strictly satisfied (or perhaps testable). It will be seen that a class of inspection problems could be solved by a semi-Markov model using the delay time concept. If we can define the degraded states of a system as the number of existing defects, we can easily define the working condition of the system as a Markov chain. Also, if we know the probability density function of delay time h, and the statistics of the defect arrival process, it is seen to be possible to determine from the probability density function of the delay time the state transition probabilities of the associated Markov inspection model. Developing these ideas, a typical semi-Markov inspection model based upon the delay time concept is presented for a complex repairable system that may fail during the course of its service lifetime. This model is contrasted with the delay time model for the same problem.

In this thesis, we are ultimately concerned with the problem of modelling to inform the task of deciding the inspection policy of equipment where the costs or the downtimes are taken into account. Setting an inspection policy includes the determination of the inspection frequencies for complex equipment. In order to study and solve the inspection problem for a complex system, first of all, we review the mathematical Markov literature in chapter 2, and develop the delay time model from the concept in chapter 3. In chapter 4, for a common inspection scenario, both a simple semi-Markov type inspection model based on the delay time concept, and a delay time model for a single component system are developed and presented, and the results are compared. Again, for a common inspection problem scenario for a complex multi-component system, in chapter 5, a semi-Markov inspection model based upon the delay time concept is developed and a comparable delay time model is constructed. The results are compared and the potential accuracy and error in using Markov models for non-Markovian problems discussed. Finally, in chapter 6, a case study of the semi-Markov inspection model and the delay time model is discussed. We believe this is the first instance of a semi-Markov inspection model being used to model a real problem in maintenance. Conclusions are presented in chapter 7.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

Over the last few decades, numerous papers have appeared in the literature which deal with the problem of finding optimal inspection policies for systems which are subject to failures. This phenomenon is indicated in various surveys of maintenance models by Pierskalla and Voelker [1976]. Sherif and Smith [1981], Christer [1984], Thomas [1986], Valdez-flores and Feldman [1989], Cho and Parlar [1991], Thomas, Gaver and Jacobs [1991], and White [1993]. The models address various aspects of inspection problems in such systems as an industrial production plant, a vehicle fleet, a housing estate, or a motor way system. The complexity of the models varies from a very simple deterministic model of a single-unit system to a very complex model of a stochastically failing multi-unit system. In general. the basic type of decision problem involved in an inspection system concentrates on determining the inspection schedules which minimises the cost per unit time or the downtime per unit time.

The general inspection model of a facility or a component may be classified into the model with two-states and the model with multi-states. Many items or systems can be described as being in one of two states, one of which is preferable to the other. This preferred state can be described as working. whilst the other might represent some form of failure. Some examples of this type are shelf life of goods, health of a human being, life testing of components, and standby systems. It is convenient to label the working state as state 0, and to define the unsatisfactory or failed state as state 1. It is assumed that a transition from state 1 to state 0, i.e. failed to working, cannot occur while the system is in service.

The inspection model with a multi-state capability normally considers an equipment that gradually deteriorates in time and whose degree of deterioration can be observed by inspections only, except for a failed state that is observed immediately upon its occurrence. Practical examples include production machines subject to stochastic breakdowns, inventory systems being depleted, and maintenance of communication systems with redundancy. An inspection is assumed to reveal the exact working condition of the system. Depending on the system's degree of deterioration, an inspection may be succeeded by a replacement or restoration. The system can be observed in one of the working conditions $0, 1, \ldots, n, f$ which describe increasing degrees of deterioration. The state 0 represents a new system, the states $1, 2, \ldots, n$ represent the degraded states, and the state f represents a failed state or a severe malfunction. It is possible to transfer from any state to a failed state f.

In this chapter, a review of the relevant literature on inspection models will be presented.

2.2 Two-State Inspection Model

2.2.1. Basic Inspection Model based upon Barlow and Proschan's Assumptions

A basic inspection model for a complex system was given by Barlow and Proschan [1965]. They considered the simplest possible case of an inspection policy, which was characterised by the following assumptions.

(1) Deterioration of a system is stochastic.

(2) The working conditions of the system are classified into states 0 and 1. State 0 denotes a good state and state 1 denotes a failed state.

(3) The condition of the system is known by inspection.

(4) An inspection takes negligible time.

(5) An inspection does not degrade the system and the system cannot fail while being inspected.

(6) Inspections are perfect in that any failure within the system will be identified.

(7) Each inspection entails a fixed cost C_1.

(8) The time elapsed between system failure and its discovery at the next inspection costs C_2 per unit of time.

(9) Repair takes place upon discovery of failure and the system is as good as new after repair.

(10) The failure distribution function $F(t)$ of the system is known.

Assumptions (1) and (2) state that the system gradually deteriorate in time. Also, assumptions (3), (4), (5), (6), and (7) are related to the general ideal inspection policy. Assumption (8) implies a system failure remain unknown until an inspection. This rules out most applications in industry. By assumption (6) and (9), an inspection will renew the system.

Under the above assumptions, they derived the expected cost up to detection of failure as

$$C = \sum_{k=0}^{\infty} \int_{x_k}^{x_{k+1}} [C_1(k+1) + C_2(x_{k+1} - t)]dF(t), \tag{2.1}$$

where $x_0 = 0 < x_1 < x_2 < \ldots$ are the successive inspection times. Assuming a density function $f(t)$ of the failure time distribution $F(t)$, a necessary condition that a sequence $\{x_k\}$ be a minimum cost inspection procedure is that $\frac{\partial C}{\partial x_k} = 0$ for all k. Hence using equation (2.1). the following equation is obtained for $k = 1, 2, 3, \ldots$.

$$x_{k+1} - x_k = \frac{F(x_k) - F(x_{k-1})}{f(x_k)} - \frac{C_1}{C_2}. \tag{2.2}$$

When $f(x_k) = 0$, $x_{k+1} - x_k = \infty$ so that no more checks are scheduled. The sequence is determined recursively once x_1 is chosen. Unfortunately, it is difficult to compute optimal inspection procedures numerically, because the computations are repeated until the procedures are determined to the required degree by changing the first check time.

To avoid this, Munford and Shahani [1972] suggested a near optimal inspection policy which depended on a single parameter p. They introduce the probability of a transition from state 0 to state 1 during the interval (x_{i-1}, x_i) given that the system was in state 0 at time x_{i-1}, which is given by

$$\frac{F(x_i) - F(x_{i-1})}{1 - F(x_{i-1})} = p, \qquad \text{for } i = 1, 2, 3, \ldots, \tag{2.3}$$

where $0 < p < 1$, $x_0 = 0$ and $F(0) = 0$. The above equation (2.3) may be rewritten as

$$F(x_i) = p + (1-p)F(x_{i-1}) \tag{2.4}$$

and

$$\begin{aligned} F(x_i) &= 1-(1-p)^i \\ &= 1-q^i, \end{aligned} \tag{2.5}$$

where $q = 1-p$. Thus for a given p, x_i can be found from

$$x_i = F^{-1}(1-q^i). \tag{2.6}$$

To choose an optimal p, let a random variable I denote the number of inspections necessary for the detection of state 1. We have

$$\Pr(I = i) = q^{i-1}p, \qquad \text{for } i = 1, 2, 3, \ldots, \tag{2.7}$$

so that

$$E(I) = \sum_{i=1}^{\infty} i q^{i-1} p = \frac{1}{p} . \tag{2.8}$$

If the transition occurs at time t and it is detected by an inspection at time x_i, then $(x_i - t)$ is the time for which the system was left in service in state 1. The mean time for which the system will be left in service in state 1 is

$$\tau = \sum_{i=1}^{\infty} \int_{x_{i-1}}^{x_i} (x_i - t) f(t) dt$$

$$= \sum_{i=1}^{\infty} x_i q^{i-1} p - E(T), \tag{2.9}$$

where

$$E(T) = \int_0^{\infty} t f(t) dt . \tag{2.10}$$

So the total expected cost until a failure is detected is given by

$$E(C) = C_1/p + C_2 \left(\sum_{i=1}^{\infty} F^{-1}(1-(1-p)^i)(1-p)^{i-1} p - E(T)\right) . \tag{2.11}$$

The optimal p can be chosen such that $E(C)$ is minimized. This policy was used for Weibull failure distribution case in Munford and Shahani [1973].

Also, Nakagawa and Yassui [1979] considered an inspection policy with periodic checking times. If it is assumed that the mean duration of undetected failure is approximately half the time between consecutive checking times, the optimum checking time p^*, which minimizes the total expected cost until a failed unit is discovered by some checking, is

$$p^* = \sqrt{2m\tau} . \tag{2.12}$$

where m is the mean of failure times of the unit and $\tau = C_1/C_2$.

Further, Nakagawa and Yassui [1980] gave an approximate calculation of optimal checking procedures which computed successive check times backward. They specified the following computing procedure for obtaining the asymptotically optimal inspection schedule :

(1) Choose an appropriate ε from among $0 < \varepsilon < C_1/C_2$.

(2) Determine a check time x_n after sufficient time has elapsed to give the required degree of accuracy.

(3) Compute x_{n-1} to satisfy

$$x_n - x_{n-1} - \varepsilon = \frac{F(x_n) - F(x_{n-1})}{f(x_n)} - \frac{C_1}{C_2} . \tag{2.13}$$

(4) Compute $x_{n-1} > x_{n-2} > \ldots$ recursively from the equation (2.2).

(5) Continue until $x_k < 0$ or $x_{k+1} - x_k > x_k$.

Luss [1983] suggested an inspection policy model for production facilities using a dynamic programming algorithm. In addition to maximizing the expected profit per cycle, he examined the problem of maximizing the expected profit per unit time and the expected profit per "good" unit time. He described a dynamic programming algorithm that found the inspection policy that maximized the expected profit per cycle. This algorithm is then imbedded within a Newton-Raphson type search that finds the optimal policies for the other objective functions.

Also, Assaf and Shanthikumar [1987] discussed an optimal group maintenance policy for a set of N machines subject to stochastic failures under continuous and periodic inspections which minimized the expected cost per unit time over an infinite horizon.

Dias [1990] analysed a new approximation for the inspection period when the failure rate is ; (a) decreasing ; (b) increasing ; (c) first increasing and then decreasing : and, finally, (d) when the failure rate has a 'bathtub' shape. From Nakagawa and Yassui's [1979] approximation p^*,

$$p^* = \sqrt{2rE(T)}, \tag{2.14}$$

where $r = C_1/C_2$, and $E(T)$ is the expected value of system of lifetime T, he derived the new approximation p^{**},

$$p^{**} = \frac{p^*}{1+0.234\sqrt{r'}}, \tag{2.14}$$

where $r' = r/E(T)$. He confirmed that p^{**} was better than p^* for the decreasing, the bathtub-shaped, and, finally, the increasing and then decreasing failure rates. Only for the case when the failure rate is increasing, he has obtained worse results. However, this is a case in which non-periodic inspections are preferable.

2.2.2. Modified Inspection Model

So far, we have cited only the literature which adopts the assumptions of Barlow and Proschan's model. However, they require restrictive assumptions because they assume an ideal state. To avoid this, the model with changed lifetime distribution is, firstly, presented by Anbar [1976], and Beichelt [1981]. Anbar considered an adaptive sequential inspection policy assuming that the lifetime distribution is known to be exponential, but with unknown parameter. He suggested the procedure for estimating the expected lifetime. This procedure yields a sequence of estimates which is strongly consistent, i.e., converges with probability 1, to the value of the unknown parameter. This in turn implies that the sequence of intervals between inspections converges to the optimal interval between inspections. Beichelt derived minimax inspection strategies for single unit systems on condition that no or only partial information on the lifetime distribution of the system is available.

Luss and Kander [1974] discussed inspection policies when the duration of checking and repair is non-negligible. Also, they assumed that the system continued operation during its inspection and could fail while being checked. The loss functions are obtained and are solved by both differentiation, which leads to efficient algorithms for IFR(Increasing Failure Rate) distributions, and by dynamic programming, which can be used for any failure rate.

Also, Jardine and Hassounah [1990] demonstrated how the relation between the mean arrival rate of breakdowns conditional upon the inspection frequency of equipment could be estimated in practice assuming that inspection times and repair times were negative exponentially distributed. This work was carried out for a large, urban transit authority operating a fleet of approximately 2000 buses undertaking about 80 million kilometres per year. A model relating total downtime of buses incurred due to inspections and repairs per unit time to inspection was developed, and the optimal inspection frequency which maximized bus availability was determined.

Further, Nakagawa [1984] considered a modified inspection policy with periodic check intervals, where the unit after check has the same age as before with probability p and is as good as new with probability q. The mean time to failure and the expected number of checks before failure are derived, forming renewal-type equations. The total expected cost and the expected cost per unit of time until detection of failure are obtained. Optimum inspection policies which minimize the expected costs are given.

Kaio and Osaki [1986] discussed a typical inspection model taking account of the following two imperfect inspection probabilities :

(a) the system might be regarded as having failed, even if it is normally operating, due to imperfect inspection,

(b) an inspection may not detect a system failure due to imperfect inspection.

For this model, they obtained the structure of the optimal policy, and discussed the optimal policy which minimized the total expected cost up to the detection of system failure.

Finally, Munford [1990] considered optimal inspection policies in which penalty costs due to the elapsed time between the failure and its detection were proportional to the duration of the inspection interval containing the failure. An algorithm for computing the optimal inspection policy is given for a wide class of failure distributions.

Two-state inspection models have been established by Barlow and Proschan [1965] by assuming the penalty costs due to the elapsed time between the failure and its detection. This assumption rules out most industrial applications. Also, the two-state inspection models do not relate to the practical industrial situation in which inspection leads to repair before failure because the models define the working condition of the system as two states, namely a good state and a failed state. So, other authors have established multi-state inspection models to be mentioned in the following section.

2.3 Multi-State Inspection Model

2.3.1. Basic Inspection Model

We have considered the papers which deal with the model to be classified into two states 0 and 1. Now we investigate the literature in which the condition of the system can be classified into multi-state $0, 1, ..., n, n+1$, being characterised typically by the following assumptions.

(1) The system gradually deteriorates in time.

(2) The working condition of the system at any time t can be completely characterised by classifying it as in one of the states $0, 1, 2, ..., n, n+1$ where 0 is the good state, $1, 2, ..., n$ are degraded states and $n+1$ is the failed state.

(3) The degree of deterioration can be observed by inspections only, except for a failed state that is observed immediately at its occurrence.

(4) An inspection takes negligible time.

(5) An inspection does not degrade the system and the system cannot fail while being inspected.

(6) An inspection reveals the exact working condition of the system.

(7) Each inspection requires a fixed cost of C_1.

(8) After each repair the system is considered to have working condition 0.

(9) If the system has working condition i at present, then one time unit later it will have working condition j with known probability r_{ij}, where r_{ij} depends only upon the current state i and the next state j.

Assumptions (1) and (2) relate to the working condition of the system which gradually deteriorate in time. Assumption (3) is related to the detection of the working condition of the system. This assumption is, particularly, different from the assumptions of Barlow and Proschan's in recognizing a failure. This implies that this model does not need the introduction of a penalty cost due to the elapsed time between the failure and its detection. Similarly with the assumptions of Barlow and Proschan, assumptions (4) to (7) are related to the general ideal inspection policy. Assumption (8) implies that a repair will renew the system. By assumption (9), it is easy to establish a decision criterion model, but it is difficult to apply to a practical industrial situation.

Under the above assumptions, Mine and Kawai [1975] discussed an inspection and replacement policy which minimized the expected total long-run average cost using a semi-Markov decision process. They, firstly, considered the following Markov degradation properties.

(a) The transition rates from one state to another are independent of time, i.e., are constant.

(b) From state i, a random transition is only possible to state $i+1$ or $n+1$.

(c) The transition rate from i to $n+1$ increases as i increases.

Under the above Markov degradation properties, they derived the transition probability $P_{ij}(t)$ which is the probability that the system was in state j at time t given the system was in state i at time 0. This transition probability $P_{ij}(t)$ is based on the transition (failure) rate from state i to the failed state $j = n+1$, α_i, and the transition (degradation) rate from state i to state $j = i+1$, β_i, with known α_i and β_i respectively. Also, defining E_i as the event that the system was in state i and an inspection had just been performed, they showed that E_0 , E_1 ,..., E_n constituted a semi-Markov process and that the process had a single imbedded Markov chain which was ergodic for every stationary policy. Then, by using the theory of semi-Markov decision processes, they found the optimal policy iteration cycle without predetermining the inspection time interval.

Further, Kander [1978] presented inspection policies for deteriorating equipment characterised by N quality levels. He assumed that the mechanism of deterioration consisted of successive Poisson transitions of the system from the prevailing state to the consecutive state and that the Poisson transition parameters were known. Considering three feasible inspection models which are pure checking, truncated checking and checking followed by monitoring, he developed optimal policies leading to minimal loss, while the system's distribution was represented by an (N+1)-state semi-Markov process.

Also, Ohnishi, Kawai and Mine [1986a] treated a continuous time Markovian deterioration system when the operating costs and the replacement costs are dependent on its state, and derived an optimal inspection and replacement policy minimizing the expected total long-run average cost. The previous literature had generally assumed that the system had a constant operating cost and a constant preventive replacement cost which was independent of the state of the system. However, the authors introduced the idea that the operating costs and replacement costs increased as the deterioration of the system increased. This increased the general applicability of this modelling, and under these assumptions. they derived an optimal inspection and replacement policy and showed that an optimal policy had monotonic properties under some reasonable conditions. Under the optimal policy, a control limit rule holds for the replacement decision. and the optimal time interval between successive inspections becomes shorter and shorter as the system undergoes deterioration.

2.3.2. A Variation on the Inspection Procedure

So far we assumed that an inspection does not adversely affect the system. We may. however. have some system which is impaired by the inspection. The system operates throughout a number of periods and is subject to failure in each period. Prior to failure the system enters a state in which it is functioning, but in a possibly impaired manner. This state can be detected only by performing an inspection, by assumption (3). Once the system is known to be in the impaired state, appropriate action may be taken to prolong its remaining life. However, the act of inspecting the system when it is not impaired may itself cause it to become impaired. In this respect inspection may be hazardous. A prime example is the inspection of nuclear reactors. Since one of the largest causes of malfunction in reactors is human error, a fundamental question is whether or not inspections by humans create more problems than they solve.

In this respect. Bulter [1979] considered a hazardous inspection model which maximized the lifetime of the system when inspections had the potential of being harmful to the system under consideration. He classified a system into one of four states which are fully functional, undetected partial failure, detected partial failure and failed. and then formulated the inspection model as a Markov decision process. Also, Chou and Butler [1983] developed an efficient computational procedures for the hazardous inspection model.

2.3.3. Inspection Model with a Catastrophic Failure Situation

Some inspection models are considered to have a renewal point after repair. In these models, the objective is usually to seek the optimal inspection schedule for a system which minimizes the total cost per renewal cycle. In some cases, however, the cost of failure is regarded as so great that it cannot be meaningfully compared to the costs of inspections and corrective actions. For example, a person develops cancer, or the failure of a crucial component of an aeroplane in flight or within a nuclear power plant.

In this respect, Milioni and Pliska [1988] sought the optimal inspection schedule for a system whose deterioration process is a semi-Markov process that progresses toward failure when the failure is catastrophic. They actually analysed two versions of this catastrophic situation. In the first version of the catastrophic failure situation, they addressed the problem of minimizing the expected maintenance costs subject to the constraint that the probability of failure be no greater than a specified value. This problem is solved using dynamic programming and Lagrange multipliers. In the second version of the catastrophic failure situation, the problem is to minimize the probability of failure subject to the constraint that the number of inspections cannot exceed a specified number. This version of the problem is formulated and solved using dynamic programming.

2.3.4. Inspection Model with Non-Negligible Inspection Time

An inspection model with non-negligible inspection time was given by Tijms and Van Der Duyn Schouten [1985]. They considered an equipment which became increasingly expensive to operate with an increasing degree of deterioration, and the following details of the problem were considered. Opportunities for inspections occur only at discrete points in time $t = 0, 1, \ldots$ and an inspection takes a fixed, integral number of T time units and costs J units. Once an inspection has revealed the exact working condition, there is an option of either doing a revision or leaving the system as it is. A subsequent revision in working condition i takes a fixed integral number of T_i

time units and involves a cost of $R_i \geq 0$, $0 < i \leq n+1$, and the system incurs an operating cost of a_i for each time unit it spends operating in working condition i, $0 \leq i \leq n$. The system must always be revised when working condition $n+1$ is found. In the absence of inspections or revisions, the working condition of the system changes according to a discrete-time Markov chain with known probability r_{ij}. It is assumed that $\sum_{j=i}^{n-1} r_{ij} = 1$ for all $0 \leq i \leq n+1$, that is, the working condition of the system cannot improve on its own. A control-limit rule is characterised by integers $\pi_0, \pi_1, \ldots, \pi_n$ and prescribes the following actions. Revision is done when the working condition $k \geq \pi_0$, and inspection is done when π_i time units have passed since the system was last known to have working condition i. Also, state space is taken as

$$I = \{i | i = 0,1,\ldots,n,n+1\} \cup \{(i,m) | i = 0,\ldots,n, m = 1,\ldots,M\},$$

where a state (i, m) corresponds to the situation of m time units having passed since the last knowledge of the system's working condition i. The possible actions are denoted by

$$a = \begin{cases} 0, & \text{leave the system as it is,} \\ 1, & \text{inspect the system,} \\ 2, & \text{revise the system.} \end{cases}$$

As to the transitions resulting from taking action a in state (i, m), the following definitions are given.

$C_{(i,m)}(a)$: The expected transition costs, which are made up of inspection, revision and operating costs.

$\tau_{(i,m)}(a)$: The expected transition time if action $a = 1$ or $a = 2$ is taken, which are made up of inspection and revision times.

$\tau_{(i,m)}(0) = 1$: The expected transition time if action $a = 0$ is taken.

$P_{s,s'}(a)$: The probability of a one-step transition from state s to state s' with $s, s' \in I$.

$q_{ij}^{(m)}$: The m-step transition probabilities of the Markov chain for m = 1, 2, ..., where $q_{ij}^{(m)}$ is the probability that m time units from now the system will have working condition j when the present working condition is i and no inspections and revisions are made.

For the semi-Markov decision model, denoting the relative operating costs $v_s(R)$, $s \in S$ associated with rule R and the long-run average costs by $g(R)$, the following set of linear equations are formulated.

$$v_i(R) = a_i - g(R) + (1 - q_{i,n+1})v_{(i,1)}(R) + q_{i,n+1}v_{n+1}(R), \qquad \text{for } 0 \le i < \pi_0, \qquad (2.16)$$

$$v_i(R) = R_i - g(R)T_i + v_1(R), \qquad \text{for } \pi_0 \le i \le n+1, \qquad (2.17)$$

$$v_{(i,m)}(R) = \sum_{j=i}^{n} \frac{q_{ij}^{(m)}}{1 - q_{i,n+1}^{(m)}} a_j - g(R) + \frac{1 - q_{i,n+1}^{(m+1)}}{1 - q_{i,n+1}^{(m)}} v_{(i,m+1)}(R) + (1 - \frac{1 - q_{i,n+1}^{(m+1)}}{1 - q_{i,n+1}^{(m)}})v_{n+1}(R),$$
$$\text{for } 0 \le m < \pi_i,\ 0 \le i < n, \qquad (2.18)$$

$$v_{(i,m)}(R) = J - g(R)T + \sum_{j=i}^{n} \frac{q_{ij}^{(m)}}{1 - q_{i,n+1}^{(m)}} v_j(R), \qquad \text{for } \pi_i \le m \le M_i,\ 1 \le i < n, \qquad (2.19)$$

augmented by putting one of the relative operating costs equal to 0, say

$$v_{n+1}(R) = 0.$$

From the above equations, the authors derived the relative operating costs by single-pass calculations and presented a special-purpose algorithm to compute the best rule within the class of control-limit rules. This algorithm generates a sequence of improving control-limit rules and it can be shown by familiar arguments from Markov decision theory that the algorithm converges after finitely many iterations.

However, Wijnmalen and Hontelez [1992] demonstrated that this algorithm, which operated on a class of control-limit rules, did not always lead to an optimal policy even within the class of control-limit rules considered. They pointed out Tijms and Van Der Duyn Schouten's algorithm was too restrictive in the sense that it adhered too much to the control-limit structure while applying the usual improvement procedure, and thus could exclude better control-limit policies. So, if it had not have been restricted to control-limit policies, the usual policy iteration procedure could have been applied without difficulty. In this respect they derived their improved algorithm which could compute the optimal control-limit rule, provided that long-run average costs were minimized.

2.3.5. Imperfect Inspection Model

The models treated so far assumed one of the following two extreme assumptions which denote a perfect inspection model.

(a) At any given time, the state of the system may be identified completely by inspection.
(b) The state of the system can be observed only through costly inspections.

On the other hand, in practical situations, many systems satisfy neither of the above two restrictive assumptions, but some intermediate characterisation of system information. That is, the decision maker obtains some information about the state of the system at each inspection time, but he needs further and costly in depth inspection to identify the exact state of the system with certainty. Under this concept, the imperfect inspection model was presented by Ohnishi, Kawai and Mine [1986b], Devooght, Dubus and Smidts [1990] and Özekici and Pliska [1991].

Ohnishi, Kawai and Mine investigated an optimal inspection and replacement problem for a discrete-time Markovian deterioration system. It was assumed that the system was monitored incompletely by a certain mechanism which gave the decision maker some information about the exact state of the system. It was noted that this information involved uncertainty and was stochastically related to the exact state of the system. The decision maker must pay an additional inspection cost to identify the exact state of the system with certainty. They assumed that the state of the system underwent deterioration according to a stationary discrete-time Markov chain having a known transition probability P_{ij} which denoted the one-step transition probability from state i to state j. They then formulated this model as a partially observable Markov decision process and showed an optimal inspection and replacement policy had monotonic structural properties. However, in this model, it was assumed that the transition probability of the deteriorating process of the system and the probabilistic relation between the system and the monitoring mechanism were completely known. This assumption does not always hold in real-world situations.

Devooght, Dubus and Smidts [1990] developed suboptimal inspection policies for imperfectly observed realistic systems. They considered that the states of the system evolve according to a continuous semi-Markovian model because repair and maintenance operations cannot be realistically described by exponential holding times, and the large number of states is a compelling reason to use supercomponents, which have subsystems, whose failure rate is a combination of exponentially failing components and, therefore, not Markovian. Also, they assumed that

(a) a large penalty was attributed to the unreliability of the whole system,
(b) states could not be ordered and arbitrary transitions were allowed.
(c) knowledge of the system state is imperfect either because of human error or because sensors give only information on the overall behaviour of subsystems and not of its detailed components,
(d) maintenance and inspection by operators are subject to human error and described by matrices which relate intended actions with actual actions, and
(e) no stationary policy is sought.

In their model, the emphasis is put on a production process with safety-related subsystems, such as in nuclear reactors, whose non-availability provokes the stopping of energy production, and therefore has a high cost associated with non-availability. Therefore the safety system is periodically inspected and eventually repaired either at fixed periods or when a critical state is entered. Under this concept, they obtained sub-optimal inspection policies using a dynamic programming algorithm based on the use of importance parameters for the components. A Markovian evolution code is combined with an optimization code using a value iteration algorithm.

Also, Özekici and Pliska developed optimal inspection schedules using a delayed Markov model with false positives and negatives under imperfect inspection. They considered a system subject to catastrophic failure which deteriorated according to a delayed Markov process and was subjected to a series of binary tests that might yield false negative and false positive outcomes. A corrective action was carried out when a true positive was observed, thereby reducing the chance of system failure. Costs of inspections, false positives, the corrective action and failure are incurred and dynamic programming is used to compute the optimal inspection schedule. The importance of this model is the fact that it is both computationally tractable and useful for several kinds of applications, especially medical screening.

Another imperfect inspection model was presented by Christer and Waller [1984a,b,c]. They suggested the optimal inspection policy minimizing the expected cost per unit time of the maintaining the plant and the expected downtime per unit time using the delay time concept. Unlike previous models, in this case the model was developed for and applied to an actual case situation within industry. This model is of a very different format to the previous ones and will be presented in detail in the next chapter along with numerous variation and developments.

2.4 Summary of the Literature Review

We have reviewed numerous models in the literature which deal with the problem of finding optimal inspection policies for systems which are subject to failures. Basic assumptions for the two-state model were given by Barlow and Proschan [1965]. Under the basic assumptions, some authors developed methods which could compute Barlow and Proschan's model easily and the other authors improved the model and changed the basic assumptions. Most of these models present the working condition of the system as being in one of two states, operating and failed, the purpose of inspection is to detect failure, and the model is based on the time to failure distribution. Importantly, the models do not relate to the practical situation in which inspection leads to repair before failure. In this respect, the introduction of Markov inspection models is viewed as a move towards reality.

The Markov inspection models are developed by numerous authors. However, most of them assume that the working condition of the system can be expressed as a discrete-time Markov chain with state, 0, 1, 2, ..., n, n+1, where the state 0 represents a good state, 1, 2, ... n are degraded states and the state n+1 is the failed state, and their transition probability is given. In practice, it is , however, difficult to define the degraded states for the deteriorated system and more difficult to qualify the transition probabilities. Assuming the unknown transition probabilities, Devooght, Dubus and Smidts [1990] tried to develop the inspection policies for imperfectly observed realistic systems, but only managed to obtain sub-optimal inspection policies.

In the sense of developing a useable model with estimable parameters, the delay time concept and model, which is described in the next chapter, represents further improvement in the modelling of inspection policies.

Chapter 3

THE THEORY OF DELAY TIME ANALYSIS

3.1 Introduction

Consider a piece of equipment subject to periodic inspection. At an inspection, faults are presumed identified and repaired so that equipment is restored to a specified condition, often regarded as new. While the equipment is being operated, faults or failures may arise and high costs will be incurred eventually due to failures. Here, a system with a failure or breakdown means a system which cannot be operated, and repair is essential, whilst a system with faults or defects means a system which needs repair but can still be operated. It is appropriate, therefore, to consider ways to identify faults at an earlier time or stage, such as inspection, though these too will incur a cost. The objective here is to devise an inspection schedule so as to strike an appropriate cost balance between the cost of inspecting and the cost of additional or more serious failures which arise through not inspecting.

To this end, there have been recently a considerable number of papers published on inspection modelling utilising and developing the application of a concept which has been of research interest for the past 14 years and known as delay time modelling. A technique called delay time analysis has been initially developed for modelling inspection policies for industrial inspection maintenance when the equipment is regularly inspected. The idea of the delay time first appeared with the context of building maintenance addressed by Christer [1982] (there called lapse time) but has since been named delay time and extended to industrial equipment.

This chapter presents a review of the delay time concept, and delay time analysis. Several models are reviewed to indicate how delay time analysis has so far been employed to model various inspection maintenance problems.

3.2 Basic Delay Time Models

In the delay time model presented by Christer and Waller [1984a], a central concept is the delay time h of a fault, which is the time lapse from when a fault could first be noticed until the time that its repair can be delayed no longer because of unacceptable consequences (see Figure 3.1).

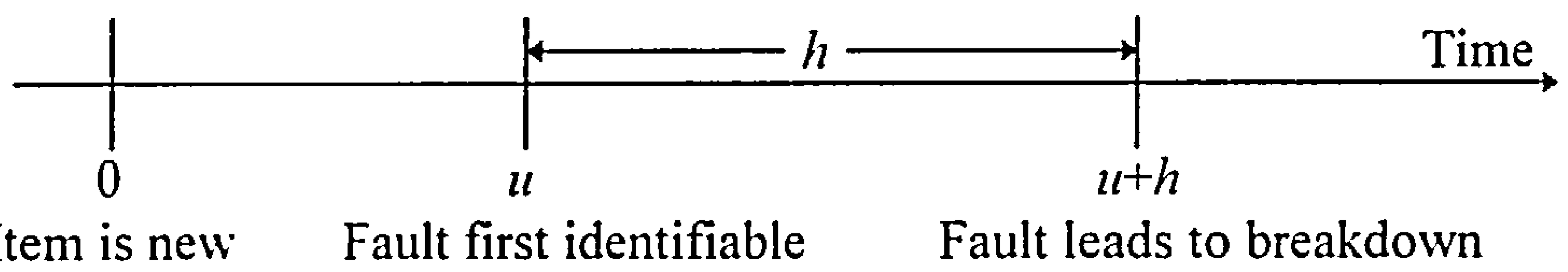


Figure 3.1. Delay time concept.

The important point about the delay time concept is that the failure process is divided into two-stages. This enables the theoretical foundation for inspection policies to be established. The difficulty in delay time modelling is, however, how to estimate the distributions of the delay time h and the initial point u, the instant at which a fault may be assumed to first arise since new or reconditioned. Two main methods for estimating these distributions have been developed, namely subjective method and the objective method. We will describe these methods in section 3.4.

If the distributions of delay time and the initial point are known, the failure behaviour of equipment can in theory be determined under any specified maintenance policy.

Consider first the simplest possible case of an inspection policy, which may be characterised by the following assumptions.

(a) An inspection takes place every T time units, costs C_i units and requires d_i time units, where $d_i << T$.

(b) Inspections are perfect in that any fault present within the system will be identified and no new fault will arise because of the inspection.

(c) Faults identified at an inspection will be repaired within the inspection period at an average cost of C_d.

(d) The initial point of a fault at time u after an inspection is independent of the delay time h.

(e) Faults arise at a constant rate of λ per unit time.

(f) Failures are repaired as soon as they arise, incurring on average d_b units of downtime, and cost C_b, where $C_b > C_d$ and $d_b << T$. Here, even though a component of the system fails, if the system can still be operated, we do not regard that the system is in condition of a 'failure' or 'breakdown'.

(g) A component of the system is as good as new after repair.

(h) The probability density function of the delay time, $f(h)$, is known.

The assumption $d_i << T$ in (a) and the assumption (c) may at first seem to be contradictory if several defects are identified. However, assumption (c) would seem to be reasonable if sufficient maintenance staff were available to perform repairs simultaneously. The assumption (e) provides an estimate of the expected number of faults arising in the period T, namely $K(T)$. This ignores the downtime due to breakdowns, during which no faults would arise since the machinery is idle. If this downtime is small compared with T, as indicated in the assumption (f), $d_b << T$, then the error will also be small. Later we consider the modelling changes necessary to relax this condition.

Under these assumptions, firstly, we determine the form of the function $b(T)$ which is the probability that a fault ultimately arises as a breakdown. Suppose that a fault arising within the period $(0, T)$ has a delay time in the interval $(h, h+dh)$. The probability that the delay time lies in this interval is $f(h)dh$. This fault will be repaired as a breakdown repair if the fault arises in period $(0, T\text{-}h)$ (see Figure 3.2), otherwise as an inspection repair.

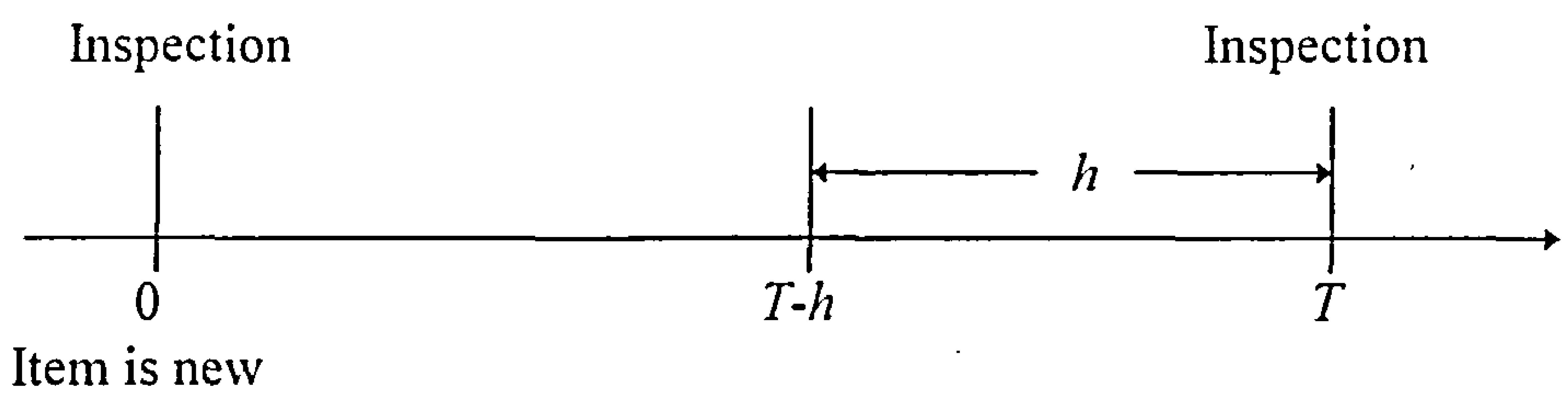


Figure 3.2. Repair process of a fault.

The probability of a fault arising before $T\text{-}h$, given that a fault will arise in $(0, T)$, is, from assumption (e), $(T\text{-}h)/T$. We have, therefore, that for small dh, the probability that a fault with delay time in the interval $(h, h+dh)$ arises as a breakdown is

$$\frac{T-h}{T} f(h)dh \ .$$

The delay time h lies in the period $(0, T)$ because the fault referred to this delay time h will be identified at the inspection time T by the assumption (b) or will arise as a breakdown in case of $h \leq T$. Allowing h to vary from 0 to T and integrating the above term over h, we have the probability of a fault arising as a breakdown, $b(T)$, is given by

$$b(T) = \int_0^T \frac{T-h}{T} f(h)dh \ . \qquad (3.1)$$

Since a components with no fault may be regarded as 'new' by the assumption (e) and all components with fault are identified and renewed after repair by the assumptions (b), (c), and (g) at an inspection point, each inspection point will become in effect a renewal or re-condition point. Accordingly, the equation (3.1) applies to each inspection period.

Using the equation (3.1), with average breakdown and inspection repair costs C_b and C_d respectively, a model of the expected cost per unit time as a function of the inspection period T may be obtained directly. The total expected cost of an inspection cycle consists of the expected cost due to failures, the expected cost of rectifying faults identified at the inspection, and the cost of the inspection itself. If the expected number of faults and breakdowns arising over $(0, T)$ is $K(T)$ and $B(T)$ respectively, then

$$K(T) = \lambda T \tag{3.2}$$

and

$$\begin{aligned} B(T) &= K(T)b(T) \\ &= \lambda T b(T). \end{aligned} \tag{3.3}$$

Therefore, the total expected cost per unit time over a full cycle of length $T + d_i$ is

$$\begin{aligned} C(T) &= \frac{\text{Total breakdown cost } + \text{ Total inspection repair cost } + \text{ Inspection cost}}{\text{Full cycle of length}} \\ &= \frac{B(T)C_b + \{K(T) - B(T)\}C_d + C_i}{T + d_i} \\ &= \frac{\lambda T\{C_b b(T) + C_d(1 - b(T))\} + C_i}{T + d_i}. \end{aligned} \tag{3.4}$$

Here, the decision variable T would be selected to minimise $C(T)$.

Again, if we are primarily interested in operating an inspection policy to reduce downtime, then the appropriate downtime model is derived by considering the expected downtime associated with failures and the downtime due to an inspection. Under the assumptions of the case being modelled, there is no additional expected downtime due to repairing faults identified at an inspection. so the total expected downtime per unit time is

$$D(T) = \frac{B(T)d_b + d_i}{T + d_i} \quad . \qquad (3.5)$$

The choice of T is made to minimise $D(T)$. In practice, the equation (3.5) may often be used due to the difficulty of establishing an agreed cost measure due to factors in the industrial plant. It may also be used because it is the appropriate model (see Christer and Waller [1984c])

Equation (3.1) to (3.5) constitute the basic inspection model. This basic model may be modified according to need.

3.3 Some Variations on the Basic Model

3.3.1 A Variation on the Downtime Model

Suppose there are insufficient staff available to complete all repairs at an inspection. Here we investigate the changes to the basic model that such a condition will make. For instance, if assumption (c) in the basic model was invalid and there were only enough maintenance staff to complete the task of identifying faults during the inspection period d_i, then further time is required to perform inspection repairs subsequent to inspection, with each inspection repair causing additional downtime. The formulation of $D(T)$ would be modified as follows. If d_d is the expected downtime due to an inspection repair, then assuming repairs are performed sequentially, the total downtime over the full cycle is

$$\lambda T d_b b(T) + d_i + \lambda T d_d \{1 - b(T)\} .$$

Also, considering d_d, the expected length of the full cycle is

$$T + d_i + \lambda T d_d\{1 - b(T)\}.$$

We have, therefore, the downtime per unit time over an inspection cycle is

$$D(T) = \frac{\lambda T d_b b(T) + d_i + \lambda T d_d\{1 - b(T)\}}{T + d_i + \lambda T d_d\{1 - b(T)\}}$$

$$= \frac{\lambda T[d_b b(T) + d_d\{1 - b(T)\}] + d_i}{T + d_i + \lambda T d_d\{1 - b(T)\}} \quad (3.6)$$

It is obvious from the equation (3.6) that as expected, d_d must be less than d_b for the inspection to be worthwhile.

3.3.2 Non-Perfect Inspection Case

So far, it has been assumed that inspections are perfect in that any fault present will be identified. It is, however, unrealistic to expect a perfect inspection every time. It is more likely, in most cases, that the probability of a fault being detected at an inspection is dependent to some extent on the duration of inspection, d_i. Such dependence would need further investigation in a particular context. Here we will be content to review the simpler model used by Christer and Waller [1984a].

We introduce a probability r that a specific fault will be identified at an inspection, and a corresponding probability $(1-r)$ that it will not. The only change that this will produce to the above models will be through the form of $b(T)$, the probability of a fault resulting in a breakdown. To find the new form of $b(T)$, consider a fault which first arises at time y after an inspection at time point 0 (see Figure 3.3). Clearly, $y \leq T$.

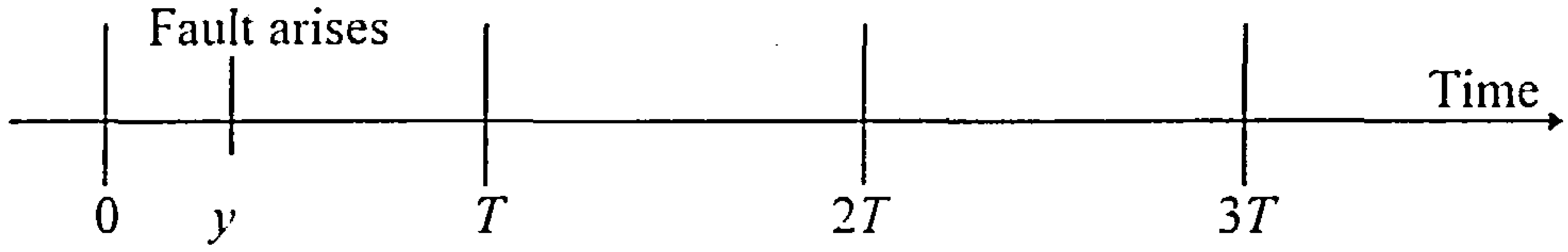


Figure 3.3. Inspection process of a fault arising at time y.

If this fault is subsequently identified at an inspection, it could be the inspection at T if $h>T-y$ or at $2T$ if $h>2T-y$, or failing this, at the inspection at $3T$ if $h>3T-y$, and so on. We have, therefore, that for the fault arising at point y,

Prob(fault identified at T)

= Prob(being identified) × Prob(not resulting in a breakdown before T)

$$= rR(T-y), \qquad (3.7)$$

where

$$R(x) = \int_x^\infty f(h)dh. \qquad (3.8)$$

Similarly,

Prob(fault identified at $2T$)

= Prob(not identified at T but identified at $2T$)

× Prob(not resulting in a breakdown before $2T$)

$$= r(1-r)R(2T-y). \qquad (3.9)$$

In general, the probability that a fault initiated at point y will be identified at the inspection at nT is

$$r(1-r)^{n-1}R(nT-y), \qquad n = 1, 2, \ldots$$

Adding all these probabilities, we obtain the probability $I(T)$, say, that a fault arising at time y will be identified at an inspection as

$$I(T)=\sum_{n=1}^{\infty} r(1-r)^{n-1} R(nT-y). \tag{3.10}$$

Since y can vary uniformly between 0 and T, summing $I(T)$ over all possible y, the probability that a fault arises as a breakdown, $b(T)$, becomes

$$b(T)=1-\int_0^T \sum_{n=1}^{\infty} \frac{r}{T}(1-r)^{n-1} R(nT-y)dy. \tag{3.11}$$

In equation (3.11), for $r=0$ or 1, $b(T)$ corresponds respectively to the failure probability for the conventional failure system, and to the basic inspection model with perfect inspection. Also, as inspection interval period T tends to 0 or ∞, $b(T)$ must intuitively converge to 0 or 1. To confirm this point, rearranging equation (3.11), we have that

$$b(T)=1-\sum_{n=1}^{\infty} r(1-r)^{n-1} \frac{1}{T}\int_0^T R(nT-y)dy. \tag{3.12}$$

If the inspection interval period T tends to 0, it follows that

$$\lim_{T\to 0} b(T)=1-\sum_{n=1}^{\infty} r(1-r)^{n-1} \lim_{T\to 0}\frac{1}{T}\int_0^T R(nT-y)dy, \tag{3.13}$$

and

$$\begin{aligned}\lim_{T\to 0} b(T) &= 1-\sum_{n=1}^{\infty} r(1-r)^{n-1} \lim_{T\to 0}\{nR(nT)-(n-1)R((n-1)T)\}\\ &= 1-r\sum_{n=1}^{\infty}(1-r)^{n-1}\\ &= 0.\end{aligned} \tag{3.14}$$

Similarly, if the inspection interval period T tends to ∞, it follows that

$$\lim_{T\to\infty} b(T) = 1 - \sum_{n=1}^{\infty} r(1-r)^{n-1} \lim_{T\to\infty} \frac{1}{T} \int_0^T R(nT - y)dy$$

$$= 1. \qquad (3.15)$$

Interestingly, in the steady state situation the only changes in permitting imperfect inspection $r \neq 1$ is that $b(T)$ changes in form, but criteria functions, such as $C(T)$ and $D(T)$ given in equations (3.4) to (3.6), remain the same. The above non-perfect inspection formulation with $r \neq 1$ was first used in an application of delay time analysis modelling of the planned maintenance for a vehicle fleet by Christer and Waller [1984b].

3.3.3 A Variation on the Instantaneous Rate of Fault Occurrence

Another assumption which may require to be relaxed is the constant rate of fault occurrence. The effect of this change in the basic perfect inspection model is to modify the formulae for $b(T)$, so producing consequential changes in the expressions for $D(T)$ and $C(T)$. In spite of this change, an inspection renews the system under the perfect inspection by the assumption (c), (f), and (g). Here we assume that the instantaneous rate of fault occurrence at time y after an inspection is not constant but is given by $g(y)$ (see Figure 3.4).

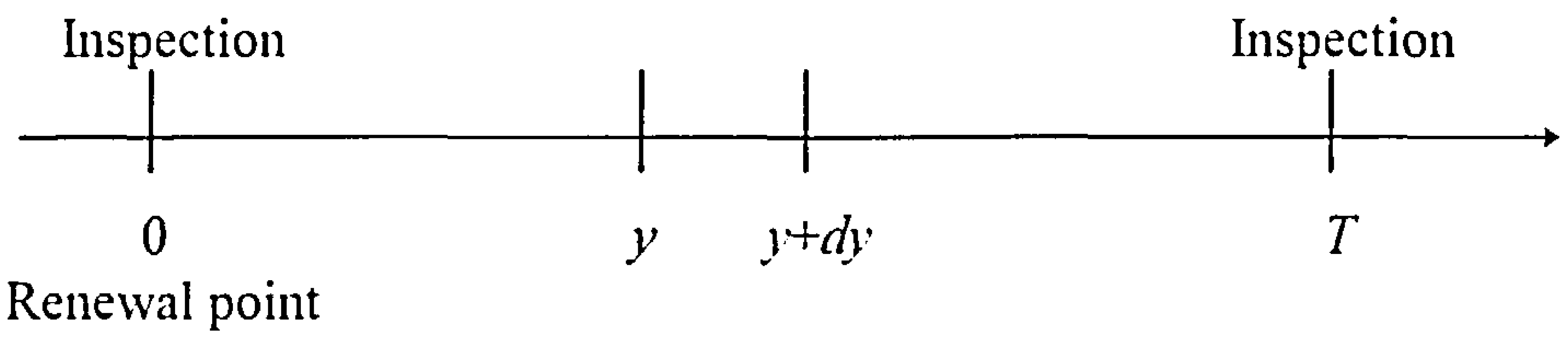


Figure 3.4. A variation on the rate of fault occurrence.

This being so, the expected number of faults arising in the small interval $(y, y+dy)$ is $g(y)dy$. Clearly, the expected number of faults arising in the interval $(0, T)$ is

$$K(T) = \int_0^T g(y)dy \ . \tag{3.16}$$

Assuming perfect inspections, a fault arising in $(y, y+dy)$ with a delay time $h<T-y$ will arise as a breakdown. Therefore the expected number of breakdowns resulting from defects arising in $(y, y+dy)$ is

$$g(y)dy \int_0^{T-y} f(h)dh = F(T-y)g(y)dy \ , \tag{3.17}$$

where

$$F(x) = \int_0^x f(h)dh \ . \tag{3.18}$$

Accordingly, the expected number of breakdowns during the time period $(0, T)$ is

$$B(T) = \int_0^T F(T-y)g(y)dy \, . \tag{3.19}$$

Since, given perfect inspection, the expected number of inspection repairs arising in $(0, T)$ is $K(T)$ - $B(T)$, the expected downtime per unit time is given by

$$D(T) = \frac{d_b B(T) + d_i}{T + d_i} \ . \tag{3.20}$$

The cost model in its simplest form is, assuming inspection repairs are performed during the inspection period d_i ,

$$C(T) = \frac{C_b B(T) + \{K(T) - B(T)\} C_d + C_i}{T + d_i} \ . \tag{3.21}$$

If inspection repair requires additional time to the inspection period of d_d per fault, then equations (3.20) and (3.21) are simply modified to the equations

$$D(T) = \frac{d_b \frac{B(T)}{K(T)} + d_d(1 - \frac{B(T)}{K(T)}) + d_i}{T + d_i + d_d(1 - \frac{B(T)}{K(T)})} \tag{3.22}$$

and

$$C(T) = \frac{C_b B(T) + \{K(T) - B(T)\} C_d + C_i}{T + d_i + (1 - \frac{B(T)}{K(T)})} . \tag{3.23}$$

Christer and Waller [1984a] give a numerical example for a comparison of the results for the three models in terms of the expected proportion of faults arising as breakdowns and the expected downtime. In terms of the expected proportion of faults arising as breakdowns, it is seen from equation (3.12) that the model for non-perfect inspections shows, as would be expected, a higher percentage of breakdowns than the basic model. Also, the model for non-constant rate of fault occurrence showed in the case considered a lower percentage of breakdowns than the basic model. This result is because the delay time distribution and the assumed fault arrival rate has a lower fault frequency earlier in the cycle. Again, the expected downtime figures show that the models with the lowest and highest occurrences of breakdowns have the lowest and highest downtimes respectively, which again is to be expected.

3.4 Parameter Estimation of the Delay Time Model

3.4.1 Subjective Estimation

A task which is vitally important in adopting a delay time model is the estimation of the delay time and initial point distribution. It is not generally possible to measure directly the delay time h associated with a fault, or the initial point u. A method which proved to be possible is to obtain subjective estimates of delay time from the repairing engineers. At a repair of a failure, or when a fault is identified at an inspection, the following questions may be asked of the repairing engineer.

(a) How long ago could the fault have first been noticed by an inspector or operator (=*HLA*)?

(b) If the repair is not carried out now, how much longer could it be delayed before a repair was essential (=*HML*)?

The delay time for each fault is estimated by $h = HLA + HML$ (see Figure 3.5).

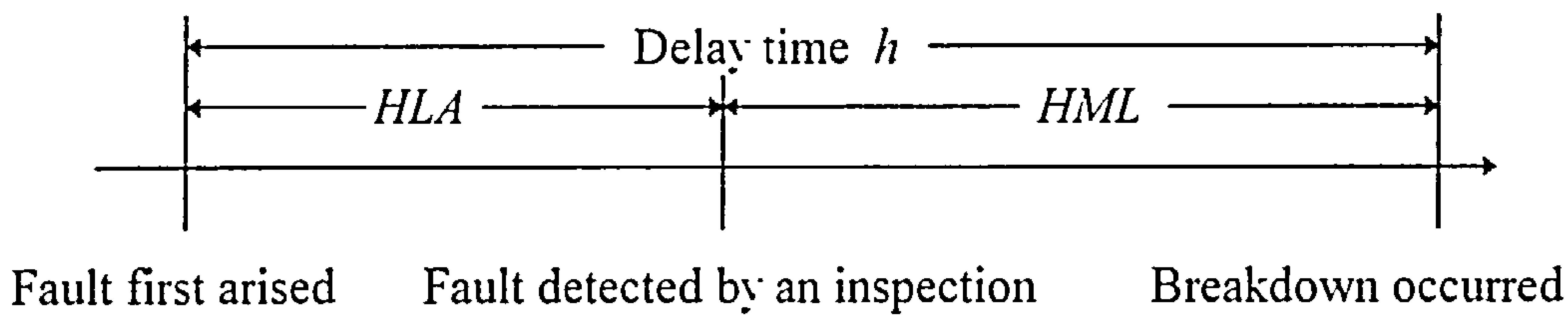


Figure 3.5. Estimation by $h = HLA + HML$.

In this way, by observing sufficient faults, the delay time distribution $f(h)$ may be obtained. Furthermore, at any point in time T when a fault is being attended to, having an estimate of HLA provides at once an estimate of the initial point u, namely $u = T - HLA$. It is the set of such estimates that enables the distribution of the initial points u to be estimated. The method of estimating the distribution parameters of delay time h and initial point u in this way is known as the subjective method. Note that, in adopting the subjective method to obtain delay time and initial point estimates, the definitions of fault and failure are important.

One of the interesting aspects of delay time modelling is that it can use a synthesis of subjectively derived data to model a maintenance situation where the variable of interest can be the expected number of breakdowns over $(0, T)$, $B(T)$, the expected downtime per unit time, $D(T)$, or the expected cost per unit time, $C(T)$. If there is a current policy of inspecting the system at point T_0, then one would expect that the relationships such as the following to hold,

$$B_0 = B(T_0), \tag{3.24}$$

$$D_0 = D(T_0), \tag{3.25}$$

$$C_0 = C(T_0), \tag{3.26}$$

where the left hand side is objective data, the currently observed number of failure, downtime and cost per unit time, and the right hand side is the output of a model based upon a synthesis of subjective assessments. However, the chance of the above relationships being satisfied is remote. The problem is simply stated in Figure 3.6.

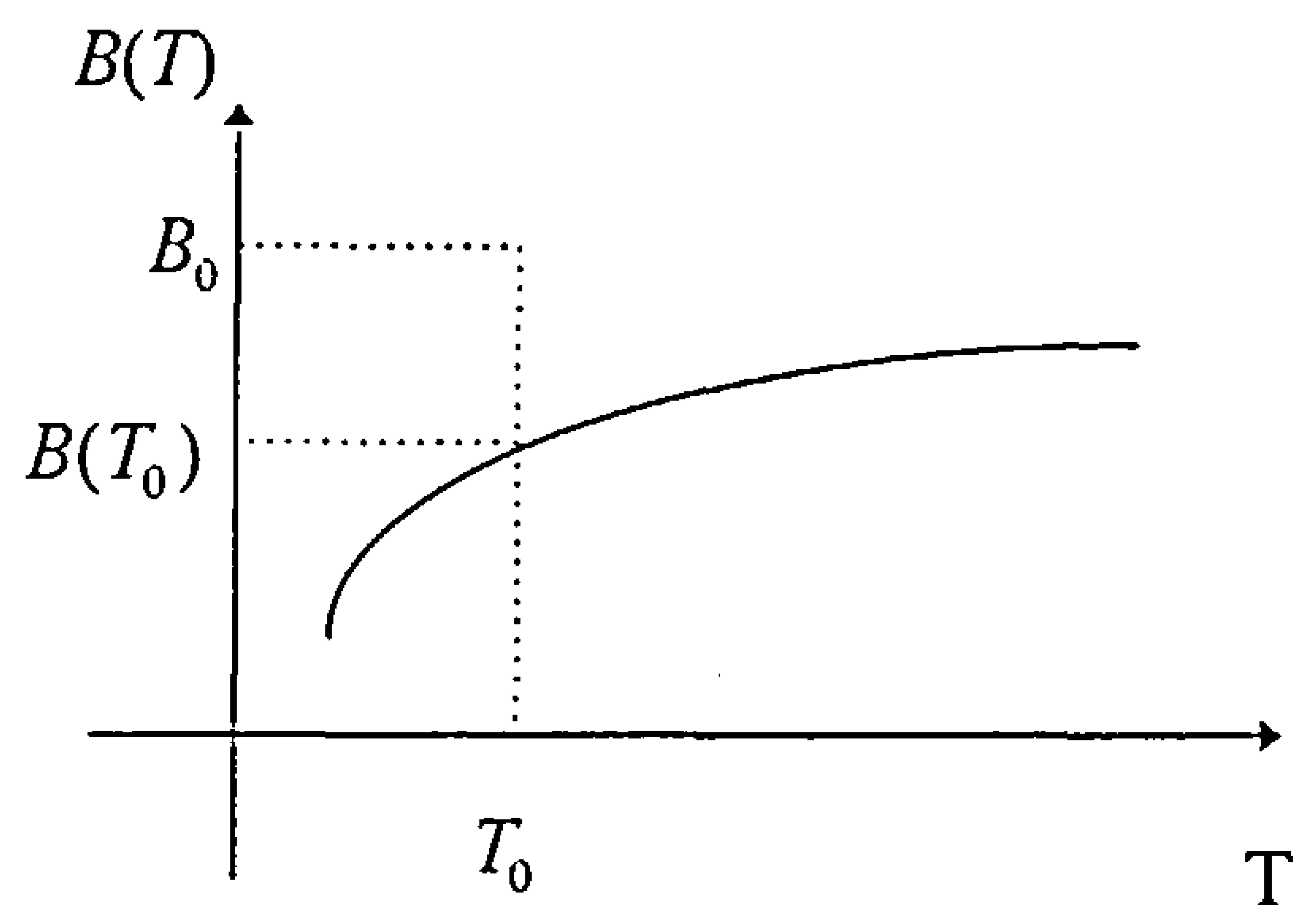


Figure 3.6. Bias of the estimation.

This suggests that a revision will be required in the modelling process, which is expected with any process of decision analysis entailing subjective assessments.

In view of this anticipated problem, a method has been developed by Christer and Redmond [1990] to formally revise or update the prior delay time distribution, $f(h)$, using the known B_0, D_0, or C_0. This is done by a shear transformation of each estimate of delay time $\hat{h}$ to $\hat{h}'$, such as $z = \alpha h + \omega$, where α and ω are the unknown parameters to be determined such that the above-mentioned relationships hold.

The argument for this type of transformation is the observed tendency for estimators to systematically underestimate delay time which, therefore, needs to be extended. In general, we would therefore expect $\alpha>1$.

Another problem which can arise from using the subjective method is one due to sampling bias. At failures, the delay time estimates obtained are $\{\hat{h}_1\}=\{HLA\}$ (since $HML=0$, by definition), which, if this is the only source of estimating data, will generally produce an underestimate of the *pdf* of h, say h_f, because shorter delay times have more chance of leading to a failure. On the other hand, at inspection repairs, the delay time estimates obtained are $\{\hat{h}_2\}=\{HLA+HML\}$, which will produce an overestimate of the *pdf* of h, say h_T, because longer delay times have more chance of spanning the inspection time point T. Christer and Redmond [1990] recognised this, established the existence of bias, and proposed a bias correction method. We examine this briefly below.

Here it will be convenient to define the instantaneous arrival rate $g(u)$, the probability density function $q(u)$ and the cumulative distribution function $Q(u)$ of the initial point u after an inspection. For the present, we suppose that inspections are perfect and inspection points are renewal points. The function $q(u)$ is given by

$$q(u)=\begin{cases} g(u)/K(T) & \text{for } 0\le u\le T, \\ 0 & \text{otherwise} \end{cases}, \tag{3.27}$$

where, as before, $K(T)$ is the expected number of faults arising in the interval $(0, T)$.

First the cumulative distribution function of delay time h_T was considered under the assumption of perfect inspection. The probability that a defect will arise in the interval $(u, u+du)$ given that it is identified at an inspection is

$$P(\text{Initial point } \in (u,u+du)|h>T-u)=\frac{q(u)(1-F(T-u))}{1-b(T)}, \tag{3.28}$$

where, as before, $F(h)$ is the cumulative distribution function of h and $b(T)$ is the probability of a fault arising as a breakdown. The distribution of this defect which has delay time h_T with the delay spanning point T is given by

$$P(h_T \le \xi | h_T > T-u) = \frac{F(\xi)-F(T-u)}{1-F(T-u)} \qquad \text{for } \xi > T-u. \tag{3.29}$$

The cumulative distribution function for h_T is obtained by integrating the product of equations (3.27) and (3.28) over all appropriate values of u (see Figure 3.7).

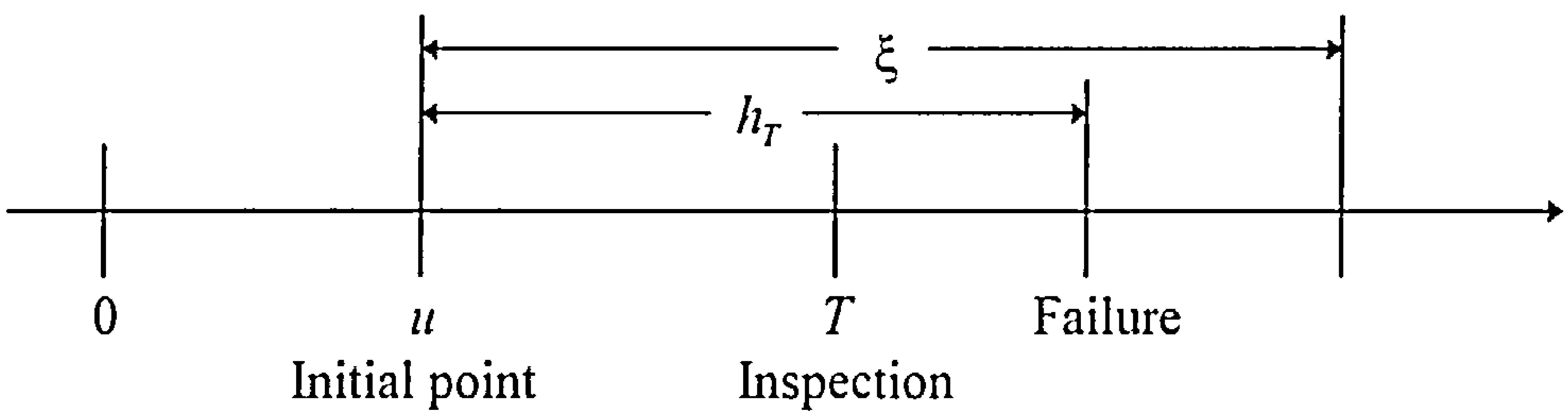


Figure 3.7. Cumulative distribution of the delay time h_T.

If $\xi > T$, u may freely range over the entire interval $(0, T)$ and still be associated with an h_T satisfying $h_T < \xi$. Otherwise, if $\xi < T$, u must be restricted to $(T-\xi, T)$ to satisfy the condition $h_T < \xi$. Noting this point, we have for the cumulative distribution function of h_T,

$$P(h_T \le \xi) = \begin{cases} \displaystyle\int_0^T \frac{q(u)\{F(\xi)-F(T-u)\}}{1-b(T)} du & \text{for } \xi \ge T, \\ \displaystyle\int_{T-\xi}^T \frac{q(u)\{F(\xi)-F(T-u)\}}{1-b(T)} du & \text{for } \xi < T. \end{cases} \tag{3.30}$$

It is noted that $P(h_T \le \xi) \ne F(\xi)$, that is, h_T is a biased estimate of h.

Secondly, we consider the cumulative distribution function for delay time h_f. To determine $P(h_f<\xi)$, it is assumed that inspection is perfect with period T, and a fault which arises at time $u<T$ leads to a failure at time $u+h_f$. If $\xi>T$, then $h_f<\xi$ and so $P(h_f<\xi)=1$. Otherwise, if $\xi<T$, two cases arises which are $u>T-\xi$ and $u<T-\xi$, see Figure 3.8a, b.

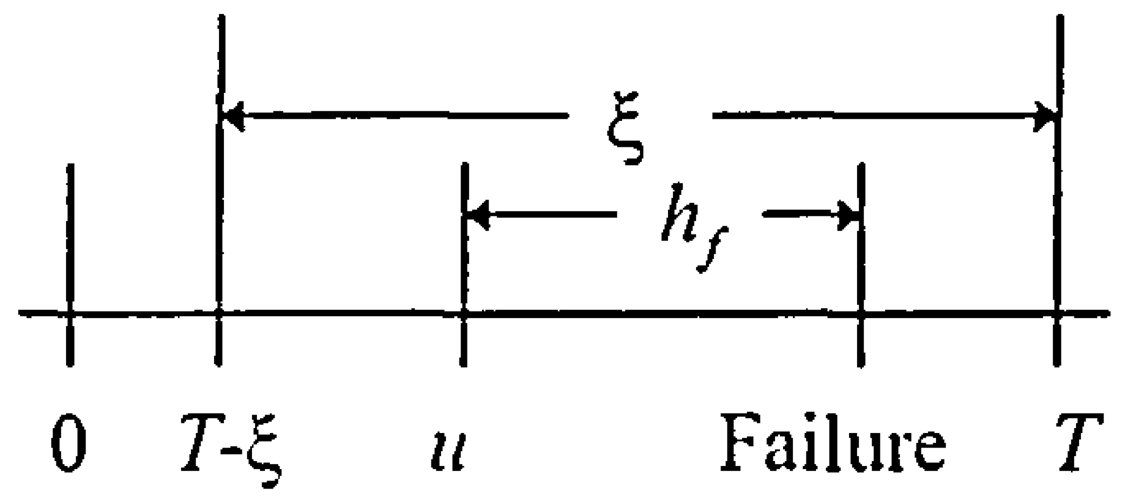


Figure 3.8a. *CDF* of h_f ($u>T-\xi$).

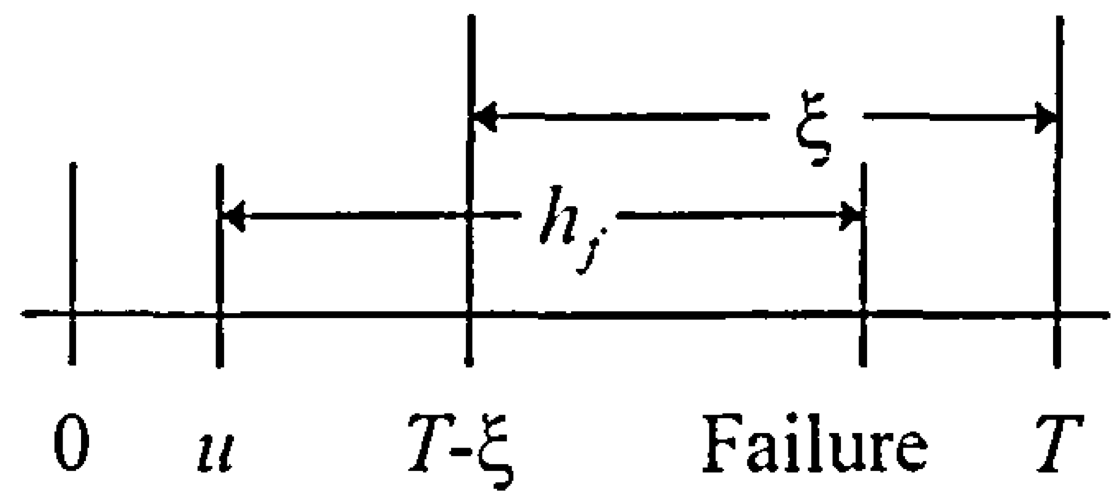


Figure 3.8b. *CDF* of h_f ($u<T-\xi$).

From the Figure 3.8a, if $u>T-\xi$, h_f must automatically satisfy $h_f<\xi$. Therefore, for $u>T-\xi$, we have

$$P(h_f<\xi \mid h_f<T-u)=1. \tag{3.31}$$

Again, for $u<T-\xi$ (Figure 3.8b), for the fault to lead to a failure after time h_f satisfying $h_f<\xi$, we require $h_f<T-u$, and the probability of this event is clearly $F(\xi)/F(T-u)$ for $u<T-\xi$. Collecting these results together, and integrating over all appropriate values of initial point u, we find the cumulative distribution function of h_f to be

$$P(h_f \le \xi)=\begin{cases}\dfrac{1}{b(T)}\{F(\xi)Q(T-\xi)+\int_{T-\xi}^{T}q(u)F(T-u)du\} & \text{for } \xi<T, \\ 1 & \text{for } \xi \ge T.\end{cases} \tag{3.32}$$

From the equation (3.32), clearly $P(h_f < \xi) \neq F(\xi)$ and as expected, $F_{T,f}(\xi) \rightarrow F(\xi)$ as $T \rightarrow \infty$, where $F_{T,f}(\xi) = P(h_f < \xi)$. Therefore the estimate of h, $\hat{h}_1$, becomes asymptotically unbiased with T.

Christer and Redmond [1990] proposed a maximum likelihood refinement for correcting the bias in estimates of delay time. Suppose that there are two data sets of estimates which are failure delay time $\{\hat{h}_j^{(1)}; j = 1,2,\ldots,n\}$ and inspection repair delay time $\{\hat{h}_k^{(2)}; k = 1,2,\ldots,m\}$. So far, the practice in case studies has been to produce a combined set $[\{\hat{h}_j^{(1)}\} + \{\hat{h}_k^{(2)}\}]$ of delay time estimates from which to establish $F(x)$. Here, adopting the same approach, a maximum likelihood refinement may be applied to compensate for the bias. Let the prior distribution for the delay-time be $F(x,\gamma)$, where γ denotes the distribution parameters. Accepting this distribution, we have, from equation (3.29), the distribution of the delay-time of observations spanning an inspection epoch T,

$$F_{T,i}(\xi,\gamma) = \begin{cases} \displaystyle\int_0^T \frac{q(u)\{F(\xi,\gamma) - F(T-u,\gamma)\}}{1-b(T)} du & \text{for } \xi \geq T, \\ \displaystyle\int_{T-\xi}^T \frac{q(u)\{F(\xi,\gamma) - F(T-u)\}}{1-b(T)} du & \text{for } \xi < T. \end{cases} \tag{3.33}$$

Again, the distribution of delay time observations made at failure epochs is, using equation (3.32),

$$F_{T,f}(\xi,\gamma) = \begin{cases} \displaystyle\frac{1}{b(T)}\{F(\xi,\gamma)Q(T-\xi) + \int_{T-\xi}^T q(u)F(T-u,\gamma)du\} & \text{for } \xi < T, \\ 1 & \text{for } \xi \geq T. \end{cases} \tag{3.34}$$

The choice of the parameter γ is made by utilising the maximum likelihood principle in the light of the observations $\{\hat{h}_j^{(1)}\}$ and $\{\hat{h}_k^{(2)}\}$, that is

$$\max_{\gamma}\{\sum_{j=1}^{n}\log f_{T,f}(h_j^{(1)},\gamma)+\sum_{k=1}^{m}\log f_{T,i}(h_k^{(2)},\gamma)\}, \tag{3.35}$$

where f denotes the probability density function of F:

$$f_{T,a}(\xi,\gamma)=\frac{\partial}{\partial\xi}F_{T,a}(\xi,\gamma) \qquad \text{for } a=i,f. \tag{3.36}$$

This optimisation process provides an appropriate fit to the parameter to enable $F(x,\gamma)$ to be defined. Of course, some form of updating adjustment will still be needed, possibly associated with an iteration between correcting for bias, process (3.34), and model and distribution adjustments to the status quo. The main point of the above discussion is that there are methods of estimating and correcting a subjectively derived delay time distribution and model.

Concerning subjective estimation of the delay time, Wang [1996] pointed out that the current practice for asking people for a point estimate or an interval estimate of point measures might not be the best method. He had recognised that the subjective probability estimation in general arose both from the fields of psychology and operational research or statistics. Then, he had proposed an alternative approach to subjective estimation of the delay time for maintenance modelling, namely, to get panel of experts and had estimated from them the probability that the mean delay time of a chosen failure type would lie in a specific time interval. This method in subjective probability and expert judgement assessment had been advocated both by psychologists and statisticians.

3.4.2 Objective Estimation

If objective data are available, Baker and Wang [1992, 1993] have recently introduced a method, now known as the objective method, to estimate the delay time distribution from objective data, that is, data from maintenance records of failures and faults found at inspections or planned maintenance. Essentially the data should include a history of breakdown (failure) times, and the results of planned maintenance or inspections which may be positive (fault found) or negative (no fault found). The objective method utilises the principle of maximum likelihood of observing a sequence of events.

Baker and Wang [1992] first consider the simple case of a single component machine, which may be characterised by the following assumptions.

(1) The time to the initial point of a defect and the subsequent time to failure of the component are independent.
(2) The distributions of initial point u and delay time h are modelled as Exponential or Weibull.
(3) Inspections are perfect.
(4) Repair times are negligible.
(5) Repairs are taken as replacements, so that the faulty component is restored to as new condition.

Assumption (2) is simply considered by Baker and Wang [1992] for convenience. Also the possible events that can contribute to the likelihood are defined as:

N : Inspection and no defect found (negative inspection),
Y : Inspection and defect found (positive inspection),
B : Breakdown (failure),
E : End of observation period.

In addition, the following notation introduced by Baker and Wang [1992] is useful.

R : Replacement on a breakdown, B, or positive inspection, Y,
X : Denotes any event.

Based upon the above assumptions and definitions, Baker and Wang [1992] establish the likelihood of observing a sequence $X_1, X_2, \ldots, X_n$ of events of types B, E, Y, and N by utilising the expression,

$$L = P_{X_1} \times P_{X_2|X_1} \times P_{X_3|X_1X_2} \times \cdots \times P_{X_n|X_1 \cdots X_{n-1}} . \tag{3.37}$$

where P_{X_1} denotes the probability of an event X_1, $P_{X_2|X_1}$ means the probability of event X_2 given that event X_1 has occurred. and so on. Since, after a replacement R, the likelihood does not depend on any event previous to R, the likelihood can be written as the product of terms conditional on events $RX_1X_2...$ starting with the last renewal. Further, since inspections are assumed to be perfect, we have

$$P_{X|RX_1 \cdots X_n} = P_{X|RX_n} . \tag{3.38}$$

Under this concept, three key probabilities can be considered for the described system.

(1) $P_{NB/R}(t_n, t)dt$ is the probability of a sequence of negative inspections of which the last occurs at time t_n from last renewal, and a breakdown at a time between t and $t+dt$ from last renewal. $P_{NB/R}(t_n, t)$ is given by

$$P_{NB/R}(t_n, t) = \int_{t_n}^{t} q(u) f(t-u) du, \tag{3.39}$$

where $q(u)$ is the *pdf* of initial point u, and $f(h)$ is the *pdf* of delay time h.

(2) $P_{NE/R}(t_n, t)$ is the probability of a sequence of negative inspections of which the last occurs at time t_n from last renewal, and no breakdown before observation ceases at time t from last renewal. This probability is given by

$$P_{NE/R}(t_n, t) = 1 - Q(t) + \int_{t_n}^{t} q(u)(1 - F(t-u)) du, \tag{3.40}$$

where $Q(u)$ is the *cdf* of initial point u, and $F(h)$ is the *cdf* of delay time h.

(3) $P_{NYR}(t_n, t)$ is the probability of a sequence of negative inspections of which the last occurs at time t_n, followed by a positive inspection at time t from last renewal. This probability is given by

$$P_{NYR}(t_n, t) = \int_{t_n}^{t} q(u)(1 - F(t - u))du. \tag{3.41}$$

Based on the above probability definitions and assumptions, Baker and Wang [1992] developed the likelihood function L of observing a sequence of events of (a) breakdowns at time t_i^B $(i = 1,2,...,n_B)$, (b) no failure before observation ceases at time t_j^E $(j = 1,2,...,n_E)$, and (c) positive inspections at time t_k^Y $(k = 1,2,...,n_Y)$, as

$$L = \prod_{i=1}^{n_B} P_{NB/R}(t_i^{B^*}, t_i^B) \prod_{j=1}^{n_E} P_{NE/R}(t_j^{E^*}, t_j^E) \prod_{k=1}^{n_Y} P_{NY/R}(t_k^{Y^*}, t_k^Y), \tag{3.42}$$

where $t_i^{B^*}$ is the time of the latest negative inspection, (or, failing that, the latest renewal) such that $t_i^{B^*} < t_i$ and similarly for $t_j^{E^*}$ and $t_k^{Y^*}$. By maximising the likelihood L in equation (3.42), estimates of parameters of the underlying initial point distribution, $q(u)$, and the delay time distribution, $f(h)$, can be obtained.

A development of and the first application within industry of the objective method for estimating delay time parameter is given in Christer *et al* [1995] in the case of a multi-component system. They present a study carried out for a copper products manufacturing company, developing and applying the delay-time modelling technique to model and thus optimise preventive maintenance (PM) of an industrial press. The data available within the plant included the dates and downtimes occurred due to both PM and failures, the nature of the occurrence, and the number of faults found at PM. To estimate the parameters of the fault arrive process and the delay-time distribution, the following assumptions were considered appropriate to their study.

(1) Faults arise according to a homogeneous Poisson process with rate λ.

(2) Faults are assumed to arise independently of each other.

(3) The delay time h of a random fault is independent of its time origin and has *pdf* $f(h)$ and *cdf* $F(h)$.

(4) Inspections carried out at PM are assumed to be imperfect in that they can only identify a fault present with probability r. Probabilities of detection of a fault at successive inspections are independent.

(5) All identified faults are rectified by repairs or replacements during the PM period. This does not influence the development of undetected faults.

(6) Failures are identified immediately, and repairs or replacements are made as soon as possible.

Under the above assumptions, the likelihood expression (3.42) need to be revised because of the non-perfect inspection. The immediate consequence of non-perfect inspection is that an inspection cannot be viewed as a renewal point. Consequently, time measures are from the 'as new' epoch.

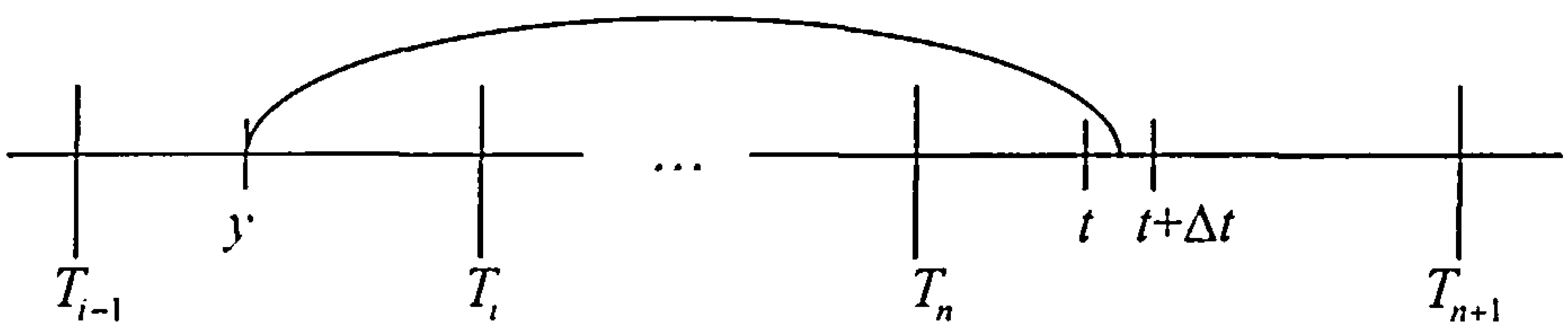


Figure 3.9. Failure process of a fault arising at time y.

If T_i is the time epoch of the ith, $i=1, 2, \ldots,$ PM from new, we have the probability of a failure in $(t, t+\Delta t)$ resulting from a fault arising at time y (see Figure 3.9), is given by

$$P(t,t+\Delta t|y) = \begin{cases} (1-r)^{n-i+1}(F(t+\Delta t-y)-F(t-y)) & \text{for } T_{i-1} < y \le T_i, \ i=1,2,\ldots,n \\ F(t+\Delta t-y)-F(t-y) & \text{for } T_n < y \le t \\ F(t+\Delta t-y) & \text{for } t < y \le t+\Delta t \\ 0 & \text{for } t+\Delta t < y \ . \end{cases} \qquad (3.43)$$

Then, for $T_n < t \le T_{n-1}$, the expected number of failures over $(t, t+\Delta t)$, $EN_f(t, t+\Delta t)$, is

$$EN_f(t, t+\Delta t) = \lambda \int_0^{\infty} P(t, t+\Delta t | y) dy$$

$$= \lambda \sum_{i=1}^{n} (1-r)^{n-i+1} \int_{T_{i-1}}^{T_i} (F(t+\Delta t - y) - F(t-y)) dy$$

$$+ \lambda \int_{T_n}^{t} (F(t+\Delta t - y) - F(t-y)) dy + \lambda \int_t^{t+\Delta t} F(t + \Delta t - y) dy. \quad (3.44)$$

Changing the integral variable and rearranging the integral sequence, and after some manipulation, we have

$$EN_f(t, t+\Delta t) = \lambda \int_t^{t+\Delta t} \sum_{i=1}^{n} (1-r)^{n-i+1} (F(x - T_{i-1}) - F(x - T_i)) dx$$

$$+ \lambda \int_t^{t+\Delta t} F(x - T_n) dx. \quad (3.45)$$

Also, we have the probability of identifying the fault at PM time T_{n+1} resulting from a fault arising at time y, is given by

$$P(T_{n+1} | y) = \begin{cases} (1-r)^{n-i+1} r (1 - F(T_{n+1} - y)) & \text{for } T_{i-1} < y \le T_i, \ i = 1,2,\dots,n \\ r(1 - F(T_{n+1} - y)) & \text{for } T_n < y < T_{n+1} \\ 0 & \text{otherwise} . \end{cases} \quad (3.46)$$

Then, the expected number of faults found at PM time T_{n+1}, $EN_p(T_{n+1})$, is

$$EN_p(T_{n-1}) = \lambda \int_0^{\infty} P(T_{n+1} | y) dy$$

$$= \lambda \sum_{i=1}^{n} (1-r)^{n-i+1} r \int_{T_{i-1}}^{T_i} (1 - F(T_{n+1} - y)) dy$$

$$+ \lambda r \int_{T_n}^{T_{n+1}} (1 - F(T_{n+1} - y)) dy. \quad (3.47)$$

Ross [1983] presented the following proposition in proposition 2.3.2:

If an event arrival process follows a Poisson process with λ, the number of events that occur by time t, $N(t)$, is an independent Poisson random variable having mean given by $\lambda\int_0^t P(s)ds$, where $P(s)$ is the probability that the event occurs independently all else at time s.

Since the fault arrival process is assumed to arise according to a Poisson process, as a generalisation of proposition 3.3.2 in Ross [1983], the number of failures in $(t, t+\Delta t)$ follows a Poisson distribution with mean $EN_f(t,t+\Delta t)$ and the number of faults found at PM follows a Poisson distribution with mean $EN_p(T_{n+1})$. Therefore, the probability of m failures over $(t,t+\Delta t)$, where $T_n < t \le T_{n+1}$, is given by

$$P(m \text{ failures in } (t,t+\Delta t)) = \frac{(EN_f(t,t+\Delta t))^m e^{-EN_f(t,t+\Delta t)}}{m!}, \tag{3.48}$$

and the probability of n faults found at T_{n-1} is

$$P(n \text{ faults at } T_{n-1}) = \frac{(EN_p(T_{n+1}))^n e^{-EN_p(T_{n-1})}}{n!}. \tag{3.49}$$

As previously indicated, the data assumed to be available are the number of failures in each working day and the number of faults identified at PM times. To formulate the likelihood function of the observed event, suppose first that n_i faults have been observed at the ith PM time (i=1,2,...,l). The PM interval (T_{i-1}, T_i) is now divided into k nonoverlapping subintervals of equal length Δt, namely

$$I_j^i = (T_{i-1} + (j-1)\Delta t, T_{i-1} + j\Delta t), \quad j = 1, 2, \ldots, k.$$

where $T_{i-1} + k\Delta t = T_i$. Let m_{ij} denote the number of failures occurring in I_j^i over (T_{i-1}, T_i).

Since we have assumed that all faults are independent of each other, it follows that a fault resulting in a failure will not have any influence on a fault which is found at PM, i.e. the number of failures since the last PM and the number of faults found at PM are also independent. This being so, the likelihood is simply

$$L = \prod_{i=1}^{l} \{P(n_i \text{ faults at } T_i) \prod_{j=1}^{k} P(m_{ij} \text{ failures in } I_j^i)\}. \tag{3.50}$$

Therefore, once $f(h)$ has been specified, it is possible to obtain maximum-likelihood estimates for any unknown parameters. This includes, for example, those inherent in the specification of $f(h)$ and the rate of occurrence of faults, λ, and the probability r that a fault will be identified at PM if it is present.

The above section is concerned chiefly with delay time analysis and the estimation of the underlying delay time distribution and defect arrival process. Once these are estimated it is possible to apply the knowledge to model maintenance problems. This is the main objective of the techniques.

3.5 Application of delay time modelling

There are many applications developing delay time maintenance modelling. Some of them have assumed that the delay time parameters are given. In real-world situation, the delay time parameters are, however, only estimated from the data which are from subjective or objective information. In this section, we are interested specifically in applications of estimating delay time parameters.

The first application of delay time modelling (DTM) was in the context of building maintenance in Christer [1982]. In this pilot application, DTM was applied to assess the potential of an inspection maintenance policy as opposed to the existing breakdown maintenance policy for a building complex. The model was of the pooled components type, (complex or multi-component type), which grouped faults from individual components. It was assumed that faults arose in a homogeneous Poisson process (HPP). Subjective and objective information were both used.

In 1984, attempts were made to apply DTM to a different maintenance context, namely, an industrial plant. In Christer and Waller [1984b], DTM along with the snap-shot modelling techniques was applied to a study to model downtime consequences of maintenance practice of a high-speed product canning process within Pedigree Petfoods Limited. In this study, the snap-shot modelling, which is a problem recognition technique introduced by Christer and Whitelaw [1983] with strong parallels to reliabilities centred maintenance. It was applied to find out which component within the canning plant develop faults most frequently, the causes of the faults, and the possible means of prevention's of these faults. In this study, the DTM was used to model the frequency of pit stops so that downtime can be reduced. The delay time distribution for the model was estimated using the subjective method, and it was found that initially the data had generally been underestimated. Re-estimation was carried out after feedback to the assessors and the result was a very encouraging, and produced modelling which actually satisfied the status quo conditions.

In another case-study in Christer and Waller [1984c], snapshot analysis and the DTM were again applied to modelling preventive maintenance for a vehicle fleet of tractor units operated by Hiram Walker Ltd. In this case study, the subjective method was again used to obtain the estimate of the delay time distribution. The recommended decrease in frequency of maintenance as a result of the modelling, was adopted by the management.

In 1988 the pooled component model applicable to the building industry appeared in Christer [1988]. Here a DTM was developed in which the probability $p(y)$ of detection of a fault at time y from the fault origin time u increased from 0 at $y = 0$ to unity at $y=h$. Repair cost now varied over the delay time as a deterministic function $C(y, h)$.

Developments of this study led to a major collaborative research project with the Concrete Research Group at QMC London into the inspection and repair modelling of concrete bridges and high-rise structures (see Christer and Redmond [1993]).

Later, Chilcott and Christer [1991] used DTM to model the maintenance practices for coal face machinery within British coal. Here, they considered the case of a non-periodic inspection process where all known defects were remedied during the next maintenance period. Also, the delay time parameters were estimated based on the subjective method which were then used to model the effectiveness of condition-based monitoring in reducing downtime. As a result, the financial consequences due to 2% downtime savings are very attractive.

Christer and Redmond [1993] studied the inspection of concrete structure of bridges. Here, it is noted that the deterioration of a component goes through a number of definable states, namely new to cracking, cracking to spalling, and spalling to failure (essential repair). The DTM approach was used, and the delay time was splitted into two phases, cracking and spalling. The delay time at the cracking and spalling phases were represented as h, and v, with *pdf*s $f(h)$ and $w(v)$, respectively. The time u is the time the component starts to deteriorate from new to cracking, and its *pdf* is $q(u)$. It is also assumed that only one type of fault can arise within a single component. The delay time parameters were estimated based on objective data at inspection, namely the age of each component and its condition.

In 1987, a model utilising the notion of delay time was used to establish the reliability consequences of inspecting a single component on different inspection periods, Christer [1987]. In the model, the inspections are assumed to be perfect and non-detrimental. After the inspection the component is returned to the as new condition. The model has recently been explored for application to model the reliability of pumping systems for the water supply in some 4,000 high rise housing in Hong Kong in Leung and Christer [1995]. Also recently, a repeat study was carried out by Christer *et al* [1995] to model preventive maintenance (PM) practices of a cooper products manufacturing in the Northwest of England. In this study, as discussed in the previous section 3.4, the DTM parameters were estimated using objective data, namely, the maintenance record data of failures and faults found at PM. The criterion of interest was the to minimise total downtime over a PM interval. Using the same case study, the subjective method was also used to estimate DTM parameters Christer *et al* [1994]. Although the results of the comparison are still to be formally reported, it is noted that both modelling techniques lead to very similar results and recommendations. This consistency is indicative of a welcome robustness of DTM.

Desa [1995] considered a bus fleet maintenance study for an inter-city bus company in developing country, namely Malaysia. He showed in this study that in a situation where data are almost totally lacking, the snap-shot modelling was both practical and valuable for problem identification and definition. Also, he showed that the use of delay time concept and modelling enables issues related to existing maintenance policy and practice to be evaluated and modelled from a starting position of basically zero data using subjective assessments. That is, in the situation where objective maintenance data are not available, subjectively derived data can be reliably used as the basis for modelling.

3.6 Conclusions

A substantial number of theoretical O.R. models developed for maintenance decision problems have been reported in the literature. On the other hand, the number of reported applications and implementations of these maintenance models to real-world problems is, though increasing, still few. The likely factors contributing to the lack of application of maintenance models have also been identified and highlighted by many authors (e.g. Pintelon and Gelders [1992], Baker and Christer [1994]).

The delay time concept has, however, provided a useful means of modelling the effect of periodic inspections on the failure rate of repairable machinery. The delay time concept defines a two-stage failure process for a component, which consists of a defect first becoming visible at time u from new with probability density function, $q(u)$, and the visible defect developing into a failure after some delay time h with probability density function, $f(h)$. Once these two distributions are known, it is possible to model the reliability, operating cost and availability functions. The distribution functions $q(u)$ and $f(h)$ and their parameters are vital to delay time modelling. Two basic approaches to solve the associated estimation problems, namely subjective and objective methods, have been developed using the information obtainable from engineers who repair the machine. Although many variations in maintenance practice are possible, the situation where an engineer decides at inspection that a component is defective, and replaces it, is very common. Since delay time models can be used for decision-making, for example choosing the interval between inspections to minimise cost or downtime, it may be natural to rely on the delay time modelling in adapting the maintenance models to real-world situation.

Chapter 4

SEMI-MARKOV AND DELAY TIME MODELS OF MAINTENANCE FOR A SINGLE COMPONENT SYSTEM: A COMPARATIVE STUDY

4.1 Introduction

In chapter 2, the literature on inspection modelling has been reviewed and in chapter 3, delay time modelling has been discussed. It was evident from the literature review that inspection models were often formulated as a Markov inspection model. The Markov inspection models assume that the working condition of the system can be expressed as a discrete-time Markov chain with degraded states and a failed state (see Tijms and Van Der Duyn Schouten [1985]). In the literature, the state transition probabilities are characteristically assumed to be given. Such an assumption may be unrealistic. In practice, it is difficult to define the working condition of a deteriorating system in terms of degraded states, and therefore, just as difficult to estimate the state transition probabilities. The literature on Markov models is notably silent on such matters at the current time.

These actual problems which have been addressed using Markov models may also be formulated and solved using the delay time techniques presented by Christer. The delay time concept provides a means of modelling the behaviour of the system, and predicting quantities of interest such as reliability, downtime or cost, and does so under various inspection polices. The delay time model defines a two-stage failure process for a component, which consists of a fault first becoming visible by some inspection techniques at time u from new, with probability density function $q(u)$, and the visible fault developing into a failure after some delay time h with probability density function $f(h)$. We have seen that the model parameters for $q(u)$ and $f(h)$ can be estimated from both subjective or objective data.

In this chapter, we will develop a Markov type inspection model for the simplest problem situation for which both Markov and delay time models are valid. The delay time and Markov models are then compared for the same problem.

4.2 System Description

Consider a repairable single failure mode machine that may become defective or suffer breakdown during the course of its service lifetime. Also, assume that the system is inspected for a visible fault at a regular periodic interval, and that the inspection pattern is re-started after the repair of a failure. For this machine, an inspection policy which minimises the expected total long-run average cost or downtime can be derived. For modelling purposes, the system considered here is assumed to have the following properties.

(1) A fault can be observed by inspections only, and a failure will be observed or repaired immediately if it occurs.

(2) An inspection is undertaken every T time units, and the inspection process restarted after a failure repair.

(3) Inspections are perfect in that any fault present within the machine will be identified at inspection, and no new fault injected because of inspection.

(4) An inspection requires C_i cost units and d_i time units, $d_i << T$.

(5) A fault identified at an inspection will be repaired within the inspection period d_i and the repair cost per defect is C_d units.

(6) The component is repaired immediately upon failures and its repair requires C_b cost units and d_b time units.

(7) The machine is as good as new after a repair.

The possible operation of this system indicated in Figures 4.1 and 4.2 is conveniently classified into two cases.

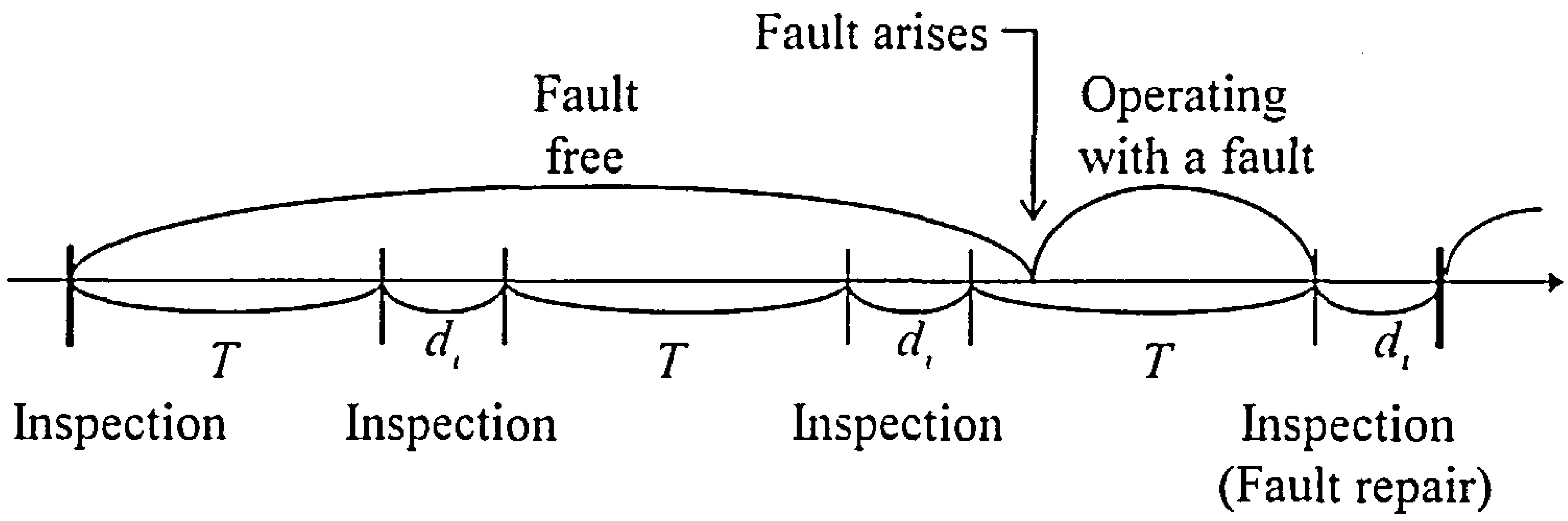


Figure 4.1. Case "1" Inspection cycle.

Case "1", Figure 4.1, is the case where a renewal point occurs after a fault is repaired within an inspection period.

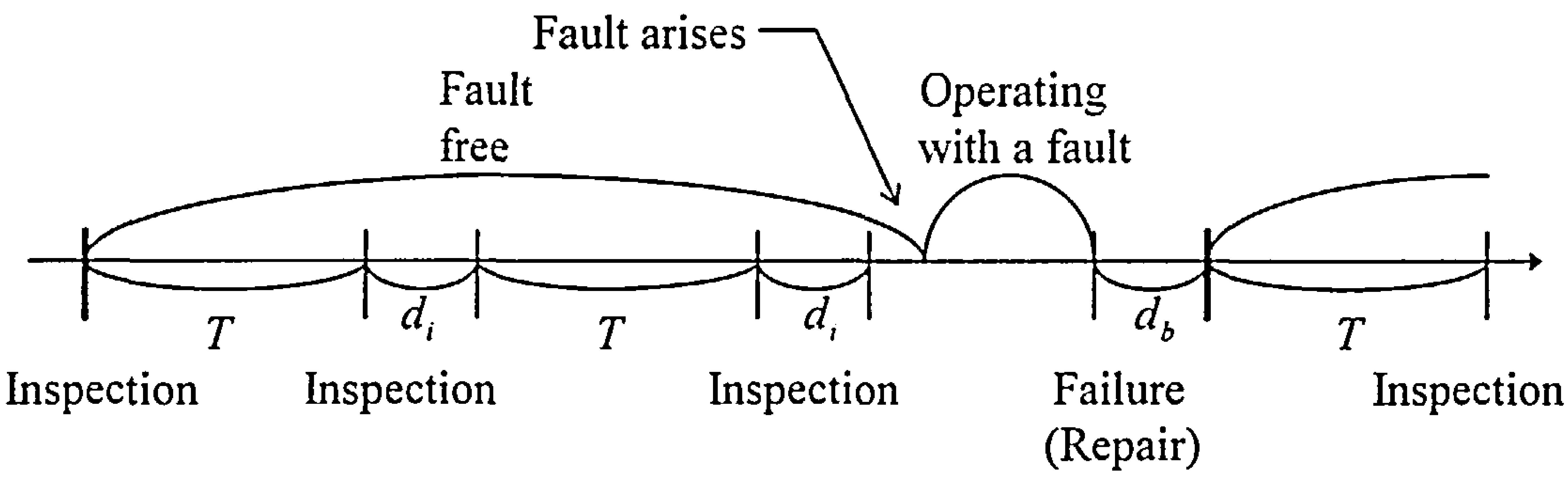


Figure 4.2. Case "2" Failure cycle.

Case "2", Figure 4.2, is the case where a renewal point occurs after a failure is repaired. Note that a renewal point of case "1" or of case "2" may occur after many inspections in which no faults are found.

For this system operating in practice. a variety of data is potentially avaiable. The data could include the inspection time point. the inspection result, and the time of failure. As an example, such characteristic data is summarised in Table 4.1 for $T = 10$ units.

Table 4.1. Characteristic structure of the data of the single component system for T=10 units.

Time point	Inspection result or failure	Action
0	Normally operated	No action
10	No fault detected	No action
20	A fault detected	Repair
27	Failure	Repair
37	A fault detected	Repair
47	No fault detected	No action
54	Failure	Repair
:	:	:
:	:	:

This data indicates fault free operation to $T = 10$, a defect arisen and being identified at $T = 20$, and a fault then arising and leading to a failure at $T = 27$. The inspection process then restarts. If it is not possible to obtain such data from actual maintenance and operating record. assuming a delay time model, it is possible to simulate the process and thereby obtain these data. Either way, having the data, it may then be used as an exercise or experiment to estimate, or recapture, the underlying delay time parameters and in this way model the real-world, or simulated real-world, situation. In section 4.5, we will describe the process of simulating data. This is important in validating the modelling methodology.

4.3 Semi-Markov Inspection Model

4.3.1 Introduction to the Semi-Markov Inspection Model

In the past, various maintenance problems assuming Markovian deterioration of systems have, as already commented, been presented in the literature. Emphasis is usually placed upon the modelling of the system, assuming parameters to be known, and upon the mathematical derivation of the model or the properties of the model. However, only a limited number of Markov papers have discussed the associated computational methods for computing maintenance policies. Examples are Tijms [1986] and Tijms and Van Der Duyn Schouten [1985] which dealt with an optimal inspection and replacement problem of a discrete-time Markovian deterioration system, and proposed a computational algorithm by modifying the policy iteration method.

In this section, we review Tijms' semi-Markov decision process and apply it to the optimal inspection problem of the discrete Markovian deterioration system described in the previous section, and use the Tijms and Van Der Duyn Schouten's computational algorithm to evaluate the results.

4.3.2 Semi-Markov Decision Process

Semi-Markov decision models are concerned with dynamic systems which at random points in time may be observed and classified into one of a possible number of states. The set of possible states is denoted by I. After observing the state of the system, a decision has to be made, and costs are incurred as a consequence of the decision made. For each state $i \in I$, a set $A(i)$ of possible actions is available. It is assumed that the state space I and the action sets $A(i)$, $i \in I$, are finite. The time between two consecutive renewal points is measured in discrete time steps Δt. A semi-Markov decision process has to satisfy the following Markovian Properties.

(1) If at a decision epoch the action a is chosen in state i, the system state at the next decision epoch depends only on the present state i and the chosen action a regardless of the past history of the system.
(2) The costs incurred until the next decision epoch depend only on the present state and the action chosen in that state. Here, cost is taken as the consequence variable of interest, which could include downtime as well as direct cost.

The long-run average cost per unit time or the long-run average downtime per unit time is often taken as the optimality criterion in the maintenance decision-making process. To establish these criterion functions in the context of a semi-Markov decision model, we define the following characteristics.

$P_{ij}(a)$: the probability that at the next decision epoch the system will be in state j if action a is chosen in the present state i.

$\tau_i(a)$: the expected time until the next decision epoch if action a is chosen in the present state i.

$C_i(a)$: the expected cost of the action a if action a is chosen in the present state i.

$D_i(a)$: the expected downtime of the action a if action a is chosen in the present state i.

It is assumed that $\tau_i(a) > 0$ for all $i \in I$ and $a \in A(i)$, that is two or more events cannot occur at the same instant in time. We define a stationary policy R as a rule which prescribes the same action $R_i \in A(i)$ whenever the system is observed in state i at a decision epoch. It can be shown that under a stationary policy, because of the finite state space, the number of decisions made in a finite time interval is finite with probability 1. Also, denoting by X_n the state of the system at the nth decision epoch from new, it follows that under a stationary policy R, the embedded stochastic process $\{X_n\}$ is a discrete-time Markov chain with one-step transition probabilities $P_{ij}(R_i)$.

Introducing and defining the random variable $Z(t)$: the total costs incurred from the initial point $t = 0$ to time t, $t \geq 0$, and denoting by $E_{i,R}$ the expectation operator when the initial state $X_0 = i$ and the stationary policy R is used, then the limit of the expected cost per unit time, $g_i(R)$, is given by

$$g_i(R) = \lim_{t \to \infty} \frac{1}{t} E_{i,R}[Z(t)] \qquad \text{for all } i \in I. \tag{4.1}$$

We can give a stronger interpretation for the average cost function $g_i(R)$. If the initial state i is recurrent under policy R, then the long-run actual average cost per unit time equals $g_i(R)$ with probability 1. In the case when the Markov chain $\{X_n\}$ associated with policy R has no two disjoint closed sets, the Markov chain $\{X_n\}$ has a unique equilibrium distribution $\{\pi_j(R), j \in I\}$, and $g_i(R) = g(R)$ independently of the initial state $X_0 = i$.

Tijms [1986] shows that if the embedded Markov chain $\{X_n\}$ associated with policy R has no two disjoint closed sets, then

$$\lim_{t\to\infty}\frac{Z(t)}{t} = g(R) \qquad \text{with probability } 1 \tag{4.2}$$

for each initial state $X_0 = i$, where the constant $g(R)$ is given by

$$g(R) = \frac{\sum_{j\in I} C_j(R_j)\pi_j(R)}{\sum_{j\in I} \tau_j(R_j)\pi_j(R)} \ . \tag{4.3}$$

Therefore, a stationary policy R^* is said to be average cost optimal if $g_i(R^*) \le g_i(R)$ for all $i\in I$ and all stationary policies R.

For computing an average optimal cost, a policy-iteration algorithm can be developed. The policy-iteration algorithm requires that for each stationary policy the embedded Markov chain $\{X_n\}$ has no more than one disjoint closed set. Suppose that $g(R)$ is the average cost and $v_i(R)$, $i\in I$, are the relative values of a stationary policy R. If a stationary policy $\overline{R}$ is constructed such that, for each state $i\in I$,

$$C_i(\overline{R_i}) - g(R)\tau_i(\overline{R_i}) + \sum_{j\in I} P_{ij}(\overline{R_i})v_j(R) \le v_i(R)\,, \tag{4.4}$$

then $g(\overline{R}) \le g(R)$.

Under these conditions, following Tijms' semi-Markov decision process [1986], we can now formulate the following policy-iteration algorithm.

Policy-iteration algorithm

Step 0 (*initialization*) : Choose an initial stationary policy R.

Step 1 (*value-determination step*) : For the current rule R, compute the average costs $g(R)$ and the relative values $v_i(R)$, $i \in I$, as the unique solution to the linear equations

$$v_i = C_i(R_i) - g(R)\tau_i(R_i) + \sum_{j \in I} P_{ij}(R_i)v_j, \qquad i \in I. \tag{4.5}$$

$$v_s = 0, \tag{4.6}$$

where s is an arbitrarily chosen state.

Step 2 (*policy-improvement step*) : For each state $i \in I$, determine an action a_i yielding the minimum in

$$\min_{a \in A(i)} \{C_i(a) - g(R)\tau_i(a) + \sum_{j \in I} P_{ij}(a)v_j(R)\}. \tag{4.7}$$

The new stationary policy $\overline{R}$ is obtained by choosing $\overline{R_i} = a_i$ for all $i \in I$ with the convention that $\overline{R_i}$ is equal to the old action R_i when this action minimizes the policy-improvement quantity.

Step 3 (*convergence test*) : If the new policy $\overline{R}$ equal the old policy, the algorithm is stopped with policy R. Otherwise, the algorithm cycles back to step 1 with R replaced by $\overline{R}$.

It can be shown that the algorithm converges in a finite number of iterations to an average cost optimal policy. Also, as a consequence of the convergence of the algorithm, there exist $g^*(R)$ and v_i^*, $i \in I$, satisfying

$$v_i^* = \min_{a \in A(i)} \{C_i(a) - g^*(R)\tau_i(a) + \sum_{j \in I} P_{ij}(a)v_j^*\}, \qquad i \in I. \tag{4.8}$$

The constant $g^*(R)$ is uniquely determined as the minimal average cost per unit time. Moreover, each stationary policy whose actions minimize the right side of (4.8) for all $i \in I$ is average cost optimal.

4.3.3 A Semi-Markov Inspection Model

The inspection system mentioned in the section 4.2 can be analysed in the framework of a semi-Markov decision process to find the long term consequences of different inspection periods when the distribution of the time to the initial defect is negative exponential. For a single component system with discrete-time Markovian deterioration, according to the degree of deterioration, the states of the system are represented as 0, 1 and f, where state 0 represents a normally operated state without a defect, state 1 represents a defective state, and state f represents a failed state. These states have been used in the paper of Tijms and Van Der Duyn Schouten' [1985] in denoting the working condition of the system. Without any maintenance activities, the state of the system at discrete-times $t = 0, \Delta t, 2\Delta t, \ldots$ is assumed to undergo deterioration according to a discrete-time Markov chain. Accepting the validity of their concept of deterioration and assuming the inspection system as outlined in section 4.2, we formulate the simplest possible case of an Markovian inspection policy, which may be characterised by the following additional assumptions.

(8) Opportunities for inspections occur only at equidistant points in time $t = 0, \Delta t, 2\Delta t, \ldots$ from new or a renewal repair.

(9) The working condition, or state, of the system cannot improve on its own.

(10) In absence of inspections and repairs, the working condition of the system follows a discrete-time Markov chain.

The decision epochs of this model are the epochs at which opportunities for inspections occur when the system is operating, and the epochs at which the exact working condition of the system is revealed by either an inspection or the immediate detection of a breakdown. We take as the state space

$$I = \{i \mid i = 0, 1, f\} \cup \{(i, m\Delta t) \mid i = 0, m = 1, 2, \ldots, M\} , \qquad (4.9)$$

where Δt is an arbitrary small time which defines the mesh over which discrete time steps are measured. State i is the working condition 0, 1 or f that describe degrees of deterioration. Also, the state $(i, m\Delta t)$ stands for the state that $m\Delta t$ time units have passed since the last inspection, or failure, where it was revealed that the working condition was i and the component has not failed currently, or $m\Delta t$ time units ago when in state i. For the current problem, we have $i = 0$ since each inspection or failure repair is a renewal point. It is noted that M is an arbitrary upper limit of m which will have no influence to the modelling if a finite solution exists. M serves to remind us of the finite nature of the state space. The possible actions a are denoted by

$$a = \begin{cases} 0, & \text{leave the system as it is,} \\ 1, & \text{inspect the system,} \\ 2, & \text{repair the system.} \end{cases}$$

To formulate the one-step transition probabilities from state i to sate j if action a is taken at state i, $P_{ij}(a)$, the one-step expected transition times $\tau_i(a)$, the one-step expected costs $C_i(a)$ and the one-step expected downtimes $D_i(a)$, we introduce the following deterioration probabilities. For $t = \Delta t, 2\Delta t, 3\Delta t, \ldots$, $i = 0$ and $j = 0, 1, f$, we define the deterioration probability r_{ij}^t, where

r_{ij}^t = the probability that t time units from now the system will have working condition j when the present working condition is i and no intervening inspections and repairs take place.

Interestingly, the deterioration probability r_{ij}^t can be easily estimated in terms of the delay time concept since the latter is a more fundamental concept. This will be discussed in subsection 4.3.4.

The one-step transition probabilities can now be readily obtained by considering the action taken at each present state based upon the deterioration probability r_{ij}^t. Firstly, at state 0, action $a = 0$, which is the appropriate action to take at state 0, is taken. If the action $a = 0$ is taken at state 0, the system will either survive until next decision epoch Δt or fail within the next decision epoch Δt. So, from the definition of the deterioration probability r_{ij}^{t}, it is obvious that

$$P_{0y}(0) = \begin{cases} r_{0f}^{\Delta t} & \text{for } y = f \\ 1 - r_{0f}^{\Delta t} & \text{for } y = (0, \Delta t) \\ 0 & \text{otherwise} \end{cases} \quad (4.10)$$

where as before, $(0, \Delta t)$ represents the state that Δt time units have passed since the last inspection, or failure. Also, if the system survives until the next decision epoch Δt, the expected time incurred at state 0 is only the time to the next decision epoch Δt. If the system fails before the next decision epoch Δt, assuming that the next decision Δt is very small, we can approximately regard the expected time incurred at state 0 as time to the next decision epoch Δt. Accordingly, we have that

$$\tau_0(0) = \Delta t. \quad (4.11)$$

When we take the action $a = 0$ at state 0, since no cost and downtime are incurred, it follows that

$$C_0(0) = 0 \quad (4.12)$$

and

$$D_0(0) = 0. \quad (4.13)$$

At state 1, action $a = 2$, which is the only possible action to take at state 1 because a fault found at an inspection has to be repaired, can be taken. If the action a=2 is taken at state 1, we have by the assumption (4) and (6) of the section 4.2 that

$$P_{10}(2) = 1. \tag{4.14}$$

Also, since the fault identified at an inspection is repaired within the inspection period and its repair costs per defect is C_d units, it is obvious that

$$\tau_1(2) = 0, \tag{4.15}$$

$$C_1(2) = C_d, \tag{4.16}$$

and

$$D_1(2) = 0. \tag{4.17}$$

At state f, action $a = 2$ is the only possible action to take because a failure must be immediately repaired. Accepting the action $a = 2$ is taken in state f, we have by assumption (6) and (7) of the section 4.2 that

$$P_{f0}(2) = 1. \tag{4.18}$$

Also, since a failure repair requires C_b cost units and d_b time units, it is obvious that

$$\tau_f(2) = d_b, \tag{4.19}$$

$$C_f(2) = C_b, \tag{4.20}$$

and

$$D_f(2) = d_b. \tag{4.21}$$

At state $(0, m\Delta t)$, the actions $a = 0$ and $a = 1$, which are the only possible actions to take at state $(0, m\Delta t)$, can be taken. Since state $(0, m\Delta t)$ represents the situation that no failure has occurred in the time period of $m\Delta t$ time units starting from state 0, all the one-step probabilities transitions from $(0, m\Delta t)$ are conditional upon no inspections or repairs being undertaken within $m\Delta t$ time units. This means that there must be no failure in the first $m\Delta t$ time units from state 0. If action $a = 0$ is taken at state $(0, m\Delta t)$ for $m = 1, 2, \ldots, M\text{-}1$, there must have been no failure in the $m\Delta t$ time units from state 0, and the system will either survive until next decision epoch $(m+1)\Delta t$ or fail before the next decision epoch $(m+1)\Delta t$. The probability that there is no failure in the $m\Delta t$ from state 0 is $1 - r_{0f}^{m\Delta t}$ and the probability that there is a failure between the present decision epoch $m\Delta t$ and the next decision epoch $(m+1)\Delta t$ is $r_{0f}^{(m+1)\Delta t} - r_{0f}^{m\Delta t}$. By the definition of the deterioration probability, we have that

$$P_{(0,m\Delta t)y}(0) = \begin{cases} \dfrac{r_{0f}^{(m+1)\Delta t} - r_{0f}^{m\Delta t}}{1 - r_{0f}^{m\Delta t}} & \text{for } y = f \\ \dfrac{1 - r_{0f}^{(m+1)\Delta t}}{1 - r_{0f}^{m\Delta t}} & \text{for } y = (0,(m+1)\Delta t) \\ 0 & \text{otherwise .} \end{cases} \tag{4.22}$$

Also, if the system survives until the next decision epoch $(m+1)\Delta t$ from the current decision epoch $m\Delta t$, the expected time incurred by taking the action $a = 0$ at state $(0, m\Delta t)$ is Δt time units from the current decision epoch $m\Delta t$ to the next decision epoch $(m+1)\Delta t$. If the system fails before the next decision epoch $(m+1)\Delta t$, assuming that Δt is very small, we can approximately regard the time of failure as $(m+1)\Delta t$. That is the expected time incurred by taking the action $a = 0$ at state $(0, m\Delta t)$ is Δt time units to the next decision epoch $(m+1)\Delta t$. Accordingly, we have that

$$\tau_{(0,m\Delta t)}(0) = \Delta t \ , \tag{4.23}$$

When the action $a = 0$ is taken at state $(0, m\Delta t)$, since no cost and downtime are incurred, it follows that

$$C_{(0,m\Delta t)}(0) = 0\,, \tag{4.24}$$

and

$$D_{(0,m\Delta t)}(0) = 0\ . \tag{4.25}$$

If action $a = 1$ is taken at state $(0, m\Delta t)$, for $m = 1, 2, \ldots, M$, assuming perfect inspection the system will be in a situation of either having a fault or having no fault at an inspection. Since the probability of finding no fault at an inspection is $r_{00}^{m\Delta t}$ and the probability of finding a fault is $r_{01}^{m\Delta t}$ by definition of the deterioration probability, under the condition that there is no failure until $m\Delta t$ from the state 0, we have that

$$P_{(0,m\Delta t)y}(1) = \begin{cases} \dfrac{r_{00}^{m\Delta t}}{1 - r_{0f}^{m\Delta t}} & \text{for } y = 0 \\ \dfrac{r_{01}^{m\Delta t}}{1 - r_{0f}^{m\Delta t}} & \text{for } y = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4.26}$$

Also, since an inspection requires d_i time units and C_i cost units, it is obvious that

$$\tau_{(0,m\Delta t)}(1) = d_i\ , \tag{4.27}$$

$$C_{(0,m\Delta t)}(1) = C_i\ , \tag{4.28}$$

and

$$D_{(0,m\Delta t)}(1) = d_i\ . \tag{4.29}$$

Now, utilising the standard semi-Markov decision model of the previous subsection 4.3.2 (see the equation (4.5)), we can formulate the cost model as

$$v_0 = -g_c(R)\Delta t + P_{0f}(0)v_f + P_{0(0,\Delta t)}(0)v_{(0,\Delta t)} \, , \tag{4.30}$$

$$v_1 = C_d + v_0 \, , \tag{4.31}$$

$$v_f = C_b - g_c(R)d_b + v_0 \, , \tag{4.32}$$

$$v_{(0,m\Delta t)} = -g_c(R)\Delta t + P_{(0,m\Delta t)f}(0)v_f + P_{(0,m\Delta t)(0,(m+1)\Delta t)}(0)v_{(0,(m+1)\Delta t)} \, , \quad \text{for } \; 0 < m\Delta t < s \, , \tag{4.33}$$

and

$$v_{(0,m\Delta t)} = C_i - g_c(R)d_i + P_{(0,m\Delta t)0}(1)v_0 + P_{(0,m\Delta t)1}(1)v_1 \, , \quad \text{for } \; s \le m\Delta t \le M\Delta t \, , \tag{4.34}$$

where $g_c(R)$ is the expected average cost per unit time given policy R, s is the time to next inspection when the working condition revealed at present inspection time is 0, and v_x , $x \in I$, are the relative costs of the various starting states when policy R is used. Using the embedded technique, by a repeated application of the above equations, we can get the expected cost per unit time. By putting $v_0 = 0$ the linear equation can determine uniquely the average cost $g_c(R)$. Once $g_c(R)$ and v_0 have been determined we can obtain all the values v_x by recursive calculations if they are required.

In a similar way, we have for the corresponding downtime model that

$$w_0 = -g_d(R)\Delta t + P_{0f}(0)w_f + P_{0(0,\Delta t)}(0)w_{(0,\Delta t)} \, , \tag{4.35}$$

$$w_1 = w_0 \, , \tag{4.36}$$

$$w_f = d_b - g_d(R)d_b + w_0 \, , \tag{4.37}$$

$$w_{(0,m\Delta t)} = -g_d(R)\Delta t + P_{(0,m\Delta t)f}(0)w_f + P_{(0,m\Delta t)(0,(m+1)\Delta t)}(0)w_{(0,(m+1)\Delta t)},$$

$$\text{for } 0 < m\Delta t < s, \qquad (4.38)$$

and

$$w_{(0,m\Delta t)} = d_i - g_d(R)d_i + P_{(0,m\Delta t)0}(1)w_0 + P_{(0,m\Delta t)1}(1)w_1,$$

$$\text{for } s \le m\Delta t \le M\Delta t, \qquad (4.39)$$

where $g_d(R)$ is the average downtime per unit time given policy R and w_x, $x \in I$, are the relative downtimes resulting from the various starting states when policy R is used. Using the same embedded technique, we can obtain the expected average downtime per unit time $g_d(R)$. By putting one of the relative downtimes equal to zero, say $w_0 = 0$, the linear equation can determine uniquely the average downtime per unit time $g_d(R)$. Once $g_d(R)$ and w_0 have been determined we can obtain all the relative downtimes w_x by recursive calculations if required.

The cost model and the downtime model can be evaluated using the following policy-iteration algorithm of the subsection 4.3.2.

Policy-iteration algorithm

Step 0 : Choose an initial policy R with the parameter s.

Step 1 : For the current rule R, compute the average costs $g_c(R)$ and the relative costs v_i, $i \in I$, or the average downtimes $g_d(R)$ and the relative downtimes w_i, $i \in I$, as the unique solution to the linear equations

$$v_i = C_i(R_i) - g_c(R)\tau_i(R_i) + \sum_{j \in I} P_{ij}(R_i)v_j, \qquad i \in I, \qquad (5.40)$$

$$v_x = 0,$$

in the cost case, or in the downtime case

$$w_i = D_i(R_i) - g_d(R)\tau_i(R_i) + \sum_{j\in I} P_{ij}(R_i)w_j, \qquad i\in I, \tag{5.41}$$

$$w_x = 0,$$

where x is an arbitrarily chosen state.

Step 2 : For each state $i\in I$, determine an action a_i yielding the minimum in

$$\min_{a\in A(i)}\{C_i(a) - g_c(R)\tau_i(a) + \sum_{j\in I} P_{ij}(a)v_j(R)\},$$

or

$$\min_{a\in A(i)}\{D_i(a) - g_d(R)\tau_i(a) + \sum_{j\in I} P_{ij}(a)w_j(R)\}.$$

The new stationary policy $\overline{R}$ is obtained by choosing $\overline{R_i} = a_i$ for all $i\in I$ with the convention that $\overline{R_i}$ is chosen as being the old action R_i when this action minimises the policy-improvement quantity.

Step 3 : If the new policy $\overline{R}$ equal the old policy, the algorithm is stopped with policy R. Otherwise, the algorithm cycles back to step 1 with R replaced by $\overline{R}$.

This algorithm generates a sequence of improving control-limit rules and it can be shown that the algorithm converges after a finite number of iterations to an average cost or downtime optimal policy (see Tijms and Van Der Duyn Schouten [1985] and Tijms [1986]). Also, as a consequence of the convergence of the algorithm, there exist a $g_c^*(R)$ and v_i^*, $i\in I$, or $g_d^*(R)$ and w_i^*, $i\in I$, where the constant $g_c^*(R)$ is uniquely determined as the minimal average cost per unit time and v_i^* as the relative cost or $g_d^*(R)$ is uniquely determined as the minimal average downtime per unit time and w_i^* as the relative downtime, when the decision variable s would be selected to minimise the average cost per unit time or the average downtime per unit time.

4.3.4 Estimating the Deterioration Probability Based upon the Delay Time Concept

The above semi-Markov inspection model formulation is dependent upon the deterioration probabilities, r_{ij}^t, where i is the working condition 0 and j is the working condition 0, 1 and f. This model assumes that the deterioration probabilities are given or readily avaiable. In practice the deterioration probabilities are not given, but need to be estimated from the available data. In this subsection we consider the estimation of the parameters related to the current finite semi-Markov decision process with deterioration probabilities. Such statistical estimating problems are of prime importance in mathematical modelling.

Since the deterioration probability r_{ij}^t is a probability that t time units from now the system will have working condition j when the present working condition is i and no inspections and repairs are undertaken, it can be easily obtained from the delay time concept. Here we are interested in the deterioration probabilities r_{00}^t, r_{01}^t and r_{0f}^t. Firstly, we consider the deterioration probability r_{00}^t, that is the probability that t time units from now the system will have the working condition 0 when the present working condition is 0 and no inspections and repairs are undertaken. According to the delay time concept, this is a case that the system has no initial point u within the period $[0,t]$. In this case, the deterioration probability r_{00}^t is given by

$$\begin{aligned} r_{00}^t &= \Pr(u \geq t) \\ &= 1 - \int_0^t q(u)du, \end{aligned} \qquad (4.42)$$

where, as before, $q(u)$ is the *pdf* of the initial point u. As already pointed, $q(u)$ is the negative exponential distribution in this case, but we keep a more general notation to assist us in subsequent robustness analysis of the model when $q(u)$ is not negative exponential. Secondly, the deterioration probability r_{01}^t is the probability that the system has an initial point u within period $[0, t]$ but is still operating normally without a failure at time t. In this case, the deterioration probability r_{01}^{t} is given by

$$\begin{aligned} r_{01}^{t} &= \Pr(u - h \geq t \text{ and } 0 \leq u \leq t) \\ &= \int_{0}^{t} q(u)(1 - F(t-u))du, \end{aligned} \tag{4.43}$$

where $F(h)$ is the *cdf* of the delay time h. Lastly, the deterioration probability r_{0f}^{t} is the probability of a failure within $[0, t]$. In this case, the deterioration probability r_{0f}^{t} is given by

$$\begin{aligned} r_{0f}^{t} &= \Pr(u - h \leq t) \\ &= \int_{0}^{t} q(u)F(t-u)du. \end{aligned} \tag{4.44}$$

Note that as required, $r_{00}^{t} + r_{01}^{t} + r_{0f}^{t} = 1$.

The deterioration probabilities are given by the distributions of the initial point u and the delay time h. If the distributions of the initial point u and the delay time h can be estimated from the collected data, or from subjective techniques, the deterioration probabilities can be easily obtained. Various available methods for estimating the initial point and the delay time distribution has been discussed in the section 3.4 of the chapter 3.

4.3.5 Satisfying the Markovian Properties required of a Semi-Markov Inspection Model

The above model has been formulated as a semi-Markov inspection model and the fact that it is such a model has been assumed. Here we establish that it is, indeed, a semi-Markov model and satisfies the necessary condition for a semi-Markov decision process given as (1) and (2) in subsection 4.3.2.

The one-step transition probabilities, $P_{ij}(a)$, of subsection 4.3.3 have been expressed as a function of the deterioration probability, r_{ij}^{t}. Accordingly, using equation (4.44) with $t=\Delta t$, equation (4.10) becomes

$$P_{0y}(0)=\begin{cases}\int_0^{\Delta t} q(u)F(\Delta t-u)du & \text{for } y=f\\ 1-\int_0^{\Delta t} q(u)F(\Delta t-u)du & \text{for } y=(0,\Delta t)\\ 0 & \text{otherwise ,}\end{cases} \tag{4.45}$$

and using equation (4.44) with $t=m\Delta t$ and $t=(m+1)\Delta t$, equation (4.22) becomes

$$P_{(0,m\Delta t)y}(0)=\begin{cases}\dfrac{\int_0^{(m+1)\Delta t} q(u)F((m+1)\Delta t-u)du-\int_0^{m\Delta t} q(u)F(m\Delta t-u)du}{1-\int_0^{m\Delta t} q(u)F(m\Delta t-u)du} & \text{for } y=f\\ \dfrac{1-\int_0^{(m+1)\Delta t} q(u)F((m+1)\Delta t-u)du}{1-\int_0^{m\Delta t} q(u)F(m\Delta t-u)du} & \text{for } y=(0,(m+1)\Delta t)\\ 0 & \text{otherwise.}\end{cases} \tag{4.46}$$

Using equations (4.42), (4.43), and (4.44) with $t=m\Delta t$, equation (4.26) for the probability the component survives $m\Delta t$ time units to be inspected at $m\Delta t$ with result $y=0$ or $y=1$ (no defect or a defect present) becomes

$$P_{(0,m\Delta t)y}(1)=\begin{cases}\dfrac{1-\int_0^{m\Delta t} q(u)du}{1-\int_0^{m\Delta t} q(u)F(m\Delta t-u)du} & \text{for } y=0\\ \dfrac{\int_0^{m\Delta t} q(u)(1-F(m\Delta t-u))du}{1-\int_0^{m\Delta t} q(u)F(m\Delta t-u)du} & \text{for } y=1\\ 0 & \text{otherwise.}\end{cases} \tag{4.47}$$

The other one-step transition probabilities, $P_{10}(2)=1$ and $P_{f0}(2)=1$ (see equations (4.14) and (4.18)). Since equations (4.14), (4.18), (4.45), (4.46), and (4.47) are seen to be dependent only upon the current state i, $i \in I$, the next state j, $j \in I$, and the chosen action a, $a = 0, 1, 2$, we have that the one-step transition probabilities, $P_{ij}(a)$, used in the semi-Markov inspection model of subsection 4.3.3 satisfy the required condition (1) of Markovian property of subsection 4.3.2.

Also, the one-step expected costs, $C_i(a)$, and the one-step expected downtimes, $D_i(a)$, are given by a constant which is dependent only upon the current state i, $i \in I$, and the chosen action a, $a = 0, 1, 2$ (see equations (4.12), (4.13), (4.16), (4.17), (4.20), (4.21), (4.24), (4.25), (4.28), and (4.29)). Accordingly, the one-step expected costs and the one-step expected downtimes used in the above semi-Markov inspection model satisfy condition (2) of Markovian property of subsection 4.3.2.

Therefore, the semi-Markov inspection model established in subsection 4.3.3 satisfy all conditions of Markovian properties of the semi-Markov decision process.

4.3.6 Some Problems of the Semi-Markov Inspection Model

The above model is concerned with a typical semi-Markov inspection model which minimises the expected total cost per unit time or the expected total downtime per unit time. It presents the model formulation and solution computation process for the maintenance problem. However, the outstanding general problem is not so much in solving the model of the problem, as in structuring the assumptions and validating them.

The model assumes that the working condition of the system can be expressed in terms of a new state, a degraded state and a failed state, and further, that these states change according to a discrete-time Markov chain. Also, the model presumes that transition probabilities, or deterioration probabilities, are given or are calculated. Such assumptions need to be established as appropriate in any particular case. In practice, as previously indicated, it is difficult to define and measure the working condition of the deteriorated system as a degraded state, and difficult to measure directly transition probabilities or deterioration probabilities. In applying the model in real-world situations. the immediate problem is to analyse and interpret the collected data. Even though the working condition of the deteriorated system is expressed as one of a set of defined states, from data we have to establish a Markovian deterioration property for the system and estimate the transition probabilities, or deterioration probabilities. However. as evidenced by the lack of actual applications to problems and data of real-world situations, it is believed difficult to justify the Markov property assumptions, excepting exponentially distributed data, and therefore difficult to determine valid transition probability or deterioration probabilities. It is not felt to be reasonable to automatically assume always that actual problems satisfy the required Markov properties. There are, of course, corresponding problems for the assumption in delay time models. However, there are techniques and case experience indicating how they may be resolved.

4.4 Delay Time Model

4.4.1 Introduction to the Delay Time Model

Markov models have the advantage of an extensive body of theory. They could, however. be improperly applied within a maintenance context. In a series of papers dating from 1984, a robust approach to solve these problems has been introduced and developed as the Delay Time Model (DTM). It has been seen in the chapter 3 that the DTM is a powerful tool when applied to the modelling of actual industrial maintenance problems.

The inspection system presented in the section 4.2 can also be analysed in the framework of a DTM to find the optimal inspection period which minimises the expected total cost per unit time or the expected total downtime per unit time. In this section, we formulate the optimal inspection problem of the system as a delay time model. There are differences in the structure of the model in that the delay time model does not require a discrete time zone for analysis, and is assumed here (though not necessary) to have a fixed but unknown inspection period T.

4.4.2 Formulation as a General Delay Time Model

To formulate the inspection system mentioned in the section 4.2 as a delay time model, in adding to the system description of the section 4.2, we adopt the following assumptions for a single repairable component:

(1) A fault can be observed by inspection only, and a failure will be observed or repaired immediately if occurs.

(2) An inspection is undertaken every T time units, and the inspection process restarted after a failure repair.

(3) Inspections are perfect in that any fault present within the machine will be identified at inspection, and no new fault injected because of inspection.

(4) An inspection requires C_i cost units and d_i time units, $d_i << T$.

(5) A fault identified at an inspection will be repaired within the inspection period d_i and the repair cost per defect is C_d units.

(6) The component is repaired immediately upon failures and its repair requires C_b cost units and d_b time units.

(7) The machine is as good as new after a repair.

Then, as we noted in the section 4.2, there are two types of renewal points, that is, at a failure or when a fault is found at an inspection. assuming that a failure repair or fault rectification at an inspection may be regarded as renewing the component. Under these conditions, first of all, we formulate the general model without the specific specification of the initial point distribution or the delay time distribution. After that, a proof is given that the model reduces to a simpler form which also represents a semi-Markov inspection model. This being so, we are able to usefully compare both models.

Firstly, we consider the expected renewal cycle cost, *E(cycle cost)*. Noting that there is a renewal point after a fault rectification at an inspection, consider a general case that negative inspections arise at T, $2T$, ... $(k\text{-}1)T$ and a positive inspection arises at kT (see Figure 4.3).

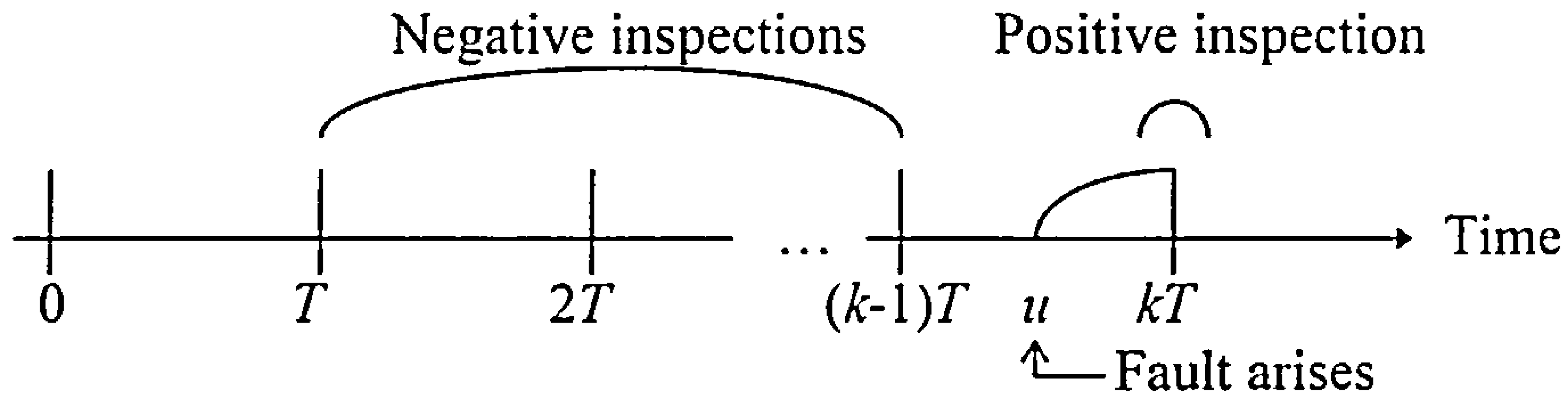


Figure 4.3. A case with a positive inspection at kT.

In this case, the total costs are k times the inspection costs plus fault repair costs at the inspection. The probability of this case arising, that is (k-1) negative inspections followed by a positive inspection at kT, $P_{k,i}$, is

$$P_{k,i} = \Pr((k-1)T < u < kT \text{ and } h > kT - u)$$
$$= \int_{(k-1)T}^{kT} q(u)(1 - F(kT - u))du, \qquad (4.48)$$

and the expected cost is given by

$$(kC_i + C_d)\int_{(k-1)T}^{kT} q(u)(1 - F(kT - u))du, \qquad k = 1, 2, 3, ... \qquad (4.49)$$

As before, C_i and C_d are inspection costs and fault repair costs at an inspection respectively, and $q(u)$ and $F(h)$ are the *pdf* of the initial point u and the *cdf* of the delay time h respectively. Summing over all possible case, we have that the expected cost up to a renewal point due to a positive inspection is given by

$$\sum_{k=1}^{\infty}(kC_i + C_d)\int_{(k-1)T}^{kT} q(u)(1 - F(kT - u))du. \qquad (4.50)$$

Again, in the case of a renewal point initiated by a failure repair, suppose that negative inspections arise at T, $2T$, ..., (k-1)T and a failure arises before kT (see Figure 4.4).

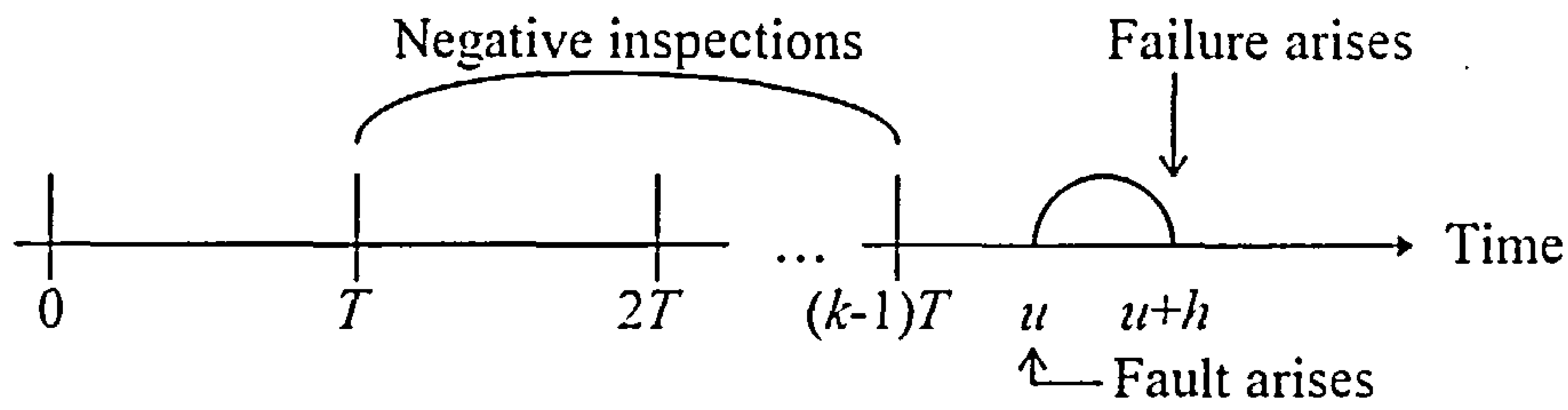


Figure 4.4. A case with a failure before kT.

In this case, since the required costs are $(k\text{-}1)$ times the inspection costs plus a failure repair cost, and the probability of this case arising, $P_{(k-1),f}$, is

$$P_{(k-1),f} = \Pr((k-1)T < u < kT \text{ and } h < kT - u)$$
$$= \int_{(k-1)T}^{kT} q(u)F(kT-u)du, \qquad (4.51)$$

it follows that the expected cost due to a failure in $((k\text{-}1)T, kT)$ is given by

$$((k-1)C_i + C_b)\int_{(k-1)T}^{kT} q(u)F(kT-u)du, \qquad k = 1, 2, 3, \ldots, \qquad (4.52)$$

where, as before, C_b is the failure repair costs. Summing over all possible values of k, the total expected cost to a renewal point caused by a failure repair is given by

$$\sum_{k=1}^{\infty}((k-1)C_i + C_b)\int_{(k-1)T}^{kT} q(u)F(kT-u)du. \qquad (4.53)$$

We have, therefore, that the expected renewal cycle cost, $E(\textit{cycle cost})$, is the sum of the expected costs resulting from inspection cycles which, from equations (4.50) and (4.53), is given by

$$E(\textit{cycle cost}) = \sum_{k=1}^{\infty}\{(kC_i + C_d)\int_{(k-1)T}^{kT} q(u)(1 - F(kT-u))du$$
$$+ ((k-1)C_i + C_b)\int_{(k-1)T}^{kT} q(u)F(kT-u)du\}$$

$$= \sum_{k=1}^{\infty} \{(kC_i + C_d) \int_{(k-1)T}^{kT} q(u)du + (C_b - C_i - C_d) \int_{(k-1)T}^{kT} q(u)F(kT-u)du\} . \quad (4.54)$$

Letting $u=(k-1)T+v$, we have that

$$E(cycle\ cost) = \sum_{k=1}^{\infty} \{(kC_i + C_d) \int_0^T q((k-1)T+v)dv$$

$$+ (C_b - C_i - C_d) \int_0^T q((k-1)T+v)F(T-v)dv\} . \quad (4.55)$$

Now, we consider the expected renewal cycle length, $E(cycle\ length)$, where

$$E(cycle\ length) = \sum_{\text{inspection cycle}} E(\text{length of inspection cycle}) \times \Pr(\text{Inspection of the cycle})$$

$$+ \sum_{\text{failure cycle}} E(\text{length of failure cycle}) \times \Pr(\text{failure of the cycle}) . \quad (4.56)$$

Similarly to the above, for a positive inspection renewal point, since the required length is of the form k times the inspection period plus downtime, and the probability of an inspection repair at kT is as given in equation (4.48), the expected length is given by

$$\sum_{k=1}^{\infty} k(T + d_i) \int_{(k-1)T}^{kT} q(u)(1 - F(kT-u))du , \quad (4.57)$$

where, as before, d_i is the downtimes for an inspection. To find the expected length to a renewal point of a failure repair, consider the random variable X, which is the time to failure from the last inspection (see Figure 4.5).

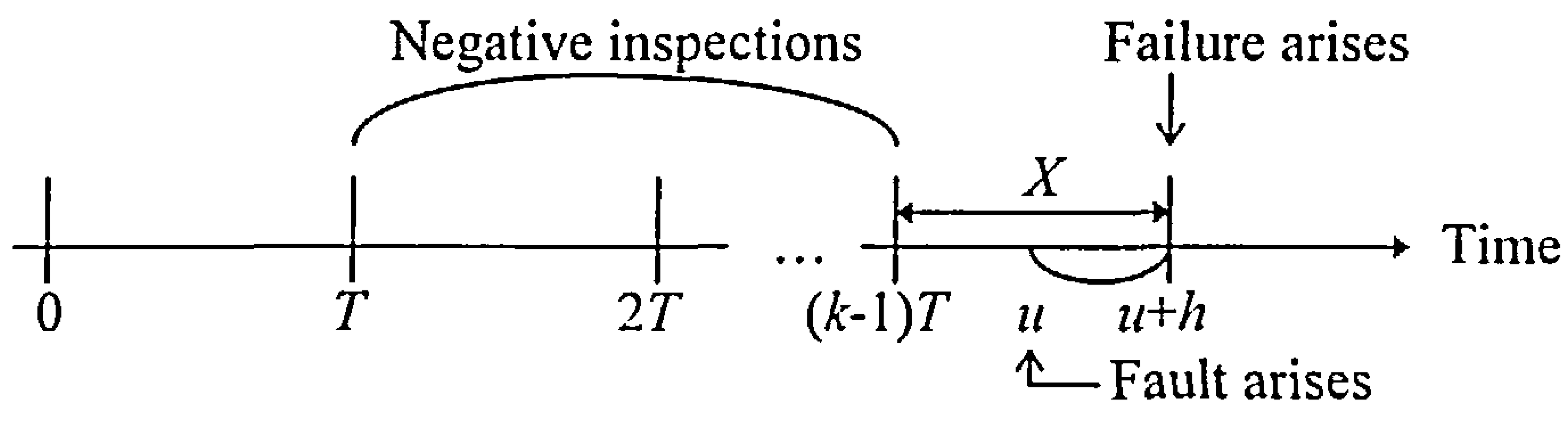


Figure 4.5. Definition of random variable X.

In this case, the required length is $(k\text{-}1)$ times the inspection period plus inspection downtime, the expected time to failure from the last inspection, $E(X)$, and the downtime for a failure repair. To formulate the expected time to failure from the last inspection, $E(X)$, if the probability density and the cumulative distribution function of X are $s_{k-1}(x)$ and $S_{k-1}(x)$ respectively, the *cdf* $S_{k-1}(x)$ is the convolution of the distribution of u and h given $u+h \le kT$ and $(k-1)T \le u \le kT$. So we have, because of perfect inspections, that the cumulative distribution of X for the interval $((k\text{-}1)T, kT)$ is given by

$$
\begin{aligned}
S_{k-1}(x) &= \Pr\{u+h \le (k-1)T+x \text{ and } (k-1)T \le u \le (k-1)T+x \\
&\qquad \text{given } u+h < kT \text{ and } (k-1)T \le u < kT\} \\
&= \frac{\Pr\{u+h \le (k-1)T+x \text{ and } (k-1)T \le u < (k-1)T+x\}}{\Pr\{u+h < kT \text{ and } (k-1)T \le u < kT\}} \\
&= \frac{\int_{(k-1)T}^{(k-1)T+x} q(u)F((k-1)T+x-u)du}{\int_{(k-1)T}^{kT} q(u)F(kT-u)du} \qquad \text{for } 0 \le x < T. \qquad (4.58)
\end{aligned}
$$

Letting $u=(k\text{-}1)T+v$ and applying the equation (4.51), the equation (4.58) becomes

$$
S_{k-1}(x) = \frac{1}{P_{(k-1),f}} \int_0^x q((k-1)T+v)F(x-v)dv \qquad \text{for } 0 \le x < T. \qquad (4.59)
$$

Also, the *pdf* of X is given by

$$
s_{k-1}(x) = \begin{cases} \dfrac{1}{P_{(k-1),f}} \displaystyle\int_0^x q((k-1)T+v)f(x-v)dv & \text{for } 0 \le x < T \\ 0 & \text{otherwise}, \end{cases} \qquad (4.60)
$$

where, as before, $f(h)$ is the *pdf* of the delay time h. The expected time to failure from the last inspection, $E(X)$ is, therefore, given by

$$E(X) = \int_0^T xs(x)dx$$

$$= \frac{1}{P_f} \int_0^T x \int_0^x q((k-1)T+v) f(x-v) dv dx. \qquad (4.61)$$

Since the probability of a renewal point initiated by a failure repair is given by equation (4.51), we have that the expected length to a renewal point caused by a failure repair is given by

$$\sum_{k=1}^{\infty} ((k-1)(T+d_i) + E(X) + d_b) \int_{(k-1)T}^{kT} q(u)F(kT-u)du, \qquad (4.62)$$

where, as before, d_b is the downtime for a failure repair. Summing over two cases, we have the expected renewal cycle length, *E*(*cycle length*), is given by

$$E(cycle\ length) = \sum_{k=1}^{\infty} \{k(T+d_i) \int_{(k-1)T}^{kT} q(u)(1-F(kT-u))du$$

$$+ ((k-1)(T+d_i) + E(X) + d_b) \int_{(k-1)T}^{kT} q(u)F(kT-u)du\}$$

$$= \sum_{k=1}^{\infty} \{k(T+d_i) \int_{(k-1)T}^{kT} q(u)du$$

$$+ (E(X) + d_b - T - d_i) \int_{(k-1)T}^{kT} q(u)F(kT-u)du\}. \qquad (4.63)$$

Letting $u=(k-1)T+v$, we have that

$$E(cycle\ length) = \sum_{k=1}^{\infty} \{k(T+d_i) \int_0^T q((k-1)T+v)dv$$

$$+ (E(X) + d_b - T - d_i) \int_0^T q((k-1)T+v)F(T-v)dv\}$$

$$= \sum_{k=1}^{\infty} \{k(T+d_i) \int_0^T q((k-1)T+v)dv + \int_0^T x \int_0^x q((k-1)T+v) f(x-v) dv dx$$

$$+ (d_b - T - d_i) \int_0^T q((k-1)T+v)F(T-v)dv\}. \qquad (4.64)$$

Thirdly, for a downtime model, we consider the expected renewal cycle downtime, *E(cycle downtime)*. Similarly, for a renewal point of a positive inspection, since the downtime is kd_i and the probability of this case arising is $P_{k,i}$, we have the expected downtime is given by

$$\sum_{k=1}^{\infty} kd_i \int_{(k-1)T}^{kT} q(u)(1-F(kT-u))du. \tag{4.65}$$

Also, for a renewal point of a failure repair, since the downtime is $(k-1)d_i + d_b$ and the probability of this case arising is $P_{(k-1),f}$, the expected downtime is

$$\sum_{k=1}^{\infty} ((k-1)d_i + d_b) \int_{(k-1)T}^{kT} q(u)F(kT-u)du. \tag{4.66}$$

Summing over two cases, we have the expected renewal cycle downtime, *E(cycle downtime)*, given by

$$\begin{aligned} E(cycle\ downtime) &= \sum_{k=1}^{\infty} \{kd_i \int_{(k-1)T}^{kT} q(u)(1-F(kT-u))du \\ &\quad + ((k-1)d_i + d_b) \int_{(k-1)T}^{kT} q(u)F(kT-u)du\} \\ &= \sum_{k=1}^{\infty} \{kd_i \int_{(k-1)T}^{kT} q(u)du + (d_b - d_i) \int_{(k-1)T}^{kT} q(u)F(kT-u)du\}. \end{aligned} \tag{4.67}$$

Letting $u=(k-1)T+v$, we have that

$$\begin{aligned} E(cycle\ downtime) &= \sum_{k=1}^{\infty} \{kd_i \int_0^T q((k-1)T+v)dv \\ &\quad + (d_b - d_i) \int_0^T q((k-1)T+v)F(T-v)dv\}. \end{aligned} \tag{4.68}$$

We can, therefore, have that, from the equations (4.55) and (4.64), the steady state expected cost per unit time $C(T)$ is

$$C(T) = \frac{E(cycle\ cost)}{E(cycle\ length)} \qquad (4.69)$$

and. from the equations (4.64) and (4.68), the steady state expected downtime per unit time $D(T)$ is

$$D(T) = \frac{E(cycle\ downtime)}{E(cycle\ length)} \ . \qquad (4.70)$$

From the above equations (4.69) or (4.70), we can determine the optimal inspection period T^* which minimises the long run expected cost per unit time or the long run expected downtime per unit time.

4.4.3 A Simpler Delay Time Model

In order to compare the delay time model with the semi-Markov inspection model, the general delay time models of the subsection 4.4.2 must be reduced to a simpler and more restricted form. For example, the initial time u is now assumed to be exponentially distributed with a mean given by $1/\lambda$ to satisfy the Markov property. If the initial point u has an exponential distribution with a mean given by $1/\lambda$,

$$q(u) = \lambda e^{-\lambda u} \ . \qquad (4.71)$$

and from equations (4.55), (4.64) and (4.68), we have that

$$\int_0^T q((k-1)T + v)dv = e^{-\lambda(k-1)T} \int_0^T q(v)dv \qquad (4.72)$$

and

$$\int_0^T q((k-1)T+v)F(T-v)dv = e^{-\lambda(k-1)T}\int_0^T q(v)F(T-v)dv. \tag{4.73}$$

So, it follows that the equation (4.55) for the expected cycle cost becomes

$$\begin{aligned} E(cycle\ cost) &= \sum_{k=1}^{\infty}\{(kC_i + C_d)e^{-\lambda(k-1)T}\int_0^T q(v)dv \\ &\quad + (C_b - C_i - C_d)e^{-\lambda(k-1)T}\int_0^T q(v)F(T-v)dv\} \\ &= C_i\int_0^T q(v)dv\sum_{k=1}^{\infty} ke^{-\lambda(k-1)T} + C_d\int_0^T q(v)dv\sum_{k=1}^{\infty} e^{-\lambda(k-1)T} \\ &\quad + (C_b - C_i - C_d)\int_0^T q(v)F(T-v)dv\sum_{k=1}^{\infty} e^{-\lambda(k-1)T}. \end{aligned} \tag{4.74}$$

Also, we have that

$$\sum_{k=1}^{\infty} e^{-\lambda(k-1)T} = \frac{1}{1-e^{-\lambda T}}, \tag{4.75}$$

$$\begin{aligned} \sum_{k=1}^{\infty} ke^{-\lambda(k-1)T} &= (\sum_{k=1}^{\infty} e^{-\lambda(k-1)T})(\sum_{k=1}^{\infty} e^{-\lambda(k-1)T}) \\ &= \frac{1}{(1-e^{-\lambda T})^2}, \end{aligned} \tag{4.76}$$

and

$$\int_0^T q(v)dv = 1 - e^{-\lambda T}. \tag{4.77}$$

Re-arranging the equation (4.74), we finally have that

$$E(cycle\ cost) = \frac{C_i + C_d\int_0^T q(v)dv + (C_b - C_i - C_d)\int_0^T q(v)F(T-v)dv}{1-e^{-\lambda T}}. \tag{4.78}$$

Also, in the equation (4.64), since

$$\int_0^T x \int_0^x q((k-1)T+v) f(x-v)\,dv\,dx = e^{-\lambda(k-1)T} \int_0^T x \int_0^x q(v) f(x-v)\,dv\,dx\,, \tag{4.79}$$

using the equations (4.72), (4.73), (4.75), and (4.76), it follows that the expected cycle length equation (4.64) becomes

$$\begin{aligned} E(cycle\ length) &= (T+d_i)\int_0^T q(v)\,dv \sum_{k=1}^{\infty} k e^{-\lambda(k-1)T} \\ &\quad + \int_0^T x \int_0^x q(v) f(x-v)\,dv\,dx \sum_{k=1}^{\infty} e^{-\lambda(k-1)T} \\ &\quad + (d_b - T - d_i)\int_0^T q(v) F(T-v)\,dv \sum_{k=1}^{\infty} e^{-\lambda(k-1)T} \\ &= \frac{T + d_i + \int_0^T x \int_0^x q(v) f(x-v)\,dv\,dx + (d_b - T - d_i)\int_0^T q(v) F(T-v)\,dv}{1-e^{-\lambda T}}. \end{aligned} \tag{4.80}$$

In the same manner, the expected renewal cycle downtime, *E*(*cycle downtime*), is given by

$$E(cycle\ downtime) = \frac{d_i + (d_b - d_i)\int_0^T q(v) F(T-v)\,dv}{1-e^{-\lambda T}}. \tag{4.81}$$

We have, therefore, from the equations (4.78) and (4.80), that the expected cost per unit time $C(T)$ is given by

$$C(T) = \frac{C_i + C_d \int_0^T q(v)\,dv + (C_b - C_i - C_d)\int_0^T q(v) F(T-v)\,dv}{T + d_i + \int_0^T x \int_0^x q(v) f(x-v)\,dv\,dx + (d_b - T - d_i)\int_0^T q(v) F(T-v)\,dv} \tag{4.82}$$

and, from the equations (4.80) and (4.81), the expected downtime per unit time $D(T)$ is given by

$$D(T)=\frac{d_t+(d_b-d_i)\int_0^T q(v)F(T-v)dv}{T+d_i+\int_0^T x\int_0^x q(v)f(x-v)dvdx+(d_b-T-d_i)\int_0^T q(v)F(T-v)dv}. \quad (4.83)$$

From the above equations (4.82) and (4.83), we can obtain the optimal inspection period T^* which minimizes the expected cost per unit time or the expected downtime per unit time respectively. In the section 4.5, using the equations (4.82) or (4.83), the numerical examples will be given to compare with the semi-Markov inspection model.

4.4.4 Evaluation of the Delay Time Model

Two basic models of the inspection process have been formulated, the delay time model and the semi-Markov model. A semi-Markov inspection model provides a means of modelling the effect of inspections on the failure rate, operating cost and downtime of repairable machinery. It is necessary to express the working condition of the system as states of a Markov chain, and establish from data or otherwise, the Markov property and associated transition probabilities. This means that this model can only be used in applied situation when the initial point of a fault has a negative exponential distribution. The advantages of a semi-Markov model include a well established theory, and a mechanism of investigating formally the structure of an optimal solution. For example, we have assumed here a constant inspection period, and thereby constrained the search space for an optimal policy. This constraint is pragmatic. since the solution needs to be workable. A delay time based investigation of non-uniform inspection strategies is possible. However, the semi-Markov theory provides a well established (though computationally demanding) means of investigation here.

In contrast to the semi-Markov inspection model, the delay time model has only a few restrictions. The distribution functions $q(u)$ and $f(h)$ and their parameters are vital to delay time modelling. Once these two distributions are known. it is possible to model the operating cost and availability functions for any inspection practice of interest. Two methods for estimating these two distributions have been developed, namely the subjective method and the objective method, which have been described in the chapter 3. Also, this model may have any distribution for u and h regardless of a Markov property and Markov transition probability. This generality of the delay time model and its basic simplicity are very useful when attempting to apply it to real-world situations.

4.5 Numerical Examples

4.5.1 Generating the Data using Simulation

When we are faced with the preventive maintenance problem of an industrial plant, the data collection is of prime importance. In attempting to apply models of maintenance in a real-world situation, the immediate problem is often not that the model does not fit the data, but that there are no data to be fitted. In such data-starved situations, one wishes to initiate collection of any available data as quickly as possible. Much thought needs to be given to data, that is, its quality, its cost, its acquisition and its use in modelling. It is moreover difficult to get the ideal data for model fitting in real-world situations. In spite of this, we need the ideal data to be fitted to the model of section 4.3 and 4.4 if we are to compare two models. It is reasonable here to generate the data using computer simulation.

From the system description for modelling of the section 4.2, we note that there are two random variables, that is, the initial point u and the delay time h. In order to carry out a simulation of a system having the initial point u and the delay time h we have to specify the probability distributions of these random variables. Then, given that these random variables follow particular distributions, the simulation proceeds by generating values of these random variables from the appropriate distribution. In this section, we select the exponential and the Weibull distributions as the distributions for these random variables for a numerical example.

If the distributions of the initial point u and the delay time h are specified, we address the issue of how we can generate random variables with these distributions in order to simulate the process. The basic ingredient needed for the method of generating random variables is a source of independent identically distributed uniform random variables with (0,1), $U(0,1)$. For this reason, it is very important that a statistically reliable $U(0,1)$ random-number generator is available. Most computer installations have a convenient random-number generator. We therefore use this computer random-number.

Using this random-number we can generate the random variables of the initial point u and the delay time h. If the exponential random variable with the mean $1/\lambda$, $\lambda>0$, is considered for an initial point u or a delay time h, we can derive the following algorithm for a exponential random variable X.

1. Generate $U \sim U(0,1)$.
2. Set $X = -\frac{1}{\lambda}\log U$ and return to the main program.

If the Weibull distribution with the shape parameter $\alpha>0$ and the scale parameter $\beta>0$ is selected, the Weibull random variable X is generated by the following algorithm.

1. Generate $U \sim U(0,1)$.
2. Set $X = \beta(-\log U)^{\frac{1}{\alpha}}$ and return to the main program.

Once the random variables of the initial point u and the delay time h are generated, the simulation can be built under the system description for modelling of the section 4.2. Figure 4.6 illustrates the simulation progress by presenting the flow chart. In Figure 4.6, as before, T is a present inspection period, d_i is a downtime for an inspection and d_b is a downtime for a failure and N is a pre-specified number of renewal points for a simulation run and RP is a present renewal point time.

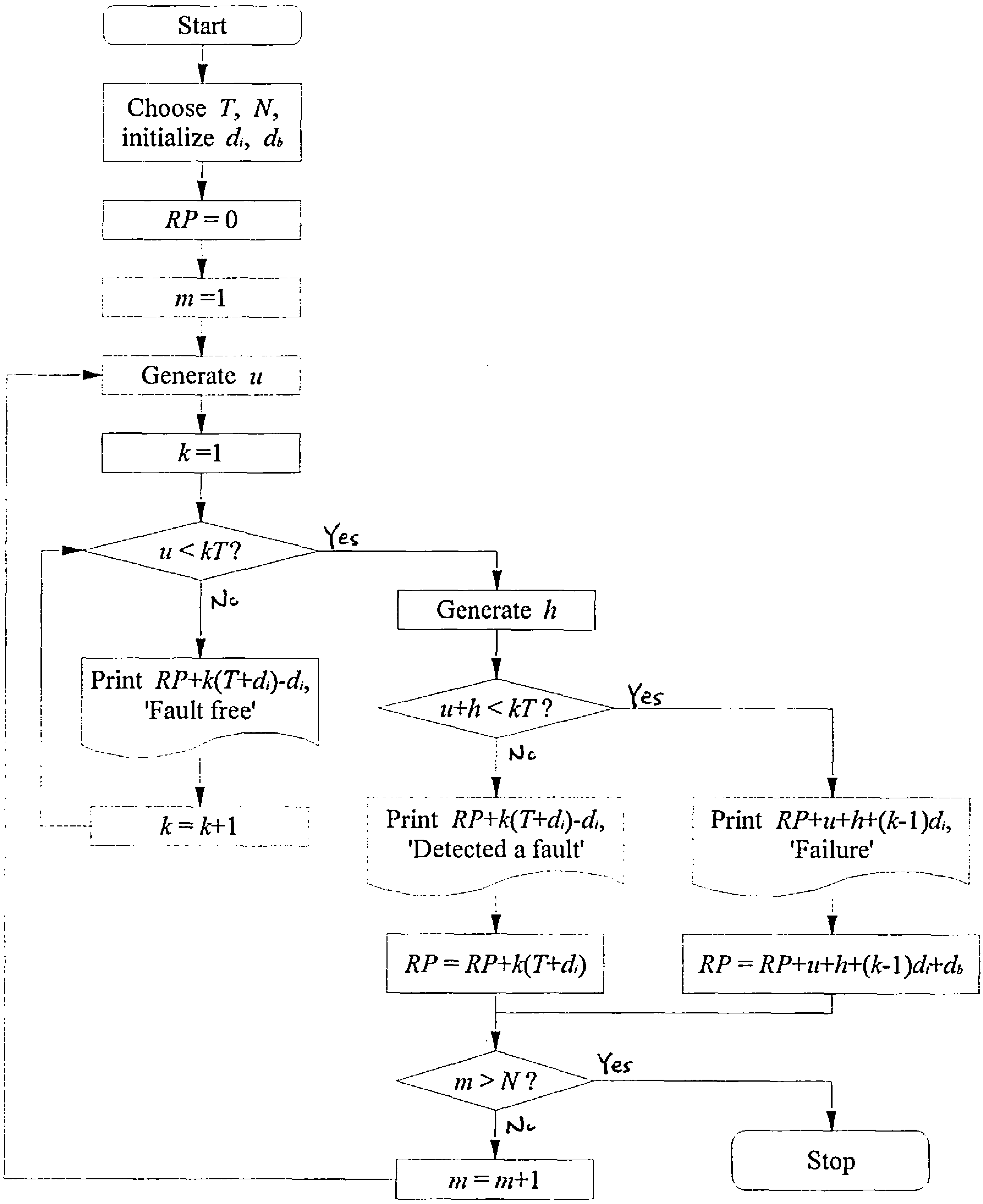


Figure 4.6. Flow chart for generating failure and inspection data for a single component system.

Including these downtimes, d_i and d_b, we can obtain the expected downtime per unit time $D(T)$. Also, If the inspection costs C_i, the inspection repair costs per fault C_d and the failure repair costs C_b are included, we can derive the expected cost per unit time $C(T)$. We will discuss the method of getting $C(T)$ and $D(T)$ using the simulation in the next subsection.

4.5.2 A Markovian Case of the Fault Arrival Rate

Numerical example of the cost model outlined in the section 4.3 and 4.4 has been evaluated for demonstration purposes. First of all, consider the case that the fault arrival rate has a Markovian property. Assuming that the fault arrival rate has been taken as λ faults per unit time, it follows that the *pdf* of the initial point u is given by

$$q(u) = \lambda e^{-\lambda u}. \tag{4.84}$$

Firstly, consider the case that the *pdf* of the delay time h has been chosen as an exponential distribution with the mean $1/\alpha$, $\alpha>0$,

$$f(h) = \alpha e^{-\alpha h}. \tag{4.85}$$

Assuming that the data have been given in the real-world situation or using simulation, λ and α have been estimated as $\lambda = 0.1$ and $\alpha = 0.05$. Also, costs are taken by $C_i = 10$ units, $C_d = 2$ units and $C_b = 100$ units and downtimes for a cost model are taken by $d_i = 0.08$ time units, $d_b = 0.18$ time units.

Then, from the semi-Markov inspection model of the section 4.2, the expected average cost per unit time $g_c(R)$ can be given as a function of the inspection period T. That is, according to the change of the inspection period T, we can obtain the expected average cost per unit time $g_c(R)$. Also, from the equation (4.82) of the simpler delay time model of the subsection 4.4.3, the expected cost per unit time $C(T)$ can be obtained. To compare the two models fairly, a simulation model can be used. It was shown in the flow chart of the Figure 4.6 of the previous subsection 4.5.1 that for a given inspection period T the required data can be generated. Changing slightly the flow chart of the Figure 4.6 into the flow chart of the Figure 4.7, we can get the expected cost per unit time $C(T)$ or the expected downtime per unit time $D(T)$ according to the inspection period T. In the Figure 4.7, TC is the total cost and TD is the total downtime.

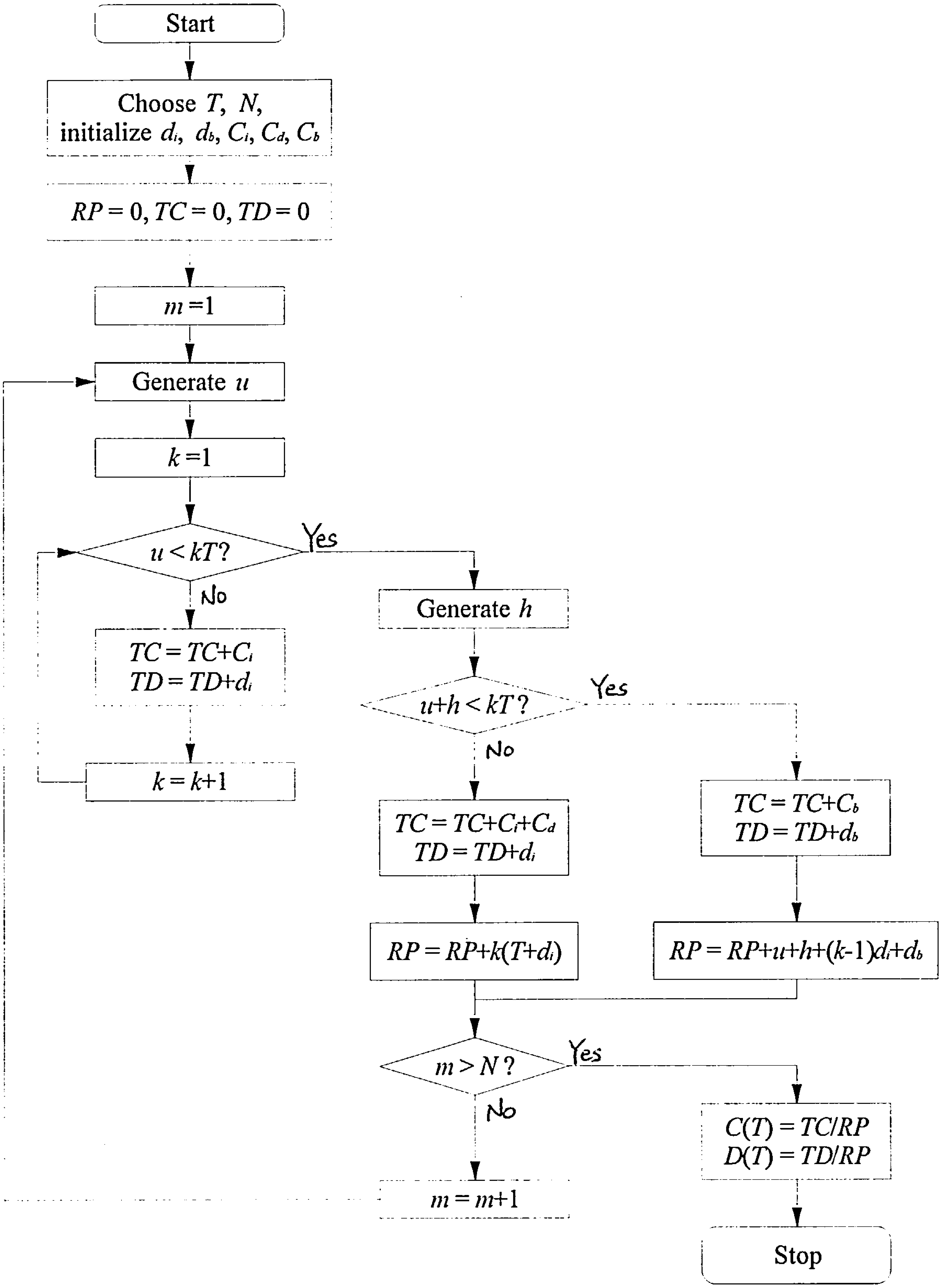


Figure 4.7. Flow chart of simulation model for generating the cost and downtime for a single component system.

The results for these models, in terms of the expected cost, are shown in Figures 4.8, 4.9, and 4.10. which represent the outcome of different small time intervals Δt of the semi-Markov inspection model.

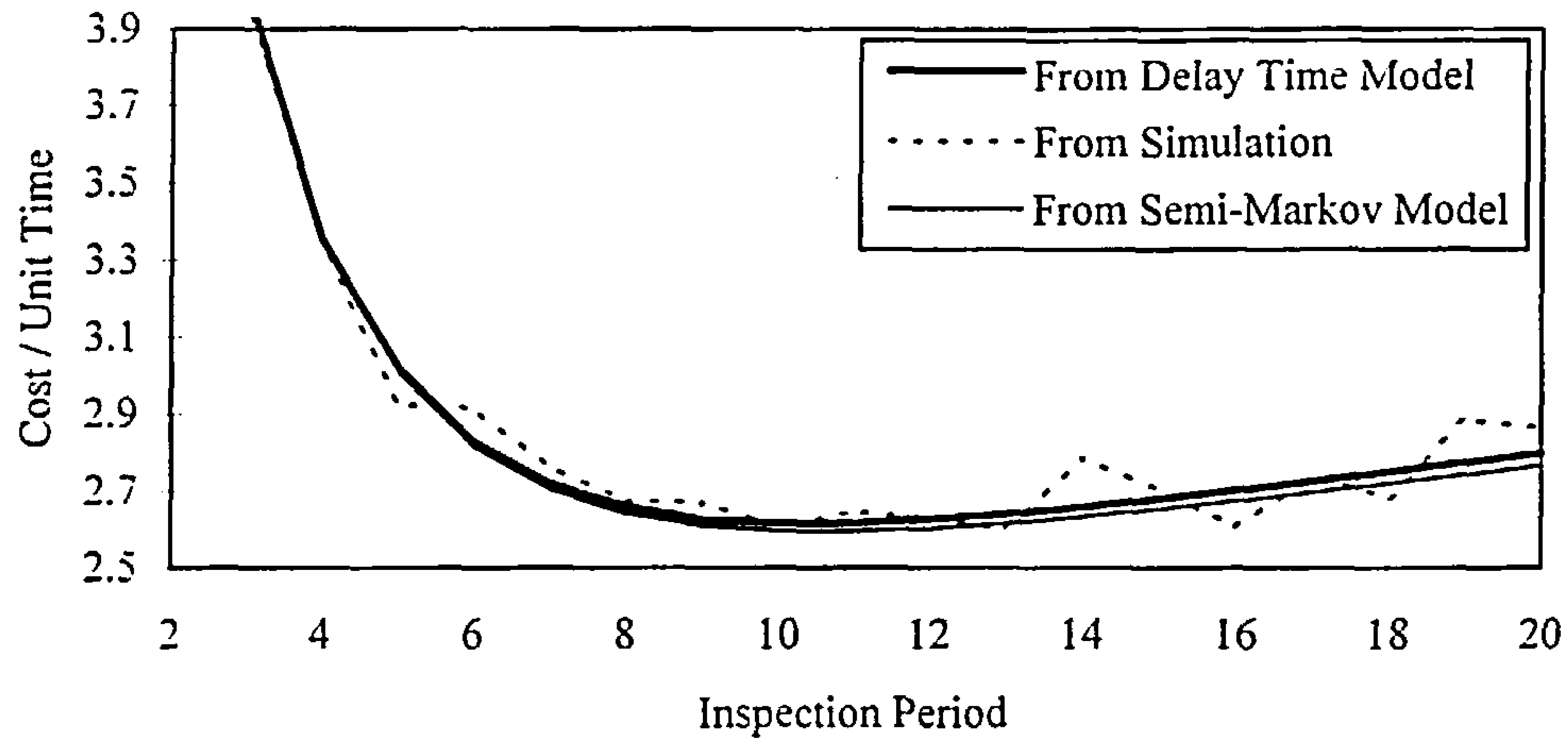


Figure 4.8. The expected cost per unit time according to the inspection period.

(This is a Markovian case of the fault arrival rate when Δt = 1, the delay time has an exponential distribution. λ=0.1, α=0.05, C_i =10, C_d =2, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

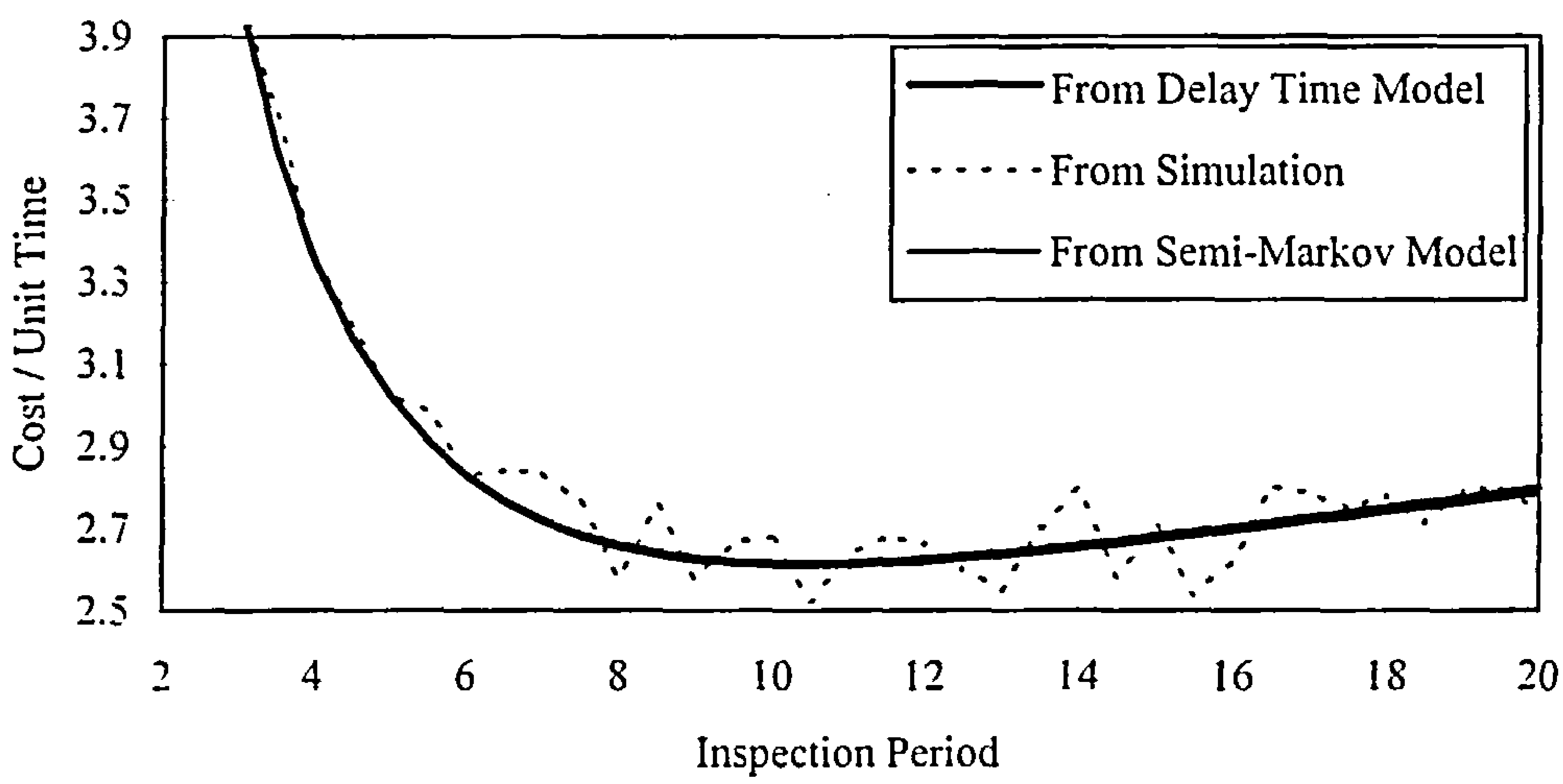


Figure 4.9. The expected cost per unit time according to the inspection period.

(This is a Markovian case of the fault arrival rate when Δt = 0.5, the delay time has an exponential distribution. λ=0.1. α=0.05, C_i =10, C_d =2, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

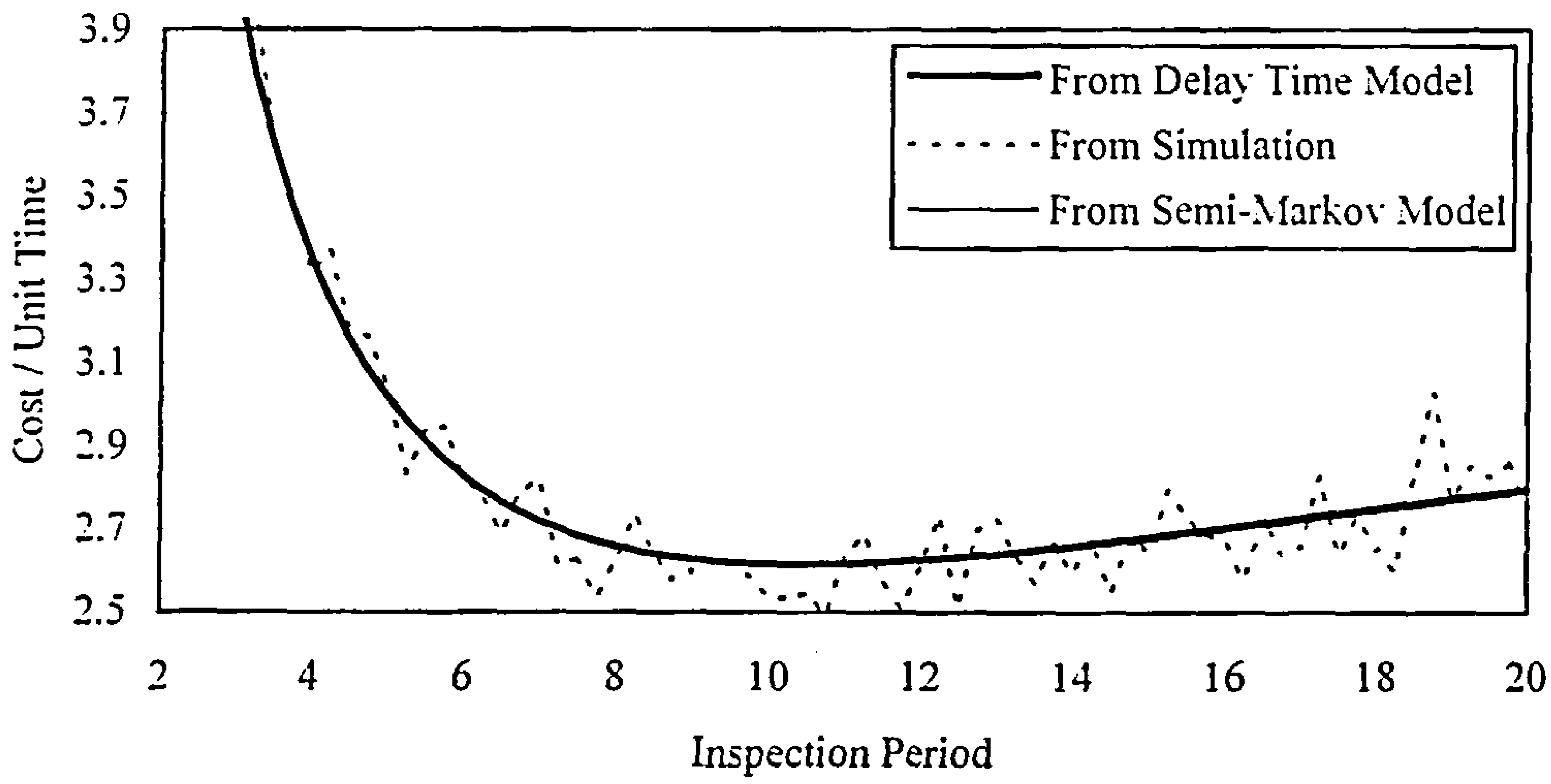


Figure 4.10. The expected cost per unit time according to the inspection period.
(This is a Markovian case of the fault arrival rate when Δt = 0.25, the delay time has an exponential distribution, λ=0.1, α=0.05, C_i =10, C_d =2, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

From Figure 4.8, there is a little difference between the delay time model curve and the semi-Markov model curve. The reason is that the delay time model is continuous and the Markov model is discrete. In establishing the semi-Markov inspection model, we have approximately calculated the average cost per unit time, $g_c(R)$, or the average downtime per unit time, $g_d(R)$, on condition that the time interval Δt is very small. Accordingly, the accuracy of the average cost per unit time, $g_c(R)$, or the average downtime per unit time, $g_d(R)$, is dependent upon the time interval Δt. As expected, we can demonstrate that as the time interval Δt decreases the semi-Markov model curve approaches to the delay time model curve (see Figure 4.8, 4.9 and 4.10). Also, the Figures show that the simulation curve is consistent with the delay time model curve and the semi-Markov model curve.

Changing the scale of the cost per unit time and removing the simulation curve of Figure 4.10, we can compare the semi-Markov model curve with the delay time model curve in detail. The result is shown in Figure 4.11.

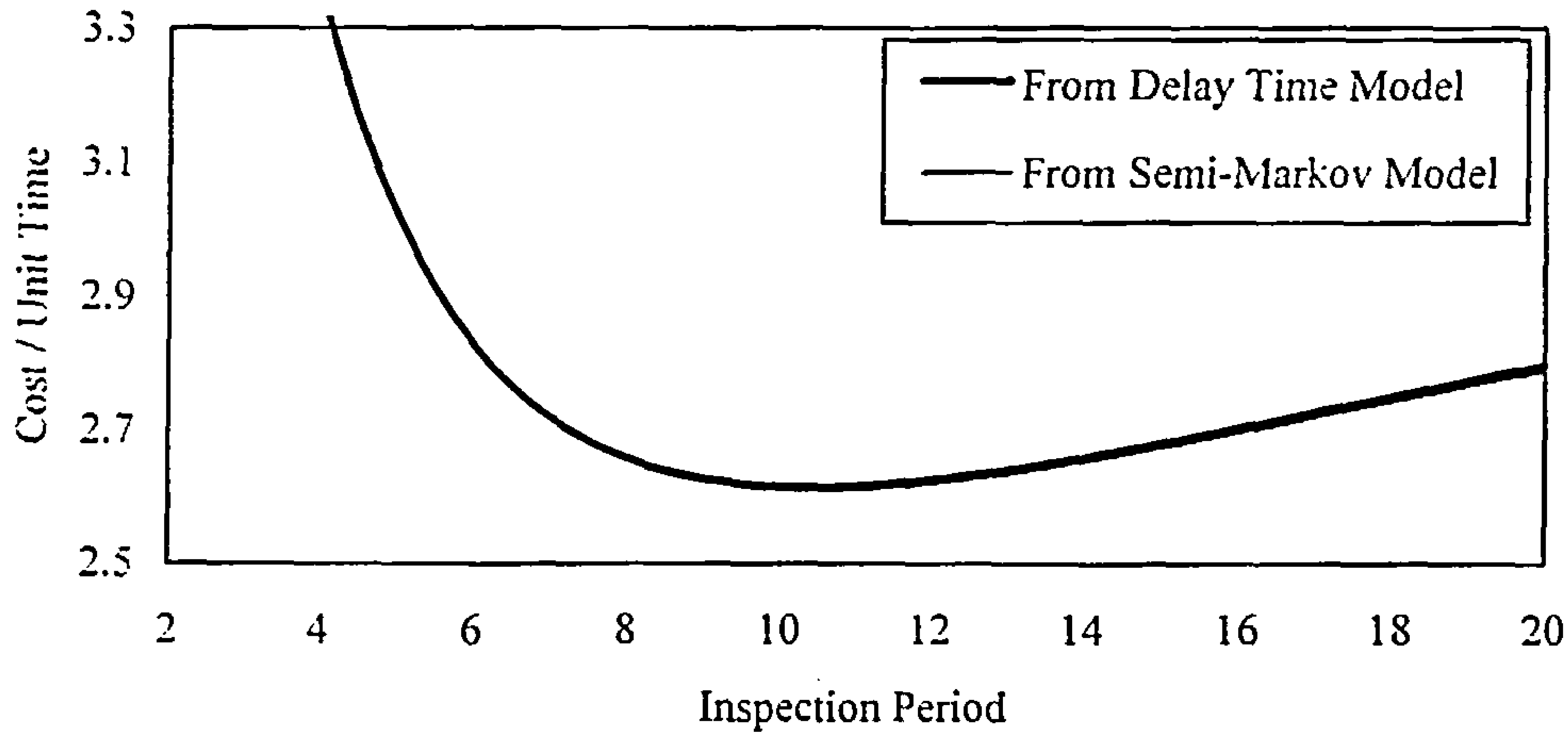


Figure 4.11. The detailed expected cost per unit time according to the inspection period.
(This is a Markovian case of the fault arrival rate when $\Delta t = 0.25$, the delay time has an exponential distribution, λ=0.1, α=0.05, $C_i = 10$, $C_d = 2$, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

Figure 4.11 shows that the semi-Markov model curve is still consistent with the delay time model curve and from the semi-Markov model curve and the delay time model curve we can obtain an optimal inspection period point which minimise the cost per unit time.

Secondly, assuming that the fault arrival rate has been taken as λ faults per unit time, consider a case where the *pdf* of the delay time h has been chosen as a Weibull distribution with the shape parameter α>0 and the scale parameter β>0,

$$f(h) = \alpha\beta^{-\alpha}h^{\alpha-1}e^{-(\frac{h}{\beta})^{\alpha}}. \tag{4.86}$$

We assume that λ, α and β have been estimated as λ=0.1, α=1.5 and β=1. Also, costs are taken by $C_i = 10$ units, $C_d = 2$ units and $C_b = 100$ units and downtimes for a cost model are taken by $d_i = 0.08$ time units, $d_b = 0.18$ time units. Then, similarly, we can get the expected cost per unit time $C(T)$ according to the inspection period T from the semi-Markov inspection model, the delay time model and from the simulation. Although this case is non-Markovian assuming Weibull distribution, transition probabilities may still be calculated using equations (4.45), (4.46), and (4.47) and a "Semi-Markov Type" model evaluated.

The results of such a process are shown in Figure 4.12, 4.13 and 4.14, which differ according to the time interval Δt of the semi-Markov inspection model.

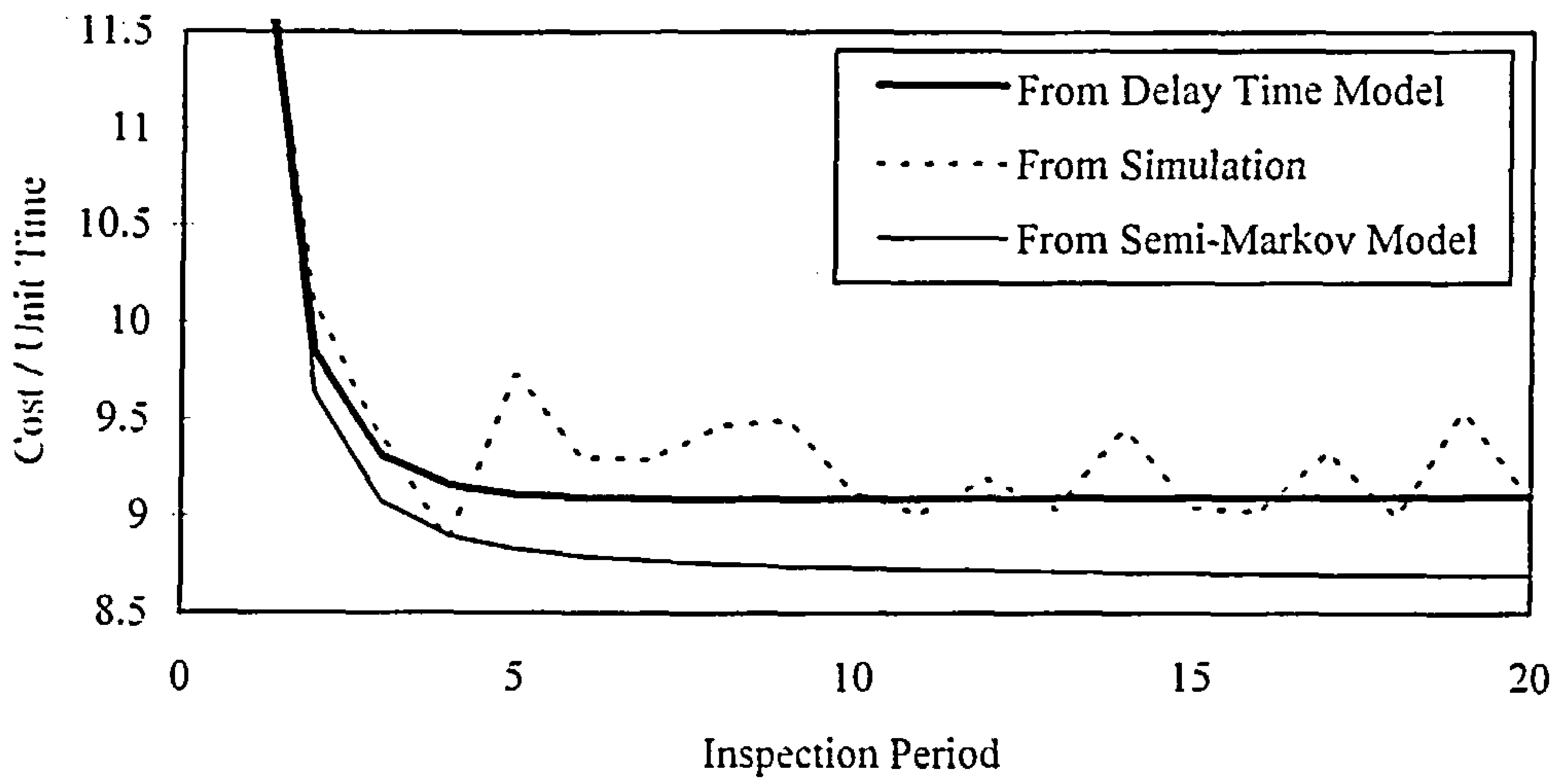


Figure 4.12. The expected cost per unit time according to the inspection period.

(This is a Markovian case of the fault arrival rate when $\Delta t = 1$, the delay time has a Weibull distribution. λ=0.1, α=1.5, β=1, C_i =10. C_d =2, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

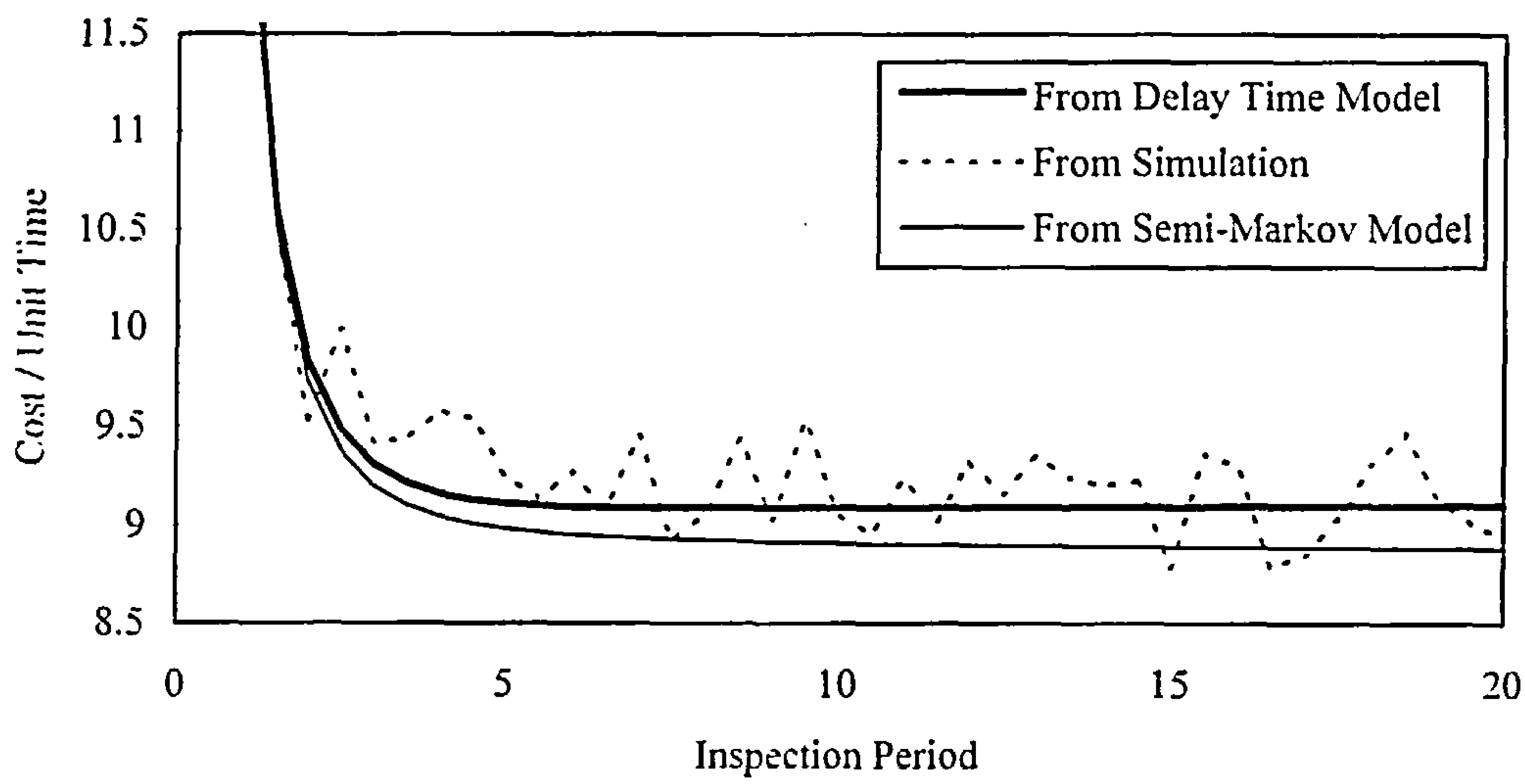


Figure 4.13. The expected cost per unit time according to the inspection period.

(This is a Markovian case of the fault arrival rate when $\Delta t = 0.5$, the delay time has a Weibull distribution. λ=0.1, α=1.5, β=1, C_i =10. C_d =2. $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

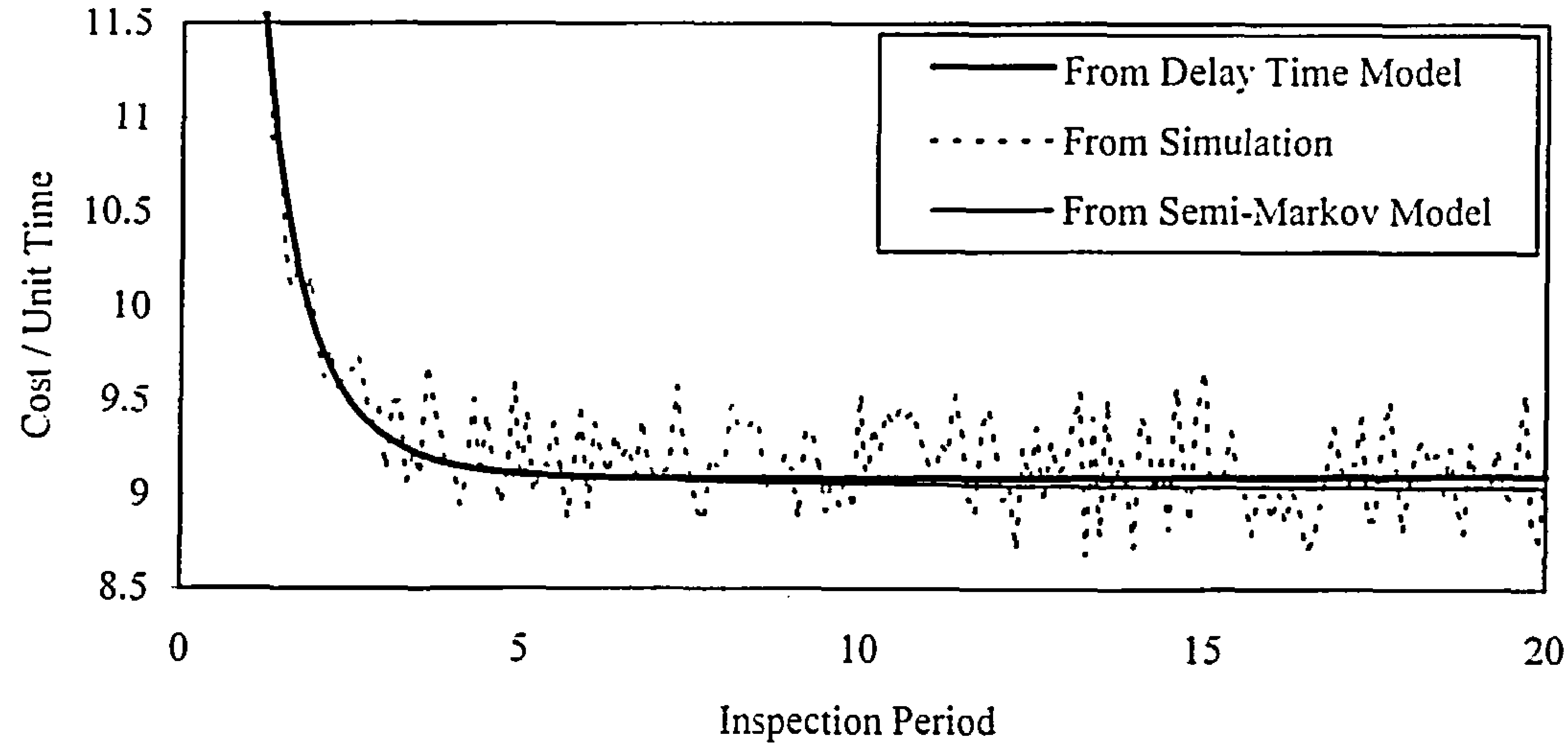


Figure 4.14. The expected cost per unit time according to the inspection period.
(This is a Markovian case of the fault arrival rate when $\Delta t = 0.1$, the delay time has a Weibull distribution, λ=0.1. α=1.5, β=1, $C_i = 10$, $C_d = 2$, $C_b = 100$, $d_i = 0.08$, and $d_b = 0.18$.)

Through Figure 4.12, 4.13 and 4.14, we can demonstrate that as the time interval Δt decreases the semi-Markov model curve is getting consistent with the delay time model curve. This means that the delay time formulae and the semi-Markov formulae for the same inspection problem lead to the same results as expected. Difference between the models is to be formed in the complications of the formulation and the extent of the numerical work required, where the delay time formulation is much simpler in this case. It is arguable that if transition probabilities for the semi-Markov inspection model had to be estimated from actual data via a delay time formulation as here, the semi-Markov inspection model would be even more demanding than the delay time model.

4.5.3 A Non-Markovian Case of the Fault Arrival Rate

In the previous subsection, a case when the fault arrival rate has a Markov property has been discussed. As a consequence of the previous subsection, we found that when the fault arrival rate has a Markov property the system can be fitted to the semi-Markov inspection model and the delay time model. If the fault arrival rate does not have a

Markov property, the system should not be fitted to the semi-Markov inspection model. However, as already stated the delay time model can be fitted to any system regardless of the Markov property of the fault arrival rate. To make this point clearly, we now consider a case when the fault arrival rate also has a non-Markov property.

Assuming that the initial point u has a Weibull distribution, i.e., a case when the fault arrival rate has a non-Markov property, it follows that the *pdf* of the initial point u is given by

$$q(u) = \alpha\beta^{-\alpha}u^{\alpha-1}e^{-(\frac{u}{\beta})^{\alpha}}, \tag{4.87}$$

where α, α>0, is the shape parameter and β, β>0, is the scale parameter. Also, assuming that the delay time h has a Weibull distribution, we have that the *pdf* of the delay time h is given by

$$f(h) = \gamma\delta^{-\gamma}h^{\gamma-1}e^{-(\frac{h}{\delta})^{\gamma}} \tag{4.88}$$

and the *cdf* of the delay time h is given by

$$F(h) = 1 - e^{-(\frac{h}{\delta})^{\gamma}}, \tag{4.89}$$

where γ, γ>0, is the shape parameter and δ, δ>0, is the scale parameter. Applying the equations (4.87), (4.88) and (4.89) to equations (4.55) and (4.64), we obtain after some algebra the expected cost per unit time formulations. Rearranging the equations (4.55) and (4.64), we have that

$$E(cycle\ cost) = C_i\sum_{k=1}^{\infty}k\int_0^T q((k-1)T+v)dv + C_d\sum_{k=1}^{\infty}\int_0^T q((k-1)T+v)dv$$

$$+(C_b - C_i - C_d)\sum_{k=1}^{\infty}\int_0^T q((k-1)T+v)F(T-v)dv\,. \tag{4.90}$$

and

$$E(cycle\ length) = (T + d_i)\sum_{k=1}^{\infty} k \int_0^T q((k-1)T + v)dv$$
$$+ \sum_{k=1}^{\infty} \int_0^T x \int_0^x q((k-1)T + v) f(x - v) dv dx$$
$$+ (d_b - T - d_i)\sum_{k=1}^{\infty} \int_0^T q((k-1)T + v) F(T - v) dv. \qquad (4.91)$$

In the equations (4.90) and (4.91), since the terms, $k\int_0^T q((k-1)T+v)dv$, $\int_0^T q((k-1)T+v)dv$, $\int_0^T q((k-1)T+v)F(T-v)dv$ and $\int_0^T x\int_0^x q((k-1)T+v)f(x-v)dvdx$, decrease as k increase, we can neglect the terms when they are less than the constant ε which does not affect the total value.

Given the data from a real-world situation or the data simulated data, we assume that the parameters for the distributions of the initial point u and the delay time h have been estimated as α=2, β=1, γ=1.5 and δ=1.3. Also, costs are taken as C_i =10 units. C_d =2 units and $C_b = 5$ units and downtimes for a cost model are taken by $d_i = 0.08$ time units, $d_b = 0.18$ time units. Then, taking $\varepsilon = 10^{-4}$, we can numerically get the expected cost per unit time $C(T)$ according to the inspection period T from the delay time model, the semi-Markov inspection model and the simulation.

The results for these models are shown in Figures 4.15, 4.16 and 4.17, which are changed according to the time interval Δt of the semi-Markov inspection model. From Figures 4.15, 4.16 and 4.17, we can see that even though the time interval Δt decreases the semi-Markov model curve is not consistent with the delay time model curve and the simulation curve. On the other hand, the delay time model curve is consistent with the simulation curve.

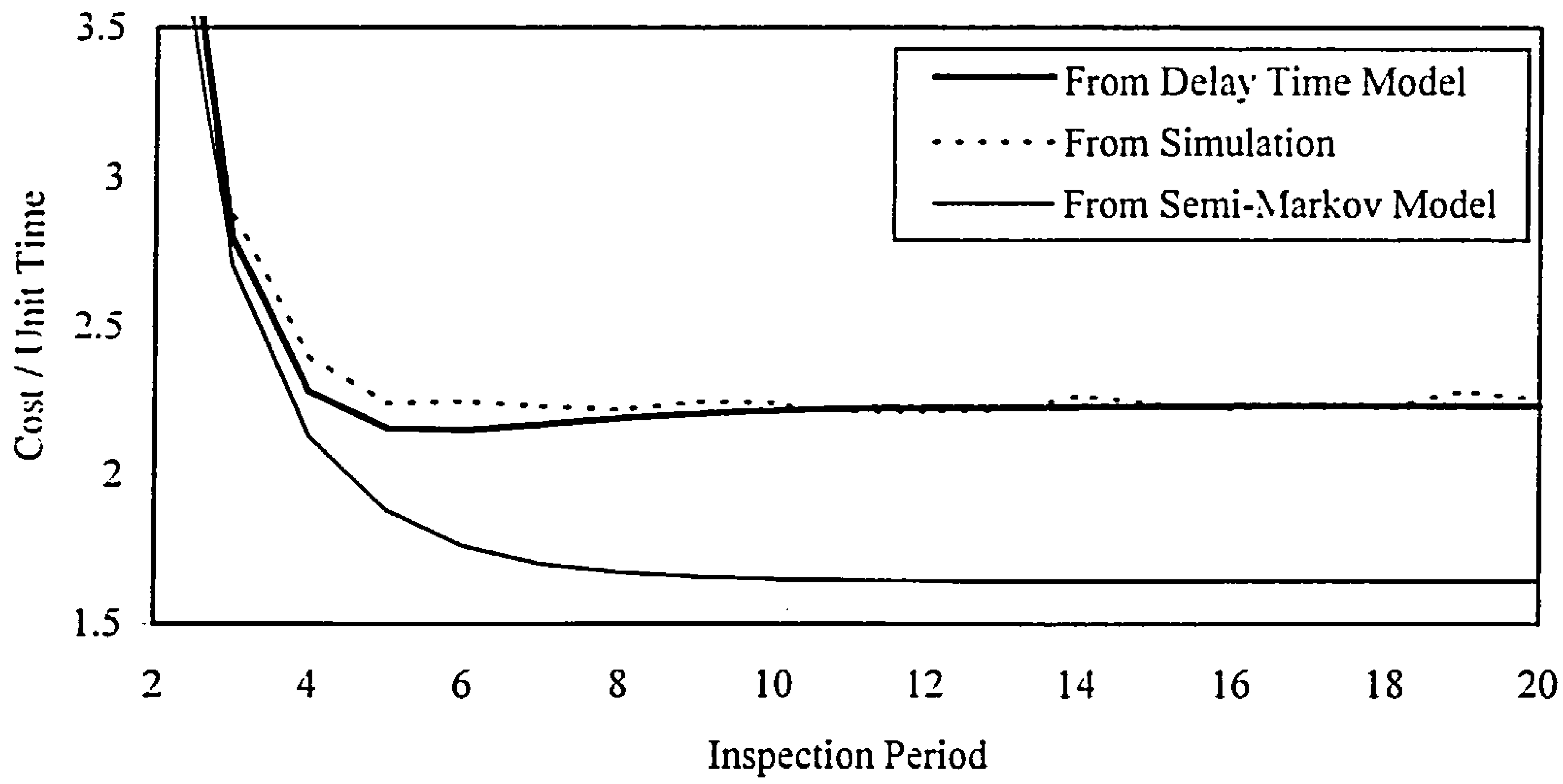


Figure 4.15. The expected cost per unit time according to the inspection period.

(This is a non-Markovian case of the fault arrival rate when Δt=1, the delay time has a Weibull distribution, α=2, β=1, γ=1.5, δ=1.3, $C_i = 10$, C_d =2, $C_b = 5$, $d_i = 0.08$, and $d_b = 0.18$.)

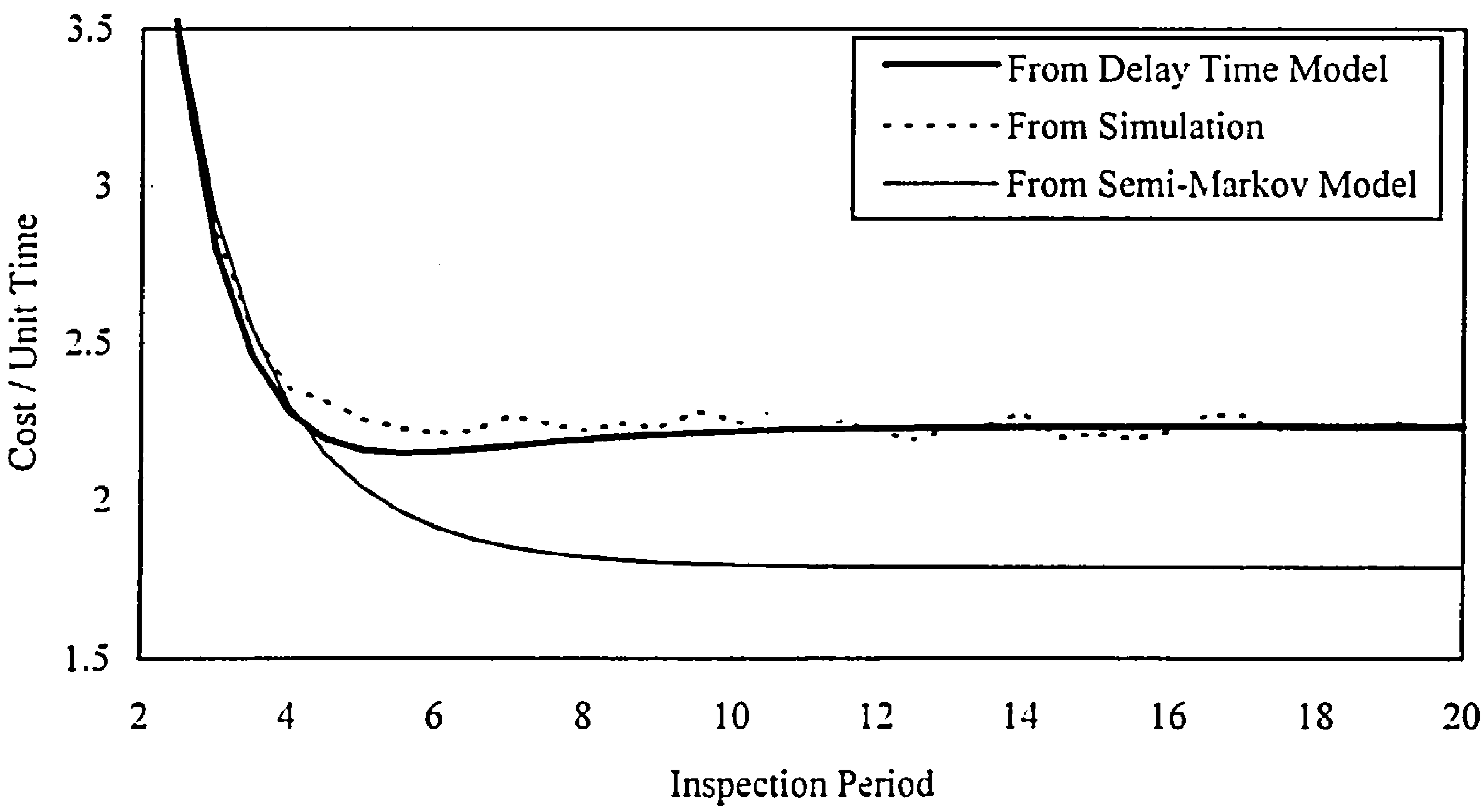


Figure 4.16. The expected cost per unit time according to the inspection period.

(This is a non-Markovian case of the fault arrival rate when Δt=0.5, the delay time has a Weibull distribution, α=2, β=1, γ=1.5, δ=1.3. $C_i = 10$. C_d =2, $C_b = 5$. $d_i = 0.08$, and $d_b = 0.18$.)

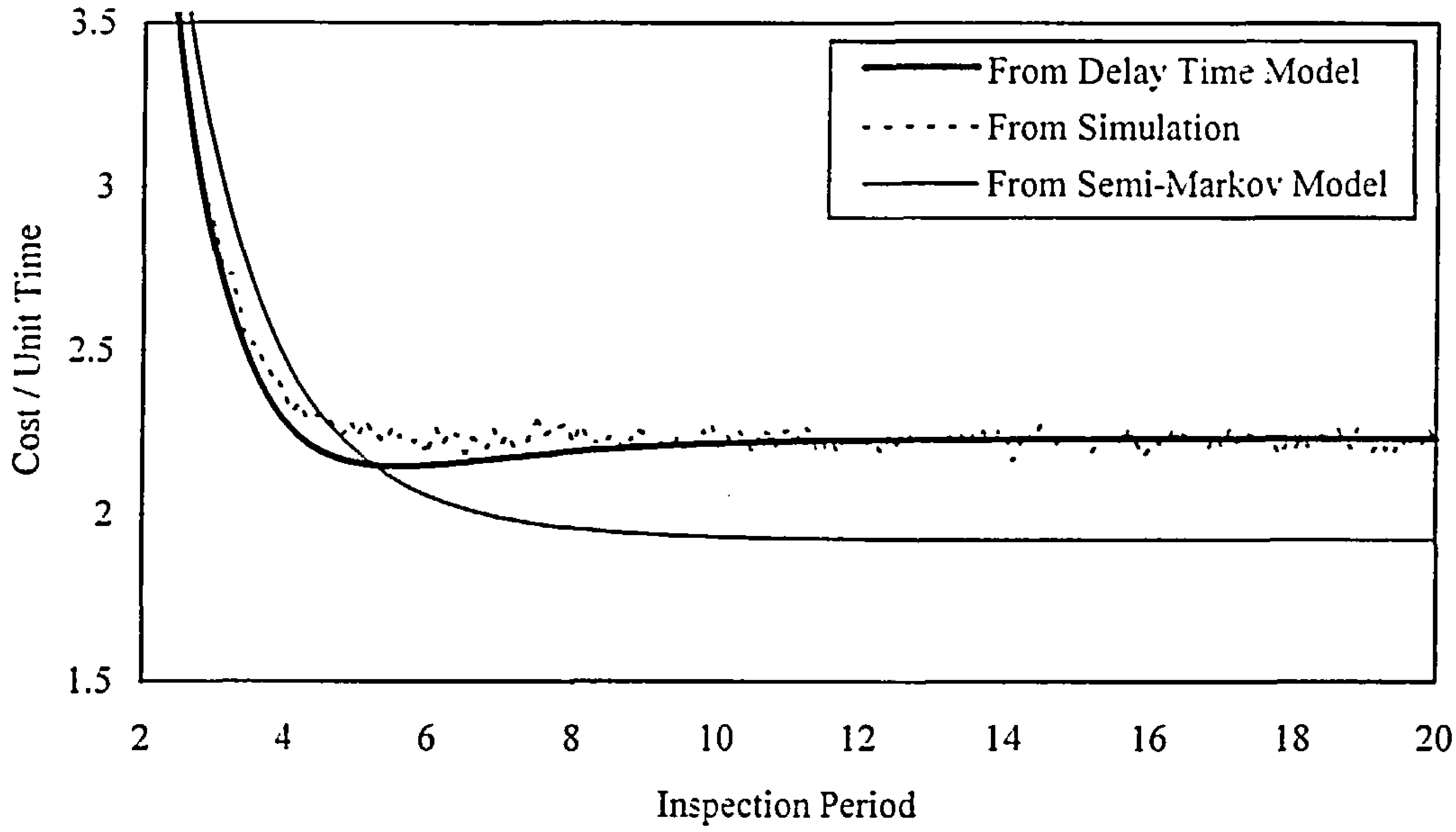


Figure 4.17. The expected cost per unit time according to the inspection period.

(This is a non-Markovian case of the fault arrival rate when $\Delta t = 0.1$, the delay time has a Weibull distribution, α=2, β=1, γ=1.5, δ=1.3, $C_i = 10$, $C_d = 2$, $C_b = 5$, $d_i = 0.08$, and $d_b = 0.18$.)

Also, to compare the semi-Markov model curve with the delay time model curve in detail, we change the scale of the cost per unit time and remove the simulation curve of Figure 4.17, see Figure 4.18. Figure 4.18 shows that the semi-Markov model curve has a notable difference with the delay time model curve as the inspection period T increases. Furthermore, from Figure 4.18, we can see that a finite optimal inspection period point which minimises the cost per unit time can be identified from the delay time model curve, but cannot be decided from the semi-Markov model curve because as the inspection period T increases the cost per unit time of the semi-Markov model curve decreases. The reason is that since the initial point u and the delay time h have a Weibull distribution which does not have a Markov property the failure process can not follow the Markov process required of the semi-Markov model. Therefore, when the system does not have a Markov property, the system cannot be fitted to the semi-Markov inspection model but can be fitted to the delay time model.

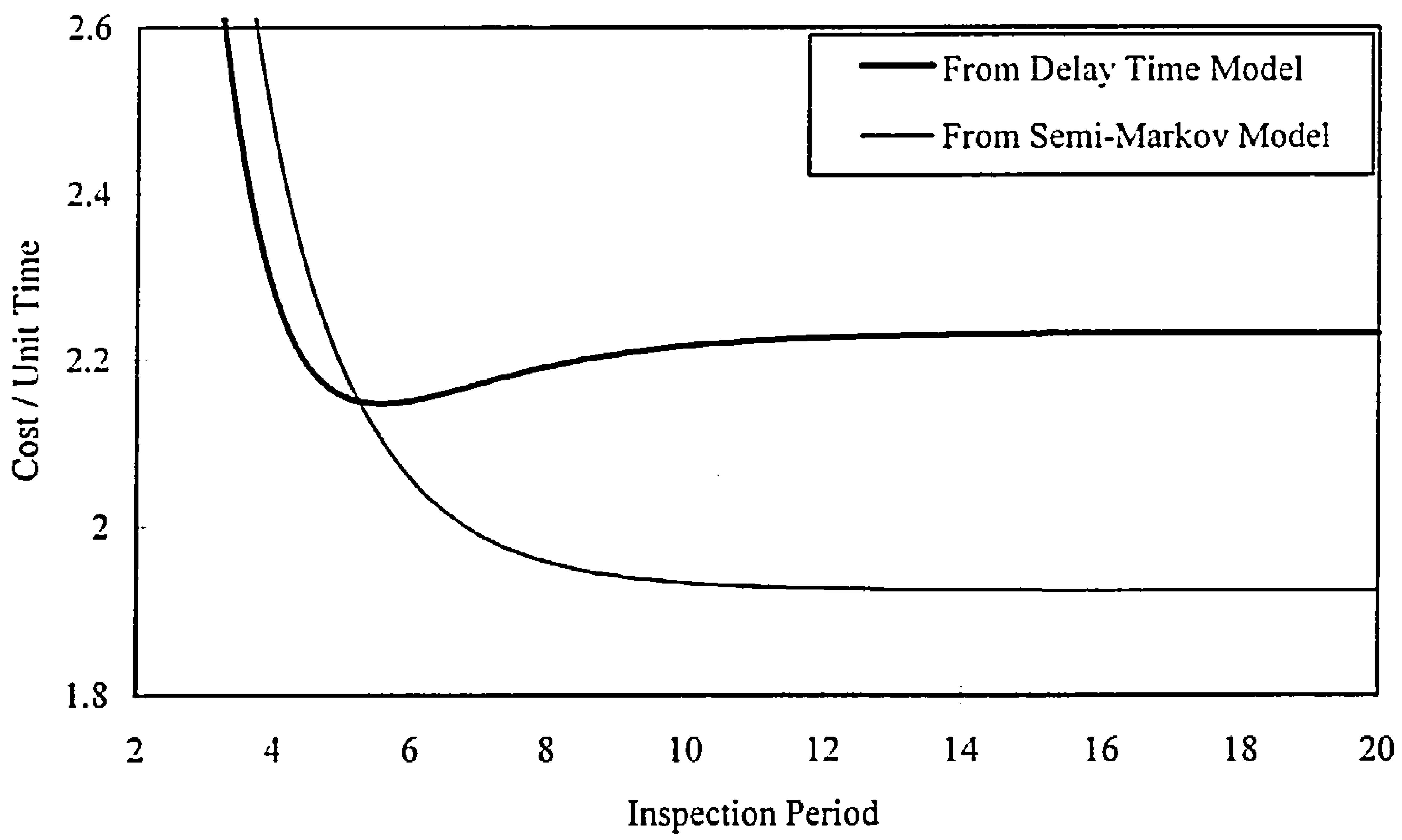


Figure 4.18. The detailed expected cost per unit time according to the inspection period.

(This is a non-Markovian case of the fault arrival rate when Δt = 0.1, the delay time has a Weibull distribution, α=2, β=1, γ=1.5, δ=1.3, $C_i = 10$, $C_d = 2$, $C_b = 5$, $d_i = 0.08$, and $d_b = 0.18$.)

4.6 Comparison and Conclusions

As expected, the semi-Markov inspection model is consistent with the delay time model in situations when the Markov assumption is valid, see Figures 4.8 to 4.10 and Figures 4.12 to 4.14 of the numerical examples. Here the key point is that the system has a Markov property because the initial point u has a negative exponential distribution which has a memoryless property. In real-world situations, however, there are few cases with data for which the Markov property can be established. Even when the data are given, it may be difficult to justify the Markov property and to get its state transition probabilities from the data. Either way, if the system has a Markov property, it was shown in section 4.3, that the deterioration probability r_{ij}^t can be estimated based upon the delay time concept; that is the state transition probabilities follow from delay time parameters. Relying on the delay time concept, the semi-Markov inspection model can now be used for the first time in practice, for the single component system.

In a single component system, as discussed in the numerical examples of the section 4.5, if the initial point u does not satisfy the Markov property, the system cannot be fitted to the semi-Markov inspection model. Figures 4.15 to 4.18 illustrate that an optimal finite inspection period point, which minimises the expected cost per unit time, may not be given by the semi-Markov inspection model because as the inspection period T increases the expected cost per unit time of the semi-Markov inspection model decreases continuously. Also, Figures 4.15 to 4.17 illustrate that the semi-Markov inspection model curve is far from the simulation curve and the delay time model curve. In contrast to the semi-Markov inspection model, the delay time model can not only give the optimal inspection period point in Figures 4.15 to 4.18, but also the delay time model curve is consistent with the simulation curve in Figures 4.15 to 4.17. We have seen that delay time model may be fitted to a large class of inspection / PM problems, and that techniques exist for parameter estimation and model validation. In the case of a semi-Markov model, the class of problems being addressed requires the Markov property to be satisfied. In delay time model terms, this means the time to initial point is negative exponential. In the case considered, the Markov model is derived under the assumption of a delay time based configuration which allows the transition probabilities to be evaluated. A process of analytic continuation enables values for notional transition probabilities to be calculated even when the Markov assumption is not valid. Figures 4.15 to 4.17 indicate the order of error introduced by this 'approximation'. Evidently the Markov model cannot be assumed to be robust to the Markov assumptions, and therefore Markov models become poor approximation for non-Markov cases. This emphasises the recognised generality of the delay time formulation for actual problem solutions.

Semi-Markov inspection models have provided a useful means of modelling the effect of inspections on the failure rate of repairable system. They enable the structure of optimal solutions to be identified. However, as already stated, in attempting to apply such models in a real-world situation, it may prove difficult to justify the Markov property assumptions and to determine the state transition probabilities or the deterioration probability from the data. When the initial point u has a negative exponential distribution, the system can be fitted to the semi-Markov inspection model by estimating the deterioration probability from the data using the delay time concept. In the case that the initial point u is not negative exponentially distributed, the system cannot be formally fitted to the semi-Markov inspection model. In contrast to the semi-Markov inspection model, the delay time model is free from such constraints. In addition, the formulation of the delay time model is seen as being much simpler than that of the semi-Markov inspection model, and the delay time model requires less computing time than the semi-Markov inspection model to compute solutions. Therefore, the delay time model is more general and practical than the semi-Markov inspection model in applying to real-world situations.

Chapter 5

SEMI-MARKOV AND DELAY TIME MODELS OF MAINTENANCE FOR A MULTI-COMPONENT SYSTEM: A COMPARATIVE STUDY

5.1 Introduction

In the previous chapter, the semi-Markov inspection model and the delay time model for a single component system have been discussed. As a consequence of the previous chapter, it was seen that the delay time model was more general than the semi-Markov inspection model in terms of application to real-world situations for a single component system. There will, however, be few systems with a single component in a real-world situation. The system usually consists of many components and is therefore more complex. Thus these models need to be extended to represent a multi-component system.

In a multi-component system, the fact is that the scientific management of planned preventive maintenance will rely on the modelling of probabilistic parameters which can change. If the process of equipment deterioration and degradation were entirely deterministic, there would be no need for frequent inspection and condition monitoring. Changes in parameters that define equipment condition are generally probabilistic. This explains why, as Christer [1984] points out, many of the published theoretical models adopt a Markov approach, where states are usually 'operating', 'operating but fault present', and 'failed'. Transitions between these states occur according to probabilistic laws, with each state being associated with the coincident occurrence of an inspection and some associated maintenance actions.

A number of such models, as Baker and Christer [1994] point out, have a seemingly arbitrary nature in their assumptions and the lack of evident conviction in applicability to the real-world. Most of them assume that the working condition of the system can be expressed as a discrete-time Markov chain with states $0, 1, 2, ..., N, f,$ where the state 0 represents a good state, the states $1, 2, ..., N$ are degraded states and the state f is the failed state. Transition probabilities are assumed given. In practice, it is, as already indicated, difficult to define the degraded states for the deteriorated system, and also difficult to determine their transition probabilities. So in general, papers on Markov modelling in maintenance do not mention the 'fit' of the model to data and no examples of actual applications or case-studies are available. This applies to both component models and complex system models. The value of these papers is to be found in an investigation of problem structure.

A robust approach to model and solve industrial maintenance problems has been introduced and developed as the Delay Time Model (DTM) in a series of recent papers dating from 1984. We will see that a semi-Markov inspection model can be established here using the delay time concept. If the degraded states selected to represent the condition of the deteriorated system in a semi-Markov inspection model can be expressed in terms of the expected number of defects, we can easily define the working condition of the system as a Markov chain state. Also, if we know the fault arrival rate λ and the probability density function (*pdf*) of the delay time h, it is possible to transform the fault arrival rate and the *pdf* of the delay time h into transition probabilities of a semi-Markov inspection model. It must be remembered that there are methods for estimating delay time parameters, the objective method and the subjective method.

In this chapter, we consider a repairable complex system with many components that may fail or suffer breakdown during the course of its service lifetime. For this system, a typical semi-Markov inspection model and a delay time model are presented and the results are compared.

5.2 Basic System Description for Modelling

Consider a system for which the number of components is very large. We assume that if any one component of the system fails, the system is considered to have failed. For this system, an inspection policy which minimises the expected total long-run average cost or downtime can be derived. For modelling purposes, consider the general case of an inspection policy, which may be characterised by the following assumptions.

(1) The condition of the system can be observed by inspections only, and a failure will be observed immediately at its occurrence.
(2) An inspection is undertaken every T time units.
(3) Inspections are perfect in that any defect present within the system will be identified at inspection, and no new fault generated because of inspection.
(4) An inspection requires C_i cost units and d_i time units.
(5) Defects identified at an inspection will be repaired within the allocated inspection time, d_i, and the mean repair cost per defect is C_d units.
(6) The component is repaired immediately upon failure and the mean repair cost and time for a failure repair are C_b cost units and d_b time units respectively.
(7) The component is as good as new after repair.
(8) Defects are independent of each other and arise as a homogeneous Poisson Process (HPP), with rate of occurrence of defects λ.
(9) The delay time h of a defect is independent of the time of origin, and all defects share a common delay time *pdf* $f(h)$ and *cdf* $F(h)$.

As a consequence of assumptions (3), (5), (7), and (8), an inspection will renew the system because all components with no fault are essentially regarded as new by assumption (8) due to the memoryless property and the fact that faults in all components are identified and renewed after repair at the inspection time, see assumptions (3), (5), and (7). It is noted that assumption (3) will be relaxed to allow imperfect inspection according to need in the subsection 5.3.2 and 5.4.2.

For a system satisfying the above assumptions, we may collect data which includes the inspection time point, the inspection result, and the time of any failures. Characteristic data is summarised in Table 5.1 as an example, when inspections are occurred every T=10 time units.

Table 5.1. Characteristic structure of the data of the multi-component system for T=10 units.

Time point	Inspection result or failure	Action
0	Normally operated	No action
10	2 faults detected	Repair
20	5 faults detected	Repair
27	Failure	Repair
30	1 fault detected	Repair
40	No fault detected	No action
44	Failure	Repair
:	:	:
:	:	:

This data indicates that 2 defects arose within the period (0, 10) and identified at T=10, 5 defects arose within the period (10, 20) and was identified at T=20, 1 defect arose within the period (20, 27) and lead to a failure at T=27, 1 defect arose within the period (20, 30) and was identified at T=30, and so on. If it is not possible to obtain such data from actual maintenance and operating records, assuming a delay time model, it is possible to simulate the process and thereby obtain these data. Either way, having the data, it may then be used to estimate the underlying delay time parameters and in this way model the real-world, or simulated real-world, situation. These parameter estimating procedures re-gain the parameter. When the data is simulated, this process provides a test of the estimating procedure. In section 5.5, we will describe the process of simulating data.

5.3 Semi-Markov Inspection Model

5.3.1 A Semi-Markov Inspection Model for a Perfect Inspection Policy

The inspection system described in section 5.2 can, as in the semi-Markov inspection model for a multi-component system, be analysed within the framework of a semi-Markov decision process to find the long term consequences of different inspection periods. For a multi-component system with discrete-time Markovian deterioration, the system can, as in the paper of Tijms and Van Der Duyn Schouten' [1985], be observed in one of the working conditions $0, 1, 2, \ldots, N, f$, where N is the upper bound of the number of defects, the working condition of the system 0 represents a normally operated state without a defect, the working conditions of the system $1, 2, \ldots, N$ represent defective states, and the working condition of the system f represents a failed state. To overcome the difficulty of expressing the degree of system deterioration as the defective states $1, 2, \ldots, N$, we assume that the defective states $1, 2, \ldots, N$ in semi-Markov inspection model represent the expected number of defects in the system, which may be determined using the delay time concept. Accepting the defined states for the working conditions of the system and assuming the inspection as outlined in section 5.2, we formulate the possible case of a Markovian inspection policy, which may be charaterised by the following additional assumptions;

(10) In the absence of inspections and repairs, the working condition of the system follows a discrete-time Markov chain.

(11) Opportunities for inspections occur only at equidistant points or epochs in time $t = \Delta t, 2\Delta t, 3\Delta t, \cdots$.

(12) The working condition of the system cannot improve on its own.

The decision epochs for this model include the epochs at which opportunities for inspections occur when the system is operating, and the epochs at which the exact working condition of the system is revealed by an inspection. We take as state space

$$I = \{i | i = 0, 1, 2, \cdots, N\} \cup \{(0, m\Delta t), (m\Delta t, f) | m = 1, 2, \cdots, M\},$$

where Δt is an arbitrary small time. Also, state i corresponds to the situation in which an inspection identifies i defects within the system, the states $(0, m\Delta t)$ corresponds to the situation in which $m\Delta t$ time units have passed since the last inspection, and the states $(m\Delta t, f)$ correspond to the situation in which a breakdown has occurred between $(m\text{-}1)\Delta t$ and $m\Delta t$. We assume a sufficiently large integer M is chosen to ensure that an inspection must always be made in the state $(0, M\Delta t)$. It is noted that the integer M will have no influence on the modelling if a finite solution exists.

The possible actions a are denoted by

$$a = \begin{cases} 0, & \text{leave the system as it is,} \\ 1, & \text{inspect the system,} \\ 2, & \text{repair the component.} \end{cases}$$

The action $a = 0$ is the only feasible action in the state 0, the action $a = 2$ is the only feasible action in the states i with $1 \le i \le N$ and in the states $(m\Delta t, f)$, actions $a = 0$ and $a = 1$ are the feasible actions in the states $(0, m\Delta t)$ except for the state $(0, M\Delta t)$, where the only action $a = 1$ is feasible.

As in the previous chapter, we define the one-step transition probabilities $P_{ij}(a)$, the one-step expected transition times $\tau_i(a)$, the one-step expected costs $C_i(a)$, and the one-step expected downtimes $D_i(a)$ given by the following.

$P_{ij}(a)$: The probability that at the next decision epoch the system will be in state j if actions a is chosen in the present state i.

$\tau_i(a)$: The expected time until the next decision epoch if action a is chosen in the present state i.

$C_i(a)$: The expected cost caused by the action a if action a is chosen in the present state i.

$D(a)$: The expected downtime caused by the action a if action a is chosen in the present state i.

As described in chapter 3, Ross [1983] presented the following proposition:

If an event arrival process follows a Poisson process with rate λ, the number of events that occur by time t, $N(t)$, is an independent Poisson random variable having mean given by $\lambda \int_0^t P(s)ds$, where $P(s)$ is the probability that a marking event occurs, independently of all else, at time s.

As a generalisation of the above proposition in Ross [1983], Christer and Wang [1995] presented the following Lemma 5.1.

Lemma 5.1

If the defect arrival process follows a HPP with the rate λ, the number of defects identified (the marking events) at time x by an inspection at time x is Poisson distributed with a mean given by

$$\mathrm{E}N_d(x) = \int_0^x \lambda(1 - F(x-u))du, \tag{5.1}$$

where, as defined in assumption (9) of section 5.2, $F(h)$ is the *cdf* of the delay time h.

Also, noting the Ross' [1983] proposition, Christer and Wang [1995] presented the following Lemma 5.2.

Lemma 5.2

If the defect arrival process follows a HPP with the rate of λ, we have the failure arrival process follows a non-homogeneous Poisson Process (NHPP) with the rate function given by

$$v(x) = \int_{j}^{x} \lambda f(x-u)du$$
$$= \lambda F(x), \tag{5.2}$$

where, as defined in assumption (9) of section 5.2, $f(h)$ is the *pdf* of the delay time h.

Based upon the above Lemmas, the one-step transition probabilities $P_{ij}(a)$, the one-step expected transition times $\tau_i(a)$, the one-step expected costs $C_i(a)$, and the one-step expected downtimes $D_i(a)$ can now be readily obtained by considering the action at each present state.

Firstly, at state 0, the action $a = 0$, which is the only sensible possible action to take at state 0, can be taken. If the action $a = 0$ is taken at state 0, the system will either survive until the next decision epoch Δt or fail within the next decision epoch Δt. The probability of a failure arising within the next very small decision epoch Δt can be obtained from Lemma 5.2. Accordingly, for very small Δt, we have that

$$P_{0j}(0) = \begin{cases} \int_0^{\Delta t} v(x)dx & \text{for } j = (\Delta t, f) \\ 1 - \int_0^{\Delta t} v(x)dx & \text{for } j = (0, \Delta t) \\ 0 & \text{otherwise}. \end{cases} \tag{5.3}$$

Also, if the system is survived until the next decision epoch Δt. the expected time incurred at state 0 is only the time to the next decision epoch Δt. If the system is failed before the next decision epoch Δt, assuming that the time interval to the next decision Δt is very small, we can approximately regard the expected time incurred at state 0 as the time to the next decision epoch Δt. Accordingly, we have that

$$\tau_0(0) = \Delta t. \tag{5.4}$$

When the action $a = 0$ is taken at state 0, since no cost and downtime are required, it follows that

$$C_0(0) = 0 \qquad (5.5)$$

and

$$D_0(0) = 0\,. \qquad (5.6)$$

At state k, $k = 1, 2, \ldots, N$, the action $a = 2$, is the only possible action to take at state k because any fault found at an inspection is repaired by assumption (3) and (5) of the section 5.2. If the action $a = 2$ is taken at state k, we have by assumptions (5) and (7) that

$$P_{k0}(2) = 1 \qquad \text{for } k = 1, 2, \ldots, N. \qquad (5.7)$$

Also, since the fault identified at an inspection is repaired within the inspection period and its repair costs per defect is C_d units, it is obvious that

$$\tau_k(2) = 0 \qquad \text{for } k = 1, 2, \cdots, N, \qquad (5.8)$$

$$C_k(2) = kC_d \qquad \text{for } k = 1, 2, \cdots, N, \qquad (5.9)$$

and

$$D_k(2) = 0 \qquad \text{for } k = 1, 2, \cdots, N. \qquad (5.10)$$

At state $(0, m\Delta t)$, $m = 1, 2, \ldots, M$, the action $a = 0$ and $a = 1$, which are the only possible actions to take at state $(0, m\Delta t)$, can be taken. If the action $a = 0$ is taken at state $(0, m\Delta t)$ with $m = 1, 2, \ldots, M\text{-}1$, the system will either survive until the next decision epoch $(m+1)\Delta t$ or fail within the next decision epoch $(m+1)\Delta t$ having survived to the present decision epoch $m\Delta t$. Since the probability that there is a failure between the present decision epoch $m\Delta t$ and the next decision epoch $(m+1)\Delta t$ for very small Δt is $\int_{m\Delta t}^{(m-1)\Delta t} v(x)dx$ from the Lemma 5.2, we have that

$$P_{(0,m\Delta t)j}(0) = \begin{cases} \int_{m\Delta t}^{(m-1)\Delta t} v(x)dx & \text{for } j = ((m+1)\Delta t, f) \\ 1 - \int_{m\Delta t}^{(m+1)\Delta t} v(x)dx & \text{for } j = (0,(m+1)\Delta t) \\ 0 & \text{otherwise .} \end{cases} \tag{5.11}$$

Also, if the system is survived until the next decision epoch $(m+1)\Delta t$ from the current decision epoch $m\Delta t$, the expected time incurred by taking the action $a = 0$ at state $(0, m\Delta t)$ is Δt time units from the current decision epoch $m\Delta t$ to the next decision epoch $(m+1)\Delta t$. Although the system is failed before the next decision epoch $(m+1)\Delta t$, assuming that Δt is very small, we can approximately regard the expected time incurred by taking the action $a = 0$ at state $(0, m\Delta t)$ as Δt time units to the next decision epoch $(m+1)\Delta t$. Accordingly, it is obvious that

$$\tau_{(0,m\Delta t)}(0) = \Delta t . \tag{5.12}$$

When the action $a = 0$ is taken at state $(0, m\Delta t)$, since no cost and downtime are required, it follows that

$$C_{(0,m\Delta t)}(0) = 0 \tag{5.13}$$

and

$$D_{(0,m\Delta t)}(0) = 0 . \tag{5.14}$$

If action $a = 1$ is taken at state $(0, m\Delta t)$ with $m=1, 2, ..., M$, assuming the perfect inspection of the system, the inspection will result in a situation of finding j, $j = 0, 1, 2, ..., N$, faults at an inspection. Since the number of defects identified at an inspection has a Poisson distribution by the Lemma 5.1 and the assumption (8) of the section 5.2, we have that

$$P_{(0,m\Delta t)j}(1) = \begin{cases} e^{-EN_d(m\Delta t)} \dfrac{(EN_d(m\Delta t))^j}{j!} & \text{for } j = 0,1,2,\cdots,N \\ 0 & \text{otherwise} \end{cases} \quad (5.15)$$

Also, since an inspection requires d_i time units and C_i cost units, it is obvious that

$$\tau_{(0,m\Delta t)}(1) = d_i\,, \quad (5.16)$$

$$C_{(0,m\Delta t)}(1) = C_i\,, \quad (5.17)$$

and

$$D_{(0,m\Delta t)}(1) = d_i\,. \quad (5.18)$$

At state $(m\Delta t, f)$, $m = 1, 2, \ldots, M$, the action $a = 2$ is the only possible action to take because a failure is immediately repaired. If the action a = 2 is taken at state $(m\Delta t, f)$, we have by assumptions (1), (6) and (7) that

$$P_{(m\Delta t,f)(0,(m+\frac{d_b}{\Delta t})\Delta t)}(2) = 1 \qquad \text{for } m = 1, 2, \ldots, M. \quad (5.19)$$

Also, since a failure requires C_b cost units and d_b time units, it is obvious that

$$\tau_{(m\Delta t,f)}(2) = d_b \qquad \text{for } m = 1,2,\cdots,M, \quad (5.20)$$

$$C_{(m\Delta t,f)}(2) = C_b \qquad \text{for } m = 1,2,\cdots,M, \quad (5.21)$$

and

$$D_{(m\Delta t,f)}(2) = d_b \qquad \text{for } m = 1,2,\cdots,M. \tag{5.22}$$

As in the single component case model, it still needs to be established that the above multi- component system model, in fact, satisfies the Markovian property requirements. We have that since equations (5.3), (5.7), (5.11), (5.15), and (5.19) are dependent only upon the current state i, $i \in I$, the next state j, $j \in I$, and the chosen action a, a =0, 1, 2, the one-step transition probabilities, $P_{ij}(a)$, satisfy condition (1) of the semi-Markov decision process of subsection 4.3.2. Also, since the one-step expected costs, $C_i(a)$, and the one-step expected downtimes, $D_i(a)$, are dependent only upon the current state i, $i \in I$, and the chosen action a, $a = 0, 1, 2$ (see equations (5.5), (5.6), (5.9), (5.10), (5.13), (5.14), (5.17), (5.18), (5.21), and (5.22)), the one-step expected costs and the one-step expected downtimes satisfy condition (2) of subsection 4.3.2 for a semi-Markov decision process. Accordingly, the above inspection system can be analysed within the framework of a semi-Markov decision process.

We have specified the basic elements of the semi-Markov decision model. Fix now a control-limit rule R with parameter value s, which is the time to next inspection when the working condition revealed at present time is 0. We first form the cost model using this control-limit rule R. Utilising the standard semi-Markov decision model of the subsection 4.3.2 (see the equation (4.5)) and the above specifications, we can have that

$$v_0 = -g_c(R)\Delta t + P_{0(0,\Delta t)}(0)v_{(0,\Delta t)} + P_{0(\Delta t,f)}(0)v_{(\Delta t,f)}, \tag{5.23}$$

$$v_k = kC_d + P_{k0}(2)v_0 \qquad \text{for } k = 1,2,\cdots,N, \tag{5.24}$$

$$v_{(0,m\Delta t)} = -g_c(R)\Delta t + P_{(0,m\Delta t)((m+1)\Delta t,f)}(0)v_{((m+1)\Delta t,f)} + P_{(0,m\Delta t)(0,(m+1)\Delta t)}(0)v_{(0,(m+1)\Delta t)}$$
$$\text{for } 0 < m\Delta t < s. \tag{5.25}$$

$$v_{(0,m\Delta t)} = C_i - g_c(R)d_i + P_{(0,m\Delta t)0}(1)v_0 + P_{(0,m\Delta t)1}(1)v_1 + \cdots + P_{(0,m\Delta t)N}(1)v_N$$
$$\text{for } s \le m\Delta t \le M\Delta t, \tag{5.26}$$

and

$$v_{(m\Delta t, f)} = C_b - g_c(R)d_b + P_{(m\Delta t, f)(0,(m-\frac{d_b}{\Delta t})\Delta t)}(2)v_{(0,(m+\frac{d_b}{\Delta t})\Delta t)}$$

$$\text{for } m = 1,2,\cdots, M, \tag{5.27}$$

where $g_c(R)$ is the average cost per unit time given policy R and v_x , $x \in I$, are the relative costs resulting from the various starting states when policy R is used. Using the embedded technique, by a repeated application of the above equations, we can obtain the average cost per unit time $g_c(R)$. By putting one of the relative costs equal to zero, say $v_0 = 0$, the linear equation can determine uniquely the average cost per unit time $g_c(R)$. Once $g_c(R)$ and v_0 have been determined we can obtain all the relative costs v_x by recursive calculations if required.

In a similar way, we have for the downtime model that

$$w_0 = -g_d(R)\Delta t + P_{0(0,\Delta t)}(0)w_{(0,\Delta t)} + P_{0(\Delta t,f)}(0)w_{(\Delta t,f)}, \tag{5.28}$$

$$w_k = P_{k0}(2)w_0 \qquad \text{for } k = 1,2,\cdots,N, \tag{5.29}$$

$$w_{(0,m\Delta t)} = -g_d(R)\Delta t + P_{(0,m\Delta t)((m+1)\Delta t,f)}(0)w_{((m+1)\Delta t,f)} + P_{(0,m\Delta t)(0,(m-1)\Delta t)}(0)w_{(0,(m+1)\Delta t)}$$

$$\text{for } 0 < m\Delta t < s, \tag{5.30}$$

$$w_{(0,m\Delta t)} = d_i - g_d(R)d_i + P_{(0,m\Delta t)0}(1)w_0 + P_{(0,m\Delta t)1}(1)w_1 + \cdots + P_{(0,m\Delta t)N}(1)w_N$$

$$\text{for } s \le m\Delta t \le M\Delta t, \tag{5.31}$$

and

$$w_{(m\Delta t, f)} = d_b - g_d(R)d_b + P_{(m\Delta t, f)(0,(m-\frac{d_b}{\Delta t})\Delta t)}(2)w_{(0,(m+\frac{d_b}{\Delta t})\Delta t)}$$

$$\text{for } m = 1,2,\cdots, M, \tag{5.32}$$

where $g_d(R)$ is the average downtime per unit time given policy R and w_x, $x \in I$, are the relative downtimes resulting from the various starting states when policy R is used. Using the same embedded technique used above, we can obtain the expected average downtime per unit time $g_d(R)$. By putting one of the relative downtimes equal to zero, say $w_0 = 0$, the linear equation can determine uniquely the average downtime per unit time $g_d(R)$. Once $g_d(R)$ and w_0 have been determined we can obtain all the relative downtimes w_x by recursive calculations if required.

These models can be evaluated using the following policy-iteration algorithm of the subsection 4.3.2.

Policy-iteration algorithm

Step 0 : Choose an initial policy R with the parameter s.

Step 1 : For the current rule R, compute the average costs $g_c(R)$ and the relative costs v_i, $i \in I$, or the average downtimes $g_d(R)$ and the relative downtimes w_i, $i \in I$, as the unique solution to the linear equations

$$v_i = C_i(R_i) - g_c(R)\tau_i(R_i) + \sum_{j \in I} P_{ij}(R_i)v_j, \qquad i \in I, \tag{5.33}$$

$$v_x = 0,$$

in the cost case, or in the downtime case

$$w_i = D_i(R_i) - g_d(R)\tau_i(R_i) + \sum_{j \in I} P_{ij}(R_i)w_j, \qquad i \in I, \tag{5.34}$$

$$w_x = 0,$$

where x is an arbitrarily chosen state.

Step 2 : For each state $i \in I$, determine an action a_i yielding the minimum in

$$\min_{a \in A(i)} \{C_i(a) - g_c(R)\tau_i(a) + \sum_{j \in I} P_{ij}(a) v_j(R)\},$$

or

$$\min_{a \in A(i)} \{D_i(a) - g_d(R)\tau_i(a) + \sum_{j \in I} P_{ij}(a) w_j(R)\}.$$

The new stationary policy $\overline{R}$ is obtained by choosing $\overline{R_i} = a_i$ for all $i \in I$ with the convention that $\overline{R_i}$ is chosen as being the old action R_i when this action minimises the policy-improvement quantity.

Step 3 : If the new policy $\overline{R}$ equal the old policy, the algorithm is stopped with policy R. Otherwise, the algorithm cycles back to step 1 with R replaced by $\overline{R}$.

This algorithm generates a sequence of improving control-limit rules and it can be shown that the algorithm converges after a finite number of iterations to an average cost or downtime optimal policy (see Tijms and Van Der Duyn Schouten [1985] and Tijms [1986]). Also, as a consequence of the convergence of the algorithm, there exist a $g_c^*(R)$ and v_i^*, $i \in I$, or $g_d^*(R)$ and w_i^*, $i \in I$, where the constant $g_c^*(R)$ is uniquely determined as the minimal average costs per unit time and v_i^* as the relative costs or $g_d^*(R)$ is uniquely determined as the minimal average downtimes per unit time and w_i^* as the relative downtimes, when the decision variable s would be selected to minimise the average cost per unit time or the average downtime per unit time.

5.3.2 A Semi-Markov Inspection Model for an Imperfect Inspection Policy

In the previous subsection 5.3.1, we have assumed that inspections are perfect in that any fault present will be identified. However, in most cases, it is more reasonable to assume that inspections are imperfect. So, in the section 5.2, the assumption (3) will be relaxed to allow an imperfect inspection policy, which may be characterised by the following changed assumption $(3')$.

$(3')$ Inspections are assumed to be imperfect in that a fault present will be identified with probability r, $0 \le r \le 1$. Probabilities of detecting a fault at successive inspections are assumed to be independent and constant.

If inspections are imperfect, as in the above assumption $(3')$, we can not directly establish the semi-Markov inspection model based upon Lemma 5.1 and 5.2 because the failures arriving in the present inspection period may be affected by the faults arising in past inspection periods. To establish the semi-Markov inspection model for an imperfect inspection policy, we need to change Lemma 5.1 and 5.2.

With an imperfect inspection policy, consider the probability of a failure arriving in $(t, t+\Delta t)$ of the period (T_{i-1}, T_i) resulting from a fault arising at time y in the period (T_{n-1}, T_n), namely $P(t, t+\Delta t | y)$, (see Figure 5.1).

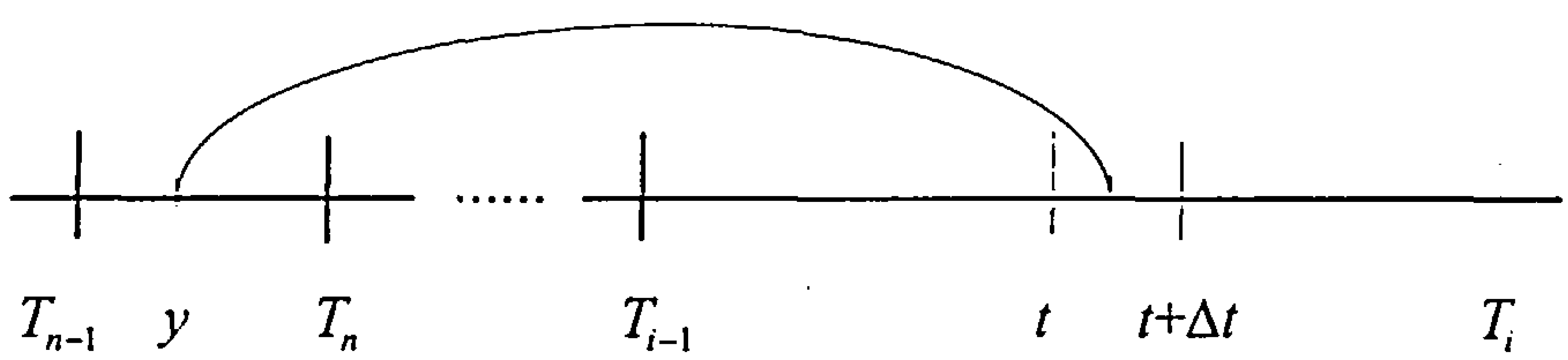


Figure 5.1. The failure process of a fault arising in (T_{n-1}, T_n)

Since the probability that a fault is identified during the inspection time is r, the probability that a fault arising in (T_{n-1}, T_n) with sufficiently long delay time will not be identified before the interval (T_{i-1}, T_i) is given by

$$c_n = (1-r)^{i-n} \qquad \text{for } n = 1, 2, \ldots, i\text{-}1. \tag{5.35}$$

Accordingly, the probability of a failure in $(t,t+\Delta t)$ resulting from a fault arising at time y in (T_{n-1},T_n) is given by

$$P(t,t+\Delta t|y)=\begin{cases} c_n(F(t+\Delta t-y)-F(t-y)) & \text{for } T_{n-1}<y<T_n,\ n=1,2,\cdots,i-1 \\ F(t+\Delta t-y)-F(t-y) & \text{for } T_{i-1}<y<t \\ F(t+\Delta t-y) & \text{for } t<y<t+\Delta t \\ 0 & \text{otherwise.} \end{cases} \qquad (5.36)$$

From equation (5.36), we can obtain the expected number of faults identified at time t if there is an inspection at time t, namely $\mathrm{E}N_d(t)$, given by

$$\mathrm{E}N_d(t)=\lambda\sum_{n=1}^{i-1}c_n r\int_{T_{n-1}}^{T_n}(1-F(t-y))dy+\lambda r\int_{T_{i-1}}^{t}(1-F(t-y))dy. \qquad (5.37)$$

Using equation (5.37), Christer *et al* [1995] presented the following Lemma 5.3.

Lemma 5.3

If the fault arrival process follows a HPP with the rate of λ, the number of faults identified at time t if there is an inspection at time t is Poisson distributed with a mean given by equation (5.37).

Also, using equation (5.36), we can derive the Lemma 5.4 instead of Lemma 5.2. The Lemma 5.4 is of value to us later.

Lemma 5.4

If the fault arrival process follows a HPP with the rate of λ, we have that the failure arrival process follows a NHPP with the rate function given by

$$v(t)=\int_0^t\lambda\lim_{\Delta t\to 0}\frac{P(t,t+\Delta t|y)}{\Delta t}dy$$

$$=\lambda\{\sum_{n=1}^{i-1}c_n(F(t-T_{n-1})-F(t-T_n))+F(t-T_{i-1})\}. \qquad (5.38)$$

Proof

To prove the lemma 5.4, we recall the definition of a Poisson process in Ross [1983]. Let $N_f(t)$, $t \geq 0$, be the number of failures which occur during $(0, t)$ and $v(t)$, $t \geq 0$, be the rate of occurrence of failures for the process. Then, to satisfy the definition of a nonhomogeneous Poisson process, we require that the process $N_f(t)$ satisfies

(1) $N_f(0) = 0$,

(2) $N_f(t)$ has independent increments,

(3) $P(N_f(t+dt) - N_f(t) = 1) = v(t)dt + o(dt)$,

(4) $P(N_f(t+dt) - N_f(t) \geq 2) = o(dt)$,

where for small dt, $o(dt)$ is defined as a function given by

$$\lim_{dt \to 0} \frac{o(dt)}{dt} = 0. \tag{5.39}$$

Condition (1), which simply states that the counting of events begins at time $t = 0$, and condition (2) can usually be directly verified from our knowledge of the process, or is otherwise assumed. We will now deduce that conditions (3) and (4) are valid for the current failure process with a regular inspection period. Should a failure arise in $(t, t+dt)$ before a regular inspection, we must have that a defect arises in some interval $(y, y+dy)$ with a delay time $h \in (t-y-dy, t-y+dt)$ (see Figure 5.2).

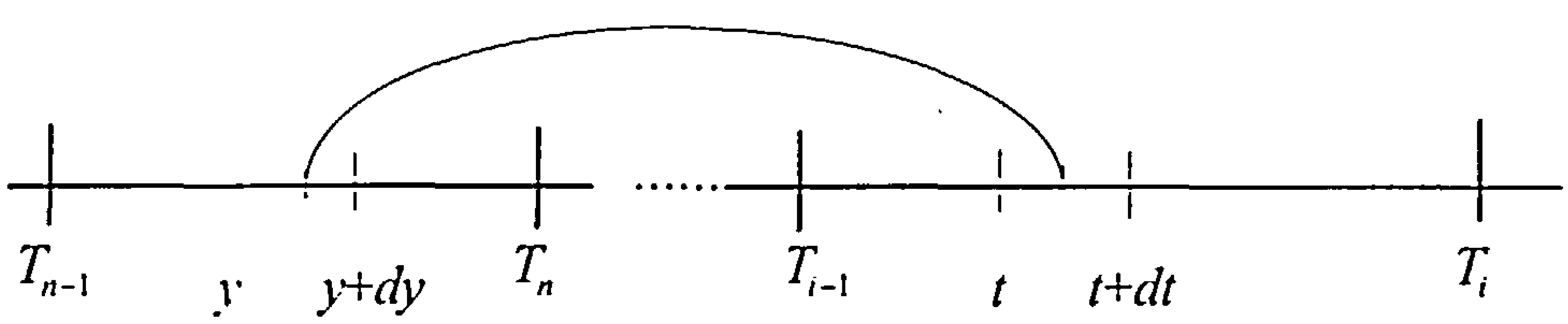


Figure 5.2. The failure process of a fault arising in $(y, y+dy)$ of (T_{n-1}, T_n).

Since the defect arrival process follows a HPP with the rate of λ, the probability that a defect arise in interval $(y, y+dy)$ is given by

$$P(N_d(y+dy) - N_d(y) = 1) = \lambda dy + o(dy), \tag{5.40}$$

where $N_d(y)$ denotes the number of defects in $(0, y)$. Also, if $p(y; t)$ denotes the probability density for a defect at time y, the probability this leads to a failure at time t equals the probability of having a delay time $h \in (t-y-dy, t-y+dt)$, that is

$$P(h \in (t-y-dy, t-y+dt)) = p(y;t)dt. \tag{5.41}$$

From equations (5.40) and (5.41), integrating over all possible y, we have

$$P(N_f(t+dt) - N_f(t) = 1) = \int_0^t \lambda p(y;t)dydt + o(dt). \tag{5.42}$$

If we define that the rate function is given by

$$v(t) = \int_0^t \lambda p(y;t)dy, \tag{5.43}$$

the equation (5.42) becomes

$$P(N_f(t+dt) - N_f(t) = 1) = v(t)dt + o(dt). \tag{5.44}$$

The equation (5.44) clearly satisfies the condition (3) for a NHPP.

For the condition (4), note that the probability of being over 1 failure, $P(N_f(t+dt) - N_f(t) \geq 2)$, is given by

$$P(N_f(t+dt) - N_f(t) \geq 2)$$

$$= 1 - P(N_f(t+dt) - N_f(t) = 1) - P(N_f(t+dt) - N_f(t) = 0)\,. \qquad (5.45)$$

Here, the probability of having no failures in $(t, t+dt)$, $P(N_f(t+dt) - N_f(t) = 0)$, is the summation of the probability of having no defects in $(0, t)$ and the probability that a defect arise in some interval $(y, y+dy)$ with a delay time $h \notin (t-y-dy, t-y+dt)$. Since the probability of no defects in $(0, t)$ is

$$P(\text{no defects in } (0,t)) = 1 - \int_0^t \lambda dy \qquad (5.46)$$

and the probability that a defect arise in some interval $(y, y+dy)$ with a delay time $h \notin (t-y-dy, t-y+dt)$ is

$$\begin{aligned} &P(1 \text{ defect in } (y, y+dy))P(h \notin (t-y-dy, t-y+dt) \\ &\quad = P(N_d(y+dy) - N_d(y) = 1)\{1 - P(h \in (t-y-dy, t-y+dt))\} \\ &\quad = (\lambda dy + o(dy))(1 - p(y;t)dt)\,, \end{aligned} \qquad (5.47)$$

integrating the equation (5.47) over all possible y, the probability of having no failures in $(t, t+dt)$ is

$$\begin{aligned} P(N_f(t+dt) - N_f(t) = 0) &= 1 - \int_0^t \lambda dy + \int_0^t (1 - p(y;t)dt)\lambda dy + o(dt) \\ &= 1 - v(t)dt + o(dt)\,. \end{aligned} \qquad (5.48)$$

Accordingly, considering the equation (5.44) for the probability of having 1 failure in $(t, t+dt)$, $P(N_f(t+dt) - N_f(t) = 1)$, the probability of more than 1 failure, $P(N_f(t+dt) - N_f(t) \geq 2)$, is given by

$$P(N_f(t+dt) - N_f(t) \geq 2) = o(dt)\,. \qquad (5.49)$$

The equation (5.49) satisfies the condition (4) which establish the NHPP for failure arrivals.

Given a defect arising at time y, the probability density function of time t to failure is given by

$$p(y;t) = \lim_{\Delta t \to 0} \frac{P(t, t+\Delta t | y)}{\Delta t}, \tag{5.50}$$

where, as before, $P(t, t+\Delta t | y)$ is the probability of having a failure over $(t, t+\Delta t)$ resulting from a fault arising at time y. We have that the failure rate function is given by

$$v(t) = \int_0^t \lambda \lim_{\Delta t \to 0} \frac{P(t, t+\Delta t | y)}{\Delta t} dy. \tag{5.51}$$

Also, for $T_{n-1} < y < T_n$, $n = 1, 2, \ldots, i\text{-}1$,

$$\lim_{\Delta t \to 0} \frac{P(t, t+\Delta t | y)}{\Delta t} = c_n \lim_{\Delta t \to 0} \frac{F(t+\Delta t - y) - F(t-y)}{\Delta t}$$
$$= c_n f(t-y), \tag{5.52}$$

and for $T_{i-1} < y < t$,

$$\lim_{\Delta t \to 0} \frac{P(t, t+\Delta t | y)}{\Delta t} = \lim_{\Delta t \to 0} \frac{F(t+\Delta t - y) - F(t-y)}{\Delta t}$$
$$= f(t-y). \tag{5.53}$$

Therefore, equation (5.51) becomes

$$v(t) = \lambda \sum_{n=1}^{i-1} c_n \int_{T_{n-1}}^{T_n} f(t-y) dy + \lambda \int_{T_{i-1}}^{t} f(t-y) dy$$
$$= \lambda \{ \sum_{n=1}^{i-1} c_n (F(t - T_{n-1}) - F(t - T_n)) + F(t - T_{i-1}) \}. \tag{5.54}$$

The proof of Lemma 5.4 is complete.

In solving a maintenance problem, our interest is in the reduction of the expected downtime per unit time over the long future period. Once the inspection process and plant has been operated for a long time period, it is expected that the behaviour will become steady state, and any initial influence of the newness at time $t = 0$ is lost. This being so, with the current assumption, we can regard an inspection as a regeneration point of the system. That is, subsequent inspection cycles become statistically identical over time.

If, after a long period of operation, an arbitrary time from an inspection T_{i-1} is x (see Figure 5.3), steady state condition can be assumed and the equation (5.36) will be changed slightly.

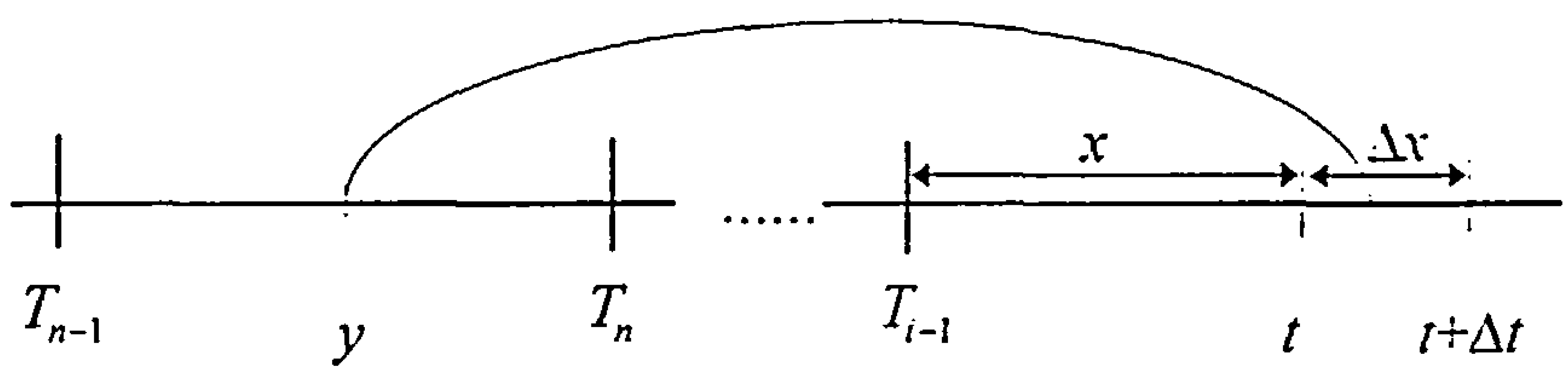


Figure 5.3. The failure process from an inspection T_{i-1}.

Since $t = T_{i-1} + x$ from Figure 5.3, using the equation (5.36), the probability of a failure in $(x, x+\Delta x)$ from an inspection T_{i-1} resulting from a fault arising at time y in (T_{n-1}, T_n) is given by

$$P(x, x+\Delta x | y)$$

$$= \begin{cases} c_n (F(T_{i-1} + x + \Delta x - y) - F(T_{i-1} + x - y)) & \text{for } T_{n-1} < y < T_n, \; n = 1,2,\cdots,i-1 \\ F(T_{i-1} + x + \Delta x - y) - F(T_{i-1} + x - y) & \text{for } T_{i-1} < y < T_{i-1} + x \\ F(T_{i-1} + x + \Delta x - y) & \text{for } T_{i-1} + x < y < T_{i-1} + x + \Delta x \\ 0 & \text{otherwise.} \end{cases} \quad (5.55)$$

From equation (5.55), the expected number of faults identified at time x from an inspection (which is a regeneration point if there is an inspection at time x), namely $EN_d(x)$, can be obtained by letting $i \to \infty$ in equation (5.37). This gives

$$EN_d(x) = \lim_{i \to \infty} \lambda r [\sum_{n=1}^{i-1} c_n \int_{T_{n-1}}^{T_n} (1 - F(T_{i-1} + x - y)) dy + \int_{T_{i-1}}^{T_{i-1}+x} (1 - F(T_{i-1} + x - y)) dy]. \quad (5.56)$$

Using equation (5.56) and Lemma 5.3, we can state the following Lemma 5.5.

Lemma 5.5

If the fault arrival process follows a HPP with the rate λ, *the number of faults identified at time* x *from an inspection (which is a regeneration point) if there is an inspection at time* x *is Poisson distributed with a mean given by equation (5.56).*

Also, from Lemma 5.4, the failure arrival rate function at time x from an inspection which is a regeneration point, namely $\rho(x)$, can be obtained by letting $i \to \infty$ in equation (5.38). Using equation (5.55), this gives

$$\rho(x) = \lim_{i \to \infty} \lambda \{ \sum_{n=1}^{i-1} c_n (F(T_{i-1} + x - T_{n-1}) - F(T_{i-1} + x - T_n)) + F(x) \}. \quad (5.57)$$

Using equation (5.57) and Lemma 5.4, we can state the following Lemma 5.6.

Lemma 5.6

If the fault arrival process follows a HPP with the rate of λ, we have that the failure arrival process follows a NHPP with the rate function given by equation (5.57).

By using Lemma 5.5 and 5.6 instead of Lemma 5.1 and 5.2 respectively, we can embed the semi-Markov inspection model for an imperfect inspection policy in the semi-Markov inspection model of the subsection 5.3.1. When computing the equations (5.56) and (5.57), since we are interested in a regular inspection policy, we have to note that

$$T_n = nT, \tag{5.58}$$

where T is the regular inspection period under the current policy R of the policy-iteration algorithm of the subsection 5.3.1. For a special delay time distribution $F(y)$, for example an exponential distribution, the expectation and failure of equations (5.56) and (5.57) can be obtained easily. Otherwise, we have to obtain the approximate value of equations (5.56) and (5.57) numerically. We will discuss this point in the subsection 5.5.3.

5.3.3 Evaluation of the Semi-Markov Inspection Model

In the previous chapter, the semi-Markov inspection model for a single component system was discussed. As a consequence of chapter 4, it was seen that when the distribution of the initial point u had a Markov property, the single component system could be modelled by a semi-Markov inspection model. However, in a real-world situation, there will be few systems with a single component. The system usually consists of many components and is therefore more complex. Thus, in contrast to the semi-Markov inspection model for a single component system, the semi-Markov inspection model for a multi-component system has the benefit of being relevant to the real-world situation. It is seen that if the fault arrival rate, regardless of the delay time distribution, satisfies the Markov property, then the multi-component system can also be modelled by a semi-markov inspection model.

However, the semi-Markov inspection model for the multi-component system may still have some problems in being applied to a real-world situation. There may be difficulties in expressing the working condition of the system as degraded states, and further difficulties in estimating the state transition probabilities for an industrial situation. It is shown in the previous subsections 5.3.1 and 5.3.2 that such difficulties may be solved using the delay time concept by expressing the degraded states of the system as the expected number of defects. State transition probabilities may then be calculated using the distribution of the delay time h and defect arrival rate λ. To succeed here, the delay time parameters with the known case histories have established that they can be estimated from real-world data or subjective techniques as described in the chapter 3, and later in the chapter 6. Either way, the multi-component system can be fitted to the semi-Markov inspection model utilsing the delay time concept.

Nevertheless, the semi-Markov inspection model for the multi-component system may still have problems in terms of its validity in a specific case. In the subsection 5.3.1, the model assumes that faults arise as a HPP with the rate of occurrence of faults as the constant λ under the perfect inspection policy. Also, from the Lemma 5.1 and 5.2, the semi-Markov inspection model has been established based upon this assumption. From the Lemma 5.1, if the system has an imperfect inspection policy, it is obvious that the number of defects identified at an inspection will be changed. Again, if the system is subject to an imperfect inspection policy, the failure arrival process may not follow a NHPP with the rate function $\nu(x)$ in equation (5.2), because failures arriving in a current inspection period may be affected by faults arising in past inspection periods. Accordingly, it may be unrealistic to apply the semi-Markov inspection model of the subsection 5.3.1 to the system with the imperfect inspection policy. It may be more reasonable in most cases to assume an imperfect inspection policy when modelling industrial situations.

To solve these problems, the semi-Markov inspection model for an imperfect inspection policy is presented in the subsection 5.3.2 under the assumption that the system will be in steady state in the long term future period. The semi-Markov inspection model for an imperfect inspection policy is established on the basis of Lemma 5.5 and 5.6 which are based upon the delay time concept where the fault arrival process follows a HPP. Also, the parameters in the model can be estimated from the delay time concept.

5.4 Delay Time Model

5.4.1 Formulation as a Basic Delay Time Model

The inspection system mentioned in the section 5.2 can also be analysed in the framework of a delay time model to find the optimal inspection period which minimises the expected total cost per unit time or the expected total downtime per unit time. Indeed, this is necessary if the two modelling methodologies are to be compared as intended. To formulate the inspection system mentioned in the section 5.2 as basic delay time model, additional to the description of the section 5.2, we need the following assumptions.

(1) An inspection takes place every T time units and requires C_i cost units and d_i time units.
(2) Inspections are perfect in that any defect present within the system will be identified at inspection, and no new fault inputted because of inspection.
(3) Defects identified at an inspection will be repaired within the allocated inspection time, d_i, and the mean repair cost per defect is C_d units.
(4) A failure will be observed immediately at its occurrence. The component is repaired immediately upon failures and the mean repair cost and time for a failure repair are C_b cost units and d_b time units respectively.
(5) The component is as good as new after repairs.
(6) Defects are independent of each other and arise as a homogeneous Poisson Process (HPP), with rate of occurrence of defects λ.
(7) The delay time h of a defect is independent of the time of origin, and all defects share a common delay time *pdf* $f(h)$ and *cdf* $F(h)$.

Then, as we noted in the section 5.2, an inspection will renew the system.

To introduce the delay time modelling of this system, it is convenient to import the basic delay time model of the chapter 3. Since the instantaneous rate of defect occurrence within the system after an inspection is λ, the number of defects arriving in the interval $(u, u+du)$ is λdu. Clearly, the expected number of defects arising over $(0, T)$ is (see the equation (3.2))

$$N(T) = \int_0^T \lambda du = \lambda T. \tag{5.59}$$

A defect arising in $(u, u+du)$ with a delay time $h<T-u$ will arise as a breakdown (see Figure 5.4).

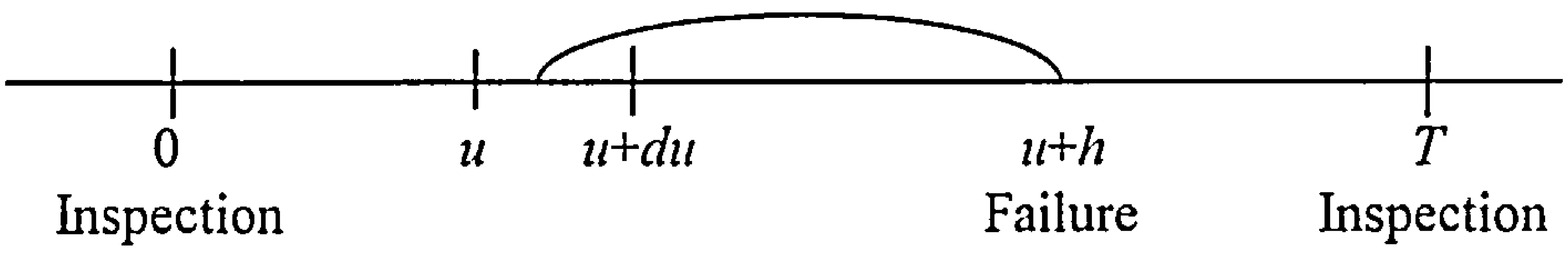


Figure 5.4. Failure process of a defect arising in $(u, u+du)$.

We have, therefore, that the expected number of breakdowns arising over period $(0, T)$ is (see the equation (3.3))

$$B(T) = \int_0^T \lambda F(T-u)du. \tag{5.60}$$

A model of the expected cost per unit time as a function of the inspection period T may be obtained directly. The expected total cost of an inspection cycle consists of the expected cost of attending to failures, the expected cost of rectifying defects identified at inspection, and the cost of the inspection itself. Since the expected number of defects and breakdowns arising over $(0, T)$ is known, namely $N(T)$ and $B(T)$, assuming that the downtime of a failure, d_b, is very small. the expected total cost per unit time over a full cycle of length $T+d_i$ is (see the equation (3.4))

$$C(T) = \frac{1}{T+d_i}(B(T)C_b + (N(T)-B(T))C_d + C_i). \tag{5.61}$$

Here, the decision variable T would be selected to minimise the expected total cost per unit time $C(T)$.

Similarly, for the downtime model, since the expected total downtime of an inspection cycle consists of the expected downtime of attending to failures and the downtime of the inspection itself, assuming that the downtime of a failure, d_b, is very small, the expected total downtime per unit time over a full cycle of length $T + d_i$ is (see the equation (3.5))

$$D(T) = \frac{1}{T + d_i}(B(T)d_b + d_i). \tag{5.62}$$

Here, the decision variable T would be selected to minimise the expected total downtime per unit time $D(T)$.

5.4.2 A Delay Time Model for an Imperfect Inspection Policy

So far, it has been assumed, as in the section 5.2, that inspections are perfect in that any fault present will be identified. However, in most cases, it is more likely that inspections are imperfect. Based upon the assumptions of the section 5.2, suppose, as before in subsection 5.3.2, that there is a probability $r \le 1$ that any fault present at an inspection will be identified at the inspection. It has been shown that , under these circumstances, the probability of a fault leading to a failure, $b(T)$, becomes (see equation (3.11))

$$b(T) = 1 - \int_0^T \sum_{n=1}^{\infty} \frac{r}{T}(1-r)^{n-1} R(nT - y)dy, \tag{5.63}$$

where $R(h)$ = 1-$F(h)$. Then, since the probability of a defect arising as a breakdown is changed, modifying equations (5.61) and (5.62), we have that the expected total cost per unit time is given by

$$C(T) = \frac{N(T)\{b(T)C_b + (1-b(T))C_d\} + C_i}{T + d_i}, \tag{5.64}$$

and the expected total downtime per unit time is given by

$$D(T) = \frac{N(T)b(T)d_b + d_i}{T + d_i}. \tag{5.65}$$

Here, the decision variable T would be selected to minimises the expected total cost per unit time $C(T)$ or the expected total downtime per unit time $D(T)$.

5.4.3 Evaluation of the Delay Time Model

In the previous section, it is shown that the system mentioned in the section 5.2 can be fitted to the semi-Markov inspection model for a multi-component system based upon the delay time concept. The delay time concept provides a means denoting the working condition of the system as the degraded states of the semi-markov inspection model, and obtaining the state transition probability from data in an industrial applications by first estimating the parameters of the delay time distribution. Using the delay time concept in this way, the semi-Markov inspection model is potentially useful in application to a real-world situation. Here we can see the modelling value of the delay time concept.

In applying the semi-Markov inspection model to an actual situation, the key point is that the fault arrival process has a HPP with the rate λ. If the fault arrival process follows a NHPP, we cannot apply the semi-Markov inspection model to the real-world situation. However, in this case, the delay time model can still be used as discussed in the chapter 3. Accordingly, the delay time model is again more robust than the semi-Markov inspection model in applying to the real-world situation. The delay time model may be fitted regardless of the HPP / NHPP status of the fault arrival process.

5.5 Numerical Examples

5.5.1 Generating the Data using Simulation

As discussed in chapter 4, when we are faced with the industrial preventive maintenance problem, the data collection is of prime importance. In attempting to apply models of maintenance to an actual industrial plant, the immediate problem is usually that there are no data to be fitted. Either way, we require data to fit to the models of section 5.3 and 5.4 in order to compare both models. Comparing the two models with each other is only part of the research process here, since both models also ideally need to be compared with the time value, which is unknown. This is a reason for generating the data using computer simulation and a known delay time model. The generated data can then be used to model the process.

By the assumption (8) of the section 5.2, defects are independent of each other and arise as a HPP with the rate λ. If X_n, $n \geq 1$, denotes the time between the $(n\text{-}1)$st and nth defect from the last inspection point (see Figure 5.5), the sequence $\{X_n, n \geq 1\}$ is called the sequence of inter-arrival times and X_n are independent identically distributed exponential random variables having mean $1/\lambda$ (see the proposition 2.2.1 of Ross' [1983]).

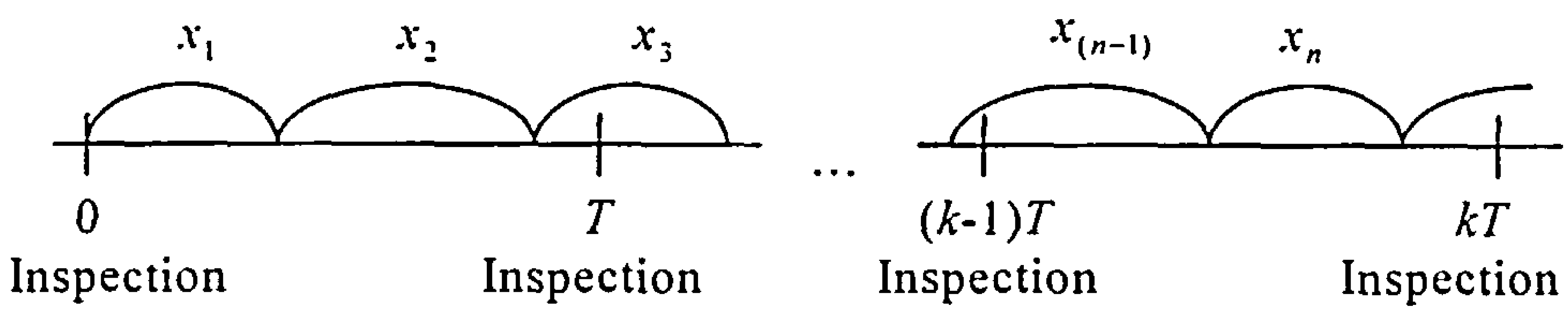
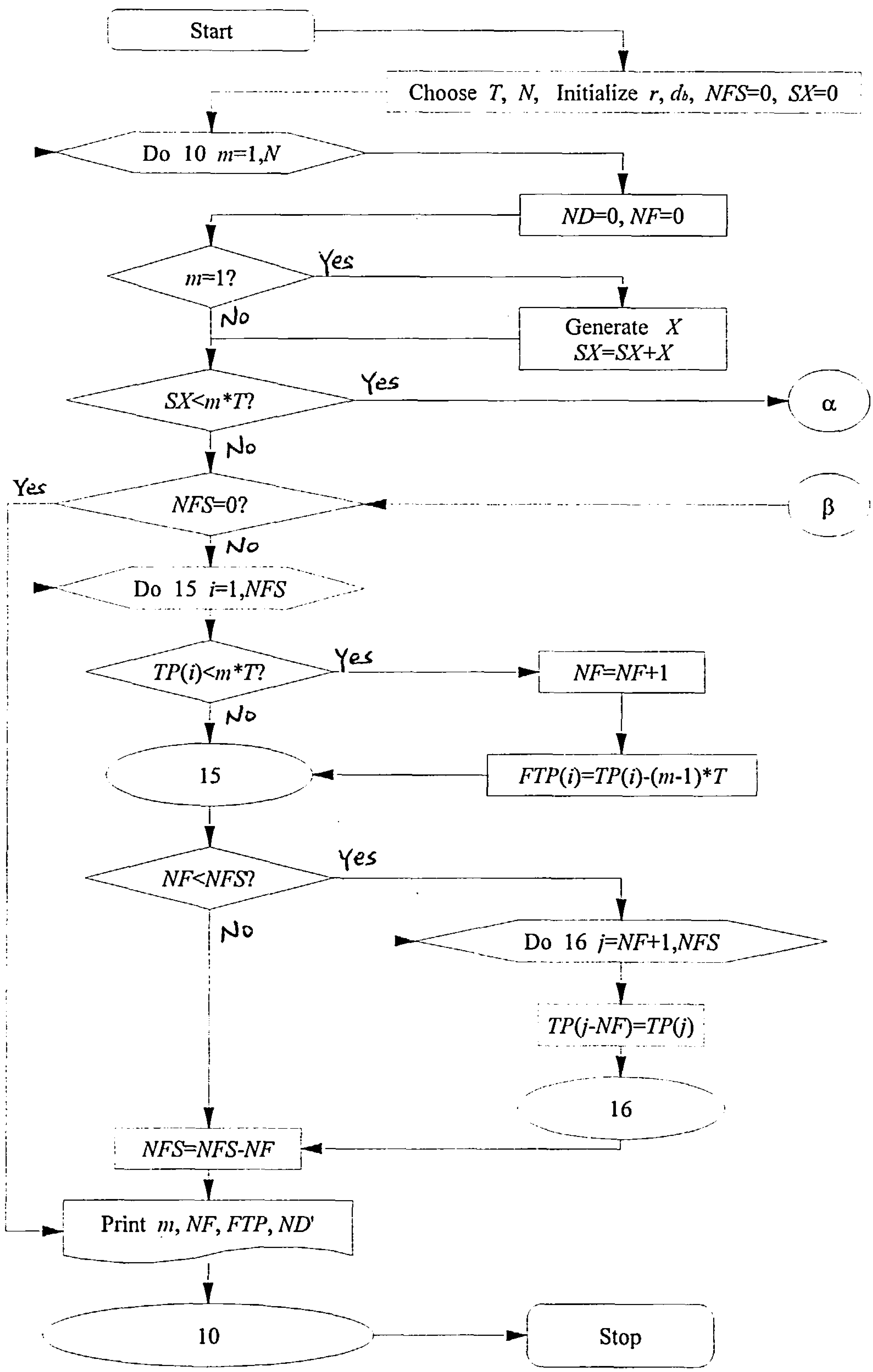


Figure 5.5. The sequence of inter-arrival times.

Noting this inter-arrival time pattern, the system description for modelling of section 5.2, and assumption (3') which is changed in assumption (3) of section 5.2, we can generate a set of synthetic data for an imperfect inspection policy corresponding to this situation using simulation. By setting the probability of identifying a fault at an inspection $r = 1$, we can generate a set of data for a perfect inspection policy. Figure 5.6

illustrate the simulation progress by presenting the flow chart which is for the characteristic structure of the data shown in the Table 5.1.



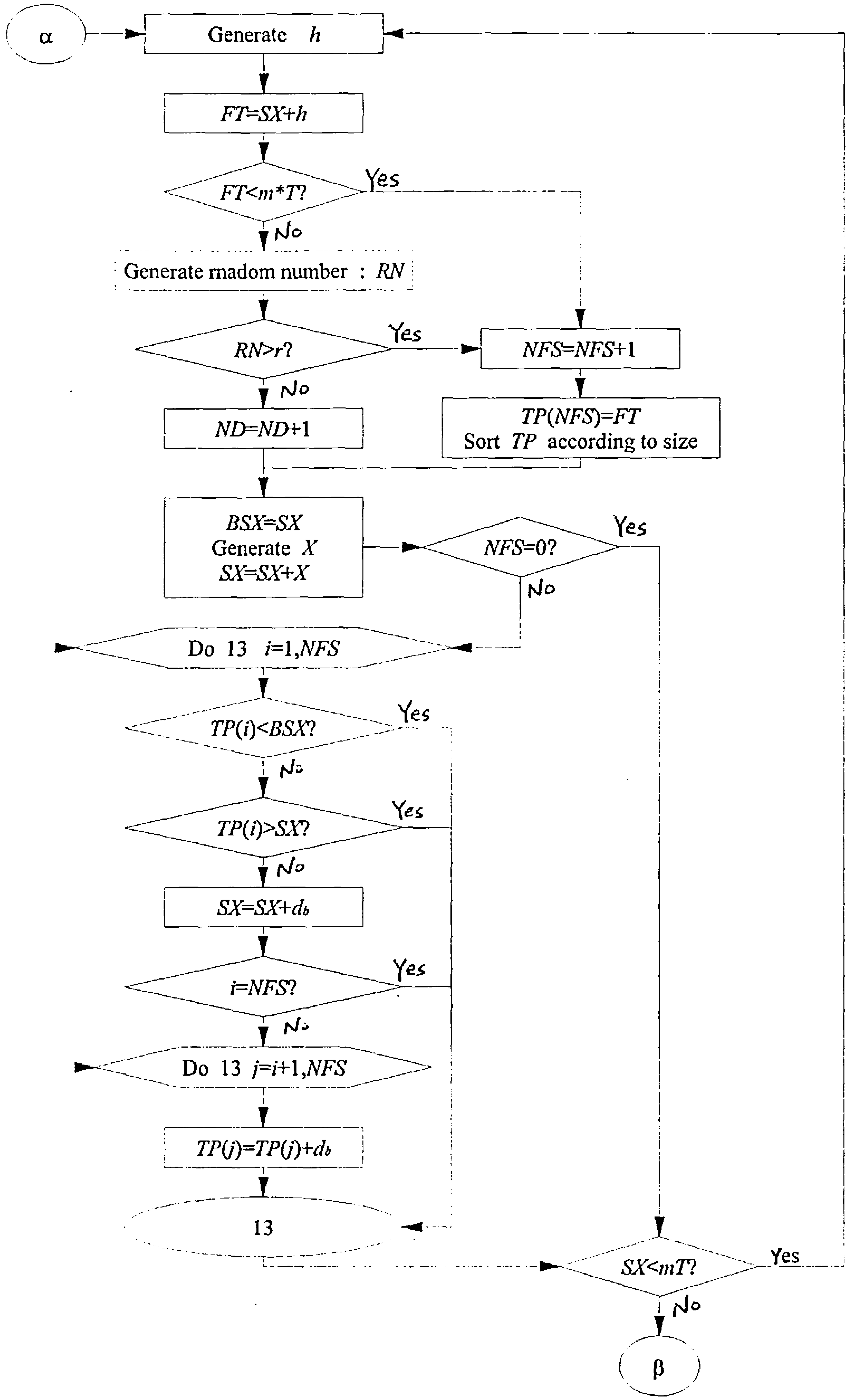


Figure 5.6. Flow chart for generating the data for a multi-component system.

In Figure 5.6, the following notation has been used.

T: A present inspection period.

N : The number of inspection for computer running.

ND : The number of defects arising within a present inspection period.

NF : The number of failures arising within a present inspection period.

X: An inter-arrival time of defect.

SX: The present summation of inter-arrival times.

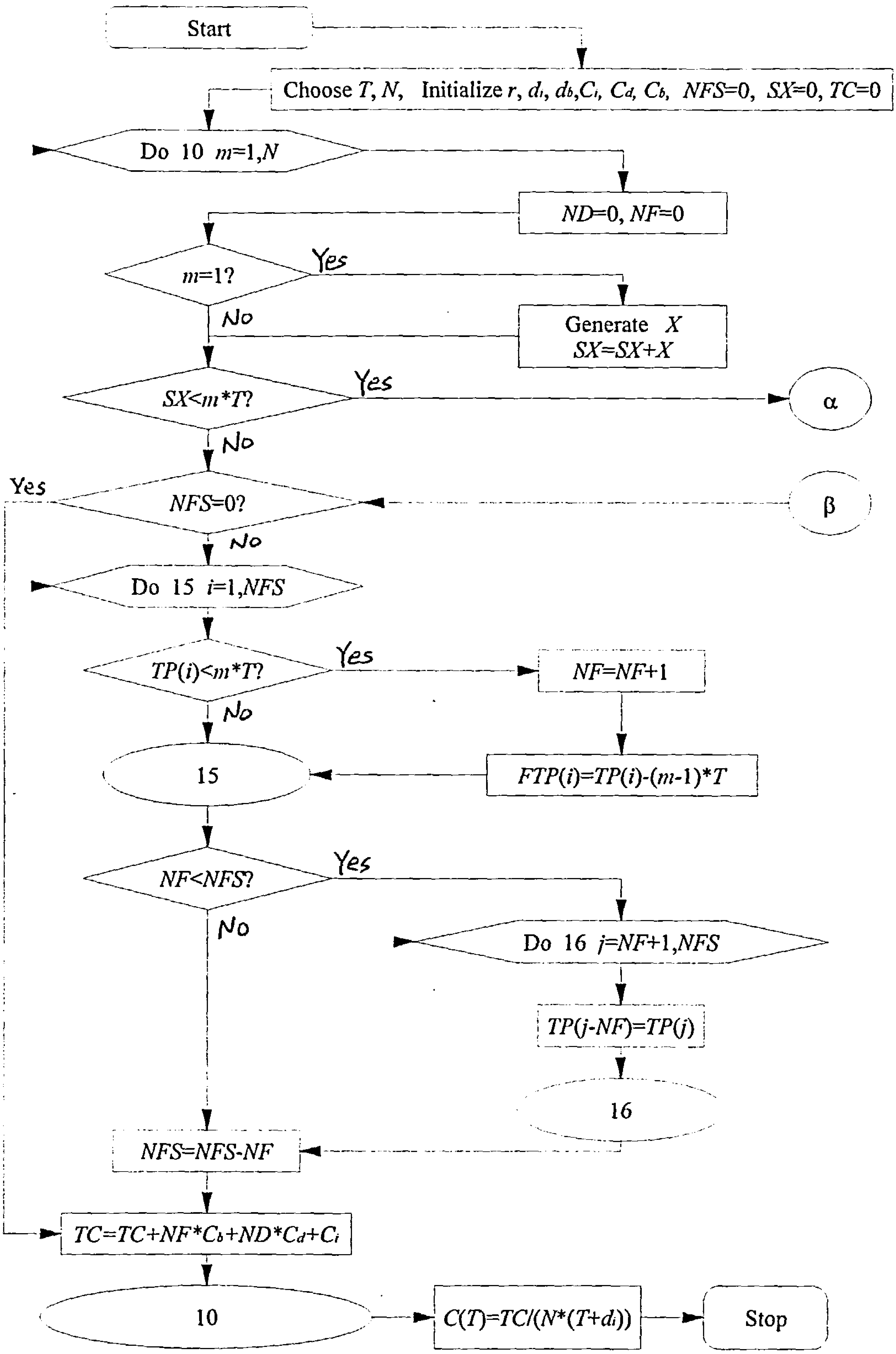
h : A delay time

$TP(j)$: The jth failure time point.

$FTP(j)$: The jth failure time point within a present inspection period.

5.5.2 A Perfect Inspection Case

A numerical example of the cost model outlined in the section 5.3 and 5.4 is evaluated for demonstration purposes. From the semi-Markov inspection model of section 5.3 and the delay time model of the section 5.4, the expected cost per unit time can be determined as a function of the inspection period T, or the decision rule $R = s^*$. This means that we can obtain the optimal inspection period T^* which minimises the expected cost per unit time. Since, however, we wish to know which is the most accurate of the models, and the extent of any difference between them, we compare the two models with a simulation model. It was shown in the flow chart of Figure 5.6 of the previous subsection 5.5.1 that the required data could be obtained for any inspection period T. A minor change to the flow chart of Figure 5.6 transforming it into the flow chart of Figure 5.7 will provide the expected cost per unit time $C(T)$ according to the inspection period T. In Figure 5.7, adding the notations of Figure 5.6, TC denotes the total expected cost. In this subsection, we can get the expected cost per unit time by setting $r = 1$ in Figure 5.7.



Start
Choose T, N, Initialize r, d_i, d_b, C_i, C_d, C_b, NFS=0, SX=0, TC=0
Do 10 m=1,N
ND=0, NF=0
m=1?
Yes
No
Generate X
SX=SX+X
SX<m*T?
Yes
α
No
Yes
NFS=0?
β
No
Do 15 i=1,NFS
TP(i)<m*T?
Yes
NF=NF+1
No
15
FTP(i)=TP(i)-(m-1)*T
NF<NFS?
Yes
No
Do 16 j=NF+1,NFS
TP(j-NF)=TP(j)
16
NFS=NFS-NF
TC=TC+NF*C_b+ND*C_d+C_i
10
C(T)=TC/(N*(T+d_i))
Stop

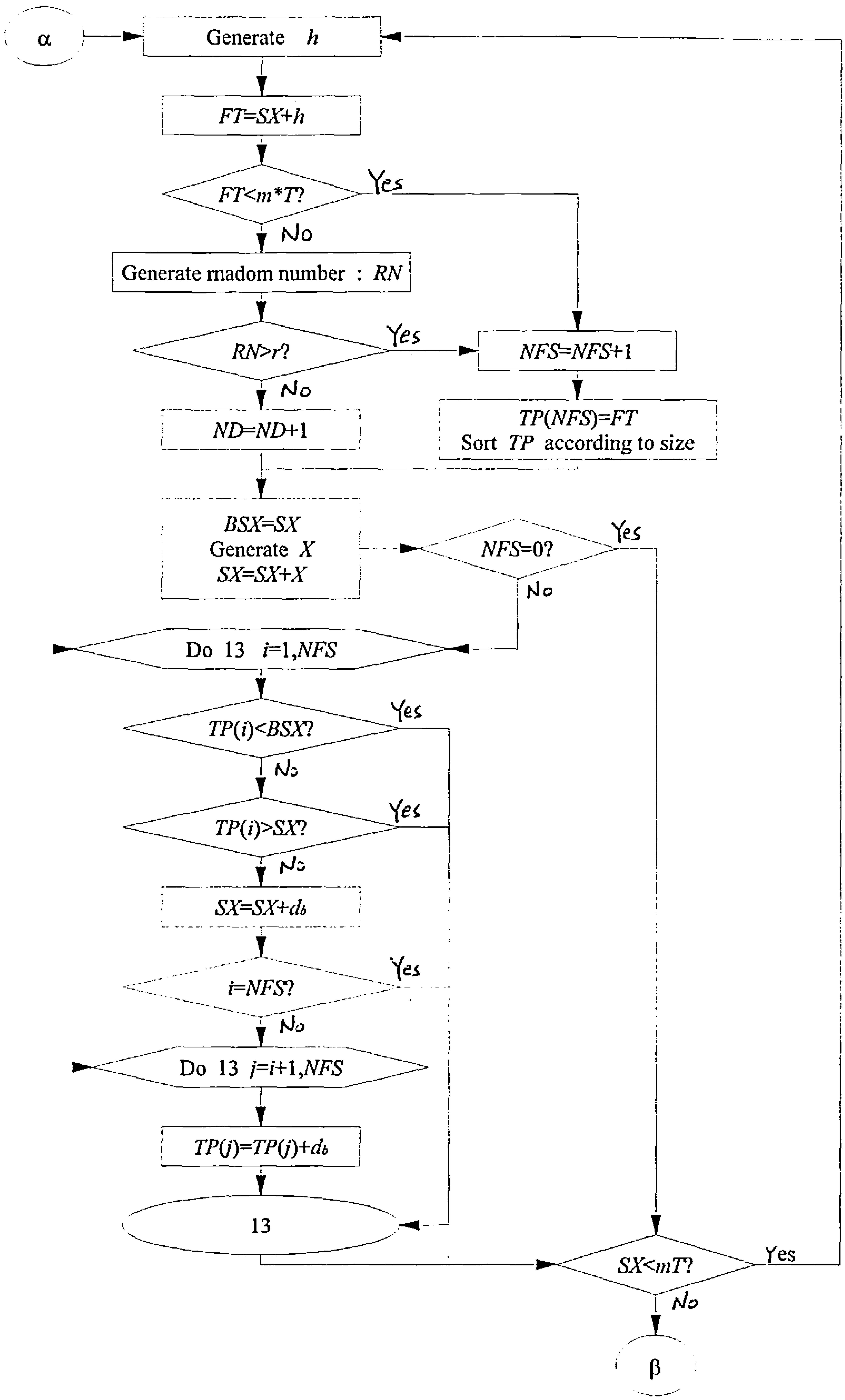


Figure 5.7. Flow chart for computing the expected cost per unit time.

To compare the semi-Markov model with the delay time model, firstly, consider the case where the fault arrival rate has been taken as λ faults per unit time. the delay time h has an exponential distribution with the mean $1/\beta$, $\beta>0$, and the downtime of the failure can be neglect, $d_b = 0$. Since the inter-arrival time X has an exponential distribution with the mean $1/\lambda$, the *pdf* of the inter-arrival time X is given by

$$q(x) = \lambda e^{-\lambda x}. \tag{5.66}$$

Also, we have that the *pdf* of delay time h is given by

$$f(h) = \beta e^{-\beta h} \tag{5.67}$$

and the *cdf* of the delay time h is given by

$$F(h) = 1 - e^{-\beta h}. \tag{5.68}$$

Assuming that the data have been given in the real-world situation, or generated using simulation, the estimates of λ and β have been estimated as $\lambda = 0.3$ and $\beta = 0.1$. Also, costs are taken as $C_i = 10$ units, $C_d = 5$ units, and $C_b = 15$ units, and downtimes for a cost model are taken as $d_i = 0.4$ time units. In the semi-Markov inspection model for the perfect inspection case of subsection 5.3.1, we have assumed that N is the upper bound of the expected number of defects and the time interval Δt is very small. For a numerical example of this subsection, we assume that $N = 30$ and the time interval $\Delta t = 1$. Under these circumstances, the result for the semi-Markov inspection model for a perfect inspection policy of subsection 5.3.1, namely equations (5.23) to (5.28) and equation (5.61) for the basic delay time model for expected cost, is shown in Figure 5.8.

Figure 5.8 shows that the semi-Markov model curve is consistent with the delay time model curve and the simulation curve and an optimal inspection period point which minimises the cost per unit time can be obtained from these three curves. This is as expected because of the Markovian nature of the modelling assumption and services as a check of our numerical procedures. Accordingly, the above mentioned system can be modelled with both a semi-Markov inspection model or a delay time model.

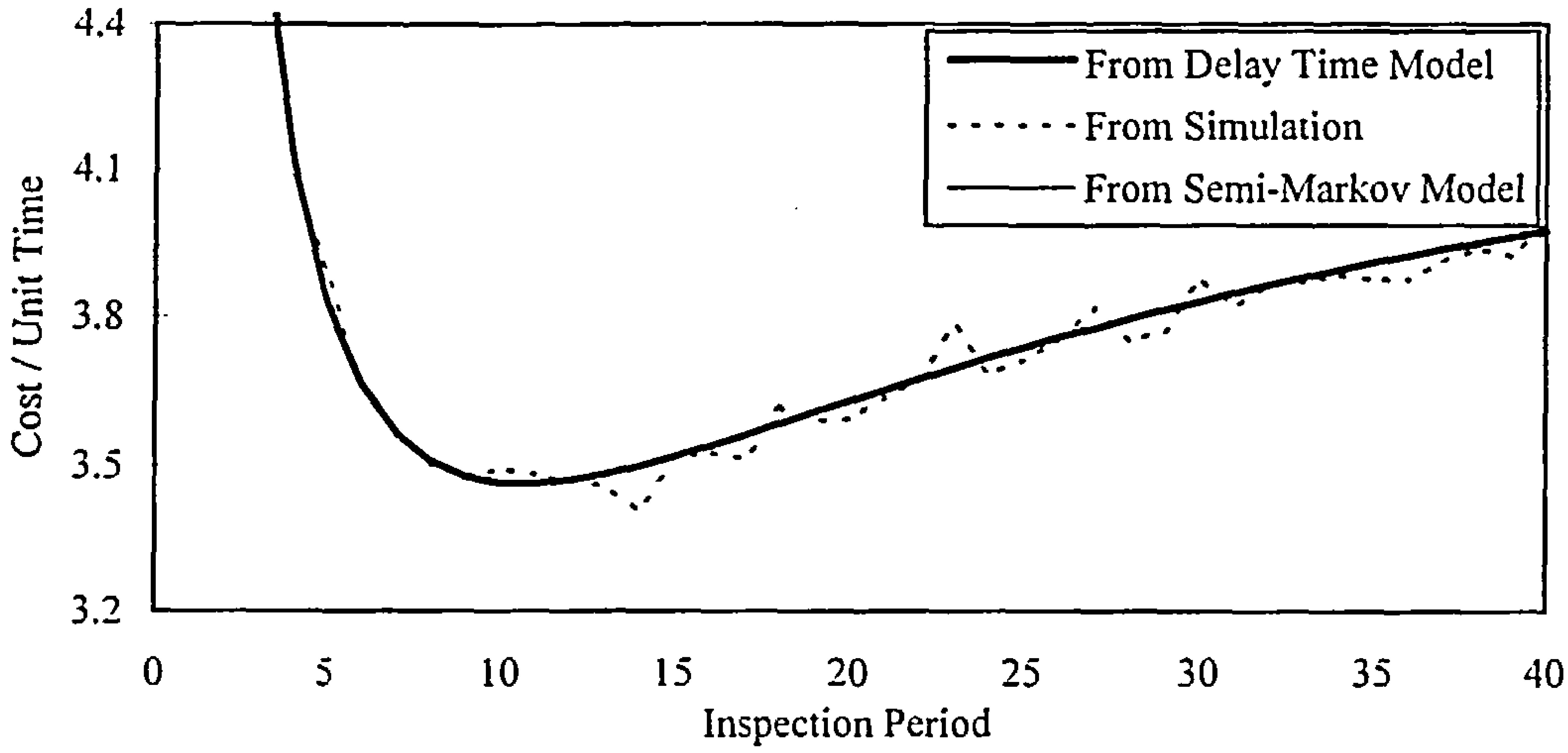


Figure 5.8. The expected cost per unit time according to the inspection period.
(This is for a perfect inspection policy when the delay time has an exponential distribution, Δt=1, λ=0.3, β=0.1, $C_i=10$, $C_d=5$, $C_b=15$, $d_i=0.4$, and $d_b=0$.)

Secondly, under the same conditions as in the above case, consider a case when the downtime of the failure cannot be neglected, $d_b \neq 0$. If we assume that $d_b = 0.05$, the result for the semi-Markov inspection model for a perfect inspection policy of the subsection 5.3.1 and the equation (5.61) of the basic delay time model can be shown in Figure 5.9.

Figure 5.9 shows that there is little difference between the semi-Markov model curve and the delay time model curve, although a difference does now exist. Further, the simulation model curve may be more consistent with the semi-Markov model curve than the delay time model curve. The reason is that the equation (5.61) of the basic delay time model neglects the downtime of a failure. This can be readily corrected if required (see Chilcott and Christer [1991]). However, in getting the optimal inspection period point which minimises the expected cost per unit time, we can obtain the optimal inspection period from the semi-Markov model and from the delay time model, and in this case the optimal inspection period from the delay time model is the same as that from the semi-Markov model (see Table 5.2).

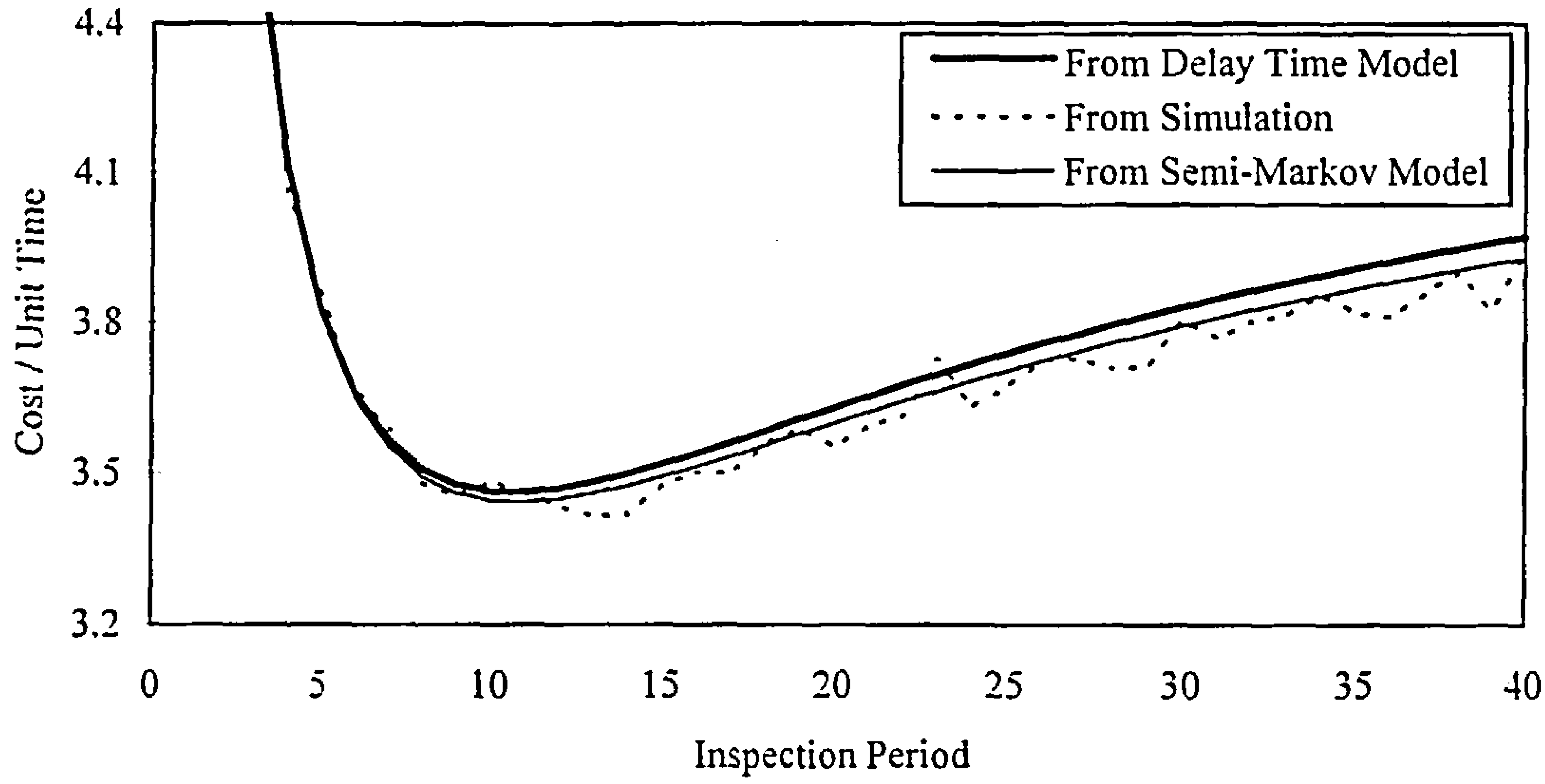


Figure 5.9. The expected cost per unit time according to the inspection period.

(This is for a perfect inspection policy when the delay time has an exponential distribution, Δt=1, λ=0.3, β=0.1, $C_i = 10$, $C_d = 5$, $C_b = 15$, $d_i = 0.4$, and $d_b = 0.05$.)

Table 5.2. Expected cost per unit time for Figure 5.9.

Inspection Period	Delay time model	Simulation model	Semi-Markov model	Remarks
:	:	:	:	:
:	:	:	:	:
7	3.567	3.588	3.553	
8	3.510	3.483	3.494	
9	3.478	3.454	3.461	
10	3.465	3.486	3.447	
11	3.464	3.462	3.444	Optimal
12	3.471	3.436	3.450	
13	3.483	3.414	3.461	
14	3.500	3.419	3.476	
15	3.519	3.477	3.494	
:	:	:	:	:
:	:	:	:	:

To confirm this point clearly, if we increase the downtime of a failure into $d_b = 0.2$, the result is as shown in Figure 5.10.

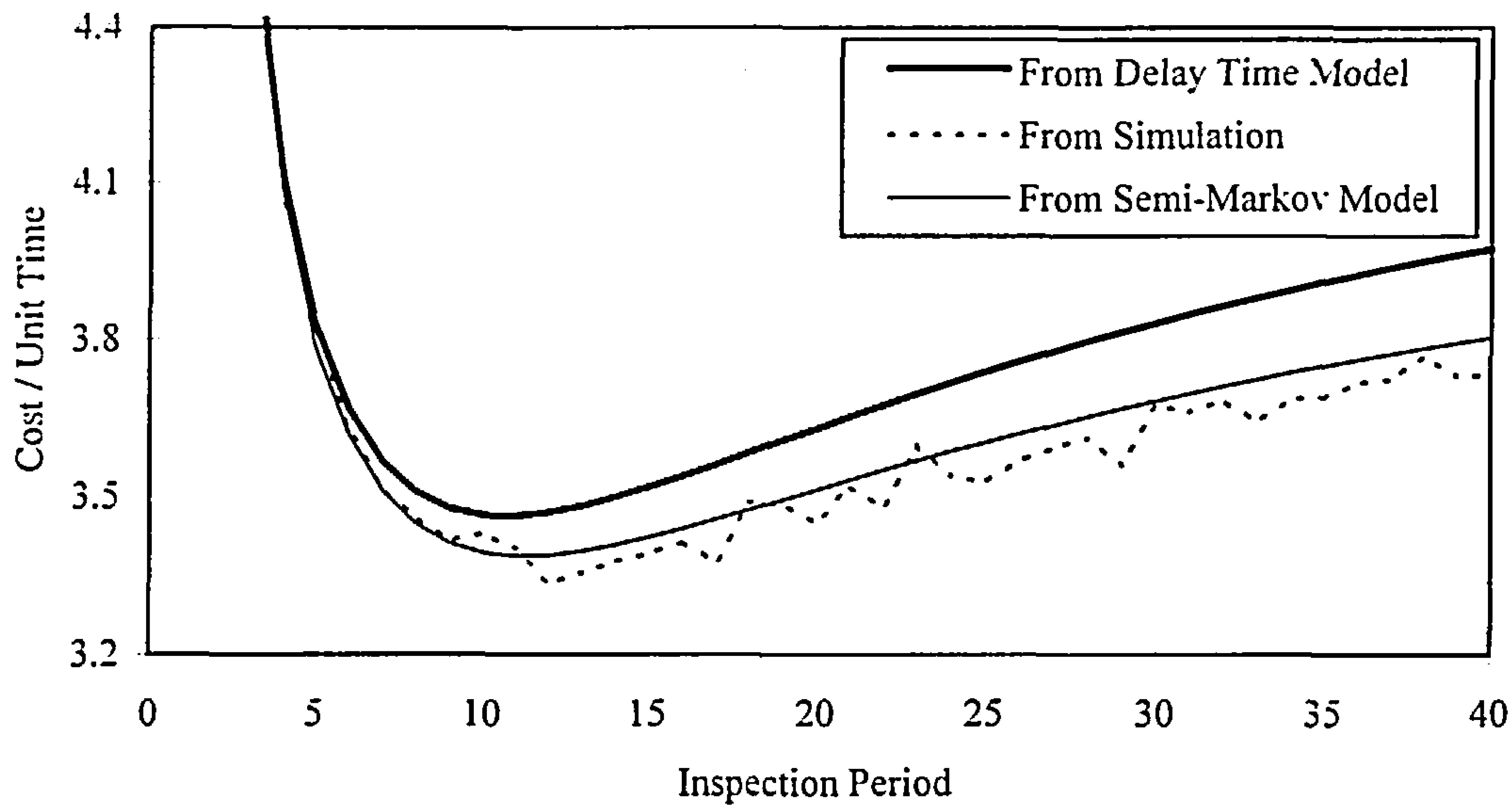


Figure 5.10. The expected cost per unit time according to the inspection period.
(This is for a perfect inspection policy when the delay time has an exponential distribution, Δt=1, λ=0.3, β=0.1, $C_i = 10$, $C_d = 5$, $C_b = 15$, $d_i = 0.4$, and $d_b = 0.2$.)

Figure 5.10 shows that there is the bigger difference between the semi-Markov model curve and the delay time model curve than Figure 5.9. Also, the simulation model output is more consistent with the semi-Markov model than the delay time model. However, from both models, the optimal inspection which minimises the expected cost per unit time can be obtained, and the inspection period choice resulting from the delay time model is the same inspection period resulting from the semi-Markov model line (see Table 5.3).

Table 5.3. Expected cost per unit time for Figure 5.10.

Inspection Period	Delay time model	Simulation model	Semi-Markov model	Remarks
:	:	:	:	:
:	:	:	:	:
7	3.567	3.515	3.511	
8	3.510	3.461	3.448	
9	3.478	3.416	3.412	
10	3.465	3.430	3.393	
11	3.464	3.400	3.387	Optimal
12	3.471	3.335	3.388	
13	3.483	3.356	3.396	
14	3.500	3.376	3.408	
15	3.519	3.392	3.423	
:	:	:	:	:
:	:	:	:	:

Thirdly, under the same conditions of the above second case, consider the case that the delay time h has a Weibull distribution with the shape parameter $\alpha>0$ and the scale parameter $\beta>0$. We have that the *pdf* of the delay time h is given by

$$f(h) = \alpha\beta^{-\alpha}h^{\alpha-1}e^{-(\frac{h}{\beta})^{\alpha}} \qquad (5.69)$$

and the *cdf* of the delay time h is given by

$$F(h) = 1 - e^{-(\frac{h}{\beta})^{\alpha}}. \qquad (5.70)$$

Assuming that the data have been given in the real-world situation or using simulation, α and β have been estimated as $\alpha = 0.9$ and $\beta = 9.0$. If we assume that $d_b = 0.05$, the result for the semi-Markov inspection model for a perfect inspection policy of the subsection 5.3.1 and the equation (5.61) of the basic delay time model can be shown in Figure 5.11.

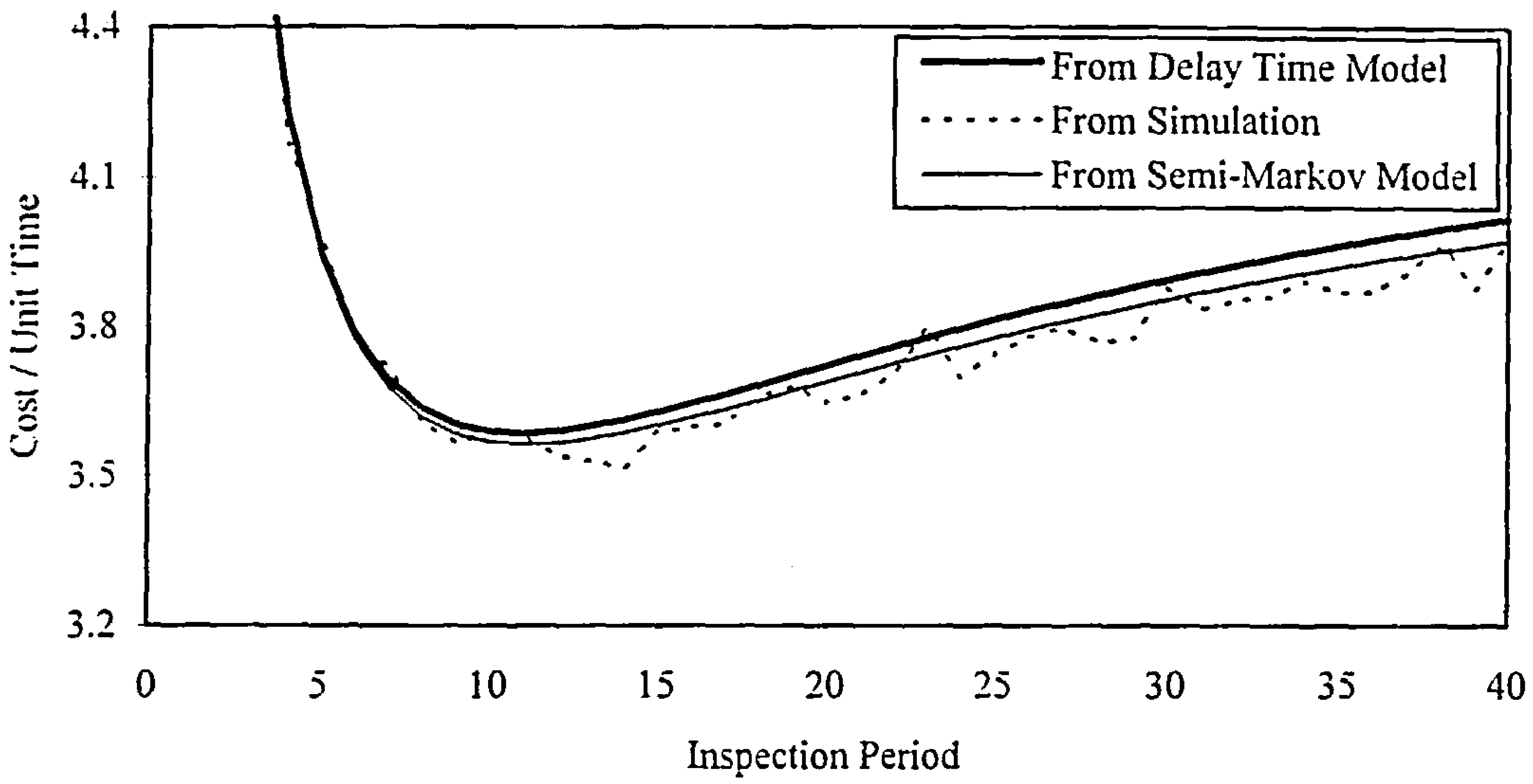


Figure 5.11. The expected cost per unit time according to the inspection period.

(This is for a perfect inspection policy when the delay time has a Weibull distribution, Δt=1, λ=0.3, α=0.9, β=9, $C_i = 10$, $C_d = 5$, $C_b = 15$, $d_i = 0.4$, and $d_b = 0.05$.)

Figure 5.11 shows that there is still a little difference between the semi-Markov model curve and the delay time model curve, and that the simulation model may be more consistent with the semi-Markov model curve than the delay time model curve. The two models would not move closer together if Δt were to decrease from $\Delta t = 1$. The simulation output confirm this. However, in determining the optimal inspection period, the delay time model and the semi-Markov model do, to the precision of the measure scale, again indicate the same decision (see Table 5.4).

Table 5.4. Expected cost per unit time for Figure 5.11.

Inspection Period	Delay time model	Simulation model	Semi-Markov model	Remarks
:	:	:	:	:
:	:	:	:	:
9	3.606	3.568	3.586	
10	3.591	3.588	3.570	
11	3.587	3.585	3.565	Optimal
12	3.591	3.543	3.567	
13	3.601	3.530	3.576	
:	:	:	:	:
:	:	:	:	:

5.5.3 An Imperfect Inspection Case

In the previous subsection, we have discussed the case when the fault arrival rate obeyed a HPP and inspections were perfect. The discussion of the previous subsection confirmed that, regardless of the distribution of the delay time h, when the fault arrival process is a HPP, for the perfect inspection case, the system can be represented by both a semi-Markov inspection model and a delay time model. In this subsection, we consider an imperfect inspection case in order to check that the system can be fitted to the semi-Markov inspection model and the delay time model. As before, if we take the simulation model to compare fairly the semi-Markov inspection model and the delay time model, we can use the flow chart of Figure 5.7 for the simulation progress.

Firstly, consider a case where the delay time has an exponential distribution of equation (5.68). In this case, for a delay time for an imperfect inspection policy of the subsection 5.4.2, the probability of a fault leading to a failure, $b(T)$, of the equation (5.63) becomes

$$b(T) = 1 - \int_0^T \sum_{n=1}^{\infty} \frac{r}{T}(1-r)^{n-1}\{1-(1-e^{-\beta(nT-y)})\}dy$$

$$= 1 - \frac{r(1-e^{-\beta T})}{\beta T(1-(1-r)e^{-\beta T})}. \qquad (5.71)$$

Applying the equation (5.64) of the delay time model for an imperfect inspection policy based upon the equation (5.71), we get the expected cost per unit time $C(T)$ according to the inspection period T. Also, for a semi-Markov inspection model for an imperfect inspection policy of subsection 5.3.2, the number of faults identified at time x from an inspection, $\mathrm{E}N_d(x)$, of the equation (5.56) becomes

$$\mathrm{E}N_d(x) = \lim_{i\to\infty} \lambda r[\sum_{n=1}^{i-1}(1-r)^{i-n}\int_{T_{n-1}}^{T_n}(1-(1-e^{-\beta(T_{i-1}+x-y)}))dy$$

$$+ \int_{T_{i-1}}^{T_{i-1}+x}(1-(1-e^{-\beta(T_{i-1}+x-y)}))dy]$$

$$= \frac{\lambda r}{\beta}\left(\frac{(1-r)(1-e^{\beta T})e^{-\beta x}}{1-r-e^{\beta T}} + 1 - e^{-\beta x}\right) \tag{5.72}$$

and the failure arrival rate function at time x from an inspection, $\rho(x)$, of the equation (5.57) becomes

$$\rho(x) = \lim_{i \to \infty} \lambda \{ \sum_{n=1}^{i-1} (1-r)^{i-n} ((1-e^{-\beta(T_{i-1}+x-T_{n-1})}) - (1-e^{-\beta(T_{i-1}+x-T_n)})) + (1-e^{-\beta x}) \}$$

$$= \lambda \{ \frac{(1-r)(1-e^{\beta T})e^{-\beta x}}{1-r-e^{\beta T}} + 1 - e^{-\beta x} \}. \tag{5.73}$$

Applying the semi-Markov inspection model for an imperfect inspection policy of the subsection 5.3.2 based upon the equations (5.72) and (5.73), we can get the expected cost per unit time $g(T)$ according to the inspection period T.

Given the data from the real-world situation or using simulation, we assume that the parameters for the distributions of the inter-arrival time X and the delay time h have been estimated as $\lambda = 0.3$ and $\beta = 0.1$ respectively and the probability of identifying a fault at an inspection has been estimated as $r = 0.7$. Also, as before, costs are taken by $C_i = 10$ units, $C_d = 5$ units, and $C_b = 15$ units and downtimes for a cost model are taken by $d_i = 0.4$ time units and $d_b = 0.05$ time units. Also, for the numerical example of this subsection, we assume that $N = 30$ and the time interval $\Delta t = 1$ in the semi-Markov inspection model for an imperfect inspection model of the subsection 5.3.2. Under these circumstances, the result for the semi-Markov inspection model for an imperfect inspection policy of the subsection 5.3.2, the delay time model of the equation (5.64), and the simulation model is shown in Figure 5.12.

Figure 5.12 shows that the semi-Markov inspection model is consistent with the delay time model and the simulation curve and an optimal inspection period point which minimises the cost per unit time can be obtained from these three curves. Accordingly, the mentioned system for an imperfect inspection policy can be fitted to the semi-Markov model or the delay time model. The difference between the presented simulation model and the analytic models of Figure 5.12 is due to the rather large average failure repair time $d_b = 0.05$ of the simulation. Had d_b been reduced, the deviation of the simulation model to the analytic models would here been much reduced for the smaller time periods.

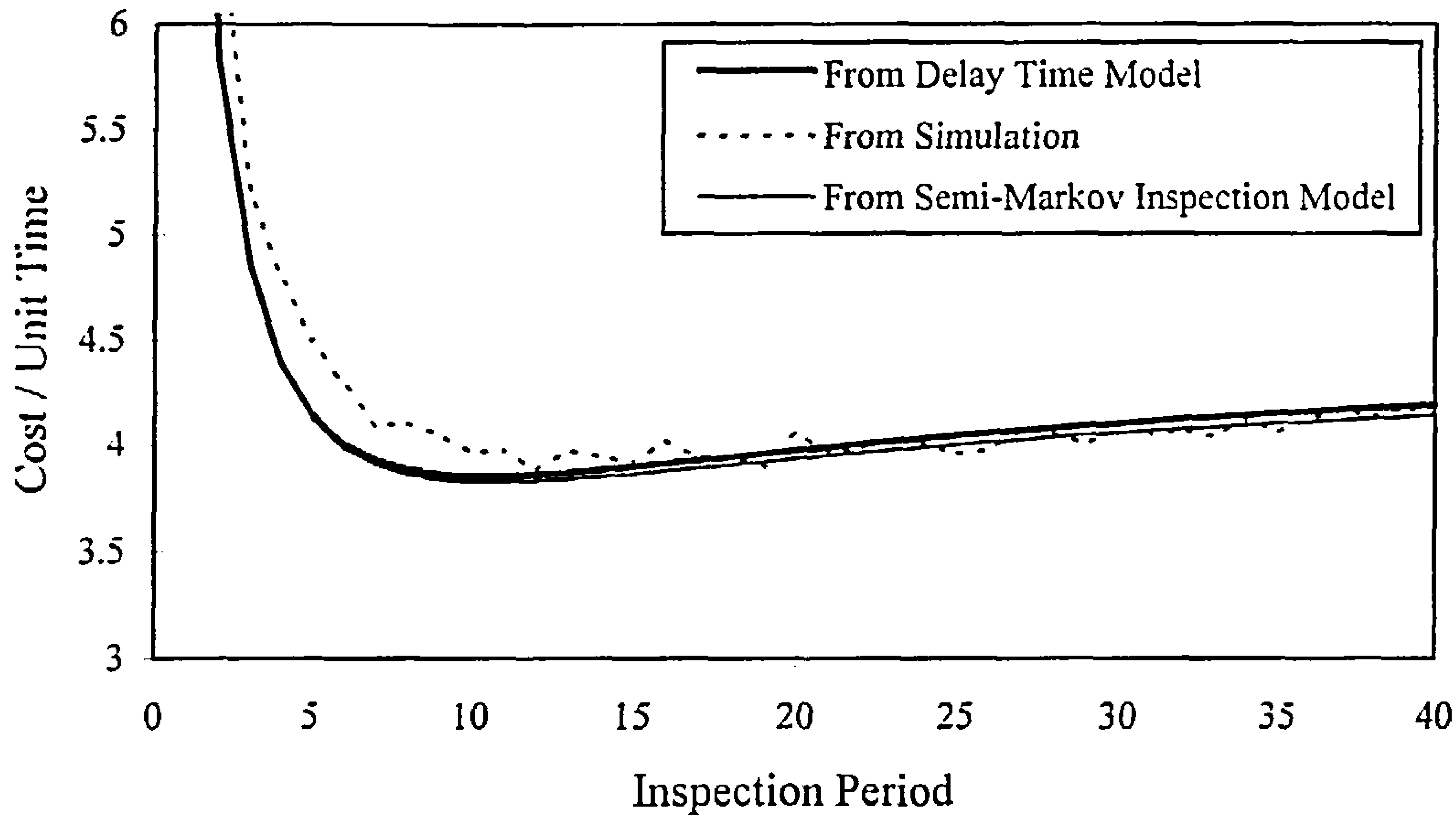


Figure 5.12. The expected cost per unit time according to the inspection period.

(This is for an imperfect inspection policy when the delay time has an exponential distribution, Δt=1, λ=0.3, β=0.1, r=0.7, $C_i = 10$, $C_d = 5$, $C_b = 15$, $d_i = 0.4$, and $d_b = 0.05$.)

Secondly, consider a case that the delay time has a Weibull distribution of the equation (5.70). In this case, for a delay time for an imperfect inspection policy of the subsection 5.4.2, the probability of a fault leading to a failure, $b(T)$, of the equation (5.63) becomes

$$b(T) = 1 - \int_0^T \sum_{n=1}^{\infty} \frac{r}{T}(1-r)^{n-1}\{1-(1-e^{-(\frac{nT-y}{\beta})^\alpha})\}dy$$

$$= 1 - \sum_{n=1}^{\infty} \frac{r}{T}(1-r)^{n-1} \int_0^T e^{-(\frac{nT-y}{\beta})^\alpha} dy. \quad (5.74)$$

Here the equation (5.74) can be obtained numerically using the computer. Applying the equation (5.64) of the delay time model for an imperfect inspection policy based upon the equation (5.74), we can get the expected cost per unit time $C(T)$ according to the inspection period T. Also, for a semi-Markov inspection model for an imperfect inspection policy of the subsection 5.3.2, the number of faults identified at time x from an inspection, $\mathrm{E}N_d(x)$, of the equation (5.56) becomes

$$\begin{aligned}
\mathrm{E}N_d(x) &= \lim_{i\to\infty} \lambda r[\sum_{n=1}^{i-1}(1-r)^{i-n}\int_{T_{n-1}}^{T_n}(1-(1-e^{-(\frac{T_{i-1}+x-y}{\beta})^{\alpha}}))dy \\
&\quad + \int_{T_{i-1}}^{T_{i-1}+x}(1-(1-e^{-(\frac{T_{i-1}+x-y}{\beta})^{\alpha}}))dy] \\
&= \lambda r \lim_{i\to\infty}[\sum_{n=1}^{i-1}(1-r)^{i-n}\int_{T_{n-1}}^{T_n} e^{-(\frac{T_{i-1}+x-y}{\beta})^{\alpha}}dy + \int_{T_{i-1}}^{T_{i-1}+x} e^{-(\frac{T_{i-1}+x-y}{\beta})^{\alpha}}dy] \qquad (5.75)
\end{aligned}$$

and the failure arrival rate function at time x from an inspection, $\rho(x)$, of the equation (5.57) becomes

$$\begin{aligned}
\rho(x) &= \lim_{i\to\infty}\lambda\{\sum_{n=1}^{i-1}(1-r)^{i-n}((1-e^{-(\frac{T_{i-1}+x-T_{n-1}}{\beta})^{\alpha}})-(1-e^{-(\frac{T_{i-1}+x-T_n}{\beta})^{\alpha}}))+(1-e^{-(\frac{x}{\beta})^{\alpha}})\} \\
&= \lambda \lim_{i\to\infty}\sum_{n=1}^{i-1}(1-r)^{i-n}(e^{-(\frac{T_{i-1}+x-T_n}{\beta})^{\alpha}}-e^{-(\frac{T_{i-1}+x-T_{n-1}}{\beta})^{\alpha}})+\lambda(1-e^{-(\frac{x}{\beta})^{\alpha}}). \qquad (5.76)
\end{aligned}$$

The equations (5.75) and (5.76) can also be obtained from the numerical method. Applying the semi-Markov inspection model for an imperfect inspection policy of the subsection 5.3.2 based upon the equations (5.75) and (5.76), we can get the expected cost per unit time $g(T)$ according to the inspection period T.

Given the data from the real-world situation or using simulated data, we assume that the parameters for the distributions of the inter-arrival time X and the delay time h have been estimated as $\lambda = 0.3$, $\alpha = 2.0$, and $\beta = 10$ respectively and the probability of identifying a fault at an inspection has been estimated as $r = 0.7$. Also, as before, costs are taken as $C_i = 10$ units, $C_d = 5$ units, and $C_b = 15$ units and downtimes for a cost model are taken as $d_i = 0.4$ time units and $d_b = 0.05$ time units. Under these circumstances, the result for the semi-Markov inspection model for an imperfect inspection policy of the subsection 5.3.2, the delay time model of the equation (5.64), and the simulation model is shown in Figure 5.13.

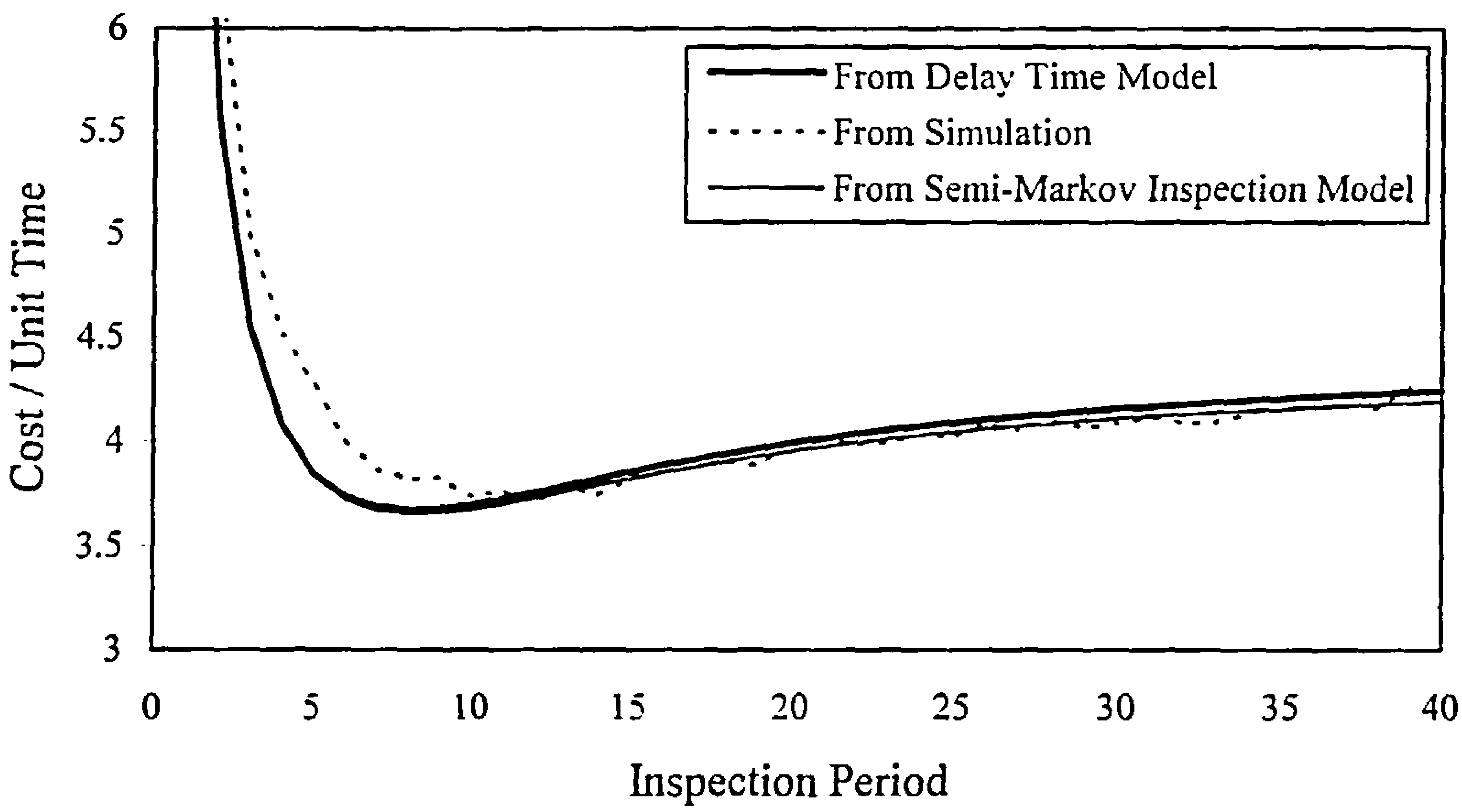


Figure 5.13. The expected cost per unit time according to the inspection period.

(This is for an imperfect inspection policy when the delay time has a Weibull distribution, Δt=1, λ=0.3, α=2.0, β=10, r=0.7. $C_i = 10$, $C_d = 5$, $C_b = 15$, $d_i = 0.4$, and $d_b = 0.05$.)

Figure 5.13 shows that the semi-Markov inspection model is consistent with the delay time model and the simulation curve and an optimal inspection period which minimises the cost per unit time can be obtained from these three curves. Accordingly, an imperfect inspection policy for the given multi-component system can be modelled by either the semi-Markov model or the delay time model.

5.6 Comparison and Conclusions

The single component system can only be fitted to the semi-Markov inspection model in the case that the initial point u has a Markov property. As well as the semi-Markov inspection model for a single component system, if the fault arriving process follows a HPP for a complex system, the semi-Markov inspection model for a multi-component system can be applied to the system regardless of the distribution of the delay time h, as discussed in the subsection 5.5.2. We note that the models here rely on the delay time concept. The delay time concept provides a means of not only denoting the working condition of the system as the degraded states of the semi-Markov inspection model, but also of obtaining the state transition probabilities from data by first estimating the fault arrival rate and the parameters of the delay time distribution. The semi-Markov inspection model based upon the delay time concept can, in this way, become useful in modelling real-world situations. The importance of the delay time concept is evident.

When we establish the semi-Markov inspection model, we have to note Lemmas 5.1 to 5.6 which are based upon the delay time concept. Then, after some complicated manipulation, we have the semi-Markov inspection model of the section 5.3 which is fitted to the multi-component system with a perfect inspection policy or an imperfect inspection policy. The equations of the semi-Markov inspection model, as shown in the section 5.3, are complicated. Therefore, when compared to the delay time model, it takes a larger time to compute the solution of the semi-Markov inspection model using the personal computer. This time depends upon the upper bound of the number of defects N and the arbitrary small time period Δt. In the numerical examples of the section 5.5, we assumed that the upper bound of the number of defects is $N = 30$ and the arbitrary small time period is $\Delta t = 1$. However, since, in the real-world situation, the multi-component system may consist of many components and need the arbitrary smaller time, there may be the more difficulties in computing the semi-Markov inspection model using the personal computer. Also, in applying the semi-Markov inspection model to the real-world situation, the key point is that the fault arrival process follows a HPP which satisfy the Lemmas 5.1 to 5.6. If the fault arrival process follows a NHPP, we cannot apply the semi-Markov inspection model to the real-world situation because Lemma 5.1 to 5.6 are not satisfied and the model is not robust to these requirements. It is noted that there is a greater applicability of semi-Markov models to complex system maintenance than to component maintenance. This is due to a fundamental difference between the consequences of a breakdown in the two cases. For a component, a breakdown represents a system renewal here, whereas for a system, the defect is repaired, and the system continues to age as before.

In real-world applications, the delay time model is consistent with the semi-Markov inspection model as discussed in the section 5.5. However, in contrast to the semi-Markov inspection model, the delay time model consists of the simpler equations as shown in the section 5.4. Therefore, it does not require a long time compared to the semi-Markov inspection model to compute the equations of the delay time model using a personal computer. Also, the delay time model provides a means of modelling the behaviour of the system and predicting such useful quantities as reliability or cost under various inspection policies. If the fault arrival rate, regardless of a HPP or a NHPP, and the parameters of the *pdf* of the delay time h, $f(h)$, regardless of any distributions, are estimated from the data of the real-world situation through either the subjective or objective estimation method, we can easily establish the delay time model which can be used to find the optimal inspection policies minimising the expected total cost per unit time or the expected total downtime per unit time. As confirmed in the numerical example of the section 5.5, it was seen that the simulation model curve and semi-Markov model are nearly consistent with the delay time model curve. This means that the delay time model can be practically applied to the multi-component system.

In conclusion, since the delay time model is free from constraints on the fault arrival process, and does not require as much time as the semi-Markov inspection model in computing the solution of the delay time model using personal computer, the delay time model may be more general than the semi-Markov inspection model in applying to the real-world situation.

Chapter 6

A CASE STUDY OF SEMI-MARKOV AND DELAY TIME MODEL

6.1 Introduction

This chapter reports on a case study of the semi-Markov and delay time models of maintenance applied to a subsystem of a complex machine used within a vehicle brake lining manufacturers, namely a Preformer in a production machine. One of the objectives and motivations behind this case study is to build a model for both identifying and quantifying ways of improving the overall efficiency of the scheme of preventive maintenance (PM) for the Preformer. The other objective in this study is to develop and check the applicability of the semi-Markov and delay time models of maintenance of the previous chapter, with particular focus on the estimation of values of parameters of the models, the compatibility of the results, and the general applications of the modelling methods.

One of the key issues in the semi-Markov and delay time models of maintenance based upon the delay time concept is the estimation of parameters which is usually the rate of occurrence of faults, the distribution of the underlying delay time h of a fault, and the probability of identifying and removing a defect at inspection. The method initially developed for this purpose is called the subjective method, because estimates are obtained from the synthesis of numerous subjective opinions of engineers collected at maintenance intervention. It has been observed that applied studies of the delay time concept were initiated by Christer & Waller [1984a, b], Chilcott & Christer [1991], and Desa [1995] within industrial situations. In all these studies, the probability density function of the delay time was established through subjective estimates derived in well structured situations.

A recent development in delay time modelling has established that these parameters can also be estimated using objective data which are from maintenance records of failures and any faults found at PM. This estimating method has been termed the objective method. In the papers by Baker & Wang [1992, 1993], the objective method was initially designed for a single component subject to failures and inspections at PM, or a system with a few key components. For complex machinery with many components, some modelling has been carried out for actual plant using the delay time concept in the paper by Christer *et al.* [1995]. They developed a model which was based upon the stochastic process of the fault-initiating process and the interval data.

Here we develop a model which is different from previous delay time models. This model deals with a case where historic data exists recording failure time points and PM times, but they are no record of the result of a PM. So, we will use the objective method for the failure data and the subjective method for the inspection data in order to estimate the parameters. Then, PM models are developed to reduce the expected total downtime using both the semi-Markov and the delay time models of maintenance. As this is a case-related study, numerical examples are given throughout to demonstrate the modelling ideas.

6.2 The Production Plant and Maintenance Practice

The formal name of the company collaborating in the research study is "Brakelinings Limited" which produce truck brake linings in the Northwest of England. The company has the production machine comprising of a set of subsystems, which are Preformer, Lift table 1, Lift table 2, Dies, Die and Trans, Hot and Press, Conveyor, and others. The production machine is key plant in the factory, which is operated 24 hours a day (3 shifts) for 5.5 days a week excluding public holidays and maintenance downtimes.

At the time of the study, in order to reduce downtime, a system called total productive maintenance (TPM) has been implemented. TPM was introduced in Japan more than ten years ago and has since found wide acceptance. TPM is implemented by all employees, and is based on the principle that equipment improvement must involve everyone in the organisation, from line operators to top management. The key innovation in TPM is that operators perform basic maintenance on their own equipment. They maintain their machines in good running order and develop the ability to detect potential problems beforê they generate breakdowns. Therefore, we can regard the TPM as a kind of preventive maintenance (PM). In the company, TPM very much takes the form of a PM activity, it requires about 6 hours, and is performed approximately once per three weeks. We have, however, no record of the performance or the results of the TPM. It is evident that the TPM as currently operated is some what short of the ideal of Nakajima [1989].

The company's objective is to reduce the downtime caused jointly by failures and TPM activities, and thereby increase the availability of the machine. The company supposed that the operators and maintenance engineers will find and rectify faults on the machine at TPM within the TPM downtime, but experience shows that there are still failures occurring immediately after TPM. The relevant questions of concern are

(1) whether TPM can identify most faults present and reduce the number of failures caused by those faults, and
(2) whether the current TPM period is the right choice, particularly, the three-week TPM cycle for the Preformer subsystem.

To establish the relationship between the downtime measure and TPM activities using the delay time concept, the first task is to estimate parameters. These are the rate of occurrence of faults, the underlying delay time distribution, and the probability of identifying and removing a fault. It is then possible build a model to describe the interplay of the failure, downtime, and TPM process.

6.3 Data Analysis

6.3.1 The Available Data

The available data over a recent 13 month period include dates and times of downtime occurrences due to both TPM and failures, along with a brief description of its nature. Production records provide the downtimes associated with each failure. The data normally relate to a 24 hour operating day (three shifts), 5.5 operating days per week, but the company sometimes has 6 operating days per week.

The production machine is divided to three sections for TPM, each of which include three subsystems. According to the recorded data, the TPMs are performed approximately every three weeks, but TPM periods are sometimes 4, 5, or 6 weeks. In an initial investigation of the effectiveness of the TPM practice, we use as a measure of plant performance the average number of breakdowns per day or week since TPM for the subsystem.

Also, since full observation of defects identified and rectified at TPM are not recorded, we need to estimate the probability of identifying and removing the defect during a TPM activity using the subjective method. Such estimates are obtained from knowledgeable engineering staff by asking the following questions;

(a) Suppose 100 defects are present at TPM. How many of them could be identified at the TPM and before they led to a failure?
(b) Suppose 100 defects are identified at TPM. How many of them would be rectified at or immediately after the TPM?

From the above questions, the probability of identifying and removing the defect during the TPM, q, can be given by the following equation.

$$\hat{q} = \frac{\text{No. ... ects identified at TPM}}{100} \times \frac{\text{No. c ...}}{...} \quad \frac{\text{... M}}{...} . \quad (6.1)$$

6.3.2 Th... umber ... According to the Subsystem

... 5. 12. ... 21. 1. ...
... bout 4... We can
...
... in Fig... 6.1.

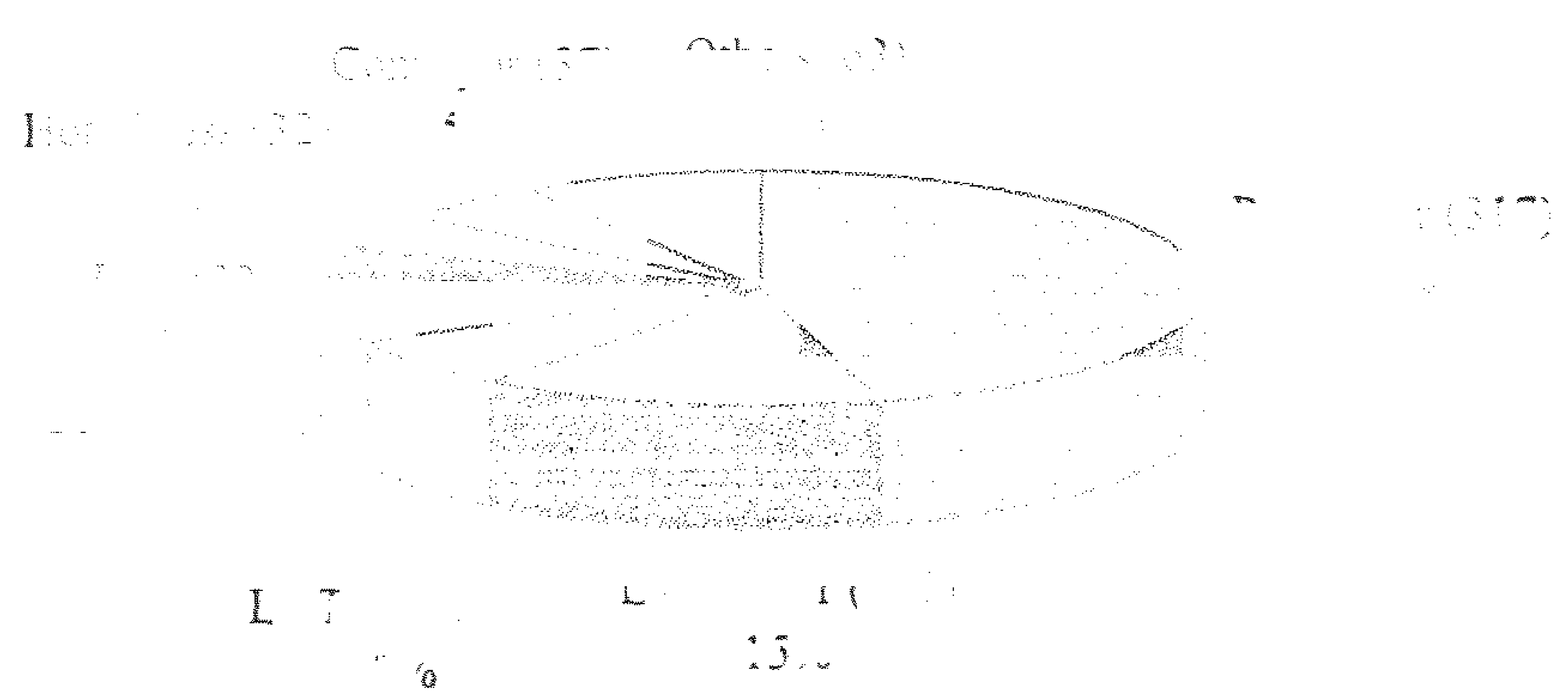


Figure 6.1. T...

F... 1 ... by the Pre... hich ... of
t... ess here ... policy fo... the Pre...
... he P...

... T... per ...
... easured ... ly

... we can ... he number ... our
... weekly for the Pr... from the data. Th... given in ... 2.

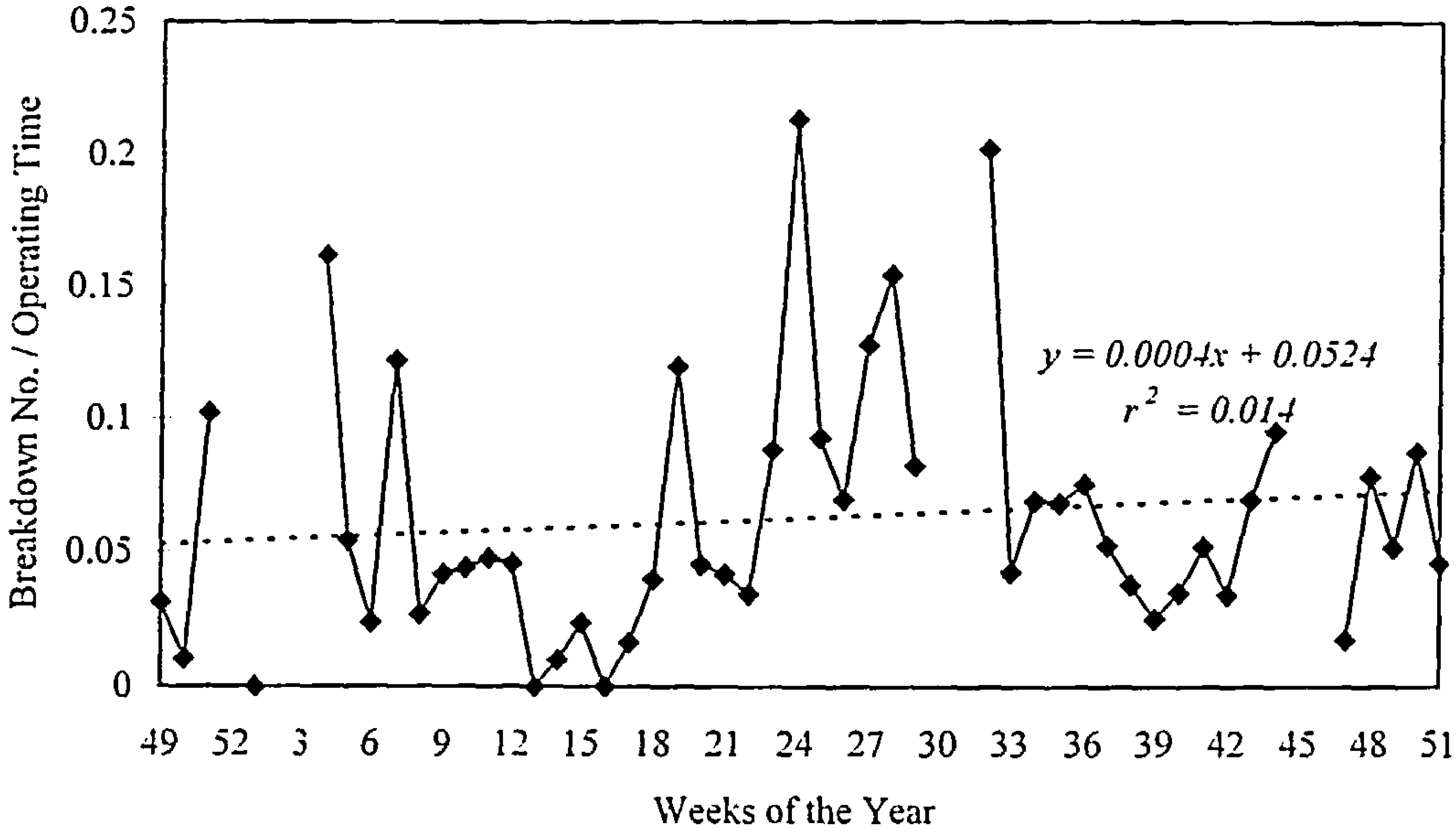


Figure 6.2. The Preformer breakdown number per operating hour measured weekly.

According to Figure 6.2, the failure occurrence rate per week for the Preformer seems to be constant through with a element of variability. A *t*-test supports the hypothesis that the performance is constant. To check the validity of this point, we can fit the graph to the regression line which is shown in Figure 6.2. Therefore, we conclude that the failure occurrence rate per week may be assumed constant.

6.3.4 The Effectiveness of TPM for the Preformer

The number of completed TPM cycles is 13 for the data collection period. Given effective TPM, it is expected that the number of failures would increase with each working day from the TPM. To explore this relationship, Table 6.1 is presented for the Preformer data and Figure 6.3 is based upon Table 6.1.

Table 6.1. The average number of failures since TPM.

Days since TPM	Total No. of Failures	Sample Size	Average No. of Failures	Days since TPM	Total No. of Failures	Sample Size	Average No. of Failures
1	2	13	0.154	27	8	3	2.667
2	14	13	1.077	28	10	3	3.333
3	7	13	0.538	29	4	3	1.333
4	9	13	0.692	30	8	3	2.667
5	16	13	1.000	31	0	3	0
6	11	12	0.917	32	4	3	1.333
7	14	10	1.400	33	4	3	1.333
8	9	8	1.125	34	1	3	0.333
9	11	7	1.571	35	0	3	0
10	11	7	1.571	36	2	3	0.667
11	3	7	0.429	37	5	3	1.667
12	5	6	0.833	38	1	3	0.333
13	10	6	1.667	39	2	3	0.667
14	2	6	0.333	40	2	3	0.667
15	12	6	2.000	41	3	3	1.000
16	11	6	1.833	42	2	2	1.000
17	11	6	1.833	43	1	2	0.500
18	6	6	1.000	44	4	2	2.000
19	10	6	1.667	45	2	2	1.000
20	3	6	0.500	46	1	2	0.500
21	4	5	0.800	47	2	2	1.000
22	7	5	1.400	48	2	2	1.000
23	7	4	1.750	49	6	2	3.000
24	8	4	2.000	50	4	2	2.000
25	2	3	0.667	51	7	2	3.500
26	7	3	2.333	52	4	2	2.000

In Table 6.1. since we cannot place reliance on the average number of failures after 21 days since TPM, because of the small number of TPMs, Figure 6.3 for the average number of failures per day is produced by the average number of failures until 20 days since TPM.

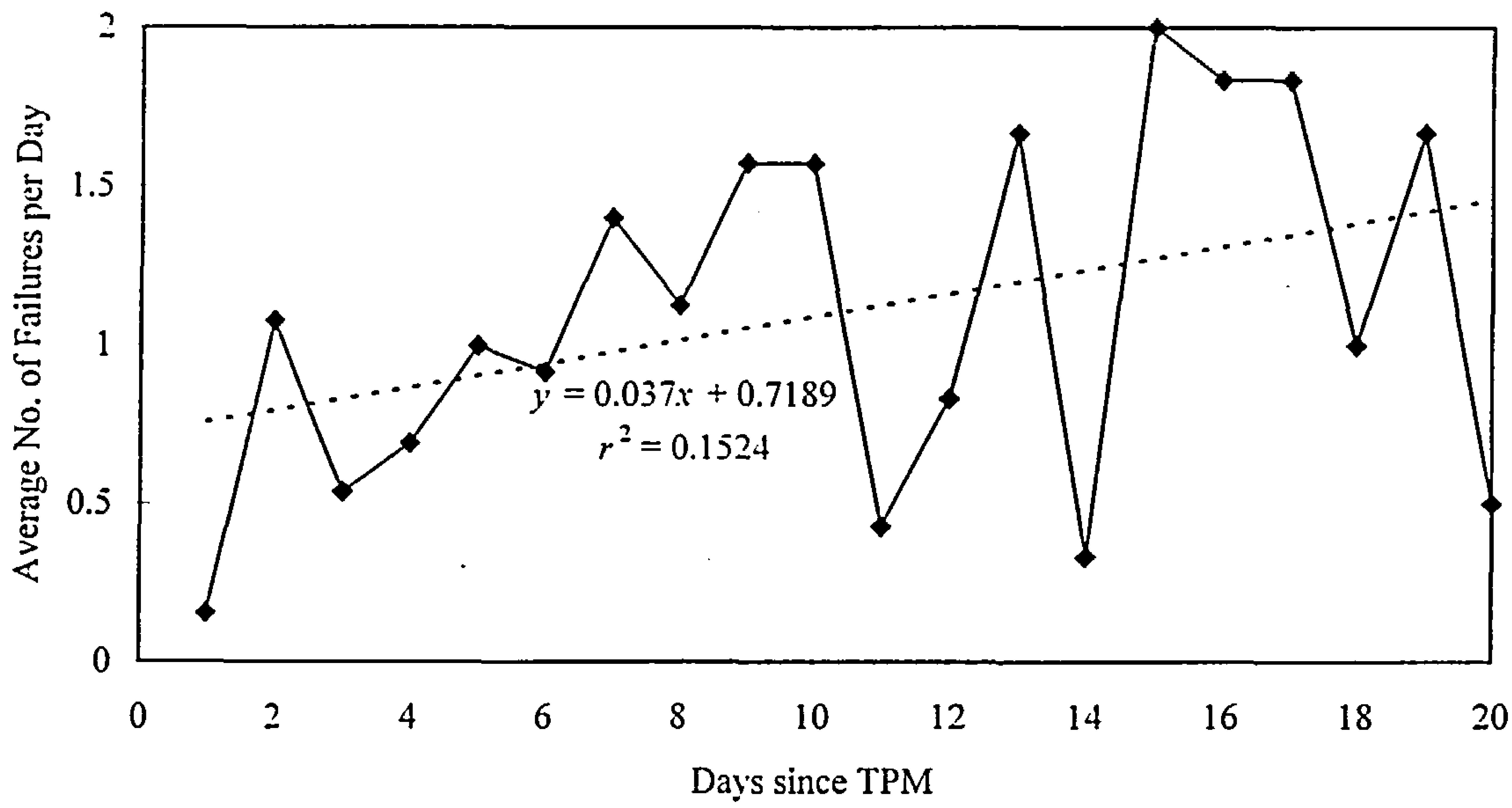


Figure 6.3. The effectiveness of TPM for the Preformer.

Figure 6.3 shows that the number of breakdowns of the Preformer slightly increase according to each working day since TPM. This may imply that the performed TPM for the Preformer might be effective in reducing the number of failures.

6.3.5 The Number of Defects Identified and Rectified during TPM

As a result of asking the production manager in the company, we have an estimate of the number of defects identified at TPM from the question (a), and the number of defects rectified during TPM from the question (b). Since the number of defects identified at TPM is given as 45 and the number of defects rectified during TPM is given as 30, which seemed low to us, from the equation (6.1), the probability of identifying and removing the defect during TPM, q, is estimated by

$$\hat{q} = \frac{45}{100} \times \frac{30}{100} = 0.135. \qquad (6.2)$$

6.3.6 The Mean Downtimes due to Failures and TPM activities

The collected data indicated 317 failures, with a total downtimes of 289.8 hours, that is a mean downtime per failure of 0.914 hours. The variance of downtime per failure repair is 2.154.

The total number of TPMs over the data collecting period is 13, with recorded total downtimes caused by TPM of 67.7 hours. Since 3 subsystems, including Preformer, are inspected for each TPM, assuming that the three subsystems take equal time to TPM, the mean duration of the TPM activity incorporating the Preformer, d_i, is given by

$$d_i = \frac{67.7}{13 \times 3} = 1.73 \text{ (hours)}. \quad (6.3)$$

6.4 Assumptions for Modelling

The first objective of the statistical modelling in this study is to estimate the parameters of the fault arriving process, the delay time distribution, and the probability of identifying and removing a fault at TPM. This will permit OR models of consequence variables such as downtime to be established based upon the estimated parameters. Observations of the data and previous experience suggest the following initial assumptions describe the operating practice over the period of data collection.

(a) Faults arise as a homogeneous Poison Process (HPP), and the instantaneous rate of occurrence of defects (ROCOD) is denoted by λ.

(b) Faults are independent of each other.

(c) The delay time h of a random fault is independent of its time of origin and has *pdf* $f(\cdot)$ and *CDF* $F(\cdot)$.

(d) Inspections carried out at TPM are assumed to be imperfect in that a fault present will be identified with probability r, $0 \le r \le 1$. Probabilities of detection of a fault at successive inspections are assumed to be independent and constant. An inspection requires d_i time units.

(e) Not all faults identified at TPM are assumed to be fixed because of limited time and resource allocated. Unattended faults may cause failures later. We assume that a fault identified at TPM is fixed during the TPM period with probability s, $0 \le s \le 1$. This does not influence the development of undetected faults. We assume that the choice of defect to rectify is arbitrary.

(f) Failures are identified immediately, and repairs or replacements are made as soon as possible. The mean downtime per failure is d_b.

In general, one would expect a nonhomogeneous Poisson process (NHPP) to be a good approximation to the fault arrival process. As plant ages, we would expect a HPP of arrival rate to be appropriate. This is because the probability of a fault arriving in a small time interval will be hardly changed by the pattern of previous fault arrivals, since those failed or defective components which are repaired or replaced are a negligible fraction of the total plant. This independence of intensity on past failure epochs characterizes a HPP. Barlow and Proschan [1965] proved that, for a complex machine with negligible repair times, the failure process does indeed in the limit follow a HPP. Furthermore, in this study, since the failure occurrence rate per week has an random pattern (see Figure 6.2), the fault arrival process may reasonably be assumed to follow a HPP. This is the reason underlying assumption (a).

When the system is so complex that it is difficult to track individual components, it is simplest when modelling to pool faults from all components. In this case $f(h)$ is the delay time *pdf* for any fault, that is assumption (b). Assumption (c) is common in delay time modelling, since it both greatly simplifies the modelling work and has been validated by real-world observation. Assumption (d) is due to the fact that the inspection work carried out on the Preformer during TPM period can be imperfect in that it has been observed that failures occur immediately after TPM, and furthermore, the number and the downtime of breakdowns are roughly the same for each working day. Assumption (e) is based upon observation and has been confirmed by the production manager. Assumptions (d) and (e) give us the probability of identifying and removing a fault, namely $q = rs$. Assumption (f) embodies the maintenance practice currently adopted.

6.5 Parameter Estimation

6.5.1 Likelihood Formulation

To form the likelihood function using the failure data including the accurate failure time, consider Figure 6.4.

Figure 6.4. The failure process of a fault arising in (T_{n-1}, T_n).

In Figure 6.4, T_n is the epoch of the nth TPM from new, $n = 1, 2, \ldots$, $t_{(i-1)j}$ is the epoch of the jth failure occurring in (T_{i-1}, T_i), $j = 1, 2, \ldots$, k_{i-1}, $t_{(i-1)k_{i-1}}$ is the time of the last failure in (T_{i-1}, T_i), and Δt is a small time interval in which only one event at most can arise.

If we can consider all observations in (T_{i-1}, T_i), these may be the TPM results at T_i and failure epochs in (T_{i-1}, T_i). The likelihood function is the product of the probabilities of these observations arising. For the TPM results, we can consider the probability of detecting and removing x_i faults from the system if they are present there, $P(x_i$ faults detected and removed at $T_i)$. For failure epochs in (T_{i-1}, T_i), we can consider the probabilities of a failure arising at times $t_{(i-1)j}$, j = 1, 2, ..., k_{i-1}, and of having no failure between failures. Therefore, the likelihood function L is given by

$$L = \prod_{i=1}^{m} \{P(x_i \text{ faults detected and removed at } T_i)$$

$$\cdot \prod_{j=1}^{k_{i-1}} P(\text{a failure at time } t_{(i-1)j}) \cdot P(\text{no failure between failures})\} \qquad (6.4)$$

and the log likelihood function l is given by

$$l = \sum_{i=1}^{m} \{\log P(x_i \text{ faults detected and removed at } T_i)$$

$$+ \sum_{j=1}^{k_{i-1}} \log P(\text{a failure at time } t_{(i-1)j}) + \log P(\text{no failure between failures})\}, \qquad (6.5)$$

where m is the number of inspections. In equation (6.5), the term, log$P$(no failure between failures), is necessary because of the complex component nature of the plant, and would not apply if it is single component item. This formulation assumes the necessary objective data is available from both inspection and failure interactive.

In the above log likelihood function, if the objective inspection data are not available, we have to use the failure data only. In this case, the log likelihood function measure is given by

$$l = \sum_{i=1}^{m} \{\sum_{j=1}^{k_{i-1}} \log P(\text{a failure at time } t_{(i-1)j}) + \log P(\text{no failure between failures})\}. \qquad (6.6)$$

To compute the above log likelihood function, firstly, we consider the probability of a failure over (t, t+Δt) resulting from a fault arising at time y, namely $P(t, t+\Delta t | y)$, (see Figure 6.5).

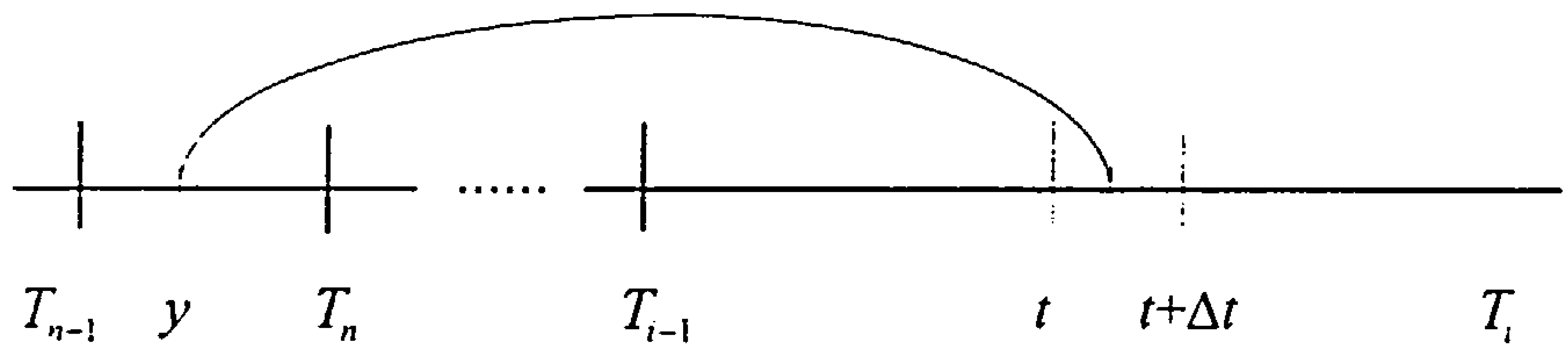


Figure 6.5. The failure process of a fault arising in (T_{n-1}, T_n)

Considering the probability q that a fault is identified and removed during the TPM period, as before $q = rs$, the probability that a fault arising in (T_{n-1}, T_n) with sufficiently long delay time will not be identified before the interval (T_{i-1}, T_i), $i > n$, is given by (see the equation (5.35))

$$c_n = (1-q)^{i-n} \qquad \text{for } n = 1, 2, \ldots, i\text{-}1. \tag{6.7}$$

Then, the probability of a failure in $(t, t+\Delta t)$ from a fault arising at time y in (T_{n-1}, T_n) is given by (see the equation (5.36))

$$P(t, t+\Delta t | y) = \begin{cases} c_n(F(t+\Delta t - y) - F(t-y)) & \text{for } T_{n-1} < y < T_n,\ n = 1,2,\cdots,i-1 \\ F(t+\Delta t - y) - F(t-y) & \text{for } T_{i-1} < y < t \\ F(t+\Delta t - y) & \text{for } t < y < t+\Delta t \\ 0 & \text{otherwise.} \end{cases} \tag{6.8}$$

From equations (6.8), we can obtain the expected number of faults found and removed at the ith TPM, namely $EN_p(T_i)$, given by (see the equation (5.37))

$$EN_p(T_i) = \lambda \sum_{n=1}^{i-1} c_n q \int_{T_{n-1}}^{T_n} (1 - F(T_i - y))dy + \lambda q \int_{T_{i-1}}^{T_i} (1 - F(T_i - y))dy. \tag{6.9}$$

Since faults are assumed to arise according to a HPP, the number of faults detected and removed at TPM follows a Poisson distribution from Lemma 5.3. Because the number of faults detected and removed at TPM follows a Poisson distribution with means defined by equation (6.9), the probability of x_i faults detected and removed at T_i is

$$P(x_i \text{ faults dectected and removed at } T_i) = \frac{(\mathrm{E}N_p(T_i))^{x_i} e^{-\mathrm{E}N_p(T_i)}}{x_i!}. \tag{6.10}$$

To find the probabilities for failure epochs in (T_{i-1}, T_i), it is of value to us to note lemma 5.4. Using the lemma 5.4, we can obtain the probability of a failure arising at time $t_{(i-1)j}$. For very small Δt,

$$P(\text{a failure at time } t_{(i-1)j}) = v(t_{(i-1)j})\Delta t, \tag{6.11}$$

where, as before, the failure arrival rate function, $v(t)$, is given by (see the equation (5.38))

$$v(t) = \lambda\{\sum_{n=1}^{i-1} c_n (F(t-T_{n-1}) - F(t-T_n)) + F(t-T_{i-1})\}. \tag{6.12}$$

Since the probability of having no failure in $(t_{(i-1)(j-1)}, t_{(i-1)j})$ is given by

$$P(\text{no failure in } (t_{(i-1)(j-1)}, t_{(i-1)j})) = e^{-\int_{t_{(i-1)(j-1)}}^{t_{(i-1)j}} v(t)dt}, \tag{6.13}$$

the total summation of the log probability of having no failure between failures becomes

$$\sum \log P(\text{no failure between failures}) = \sum_{j=1}^{k_{i-1}} (-\int_{t_{(i-1)(j-1)}}^{t_{(i-1)j}} v(t)dt) - \int_{t_{(i-1)k_{i-1}}}^{T_i} v(t)dt. \tag{6.14}$$

If we define $t_{(i-1)0} = T_{i-1}$, without loss of generality, the equation (6.14) becomes

$$\sum \log P(\text{no failure between failures}) = -\int_{T_{i-1}}^{T_i} v(t)dt. \tag{6.15}$$

If failures were not occurred in (T_{i-1}, T_i). from equation (6.13). the probability of having no failure in (T_{i-1}, T_i) is given by

$$P(\text{no failure in } (T_{i-1}, T_i)) = e^{-\int_{T_{i-1}}^{T_i} v(t)dt} . \tag{6.16}$$

Accordingly, by taking the log of equation (6.16), we can see that equation (6.15) is satisfied.

Dividing the product of the equations (6.10), (6.11), and (6.15) by Δt and taking a log, the log likelihood function becomes

$$l = \log L = \sum_{i=1}^{m} \log P(n_i \text{ faults at } T_i) + \sum_{i=1}^{m} \log P(p_i \text{ faults rectified at } T_i)$$

$$+ \sum_{i=1}^{m} \{\sum_{j=1}^{k_i - 1} \log \frac{P(\text{a failure at time } t_{(i-1)j})}{\Delta t} + \sum \log P(\text{no failure between failures})\}$$

$$= \sum_{i=1}^{m} (x_i \log EN_p(T_i) - EN_p(T_i) - \log x_i!) + \sum_{i=1}^{m} \{\sum_{j=1}^{k_i - 1} \log v(t_{(i-1)j}) - \int_{T_{i-1}}^{T_i} v(t)dt\} . \tag{6.17}$$

Also, for the case where objective inspection data are not available, information is lost and the log likelihood function is given by

$$l = \sum_{i=1}^{m} \{\sum_{j=1}^{k_i - 1} \log v(t_{(i-1)j}) - \int_{T_{i-1}}^{T_i} v(t)dt\} . \tag{6.18}$$

Using the above likelihood equations, we can estimate the parameters of the process.

6.5.2 Simulation Test

Since we know that $q = 0.135$, the number of TPMs is 13, and the sample size for the Preformer is about 300, we can check the validity of the likelihood formulation given by the equations (6.17) and (6.18) using a simulation test. Firstly, we generate a set of data by simulating the process represented by the set of assumptions in section 6.4 with the known parameters using Figure 5.6. If these parameters can be recovered by the maximum likelihood method based upon the equations (6.17) and (6.18), we will have confirmed the log likelihood formulation and use it.

FORTRAN is chosen as a suitable language for simulating, mainly because of the excellent NAG library of numerical routines available for the Pentium-PC. The NAG function minimizer E04JAF was used to minimise minus the log likelihood.

To simplify the problem, we firstly consider a case that the delay time is exponentially distributed,

$$F(h) = 1 - e^{-\beta h}, \qquad (6.19)$$

where β is the arrival rate. Firstly, we assume that the probability of identifying and removing a fault, namely q = 0.135, is given by the subjective method. Also, to generate a set of data using a simulation, we assume that λ = 0.05 and β = 0.02. Next. we can generate a set of data using Figure 5.9. From the simulated set of data, the result of the parameter estimation process based upon the equations (6.17) and (6.18) is given by Table 6.2.

Table 6.2. Estimation result for an exponential distribution when the probability q is fixed as 0.135.

Number of inspections	Sample size	Estimation of case with inspection data		Estimation of case without inspection data	
		λ	β	λ	β
13	259	0.052	0.014	0.051	0.023
30	623	0.054	0.015	0.053	0.020
60	1215	0.052	0.018	0.051	0.024
100	2039	0.052	0.020	0.052	0.021
150	3006	0.051	0.022	0.050	0.023

* λ is the rate of occurrence of faults and β is the scale parameter of the exponential distribution.
* True values are λ = 0.05 and β = 0.02.

Table 6.2 shows that the maximum likelihood estimates recover quite well the underlying process of failure and fault origination based upon equations (6.17) and (6.18).

If the probability of identifying and removing a fault q is not given by the subjective method, we have to estimate the probability of identifying and removing a fault q by the objective method. After simulating a set of data, the estimation results based upon the equations (6.17) and (6.18) for the probability of identifying and removing a fault q, for λ and β, as function of the number of inspections are given by Table 6.3 to 6.7.

Table 6.3. Estimation result for an exponential distribution when $q = 0.135$.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	β	q	λ	β	q
13	259	0.052	*0.031*	*0.400*	0.053	0.025	*0.500*
30	623	0.053	*0.028*	*0.305*	0.055	0.020	*0.398*
60	1215	0.052	*0.044*	*0.363*	0.052	*0.040*	*0.374*
100	2039	0.052	0.026	0.178	0.052	*0.036*	0.163
150	3006	0.051	0.021	0.129	0.050	*0.069*	0.151

* True values are λ=0.05 and β=0.02.

Table 6.3 shows that if a set of inspection data is available, the maximum likelihood estimates recover the underlying process of failure and fault origination parameters based upon equation (6.17) in the case of the large number of inspections. Also, Table 6.3 shows that if inspection data is not available, the maximum likelihood estimates do not recover the underlying process of failure and fault origination based upon equation (6.18) in spite of the large number of inspections. This is to be expected since information on q is contained within inspection data, and if q is small, very little information on q will be available from failure data. This will change as q increases.

Table 6.4. Estimation result for an exponential distribution when $q = 0.3$.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	β	*q*	λ	β	*q*
13	255	0.052	0.024	*0.488*	0.052	0.022	*0.527*
30	613	0.052	0.025	*0.425*	0.055	0.019	*0.479*
60	1197	0.052	*0.035*	*0.484*	0.052	*0.037*	*0.482*
100	2006	0.052	0.021	0.273	0.051	0.029	0.266
150	2949	0.051	0.016	0.224	0.050	0.017	0.270

* True values are λ=0.05 and β=0.02.

Table 6.4, in case of $q = 0.3$, shows that if the number of inspections is large, the maximum likelihood estimates recover the underlying process of failure and fault origination based upon equations (6.17) and (6.18) regardless of a set of inspection data.

Table 6.5. Estimation result for an exponential distribution when $q = 0.5$.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	β	*q*	λ	β	*q*
13	250	0.052	0.020	0.514	0.052	0.020	0.539
30	598	0.053	0.020	0.545	0.055	0.016	0.573
60	1157	0.051	0.021	0.559	0.051	0.024	0.557
100	1946	0.052	0.016	0.431	0.051	0.019	0.496
150	2855	0.051	0.015	0.488	0.050	0.023	0.495

* True values are λ=0.05 and β=0.02.

Table 6.6. Estimation result for an exponential distribution when $q = 0.7$.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	β	*q*	λ	β	*q*
13	247	0.052	0.020	0.607	0.052	0.020	0.603
30	589	0.053	0.020	0.681	0.055	0.017	0.690
60	1133	0.052	0.020	0.691	0.051	0.022	0.691
100	1897	0.052	0.017	0.675	0.051	0.019	0.677
150	2783	0.051	0.016	0.662	0.050	0.020	0.672

* True values are λ=0.05 and β=0.02.

Table 6.7. Estimation result for an exponential distribution when $q = 0.9$.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	β	q	λ	β	q
13	238	0.052	0.018	0.855	0.052	0.018	0.857
30	576	0.053	0.020	0.834	0.055	0.016	0.836
60	1165	0.052	0.021	0.898	0.052	0.021	0.899
100	1844	0.052	0.018	0.824	0.052	0.020	0.828
150	2705	0.051	0.018	0.830	0.050	0.021	0.825

* True values are λ=0.05 and β=0.02.

Table 6.5, 6.6, and 6.7, in case of $q > 0.5$, show that the maximum likelihood estimates recover the underlying process of failure and fault origination based upon equations (6.17) and (6.18) regardless of a set of inspection data and the number of inspections.

From Table 6.3 to 6.7, we can see that as the probability of identifying and removing a fault q increases, the estimation of parameters based upon equations (6.17) and (6.18) is more accurate. If q is over 0.5, the estimation of parameters based upon equations (6.17) and (6.18) is accurate in spite of a low number of inspections. If the probability of identifying and removing a fault q is small, we need more information about inspections in order to estimate the parameters accurately. Referring to Table 6.3 and 6.4, in case of q=0.135 and q=0.3, show, if the number of inspections is less than 100, we cannot estimate the parameters with any confidence. In comparison with the case of Table 6.3, where the number of parameters to estimate is 2, if the number of parameters to estimate is 3, we need more information about inspections.

In conclusion here, when the inspection parameter q is known, we can estimate the parameters accurately in spite of a small number of inspections. However, when the inspection parameter q also needs to be estimated, we need more information about inspections in order to obtain the accurate parameter estimations.

Next, assuming that the delay time has a Weibull distribution.

$$F(h) = 1 - e^{-(\beta x)^{\alpha}}, \tag{6.20}$$

where α is the shape parameter and β is the scale parameter. the process is again simulated. Firstly, we assume that the probability of identifying and removing a fault q is given by the subjective method. Also, to generate a set of data using a simulation, we assume that $\lambda = 0.05$, $\alpha = 1.5$, and $\beta = 0.02$ and the probability of identifying and removing a fault q is fixed. We then generate a set of data using Figure 5.6. From the simulated set of data, the results of the parameter estimation process based upon the equations (6.17) and (6.18), various values of the probability of identifying and removing a fault q and the number of inspections are given by Table 6.8 to 6.12.

Table 6.8. Estimation result for a Weibull distribution when the probability q is fixed as 0.135.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	α	β	λ	α	β
13	261	Cannot find optimal values			Cannot find optimal values		
30	609	Cannot find optimal values			Cannot find optimal values		
60	1237	0.053	*6.179*	0.017	0.053	*6.018*	0.017
100	2064	0.053	*5.880*	0.017	0.053	*5.912*	0.017
150	3058	0.052	*5.444*	0.018	0.052	*5.661*	0.017

* λ is the rate of occurrence of faults and α is the shape parameter of the Weibull distribution and β is the scale parameter of the Weibull distribution.
* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.9. Estimation result for a Weibull distribution when the probability q is fixed as 0.3.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	α	β	λ	α	β
13	257	Cannot find optimal values			Cannot find optimal values		
30	596	0.052	*5.542*	0.017	Cannot find optimal values		
60	1214	0.053	*5.715*	0.017	Cannot find optimal values		
100	2026	0.053	*4.566*	0.017	0.052	*5.154*	0.016
150	3007	0.052	*3.028*	0.019	0.052	*4.531*	0.017

* True values are λ=0.05, α=1.5 and β=0.02.

Table 6.8 and 6.9, in case of q = 0.135 and q = 0.3, show that apart from the parameter λ. the maximum likelihood estimates do not recover the underlying process of failure and fault origination based upon equation (6.17) and (6.18) in spite of the large number of inspections.

Table 6.10. Estimation result for a Weibull distribution when the probability q is fixed as 0.5.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	α	β	λ	α	β
13	253	0.051	*57.90*	0.021	0.052	*52.95*	0.021
30	585	0.052	*5.083*	0.018	0.053	*4.804*	0.016
60	1193	0.053	2.475	0.018	0.054	2.277	0.014
100	1988	0.053	2.672	0.019	0.053	2.472	0.016
150	2945	0.052	2.191	0.020	0.052	2.231	0.018

* True values are λ=0.05, α=1.5 and β=0.02.

Table 6.11. Estimation result for a Weibull distribution when the probability q is fixed as 0.7.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	α	β	λ	α	β
13	253	0.051	*25.81*	0.022	0.053	*34.40*	0.021
30	576	0.052	*3.185*	0.020	0.053	*2.819*	0.016
60	1167	0.053	2.073	0.019	0.054	1.841	0.015
100	1943	0.053	2.376	0.019	0.053	2.190	0.018
150	2868	0.052	2.147	0.020	0.052	2.116	0.019

* True values are λ=0.05, α=1.5 and β=0.02.

Table 6.10 and 6.11, in case of q = 0.7 and q = 0.9, show that if the number of inspections is large. the maximum likelihood estimates recover the underlying of failure and fault origination based upon equations (6.17) and (6.18) regardless of a set of inspection data.

Table 6.12. Estimation result for a Weibull distribution when the probability q is fixed as 0.9.

Number of inspections	Sample size	Estimation of case with inspection data			Estimation of case without inspection data		
		λ	α	β	λ	α	β
13	249	0.051	1.647	0.024	0.054	*8.088*	0.018
30	564	0.052	2.357	0.019	0.054	2.154	0.016
60	1144	0.053	1.837	0.019	0.055	1.616	0.016
100	1904	0.053	2.110	0.020	0.054	1.903	0.018
150	2811	0.052	1.924	0.020	0.052	1.857	0.019

* True values are λ=0.05, α=1.5 and β=0.02.

Table 6.12 shows that if a set of inspection data is available, the maximum likelihood estimates recover the underlying process of failure and fault origination based upon equation (6.17) regardless of the number of inspections. Also, Table 6.12 shows that if a set of inspection data is not available, the maximum likelihood estimates recover the underlying process of failure and fault origination based upon equation (6.18) only in the case of the large number of inspections.

From Table 6.8 to 6.12, we can see that as before, as the probability of identifying and removing a fault q increases, the estimation of parameters based upon equations (6.17) and (6.18) is more accurate. However, the above tables show that although q is over 0.5. we can estimate the parameters accurately in the cases where the number of inspections is over 60. If the probability of identifying and removing a fault q is small, it may be difficult to estimate the parameters accurately. In case of q = 0.135 and q = 0.3, if the number of inspections is less than 150, we cannot estimate the parameters accurately. When the probability q is small, we may need more information about inspections. However, to obtain more information about inspections, we have to increase the number of inspections. Also, it requires more times to estimate the parameters using the NAG library. In comparison with the case of the exponential distribution. the Weibull distribution case requires more information on inspections.

Secondly, consider the case where the probability of identifying and removing a fault q is not known. If the probability of identifying and removing a fault q cannot be estimated by the subjective method, we have to estimate the probability of identifying and removing a fault q by the objective method. After simulating a set of data, the estimation results based upon the equations (6.17) and (6.18), for different values of the probability of identifying and removing a fault q and various number of inspections are given by Table 6.13 to 6.17.

Table 6.13. Estimation result for a Weibull distribution when $q = 0.135$.

Number of inspections	Sample size	Estimation of case with inspection data				Estimation of case without inspection data			
		λ	α	β	q	λ	α	β	q
13	261	Cannot find optimal values				Cannot find optimal values			
30	609	Cannot find optimal values				Cannot find optimal values			
60	1237	Cannot find optimal values				Cannot find optimal values			
100	2064	Cannot find optimal values				0.053	*5.019*	0.017	0.245
150	3058	0.052	*5.426*	0.018	0.176	0.052	*5.340*	0.017	0.188

* λ is the rate of occurrence of faults and α is the shape parameter of the Weibull distribution and β is the scale parameter of the Weibull distribution.
* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.14 Estimation result for a Weibull distribution when $q = 0.3$.

Number of inspections	Sample size	Estimation of case with inspection data				Estimation of case without inspection data			
		λ	α	β	q	λ	α	β	q
13	257	Cannot find optimal values				Cannot find optimal values			
30	596	0.052	*5.547*	0.017	0.300	Cannot find optimal values			
60	1214	0.053	*6.026*	0.016	0.283	0.054	*2.667*	0.013	0.378
100	2026	0.053	*4.420*	0.018	0.305	0.054	*2.649*	0.016	0.379
150	3007	0.052	*4.598*	0.017	0.260	0.052	*4.699*	0.016	0.293

* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.13 and 6.14, in case of $q = 0.135$ and $q = 0.3$, show that the maximum likelihood estimates do not recover the underlying process of failure and fault origination based upon equation (6.17) and (6.18) in spite of the large number of inspections.

Table 6.15. Estimation result for a Weibull distribution when $q = 0.5$.

Number of inspections	Sample size	Estimation of case with inspection data				Estimation of case without inspection data			
		λ	α	β	q	λ	α	β	q
13	253	Cannot find optimal values				Cannot find optimal values			
30	585	0.052	*5.180*	0.017	0.455	Cannot find optimal values			
60	1193	0.053	*5.770*	0.016	0.406	0.054	2.293	0.014	0.498
100	1988	0.053	*2.854*	0.019	0.487	0.054	2.002	0.016	0.547
150	2945	0.052	2.296	0.020	0.490	0.052	2.027	0.019	0.519

* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.16. Estimation result for a Weibull distribution when $q = 0.7$.

Number of inspections	Sample size	Estimation of case with inspection data				Estimation of case without inspection data			
		λ	α	β	q	λ	α	β	q
13	253	Cannot find optimal values				Cannot find optimal values			
30	576	0.052	*3.949*	0.019	0.636	0.053	*2.800*	0.016	0.702
60	1167	0.053	2.336	0.018	0.664	0.054	1.712	0.015	0.726
100	1943	0.053	2.023	0.021	0.773	0.054	1.689	0.019	0.801
150	2868	0.052	1.721	0.022	0.798	0.052	1.590	0.021	0.811

* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.17. Estimation result for a Weibull distribution when $q = 0.9$.

Number of inspections	Sample size	Estimation of case with inspection data				Estimation of case without inspection data			
		λ	α	β	q	λ	α	β	q
13	249	Cannot find optimal values				Cannot find optimal values			
30	564	0.052	*3.926*	0.018	0.781	0.053	*2.913*	0.016	0.826
60	1144	0.053	2.475	0.018	0.808	0.054	1.922	0.015	0.847
100	1904	0.053	2.220	0.019	0.879	0.054	1.938	0.018	0.894
150	2811	0.052	1.924	0.020	0.900	0.052	1.827	0.019	0.906

* True values are $\lambda = 0.05$, $\alpha = 1.5$ and $\beta = 0.02$.

Table 6.15, 6.16, and 6.17, in case of $q = 0.5$. $q = 0.7$, and $q = 0.9$, show that if the number of inspections is large. the maximum likelihood estimates recover the underlying failure and fault origination based upon equations (6.17) and (6.18) regardless of a set of inspection data.

From Table 6.13 to 6.17. we can see that as before, as the probability of identifying and removing a fault q increases, the estimation of parameters is more accurate. However, Table 6.15 shows that in the case of q=0.5, we can estimate the parameters reasonably well if the number of inspections is over 150. In case of q=0.7 and q=0.9, we can estimate the parameters well if the number of inspections is over 60. If the probability of identifying and removing a fault q is small, we need more information about inspections in order to estimate the parameters accurately. In the case where the probability of identifying and removing a fault q is very small, for example q=0.135 and q=0.3, we cannot estimate the parameters accurately. In comparison with the case where the number of parameters to be estimate is 3, this case need more information about inspections. When q is small, it seems appropriate to obtain a subjective estimation of q, and then estimate the other parameters conditional on this estimate. It is possible, of course, that the value of q can be increased through engineering and manpower management means including better supervision and control.

In conclusion, when the number of parameters to estimate is 2, we can estimate the parameters accurately in spite of a small number of inspections. If the number of parameters to estimate increases, we need more information about inspections. This requires more time to estimate parameters using the NAG library. For example, in the case that the number of parameters to estimate is 3, the computing process using the NAG library took approximately 1 hour using a Pentium-PC. Also, as expected, the required volume of information about inspections is dependent upon the distribution of the delay time. When the delay time has a Weibull distribution, if the probability of identifying and removing a fault q is large, the maximum likelihood estimates recover the underlying process of fault origination and failure based upon equations (6.17) and (6.18), otherwise it does not. With the actual data set, since the probability of identifying and removing a fault during the TPM. estimated as q=0.135, is very small and the number of TPMs is 13, we are unable to use the Weibull distribution as a candidate distribution.

6.5.3 Selecting the Distribution of the Model

Before fitting a model to the real data, the functional form of the delay time distribution must be specified. The best choice of the distribution from a family of plausible distributions for h is made, using the criterion of minimum Akaike Information (AIC) which is a method of evaluating the goodness of fit of a model. AIC is derived under the assumption that the true distribution can be described by the given model when its parameters are suitably adjusted.

In maximum log likelihood estimation, the goodness of parameters of a specific model can be measured by the expected log likelihood, namely, the larger the expected log likelihood, the better the parameters values. The log likelihood can be regarded as an estimator of the expected log likelihood. We now introduce the mean expected log likelihood as a measure for the goodness of fit of a model. This quantity is defined as the mean, with respect to the data x, of the expected log likelihood of the maximum likelihood model. The larger the mean expected log likelihood, the better the fit of the model. At first sight, it would seem that the mean expected log likelihood can be estimated by the maximum log likelihood. The maximum log likelihood, however, is shown to be a biased estimator of the mean expected log likelihood (see Sakamoto *et al* [1986]). The maximum log likelihood has a general tendency to over estimate the true value of the mean expected log likelihood. This tendency is more prominent for models with a larger number of free parameters. This means that if we choose the model with the largest maximum log likelihood, a model with an unnecessarily large number of free parameters is likely to be chosen.

By a close examination of the relationship between the bias and the number of free parameters of a model, we will find that (maximum log likelihood of a model) - (number of free parameters of the model)

is an asymptotically unbiased estimator of the mean expected log likelihood (see Sakamoto *et al* [1986]). Taking historical reasons into account, this is presented as minus twice of this value

$$AIC = -2\times(\text{maximum log likelihood of the model}) \\ +2\times(\text{number of free parameters of the model}) \qquad (6.21)$$

is proposed as the criterion for model selection in Sakamoto *et al* [1986]. A model which minimises the AIC (minimum AIC estimate) is considered to be the most appropriate model. Equation (6.21) implies that when there are several models whose values of maximum likelihood are about the same level, we should choose the one with the smallest number of free parameters.

6.5.4 Results of the Model Fit

We now consider the Preformer data. Since the probability of identifying and removing a fault during the TPM is very small, possible candidates for $F(h)$ are (1) exponential and (2) mixed delta-exponential distribution. There may be some faults that have zero delay time, which can be modelled with a mixed delay time distribution with *pdf* given by $(1-P)f(h)+P\delta(h)$, where $f(h)$ is the *pdf* of delay time h, $\delta(h)$ is the Dirac delta function, and P is the proportion of faults that have zero delay time.

Assuming the above choice of delay time distribution, the fitted values of parameters are shown in Table 6.18. From Table 6.18, it can be seen that there is not much difference in AIC values between model choice of delay time distribution. This means that the delay time models are not very sensitive to the data. Mixed delta-exponential distribution is selected as having the lowest AIC value.

Table 6.18. Fitted values of parameters from the Preformer data.

Models	Exponential distribution	Mixed delta-exponential distribution
ROCOD ($\hat{\lambda}$)	0.117	0.167
Scale parameter ($\hat{\beta}$)	0.00047	0.000226
$\hat{P}$		0.053
Maximum log likelihood	-1220.92	-1213.09
AIC	2445.84	2432.18

*Notes : ROCOD is the rate of occurrence of faults.

6.5.5 Test for Goodness of Fit

The chi-squared test statistic is

$$\chi^2 = \sum_{i=1}^{K} \frac{(n_i - \hat{n}_i)^2}{\hat{n}_i}, \tag{6.22}$$

where the range of data is divided into K suitable classes, n_i is the number of the events in the ith class, and $\hat{n}_i$ is the expected number of events in the ith class calculated from the fitted model.

If we group the failure times in our data into classes, then there will, however, be two difficulties in the calculation of the expected failures in each class: (a) each failure epoch may have a different number of previous TPMs, and (b) the data set show that TPM intervals are not equally spaced. To overcome these two difficulties, let Δt denote the length of the class interval of the histogram of failure times and let $EN_f(T_1,...,T_{i-1},I_j^i)$ denote the expected number of failures over the interval

$$I_j^i = (T_{i-1} + (j-1)\Delta t, T_{i-1} + j\Delta t), \tag{6.23}$$

given that the previous TPM history is $(T_1,....T_{i-1})$. Then, we have

$$\mathrm{E}N_f(T_1,\ldots,T_{i-1},I_j^i)=\lambda\int_{T_{i-1}+(j-1)\Delta t}^{T_{i-1}+j\Delta t}\sum_{n=1}^{i-1}c_n(F(x-T_{n-1})-F(x-T_n))dx$$
$$+\lambda\int_{T_{i-1}+(j-1)\Delta t}^{T_{i-1}+j\Delta t}F(x-T_{i-1})dx\,, \tag{6.24}$$

where, as before, λ is the instantaneous rate of occurrence of defects (ROCOD) and c_n is the probability that a fault arising in (T_{n-1},T_n) with sufficiently long delay time will not be identified before the interval (T_{i-1},T_i).

Given that the expected number of failures over I_j^i within (T_{i-1},T_i) is available, the next task is to group the expected number of failures over I_j^n, (n = 1, …, l), where l is the number of TPMs, in different TPM intervals together. To do this, we introduce a step function (θ) to give effect to the mechanism of different TPM intervals. It follows that the expected number of failures over

$$I_j=\bigcup_{n=1}^{l}I_j^n \tag{6.25}$$

is given by

$$\mathrm{E}N_f(I_j)=\sum_{n=1}^{l}\mathrm{E}N_f(T_1,\ldots,T_{n-1},I_j^n)\theta_{n-1,j}\,, \tag{6.26}$$

where

$$\theta_{n-1,j}=\begin{cases}1 & \text{if } T_{n-1}+j\Delta t\le T_n,\\ 0 & \text{otherwise.}\end{cases} \tag{6.27}$$

If $N_f(I_j)$ denotes the observed number of failures in the jth class, using the equation (6.22), the chi-squared test statistic in our case is simply

$$\chi^2 = \sum_{j=1}^{K} \frac{(N_f(I_j) - EN_f(I_j))^2}{EN_f(I_j)} \tag{6.28}$$

and the number of degrees of freedom is $l + K - v$, with v, namely the number of model parameters.

Using the equation (6.28) on the Preformer data with the model parameters of Table 6.18, the model predictions and actual data for the number of failures of the Preformer are shown in Figure 6.6.

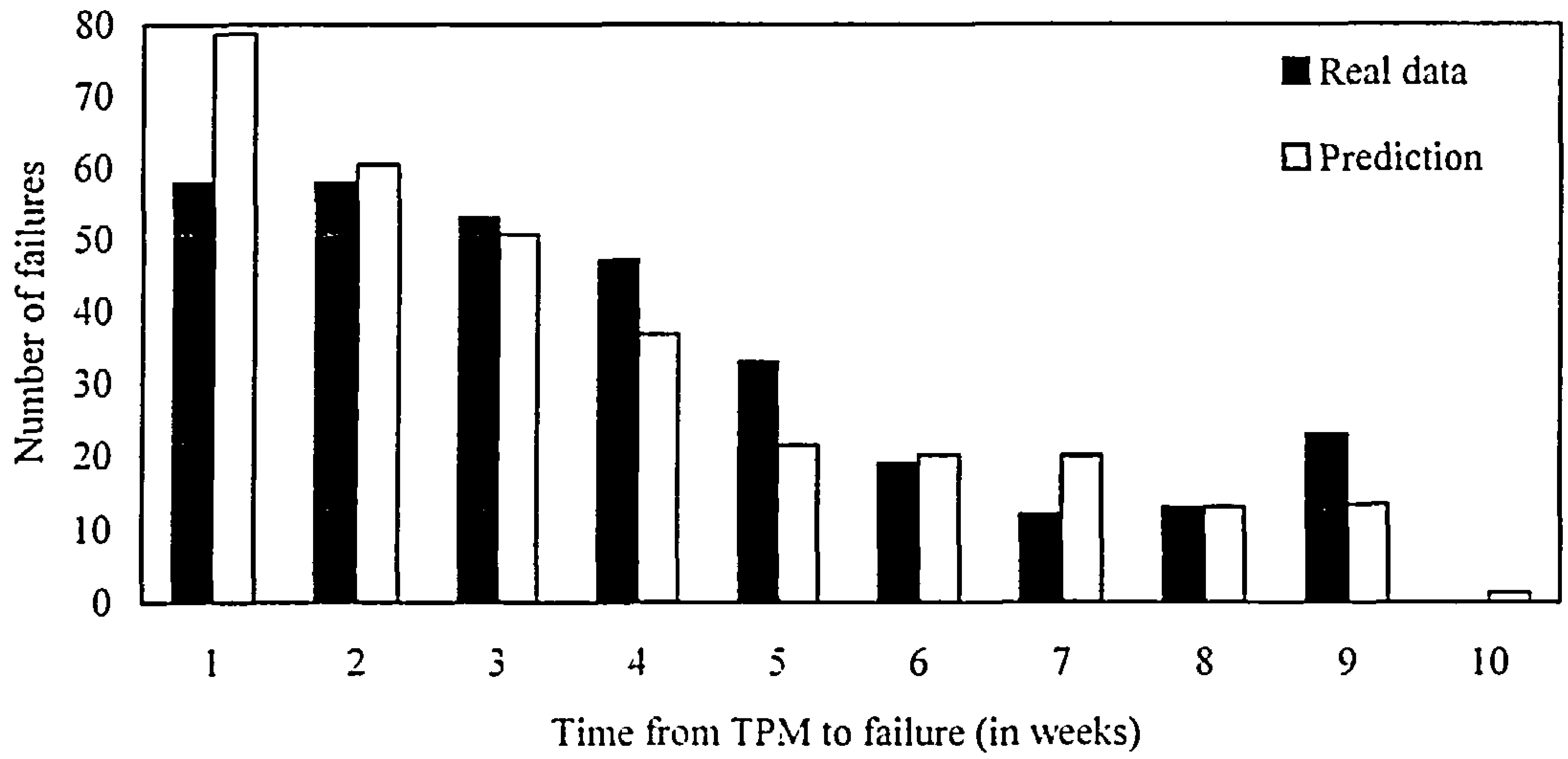


Figure 6.6. Histogram of failures for Preformer.

Figure 6.6 shows that the model fit seems to be adequate. To check the validity of this point, we can compare the chi-squared test statistics with the critical value of the chi-squared distribution table. Since the number of degrees of freedom is

$$df = l + K - v = 20, \tag{6.29}$$

the critical value of the chi-squared distribution is given by

$$\chi^2_{(\alpha, df)} = 31.4 \tag{6.30}$$

with the significance level of $\alpha = 0.05$. Since the chi-squared test statistics is given by $\chi^2 = 26.072$, we can conclude that the model fit is acceptable at the significance level of $\alpha = 0.05$.

6.6 Downtime Model

6.6.1 Delay Time Model

Given an acceptable model for the failure and TPM process of the Preformer, a TPM model of maintenance practice of the machine may now be established. We model downtime, since the major concern of the company is to reduce the downtime caused by failures and TPM activities. The conventional downtime measure is the expected downtime per unit time over a long future period. The key issue in the model is the expected number of failures over different TPM cycles.

Since we have assumed that faults arise according to an HPP, the expected number of failures over $(t, t+\Delta t)$ can be given by

$$\begin{aligned} EN_f(t, t+\Delta t) &= \lambda \int_0^{\infty} P(t, t+\Delta t \mid y) dy \\ &= \lambda \sum_{n=1}^{i-1} C_n \int_{T_{n-1}}^{T_n} (F(t+\Delta t - y) - F(t-y)) dy \\ &\quad + \lambda \int_{T_{i-1}}^{t} (F(t+\Delta t - y) - F(t-y)) dy + \lambda \int_t^{t+\Delta t} F(t+\Delta t - y) dy, \quad (6.31) \end{aligned}$$

where, as before, $P(t, t+\Delta t \mid y)$ is the probability of a failure in $(t, t+\Delta t)$ from a fault arising at time y in (T_{n-1}, T_n) (see Figure 6.5 and equation (6.8)) and $EN_f(t, t+\Delta t)$ denotes the expected number of failures over $(t, t+\Delta t)$. Changing the integral variable and rearranging the integral sequence, after some manipulation we have

$$\mathrm{E}N_f(t,t+\Delta t) = \lambda \int_t^{t+\Delta t} \sum_{n=1}^{i-1} C_n (F(x-T_{n-1}) - F(x-T_n))dx + \lambda \int_t^{t+\Delta t} F(x-T_{i-1})dx \,. \quad (6.32)$$

Also, since the delay time follows a mixed delta-exponential distribution, the expected number of failures over a TPM cycle when in a steady state under various TPM cycles T can be obtained by setting $t = T_{i-1}$, $t + \Delta t = T_{i-1} + T = T_i$, and letting $i \to \infty$ in the equation (6.32), and summing the resultant geometric series of equation (6.32) with respect to n. This gives

$$\mathrm{E}N_f(T) = [\lambda(e^{-\beta T} - 2 + e^{\beta T}) \frac{1-q}{\beta(e^{\beta T} - 1 + q)}](1-P) + \lambda(T + \frac{1}{\beta}(e^{-\beta T} - 1)(1-P)), \quad (6.33)$$

where T is the TPM cycle length and $\mathrm{E}N_f(T)$ denotes the expected number of failures over T. In equation (6.33), for $q = 1$, $\mathrm{E}N_f(T)$ is corresponding to the inspection model with a perfect inspection policy and is given by

$$\mathrm{E}N_f(T) = \lambda(T + \frac{1}{\beta}(e^{-\beta T} - 1)(1-P)). \quad (6.34)$$

Equation (6.34) can be obtained by letting $q = 1$, $t = T_{i-1}$, and $t + \Delta t = T_{i-1} + T = T_i$ in the equation (6.32) as expected.

Since d_b denotes the mean downtime per failure when the TPM cycle length is T and d_i denotes the mean duration of the TPM activity, it follows that the long term measure of the expected downtime per unit time, $\mathrm{E}D(T)$, is

$$\mathrm{E}D(T) = \frac{d_b \mathrm{E}N_f(T) + d_i}{T}. \quad (6.35)$$

Since $d_b = 0.914$, using the fitted model parameters of Table 6.18, we can obtain the resulting model output shown in Figure 6.7.

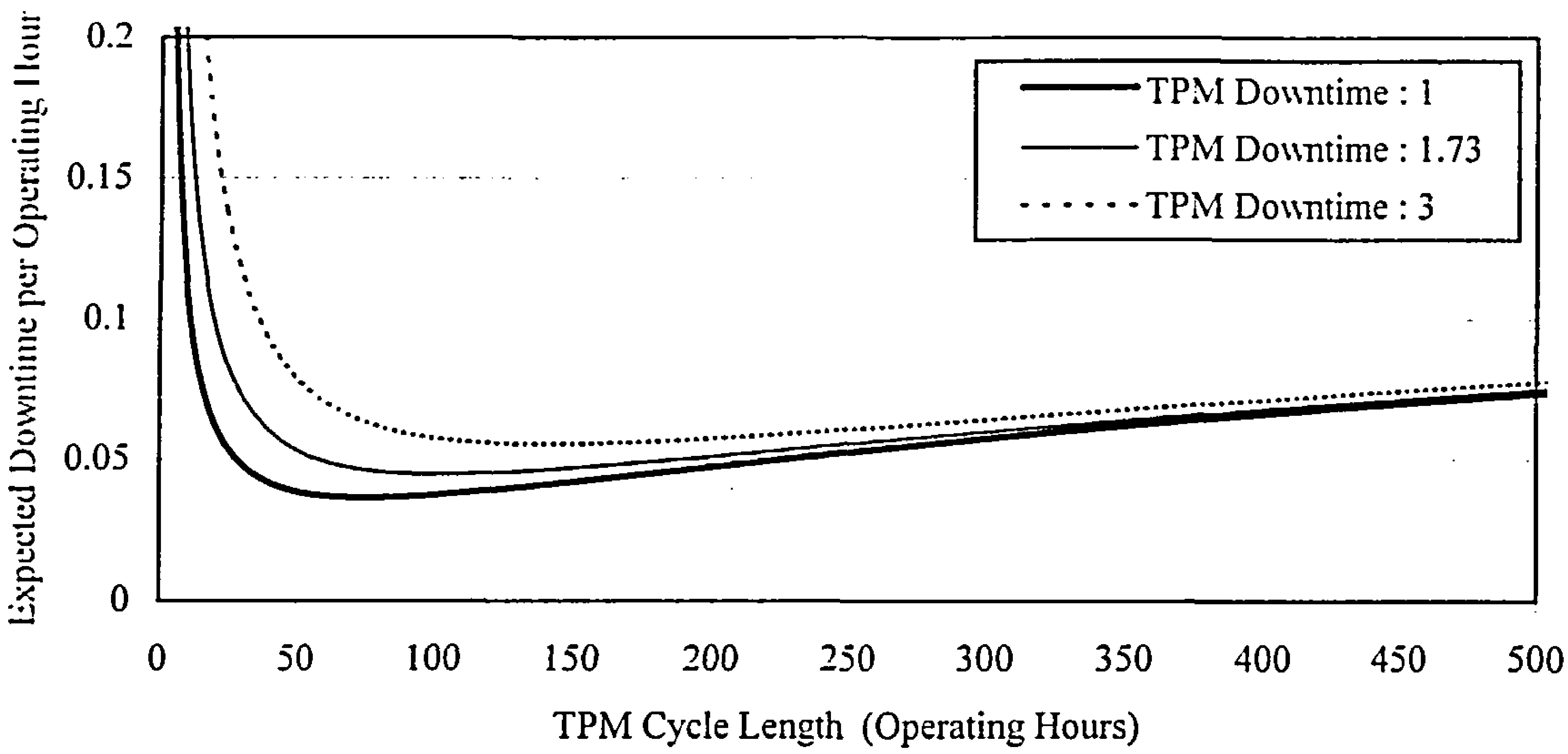


Figure 6.7. Expected downtime per unit operating time and TPM cycle length.

From Figure 6.7. it can be seen that, under the current TPM policy, $d_t = 1.73$, the model suggests that the optimal TPM cycle length is 100 hours and the expected downtime per hour is 0.045 hours. Since the production times per week are about 120 hours, the optimal TPM interval is about 0.83 weeks, but the practical solution with virtually no extra downtime is once a week, and the expected downtime per week is 5.4 hours. Since the company has performed the TPM every 3 weeks, Figure 6.7 shows that the expected downtime per week is 7.8 hours. Accordingly, if the company takes the TPM once per week instead of every 3 weeks, the company can reduce the expected downtime of 2.4 hours per week. Also, if the downtime due to TPM is changed to 1 or 3 hours, the mathematical optimal TPM interval becomes 0.62 or 1.2 weeks respectively.

Questions (a) and (b) of the subjective data survey indicated that the number of defects identified at a TPM is about 45 and the number of defects rectified during a TPM is about 30 given that 100 defects are present, which as commented, seemed very low. If the number of defects identified at a TPM increased to 60 and the number of defects rectified during TPM increase to 60 and 80 under the current TPM policy, $d_t = 1.73$, the probability of identifying and removing a fault during a TPM is $q = 0.36$

and q = 0.48 respectively. A graph indicating the consequences to the downtime model of this change is presented by Figure 6.8.

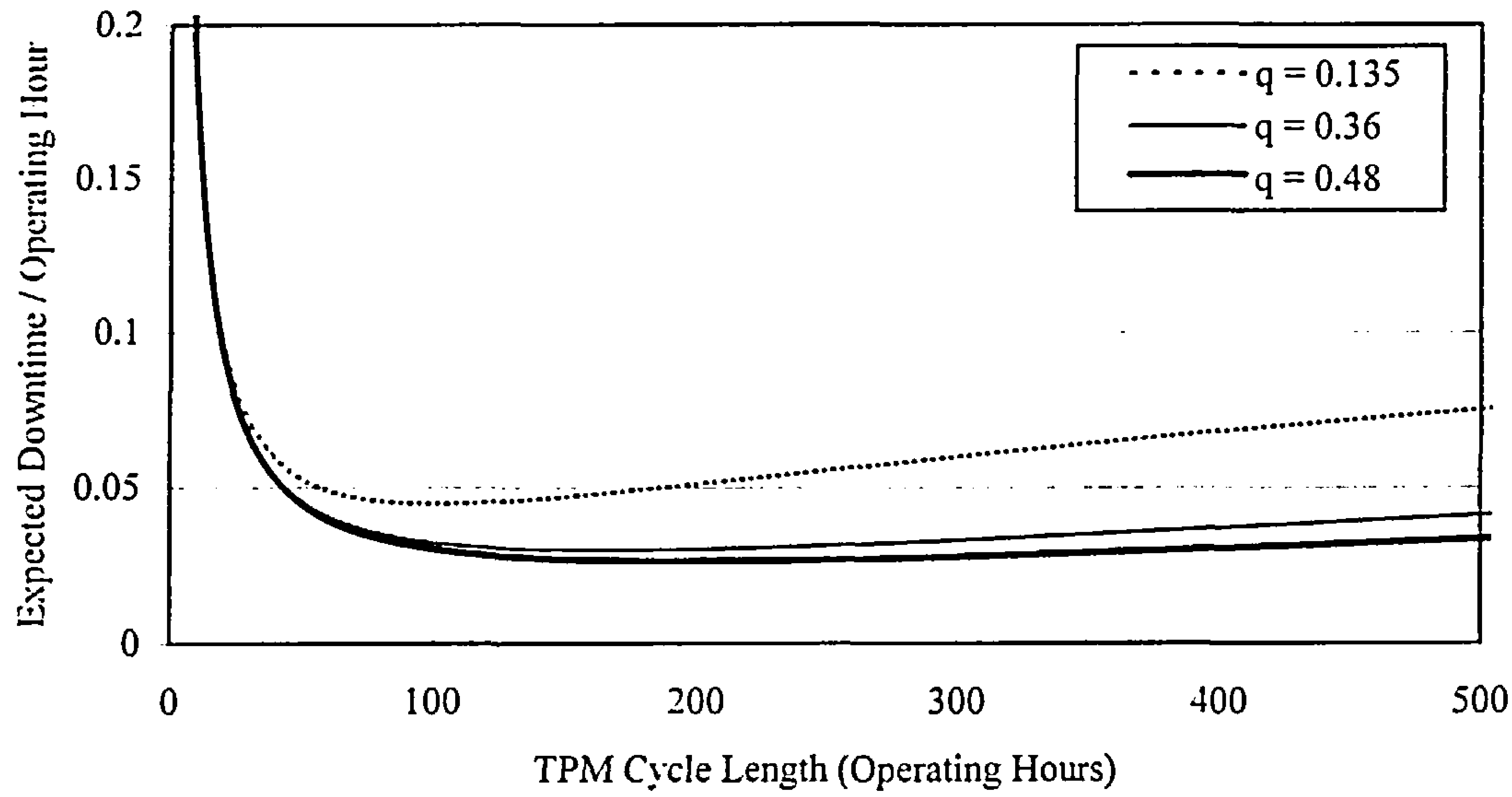


Figure 6.8. Expected downtime per unit time according to TPM cycle length from the delay time model when the probability q changes.

Figure 6.8 produces Table 6.19.

Table 6.19. The optimal TPM interval and the expected downtime per week from the delay time model when the probability q changes.

Probability	Optimal TPM interval		Expected downtime (hours)	
q	in hour	in week	per hour	per week
0.135	100	0.83	0.045	5.40
0.36	165	1.38	0.030	3.48
0.48	195	1.63	0.026	3.12

* The current policy is q = 0.135.

From Table 6.19. we can see that as the probability q increases, the optimal TPM interval increases and the expected downtime per hour reduces. This is as expected since with q increasing, more defects and therefore potential failure, and being removed from the system. It is noted that the modelling here gives the management quantitative insight into the value of improved training and tooling to increase the q measure.

6.6.2 Semi-Markov Inspection Model

In the chapter 5, the semi-Markov inspection model has been discussed. As a consequence of the discussion, we found that the semi-Markov inspection model utilising the delay time concept for parameter estimation can be applied for a multi-component system. However, in applying the semi-markov inspection model to a multi-component system of the real-world situation, the key point is that the fault arriving process must follow a HPP which satisfy the Lemma 5.1 to 5.6. Also, in the case of imperfect inspection, as discussed in the subsection 5.3.2, we can approach the modelling by assuming that the system failure process is in a steady state between inspections and that the defect detecting process at inspection has operated for a long period. Since we assume that the fault arriving process follows a HPP by the assumption (a) of the section 6.4. and our interest is in the reduction of the downtime due to failures and the TPMs over the long term future, we may adopt the semi-Markov inspection model of the subsection 5.3.2 in this case study.

Since the delay time has a mixed delta-exponential distribution, using the equation (5.56), the number of defects identified at time x if there is an inspection at time x, namely $\mathrm{E}N_d(x)$, is given by

$$\mathrm{E}N_d(x) = \frac{\lambda q(1-P)}{\beta}(1+\frac{qe^{-\beta(x-T)}}{1-q-e^{\beta T}}) \tag{6.36}$$

and, using the equation (5.57), the rate function of the failure arrival process, namely $\rho(x)$, is given by

$$\rho(x) = \lambda(1 + \frac{(1-P)qe^{-\beta(x-T)}}{1-q-e^{\beta T}}). \qquad (6.37)$$

Based upon the equations (6.36) and (6.37), we can apply the semi-Markov inspection model of the section 5.3 to the Preformer system.

In section 5.3. the state space was given by

$$I = \{i | i = 0,1,2,\cdots,N\} \cup \{(0, m\Delta t), (m\Delta t, f) | m = 1,2,\cdots,M\},$$

where state i corresponds to the situation in which an inspection identifies i defects within the system, the states (0, $m\Delta t$) corresponds to the situation in which $m\Delta t$ time units have passed since the last inspection, and the states $(m\Delta t, f)$ correspond to the situation in which a breakdown has occurred between $(m-1)\Delta t$ and $m\Delta t$. Also, Δt is an arbitrary small time, N is the number of components of the system, and M is a sufficiently large integer. For numerical solution, we set the upper bound of the number of defects is $N = 30$, the arbitrary small time is $\Delta t = 4$, and the sufficiently large integer is $M = 125$ in this subsection. The possible actions a were denoted by

$$a = \begin{cases} 0, & \text{leave the system as it is,} \\ 1, & \text{inspect the system,} \\ 2, & \text{repair the component.} \end{cases}$$

Based upon the above state space and possible action, the downtime model was given by

$$w_0 = -g_d(R)\Delta t + P_{0(0,\Delta t)}(0)w_{(0,\Delta t)} + P_{0(\Delta t,f)}(0)w_{(\Delta t,f)}, \qquad (6.38)$$

$$w_k = P_{k0}(2)w_0 \qquad \text{for } k = 1,2,\cdots,N, \qquad (6.39)$$

$$w_{(0,m\Delta t)} = -g_d(R)\Delta t + P_{(0,m\Delta t)((m+1)\Delta t,f)}(0)w_{((m+1)\Delta t,f)} + P_{(0,m\Delta t)(0,(m-1)\Delta t)}(0)w_{(0,(m-1)\Delta t)}$$

$$\text{for } 0 < m\Delta t < s, \qquad (6.40)$$

$$w_{(0,m\Delta t)} = d_i - g_d(R)d_i + P_{(0,m\Delta t)0}(1)w_0 + P_{(0,m\Delta t)1}(1)w_1 + \cdots + P_{(0,m\Delta t)N}(1)w_N$$

$$\text{for } s \le m\Delta t \le M\Delta t, \qquad (6.41)$$

and

$$w_{(m\Delta t,f)} = d_b - g_d(R)d_b + P_{(m\Delta t,f)(0,(m+\frac{d_b}{\Delta t})\Delta t)}(2)w_{(0,(m+\frac{d_b}{\Delta t})\Delta t)}$$

$$\text{for } m = 1,2,\cdots,M, \qquad (6.42)$$

where $g_d(R)$ is the average downtime per unit time given policy R with parameter value s, w_x, $x \in I$, are the relative downtimes resulting from the various starting states when policy R with parameter value s is used, and $P_{ij}(a)$ is the probability that at the next decision epoch the system will be in state j if actions a is chosen in the present state i.

Using equations (6.36) and (6.37), we can obtain the one-step transition probabilities $P_{ij}(a)$. Firstly, if the action $a = 0$ is taken at state 0, the system will either survive until the next decision epoch Δt or fail within the next decision epoch Δt. In equation (6.38), for very small Δt, the one-step transition probabilities $P_{0j}(0)$, for $j = (0,\Delta t),(\Delta t,f)$, are given by (see equation (5.3))

$$P_{0j}(0) = \begin{cases} \int_0^{\Delta t} \rho(x)dx & \text{for } j = (\Delta t, f) \\ 1 - \int_0^{\Delta t} \rho(x)dx & \text{for } j = (0, \Delta t) \\ 0 & \text{otherwise}. \end{cases} \qquad (6.43)$$

In equation (6.39), since the action $a = 2$ is taken at state k, $k = 1, 2, \ldots, N$, we have that (see equation (5.7))

$$P_{k0}(2) = 1 \qquad \text{for } k = 1, 2, \ldots, N. \qquad (6.44)$$

In equation (6.40), since the action $a = 0$ is taken at state $(0, m\Delta t)$ with $m = 1, 2, \ldots, M\text{-}1$, the system will either survive until the next decision epoch $(m+1)\Delta t$ or fail within the next decision epoch $(m+1)\Delta t$ having survived to the present decision epoch $m\Delta t$. Accordingly, for very small Δt, the one-step transition probabilities $P_{(0,m\Delta t)j}$, for $j = (0,(m+1)\Delta t),((m+1)\Delta t, f)$, are given by (see equation (5.11))

$$P_{(0,m\Delta t)j}(0) = \begin{cases} \int_{m\Delta t}^{(m+1)\Delta t} \rho(x)dx & \text{for } j = ((m+1)\Delta t, f) \\ 1 - \int_{m\Delta t}^{(m+1)\Delta t} \rho(x)dx & \text{for } j = (0,(m+1)\Delta t) \\ 0 & \text{otherwise .} \end{cases} \qquad (6.45)$$

In equation (6.41), if action $a = 1$ is taken at state $(0, m\Delta t)$ with $m=1, 2, \ldots, M$, the inspection will result in a situation of finding j, $j = 0, 1, 2, \ldots, N$, faults at an inspection. Since the number of defects identified at an inspection has a Poisson distribution by the Lemma 5.3, we have that (see equation (5.15))

$$P_{(0,m\Delta t)j}(1) = \begin{cases} e^{-\mathrm{E}N_d(m\Delta t)} \dfrac{(\mathrm{E}N_d(m\Delta t))^j}{j!} & \text{for } j = 0,1,2,\cdots,N \\ 0 & \text{otherwise .} \end{cases} \qquad (6.46)$$

Lastly, in equation (6.42), since the action $a = 2$ is taken at state $(m\Delta t, f)$, $m = 1, 2, \ldots, M$, we have that (see equation (5.19))

$$P_{(m\Delta t, f)(0,(m+\frac{d_s}{\Delta t})\Delta t)}(2) = 1 \qquad \text{for } m = 1, 2, \ldots, M. \qquad (6.47)$$

Using the embedded technique, we can obtain the expected average downtime per unit time $g_d(R)$ from equations (6.38) to (6.42). By putting one of the relative downtimes equal to zero, say $w_0 = 0$, the linear equation can determine uniquely the average downtime per unit time $g_d(R)$.

Using the fitted model parameters of the Table 6.18 and the above mentioned downtime model, the resulting model output of downtime as a function of TPM period is shown in Figure 6.9.

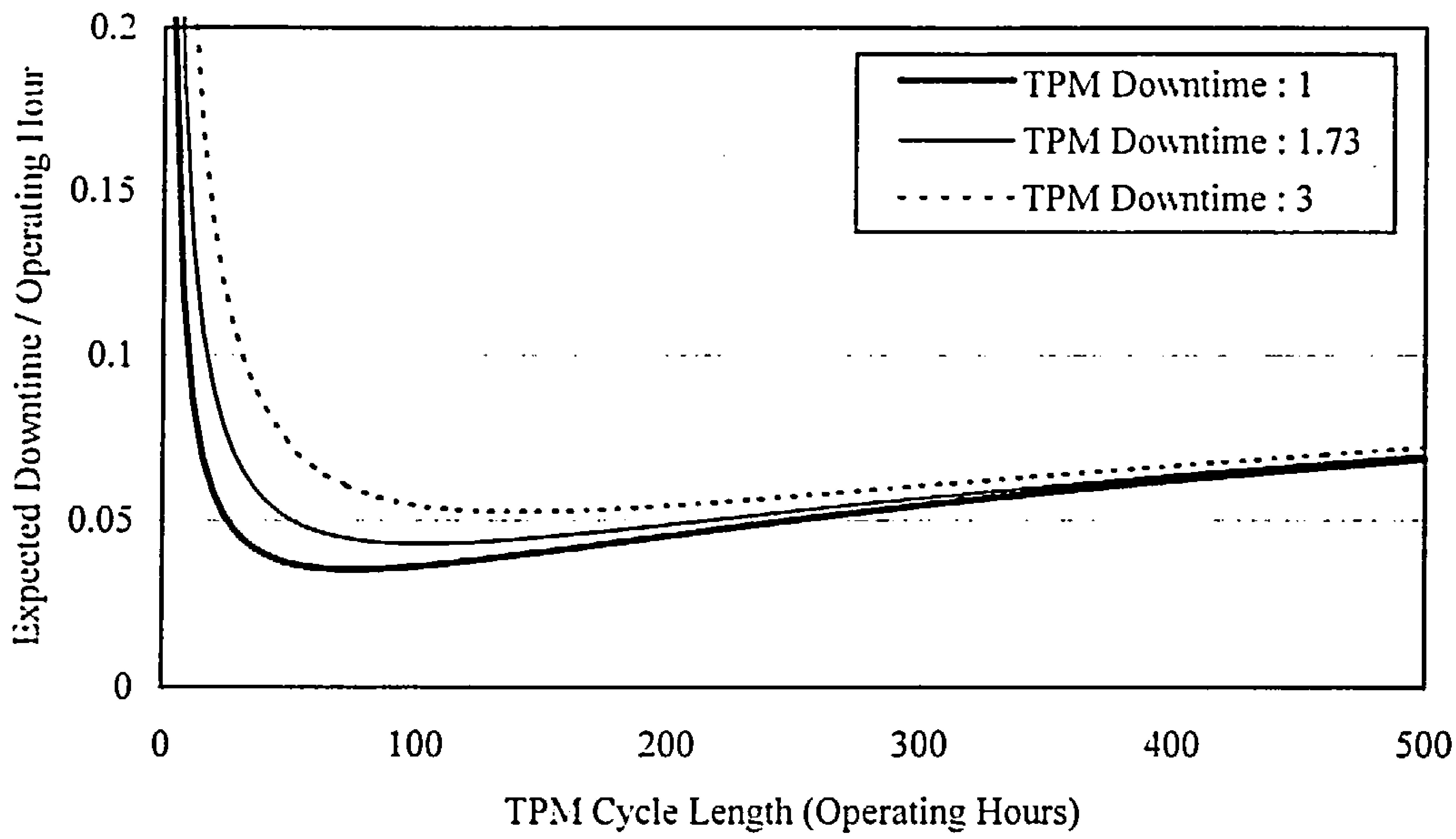


Figure 6.9. Expected downtime per unit time according to TPM cycle length from the semi-Markov inspection model with step size $\Delta t = 4$.

From Figure 6.9, it can be seen that, under the current TPM policy, $d_l = 1.73$, the model suggests that the optimal TPM cycle length is 100 hours and the expected downtime per hour is 0.043 hours. Since the production times per week are about 120 hours, the optimal TPM interval is about 0.83 weeks and the expected downtime per week is 5.16 hours. Since the company has performed the TPM every 3 weeks, Figure 6.8 shows that the expected downtime per week is 7.32 hours. Accordingly, if the company takes the TPM once per week instead of every 3 weeks, the company can reduce the expected downtime of 2.16 hours per week. Also, if the downtime due to TPM is about 1 and 3 hours, the optimal TPM interval should be 0.63 and 1.2 weeks respectively. These results of the semi-Markov inspection model are as expected, very similar to the results of the delay time model.

Also, as with the delay time model, if the number of defects identified at TPM increase to 60 and the number of defects rectified during TPM increase to 60 and 80 under the current TPM policy, $d_i = 1.73$, the probability of identifying and removing a fault during a TPM is q=0.36 and q=0.48 respectively. A graph for this change is presented by Figure 6.10.

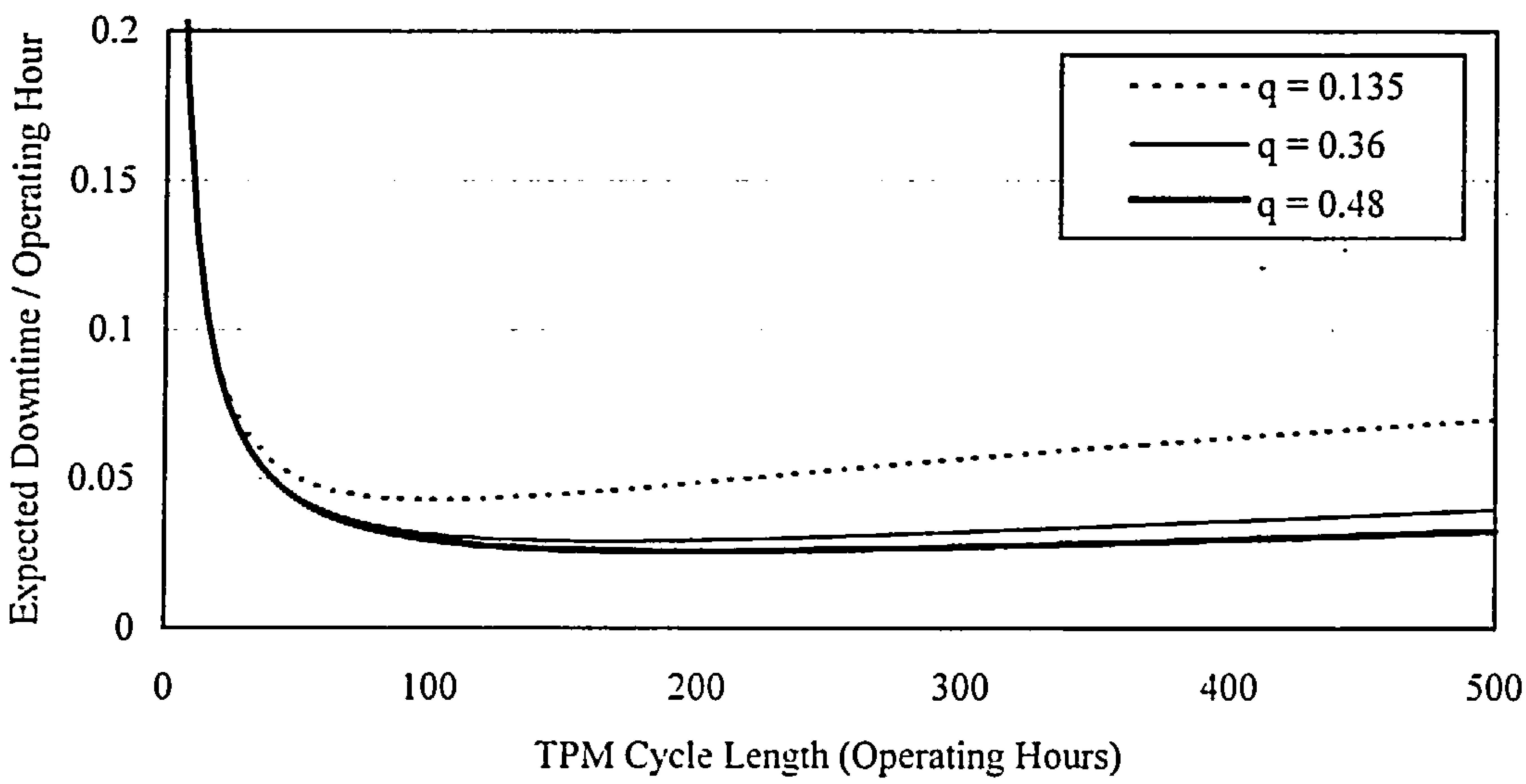


Figure 6.10. Expected downtime per unit time according to TPM cycle length from the semi-Markov model when the probability q changes.

Figure 6.10 produces Table 6.20.

Table 6.20. The optimal TPM interval and the expected downtime per week from the semi-Markov model when the probability q changes.

Probability	Optimal TPM interval		Expected downtime (hours)	
q	in hour	in week	per hour	per week
0.135	100	0.83	0.043	5.16
0.36	164	1.37	0.029	3.48
0.48	196	1.63	0.026	3.12

* The current policy is $q = 0.135$.

From Table 6.20. we can see that as the probability q increases. the optimal TPM interval increases and the expected downtime per hour reduces. This is as expected since with q increasing, more defects and therefore potential failure, and being removed from the system. These results of the semi-Markov inspection model are very similar to the results of the delay time model.

6.7 Comparison and Conclusions

To compare the semi-Markov inspection model with the delay time model in detail, we can take the curve of the current TPM policy, $d_i = 1.73$, from Figure 6.7 of the delay time model and the curve of the current TPM policy, $d_i = 1.73$, from Figure 6.9 of the semi-Markov inspection model respectively. From both models, Figure 6.11 is presented.

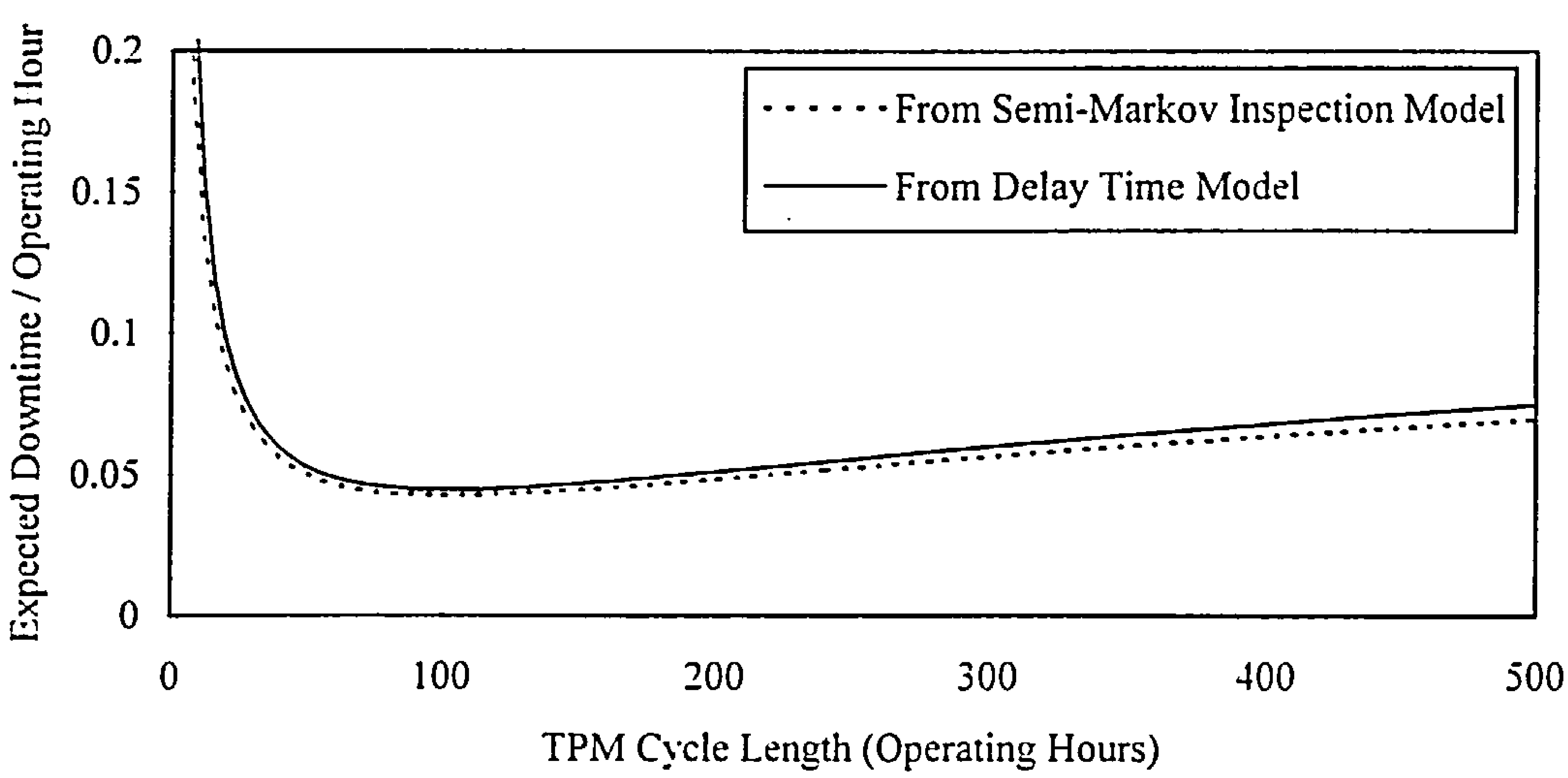


Figure 6.11. Expected downtime per unit time according to TPM cycle length for the comparison.

Figure 6.11 shows that the semi-Markov inspection model curve is consistent with the delay time model curve. From both these models, an optimal TPM period point which minimises the expected downtime per unit time can be obtained.

From the perspective of the output from both models, we can now address the two questions outlined in the introduction, namely (1) whether TPM can identify most faults present and thereby reduce the number of failures caused by those faults, and (2) whether the current TPM period is the right choice, particularly the current three-week TPM cycle. Clearly, as shown in Figure 6.3, TPM is effective for the Preformer because it can reduce the total downtime due to failures. However, since the downtime caused by TPM of the Preformer is currently about 1.73 hours, for the current TPM policy, the optimal TPM cycle length is about weekly. Should the downtime due to TPM of the Preformer increases to 3 hours, the optimal TPM cycle length is still less than 2 weeks. This indicates that the current TPM policy for the Preformer, which is about three weeks, is not appropriate. Accordingly, we suggest that if it is possible, the company need to reduce the TPM cycle length of the Preformer to weekly. If the company undertakes TPM of the Preformer about weekly, the company can reduce the expected downtime by 2.4 hours per week, from 7.8 hours per week to 5.4 hours per week according to the delay time model and 2.16 hours per week from 7.32 hours per week to 5.16 hours per week based upon the semi-Markov inspection model.

Also, if the probability of identifying and removing a fault during TPM q increases by improving the engineering ability in identifying and rectifying faults during TPM, perhaps through training, the expected downtime can furthermore be reduced. As shown in Table 6.19 and 6.20, if the probability of identifying and removing a fault during TPM q increases from 0.135 to 0.36, the company can reduce the expected downtime of 4.32 hours per week from 7.8 hours per week to 3.48 hours per week in the delay time model and 3.84 hours per week from 7.32 hours per week to 3.48 hours in the semi-Markov inspection model respectively for current TPM policy of every 3 weeks. Also, if the probability of identifying and removing a fault during TPM q increases from 0.135 to 0.48, the company can reduce the expected downtime by 4.68 hours per week, from 7.8 to 3.12 hours per week based upon the delay time model and by an expected 4.2 hours per week, from 7.32 to 3.12 hours per week according to the semi-Markov inspection model, all assuming the current TPM policy of having every 3 weeks.

In this case study, we propose an OR and statistical approach to model TPM practice for a Preformer in Brakelining Ltd. Modelling is based upon the delay time concept. A statistical model based upon failure and TPM data has been established to give the estimated values for model parameters, and a TPM model has been derived to find the optimal TPM cycle length in terms of minimising the total expected downtime caused by failures and TPMs from the delay time model and the semi-Markov inspection model. It has proved to be successful in that it recovered the underlying delay time distribution using conventional maximum likelihood method, and the consequential downtime model reflected the current and previous operating downtime level adequately.

The problems encountered in practical modelling, in establishing the estimation procedure for parameters and in validating the modelling, have been highlighted. The recommendations to the company were based upon the modelling reported here. There is an attempt being made to move to weekly TPM's and to monitor results. Other recommendations arise from the discovery of data that were not available, namely TPM data on condition found and faults rectified. The need to collect such information if the maintenance process is to be managed cost effectively has been outlined, and how the data will or may be used indicated. As in other applied studies, this modelling exercise will, we hope, initiate a cultural change in the maintenance management process within the company and see a move to greater qualification and, therefore, control.

Chapter 7

CONCLUSION

In this thesis, we have considered the inspection policy of facilities which gradually deteriorate in time and eventually fail. The inspection policy is some activity carried out at intervals, with the intention of reducing or eliminating the number of failures occurring, or of reducing the consequences of failure in terms of downtime or operating cost. It has been seen in the literature review of the chapter 2 that there are a great many models addressing the problem of finding the optimal inspection policy. Some authors developed the two-state models in which the working condition of the system was expressed as one of two states, operating or failed, based upon the time to failure. Since the two-state models may not allow for an inspection, which leads to repair before a possible failure of the system, most of the published theoretical models for finding the optimal inspection policy adopt a multi-state Markov approach where the states are operating, operating but fault present, and failed. Generally speaking, most such models have assumed that the working condition of the system can be expressed as a discrete-time Markov chain with a new state, degraded states, and a failed state, and the transition probabilities are assumed to be given. In practice, it is, however, difficult to define the degraded states for the deteriorated system, and more difficult to determine the state transition probabilities. So, most authors do not mention the fit of their model to data, and present no examples of actual applications or case studies utilising their model.

In contrast to the Markov models, the delay time concept has, however, provided a useful means of modelling the effect of periodic inspections on the failure rate of repairable machinery. The delay time concept regards the failure process as basically a two-stage process, but three if one includes failure. First, a defect can be first identified at time u if an inspection is carried out at that time. If the defect is not identified, the faulty component subsequently fails after further interval h which is called the delay time of the defect. As seen in chapter 3, the delay time concept has been increasingly used in inspection modelling. Also, it has been noted that the introduction of the delay time concept in inspection modelling has provided a powerful tool in modelling and validating the relationship between inspection actions and the consequences of these actions. Obviously, the successful use of the delay time concept in maintenance modelling depends upon how well the underlying delay time distribution can be estimated from available information sources. Two methods for estimating the parameters of the delay time modelling, namely the subjective method and the objective method, have been presented. With such a method, numerous applied studies of the delay time concept have been developed since the first was published in 1984. Since delay time models can be used for decision-making, for example choosing the interval between inspections to minimise cost or downtime, it may be natural to rely on the delay time modelling in adapting the maintenance models to real-world situation.

To overcome the restrictions of the Markov models, we have used the delay time concept in the chapter 4 and 5. By defining the degraded states of the Markov chain as the number of defects within the context of the delay time concept, we can readily define the working condition of the system as a Markov chain. Also, by using the parameters of the delay time model which can be estimated by the subjective method or the objective method, the state transition probabilities of the Markov model can be calculated. Under these conditions, a typical semi-Markov inspection model based upon the delay time concept for a component and for a complex repairable system that may fail during the course of its service lifetime has been established.

Firstly, for a single component system, we have shown that the semi-Markov inspection model is consistent with the delay time model in the system with a Markov property of the initial point u. In real-world situation, however, it is expected that there are relatively few cases satisfying the Markov property and inadequate data fit the model. Although data may be available in application studies, it may be difficult to confirm the Markov property and to determine the state transition probabilities from them. Either way, if the system has a Markov property, by utilising the deterioration probability $r_{ii}^{\prime}$ which, as we have shown, can be estimated based upon the delay time concept, the state transition probabilities can be calculated. Relying on the delay time concept, the semi-Markov inspection model can now be used for the first time in practice for the single component system. If the initial point u does not satisfy the Markov property, the semi-Markov inspection model can become a very poor approximation to the actual inspection process. In contrast to the semi-Markov inspection model, the delay time model can not only give the optimal inspection period point, but also the delay time model is consistent with the simulation procedure. We can, therefore, see that the delay time model can be fitted to any system regardless of a Markov property.

As well as the semi-Markov inspection model for a single component system, the semi-Markov inspection model for a multi-component system can be applied to the system regardless of the Markov property of the delay time h. We need to note that these models rely on the delay time concept in order to determine parameters. The delay time concept provides a means of not only denoting the working condition of the system as the degraded states of the semi-Markov inspection model, but also of obtaining the state transition probability from data of the real-world situation through estimating the fault arrival rate and the parameters of the delay time distribution. The semi-Markov inspection model based upon the delay time concept is, perhaps for the first time, available to usefully apply to real-world situations. Here we can see the importance of the delay time concept. In practice, having delay time parameters, a delay time would, of course, usually be preferred to a semi-Markov model.

When we establish the semi-Markov inspection model for a multi-component system, we have to note Lemmas 5.1 to 5.6 which are based upon the delay time concept. Then, after some complicated manipulation, we can formulate the semi-Markov inspection model of the section 5.3 which is fitted to the multi-component system with a perfect inspection policy or an imperfect inspection policy. Since the equations of the semi-Markov inspection model are complicated, it takes a relatively long time to compute the equations of the semi-Markov inspection model. Also, in applying the semi-Markov inspection model to the real-world situation, the key point is that the fault arrival process follows a HPP which satisfy the Lemmas 5.1 to 5.6. If the fault arrival process follows a NHPP, we cannot apply the semi-Markov inspection model to the real-world situation because Lemma 5.1 to 5.6 are not satisfied.

In applying to a real-world situation, the delay time model is consistent with the semi-Markov inspection model as discussed in the chapter 5. However, in contrast to the semi-Markov inspection model, the delay time model consists of the simpler equations. Thereby, it does not require a long computing time compared to the semi-Markov inspection model for computing the equations of the delay time model. Also, the delay time model provides a means of modelling the behaviour of the system and predicting such useful quantities as reliability, cost or downtime under various inspection policies. If the fault arrival rate regardless of a HPP or a NHPP and the parameters of the *pdf* of the delay time h, $f(h)$, regardless of any distributions, are estimated from the data of the real-world situation using the subjective or objective estimation method, we can easily establish the delay time model which can find the optimal inspection policies minimising the expected total cost per unit time or the expected total downtime per unit time. As confirmed in the numerical example of the chapter 5, it was shown that the simulation model and semi-Markov model are nearly consistent with the delay time model. This means that the delay time model and semi-Markov model can both apply practically to the multi-component system.

Also, it was seen in the case study of the chapter 6 that the semi-Markov inspection model is consistent with the delay time model. From both the models, an optimal TPM period which minimises the expected downtime per unit time were obtained which were consistent. This is the first known case of a Markov inspection model being built for an actual application. In this case study, we proposed an OR and statistical approach to model TPM practice on the Preformer machine operated within Brakelining Ltd, and based upon the delay time concept. A statistical model based upon failure and TPM data can be established to give the estimated values for model parameters, and a TPM model can be derived to find the optimal TPM cycle length in terms of minimising the total expected downtime caused by failures and TPMs from the delay time model and the semi-Markov inspection model. It has proved to be successful in that it recovered the underlying delay time distribution using conventional maximum likelihood method, and the consequential downtime model reflected the current and previous operating downtime level adequately.

In conclusion. since not only is the delay time model free from the requirement of a HPP fault arrival process, but it also requires much less time than the semi-Markov inspection model for computing the equations, the delay time model is considered to be more general and practicable than the semi-Markov inspection model in applying to real-world situations. Furthermore, the delay time concept has an evident contribution in rending semi-Markov inspection models more applicable. When establishing a Markov model for a component inspection problem or a complex system, it is clearly very important to establish the validity of the Markov assumption and appropriate measures of the parameters.

REFERENCES

[1] Anbar. D. [1976], "An asymptotically optimal inspection policy", *Naval Research Logistics Quarterly*, Vol.23, pp.211-218.

[2] Assaf. D. & Shanthikumar, J. G. [1987], "Optimal group maintenance policies with continuous and periodic inspections", *Management Science*, Vol.33, pp.1440-1452.

[3] Baker. R. D. & Christer, A. H. [1994], "Review of delay-time OR modelling of engineering aspects of maintenance", *European Journal of Operational Research*, Vol.73, pp.407-422.

[4] Baker, R. D. & Wang, W. [1992], "Estimating the delay-time distribution of faults in repairable machinery from failure data", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.3, pp.259-281.

[5] Baker. R. D. & Wang, W. [1993], "Developing and testing the delay time model", *Journal of the Operational Research Society*, Vol.44, pp.361-374.

[6] Barlow, R. E. & Proschan, F. [1965], *Mathematical theory of reliability*, John Wiley, New York.

[7] Beichelt, F. [1981], "Minimax inspection strategies for single unit systems", *Naval Research Logistics Quarterly*, Vol.28, pp.375-381.

[8] Butler, D. A. [1979], "A hazardous-inspection model", *Management Science*, Vol.25, pp.79-89.

[9] Chilcott, J. B. & Christer, A. H. [1991], "Modelling of condition-based maintenance at coal face", *International Journal of Production Economics*, Vol.22, pp.1-11.

[10] Cho. D. I. & Parlar, M. [1991], "A survey of maintenance models for multiunit systems", *European Journal of Operational Research*, Vol.51, pp.1-23.

[11] Chou. C. K. & Butler, D. A. [1983], "Assessment of hazardous-inspection polices", *Naval Research Logistics Quarterly*, Vol.30, pp.171-177.

[12] Christer, A. H. [1982], "Modelling inspection policies for building maintenance", *Journal of the Operational Research Society*, Vol.33, pp.723-732.

[13] Christer, A. H. [1984], "Operational research applied to industrial maintenance and replacement", *Developments in Operational Research*, Pergamon Press, Oxford.

[14] Christer, A. H. [1987], "Delay time model of reliability of equipment subject to inspection monitoring", *Journal of the Operational Research Society*, Vol.38, pp.329-334.

[15] Christer, A. H. [1988], "Condition-based inspection models of major civil-engineering structures", *Journal of the Operational Research Society*, Vol.39, pp.71-82.

[16] Christer, A. H. & Redmond, D. F. [1990], "A recent mathematical development in maintenance theory", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.2, pp.97-108.

[17] Christer, A. H. & Redmond, D. F. [1992], "Revising models of maintenance and inspection", *International Journal of Production Economics*, Vol.24, pp.227,234.

[18] Christer, A. H. & Redmond, D. F. [1993], "Modelling the deterioration and maintenance of concrete structures", *Paper Presented at the 12th IFORS Conference*, Lisbon.

[19] Christer, A. H. & Waller, W. M. [1984a], "Delay time models of industrial inspection maintenance problems", *Journal of the Operational Research Society*, Vol.35, pp.401-406.

[20] Christer, A. H. & Waller, W. M. [1984b], "An operational research approach to planned maintenance : Modelling P.M. for a vehicle fleet", *Journal of the Operational Research Society*, Vol.35, pp.967-984.

[21] Christer, A. H. & Waller, W. M. [1984c], "Reducing production downtime using delay-time analysis", *Journal of the Operational Research Society*, Vol.35, pp.967-984

[22] Christer, A. H. & Wang, W. [1995], "A delay-time-based maintenance model of a multi-component system", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.6. pp.205-222.

[23] Christer, A. H.. Wang, W., Baker. R. D. & Sharp, J. [1994], "Operational modelling of preventive maintenance of production plant", In *Proceeding of IFRIM workshop 1994*, Eindhoven, The Netherlands.

[24] Christer, A. H.. Wang, W., Baker. R. D. & Sharp, J. [1995], "Modelling maintenance practice of production plant using the delay time concept", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.6, pp.67-84.

[25] Christer, A. H. & Whitelaw, J. [1983], "An operational research approach to breakdown maintenance: Problem recognition", *Journal of the Operational Research Society*, Vol.34, pp.1041-1052.

[26] Desa, M. I. [1995], "Bus fleet maintenance modelling in a developing country", *Ph.D. Thesis*, Department of Mathematics and Computer Science, University of Salford.

[27] Devooght, J., Dubus, A. & Smidts, C. [1990], "Suboptimal inspection policies for imperfectly observed realistic systems", *European Journal of Operational Research*, Vol.45, pp.203-218.

[28] Dias, J. R. [1990], "A new approximation for the inspection period of systems with different failure rates", *European Journal of Operational Research*, Vol.45, pp.219-223.

[29] Hopkins, T. & Phillips, C. [1988], *Numerical methods in practice using the NAG library*, Addison-Wesley, Wokingham.

[30] Jardine, A. K. S. & Hassounah, M. I. [1990], "An optimal vehicle-fleet inspection schedule", *Journal of the Operational Research Society*, Vol.41, pp.791-799.

[31] Kaio, N. & Osaki, S. [1986], "Optimal inspection policy with two types of imperfect inspection probabilities", *Microelectronics and Reliability*, Vol.26, pp.935-942.

[32] Kander, Z. [1978], "Inspection policies for deteriorating equipment characterised by N quality levels", *Naval Research Logistics Quarterly*, Vol.36, pp.234-255.

[33] Law, A. M. & Kelton, W. D. [1982], *Simulation Modelling and Analysis*, McGraw-Hill Book Company, New York.

[34] Leung, F. K. & Christer, A. H. [1995], "The reliability of water supply in High Rise Housing in Hong Kong", *Paper Presented at the 2nd IMA Iternational Conference on Mathematical Models in Maintenance*, University of Salford.

[35] Luss, H. [1983], "An inspection policy model for production facilities", *Management Science*, Vol.29, pp.1102-1109.

[36] Luss, H. & Kander, Z. [1974], "Inspection policies when duration of checkings is non-negligible", *Operational Research Quarterly*, Vol.25, pp.299-309.

[37] Milioni, A. Z. & Pliska, S. R. [1977], "Optimal inspection under semi-Markovian deterioration : the catastrophic case", *Naval Research Logistics Quarterly*, Vol.35, pp.393-411.

[38] Mine, H. & Kawai, H. [1975], "An optimal inspection and replacement policy", *IEEE Transactions on Reliability*, Vol.R-24, pp.305-309.

[39] Munford, A. G. [1990], "Optimal inspection policies where penalty costs are proportional to the duration of the inspection interval containing the failure", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.2, pp.247-254.

[40] Munford, A. G. & Shahani, A. K. [1972], "A nearly optimal inspection policy", *Operational Research Quarterly*, Vol.23, pp.373-379.

[41] Munford, A. G. & Shahani, A. K. [1973], "An inspection policy for the Weibull case", *Operational Research Quarterly*, Vol.24, pp.453-458.

[42] Nakagawa, T. [1984], "Periodic inspection policy with preventive maintenance", *Naval Research Logistics Quarterly*, Vol.31, pp.33-40.

[43] Nakagawa, T. & Yasui, K. [1979], "Approximation calculation of inspection policy with Weibull failure times", *IEEE Transactions on Reliability*, Vol.R-28, pp.403-404.

[44] Nakagawa, T. & Yasui, K. [1980], "Approximation calculation of optimal inspection times", *Journal of the Operational Research Society*, Vol.31, pp.851-853.

[45] Nakajima, S. [1989], *TPM Development Program*, Productivity Press, Cambridge.

[46] Ohnishi, M., Kawai, H. & Mine, H. [1986a], "An optimal inspection and replacement policy for a deteriorating system", *Journal of Applied Probability*, Vol.23, pp.973-988.

[47] Ohnishi, M., Kawai, H. & Mine, H. [1986b], "An optimal inspection and replacement policy under incomplete state information", *European Journal of Operational Research*, Vol.27, pp.117-128.

[48] Özekici, S. & Pliska, S. R. [1991], "Optimal scheduling of inspections : a delayed Markov model with false positives and negatives", *Operations Research*, Vol.39, PP.261-273.

[49] Pierskalla, W. P. & Voelker, J. A. [1976], "A survey of maintenance models : the control and surveillance of deteriorating systems", *Naval Research Logistics Quarterly*, Vol.23, pp.353-388.

[50] Pintelon, L. M. & Gelders, L. F. [1992], "Maintenance management decision making", *European Journal of Operational Research*, Vol.58, PP.301-317.

[51] Ross, S. M. [1983], *Stochastic Processes*, Wiley, New York.

[52] Sakamoto, Y., Ishiguro, M., & Kitagawa, G. [1986], *Akaike Information Criterion Statistics*. KTK Scientific Publishers. Tokyo.

[53] Sherif. Y. S. & Smith, M. L. [1981]. "Optimal maintenance models for systems subject to failure - a review", *Naval Research Logistics Quarterly*, Vol.28, pp.47-74.

[54] Thomas, L. C. [1986], "A survey of maintenance and replacement models for maintainability of multi-item systems". *Reliability Engineering*, Vol.16, pp.297-309.

[55] Thomas. L. C., Gaver, D. P. & Jacobs, P. A. [1991], "Inspection models and their application", *IMA Journal of Mathematics Applied in Business and Industry*, Vol.3, pp.283-303.

[56] Tijms, H. C. [1986], *Stochastic Modelling and Analysis: A Computational Approach.* Wiley, New York.

[57] Tijms, H. C. & Van der Duyn Schouten, F. A. [1985], "A Markov decision algorithm for optimal inspections and revisions in a maintenance system with partial information", *European Journal of Operational Research*, Vol.21. pp.245-253.

[58] Valdez-Flores, C. & Feldman, M. [1989], "A survey of preventive maintenance models for stochastically deteriorating single-unit systems", *Naval Research Logistics Quarterly*, Vol.36, pp.419-446.

[59] White, D. J. [1993], "A survey of applications of Markov decision processes", *Journal of the Operational Research Society*, Vol.44, pp.1073-1096.

[60] Wijnmalen, D. J. D. & Hontelez, J. A. M. [1992], "Review of a Markov decision algorithm for optimalilnspections and revisions in a maintenance system with partial information", *European Journal of Operational Research*, Vol.62. pp.96-104

[61] Wang, W. [1996], "Subjective estimation of the delay time distribution in maintenance modelling", To be appeared in *European Journal of Operational Research*.