Modelling and simulation for the joint optimisation of inspection maintenance and spare parts inventory in multi-line production settings

<u>Farhad Zahedi-Hosseini</u>

PhD Thesis

Salford Business School

University of Salford

2017

Preface

The studies developed in this PhD thesis form the basis for three journal papers, as detailed below:

1. The material in Chapter 3 has been developed into a paper for the *European Journal of Industrial Engineering.* The original manuscript was submitted on 6th October 2016. The revised manuscript was submitted on 29th April 2017; this is currently under review.

2. The study in Chapter 4 has been developed into Zahedi-Hosseini et al. (2017), which was published online on 12th March 2017, and appeared in the Journal of *Reliability Engineering and System Safety* 168: 306-316.

3. Finally, the work in Chapter 5 is being prepared for a paper to be submitted to the *Journal of Manufacturing Systems*.

Abstract

A simulation methodology is developed to model the joint optimisation of preventive maintenance and spare parts inventory in multi-line settings. The multi-line machines are subject to failure, based on the delay-time concept, and a selection of policies are used for the replenishment of the machines' critical component. Production lines of varied configurations are modelled and described in three principal chapters.

Firstly, the optimisation of preventive maintenance for a multi-line production system is developed in the context of a case study. The policy proposed indicates that consecutive inspection with priority for failure repair is cost-optimal, which suggests a substantial maintenance cost reduction of 61% compared to the *run-to-failure* policy. The contribution of this study is first and foremost in narrowing the gap between the theory and practice of managing multi-line systems, and in particular, that the scenarios and policies considered have important economic and engineering implications.

In a second study, spare parts provisioning for a single-line system is considered, given that the demand for industrial plant spare parts should be driven, at least in part, by maintenance requirements. A paper-making plant provides a real context, for which simulation models are developed to jointly optimise the planned maintenance and the associated spare part inventory. This challenge is addressed in the context of the failure of parts in service and the replacement of defective parts at inspections of period T, using the delay-time concept, and a selection of replenishment policies. The results indicate that a periodic review policy with replenishment twice as frequent as inspection is cost-optimal. Further discussions and sensitivity analysis give insights into the characteristics and features of the policies considered.

Finally, in the third study, the joint optimisation of preventive maintenance and the associated spare parts inventory for a multi-line system is developed using an idealised context. It is found that a review policy with inspection as frequent as replenishment using *just-in-time* (JIT) ordering is cost-optimal, and also the lowest risk policy; it is associated with the lowest simultaneous machine downtime and low stock-out cost-rates. This is a significant contribution to the literature.

An implication of the proposed methodology is that, where mathematical modelling is intractable, or the use of certain assumptions make them less practical, simulation modelling is an appropriate solution tool. Throughout this thesis, the long-run average cost per unit time or cost-rate is used as the optimality criterion. In other contexts, one may wish to use availability or reliability instead. To do so would not change the methodology that is presented here.

IN THE NAME OF GOD

This PhD thesis is dedicated to the following people:

To my dear **parents**,
for their complete support during my long study years in the UK,
especially, to my late beloved **father**,
whose wish was for me to complete my PhD - may God bless his soul;

To my beautiful wife, **Houri**,
for her love, kindness, understanding, patience, and hard work;

To my gorgeous daughter, Dr. **Delaram** Zahedi, MBChB,
who makes us so proud, and whose love and kindness is eternal.

Acknowledgements

I would like to thank the following people whose help, guidance, advice, and encouragement, were pivotal in the completion of this part-time PhD:

- To my main supervisor, **Professor Philip Scarf** - a true expert in maintenance and mathematical modelling, whose vision, help, encouragement and deep knowledge of the maintenance subject was invaluable in enabling me to better understand, analyse, and interpret challenging issues in maintenance. Phil's email discussions, sometimes during late evenings and weekends, were invaluable, helpful, supportive, and highly appreciated.

- To my original supervisor, the late **Professor Wenbin Wang**, who sadly passed away last year. His enthusiasm, help, guidance, and encouragement was invaluable in the early stages of my research work. May God bless his soul. To my second supervisor, **Professor Aris Syntetos**, whose help and advice in later years of my PhD was very helpful.

- To **Professor David Howard**, my colleague, mentor, and a true friend, who helped and encouraged me in various ways to start and continue my PhD. His encouragements and long discussions gave me the audacity to start my PhD despite my family commitments and full-time work. I'm very thankful to him for the continuation of help and encouragement all throughout my PhD.

- To **Professor Tim Ritchings**, our previous Head of School, **Professor Sunil Vadera**, our current Dean of School, and **Dr Tony Jones** for their support and authorising the fee waiver for my PhD.

Table of Contents

List of Tables

List of Figures

Figures associated with Chapter 4

Figures associated with Chapter 5

List of Appendices

Appendices associated with Chapter 5

List of Abbreviations

CM	Corrective Maintenance
PM	Preventive Maintenance
CBM	Condition-Based Maintenance
DTM	Delay-Time Modelling
DES	Discrete-Event simulation
SKU	Stock-Keeping Unit
JIT	Just-in-time (such that the delivery of parts coincides with maintenance interventions)

Chapter 1

Introduction

1.1. Background and methodology

Many research papers have been published during the past decades contributing to the ever-growing interest in using maintenance analysis in the area of Production and Operations Management, and to guide the decision-making process (Wang, 2012a). In particular, the issue of equipment downtime and the need for the reduction of its associated costs including spare parts inventory, has been the subject of intense research.

Imagine a plant with one or more failure modes, which has a maintenance policy of repairing failures as they arise and inspecting every T time units. The objective of the inspection would be to identify and remove any defects before they cause machine failure. Clearly, in this context, the aim would be to minimise the plant operational downtime by reducing the effects of failures and inspection stoppages. Therefore, the main decision variable is the optimal inspection interval, T. If a short interval is used for T, the percentage of time that the plant would potentially be operational will be reduced since there would be frequent inspection activities. Alternatively, if a large T is used, then one would not distinguish between this policy and running the plant under a breakdown maintenance regime. There are five factors that would influence the determination of the optimum inspection interval and thus minimizing the cost of downtime: (i) the timing and the rate of arrival of defects; (ii) the time it takes for defects to cause failures; (iii) the pace at which inspections are undertaken; (iv) the cost and downtime associated with inspections and defect removal (by replacing/repairing parts); and finally (v) the cost and downtime associated with replacing/repairing failures. Thus, using a modelling tool for determining the optimal period for T would be beneficial in guiding the decision-making process.

Many methodologies have been proposed and several concepts have been developed to test and establish the optimum inspection interval, which would minimise the expected downtime and hence the overall cost of production. One of these inspection methodologies is the delay-time modelling (DTM) concept, which describes the failure of industrial equipment in two separate, but linked stages. The first stage defines the time lapse between the new (or as new) up to such a time that a defect arrives - the *time-to-defect*. The second stage describes the time during which the defect continuously deteriorates, up to the point where it finally fails - the *delay-time*. It is this second stage or *delay-time* which opens a window of opportunity for the inspection of plant, identification of defects, and replacement/repair of parts, before downtime occurs.

A number of review papers are present in the literature addressing the issue of optimising the preventive maintenance interval to maximise the operations of industrial plant. These include Thomas (1986), Cho and Parlar (1991), Dekker (1996), Wang (2002), Nicolai and Dekker (2008), Van Horenbeek et al. (2013) and Ding and Kamaruddin (2015). It is noted from these reviews that most, if not all, analytical models are based on assumptions which simplify real life situations and make them less practical. In practical situations, simplifying assumptions is undesirable but permissible to some extent for converging, as far as possible, the application of theory into practice. To relax or eliminate some assumptions of these models, will make them less practical to be implemented. Scarf (1997) is an "appeal to maintenance modellers to work with maintenance engineers and managers on real problems". The author acknowledges, "too much attention is paid to the invention of new models, with little thought, it seems, as to their applicability". It is interesting to note that the same observation still seems valid since evidence suggests that little research is conducted on the optimisation of maintenance in industrial systems (Alrabghi et al., 2017).

For certain industrial situations such as multi-line settings, developing analytical models might prove intractable or mathematically untraceable. The other avenue, which is followed in this PhD thesis, is to replace the analytical approach by simulation modelling.

In developing the simulation models in this thesis, *ProModel* (ProModel, 2016), a process-based discrete-event simulation language, (see for example, Harrell et al., 2011), one of many proprietary simulation packages available in the market, was used (see Appendix 1.1, for the procedure to develop a simple model using *ProModel*). The models, composed of n machines ($n \geq 1$), were developed as continuous production lines. To ensure that the optimal cost is achieved, *SimRunner* (see ProModel, 2010), a simulation optimisation tool, is integrated with the simulation models, which performs sophisticated analysis to determine the optimal value of decision variable(s). The optimisation tool automatically runs multiple combinations of certain variables (if needed) to find the unique combination, which provides the optimal value of the objective function - the long-run average expected cost (or cost-rate). When optimising a particular system, one might use either exact solution methods (analytical) or heuristic methods to find near optimal values for the decision variables. Safety factors, environmental impact, various service levels, system downtime or costs, to mention only a few, are examples which could be used as a focus in an optimisation study. The minimisation of the costs is most common in the optimisation of maintenance-inventory problems (Van Horenbeek et al., 2010), which is also used for the models in this thesis. An optimisation study of a different context is the joint age-usage maintenance strategy by Shafiee et al. (2016) applied to railway tracks, for which the maintenance cost-rate is also minimised.

1.2. Aims and objectives

The main aim of this research is to develop simulation models in order to jointly optimise preventive maintenance and spare part provisioning for multi-line production systems. This is in order to eliminate, or at least to minimise, the occurrence of simultaneous downtime in systems where there are parallel production machines. Simultaneous machine downtime may halt production, which will have a significant adverse effect on profitability or other performance measures. The research uses contexts for which analytical models cannot be developed due to the underlying difficulty in mathematical analysis and intractability. The aim is therefore reflected in the following objectives:

- To undertake a comprehensive literature review of maintenance methodologies and policies for determining the optimum inspection interval for different production line configurations – measured by searching through published and review papers and key literature on maintenance optimisation using analytical and simulation models, and compiling a comprehensive literature;

- To develop discrete-event simulation models for multi-line production systems, and for the interface between maintenance and production management activities – measured by developing working simulation models, the results of which are optimised by the use of an optimisation tool;

- To make recommendations for practitioners to manage effectively the maintenance of their industrial plant – measured by producing a set of results and performance measures, to compare and contrast the key differences and characteristics of each policy;

- To assess the viability of the proposed models in real-world situations – measured by applying the simulation models to: (i) real-life case studies, (ii) reported case-studies in the literature; and finally (iii) idealised contexts documented in journal papers;

- To develop solution tools for gaps identified in the literature and make a significant contribution – measured by demonstrating and presenting findings at conferences and publishing papers in journals on the joint modelling and simultaneous optimisation of multi-line production systems.

1.3. Structure of this thesis

This chapter described the background information for helping to understand matters which will be discussed in more detail in the subsequent chapters. It also described the simulation methodology used in this PhD thesis. Chapter 2 reviews the general literature on three important topics of maintenance systems (including delay-time modelling), inventory control systems, and discrete-event simulation, all of which are pertinent to the topic of this thesis. However, more specific literature review is presented in each of the three principal Chapters, 3 to 5. Chapter 3 discusses in detail, the optimisation of preventive maintenance for a multi-line production system in the context of a case study by developing a number of simulation models (model 1 in Figure 1.1). In Chapter 4, a single-line production plant provides a real context for jointly optimising the planned maintenance of a paper making machine and the associated spare part inventory using the delay-time concept and a number of spare replenishment policies (model 2 in Figure 1.1). The detailed discussion in this chapter gives insights into the characteristics of each policy considered. Chapter 5 describes the development of several simulation models which aim to jointly optimise the planned maintenance and spare part provisioning for an industrial plant comprising a two-machine parallel system (model 3 in Figure 1.1). Chapter 6 summarises the findings and conclusions of the work carried out in this PhD. In addition, it describes the proposed extensions to the simulation models already developed. Detailed appendices related to various simulation models described in Chapters 3 to 5, and the list of references used in this thesis are given in the final two Sections, 7 and 8.

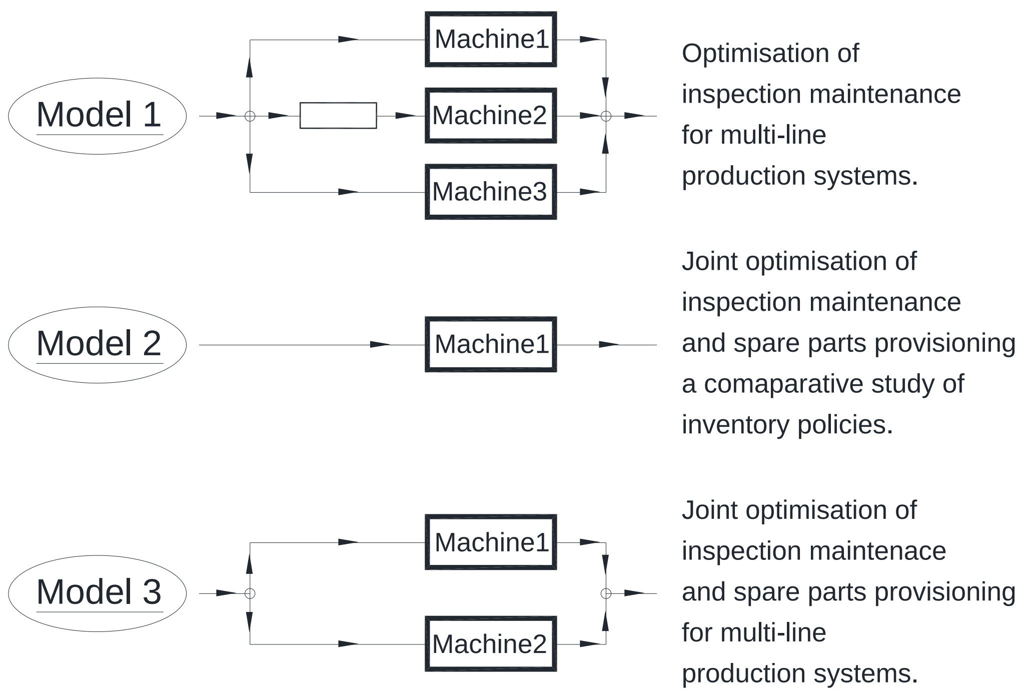


Figure 1.1. The schematic diagram of production configurations used in developing the three sets of simulation models in Chapters 3, 4 and 5:

Model 1 → Chapter 3 → Data from Akbarov et al. (2008);

Model 2 → Chapter 4 → Data from survey and Wang (2012b);

Model 3 → Chapter 5 → Data from survey and Wang (2012b).

Chapter 2

Literature review and Notation

This chapter presents the general literature review on three important topics of maintenance, inventory control, and discrete-event simulation. However, more specific literature is reviewed in each of the three principal Chapters, 3 to 5.

2.1. Maintenance systems

Over the past half a century, many review papers have appeared in the maintenance literature including: McCall (1965); Thomas et al. (1991); Cho and Parlar (1991); Dekker (1996); Dekker and Scarf (1998); Wang (2002); Pierskalla and Voelker (2006); Nicolai and Dekker (2008); Pophaley and Vyas (2010); Das and Sarmah (2010); and Van Herenbeek (2013). For many years, mathematical models have been used for quantifying maintenance functions of industrial plant using various optimisation techniques (Pierskalla and Voelker, 2006). The primary purpose of maintenance optimisation is to find an effective implementation of maintenance policies to minimise maintenance costs or system downtime, or maximise system availability, to mention only a few examples for the focus in the objective function. Maintenance models, if developed appropriately and applied correctly under prescribed conditions, can prove to be very cost-effective in practice (Wang, 2012a).

When a system is to be maintained or restored, the consequence of the maintenance actions undertaken can result in different outcomes. Pham and Wang (1996) give a comprehensive review of the degree of maintenance for repairable parts, as described below:

- Perfect maintenance - after the maintenance actions are carried out, the fixed system is as good as new.

- Imperfect - the system is restored to a state between as good as new and as bad as old.

- Minimal - the system is restored to an as bad as old state with the same failure rate as before.

- Worse - the system's condition is worse than just before the maintenance actions were undertaken.

- Worst - the system breaks down completely after the maintenance actions are carried out.

Some of the terminologies used here may also apply to the replacement of parts. Perfect replacement, as opposed to repair, may occur if the correct parts are installed and the system's state is restored to as good as new. Conversely, replacements can be imperfect if the wrong installation of parts have taken place. The perfect and imperfect analogy may also be applied to inspection. If all faults or defects are identified at an inspection, the inspection is said to be perfect, and anything less is thus imperfect. Van Horenbeek et al. (2013) state that the vast majority of the papers in the literature assume perfect inspection for the restoration of their systems. The authors give a detailed account of the three main maintenance strategies, namely: (i) corrective; (ii) preventive; and (iii) predictive maintenance. Wang (2002) also reviews the most important preventive maintenance policies for both single and multi-unit systems. It should be noted that, a part or a component of a machine (or equipment) is called a unit, which may be repaired or replaced upon failure, or identified as defective and repaired/replaced at inspection. The three terms: unit; part; or component are used interchangeably by different authors.

Under the corrective maintenance, or sometimes referred to as failure-based maintenance, whenever a unit fails, it is immediately repaired or replaced by a new one, provided spares are available. Consequently, if no spare is available, equipment downtime will normally occur and the system will have to await the delivery of new parts while the emergency parts are in transit.

In the manufacturing sector, for example, bearings used extensively in a production plant can fail unexpectedly and catastrophically (Folger et al., 2014a; 2014b) which will need to be repaired or replaced. In other situations, unexpected failure of components may cause disruption to services or accidents (Dinmohammadi et al., 2016). Corrective maintenance is the most reactive of all maintenance strategies. Figure 2.1 illustrates that as corrective maintenance is not planned, the demand requirements (the arrival of defects) is stochastic, yet its size is deterministic (normally single-units) (Wang and Syntetos, 2011).

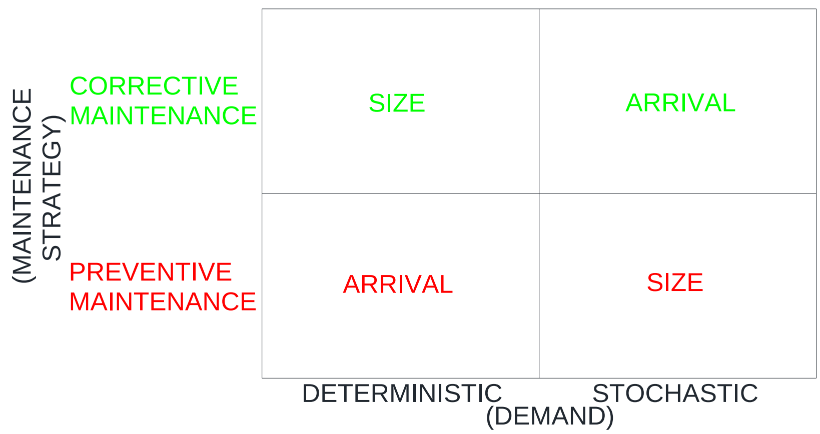


Figure 2.1. Demand behaviour of and spare part requirements under corrective and preventive maintenance strategies.

Systems may also be maintained under a preventive strategy where equipment is inspected at regular intervals, with a view to identifying and replacing all defective (faulty) parts before they cause failures (for example, Wang, 2008). Evidently, there is a strong link between the preventive maintenance inspection interval and the spare part inventory. If the inspection interval is too short, then the 'lumpy' demand effect is created. This is the result of replacing multiple defective but still working parts to reduce the risk of failure at a later stage. Equally, if the inspection interval is too long, then the number of single-unit parts randomly failing is increased, adding to the overall downtime. As shown in Figure 2.1, the timing of preventive maintenance, by its nature, is deterministic as it is planned in advance. However, under this policy, although the demand for spare parts is deterministic, its size is stochastic. There are a number of major policies that fall under this category which may be implemented depending on whether the system under consideration is single or multi-unit. If the asset, machine or equipment lends itself to being maintained based on the age of the unit, then the well-known age-based preventive maintenance, first suggested by Barlow and Hunter (1960), may be used. Under this policy, apart from the units that have failed in service, the rest are replaced whenever they reach their predefined age. Sequential maintenance may also be considered as age-based preventive maintenance since the frequency of maintenance will be increased as the machine and/or units become older. In comparison, under the periodic block-based strategy, failed units are replaced too, but all units are also 'block-replaced' at constant intervals regardless of their history, current condition, and age. Finally, under the failure limit policy, units are replaced when the failure reaches a predetermined rate. Units may or may not be independent or identical. In systems where multiple units exist, parts may be maintained using group or opportunistic strategies. The group maintenance policy combines the same features of the age-based and block-based strategies described for the single-unit systems but as a group replacements for multiple unit of parts. If dependencies exist between the units, one could 'opportunistically' replace other units when a failure occurs. It is important to note that under all policies, failed units are immediately replaced by new ones provided spares are available.

Finally, under the predictive maintenance strategy, better known as condition-based maintenance (CBM), the state of the system is continuously observed and monitored, and where certain or a combination of 'signals' such as product quality, tolerances, excessive vibration, heat, odour, noise etc., reach a prescribed limit, maintenance action is undertaken and units may be replaced (for example, Shafiee et al., 2015). CBM was introduced in order to ensure that PM is only triggered when required, either through scheduled inspections or with smart assistance of sensors, providing data for specialised maintenance software (see, Olde Keizer et al., 2017 for the latest review paper).

Whichever maintenance strategy is used to restore the system under consideration, different costs will be incurred. These costs could include inspection, downtime, labour and spare replacements, for example. A distinction must also be made between failure replacement and preventive replacement, which will have a different cost element for the associated labour and downtime costs.

Alrabghi and Tiwari (2015) observe that the vast majority of journal papers in the maintenance literature, make use of only a limited number of maintenance strategies and policies rather than comparing different alternatives for particular contexts. They conclude that the research in the literature is also limited in terms of comparing and selecting the optimum maintenance policies in multi-component systems.

2.1.1. Delay-time modelling

Delay-time modelling (DTM) was first introduced by Christer (1976) in the context of building maintenance. It was eight years later when Christer and Waller (1984a) applied the same concept to an industrial maintenance problem. Since then, many research papers have appeared in the literature with regard to the concept of this methodology and many more have been published to describe several industrial applications. Since its conception in 1976, a few detailed review papers have been published on delay-time modelling, by Baker and Christer (1994), Christer (1999) and Wang (2012a). Also, a textbook chapter by Wang (2008) comprehensively discusses different aspects of the methodology.

Delay-time modelling, describes the evolution of defects in industrial equipment in two separate but linked stages, as illustrated in Figure 2.2. The first stage is the time lapse from new (or as new) until a defect (or fault) arrives. This is the *time-to-defect* arrival, u. Equivalently, it is the sojourn in the *good* state. The second stage is the time lapse from defect arrival to the point at which this defect causes the equipment to fail. This is the *delay-time*, h. Equivalently, it is the sojourn in the *defective* state. The second stage opens a window of opportunity for inspection, identification of defects, and remedial maintenance intervention (component repair or replacement) before a defect causes failure. By the definition of the *delay-time*, the plant state before failure is binary: *good* or *defective* (Wang, 2012a). Thus, the ‘change point’ from the *good* state to the *defective* state occurs at a random time, failure occurs some random time later, and the time of transition from the *good* to *defective* state is only observable by inspection. By using failure times and counting instances of defects found at inspection, the distributions of the *time-to-defect* and *delay-time* may be estimated, and the relationship between the number of failures and the inspection interval can be established, as discussed by Baker and Wang (1992).

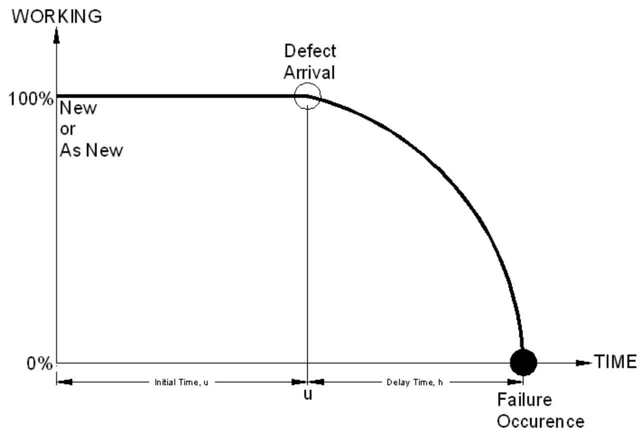


Figure 2.2. The delay-time concept.

In modelling a system, if λ denotes the rate of arrival of defects from all components within an industrial plant, and $F(h)$ denotes the *delay-time* distribution of all failures, then the expected number of failures, $EN_f(T)$, within an inspection interval t (over $0, T$), is given by (Christer and Waller, 1984a):

$$EN_f(T) = \int_0^T \lambda F(h)dh \qquad \text{-- (1)}$$

The above formula is derived under a perfect inspection assumption, and it is the fundamental expression used in all delay-time-based models. This formula explicitly establishes the desired relationship between the expected number of failures and the inspection interval. The probability, $b(T)$, that a fault arising causes a failure is:

$$b(T) = \frac{E[N_f(T)]}{\lambda T} \qquad \text{-- (2)}$$

(Christer, 1999), which increases from 0 to 1 as T increases from 0 to ∞. Accepting the following basic delay-time-modelling assumptions: (i) the plant is running under steady-state conditions; (ii) defects only arise whilst plant is operating and according to a homogeneous

Poisson process (HPP); (iii) all defects are identified at inspection, every T time units, and $d_s <<$ T; and finally (iv) failures are repaired/replaced immediately; then the expected number of failures over an inspection period is $\lambda T.b(T)$, and the expected downtime per unit time $D(T)$ becomes:

$$D(T) = \frac{d_f.E\left[\,N_f\left((i-1)T,T\right)\right]+d_s}{T+\,d_s} = \frac{d_f.\lambda\mathrm{T.b(T)}+d_s}{T+\,d_s} \qquad \text{--} \ (3)$$

(Christer and Wang, 1995). These expressions clearly exhibit the expected characteristics of having large values for small T, and where d_f is the mean downtime per failure and d_s is the mean downtime per inspection. Equation (3) can be minimised in terms of, t if the expected number of failures can be computed and d_f and d_s are known. Equation (3) is established assuming that the defects identified at an inspection will always be removed without costing any extra downtime or cost. This assumption can be relaxed as shown in equation (4) below (Wang, 2008):

$$D(T) = \frac{d_f.E\left[\,N_f\left((i-1)T,T\right)\right]+d_s+d_r.E[\,N_r(iT)]}{T+\,d_s+\,d_r.E[\,N_r(iT)]} \qquad \text{--} \ (4)$$

Clearly, the form of distributions regarding the failure time and the associated parameters must be selected and estimated. Christer and Waller (1984a) state that, when a defect is identified at an inspection, the following questions may be asked: how long ago (HLA) could an inspection or operator have first noticed the fault? And, if the defect is not removed, how much longer (HML) could it be delayed before it causes downtime? The delay time for each fault is then estimated by $h = HLA + HML$. In this way, by observing sufficient defects, a prior distribution for $F(h)$ may be obtained. And, if the inspection identifying a defect is made at time t, then $u = t - HLA$.

Wang (2011b) extended the delay-time concept from a two-stage (Figures 2.3 (a) and 2.3(b)) into a three-stage (Figure 2.3(c)) failure process, where the *delay-time* stage is divided into two stages, corresponding to a *minor* and a *severe* defective stage. This means that, at any one time, the plant item can be in one of the four states of: good; minor defective; severe defective; and failed. The three-stage delay-time concept seems to better reflect the true states of a plant item in reality, however, the extended model is more complicated to develop and will require more information to enable the parameter estimation procedure in practice (for example, Baker and Wang, 1992).

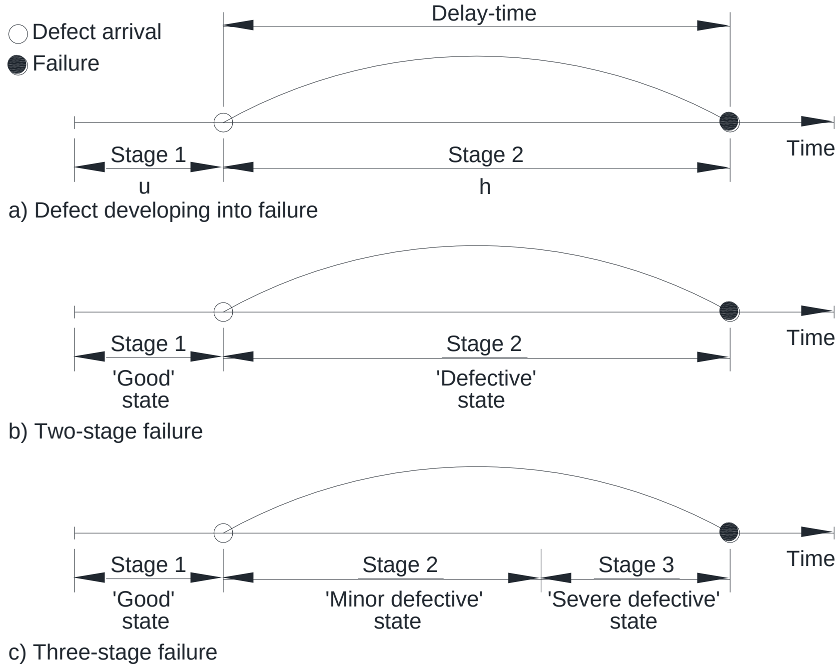


Figure 2.3. A depiction of development of the three-stage delay-time model.

Delay-time modelling captures the relationship between failures of items in service, inspection at constant PM epochs, and PM replacement of defective items under the assumption that all defective items are always identified and replaced (provided spares are available) at inspections. The fundamental difference between DTM and other inspection strategies is that under the former, only defective items (if any) are replaced at inspection intervals. In comparison, under the age-based policy, items are only replaced according to their age, or irrespective of their age and condition under the block-based replacement policy. Apart from inspection, and repair/replacement at PM interventions and failure events, there may be other activities, such as, removing metal burrs, lubricating components, and changing engine oil, for example. Depending on circumstances and if necessary, these events may be modelled using a variable rate for entity arrivals.

DTM is a modelling methodology that can be used to determine an inspection-based or block-based PM policy, where all items are inspected at constant inspection intervals (a decision variable to be determined) and defective items are 'block' replaced. Under an inspection or block-based PM policy, inspection identifies all defective items that will be replaced, whereas in a normal block-based replacement policy all items are replaced regardless of their age and conditions. Generally, failures of the items in service generate intermittent single-unit demand. In addition, the inspection process generates multiple (lumpy) demand as a result of identifying and replacing all defective items at PM intervals. Furthermore, the timing of demand for spare parts is stochastic at failures, but deterministic at the times of preventive replacements.

There are two distinct types of DTM systems: (i) single-component or component tracking; and (ii) multi-component or complex system. In a single-component system, as shown in Figure 2.4, there may be a single dominant failure mode, and the system may be renewed upon failure (Wang, 2008). Under the inspection-based policy and the instance shown in Figure 2.4, inspection at the first and third epochs will identify and remove the defects and the system is thus renewed. However, before the 2nd and 4th inspection epochs, component failures occur and the system is renewed again upon replacements. Examples of single-component systems are reported in: Baker and Wang (1992, 1993); Wang and Christer (1997); and Yang et al. (2016).

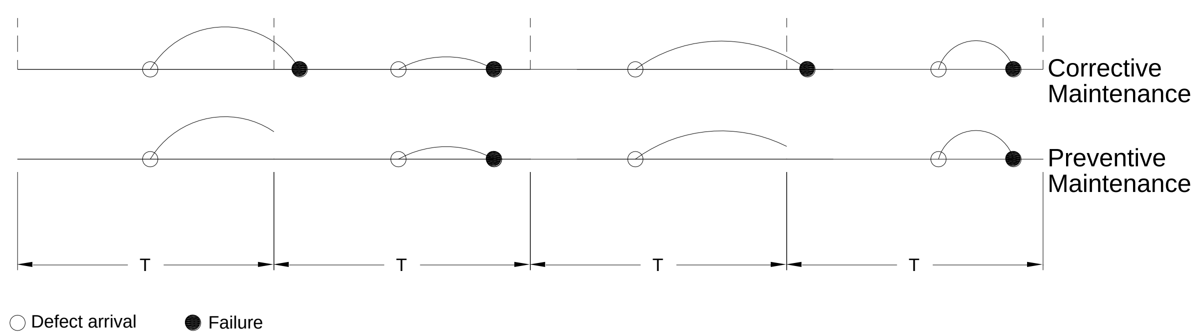


Figure 2.4. Defect arrivals and failure occurrences in a single-component system.

In contrast, a complex system is one in which many failure modes could arise, and the correction of one failure or the replacement of one defect will have nominal impact upon the overall plant failure characteristics or the steady state of the system. Figure 2.5(i) depicts an example of a complex system where six defects (1, 2, etc.) arrive over time. With the assumption of perfect inspection, if regular inspection takes place, for example at points A, B and C, then some defects will be identified and removed before failures occur, as shown in Figure 2.5(ii). Considering Figure 2.5(ii) further, at inspection point A, two defects have already arrived and are currently in their *delay-times*. Thus, both defects 1 and 2 will be identified and removed at inspection point A, either by replacing or repairing before failures occur. Defect 3 arrives in the middle of the period between scheduled inspections A and B and will be identified and removed at inspection point B. Before inspection C, one failure occurs as a result of defect 5. However the inspection at point C, identifies and removes both defects 4 & 6 before they cause downtime. Therefore, in this instance, with a suitable length for the inspection interval, 5 out of 6 defects (83%) will be identified and removed. The system may thus be renewed at inspection points A, B and C if the rate of arrival of defects is constant and the inspections are perfect. Most delay-time-based models reported in the literature are models of complex systems, and examples include: Christer and Waller (1984a, 1984b); Christer et al. (1995); Akbarov et al. (2008); Jones et al. (2010); Lu and Wang (2011); and Pietruczuk and Werbinska-Wojciechowska (2017).

There are many delay-time-based case study applications reported in the literature. Some examples include: Christer and Waller (1984a); Christer (1987); Baker and Wang (1992, 1993); Christer et al. (1995); Pillay et al. (2001a, 2001b), Arthur (2005); Akaborov et al. (2008); Jones et al. (2009, 2010); Liu et al. (2015); and Emovon et al. (2016).

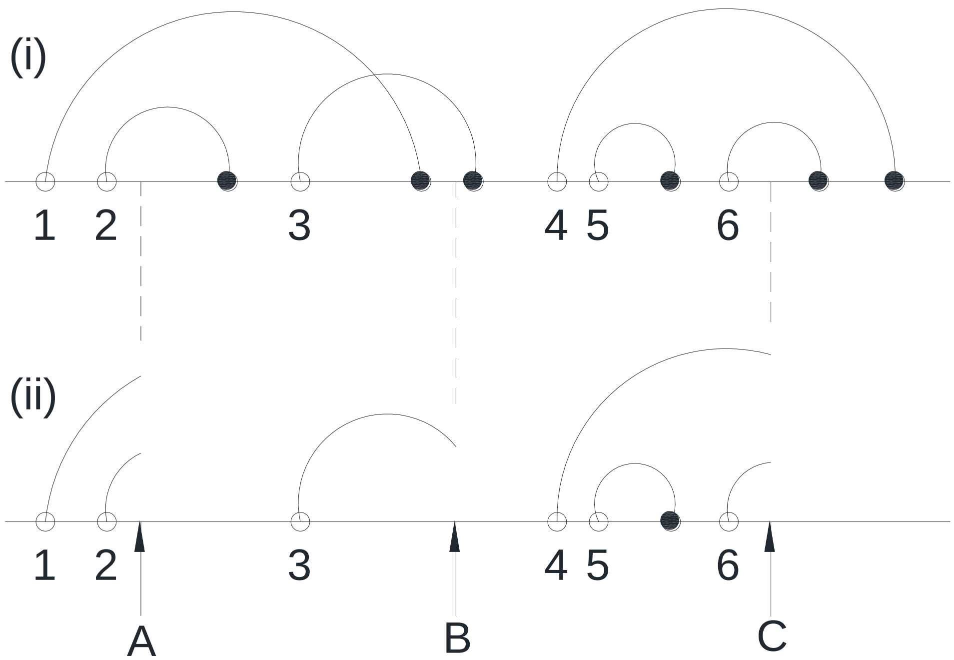


Figure 2.5. Defect arrivals and failure occurrences
in a complex system of multiple components.

2.2. Inventory control systems

Maintenance costs are clearly dependent on the availability of spare parts. However, many models assume there is an infinite inventory of spare parts at all times, which makes their use unrealistic in practice. The inventory for spare parts is normally controlled by a particular replenishment policy. The overall objective is always to find the optimal policy. Keeping too many spares will increase the holding cost, which will have financial implications on the company's cash flow and/or borrowing, or will increase the risk of spare parts' obsolescence. Conversely, keeping too few parts might result in the plant's unavailability at critical times. The cost associated with the unavailability of spare parts is twofold: (i) the cost of equipment downtime while awaiting spare delivery; and (ii) the cost of expediting the delivery of parts in emergencies.

For spare parts classification, the usual approach is to categorise according to a part's service criticality. "Alternatively, an ABC classification is used, which lists all stock-keeping units

(SKU) in descending order, by total volume, or total value of sales, with the A items being assumed to be the most critical and requiring the highest service levels" (Boylan and Syntetos, 2010). However, the process may be guided by criticality and cost considerations, as well as the ABC classification. Molenaers et al. (2012), classified spares based on attributes like demand pattern, unit price and inventory costs. Later, Hu et al. (2017) classified spare parts based on multiple criteria of criticality, price, demand, lead time, and obsolescence, since they noted that "single objective of price is generally misleading". The authors state that their approach does not offer optimisation, but they intend to add this extension as an enhancement in the future.

There are two distinct approaches for the replenishment and management of spare parts (see, for example, Muller, 2011). Stock may be reviewed: (i) periodically; or (ii) continuously (See, Kennedy et al., 2002, and Santos and Bispo, 2016, for example). Under the periodic review policy, there are at least three methods by which parts may be replenished: (i) periodically (R), at the beginning of each cycle, raising the inventory position to a pre-defined level S, based on the forecasted demand for the next period for example - the (R, S) policy; (ii) periodically raising the inventory position to level S if the stock level has reached or dropped below a certain level s - the (R, s, S) policy; and (iii) periodically raising the inventory position by ordering a fixed quantity Q of stock if the inventory position has reached or dropped below s – the (R, s, Q) policy (see, for example, Silver et al., 2016).

In comparison, under the continuous review policy, every time the stock level is depleted, the inventory levels are checked. Then, either a sufficient quantity, up-to-level S is ordered if the inventory position reaches or drops below s – the (s, S) policy, or a fixed quantity of parts is ordered when the inventory position reaches or drops below s – the (s, Q) policy. When there is a per unit demand, both the (s, S) and (s, Q) policies give the same result when $Q = S - s$. A special case of the (s, Q) replenishment policy is famously called a two-bin policy where a replenishment order, sufficient to fill up the bin, is immediately placed when the first bin is empty. The second bin is then used during the replenishment lead-time. This policy is mainly used for low cost and high demand spare parts.

There are three major costs associated with all stock ordering policies, namely: (i) ordering; (ii) holding; and (iii) shortage costs. Firstly, the fixed ordering cost is either for the unit purchase cost under normal circumstances or for the replenishment of parts in emergencies. Secondly, holding inventory is expensive since it will have capital and space cost implications. And finally, shortage costs will be incurred if the number of spares in stores is insufficient to meet the demand. Different policies aim to balance these costs in order to produce an overall optimum cost. Stock replenishment quantities depend on whether the system under consideration is single or multi-unit. However, when failure frequencies are high or spare replenishment lead-time is long, it might prove wise to keep more than one part in stock, even for a single-unit system. On the other hand, keeping multiple units of spare parts increases the cost of inventory and the risk of obsolescence, which is a major issue and has cost implications too.

2.3. Discrete-event simulation

Simulation has been used for many years to understand and experiment with systems under study, especially in the production and manufacturing industry where the use of discrete-event simulation (DES) has been very effective. The use of simulation has grown dramatically since modern manufacturing systems have become more complex as a result of dependencies and interactions between system components. Gupta and Lawsirirat (2006) highlight the fact that the term *component* has a different meaning in different contexts. Since it is not possible to model every part in a complex system, it is practical to consider only the components that have significant impact on the performance of the system. Compared to the traditional and mainstream discrete-event simulation, agent-based modelling and simulation (ABMS) is a relatively new technique for simulation (Macal, 2016), for which the number and breadth of applications has had a huge expansion during the past 10 years (Cheng et al., 2016). A very important step forward in the world of simulation is the obvious and essential procedures for verification and validation, which can only lead to credibility of simulation models and the results achieved from them (Rabe and Dross, 2015). The gap between research in optimization via simulation and the development of algorithms that can be applied to real-life problems has narrowed substantially in the last ten years. One factor influencing this issue is the ever-growing use of parallel simulation, which is becoming easy to do, and any simulation study that requires multiple replications or multiple scenarios will benefit from this advancement (Nelson, 2016).

During the past three decades, simulation software packages or tools have gradually replaced simulation languages. Dias et al. (2016) ranked the top worldwide most popular and used simulation software, based on intensity of usage or level of presence in different sources - “popularity”. The survey findings clearly identify *Arena* software package as the most ‘popular’ and widely used discrete-event simulation tool, followed by *ProModel*, which is the tool used for the development of simulation models in this thesis. The next three most used tools, namely *FlexSim*, *Simul8*, and *WITNESS*, making up the first cluster, have more or less similar rankings in the comparison table. It is important to note that the contexts of these simulation tools are constantly changing, whether in industry or academia, and the ranking may indeed change in the next survey published in the literature.

Simulation delivers an advantage over analytical approaches since many maintenance policies are not analytically traceable (Nicolai & Dekker, 2008). Analytical and mathematical approaches are limited in solving such complex maintenance problems.

To optimise their maintenance problem, Allaoui and Artiba (2004) preferred the use of simulation over the analytical approach and stated that the immediate availability of maintenance resources is a major assumption in many studies. Similar conclusions were reached by Rezg et al. (2005) who used both analytical and simulation approaches (using *ProModel*) to develop a model of a system with storage buffers, which helped in reducing the effect of lower production rate while maintenance interventions were taking place. Gharbi and Kenné (2005) developed a discrete-event simulation (DES) model in which the degradation of the machine was modelled as a continuous process to reflect the reality that as time goes by machines are automatically aged (see also, Roy et al., 2016).

Many authors use discrete-event simulation packages for their maintenance studies. However, others including Cavory et al. (2001) used a general-purpose programming language to develop their models. The authors found that the development of the analytical model for the given context was difficult and thus used simulation for the resolution of their problem. Ilgin and Tunali (2007) developed a simulation optimization model, which they believed, gives the ability to describe multivariate non-linear relations that is difficult to express in an analytical form. They concluded that a 53% reduction in total annual maintenance cost and 6% improvement in average monthly production were achieved.

Alrabghi and Tiwari (2015) surveyed the literature and reported on the state-of-the-art simulation-based maintenance optimisation, which resulted in 59 journal papers since the year 2000. Discrete-event simulation is the most reported technique for modelling maintenance systems. In comparison with general-purpose languages, specialised simulation software provide several advantages such as, rapid modelling, animation, automatically collected performance measures, and statistical analysis tools (Banks, 2010).

Minimising cost is reported as an objective in the majority of simulation studies in the literature. Kuntz et al. (2001) used an inspection-based model and included machine downtime in the cost function. Instead of minimising maintenance cost, Roux et al. (2013) argue that maximising machines' availability should be the objective function since production costs tend to be higher. However, other authors believe that the significance of maintenance costs cannot be underestimated. The use of PM and CBM were investigated in a study by Xiang et al. (2012). They observe that as sensors get less expensive, the use of CBM strategy will become more popular, which has potential cost-savings.

The main assumptions common in the majority of simulation-based studies in the literature include: (i) perfect maintenance in maintaining identical and independent units; (ii) failures are detected immediately; (iii) costs of maintenance actions are constant, but the cost of CM is always higher than PM; (iv) duration of maintenance activities are constant or take zero time; and finally, maintenance resources are always immediately available when required. Considering the points listed above highlight the limitations of studies in the literature.

Since it is difficult to obtain accurate cost data for conducting maintenance and inspection activities in simulation studies, sensitivity analysis is used in several publications in the literature, which test the robustness of a suggested model by varying inputs and investigating if the results are in line with the expected outcome (Boulet et al., 2009). Surprisingly, many researchers do not disclose the simulation technique or the specific software used in their research, which will limit the replicability of the experiments by independent researchers.

Over the years, discrete-event process simulation has steadily grown in power due to the advancement of hardware, ease of application due to software sophistications, availability of expertise due to the growth of simulation as a business-improvement tool, and breadth of applications to business challenges, especially in manufacturing. There are different reasons and motivations for the use of simulation to initiate a manufacturing-context simulation project. Khalili and Zahedi (2013) used simulation to prepare a mattress production line for anticipated demand increases over a five-year planning horizon (Williams, 2014). Natarajan (2016) reports on a simulation-based case study of a production line at an automotive ancillary manufacturing plant. In the first phase, the existing system is simulated to identify the critical operations in the system. Then, the existing system is modified based on the suggestions of the finding of the initial phase of this study. Finally, the revised model is simulated, which produces improvements in the production volume for the engine thrust bearing line. Rozen and Byrne (2016) examine preventative maintenance segregation with the aim of determining the optimum PM frequency that results in reduced cycle time. The resulting solutions from many years of maintenance modelling have proven to be very effective. However, to improve on those simple solutions, complex and time-consuming simulation modelling is required, and reliable input data is even more important, which is driven by Big Data and the Internet of Things technologies (Volovoi, 2016b). In order to be successful in this path, the inner workings of DES, so far hidden from decision makers, have to be highlighted to the users (Volovoi, 2016a) (also see, Alrabghi and Tiwari, 2016).

In many studies, simulation is used as a solution tool. Sarker and Haque (2000) used simulation since the development of mathematical models was “extremely difficult”. In their model, they considered maintenance resources, such as spare parts, and concluded that the results of their jointly optimized policy was superior over the combination of separately or sequentially optimized policies. Tateyama et al. (2010) is one of few research articles that considers maintenance resources, such as technicians, as decision variables.

Maintenance plays a key role to sustain the operations of manufacturing systems under high production throughput, reliability and safety requirements. Opportunistic maintenance gives staff the chance to replace or repair those items that are found to be defective or need replacement during the maintenance of another machine or component. However, components are usually assumed to be independent. Lung et al. (2016) develop an opportunistic maintenance policy, considering both economic and structural dependence between different components. In Babishin and Taghipour (2016), a system consisting of components subject to soft and hard failures is modelled using simulation together with an optimisation tool. Hard failures are self-announcing and are fixed immediately (similar to failures under DTM) and provide an opportunity for inspection (opportunistic inspection). Soft failures (which may be thought of as defects under DTM) are only identified and revealed at periodic inspections, which are then corrected (repaired or replaced).

A specific application of simulation optimisation is in the area of opportunistic maintenance of wind farms. Corrective and time-based preventive policies are widely employed for the maintenance optimisation of wind turbines. In Zhang et al. (2017), an opportunistic maintenance approach is proposed, considering imperfect maintenance based on reliability. The simulation results highlight the economic advantage of using an opportunistic maintenance strategy. The cost of energy generated from offshore wind is dependent on maintenance cost to a great extent, which in turn depends on the strategy for performing maintenance. In Sarker and Faiz (2016) model, opportunistic maintenance is performed on other components in the system while a failed component is replaced. The model in Irawan et al. (2017) aims to optimise the maintenance schedule, the routes for the crew transfer vessels to service the turbines, and the number of technicians required for each vessel. The proposed approach was tested using data from a case study reported in the literature as well as for a context generated randomly. The use of opportunistic maintenance strategy is becoming very popular in the literature for reducing the cost of wind power generation. Policies are developed and decisions are made based on static reliability or age thresholds. In Erguido et al. 2017, the dynamic reliability thresholds are allowed to vary according to the weather conditions. The authors claim that the results improve the ones proposed by the static reliability thresholds, both in terms of wind farm production and life cycle cost. Abdollahzadeh et al. (2016) argue that almost all optimisation models for the opportunistic maintenance of wind turbines focus on a single objective in the optimisation process. In their paper, a discrete-event simulation model is developed to simultaneously maximise the expected rate of energy and minimise the total expected maintenance costs. Particle swarm algorithm is used to search for cost effective solutions in the multi-objective optimisation process.

Boschian et al. (2009) considered a two-machine parallel system and discuss the complexity of analytical modelling. Their study uses simulation but does not consider joint optimisation. Lynch et al. (2013) considered the joint optimisation of spares inventory and preventive maintenance and concluded that developing such aspects separately will lead to sub-optimal performance due to trade-offs between different cost components. Nguyen et al. (2016) deals with the problem of maintenance optimization of a two-component system, in series, which are dependent in such a way that the lifetime parameters of component subject to shocks (causing sudden failures) depend on the degradation level of the gradually deteriorating component. They show that both stochastic and economic dependence have a significant influence on the performance of various policies, and that joint optimisation reduces the cost of components up to 16% when compared to individual sequential optimization. The model in Zhang et al. (2016) aims to optimise a maintenance policy considering minimal repairs and imperfect maintenance for a two component load-sharing system, in which the failure of one component may increase the failure rate of the other. The authors claim that the maintenance policy can be generalised to include multi-component systems.

Gopalakrishnan et al. (2016) proposes an approach to determine the machine priorities for dynamic scheduling of maintenance work orders by identifying buffer utilization, which improves throughput of the system in comparison to a first-come-first-served approach for executing maintenance work orders. Machine breakdowns and improper maintenance management cause production systems to function inefficiently. The approach is applied in an industrial case study, which suggest that systems view for maintenance prioritization can be a powerful decision support tool for maintenance planning. Alrabghi et al. (2017) suggest that not enough research is conducted on the optimisation of maintenance of industrial systems. The authors present two case studies: a tyre re-treading factory and a petro-chemical plant. They investigate optimising various maintenance strategies simultaneously including Corrective Maintenance, Preventive Maintenance, Opportunistic Maintenance and Condition-Based Maintenance. The study states that over-looking the optimisation of maintenance on the strategic level may lead to sub-optimal solutions. One of the general findings suggest that the high cost and time of computation associated with simulation-based maintenance optimisation of complex systems is still an issue to be resolved. A possible area of future research would be to investigate approaches for reducing computational expenses. Linneusson and Aslam (2016) argue that the inherent complexity of maintenance and its dependency on production is a complex equation. Simulation facilitates the testing of different strategies and scenarios, which in this case has proved the value of condition-based maintenance over reactive unplanned maintenance.

Unlike Roux et al. (2008), who considered three different maintenance policies, many researchers evaluate only one policy, which seems to be a limitation of the research in the literature (Alrabghi and Tiwari, 2015). Zahedi-Hosseini et al. (2017) analysed the characteristics of ten policies in their joint optimisation study of preventive maintenance and spare part inventory for a production plant. The authors concluded that in the context of a paper making plant, a periodic policy with ordering that is twice as frequent as inspection is cost optimal. This study (Chapter 4 in this thesis) directly addresses the limitation of some of the papers in the literature for the lack of comparison of different policies.

As observed by Van Horenbeek et al. (2013), the research on optimising a system composed of several machines is limited and most do not consider the production configuration. The authors concluded that single-machine systems are therefore oversimplified that do not reflect the interactions in real manufacturing systems. Chapters 3 and 5 in this thesis address directly this limitation, which has been identified in the literature.

2.4. Gaps in literature

As a result of reviewing the relevant material on maintenance systems, delay-time modelling, inventory control systems, and joint optimisation of maintenance and inventory, the following gaps have been identified in the literature:

- The research on optimising a system composed of several machines is limited;

- Analytical models are only applicable for the maintenance activities of single-line production systems, which may prove to be oversimplified. These models do not reflect the dependencies and interactions inherent in complex manufacturing systems under different configurations for practical industrial situations. Simulation may constitute an alternative solution tool in situations such as multi-line production systems where mathematical models are limited due to the difficulty of mathematical analysis;

- The vast majority of journal papers in the literature evaluate only a single maintenance policy rather than comparing different alternatives for particular contexts. This is also a limitation in terms of comparing and selecting the optimum maintenance policy in multi-component systems;

- The joint modelling and simultaneous optimisation of preventive maintenance and spare parts inventory for multi-line production systems.

2.5. Notation

2.5.1. Notation associated with Chapter 3

C_d	Cost-rate of production downtime
C_s	Cost-rate of inspection maintenance (including repair)
$C(T)$	The long-run cost per unit time (cost-rate)

$C_{con}(T)$	The cost-rate for consecutive inspection of parallel lines
$C_{mod}(T)$	The cost-rate for a modified system of inspection of parallel lines
$C_{sim}(T)$	The cost-rate for simultaneous inspection of parallel lines
$D(T)$	The long-run downtime per unit time (downtime-rate)
d_f	Failure stoppage duration
d_s	Inspection duration (where $d_s << T$)
d_v	The downtime for a concurrent occurrence of two failure stoppages on separate parallel lines
$d_{v\prime}$	The downtime for a concurrent occurrence of a failure stoppage and an inspection on separate parallel lines
$EN_f(T)$	The expected number of failures over the interval (0, *T*)
F	Failure repair process
I	Inspection process
λ	Defect arrival intensity, per unit time
T	Inspection interval
U	Initial time from new (or as new) until a defect that could be identified by inspection arises (the *time-to-defect*), a random variable
$F_U(u)$	Cumulative distribution function (cdf) of *U*
u	Particular realisation of *U*
H	Time between a defect arising and the subsequent failure if left unattended (the *delay-time*), a random variable
$F_H(h)$	Cumulative distribution function (cdf) of H independent of *U*
h	Particular realisation of *H*

2.5.2. Notation associated with Chapter 4

C_d	Cost-rate of machine downtime
C_h	Cost-rate of inventory holding
C_i	Cost of inspection
C_m	Cost-rate of one maintenance technician

C_o Order cost including normal delivery

C_f Cost of failure replacement (per item)

C_r Cost of preventive replacement (per item)

C_{sh} Order cost for stock-out including emergency delivery

C_u Cost of one item

$C(T)$ Long-run expected cost per unit time, or cost-rate

d_f Downtime due to failure (per item)

d_r Downtime due to preventive replacement (per item)

$D(T)$ Expected downtime per unit time

L Normal delivery lead-time

L_{sh} Emergency delivery lead-time

Q Order quantity; a decision variable - dependent on the replenishment policy

R Order review period; a decision variable - dependent on the replenishment policy

s Re-order level; a decision variable - dependent on the replenishment policy

S Order-up-to-level; a decision variable - dependent on the replenishment policy

T Inspection interval; a decision variable; $T = k\ R$, for any $k > 0$.

U Time-to-defect arrival; initial time from new (or as new) until a defect that could be identified by inspection arises; a random variable

$F_U(u)$ Cumulative distribution function (cdf) of U

u Particular realisation of U

H *Delay-time*; time between a defect arising and the subsequent failure if left unattended; a random variable

$F_H(h)$ Cumulative distribution function (cdf) of H, independent of U

h Particular realisation of H

λ Defect arrival rate (intensity)

2.5.3. Notation associated with Chapter 5

$C_{d(ind)}$ Cost-rate of individual machine downtime

$C_{d(sim)}$	Cost-rate of simultaneous machine downtime
C_f	Cost of failure replacement (per item)
C_r	Cost of preventive replacement (per item)
C_h	Cost-rate of inventory holding
C_i	Cost of inspection
C_m	Cost-rate of one maintenance technician
C_o	Order cost including normal delivery
C_{sh}	Order cost for stock-out including emergency delivery
C_u	Cost of one item
$C(T)$	Long-run expected cost per unit time, or cost-rate
d_f	Downtime due to failure (per item)
d_r	Downtime due to preventive replacement (per item)
$D(T)$	Expected downtime per unit time
L_o	Normal delivery lead-time
L_{sh}	Emergency delivery lead-time
R	Order review period; a decision variable
S	Order-up-to-level; a decision variable
T	Inspection interval; a decision variable; $T = kR$, for $k > 0$.
U	*Time-to-defect* arrival; initial time from new (or as new) until a defect that could be identified by inspection arises; a random variable
$F_U(u)$	Cumulative distribution function (cdf) of U
u	Particular realisation of U
H	*Delay-time*; time between a defect arising and the subsequent failure if left unattended; a random variable
$F_H(h)$	Cumulative distribution function (cdf) of H, independent of U
h	Particular realisation of H
λ	Defect arrival rate (intensity)

Chapter 3

Optimisation of preventive maintenance for multi-line production systems

3.1. Summary

This chapter develops simulation models to determine the cost-optimum inspection policy for a number of multi-line production systems. Analytical models are complex and intractable for determining the optimal inspection interval in such a setting. An approach is developed using a well-known discrete-event simulation environment. The machines in the multi-line system are subject to a two stage failure process based on the delay-time concept. The policy which is studied indicates that consecutive inspection of lines with priority for failure repair is cost-optimal and suggests a substantial cost reduction of 61% compared to a 'run-to-failure' policy. The implication of this pragmatic policy is that maintainers need to be responsive to operational requirements. These ideas are developed in the context of a case study of a plant with three lines, one of which is on cold-standby.

3.2. Introduction

Many studies highlight the importance of maintenance within the production context. In an early paper, Geraerds (1978) concluded that The Netherlands spent 14% of GDP on maintenance activities, 34% of which was associated with expenditure for the industrial plant. Komonen (2002) reported that in Finland maintenance cost is typically 5.5% of the turnover of a company, but could be as much as 25%. Generally, organisations have become increasingly aware that proper maintenance of their production facilities is a vital part of their everyday business (Cholasuke et al., 2004). Analysing the results from a case study, Alsyouf (2006) showed that companies have the potential to improve their return on investment (ROI) by 9% through the planned and timely use of maintenance. Alsyouf (2009) observes that the maintenance function contributes a potential improvement of 14% in ROI, and in a later work, Alsyouf et al. (2016) show that good maintenance planning can reduce maintenance costs significantly. There is, therefore, a great deal of financial interest to optimise maintenance operations and thus reduce the effect of plant downtime by identifying and removing defects (faults) before they cause plant to fail.

However, studies show that little research is directed towards the realistic scenario of optimising maintenance for a system composed of several machines and the focus is instead on optimising a single machine without considering the production configuration (van Horenbeek et al., 2013). This view is supported by considering the many review papers that address the optimisation of preventive maintenance (e.g. Cho and Parlar, 1991; Dekker, 1996; Wang, 2002; Nicolai and Dekker, 2008; Das and Sarmah, 2010; and Ding and Kamaruddin, 2015). Here all analytical models relate to single-line production facilities. Furthermore, most, if not all these models, are based on assumptions which simplify real life situations and make them less practical. In practical situations, simplifying assumptions are undesirable but necessary to some extent for the convergence, as far as possible, of theory and practice. To relax or eliminate some assumptions from these models, one would require more complex and detailed modelling, which may not be amenable to analytical solution. An alternative avenue, using simulation to replace intractable, analytical models, is followed in this study for the optimisation of preventive maintenance of multi-line parallel production facilities in particular.

Simulation has the potential to address the increasingly complex and dynamic nature of maintenance optimisation problems. Alrabghi and Tiwari (2015) surveyed the literature and reported the state-of-the-art in simulation-based optimisation of preventive maintenance research, with 59 research articles since the year 2000. Discrete event simulation seems to have been the most reported technique for modelling maintenance systems. Specialised simulation software provides several advantages over general-purpose programs such as rapid modelling, animation, automatically collected performance measures and statistical analysis (Banks et al., 2013). Boschian et al. (2009) discuss the complexity of analytical modelling in optimising the maintenance strategies for a ‘small size’ production system (two machines working in parallel) that is prone to random failures and undergoes preventive and/or corrective maintenance. 'To get round this complexity' they also chose an approach based on simulation.

For single-line systems, a number of models that aim to optimise an inspection interval have been proposed and tested. Early models include those due to, for example, Barlow and Proschan (1965), Luss (1983), Kaio and Osaki (1989), Jardine and Hassounah (1990) and Abdel-Hameed (1995). Other authors have integrated production quality into the inspection problem (e.g. Lu et al., 2016) and considered preventive maintenance planning in job shop scheduling (e.g. Thörnblad et al., 2015). In this study, the delay-time concept is used, which was first introduced by Christer (1976), applied to the maintenance of industrial equipment by Christer and Waller (1984a), and later developed further by many (e.g. Baker and Christer, 1994; Okumura et al., 1996; Jones et al., 2009; Van Oosterom et al., 2014; Flage, 2014; and Chellappachetty and Raju, 2015). Related case studies include those due to Jones et al. (2010) and Emovon et al. (2016). The delay-time concept has the advantage that it explicitly models the relationship between plant failures and the inspection interval. The latest review of the advances in delay-time-based maintenance modelling including applications is Wang (2012a).

Based on the literature review undertaken, two contributions are made in this chapter: (i) the delay-time concept is used for the first time to describe the inspection of multi-line production facilities; and (ii) an attempt is made to bring this theory closer to practice by minimising the production downtime for a multi-line parallel facility as a whole, using simulation to do so. The importance of this work lies in its implications for the design of preventive maintenance for multi-line production systems and the contribution that good maintenance can make to economic performance.

This chapter is structured as follows. Section 3.3 describes the delay-time concept, the modelling methodology, discusses practical circumstances in which such models are intractable, and introduces models of multi-line production systems with focus upon how downtime affects production. In Section 3.4, a case study is described and the focus is solely on the development of the simulation models and analysing the results of several alternative policy scenarios, beginning with a single-line packing facility, and developing to several model extensions for multi-line production systems. In the final section, detailed conclusions are drawn, their implications are discussed, and the direction of future research is suggested.

3.3. Modelling methodology

3.3.1. Notation

See Section 2.5.1., for a list of Notation associated with this chapter.

3.3.2. The delay-time model development

As illustrated in Figure 3.1, delay-time modelling describes the evolution of defects in industrial equipment in two separate, but linked stages. The first stage is the time lapse from new (or as new) until a defect (fault) arrives. This is the *time-to-defect arrival*, *U*. Equivalently, it is the sojourn in the *good* state. The second stage is the time lapse from defect arrival to the point at which this defect causes the equipment to fail. This is the *delay-time, H*. Equivalently, it is the sojourn in the *defective* state. This second stage is the window of opportunity for inspection, identification of defects, and remedial maintenance intervention (repair or component replacement) before a defect causes failure of operational function. Thus, the 'change point' from the good state to the defective state occurs at a random time, failure occurs some random time later, and the time of transition from the good to defective state is unobservable. Nonetheless, using failure times and counting instances of defects found at inspection, the distributions of time of defect arrival and *delay-time* may be estimated (Baker and Wang, 1992).

Consider now the complex or multi-component plant maintenance modelling scenario in this study, shown in Figure 3.2, with the associated notations and assumptions shown in Section 3.3.1. Here, multiple concurrent defects are possible. The 'first to fail' determines the failure time, and thus a series system is assumed. Failures are repaired immediately but not instantaneously, so that a downtime of d_f time units is incurred at a cost of C_s per unit time.

Inspections are carried out every T time units, requiring d_s time units and costing C_s per unit time, where $d_s << T$. All defects identified at inspection are repaired during the inspection process, I. During the failure stoppage process, the system is returned to the operational state, but any defects present are not removed. During inspection and failure stoppage (repair) processes, plant components are assumed to be in a state of suspension, so that the system is then not ageing, and thus defects and failures can only arise whilst plant is operating, and defects are not 'growing'. It is assumed that the plant has operated sufficiently long to be considered working under a steady state condition.

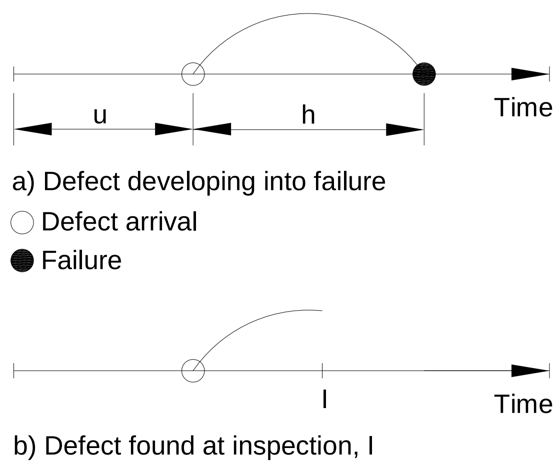


Figure 3.1. The delay-time concept, depicting the arrival of a defect and its evolution into: (a) failure; or (b) not, as inspection intervenes.

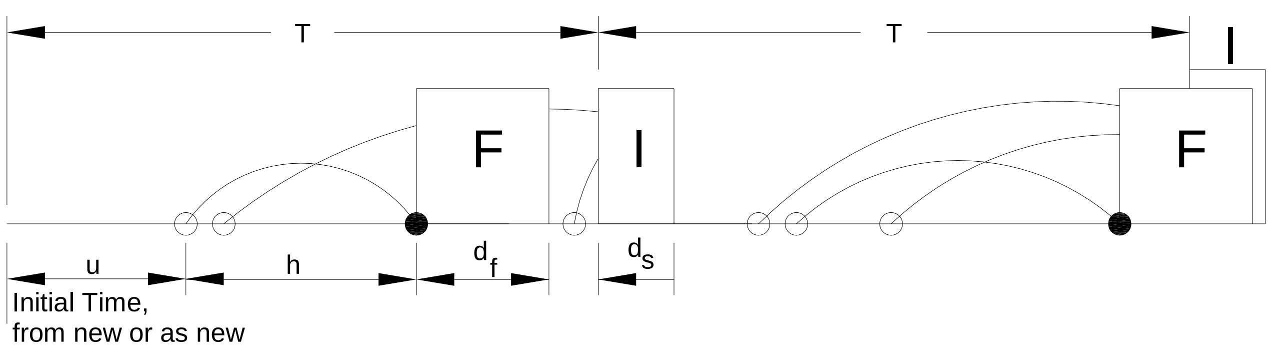


Figure 3.2. Defect arrivals, failures, failure repair F, and inspection I in this complex system of multiple components.

These assumptions represent a relatively simple inspection problem in the class reported by Christer (1999), whom under these assumptions, gives the expected number of failure breakdowns over the interval $(0, T)$, $EN_f(T)$, and the expected downtime per unit time, $D(T)$. Provided that $F_U(u)$ and $F_H(h)$ can be estimated, either through the consideration of data or subjective, expert opinion or both, $D(T)$ and $C(T)$ equivalently can be calculated and then the T that minimises $D(T)$ or $C(T)$ can be determined. It is this optimisation step that links the inspection frequency to the defect arrival and failure rates, and the cost and downtime parameters. However, these models are only applicable to the maintenance activities for single-line plant. Hence, there are many practical industrial situations, like this one with multiple lines, where their use is limited. For example, for a production system consisting of a two-out-of-three set up with an inventory buffer (storage) facility, the mathematical analysis is very difficult and may be intractable. Thus a different approach is considered in this study.

3.3.3. Modelling multi-line production systems

The main objective of this research is to determine the downtime per unit time and thus the optimum inspection policy for a multi-line production system. In this case, downtime is defined as the duration of a stoppage to the downstream and/or the upstream processes. Consider the three-line scenario with no standby line in Figure 3.3. Under our definition, downtime occurs only when the individual stoppages coincide (period of length z in that figure). In other situations, upstream and downstream downtime may have different consequences and upstream and/or downstream inventory buffers may exist. For example, in a production system with a two-out-of-three line set up (see e.g. Smith and Dekker, 1997, or De Smidt-Destombes et al., 2007, for a general discussion of *k-of-out-n* systems) and one line used as standby (Figure 3.4), the definition of downtime should depend on the way the management operates the facility.

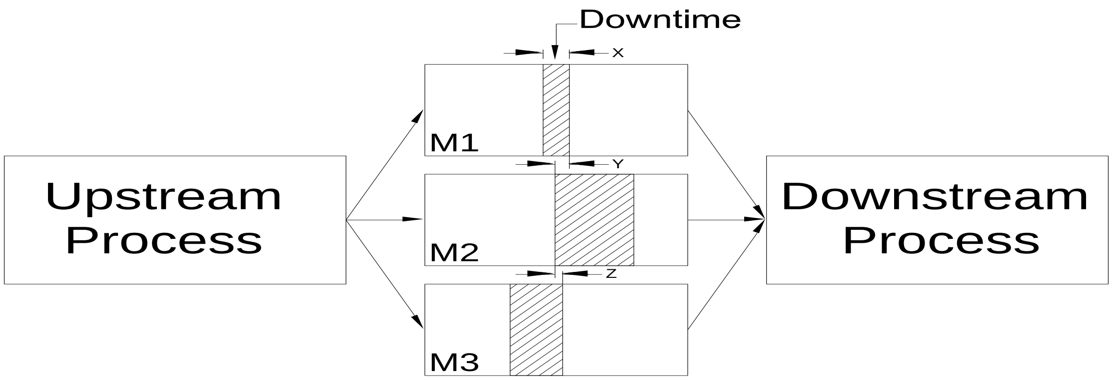


Figure 3.3. Plant downtime in a simple multi-line production system, indicating downtime for M1 of duration *x*, downtime of M2 that is concurrent with M1 of duration *y*, and complete system downtime of duration *z*.

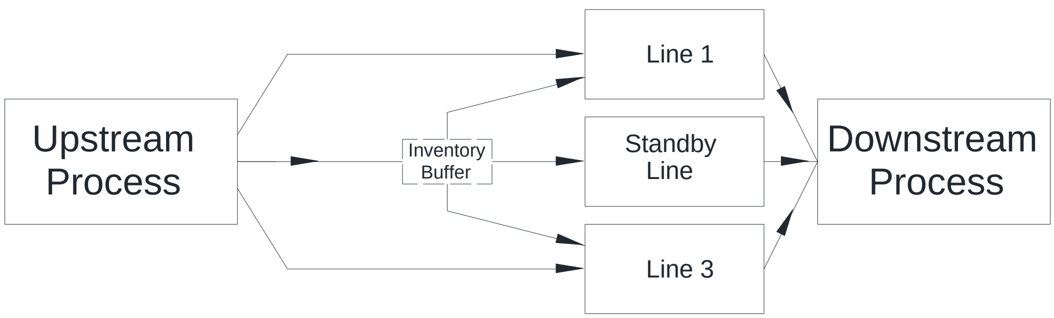


Figure 3.4. A multi-line production system with a two-out-of-three line set up and inventory buffer.

There are two principal ways in which inspection can be performed for the system in Figure 3.4, namely, simultaneous (concurrent) inspection of all parallel lines, or consecutive inspection, inspecting each in sequence. If inspection is performed simultaneously, assuming that the required resources (spares, personnel) are available, then the inspection time itself is downtime (similar to a single-line scenario), and the long-run cost per unit time (cost-rate) for the realisation shown in Figure 3.5(a) is:

$$C_{sim}(T) = \big(3(\,d_s + d_f\,)C_s + (d_s + d_v)C_d\big)/T. \quad (1)$$

With consecutive inspection, the cost-rate for the realisation shown in Figure 3.5(b) is:

$$C_{con}(T) = \big(3(\,d_s + d_f\,)C_s + (d_{v'} + d_v)C_d\big)/T < C_{sim}(T). \quad (2)$$

since $d_{v'} < d_s$ and $d_v < d_f$. In practice, it may be possible to reduce the cost of downtime further by modifying the policy so that if a failure occurs while another line is being inspected, the inspection is suspended until the failed line becomes operational. Then for the realisation shown in Figure 3.5(c), for example, the cost-rate is:

$$C_{mod}(T) = \big(3(\, d_s + \, d_f\,)C_s + d_v\, C_d\big)/T < C_{con}(T). \tag{3}$$

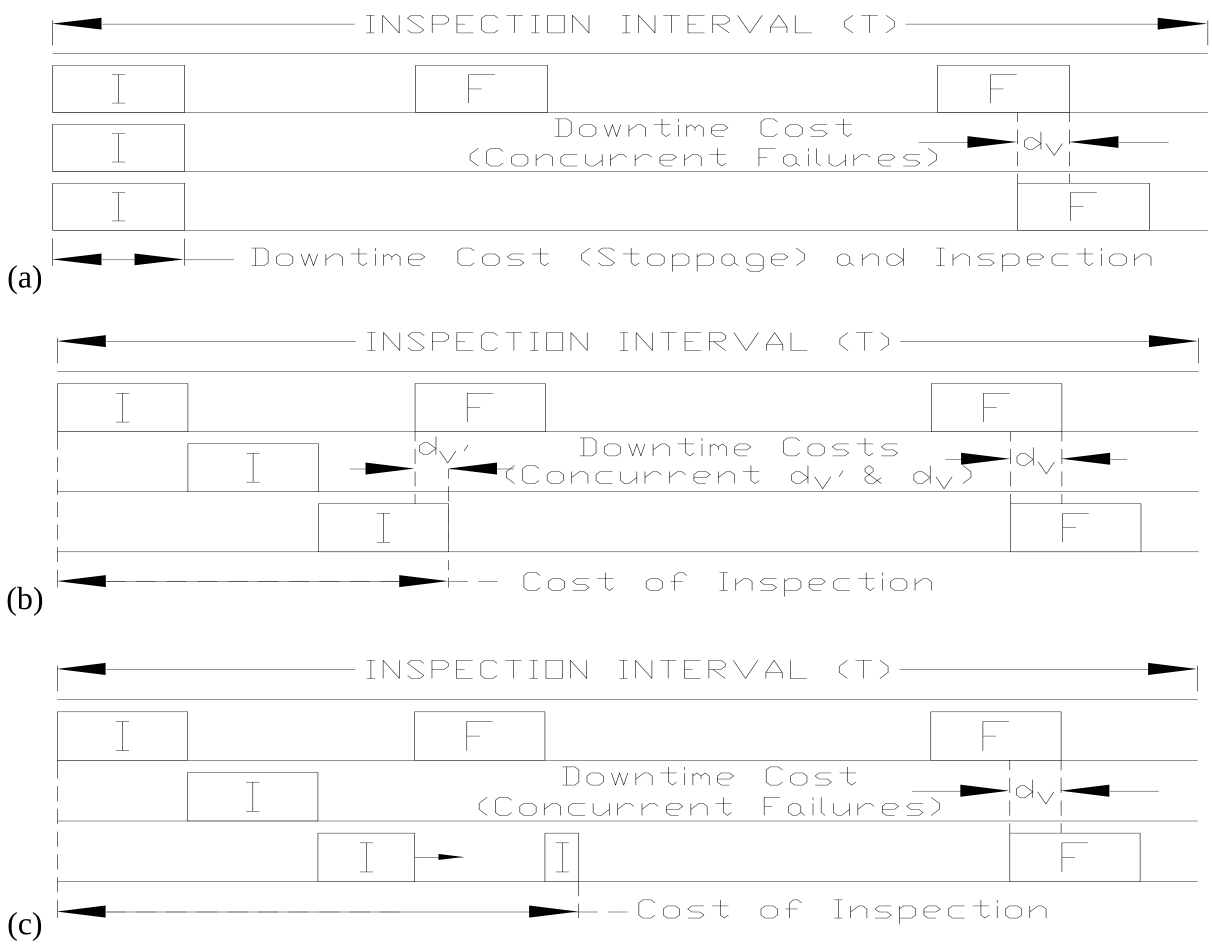


Figure 3.5. Policy schematic for two-out-of-three line system

(inspection I; failure repair F; concurrent occurrence of two failure stoppages d_v;

concurrent occurrence of a failure stoppage and an inspection $d_{v'}$):

(a) simultaneous inspection; (b) consecutive inspection; and

(c) consecutive inspection prioritising failure repair.

3.4. Case study

3.4.1. Problem description

Akbarov et al. (2008) studied a chocolate cake manufacturing plant with production downtime issues on its packing lines. They determined the optimum inspection interval for a single-line packing system using analytical modelling. In practice, the 'existence' of a defect may be identifiable by some operational "signal", such as, excess heat or vibration, and in this particular case study, a defect was observable as the presence of significant chocolate contamination on the production line. This was the direct cause of several major failure modes on the packing lines, so that preventive cleaning procedures (regular inspection and removal of chocolate contamination if required) were considered to be of value. Figure 3.4 is a schematic representation of the real production process at this plant, in which the upstream process bakes cakes and the two-out-of-three system packs them. A stoppage of the upstream baking process is considered as *downtime* (as it leads to lost revenue) whilst a stoppage of one of the packing lines is seen as *lost time* (as the packing process can still continue). Under normal production conditions, baked cakes are packed on lines 1 and 3, the inventory buffer is empty, and line 2 is on cold-standby. If there is a stoppage to either line 1 or 3, line 2 (the standby) is started and the cakes are routed through the inventory buffer to this line. The inventory buffer storage area is designed to provide sufficient capacity to start line 2 without having to stop upstream production. When normal production is resumed after a stoppage, there is sufficient capacity in lines 1 and 3 to empty the inventory buffer storage. Production downtime accrues from the instant the inventory buffer is full and two lines are down (see Crespo et al., 2003 for a detailed study on maintenance policies for a production system including inventory buffer capacity).

Although the real system has three parallel lines, to use the delay-time model analytically, Akbarov et al. (2008) consider this system as a single-line packing facility. It is interesting to know what are the implications (both for maintenance management of the plant and for modelling of the system) of this limiting assumption. To model the three-line system would invite a more complicated analysis which simulation would be able to offer. In fact, in the literature, the modelling and maintenance optimisation for parallel line systems is very rare. In this study, the process was begun by simulating the single-line model proposed by Akbarov et al. (2008) (Section 3.4.4) as a complex system of multiple components. In doing so, it was ensured that the base model was validated against known results, as did Boschian et al. (2009) for their case study, and not simply based on an arbitrary situation. With this impetus, the real practical situation for which analytical modelling is intractable (Section 3.4.5) was simulated. This scenario, which is precisely the system operated by the company's management, is called a *modified* two-out-of-three parallel line system here. Thirdly, a further simulation model was developed for a *standard* two-out-of-three packing parallel system, in which any two lines are operational at any one time. Although the company did not operate the packing facility in this way, the development of such a model was useful for comparison purposes (Section 3.4.6). Finally, the scenario in which all three parallel packing lines are operated concurrently was also simulated (Section 3.4.7).

The data for the base model were taken from Akbarov et al. (2008). Defect arrivals were described by the exponential distribution $F_U(\mathrm{u}) = 1 - \exp(-\lambda u)$ with rate $\lambda = 3$ per day; *delay-times* were described by the Weibull distribution $F_H(\mathrm{h}) = 1 - \exp(-(h/\alpha)^{\beta})$ with $\beta = 6.27$ and $\alpha = 0.193$ days, implying a mean *delay-time* of 4.3 hours and a standard deviation of 0.8 hours. Many previous studies have proposed, in detail, ways to select and estimate these parameters in practice (see, for example, Wang, 2008). The duration of a stoppage of a line due to failure was $d_f = 10$ minutes; and due to inspection was $d_s = 2$ minutes. Both inspection and repair in practice corresponded to removal of chocolate contamination, the former being carried out preventively, the latter correctively. The cost-rates of "inspection" and "repair" were thus assumed equal, and assigned as $C_s = £30$ per hour. The production downtime cost-rate was $C_d = £1{,}000$ per hour, based on the value of product output per unit time. The cost of a single failure event is then $d_f C_s$ plus the cost of the production downtime (if any) resulting from the failure.

3.4.2. Numerical example

The cost per unit time for simultaneous inspection for the realisation shown in Figure 3.5(a) is:

The cost per unit time = Cost of inspection for each of the three separate lines + Cost of downtime due to all three lines being inspected simultaneously + Cost of stoppages due to three failures + Cost of downtime due to two concurrent failures. Therefore:

$C_{sim}(T) = \left(3\ d_s\, C_s + \ d_s\, C_d + \ 3\ d_f\, C_s + \ d_v\, C_d\right)/T.$

With all the costs and durations given in Section 3.4.1, and with an arbitrary value of 3 minutes for d_v, since $d_v < d_f$, the cost is as follows (assuming that *T*=1 hour):

$C_{sim}(T)$ = (3 (2 + 10)/60 (30) + (2+3)/60 (1000)) = £101.33

For consecutive inspection, the cost per unit time for the realisation shown in Figure 3.5(b) for example is:

The cost per unit time = Cost of inspection for each of the three separate lines + Cost of downtime due to a concurrent failure and inspection process + Cost of stoppages due to three failures + Cost of downtime due to two concurrent failures:

$C_{con}(T) = \left(3\ d_s\, C_s + d_{v'}\, C_d + \ 3\ d_f\, C_s + \ d_v\, C_d\right)/T.$

With an arbitrary value of 0.5 minute for $d_{v'}$ (since $d_{v'} < d_s$), the cost is as follows:

$C_{con}(T)$ = (3 (2 + 10)/60 (30) + (0.5+3)/60 (1000)) = £76.33, which is $< C_{sim}(T)$.

For consecutive inspection prioritising failure repair, the cost per unit time for the realisation shown in Figure 3.5(c) for example is:

The cost per unit time = Cost of inspection for each of the three separate lines + Cost of stoppages due to three failures + Cost of downtime due to two concurrent failures:

$C_{mod}(T) = \left(3\ d_s\, C_s + \ 3\ d_f\, C_s + \ d_v\, C_d\right)/T.$

With the same arbitrary values for $d_{v'}$ and d_v as above, the cost is as follows:

$C_{mod}(T)$ = (3 (2 + 10)/60 (30) + (3)/60 (1000)) = £68.00, which is $< C_{con}(T)$.

3.4.3. Simulation modelling

ProModel (ProModel, 2016), a process-based discrete-event simulation language (see for example, Harrell et al., 2011), was used for developing the *base* model and the various model extensions. The models, composed of n machines (packing lines in this case study), denoted by M_n (for $n \geq 1$), as shown in the general schematic diagram (Figure 3.3) were developed as continuous production lines. The model was developed in three stages.

Stage 1: Construction of the overall model framework comprising the minimum system requirements

The development of any model using this programming environment requires, at least, the use of the paradigm 'LEAP'; Locations, Entities, Arrivals, and Processing. 'Locations', which may be single or multiple capacity, are generally fixed positions in the system, where entities wait to be processed, such as, machines, queues, or storage areas (buffers). 'Entities' are the objects that enter into, flow through and depart from the system as complete objects, such as parts, or even defect arrivals and failure occurrences. 'Arrivals' describe the precise pattern: timing; quantity; frequency; and location of Entities (defects, failures) entering into the system. And finally, 'Processing' defines the exact route that an Entity follows, from entering into, to leaving the system. This includes any activity that happens at a Location such as the required operations that need to be performed, the amount of time an Entity spends at a Location, and the Resources it needs to complete Processing. Although, the most simple model in this environment needs to have 'LEAP' described, any further sophistication needed almost certainly will require the use of other 'modules' and/or development of special programming routines.

Stage 2: Detailed programming of the maintenance strategy

The arrival time of defects (faults) and their evolution into failures over the *delay-time* period are generated and scheduled based on their respective distribution functions. The maintenance strategies are programmed by scheduling inspection intervals or failures occurrences, whichever occurs first, at which time the production of machine M_n is interrupted by the downtime process and is terminated after d_s or d_f periods respectively. All the relevant costs, system variables and attributes are constantly updated to determine the expected cost per unit time or cost-rate.

Stage 3: Development of the model scenarios, input data, output analysis, and optimisation

The developed simulation models are non-terminating and the unit of time is days. *Macros* were set up to be able to instantly change input data, such as C_d, C_s, d_f, d_s, $F_H(h)$, $F_U(u)$, λ, T, *warm-up* period, number of *replications*, and simulation *run length*. The continuous onscreen data for each model *replication* includes updating inspection duration, failure duration, downtime duration, total expected cost per unit time, number of defects present, number of defects removed, number of failure occurrences, and number of inspections taken place. The simulation output report includes various data and graphs, including the average cost per unit time and the average downtime per unit time. The models were *warmed-up* for the system to reach *steady state* before experiments could begin, with a suitable *warm-up* period determined using Welch's graphical procedure (e.g. see Banks et al., 2013). To achieve *steady state* in the output results, each of the simulation experiments were continued with a *run length* of 1,000 days and results from the first 10 days (*warm-up* period) were excluded to eliminate the transient components of the results, thus achieving *steady state*. Simulation experiments were conducted for 10 replications to achieve sufficient narrow 95% confidence intervals in the output data. The models were *run* through various simulation scenarios with different values of the inspection interval T. Finally, the optimum value of T was determined.

Figure 3.6 shows the flow chart of the base model, developed for the first simulation, representing a single-line packing facility. The graphical representation refers to eight different processing routines (modelling routines) which were developed for different aspects of the model conceptualisation. Table 3.1 displays a sample *ProModel* code written for the Failure Occurrence routine (see the flowchart), the time between a defect arising and the subsequent failure. Appendix 3.1 illustrates a summary of some of the main modelling routines, specifically for the modified two-out-of-three parallel system with consecutive inspection prioritising failure repair. Appendix 3.2 depicts the general layout for one of the simulation models. Appendix 3.3 illustrates the percentage time packing lines are either working or idle, under the modified 2-out-of-3 parallel system. Appendix 3.4 displays the average number of defects arriving at packing line 1, which confirms that the system indeed stabilises and reaches *steady state* by running 10 *replications*. Finally, Appendices 3.5 and 3.6 display detailed sample source data for the single-line packing system and the modified two-out-of-three parallel system, respectively.

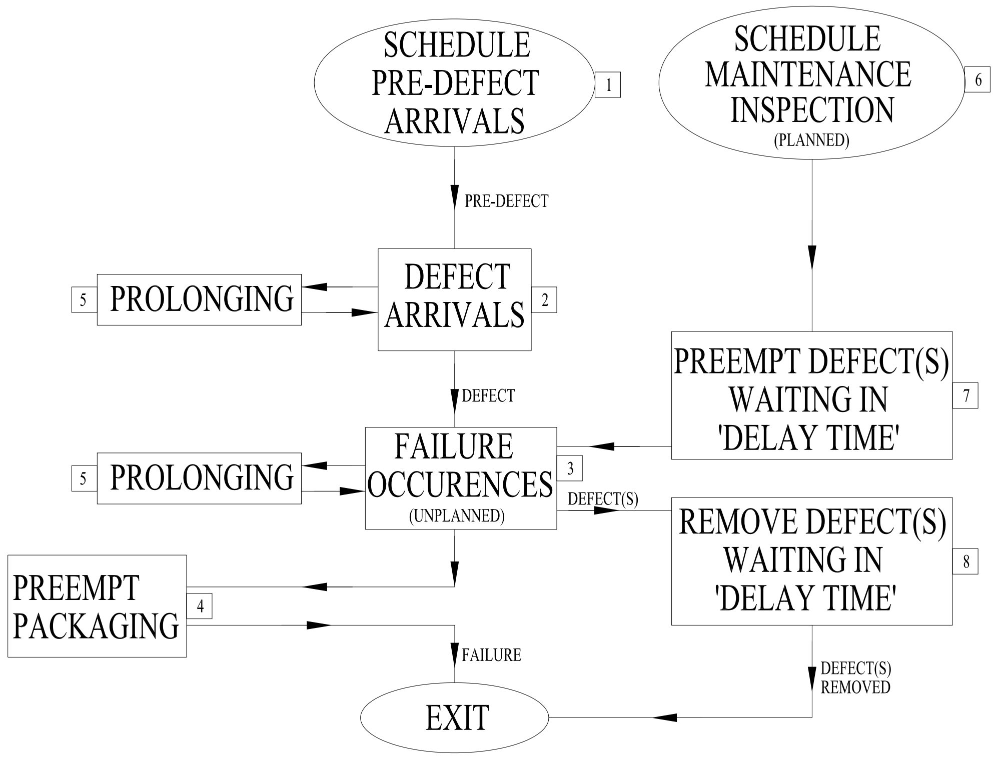


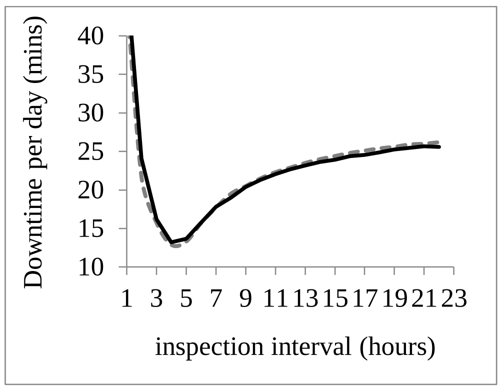
Figure 3.6. The *base* model for the single-line packing facility, showing eight programming algorithms.

Table 3.1. Sample *ProModel* Code for the Failure Occurrence routine.

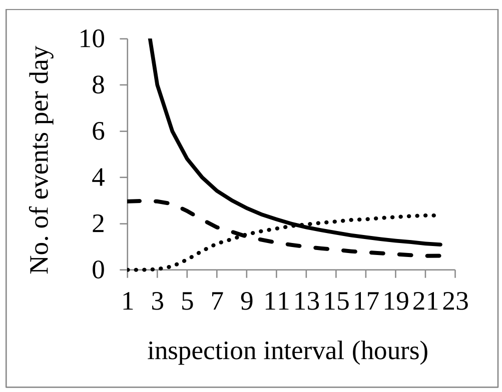
```
ORDER 1 PreDefect_E TO Leg1DefectArrival_L
REAL FailureRandom, FailureTimelapse
INC Leg1FailureGenerated_V
FailureRandom=RAND(1)
FailureTimelapse=277.43*(-LN(1-FailureRandom))**(1/6.27)
StartTime_A=CLOCK(MIN)
WAIT FailureTimelapse
ProlongingLeg1_Sub
IF Leg1DowntimeOnFailure_V=1 THEN
        WAIT Leg1DowntimeDur_V
IF Leg1InspectionOnFailure_V=1 THEN
        WAIT
(Leg1InspectionDur_V)+(CONTENTS(Leg1FailureOccurence_L)*Leg1RepairDur_V)
INC Leg1FailureArrived_V
DOWN Leg1ToGoDown,999
```

3.4.4. Base model (validation)

Figure 3.7(a) compares the results obtained from this simulation model with that of the analytical model from the Akbarov et al. (2008) study. It shows downtime per day, in minutes, against inspection interval, in hours. The results are clearly very close. Akbarov et al. (2008) recommended the same optimal inspection interval of 4 hours for removing the chocolate build-up from the packing machinery, with an expected production downtime of 12.3 minutes per day against a simulated downtime of 13.2 minutes per day.



(a)



(b)

Figure 3.7. Comparison of results for the single-line packing system:-

(a) expected downtime per day against preventive inspection interval:

▬▬ simulation; ▬ ▬ Akbarov et al. (2008);

and (b) expected number of events against preventive inspection interval:

▬▬ inspection; ▬ ▬ defects removed; ●●●● failures.

Two factors may have contributed towards the difference of 6.7% between the results of the two approaches: (i) the suspension of aging of plant components during inspection and failure; (ii) the possible overlap time of inspection and failure processes, both of which were included in the development of this simulation model and ignored in the previous study. With the defect arrival rate of 3 per day and the duration of a stoppage of a line due to failure of 10 minutes, there will be an expected downtime of 30 minutes/day when no preventive maintenance is carried out. The results suggest that regular inspection can reduce production downtime by 56%, with the number of failures per day reduced from 3 to almost zero (Figure 3.7(b)). The fact that the optimal inspection interval of 4 hours corresponds closely with the mean *delay-time* of 4.3 hours is not surprising given the relatively small *delay-time* standard deviation of 0.8 hours (since $\beta = 6.27$). Appendix 3.5 displays detailed simulation source data for this model.

3.4.5. 'Modified' two-out-of-three parallel system

If a policy of simultaneous inspection were to be followed for maintaining the *modified* two-out-of-three parallel line facility, then the simulation results shown in Figure 3.8(a) suggest that there is no optimal inspection interval; run-to-failure is then optimal. Here essentially the cost of lost production due to the stoppage of the upstream process during simultaneous inspection outweighs the cost of stoppages due to failure. In contrast, under a consecutive inspection policy, there is no planned downtime. There are, however, occasions when downtime may occur: (i) at least one failure and one inspection process occurring concurrently; (ii) two or more simultaneous failures. Furthermore, under a consecutive inspection policy prioritising failure repair, if a failure occurs while the inspection of another line is taking place, the inspection operation is stopped and then re-started once the failed line becomes operational.

The simulation results for these consecutive policies are also shown in Figure 3.8(a), but are shown again in Figure 3.8(b) to resolve the cost-rates for the policies of interest. Figure 3.8(b) suggests the optimal inspection interval is every 4 hours for consecutive inspection and 5 hours if failure repair is prioritised. The advantage of following the latter policy is less frequent inspections and a cost-rate reduction of 8.3%. In Figure 3.8(a), it can be seen that as T increases, the cost-rates converge fairly quickly (as expected since the *delay-time* variance is small). Finally, the cost-rate reduction for the best policy relative to a run-to-failure policy is of the order of 60%. Appendix 3.6 displays detailed simulation source data for this model.

(a)
(b)

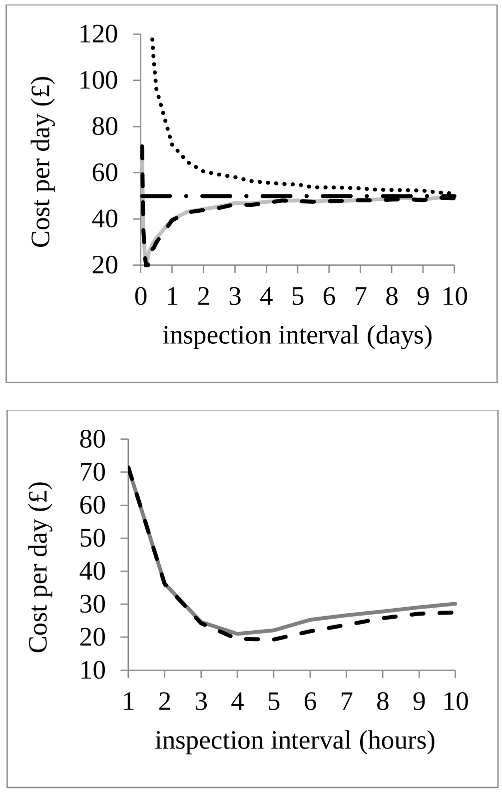


Figure 3.8. Cost-rate as a function of inspection interval:

(a) all policies with inspection interval up to 10 days; and (b) consecutive policies only.

●●●● simultaneous inspection; ▬▬ consecutive inspection;

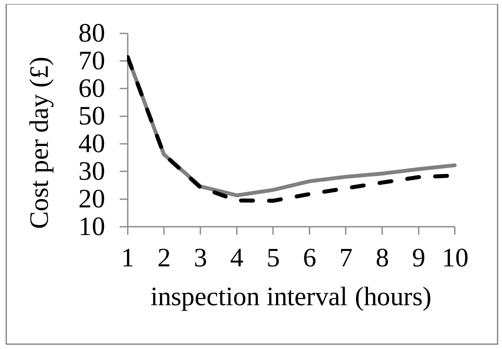
▬ ▬ consecutive inspection prioritising failure repair; ▬▬ ● ▬▬ run-to-failure.

3.4.6. ‘Standard’ two-out-of-three parallel system

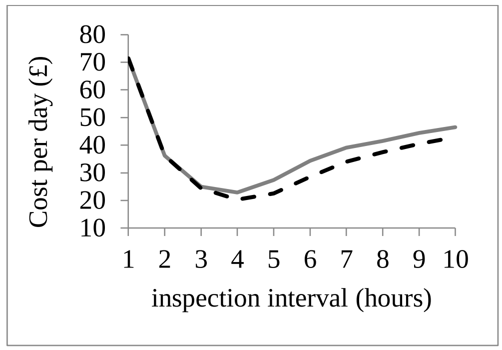
A *standard* two-out-of-three parallel line configuration was also investigated in order to compare results with the *modified* two-out-of-three parallel line system discussed above. For such a system, any two parallel lines would be operational at any one time, so that all packing lines would be equally utilised in the long run. A failed line would be repaired and ready to use at the next line failure. For the consecutive inspection policy, the cost-rate appears to be either equal or higher than that for the *modified* parallel system (Figure 3.9(a)). Similarly, for the consecutive inspection prioritising failure repair policy, the cost-rate appears to be equal or slightly higher than that for the *modified* parallel system. This is due to all three lines having been utilised more uniformly and hence causing more simultaneous failures. For this system, the optimal interval remains the same, at 4 and 5 hours, for the consecutive inspection policy and the consecutive inspection prioritising failure repair policy, respectively. However, there will be 1.6% and 0.6% increases in the cost-rate for these policies when compared to the *modified* two-out-of-three mode of operation.

3.4.7. Three-parallel lines system

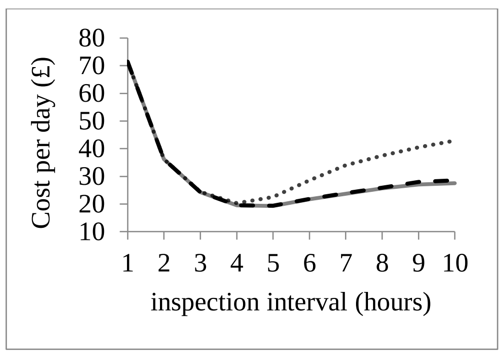
The final modelling scenario considered the system with three parallel lines. Although there cannot be any direct comparison between this and the previous two systems, looking at the related results alongside the two-out-of-three parallel systems is useful in case production needs to be increased. For the three-line system, downtime will necessarily be greater than under both two-out-of-three systems (Figure 3.9(b)) because there is more chance of at least two failures occurring simultaneously. However, production output will also be higher. As discussed before, the most sensible policy applicable in practice will be consecutive inspection prioritising failure repair. Figure 3.9(c) compares the costs for this policy under all three systems. The cost of downtime (downtime-rate) is higher if all three lines are operated at the same time.



(a)



(b)



(c)

Figure 3.9. Cost-rate as a function of inspection interval:-

(a) *standard* two-out-of-three parallel system:

━━━ consecutive; ━ ━ consecutive prioritising failure repair;

(b) three-line system:

━━━ consecutive; ━ ━ consecutive prioritising failure repair;

and (c) consecutive inspection prioritising failure repair for each system:

━━━ *modified* two-out-of-three; ━ ━ *standard* two-out-of-three; ●●●● three-line.

3.4.8. Sensitivity analysis

The sensitivity of the consecutive inspection prioritising failure repair policy was investigated for the principal mode of operation of interest (*modified* two-out-of-three system) to parameter values. Figure 3.10(a) shows the sensitivity to inspection duration, d_s. The behaviour is as expected here, with the cost-rate of the optimum policy for $0.5d_s$ and $2d_s$ at respectively 54% and 173% of the baseline.

Sensitivity to variation in the failure stoppage duration (Figure 3.10(b)) shows a somewhat different pattern (optimum cost-rate is 84% and 110% of the original cost for $0.5d_f$ and $2d_f$, respectively). Varying d_f has the greatest effect when inspection is infrequent; varying d_s has the greatest effect when inspection is frequent, again as it would be expected, since failure stoppage duration dominates when inspection is infrequent, and downtime due to inspection duration dominates when inspection is frequent. Sensitivity to the rate of arrival of defects (formation of chocolate contaminations), Figure 3.10(c), shows anticipated effects although the doubling of the failure rate is not sufficient to increase the optimum inspection frequency.

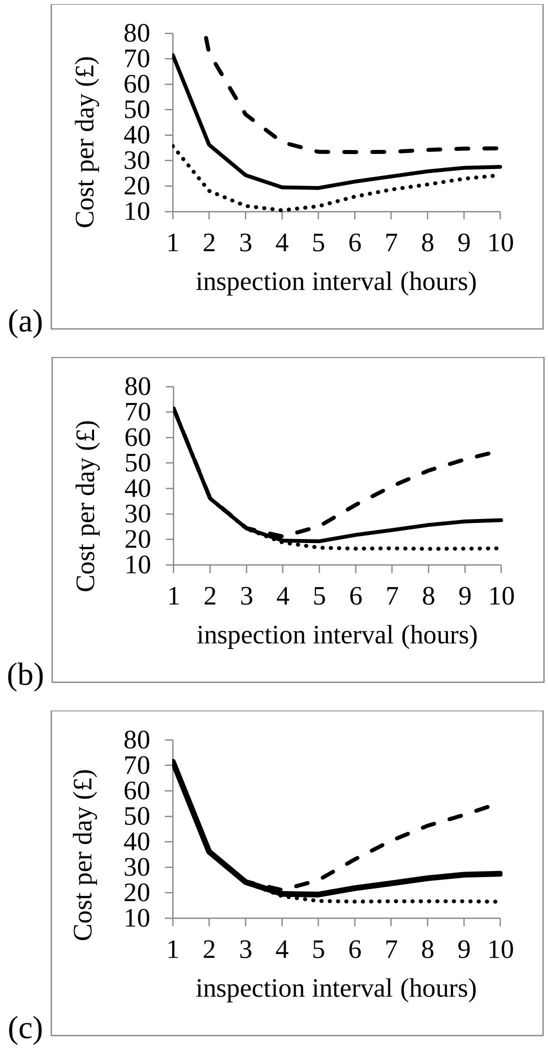


Figure 3.10. Sensitivity analysis of parameters:-

(a) inspection duration, d_s: ●●●● $0.5d_s$; —— d_s; – – $2d_s$;

(b) failure stoppage duration, d_f : ●●●● $0.5d_f$; —— d_f; – – $2d_f$;

and (c) defect arrival intensity, λ: ●●●● 0.5λ; —— λ; – – 2λ.

3.5. Conclusions and further work

Almost all previous delay-time inspection models in the literature are concerned with single-line single-component systems or series systems with multiple components, with restrictions. This study uses simulation to determine optimal inspection policy for a number of multi-line production facility scenarios using the delay-time concept. In the first scenario, a single-line facility is simulated to validate earlier analytic results. In the second, a *modified* two-out-of-three parallel system is analysed to help address the issue of plant downtime under the *actual* operating conditions in the case study. Two further model extensions are developed and analysed to consider whether modifications to either the operation of the system or the design of the system in the case study would be worthwhile. The latter three models extend the study by Akbarov et al. (2008), in which the multi-line production facility is modelled as if it is a single line. In fact, in their survey, Alrabghi and Tiwari, (2015) found that studies that dominate the literature, such as cases of single machines producing single products, are oversimplified and do not reflect the complexity and interactions of real systems in practice.

It is found that: 1) the simulation of the single-line system validates (reproduces) earlier results; 2) consecutive inspection with prioritised failure repair lowers the cost-rate (by 8.3%) and reduces the frequency of inspections (by 20%) compared to consecutive inspection; 3) the standard two-out-of-three design configuration increases the cost-rate by 1.6% and 0.6% for the consecutive inspection and consecutive inspection prioritising failure repair policies respectively compared to the modified two-out-of-three configuration operated by the management; and 4) the three parallel-line design configuration increases the frequency of inspections (by 25%) and increases the cost-rate (by 5.2%) for the consecutive inspection prioritising failure repair. This is clearly due to the third line being used permanently and not as a *standby*, which would naturally increase the number of inspections and the possibility of further failure occurrences. However, it should be noted that the production throughput would increase as well, increasing the revenue.

The solution proposed in this chapter may seem rather obvious as it recommends the consecutive inspection policy with priority given to failure repair for the maintenance management of multi-line production systems. However, the implications for this case study are substantial as the policy proposition suggests a cost reduction of 61.3% compared to the 'run-to-failure' policy. Furthermore, analytical models are complex and intractable for determining the optimal inspection interval in a parallel multi-line setting. In this respect, it can be argued that the contribution of this study is first and foremost in narrowing the gap between the theory and practice of managing multi-line systems. The advantage of using a case study is that it illuminates the practical problems that operations managers face in everyday real-world situations and the complexities that may exist in developing pragmatic solutions. This work demonstrates, in particular, that the scenarios and policies considered have economic and engineering implications for the management of production lines and that maintenance planning and execution first and foremost needs to be responsive to operational requirements. Finally, real-time decision making using simulation would be very useful in dynamic situations where the condition of the system state is monitored.

Regarding the scalability of the model, it should be noted that the simulation model for this study has been specifically developed to address the optimisation of the inspection interval for a very specific two-out-of-three parallel production system. However, the model is easily scalable. It takes 55 seconds to run the model through one replication for the current scenarios. It takes approximately 10 minutes to simulate 10 independent replications of 1000 working days on a standard desktop computer.

Extensions to the work presented in this chapter may be developed in several directions. The assumption of perfect inspection may be relaxed. Indeed, Wang (2008) noted that in all the case studies conducted using the delay-time concept, none of them supported the perfect inspection assumption. Other models might consider imperfect repair, multi-level inspections, delayed replacement or repair (of defective components identified at inspection), manpower planning and spare parts planning, although not all extensions are relevant to the particular case study described here. Simulation is the ideal tool for extending delay-time modelling research in these areas. The natural future extension to this work will be the joint optimisation of planned maintenance and various spare parts inventory control policies for multi-line parallel systems.

Chapter 4

Joint optimisation of inspection maintenance and spare parts provisioning: a comparative study of inventory policies

4.1. Summary

The demand for industrial plant spare parts is driven, at least in part, by maintenance requirements. It is therefore important to jointly optimise planned maintenance and the associated spare parts inventory using the most appropriate maintenance and replenishment policies. In this simulation-based study, this challenge is addressed in the context of the random failure of parts in service and the replacement of defective parts at inspections of period T. Inspections are modelled using the delay-time concept. A number of simultaneous periodic review and continuous review ordering policies are compared. A paper making plant provides a real context for the presentation of these ideas. Practitioners working with such plant are surveyed in order to collect real data that informs the values of parameters in the models. The simulation results indicate that the periodic review (R, S) policy with $T = 2R$ is cost optimal. The discussion also gives insights into the characteristics of the policies considered, including that at the optimal inspection interval, the $(R, s, Q, T = R)$ policy has lower ordering cost per unit time than the $(R, s, S, T = R)$ policy, but higher holding cost per unit time than the $(R, S, T = R)$ policy.

4.2. Introduction

The demand for spare parts for industrial plant is predicated on the operation and maintenance of the plant. Therefore, the planning of spare parts inventory should be driven by operational and maintenance requirements rather than the observation of demand. This is because operation and maintenance schedules provide partial information about the demand for spare parts in advance, and the forecasting of spare parts demand based on historical usage is sub-optimal (Ghobbar and Friend, 2003; Boylan and Syntetos, 2010). Furthermore, maintenance planning that assumes 100% availability of spare parts is also sub-optimal (Sharma and Yadava, 2011). Consequently, it is important to coordinate the planning of operation, maintenance and spare parts inventory (Wang and Syntetos, 2011). Many researchers have tackled this coordination of maintenance and inventory separately or sequentially (De Almeida, 2001; Marseguerra et al., 2005; Cheng and Tsao; 2010, De Almeida et al., 2015). However, it has been demonstrated that joint optimisation is superior (in a cost sense) to separately or sequentially optimised policies (Sarker and Haque, 2000).

Therefore, it is important not only to coordinate operation and maintenance planning and spare parts inventory control but also to carry out optimisation jointly. This is precisely what is done in this chapter, whilst setting aside the question of coordination with operation by supposing that a plant is continuously or regularly operated. Focus is put on the joint cost-optimisation of planned, periodic inspection maintenance and each of several periodic and continuous review replenishment policies. The delay-time concept is used to model inspection maintenance. A simulation model is developed in the context of a paper machinery plant. Further, data is collected that inform the values of parameters in the simulation using a survey of practitioners working with such plant. Several replenishment policies are considered so as not to select one arbitrarily, as is the case in almost all previous joint optimisation studies (Van Horenbeek et al., 2013). Thus, two contributions are made in this chapter: (i) this is the first study to consider a range of inventory replenishment policies in joint optimisation with maintenance planning; and (ii) it develops insights into the characteristics of each policy, not previously addressed in the joint optimisation studies. The use of simulation allows the models to make less simplifying assumptions than is usual with more analytical papers including: (i) the selection and use of many statistical distributions, other than the exponential distribution, such as Weibull; (ii) the prolonging of defects' *delay-times* during system downtime; (iii) the overlap between inspection and failure activities; and finally (iv) the stochastic behaviour of spare replenishment lead-time, emergency lead-time, and the availability of various maintenance resources.

This chapter is organised as follows. Section 4.3 critically reviews the joint optimisation literature. Here, the aim is to justify the use of the delay-time concept and the selection of replenishment policies. In Section 4.4, an industrial problem and the modelling of a complex system of multiple components (bearings) is described. In Section 4.5, the simulation models are discussed, including their assumptions and cost factors. The results are analysed and discussed in Section 4.6, including sensitivity analysis. In the final section, conclusions are drawn and proposals are developed for future research.

4.3. Literature review

There are numerous industrial situations where either the replacement of multiple plant items is too costly, prohibiting the block replacement policy, or too critical to bear the risk of replacements at a pre-specified item age (items/units/components are parts of machines or equipment, which may be repaired or replaced). In these circumstances, the reasonable and rational maintenance strategy is to inspect periodically and replace only the defective (faulty) parts while retaining the good ones. Such an inspection strategy can be modelled with the delay-time concept (Wang, 2012a) used here and utilised in numerous case studies (e.g. Christer and Wang, 1995; Pillay et al., 2001a; Akbarov et al., 2008; Jones et al., 2010; and Emovon et al., 2016). The delay-time concept has the advantage that it explicitly models the relationship between plant failures and the inspection interval (see Section 4.4.3).

Classic maintenance policies (e.g. Barlow and Proschan, 1965) assume 100% availability of spares, implying that spare parts are either highly standardized for easy procurement from suppliers, or so inexpensive that large quantities may be stored to protect against possible stock-outs. Nevertheless, parts are in fact highly customized (and potentially very expensive), and their procurement lead-time cannot be neglected (Brezavscek and Hudoklin, 2003).

A specific question that motivates the publication of this study is the extent to which particular maintenance policies and particular inventory policies have been jointly optimised. In a way to answer this question, the papers reviewed by Van Horenbeek et al. (2013) and others that have appeared since 2013 have been classified in Table 4.1.

Table 4.1. Characteristics of the joint maintenance and inventory control models in journal papers.

Author (Year)		Maintenance policy		Replenishment policy		Model development		Components in system	
		Age-based	Block-based	Periodic Review	Continuous Review	Analytical	Simulation	Single	Multiple
Sarker and Haque (2000)			♦ Periodic		♦ (*s*,*S*)		♦		♦
Chelbi and Ait-Kadi (2001)			♦ Periodic	♦ (*R*,*S*)		♦			♦
Yoo et al. (2001)			♦ Periodic	♦ (*R*,*S*)		♦			♦
Brezavscek and Hudoklin (2003)			♦ Periodic	♦		♦			♦
Vaughan (2005)	*		♦ Inspection		♦ (*s*,*S*)	♦			♦
Ilgin and Tunali (2007)			♦ Periodic		♦ (*s*,*S*)		♦		♦
Huang et al. (2008)	*		♦ Periodic	♦		♦			♦
De Smidt-Destombes et al. (2009)			♦ Periodic	♦		♦		♦	
Wang (2011a)	#		♦ Inspection (DTM)	♦ (*R*,*s*,*Q*)		♦			♦
Wang (2012b)	#		♦ Inspection (DTM)	♦ (*R*,*s*,*Q*)		♦			♦
Chen et al. (2013)	#	♦			♦ (*s*,*Q*)	♦			♦
Panagiotidou (2014)	#		♦ Inspection	♦ (*R*,*S*)	♦ (*s*,*S*)	♦			♦
Gan et al. (2015)	#	♦		♦		♦		♦	
Jiang et al. (2015)	#		♦ Periodic	♦ (*R*,*S*)		♦			♦
Samal and Pratihar (2015)	#		♦ Periodic	♦		♦			♦
Alrabghi and Tiwari (2016)	#		♦ Periodic		♦ (*s*,*Q*)		♦		♦
Zahedi-Hosseini et al. (2017)	#		♦ Inspection (DTM)	♦ (*R*,*S*); (R,s,S); (R,s,Q)	♦ (*s*,*S*); (s,Q)		♦		♦

* Not listed in the Van Horenbeek et al. (2013) review as one of the joint optimisation of maintenance and inventory control papers.
# Not included in the latest review paper by Van Horenbeek et al. (2013) since they were published after its submission date (2010).

It can be seen that only Panagiotidou (2014) uses more than one inventory replenishment policy for comparison. This then reinforces the view that it is timely to consider more than one potential candidate replenishment policy with a particular maintenance policy. While the author also considers inspection maintenance, the model employed (that considers two types of failure: minor and major) is different to this current study. Apart from Panagiotidou (2014), the closest in scope to this study are Wang (2011a, 2012b), which use the delay-time concept but only consider the (R, s, Q) inventory policy. The quantities $R, s,$ and Q, and S in Table 4.1 are the standard inventory control policy parameters, defined in the notation list in Section 4.4.1.

Furthermore, the literature review carried out for this study indicates that in joint optimisation studies: (i) few researchers have considered age-replacement, most have considered block replacement and the literature on inspection-based maintenance is growing; (ii) many researchers continue to use analytical models with restrictive assumptions rather than simulation; (iii) interest in models for multi-unit systems is growing considerably; and finally (iv) both periodic and continuous review replenishment policies, almost equally, are used in these studies.

4.4. Modelling methodology

4.4.1. Notation

See Section 2.5.2., for a list of Notation associated with this chapter.

4.4.2. Problem description

A specific industrial plant situation (a paper mill) is considered, and in particular, several simulation models are developed for jointly optimising the inspection maintenance and the inventory policy for bearings, which are critical components in the plant. Bearings are used extensively in many types of plant and can result in very high costs due to unexpected and catastrophic failures. Folger et al. (2014a, 2014b) describe four conditions under which a bearing might not reach its maximum life: (i) improper handling and installation; (ii) inadequate lubrication; (iii) contamination; and (iv) force, speed, and temperature overload. The consequences of damage to a bearing system in industrial machinery can be very significant, including general risks to safety, cost to repair or expensive replacements, and unplanned machine downtime.

The situation which is described is not a case study. Rather, it is an idealised context that is closely informed by a survey. In this survey, maintenance experts and paper manufacturers were asked about their experience of paper making machinery in general and the critical components of these machines. A questionnaire, with fifteen questions, was used to obtain information about: inspections; replacements; failure replacements; costs and lead-times; possible defect arrival patterns, *delay-times*, and their distributions; current maintenance policy(s) used; and finally, the policy(s) used for the replenishment of spares for critical components. Nine questionnaires of the 15 distributed were returned. The respondents were three experienced maintenance and inventory control researchers and six paper machine manufacturers. Where available, the range of values of the parameters that were provided have been indicated. The information obtained from the survey ensured that the models and simulation experiments were realistic, and were the basis for the costs and parameter values used in the models. See Appendix 4.1 for a copy of the survey questionnaire that is described below.

In the models (the idealised context) that are informed by this survey, it is supposed that:

- The monitoring of bearing condition is carried out by a third party (external specialists).

- Bearing condition is only reported periodically by the third party following processing of the raw condition data. This is typical of condition monitoring arrangements in modern plant (Wang and Wang, 2015). In this way, condition monitoring and the reporting of bearing condition in particular incurs cost but no plant downtime.

- Bearings that are reported as defective are replaced; this intervention has a non-zero downtime.

- Failed bearings are replaced immediately since it is supposed that such events cause an immediate unscheduled stop to the plant and production downtime.

- At failure, only failed bearings are attended to, in order to return the plant to operation as quickly as possible. Thus there is no inspection of bearings at failure events.

These assumptions describe the system in broad terms. Next, the inspection model is described in detail, followed by the inventory model and then the model parameters.

Note, in this chapter, the terms bearings, components, items, and spare parts are used interchangeably.

4.4.3. The inspection maintenance model and its assumptions

It is supposed that the system has *n* identical bearings that are subject to deterioration. In this (complex) system model, multiple concurrent defects are possible and the failure process of a bearing has two-stages, according to the delay-time concept. In the first stage, a bearing is good and working normally. Then in the second stage a bearing is working but defective. The second stage terminates with a failure. The length of the first stage is the *time-to-defect arrival*, and the length of the second stage is the *delay-time*. If inspection is carried out during this second stage, it is assumed that defective items are identified and replaced, as depicted in Figure 4.1. Only defective items identified at inspections are replaced, rather than replacing all items regardless of their age or conditions.

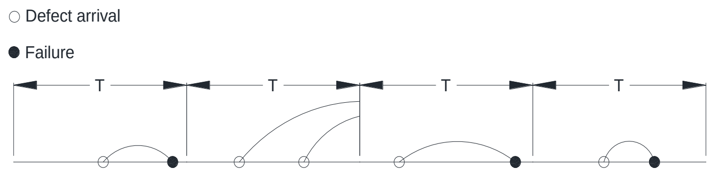


Figure 4.1. Defect arrivals and failure occurrences
in this complex system of multiple components.

The maintenance policy that is considered inspects all items in parallel every T time units, and defective bearings are replaced preventively. On failure, only failed bearings are replaced. It is assumed that the system is in a state of suspension while replacement is carried out. Therefore, defects do not grow and the bearings do not age during replacement downtime, and defects and failures can only arise whilst the plant is operating. The system is assumed to be operating under *steady state* conditions. Any operational loss due to the presence of defects other than inspection, replacement and failure are ignored. These are standard assumptions in inspection models (Wang, 2011a).

Times between defect arrivals are assumed to be independent, exponentially distributed, consistent with the delay-time model of a complex system (see for example, Wang, 2012b), with a defect arrival rate (intensity) of 0.05 per week. This value and the others that follow are based on the survey. Note, it is only necessary to specify defect arrival rate for the collection of identical bearings as a whole, as it is done here. Nonetheless, this quantity can be determined as the product of the number of bearings and the "per-bearing" defect arrival rate. The number of identical bearings in a paper rolling machine is typically large (>100) (Wang, 2011a). The *delay-time* follows the Weibull distribution with scale and shape parameters, $\alpha = 10$, $\beta = 3$ respectively (implying a mean *delay-time* of 8.93 weeks). The downtimes due to each replacement and failure are $d_r = 4$ hours = 0.024 weeks (survey range 1-6 hours) and $d_f = 9$ hours = 0.054 weeks (survey range 1-36 hours).

4.4.4. The inventory control model and its assumptions

Paper machinery typically have many identical bearings. In the model, it is supposed that inventory planning is concerned only with these bearings. That is, inventory policy for a single stock keeping unit will be considered. The following assumptions are made. The demand for bearings is generated through two routes. Failures of parts in service occur between inspections, which generate intermittent single-unit demands. And, every T time units, at scheduled inspections, all defective bearings are identified and preventively replaced (provided there are enough spares), generating ‘lumpy’ demand. Demand arising from failure replacements and due to preventive maintenance at inspections are satisfied from the existing inventory or by expediting an emergency order.

Using the simulation models developed, several periodic and continuous review inventory policies will be compared (see, for example, Muller, 2011 and Silver et al., 2016). As depicted in Figure 4.2, these policies include: (i) the periodic (R, S) policy, where every R time units (the review period) an order is placed to raise the inventory position to level S; (ii) the periodic (R, s, S) policy, where every R time units, an order is placed to raise the inventory position to level S provided the inventory position has reached or fallen below the re-order level s; (iii) the periodic (R, s, Q) policy, where every R time units an order of Q units is placed provided the inventory position is less than or equal to s; (iv) the continuous (s, S) policy, where an order is placed to raise the inventory position to level S when the inventory position falls to or below level s; and finally (v) the continuous (s, Q) policy, where Q units are ordered when the inventory position falls to or below level s.

In the five policies illustrated in Figure 4.2, orders are placed at points A, C, E and G, for example, and arrive at points B, D, F and H respectively, after a lead-time, L. It is important to note that although the same arbitrary demand profile has been used for all five replenishment policies in Figure 4.2, the ordering outcome is different in each case. Also, in this illustration $L < R$ for simplicity, but this restriction is not imposed in the model.

Here, values of the decision variables that minimise the long-run expected cost per unit time or cost-rate, $C(T)$ are sought. The set of decision variables depends on the exact inventory policy considered. In all cases, the joint policy contains the decision variable T, the inspection interval.

Throughout the analysis, the unit of time is one week. This is an arbitrary unit that is convenient for the reporting of the results.

Based on the survey, the lead-time L is set at 3 weeks (survey range 2-6 weeks) and the shortage emergency delivery lead-time, L_{sh}, fixed at 1 day (survey range 1-10 days). Further it is assumed that orders are placed at the end of each order placing day. Orders arrive at the beginning of each order receipt day but before reviewing the current inventory if it coincides with an order placing day. The last two assumptions are for modelling purposes. The order interval is flexible since there might be times that based on the current inventory level the optimal cost will be achieved by ordering no spares. The preventive maintenance interval $T = kR$ for any positive value of k. In these models, based on the survey data, $C_o = £100$ and fixed, including the cost of delivery. C_h is costed at 1% of item cost per week. $C_u = £2{,}000$ per item (survey range £1000-4000), and finally, $C_{sh} = £1{,}000$ per emergency shipment (survey range £500-1200).

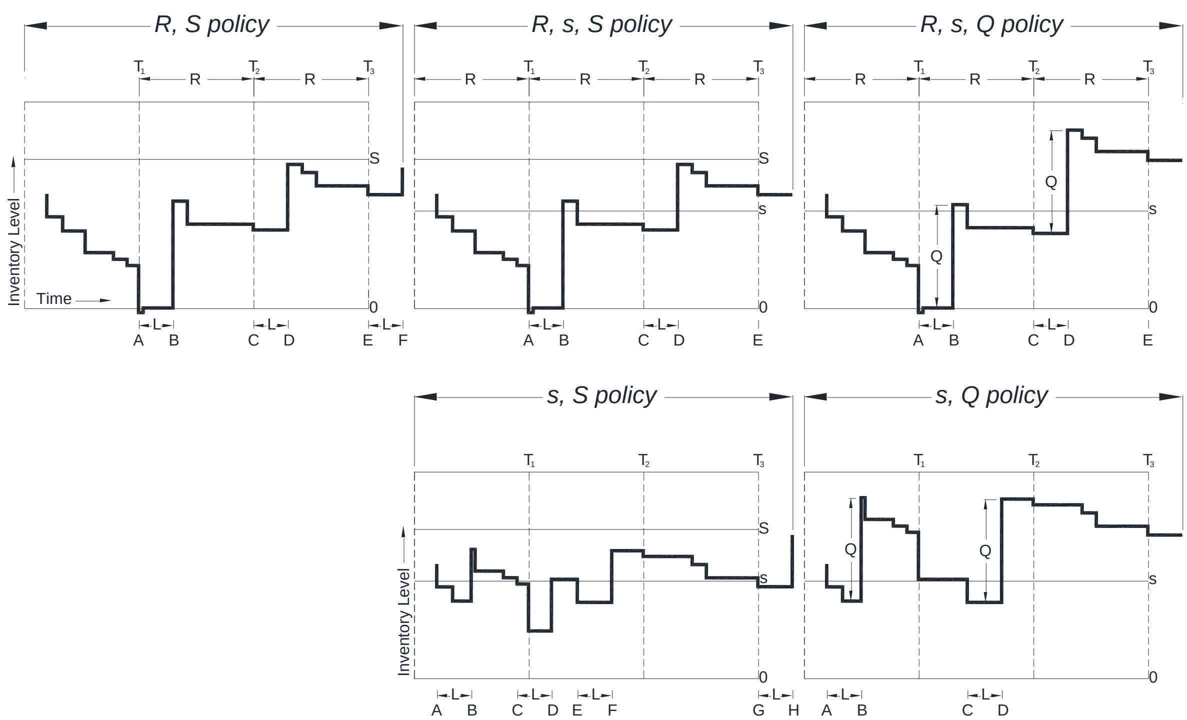


Figure 4.2. Characteristics of the periodic and continuous review inventory replenishment policies.

4.4.5. Costs and downtime specifications

The parameters C_r and C_f are costed on the basis of 3 maintenance technicians at C_m = £60 per maintenance technician per hour. Further, C_d = £1,000 per hour. C_i = £1,000 and is fixed.

One maintenance technician assists the external consultants during an 8 hour shift, performing data analysis and reporting. Therefore: $C_i = 1{,}000 + 8\ C_m = £1{,}480$.

There are two preventive maintenance (PM) inspection renewal costs depending on whether spare parts are immediately available. PM replacement cost (C_r) includes 3 maintenance technicians for 4 hours. If spare parts are available, then the cost includes downtime and manpower, hence: $C_r = 4\ C_d + 12\ C_m = £4{,}720$. If spare parts are not immediately available, there are two extra costs, namely, the shortage shipment cost and the downtime cost while the emergency shipment is in transit, hence: $C_r = 4\ C_d + 12\ C_m + C_{sh} +\ C_d\ L_{sh} = £29{,}720$.

There are also two failure renewal costs depending on whether spare parts are immediately available. Failure replacement cost (C_f) includes 3 maintenance technicians for 9 hours. If spare parts are available, then the cost includes the costs of downtime and manpower, hence: $C_f = 9\ C_d + 27\ C_m = £10{,}620$. If spare parts are not immediately available, then there are two extra costs, namely, the shortage shipment cost and the downtime cost while the emergency shipment is in transit, hence: $C_f = 9\ C_d + 27\ C_m + C_{sh} +\ C_d\ L_{sh} = £35{,}620$.

4.5. Simulation modelling

Using a modular approach, simulation models were developed for the joint optimisation of the inspection maintenance and the periodic and continuous review inventory control policies of interest. *ProModel*, a process-based discrete-event simulation language (see, for example, Harrell et al., 2011) was used to model a system with a single machine (but extendible to multiple machines or lines) as a continuous production system with the consideration of all major assumptions and cost figures given in Sections 4.4.1 to 4.4.5.

The construction of the overall simulation model framework, the modelling methodology, and finally the input parameters, output analysis, model scenarios and the optimisation technique are discussed in Sections 4.5.1 to 4.5.3.

4.5.1. Construction of the model framework and the minimum system requirements

The development of any model using the *ProModel* programming environment requires, at least, the use of *LEAP*: *L*ocations, *E*ntities, *A*rrivals, and *P*rocessing. *Locations*, which may be single or multiple capacity, are fixed positions in the system where *entities* may wait to be processed, such as, machines, queues, or buffer storage areas. *Entities* are objects that enter into, flow through and depart from the system, such as parts, defects and failed items. *Arrivals* describe the precise pattern: timing; quantity; frequency; and location of *entities* entering into the system. And finally, *processing* defines the exact route that an *entity* follows, from entry, to leaving the system, including all programming logic. The *processing* also includes operations, their durations, and the *resources* needed. Although, the simplest model needs to have at least *LEAP* described, any further model sophistication almost certainly requires the use of other *modules* and/or development of special programming routines. In these models *variables*, *attributes*, *resources* and *path networks* were used extensively. S*ubroutines* were developed for generating *delay-times* for the failure and replacement processes, and for calculating various costs in the model, for example. Various *macros* were also set up to enable the easy alteration of decision variables. The following section describes in detail the approach taken in developing the various simulation models.

4.5.2. Simulation modelling methodology

Appendix 4.2 illustrates the flow chart of the general simulation procedure, depicting the nine modelling routines developed for different aspects of the model. Referring specifically to the numbers stated in each box in Appendix 4.2, a simple description of the flowchart is given below:

1. Defects are scheduled to arrive with the appropriate arrival intensity based on the given statistical distribution;

2. Defects arrive and wait in their *delay-time*;

3. Machine downtime is scheduled, either due to defects transforming into failures (unplanned process), or defects identified at preventive maintenance instances (planned process), triggering the replacement of bearings (if needed);

4. Machine downtime occurs;

5. Defects are *delayed* or *prolonged* during the *time-to-defect* and *delay-time*, if necessary, since defects and failures can only arise whilst the plant is operating, and defects do not grow and the bearings do not age during replacement downtime;

6. Planned maintenance is scheduled;

7. Defects (if any) are identified so that bearings may be replaced;

8. Defects (if any) are removed by the replacement of bearings (provided enough spare parts are in stock);

9. Planned orders for spare parts is scheduled based on the inventory control policy.

4.5.3. Input parameters, output analysis, model scenarios and optimisation

The simulation models are non-terminating. Parameter values are inputted, including of course the model parameters *warm-up* period, number of *replications* and simulation *run time*. *Macros* are set up to enable the instant changing of decision variables. The output report includes various data and graphs, most importantly the average expected cost and downtime, per unit time.

Before analysis can begin, the *warm-up* period, the number of *replications*, and the simulation *run length* had to be determined to ensure the validity of the simulation results and that the quality of the output achieves the normal 95% confidence. The Time Series method (see Appendix 4.3(a)) based on the weekly cost mean value and Welch's method (see Appendix 4.3(b)) based on the weekly cost moving average with a window length of 5 (see, for example, Robinson, 2004; and Banks, 2010) were used to determine the *warm-up* period. A conservative 300 weeks *warm-up* was used to provide a very high degree of confidence that the system indeed stabilised and reached the *steady state* beyond the *transient* period as shown in Appendix 4.3, plots (a) & (b). Further analysis showed that running the simulation with seven replications provides a consistent confidence interval of 95% as shown in Appendix 4.3(c). Appendix 4.4 (plots (a) and (b)), display a sample analysis for the determination of the number of *replications* required, specifically for the $(R, S, T = R)$ policy. The plots confirm the decision reached by analysing Appendix 4.3(c).

Finally, to ensure that the simulation is run long enough for convergence, various *run lengths* were explored, and a 5,000-week length was deemed appropriate to ensure a high degree of confidence in the simulation results. The computation time to run through all seven replications takes approximately 17 minutes. To find the optimal policy, the simulation models were integrated with *SimRunner* (see ProModel, 2010). This tool uses sophisticated search algorithms (Kim et al., 2012), running multiple combinations (where applicable) of the decision variables to find the unique one that is optimal.

4.6. Results analysis and discussion

4.6.1. Joint optimisation

Five principal joint inventory-maintenance policies were considered: three models with periodic review $(R, S, T = R)$, $(R, s, S, T = R)$ and $(R, s, Q, T = R)$ and two with continuous review (s, S, T) and (s, Q, T). In addition, three variant policies were modelled, namely: $(R, S, T = 2R)$; $(R, s, S, T = 2R)$, and $(R, s, Q, T = 2R)$. While the presentation in this chapter is restricted to the cases $T = kR$ with $k = 1,2$, many other values for k were investigated but found to be cost-sub-optimal for the range of parameter values used here. Figure 4.3 and Table 4.2 illustrate that among all joint policies modelled, the $(R, S, T = 2R)$ policy has the lowest total cost per unit time (cost-rate), maintaining and inspecting the bearings in the plant every 10 weeks and ordering spares every 5 weeks (see Appendix 4.5 for a depiction of the simulation layout for one of the models). Note that the $(R, s, S, T = 2R)$ policy is equivalent to the cost-minimal policy, because $S^* - s^* = 1$. (Here, S^* and s^* are the optimum values of S and s respectively.) Globally, the second and third best policies (lowest cost-rate) are also the $(R, S, T = 2R)$ policy, inspecting every 11 weeks and ordering every 5.5 weeks, and the (s, S, T) policy inspecting the machinery every 11 weeks. Clearly, under the cost-optimal policy, more frequent inspections are performed (every 10 compared to every 11 weeks), which will potentially identify more defects (if any) in the system and trigger their replacements, thus reducing failures and ultimately reducing cost. This trade-off between the marginal decreased cost of additional failures and the marginal increased cost of additional inspection lies at the heart of the maintenance decision problem.

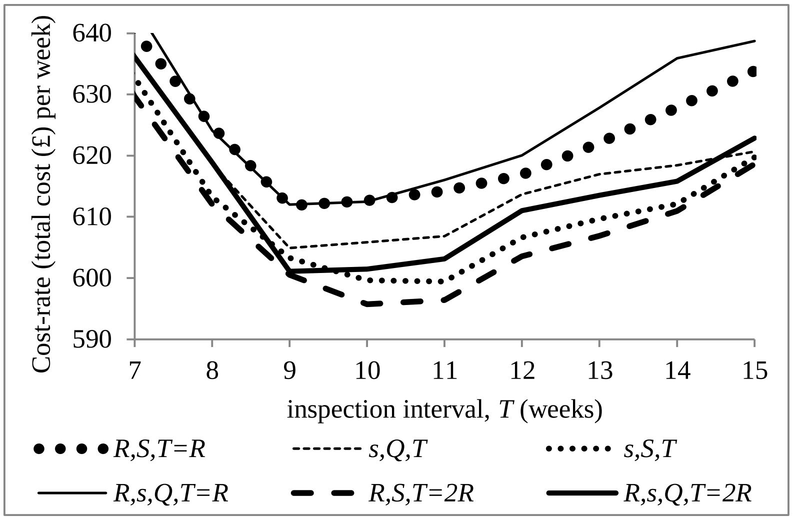


Figure 4.3. The effect of different inventory replenishment policies on the joint optimisation cost.

Table 4.2. Comparison of the joint optimisation cost, based on parameter values for different policies. Shaded results are cost-optimal for each policy. **Bold, shaded** is overall the policy with lowest cost.

	(R,S,T=R)		(R,s,S,T=R)			(R,s,Q,T=R)			(s,Q,T)			(s,S,T)			(R,S,T=2R)		(R,s,S,T=2R)			(R,s,Q,T=2R)		
T	cost-rate	S	*cost*-rate	s	S	cost-rate	s	Q	cost-rate	s	Q	cost-rate	s	S	cost-rate	S	cost-rate	s	S	cost-rate	s	Q
5	698.07	3	698.07	2	3	704.70	2	2	693.30	1	1	689.47	1	2	704.72	3	704.72	2	3	709.52	2	2
6	660.35	3	660.35	2	3	664.08	2	2	654.05	1	2	652.92	1	2	651.64	2	653.44	2	3	654.22	1	2
7	640.27	3	640.27	2	3	644.57	2	2	636.19	1	2	632.76	1	3	629.64	3	629.64	2	3	636.09	1	2
8	624.79	3	624.79	2	3	624.12	2	2	618.60	1	2	613.26	1	3	612.04	3	612.04	2	3	618.88	1	2
9	611.81	3	611.81	2	3	612.00	2	2	604.92	1	2	603.27	1	3	600.53	3	600.53	2	3	601.08	1	2
10	612.64	4	612.64	2	3	612.48	2	2	605.86	2	2	599.65	2	3	**595.69**	**3**	**595.69**	**2**	**3**	601.45	2	2
11	614.21	4	614.21	2	3	616.04	2	2	606.83	2	2	599.41	2	3	596.38	3	596.38	2	3	603.15	2	2
12	616.87	4	616.87	2	3	620.06	2	2	613.68	2	2	606.60	2	3	603.55	3	603.55	2	3	611.00	2	2
13	622.11	4	622.11	2	3	627.82	2	2	616.96	2	2	609.64	2	3	606.88	3	606.88	2	3	613.53	2	2
14	627.84	4	627.84	2	3	635.90	2	2	618.43	2	2	612.12	2	3	610.97	3	610.97	2	3	615.81	2	2
15	633.85	4	633.85	2	3	638.71	2	2	620.66	1	2	619.71	2	3	618.62	3	618.62	2	3	622.85	2	2

Similarly in the inventory decision problem, the cost of stock-outs is traded-off with the cost of inventory (as is shown in Figures 4.4(c) and 4.4(e), which will be discussed in more detail in Section 4.6.2). Thus, more orders might be placed or more stock might be held to reduce the possibility of stock-outs, depending on the relative sizes of the order cost and the holding cost. In the joint optimisation problem, where the inspection period is an integer multiple of the order period, more frequent ordering can by implication potentially reduce the frequency of bearing failures, and this appears to be the case here.

4.6.2. Insights into the characteristics of different replenishment policies

To obtain insights into the replenishment characteristics, the simulation results for each policy shown in Table 4.2 were analysed and a selection of these policy types are illustrated in Figure 4.4, for which the data was collected over a simulation of 5,000 weeks. The optimal inspection interval for the different policies ranges from 9 to 11 weeks. Considering the ordering cost per unit time, Figure 4.4(a) demonstrates that in general the (s, S, T) policy has the highest ordering cost-rate since it can potentially place more orders at both inspections and at failures. This is also partly due to $s = 2$ and $S = 3$, thus always triggering an order when the stock level drops by one unit. This is the $(S - 1, S)$ policy often used in a maintenance context, which becomes a special case of the (s, S) policies. This conclusion is supported by Figure 4.4(f) since the mean number of spares ordered per order is lowest for the (s, S, T) policy. Note, while the (s, Q) and (s, S) inventory policies are equivalent for unit sized transactions (whence $Q = S - s$), it can be seen in Table 4.2, for example, that the cost-rate for $(s = 1, S = 2, T)$ is less than the cost-rate for $(s = 1, Q = 1, T)$ for the same T. This is because in this model, while failure demands are always unit sized (the probability of two or more failures occurring together is zero), demand at an inspection may be greater than unit-sized (when more than one bearing is found to be defective).

Moving to a discussion of mean spares per order, furthermore, at the optimal interval, the $(R, s, Q, T = R)$ policy orders two spares per order every time (since $Q = 2$), compared to the $(R, s, S, T = R)$ $(R, S, T = R)$ policy which orders on average a little over one spare per order. So the former policy must have a lower order cost per unit time due to its placing fewer orders. It is also expected that the $(R, S, T = 2R)$ policy to have a higher order cost per unit time than the $(R, S, T = R)$ policy (Figure 4.4(a)) since the former can potentially place orders twice as frequently as the latter. Due to the nature of the continuous review policies, one would expect that they would generate at least as many opportunities as periodic review policies for placing orders. This is demonstrated in Figure 4.4(b), except for one policy. The number of opportunities for placing orders for the (s, S, T) policy is lower than the $(R, s, Q, T = R)$ and $(R, S, T = R)$ policies only because its optimal inspection interval is longer - 11 weeks compared to 9 weeks for the latter two policies. Figure 4.4(d) illustrates that the number of spares replaced at PM intervals was lowest for the (s, S, T) policy and highest for both the (s, Q, T) and the $(R, s, Q, T = R)$ policies. There is understandably a direct association between the number of spares replaced and the number of PM instances carried out.

Similar observations to those made about order cost-rates can be made in relation to holding cost-rates and stock-out cost-rates (Figures 4.4(c) and 4.4(e) respectively). A more interesting observation is that inventory costs seem to be traded off. Thus considering Figures 4.4(a), (c) and (e), it can be seen that where the holding cost-rate is high, the stock-out cost-rate is low, and vice versa, except for policy $(R, s, Q, T = R)$, where they are both high but compensated by the low order cost-rate for this policy. Where policy $(R, S, T = 2R)$ does appear to outperform the other policies is in the additional opportunities it offers for replenishment, even if it does not necessarily use them (Figure 4.4(b)).

Figure 4.4(d) confirms that the usage rate of spare parts is similar for all policies, as expected, since in the long run (at steady state) the consumption of parts is most influenced by the rate of arrival of defects. This implies the rather obvious but important observation: if one wants to reduce inventory costs, then first and foremost, one should use more reliable (better quality) parts. Thus, quality of spare parts is another factor than impinges on both maintenance and inventory, a point made in Scarf and Cavalcante (2012).

Figure 4.4(d) also shows variation between the policies in the number of spares used at inspections. Thus, those policies with more positive inspections per unit time (an inspection is deemed to be positive if at least one defect is found) use more spares at inspections (and the mean number of spares used per inspection is approximately 1.2). This variation is balanced by the variation in failure: while (s, S, T) and $(R, S, T = 2R)$ have the lowest positive inspection rates, their failure rates are the highest. A higher cost of failure might then lead to a different policy ranking. Also, if one were remanufacturing spare parts, one might prefer a policy with fewer failures per unit time.

This then brings the discussion to failures and stock-outs and policy risks. The (s, S, T) policy might be perceived as a low risk policy since it has the lowest stock-out cost. However, it has the largest failure rate. The optimal (lowest cost-rate) policy, $(R, S, T = 2R)$, has a very low stock-out rate and a moderate failure rate, although its stock-out rate is much lower than inferior policies (70% lower) and its failure rate is marginally higher (23% higher). The variation in failure rates is due almost entirely to the variation in the inspection interval; a longer inspection interval will result in more failures. The (s, S, T), the $(R, S, T = 2R)$, and the other three policies in Figure 4.4(g) with optimal inspection intervals of 11, 10, and 9 weeks have failure rates of 1.14, 0.9 and 0.73 per 100 weeks respectively. The other much smaller contribution will come from stock-out rate variation, where a policy with more stock-outs will ceteris paribus have fewer failures because of the plant stoppages due to stock-outs.

Now two final points are made. The first is that the variation in stock-out cost-rates across the policies is relatively large (Figure 4.4(e)). The $(R, S, T = R)$ policy has a stock-out cost that is four times that of the optimal policy, $(R, S, T = 2R)$. This would be expected since the latter has the potential to place twice as many orders. The second point is that generally the stock-out cost-rates are much lower than the failure cost-rates (Figure 4.4(e) vs Figure 4.4(h)).

Thus, in summary, first and foremost it is the failure rate (and equivalently defect arrival rate) that has the greatest influence on the choice of policy, followed by the emergency order (stock-out) cost. Furthermore, although the cost-rates for the jointly optimised policies are quite similar across the range of policies, the optimal values of decision variables for each policy can be quite different, so that the components of the cost-rate can be different. Thus, the different policies, at their optimal settings, place different demands on inventory.

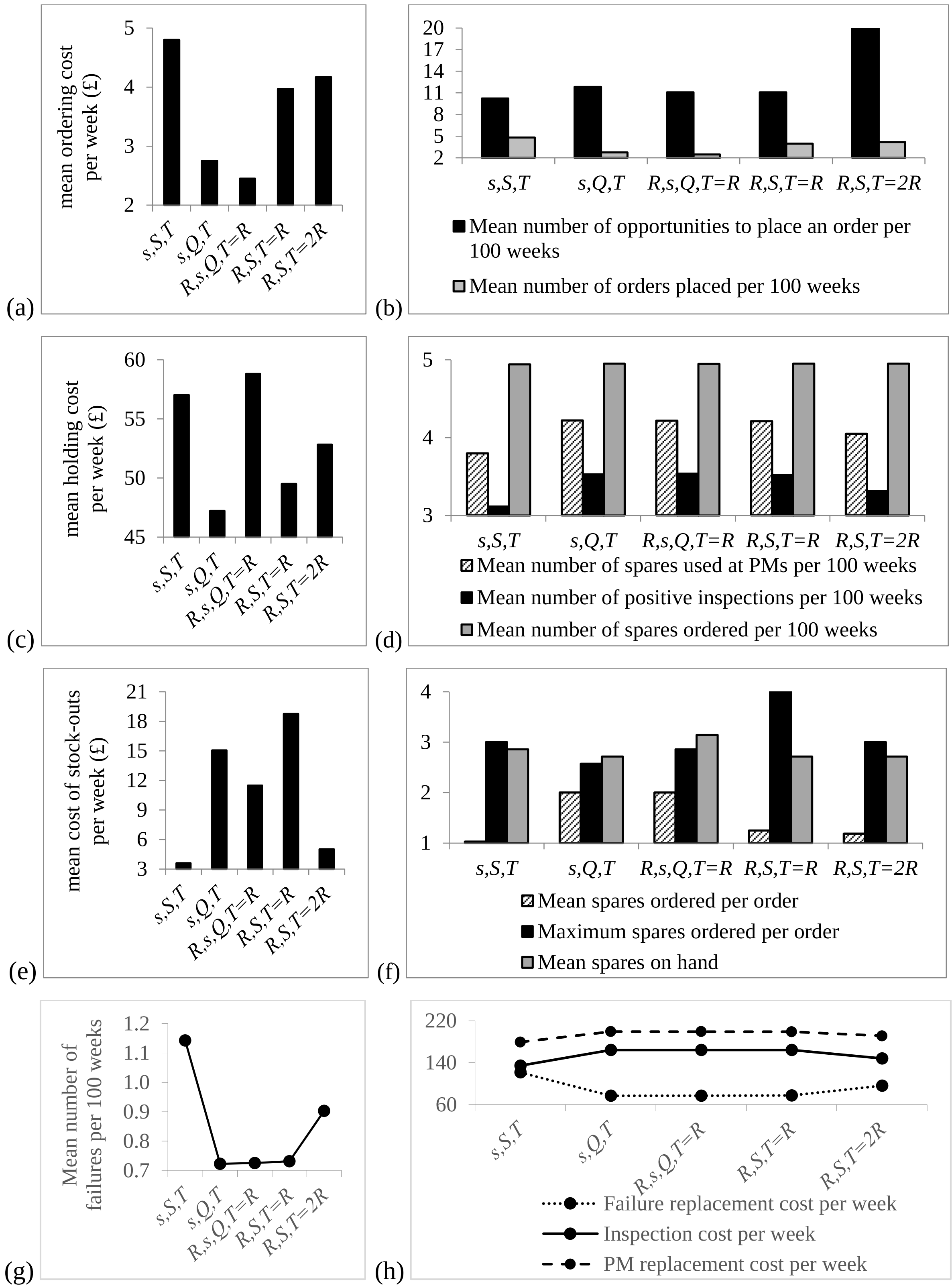


Figure 4.4. For the optimum policy in each class of inventory policies:
(a) order cost-rate; (b) order point statistics; (c) holding cost-rate; (d) PM statistics;
(e) stock-out cost-rate; (f) order size statistics; (g) failure rate; and (h) maintenance cost-rates.

4.6.3. Sensitivity analysis

Sensitivity analysis of various parameters for the optimal policy $(R, S, T = 2R)$ are shown in Figure 4.5. The behaviour of the cost-rate with respect to the defect arrival intensity is as expected here since the cost-rate of the optimal policy for 0.5λ and 2λ are at 63% and 166% of the baseline respectively (Figure 4.5(a)). The optimal times between inspections behave as expected, that is, as the intensity of defect arrivals increases, more frequent inspection is expected to yield the optimal interval and vice versa. However, a reduction of 50% in the defect arrivals has a more sustained impact on the optimal interval. Similarly, as the scale parameter of the Weibull *delay-time* distribution α is reduced (Figure 4.5(b)), the cost-rate rises as expected since the mean *delay-time* decreases and defects develop into failures more quickly. The cost-rate of the optimal inspection interval for 0.5α and 2α are 116% and 89% of the baseline, respectively. However, in extending the *delay-time*, minimal effect is displayed when inspection is very frequent. Figure 4.5(c) displays the cost-rate of the optimal inspection interval for $0.5\mathrm{C}_d$ and $2\mathrm{C}_d$ at 78% and 141% of the baseline, respectively. The optimal times between inspections also behave as expected, but with the greatest impact for the 100% increase in the cost of failure and when inspection is infrequent. Overall, the greatest impact is evident when inspection is less frequent. And finally, the cost-rate of the optimal inspection interval for $0.5\mathrm{C}_i$ and $2\mathrm{C}_i$ are at 91% and 115% of the baseline respectively (Figure 4.5(d)). Varying C_i has the greatest effect when inspection is frequent and the optimal intervals move in the expected direction.

The cost-rates for different unit costs, $0.5\mathrm{C}_u$ and $2\mathrm{C}_u$ are 87% and 124% of the baseline, respectively. The reduction in the unit cost does not seem to have any effect on the frequency of inspection (Figure 4.5(e)). Figures 4.5(f) and 4.5(g) suggest that halving or doubling the costs of ordering or shortage shipment have minimal effect on the overall cost-rates. In fact, for $0.5\mathrm{C}_{sh}$ and $2\mathrm{C}_{sh}$ this is true for every inspection interval. Further, the sensitivity to the order lead-time (Figure 4.5(h)) suggests that a change in the order lead-time has the smallest effect when the lead-time is halved compared to the baseline, but has the greatest effect when it is doubled. However, in the latter case, the effect becomes negligible when inspection is infrequent. This is because when the inspection interval is very large, it matters little if the lead-time is large.

Finally, the sensitivity of the order-up-to-level S was investigated for the decision variables $T = 10$ and $R = 5$. The original cost for this scenario was £595.69 per week. When S was reduced from 3 (the cost optimal quantity) to 2, the weekly cost was increased, as expected, but only by 3.0%. Similarly, when S was increased from 3 to 4, the weekly cost was increased, again as expected, but only by 2.6%.

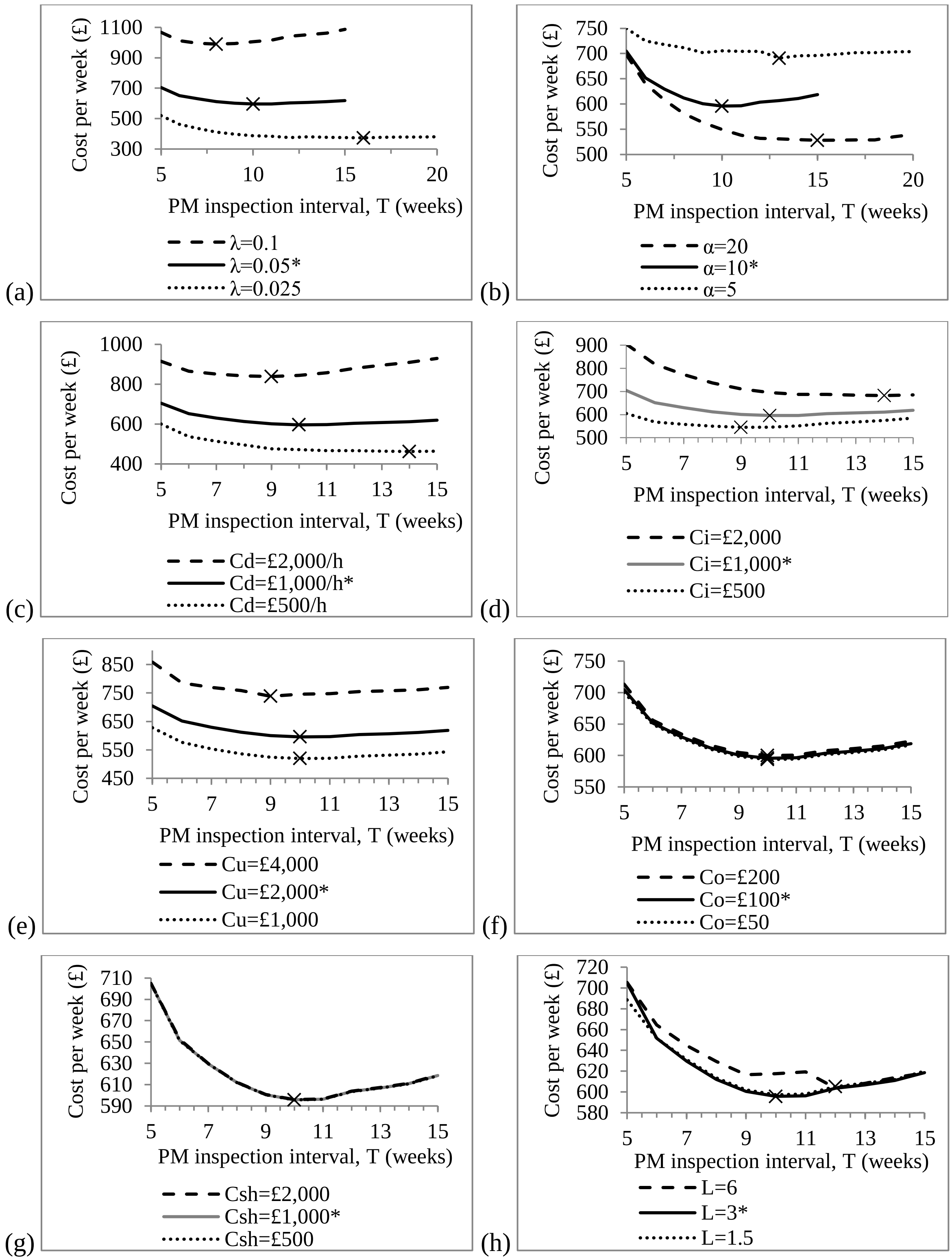


Figure 4.5. The sensitivity on the joint optimisation cost for the optimal $(R, S, T = 2R)$ policy for various parameters (x=minimum; *=baseline):
(a) defect arrival rate; (b) failure *delay-time*; (c) cost-rate of machine downtime;(d) inspection cost; (e) unit cost; (f) ordering cost; (g) shortage shipment cost; and (h) order lead-time.

4.7. Conclusions and further work

Several simulation models were developed for a complex system with multiple identical bearings, in the context of a paper machinery plant. The planned maintenance inspection interval T, based on the delay-time concept, and the spare parts inventory policy, considering various policies, have been jointly optimised. This is the first study that compares a number of periodic and continuous review replenishment policies, and analyses their efficacy when joined to the inspection policy. The objective was not only to find the cost-minimal policy across the range of policies, but also to illustrate the characteristics of each policy, so that engineers might be guided about the suitability of these policies across a range of criteria that may be appropriate for particular industrial contexts. The particular context that motivated this study is the maintenance of bearings in a paper mill.

Joint models require complex mathematical formulations and it may not be possible to solve these analytically except for limited situations and/or with simplifying assumptions. Simulation has been used, which compared to mathematical modelling, has the advantage to describe multivariate non-linear relations. However, since simulation is not an optimization technique, so as to find the global optimal policy, simulation was integrated with an optimization tool.

For the policies considered here, the following conclusions are made:

- $(R, S, T = 2R)$ and $(R, s, S, T = 2R)$ are the minimum cost policies.

- The additional cost of more frequent ordering, and hence inspection, under the $(R, S, T = 2R)$ policy is compensated by the reduction of bearing failures.

- At the optimal interval, the $(R, s, Q, T = R)$ policy results in ordering more spares every time, compared to the $(R, S, T = R)$ policy, resulting in a lower ordering cost-rate but a higher holding cost-rate.

- The $(R, S, T = 2R)$ optimal policy is a relatively low risk one as it is associated with one of the lowest stock-out cost-rate.

- The $(R, S, T = R)$ policy is associated with the highest stock-out cost-rate at its optimal settings.

- The $(R, s, Q, T = R)$ policy has similar overall cost-rate to the $(R, S, T = R)$ policy, suggesting that the holding cost is not necessarily a decisive variable.

- The sensitivity analysis to different parameters for the optimum policy, $(R, S, T = 2R)$, gives results that are expected. Varying C_d and C_i has the greatest impact when inspection is infrequent and frequent, respectively. The optimal policy at its optimum settings is not sensitive to the order-up-to-level S.

- The defect arrival rate of the unit (part) in service is the principal determinant of policy, followed by the emergency order (stock-out) cost.

- Finally, while the cost-rates are similar across the range of policies, the components of the cost-rates are quite different because the policies' decision variables are different, and so the different policies, at their optimal settings, place different demands on inventory.

The models that have been developed can be extended in several ways to model more realistic industrial situations, including: imperfect inspection whereby false positive and false negative inspections are present (Berrade et al., 2012); postponed replacement of defective components; variable replenishment lead-times; modelling of dependent and/or non-identical multi-unit systems; the formulation of a cost-effective spare parts ordering policy based on historic data and *dynamic* forecasting to predict spare parts demand; and multi-line parallel non-identical systems.

Chapter 5

Joint modelling and simultaneous optimisation of preventive maintenance and spare parts inventory for multi-line production systems

5.1. Summary

In parallel production settings where interactions and dependencies exist between system components, maintenance policies are, for the most part, analytically intractable. In this study, a simulation language and a numerical optimisation tool are used to determine the cost-optimal joint inspection and replenishment policy for maintaining machines in a specific multi-line production system. This optimisation is performed in order to eliminate, or at least minimise, the occurrence of simultaneous machine downtime in a system with parallel machines. The occurrence of simultaneous downtime may halt production, which will have a significant adverse effect on profitability or other performance measures. An industrial setting provides the idealised context for modelling the plant maintenance. The demand and replenishment of spare parts is considered with several variants of a periodic review policy. A number of simulation models are developed for the joint optimisation of maintenance-inventory. The results indicate that among several joint policies considered, the policy that uses the same frequency of maintenance and replenishment is cost-optimal when *just-in-time* (JIT) ordering is such that the delivery of parts coincides with maintenance interventions. A sensitivity analysis offers insights to practitioners for the management of their multi-line systems.

5.2. Introduction

In the research and practitioners' literature, maintenance is being increasingly highlighted as an integral part of production (see, for example, Alsyouf, 2009; and Wall, 2013). Generally, organisations have become increasingly aware that the maintenance function is an integrated part of their business (Ding and Kamaruddin, 2015). There is therefore a great deal of financial interest in optimising maintenance operations and thus reducing the effect of plant downtime. Although there is significant cost associated with planned maintenance, Alsyouf et al. (2016) show that good maintenance planning can reduce overall maintenance costs.

Extensive research is evident in the literature addressing the problems of maintenance and inventory control separately or sequentially (for example, De Almeida, 2001; Marseguerra et al., 2005; and Cheng and Tsao, 2010). Typically, in the maintenance literature, the optimisation issue has been tackled by determining the optimal inspection interval which yields the minimum cost, assuming infinite availability of spare parts (for example, Sharma and Yadava, 2011). For single-line systems, early models include those due to, for example, Barlow and Proschan (1965). Others have integrated production quality into the inspection problem (e.g. Lu et al., 2016) and considered preventive maintenance planning in job shop scheduling (e.g. Thörnblad et al., 2015). These studies assume that spare parts are readily available, which implies that parts are either highly standardized that can be readily bought from a supplier, or are so inexpensive that large quantities can be stored. However, parts are usually highly customized and their procurement lead-time cannot be neglected (Panagiotidou, 2014). Therefore, maintenance analysis without spare parts consideration will result in misleading decisions. To carry out maintenance effectively, spare parts need to be available immediately, in order to replace both failing items in service and faulty parts at inspections. The operational effectiveness of the inspection process is also dependent upon the availability of spare parts. Clearly, maintaining a sufficient amount of spare part inventory is indeed the challenge faced by plant managers in order to minimise the holding cost and the risk and cost of stock-outs. Note that the term component, item, or part may be used interchangeably to refer to the critical component that needs to be replenished as the spare part.

Many review papers address the optimisation of preventive maintenance, for example, Ding and Kamaruddin (2015). Maintenance models are broadly developed for block-replacement of plant items or replacements based entirely upon some pre-specified item age, most of which are concerned with one-unit systems. In many industrial situations, the replacement of multiple plant components is either too expensive, which would make the block replacement policy very costly, or the parts are too critical for taking the risk of replacements using an age-based policy. Alternatively, parts may be inspected at optimal intervals and replaced only if found to be defective (faulty), which seems to be a reasonable and rational maintenance strategy. One of these inspection methodologies is delay-time modelling (DTM), which was first introduced by Christer (1976), and developed further by many including Flage (2014). DTM, which is also used in this study, has been extensively utilised in numerous case studies, for example, Emovon et al. (2016). The concept has the advantage that it explicitly models the relationship between plant failures and the inspection interval. Wang (2012a) gives the latest review of the delay-time advances including industrial applications.

Studies show that almost all maintenance models relate to single-line production facilities and little research is directed towards the realistic scenario of optimising maintenance for a system composed of several production lines (Van Horenbeek et al., 2013). Furthermore, most if not all of these models are analytical, which are generally complex and, for multi-line production systems, they are intractable for determining the optimal maintenance inspection interval. Simulation is well suited and has the flexibility to address the increasingly complex and dynamic nature of maintenance optimisation problem. In the latest literature survey by Alrabghi and Tiwari (2015), reviewing 59 journal papers since the year 2000, the authors report on the state-of-the-art simulation-based maintenance optimisation. They observe that discrete-event simulation (DES) is the most reported technique for modelling maintenance systems. This study uses discrete-event simulation as an alternative approach to model the operations of a plant comprising two parallel machines. This is in order to eliminate, or at least to minimise, the occurrence of simultaneous machine downtime, which may halt production and thus cause a significant adverse effect on profitability or other performance measures. Apart from this chapter and Chapter 3, the only other study that considers a parallel system of multiple machines is Boschian et al. (2009), which discusses the complexity of analytical modelling and also uses simulation. However, unlike this current study, Boschian et al. (2009) do not consider joint optimisation.

Van Horenbeek et al. (2013) carried out a comprehensive review of literature on joint maintenance and inventory optimization for non-repairable units. Among the literature, Sarker and Haque (2000) used simulation since they considered the development of analytical models to be “extremely difficult”. They showed that the jointly optimized policy was superior to the combination of separately or sequentially optimized policies. In comparison, Chelbi and Ait-Kadi (2001) simultaneously optimised the block replacement interval, the optimal inventory level, and the replenishment cycle. Yoo et al. (2001) developed an analytical model for a system of N identical units. Brezavscek and Hudoklin (2003) formulated a stochastic mathematical model and found that it was relatively insensitive to moderate changes of the parameter values, but they also showed a 97.4% increase in the value of the objective function (the expected total cost of system maintenance per unit time) when the order-up-to-level S was decreased by 10%. Using a different approach, Vaughan (2005) treated the time between inspection operations as fixed (not optimised) and used a stochastic dynamic programming model to develop a policy for ordering parts due to both sources of demand, rather than addressing them separately. The maintenance model in Vaughan (2005) best resembles delay-time modelling since the n units in service are inspected at scheduled maintenance intervals T, during which some or all are replaced, if needed. Later, Ilgin and Tunali (2007) developed a simulation optimization model, which they believed gives the ability to describe multivariate non-linear relations that are difficult to express in an analytical form. They concluded that a reduction in total annual maintenance cost and an improvement in average monthly production were achieved. Huang et al. (2008) developed a mathematical model and generalised the study by Brezavscek and Hudoklin (2003) in their joint optimisation with random lead-time.

In a different approach, De Smidt-Destombes et al. (2009) developed heuristics for the joint optimisation of preventive maintenance frequency, spare parts inventory levels and spare parts repair capacity for a single k-out-of-N system under block replacement. Apart from this current study, Wang (2011a, 2012b) is the only other author who has used delay-time modelling (DTM) in joint optimisation studies, but only for a single-line production system. The author assumes order lead-time negligible “for model simplicity” which is “used extensively in inventory literature”. Chen et al. (2013) developed an analytical model in which the procurement lead-time is assumed to be constant but state that it is necessary to explore the policy with random lead-time as an extension. A year later, Panagiotidou (2014) studied the joint maintenance and spare parts ordering problem for both (R, S) and (s, S) replenishment policies. The inspection model of this author has some similarities with delay-time modelling in that minor failures (which could be considered as defective components) are only detectable through inspection. However, it differs from DTM in that random failures are not self-announcing and may only be identified at inspections. Jiang et al. (2015) developed an analytical model based on the same assumptions as Brezavscek and Hudoklin (2003) but taking into account the cost of spare part deterioration. Finally, Samal and Pratihar (2015) used particle swarm algorithms in their study of electric overhead travelling cranes, which they claim gives a 'better global solution' compared to other optimisation methods. The authors extended the maintenance interval from 1 to 1.5 years. However, they assumed that spare parts are replenished instantaneously, which may prove unrealistic in practice.

Zahedi-Hosseini et al. (2017) have classified the characteristics of joint maintenance and spare parts inventory control models in the literature using several categories of: (i) maintenance policy (age-based or block-based); (ii) replenishment policy (periodic or continuous review); (iii) model development (analytical or simulation); and (iv) components in system (single or multiple) - (see Section 4.3 including Table 4.1 in this thesis). Compared to Van Horenbeek et al. (2013) review, a number of new insights have thus emerged. First, few authors consider an age-replacement policy; most consider a periodic block replacement policy instead; and the use of inspection-based preventive maintenance is growing. Second, despite the complexity of joint optimisation models, many researchers continue to use analytical models with restrictive assumptions rather than making use of simulation. Finally, model development for multi-unit series systems has grown considerably. Maintenance models for multi-line parallel systems are very rare in the literature, except for studies that mainly integrate maintenance with production scheduling (for example, Wang and Liu, 2015). However, there are no studies addressing the integration of maintenance and spare parts inventory for such systems.

Thus, based on the detailed literature review undertaken here, this current study makes two significant contributions: (i) it is the first study to consider the joint optimisation of preventive maintenance and spare parts provisioning for multi-line production systems; and (ii) it provides insights into the characteristics of the *best* joint policies for multi-line production studies. Clearly, the two critical issues of maintenance and spare parts need to be jointly addressed if models are to be realistically implemented in practice.

In this study, simulation models are therefore developed for the joint modelling and simultaneous optimisation of an inspection interval T in a delay-time model, and spare parts provision using a (R, S) periodic review replenishment policy. Thus, the decision variables T, R, S are optimised simultaneously.

The chapter is organised as follows. Section 5.3 discusses the methodology, assumptions, and cost factors for modelling a complex system with multiple identical components (bearings) in the context of a paper making plant comprising parallel machines. In Section 5.4, the details of the simulation models are described. The results are analysed and discussed in Section 5.5, including a sensitivity analysis of the parameters affecting the cost-optimal policy. In the final section, conclusions are drawn and proposals are developed for the future direction of this research.

5.3. Modelling methodology

5.3.1. Notation

See Section 2.5.3., for a list of Notation associated with this chapter.

5.3.2. Problem description

The specific industrial plant situation and the idealised context considered here is a paper mill consisting of two machines working in parallel. Beside the relatively low-cost cutting blades, expensive bearings are the *critical* components in this plant.

Bearings are used extensively in paper making machines and, apart from general risks to safety, their failure can incur costs due to repair or replacement, and unplanned machine downtime. Folger et al. (2014a, 2014b) describe several conditions under which bearings can fail unexpectedly and catastrophically including: improper handling and installation; contamination; inadequate lubrication; and various types of overload. Bearing life, often referred to as the *L10 life*, is a method of specifying a bearing's useful life before it shows the first signs of fatigue, before failure. The term *L10* is used to denote the life that 90 percent of seemingly identical bearings, operating under identical conditions, can operate before fatigue occurs, or 'defects arrive'. Therefore, there is a 90 percent reliability that the bearing will achieve the specified life. The calculation of *L10 life* for ball bearings as the load-carrying elements, is given by:

$$L = \left[\frac{C}{F}\right]^{n}$$

(Collins, 2017), where, L = basic rating life 10^6 (revolutions); C = bearing dynamic load capacity (N); F = applied dynamic load (N); and n = 3 for ball bearings & n = 10/3 for roller bearings (also see, for example, Jacobs et al., 2016).

In this study, several simulation models are developed for the joint optimisation of preventive maintenance and inventory replenishment for the replacement of bearings in the plant. The models are informed by a survey conducted by the author (described in Section 4.4.2, second paragraph onwards), collecting information from maintenance/inventory control experts and paper manufacturers about their experience of paper making machinery and their critical components. A questionnaire was used to obtain information about: possible defect arrival patterns, *delay-times*, and their distributions; inspections; preventive maintenance replacements; failure replacements; current maintenance and replenishment policies for replacing critical components; lead-times; and finally costs (see Appendix 4.1(a) for a copy of the survey questionnaire and details of the questions, and Appendix 4.1(b) for a summary of the responses). Sixty per cent (nine out of 15) of the questionnaires were returned: three from experienced maintenance and inventory control researchers; and six from paper machine manufacturers. The information obtained from the survey were the basis for the costs and parameter values used in the models, which ensured that the models and simulation experiments were realistic and were not based on some arbitrary data. The parameters' value ranges are indicated in the relevant Sections.

In the analysis process, the data from the six paper machine manufacturers were generally consistent since they were all referring to the same type of plant. However, the data from the maintenance and inventory control experts varied, depending on the projects they had previously undertaken, and not necessarily anything to do with paper rolling plant. There were two areas, in which the data seemed to vary considerably. The first is the inspection duration. However, as will be discussed in more detail in Section 5.3.3, the assumption in this study is that the inspection activity (or the data analysis conducted by specialists - original equipment manufacturer (OEM)) to identify defects at inspection intervals has zero downtime because it takes place off-line (Maintenance & Engineering, 2017). Therefore, the variability of the inspection duration in the survey data was not an issue. The second area where the data showed high variability was the failure distribution and its parameter values. In this case, it was decided to use the same data used in Wang (2012b), which also considered the maintenance optimisation for a paper making plant.

5.3.3. The preventive maintenance model and its assumptions

In the idealised context considered here, each machine has n identical bearings that are subject to deterioration. The two-machine parallel system is assumed to be operating under steady-state conditions. In this (complex) system model, multiple concurrent defects are possible and the failure process of a bearing is based on the two-stage delay-time concept. During the first stage (the *time-to-defect* arrival, u), a bearing is good and working normally until it becomes defective. Then in the second stage (the *delay-time*, h), a defective bearing deteriorates progressively and fails eventually after h time units. If inspection is carried out during this second stage, it is assumed that *all* defective items are identified and replaced, as depicted in Figure 5.1. This is very similar to the original Barlow and Proschan (1965) block-based replacement policy at fixed intervals. However, the major difference is that under the delay-time modelling, only defective items identified are replaced, rather than replacing all items regardless of their age or conditions.

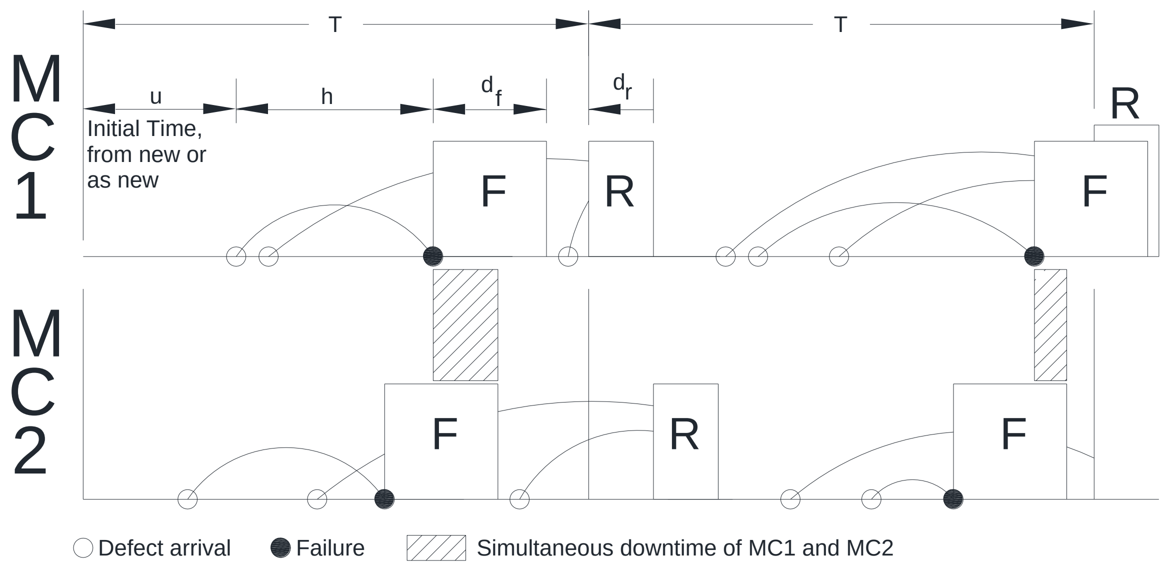


Figure 5.1. Defect arrivals, failures, failure stoppage duration F, and replacement stoppage duration R in this complex system of multiple components for a two-machine (MC1 & MC2) parallel system.

For the identification of defective bearings, a third party (external specialists) are employed to monitor bearing condition and compile reports periodically, following processing of the raw condition data. Therefore, in this way, *inspection* is replaced by condition monitoring and the reporting of bearing condition, which incurs cost but no plant downtime since this arrangement does not interfere with the operation of the plant. Wang and Wang (2015) state that this is typical of condition monitoring arrangements in modern plant.

The maintenance policy that is considered *inspects* all items across both machines in parallel every T time units, and defective bearings identified are replaced preventively. This intervention has d_r downtime per item. Concurrent preventive replacement of items on both machines does not take place in order to minimise simultaneous downtime. In addition, preventive replacement is not carried out if the other machine is already down due to failure, until the failed machine becomes operational again.

On failure, failed bearings are replaced immediately in order to return the plant to operation as quickly as possible, since such events cause downtime and an immediate unscheduled stop to the plant production. But only failed bearings are replaced and the downtime takes d_f time units per item. Thus, there is no inspection of bearings at failure events.

It is assumed that the system is in a state of suspension while both preventive or failure replacement is carried out. Therefore, defects grow, bearings age, and defects and failures can arise only whilst the plant is operating. Any operational loss due to the presence of defects other than inspection, replacement and failure are ignored. These are standard assumptions in inspection models (Wang, 2011a).

The majority of parameter values are based on the survey and the rest from Wang (2012b). Times between defect arrivals are assumed to be independent, exponentially distributed, consistent with the delay-time model of a complex system (see for example, Wang, 2012b), with a defect arrival rate (intensity) of 0.125 per week. The number of identical bearings in a typical paper rolling machine is greater than 100 (Wang, 2011a). The *delay-time* follows the Weibull distribution with scale and shape parameters, $\alpha = 10$, $\beta = 3$ respectively (implying a mean *delay-time* of 8.93 weeks). The downtimes due to each replacement and failure are $d_r =$ 4 hours = 0.024 weeks (survey range 1-6 hours) and $d_f = 24$ hours = 0.143 weeks (survey range 1-36 hours).

5.3.4. The inventory control model and its assumptions

There are typically over 100 identical bearings in a paper making plant. In the model, only the inventory planning for a single stock keeping unit (bearings) for this particular plant is considered. The demand for the bearings is generated through failure of parts in service occurring between inspections, and the identification and preventively replacing all defective bearings at inspections. Any demand is satisfied from the existing inventory or by expediting an emergency order.

Using the simulation models developed, several variants of the (R, S) periodic review replenishment policy (see, for example, Muller, 2011 and Silver et al., 2016) are compared. The $(R, S, T = R)$ policy using *standard* ordering (only one of the policy variants considered in this study), is depicted in Figure 5.2, where every R time units (the review period) an order is placed to raise the inventory position to the order-up-to-level S. In Figure 5.2 illustration, orders are placed at points A, C, and E, for example, and arrive at points B, D, and F respectively, after a lead-time, L_o. It is important to note that in this illustration: (i) an arbitrary demand profile has been used; (ii) $L_o < R$ for simplicity, which is not a restriction in the model; and (iii) the preventive maintenance and ordering events coincide, which is not the case for all policy variants considered here. However, if the maintenance intervention and ordering events coincide, then the order quantity will take into account the replacement of defective bearings at those events. This will be further discussed in Section 5.5.

For all policy variants, the joint optimisation policy contains the inspection interval T, the review period R, and the order-up-to-level S. The preventive maintenance interval $T = kR$ for $k > 0$. Here, values of the decision variables that minimise the long-run expected cost per unit time or cost-rate, $C(T)$, are sought.

In these models, based on the survey data, $C_o = £100$ and fixed, including the cost of delivery. C_h is costed at 1% of item cost per week. $C_u = £1{,}000$ per item (survey range £1000-4000), and $C_{sh} = £1{,}000$ per emergency shipment (survey range £500-1200). The lead-time, L_o is set at 4 weeks (survey range 2-6 weeks) and finally the shortage emergency delivery lead-time, L_{sh}, is fixed at $7\ days = 1$ week (survey range 1-10 days). Further, it is assumed that orders are placed at the end of each order-placing day and arrive at the beginning of each order-receipt day but before reviewing the current inventory if it coincides with an order-placing day. These assumptions are for modelling purposes. The order interval is flexible since there might be times that based on the current inventory position the optimal cost will be achieved by ordering no spares.

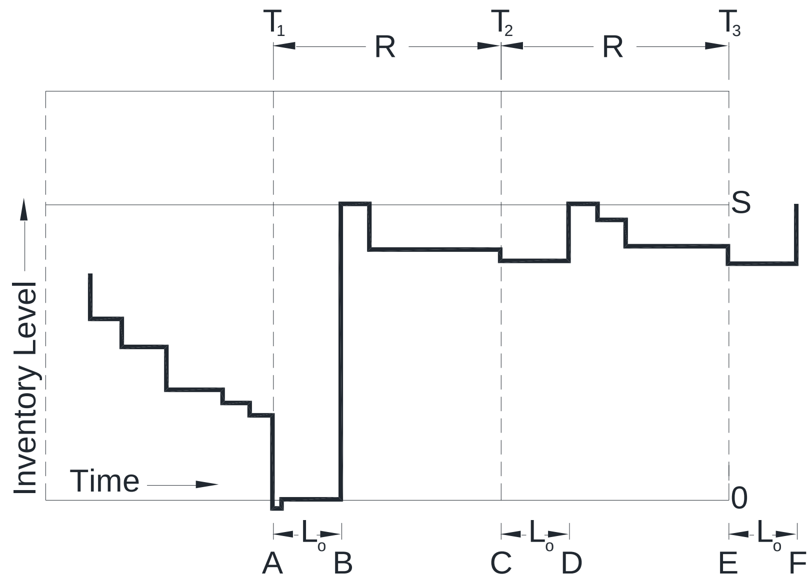


Figure 5.2. Characteristics of the (R, S) inventory control policy using *standard* ordering.

5.3.5. The order of events in the joint policies

All bearings are simultaneously inspected across both machines every T time units. In order to minimise simultaneous downtime, replacement of items on both machines does not take place concurrently. Preventive replacements are carried out in sequence, first machine 1 and then machine 2. In addition, replacement is not carried out if the other machine is already down due to a failure, until the failed machine becomes operational again. Orders are placed to raise the inventory position to the order-up-to-level S. If the maintenance intervention and ordering events coincide, then the order quantity will take into account the replacement of defective bearings at those events. Orders arrive at the beginning of each order-receipt day but before reviewing the current inventory if it coincides with an order-placing day.

5.3.6. Costs and downtime specifications

- Preventive maintenance replacement cost (C_r) and failure replacement cost (C_f) are based on 3 maintenance technicians at C_m = £60 per maintenance technician per hour.

- The individual machine downtime cost-rate, $C_{d(ind)}$ = £1,000 per hour.

- There are two PM inspection renewal costs depending on whether spare parts are immediately available:
  - If spares are available, then the cost includes downtime and manpower, hence: $C_r = 4\ C_{d(ind)} + 12\ C_m = £4{,}720$ since it takes 4 hours to replace each item.
  - If spare parts are not immediately available, there are two extra costs, namely, the shortage shipment cost and the downtime cost while the emergency shipment is in transit, hence: $C_r = 4\ C_{d(ind)} + 12\ C_m + C_{sh} + \ C_{d(ind)}\ L_{sh} = £173{,}720$.

- There are also two failure renewal costs depending on whether spare parts are immediately available:
  - If spare parts are available, then the cost includes the costs of downtime and manpower, hence: $C_f = 24\ C_{d(ind)} + 72\ C_m = £28{,}320$ since it takes 24 hours to replace each failed item.
  - If spare parts are not immediately available, then there are two extra costs, namely, the shortage shipment cost and the downtime cost while the emergency shipment is in transit, hence: $C_f = 24\ C_{d(ind)} + 72\ C_m + C_{sh} + \ C_{d(ind)}\ L_{sh} = £197{,}320$.

- The inspection cost C_i is fixed at £1,000 with one maintenance technician assisting the external consultants during an 8 hour shift, performing data analysis and reporting. Therefore: $C_i = 1{,}000 + 8\ C_m = £1{,}480$.

- Finally, the simultaneous machine downtime cost-rate, $C_{d(sim)}$, is £10,000 per hour.

5.4. Simulation modelling

Using a modular approach, simulation models were developed for the joint optimisation of the inspection maintenance and spare parts provisioning for a multi-line production facility. *ProModel*, a process-based discrete-event simulation language (see, for example, Harrell et al., 2011) was used to model a two-machine parallel system (but extendible to several machines or lines) as a continuous production system with the consideration of all cost figures and major assumptions given in Sections 5.3.1 to 5.3.6.

The construction of the overall simulation model framework, the modelling methodology, and finally the input parameters, output analysis, model scenarios and the optimisation technique are discussed in Sections 5.4.1 to 5.4.3.

5.4.1. Construction of the model framework and the minimum system requirements

The basic model framework using the *ProModel* programming environment requires, at least, the use of four *modules*: *L*ocations; *E*ntities; *A*rrivals; and *P*rocessing (*LEAP*). *Variables*, *attributes*, *subroutines*, *resources* and *path networks* are used extensively for further model sophistication. *Macros* are also set up to enable the easy alteration of decision variables.

5.4.2. Simulation modelling methodology

This section describes the approach taken in developing the simulation models and presents a few of the main flowcharts. However, a complete set of flowcharts will be found in the Appendices. Figure 5.3 (also shown in Appendix 5.1 for the completeness of information in the Appendices) illustrates the flow chart of the general simulation procedure, depicting the nine modelling routines developed for different aspects of the model. The model description and the algorithms defined in this section (and its sub-section 5.4.2.1) are specific to the cost-optimal $(R, S, T = R)$ policy, using *just-in-time* ordering (such that the delivery of parts coincides with maintenance interventions). Referring specifically to the numbers stated in the right hand side of each routing box in Figure 5.3, a brief description is given below for each, which applies to each machine, except for box 9 that applies to both machines:

1. Defects are scheduled to arrive with the appropriate arrival intensity based on the given statistical distribution;

2. Defects actually arrive and wait in their *delay-time*;

3. Machine downtime is scheduled, either due to defects transforming into failures (unplanned process), or defects identified at preventive maintenance instances (planned process), triggering the replacement of bearings (when needed);

4. Machine downtime occurs;

5. Defects may be *delayed* or *prolonged* during the *time-to-defect* and *delay-time*, if necessary, since defects and failures can only arise whilst the plant is operating, and defects do not grow and the bearings do not age during replacement downtime;

6. Planned maintenance is scheduled;

7. Defects (if any) are identified so that bearings may be replaced;

8. Defects (if any) are removed by the replacement of bearings (provided enough spares are in stock);

9. Planned orders for spare parts are scheduled for both machines at regular intervals.

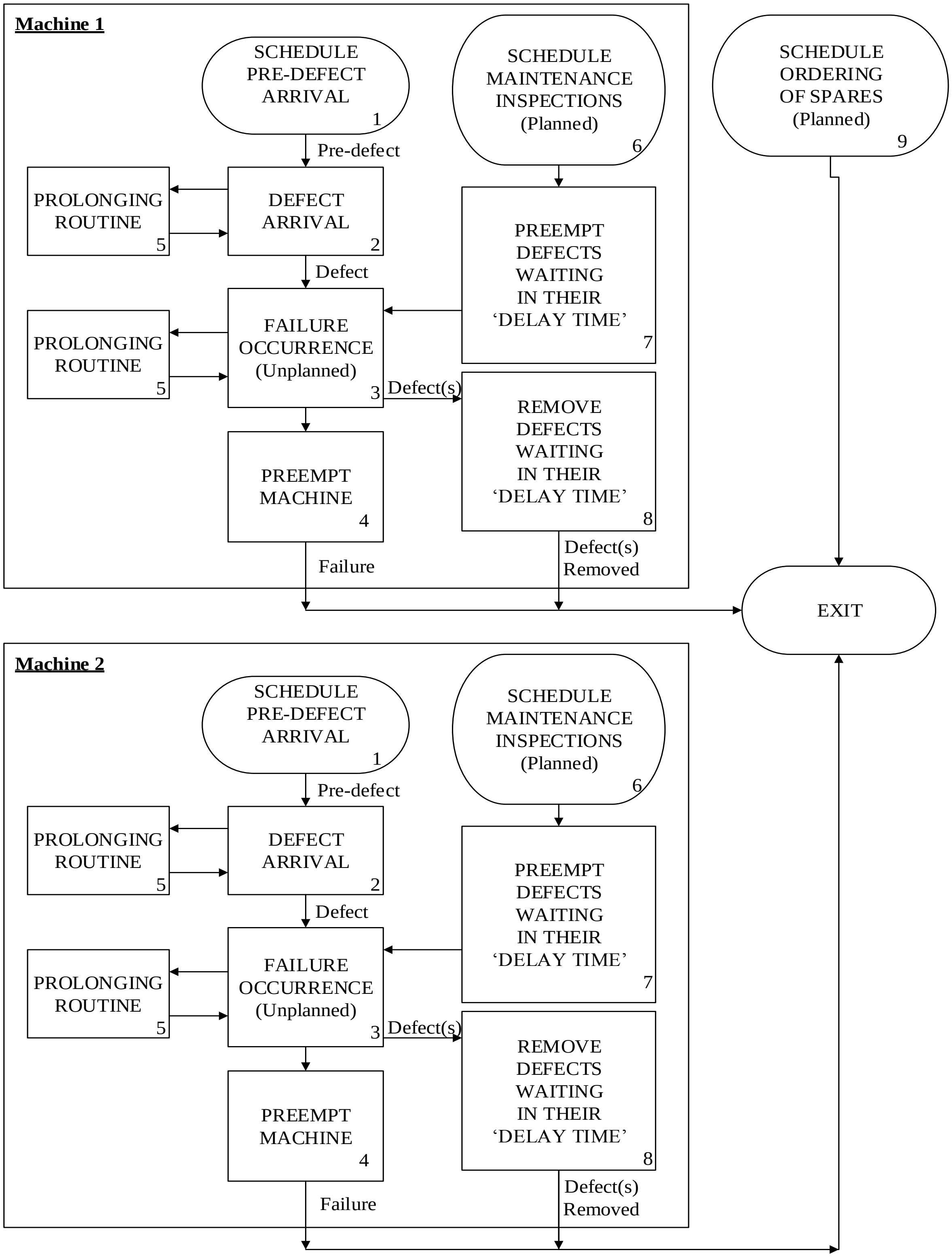


Figure 5.3. Flowchart of the general simulation procedure, showing the flow of entities from one modelling routine to another.

5.4.2.1. ProModel 'build' modules

The following modules and programming logic were specifically developed for the model:

(a) Locations : machines, queues, buffer storage areas, etc. Appendix 5.2(a) shows the details of the *locations* created for this model. For the 'machine process' and 'failure occurrence' *locations*, *clock* and/or *called* downtimes are specified, as described below:

- *Clock* (planned/scheduled) downtime. This routine is repeated at the beginning of every PM interval, causing machine downtime due to inspection (if any) and/or replacement of bearings (if any). Figures 5.4 to 5.6 (also shown in Appendices 5.3 to 5.5 for the completeness of information in the Appendices) display the main algorithms and sub-processes for the *clock* downtime;
- *Called* (unplanned/unscheduled) downtime. This routine is 'called' whenever there is machine downtime due to bearing failures. Appendices 5.6 and 5.7 show the main algorithms and sub-processes for the *called* downtime.

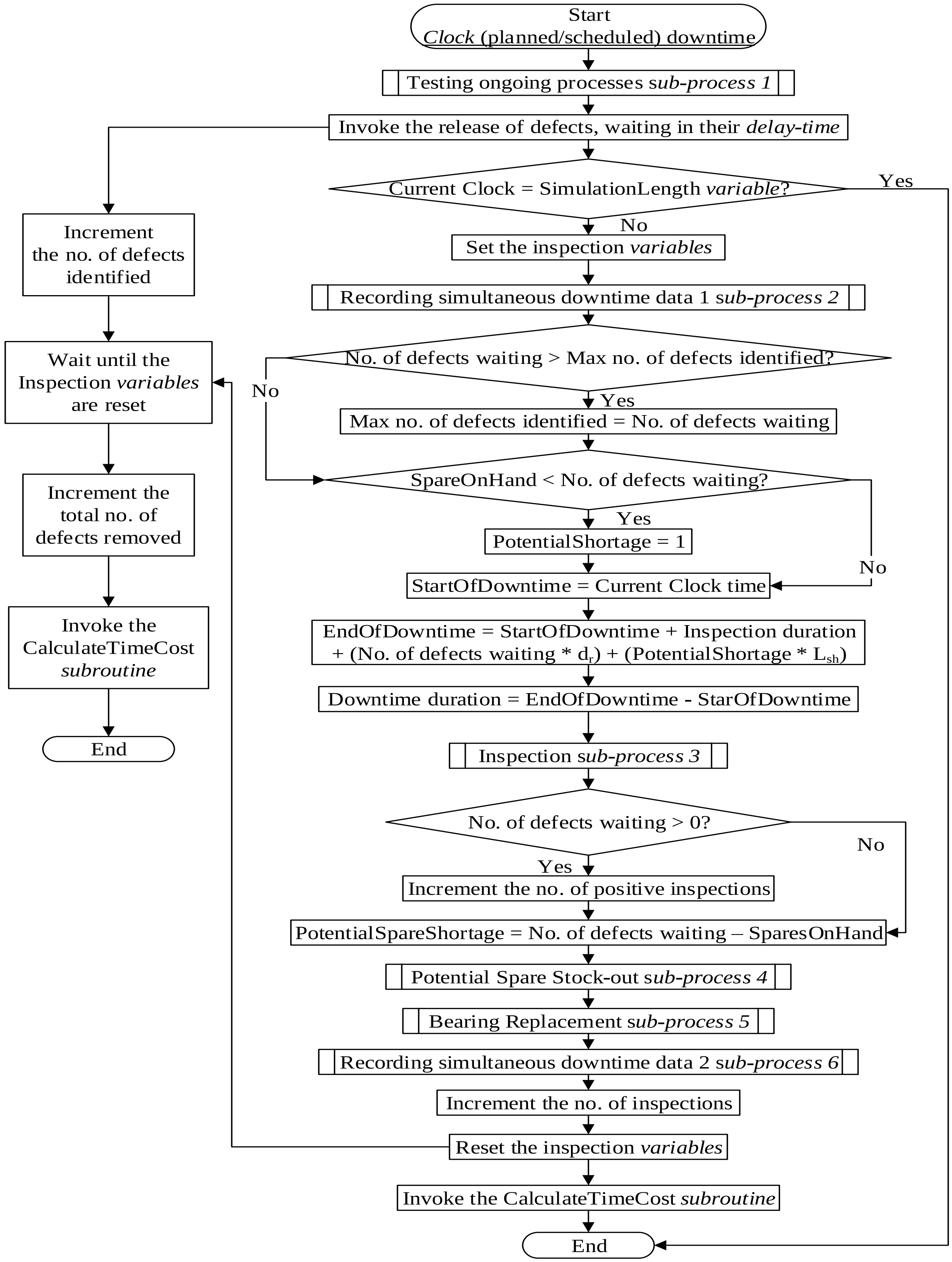


Figure 5.4.

Flowchart of the *Clock* (planned/scheduled) downtime routine for each machine.

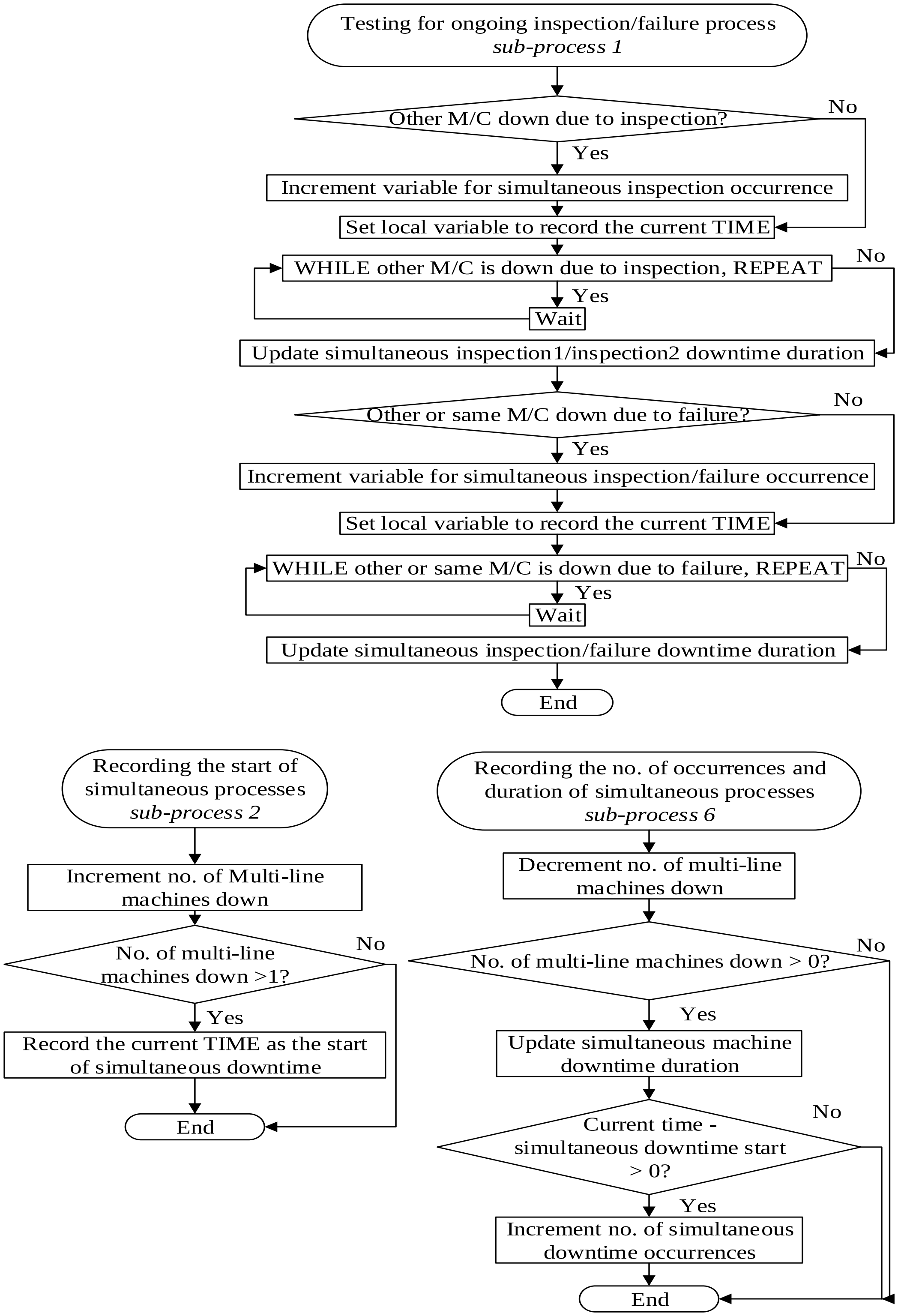


Figure 5.5. Flowchart of the *Clock* (planned/scheduled) downtime *sub-processes 1, 2 & 6* routines for each machine.

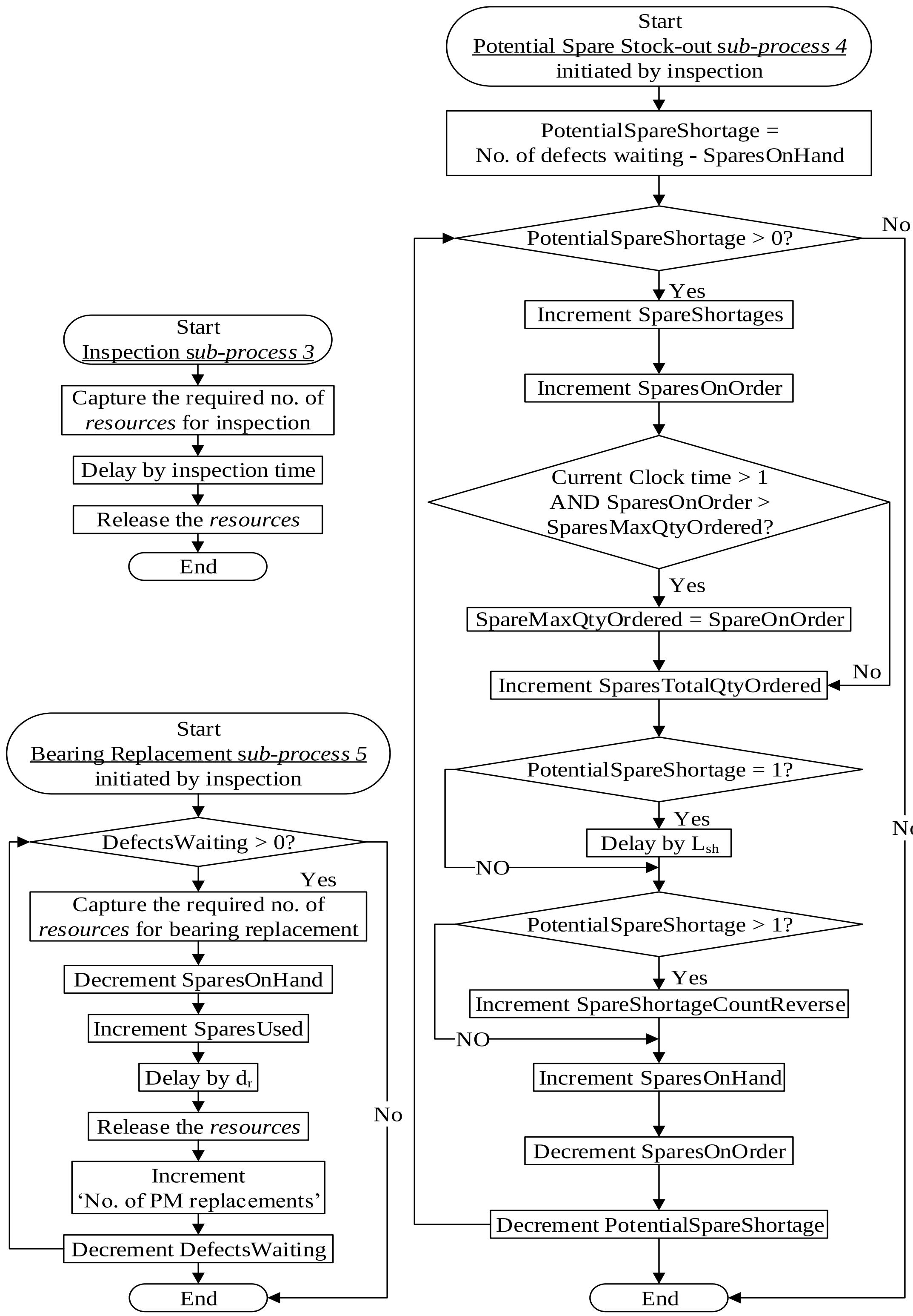


Figure 5.6. Flowchart of the *Clock* (planned/scheduled) downtime *sub-processes 3, 4 & 5* routines for each machine.

(b) *Entities*: jobs, products, defects, failures, etc. Appendix 5.2(b) displays the details of the *entities* created for this model.

(c) *Path Networks*: physical routes between *locations*.

(d) *Resources*: *resources* used for the operation of machines, such as, operators, maintenance technicians, etc. Appendix 5.8(a) shows the details of the *resources*.

(e) *Processing*: detailed programming logic for *entities* moving from one *location* to another, including processing times and move logic. Appendix 5.8(b) displays the details of the *processing* created for this model with further information given in Appendices 5.9 to 5.12.

(f) *Arrivals*: creation of *entities* such as, defects and failures etc., and their pattern of arrival into the system. Appendix 5.8(c) shows the details of the *arrivals*.

(g) *Attributes*: information that is stored in each entity and moves with that same entity. Appendix 5.13(a) displays the details of the *attributes* created for this model.

(h) *Macros*: for easy alteration of model and decision variable values. Appendix 5.13(b) shows the details of the *macros* created for this model.

(i) *Subroutines*. Appendix 5.13(c) displays the details of the *subroutines* created for this model with further detailed information given in Appendices 5.14 and 5.15. Figure 5.7 (also shown in Appendix 5.16 for the completeness of information in the Appendices) shows the flowchart of "CalculateTimeCostMCS" *subroutine* (multi-line), for calculating $D(T)$ and $C(T)$.

(j) *Variables*: global variables used in the model.

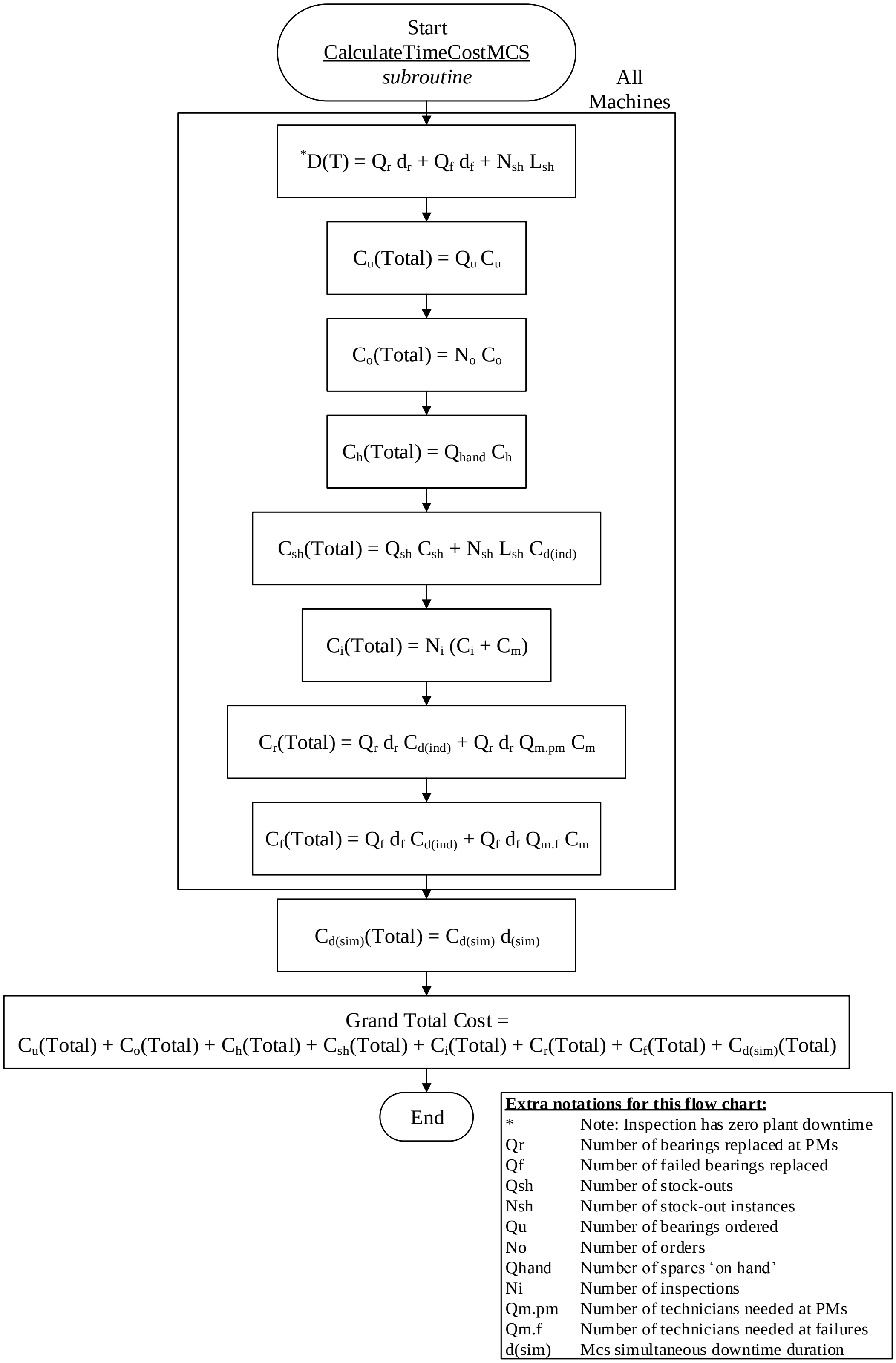


Figure 5.7. Flowchart of the CalculateTimeCostMCS *subroutine* (multi-line).

5.4.3. Input parameters, output analysis, model scenarios and optimisation

Before analysis can begin, the *warm-up* period, the number of *replications* and the *length of run* had to be determined to ensure the quality of the output data and the validity of the simulation results. The Time Series method based on the weekly cost mean value and Welch's method based on the weekly cost moving average with a window length of 5 (see, for example, Banks, 2010; and Law, 2015) were used to determine the *warm-up* period for the model. Although detailed analysis showed that a *warm-up* period of 600 weeks was sufficient, a conservative 1,000 weeks of *warm-up* was used to provide a very high degree of confidence that the system indeed stabilised and reached the steady-state beyond the transient period. Robinson (2004) states that a single long run may be performed instead of using multiple replications since "if the replications were run for an infinite period, they would produce exactly the same results". Common practice is to use a *warm-up* period and a long *run-length* for non-terminating simulations (Robinson, 2004). Banks (2010) recommends that the *run-length* should be at least 10 times the length of the *warm-up* period. Since rare simultaneous machine downtime occurrences are to be observed, it was decided that the model should be run for 500,000 weeks to ensure that the simulation is run long enough for convergence. The simulation programming is very efficient and the computation time takes only 5 minutes. Figure 5.8 (also shown in Appendix 5.17 for the completeness of information in the Appendices) depicts four situations (a, b, c, and d) of machine downtime for machines 1 and 2, two of which illustrate simultaneous downtime of both machines. The flowcharts in Figure 5.8 capture the process of recording and accumulating the time duration for simultaneous machine downtime.

To run the models, parameter values are inputted, including the model parameters: *warm-up* period and simulation *run-length*. *Macros* enable the instant changing of model and decision variables. The simulation models are non-terminating. Appendix 5.18 shows a depiction of the simulation model layout. The detailed output report includes various data and graphs, most importantly the average expected cost and downtime per unit time.

To find the cost-optimal policy for each policy variant, the simulation models were integrated with *SimRunner* (see ProModel, 2010). This tool uses sophisticated search algorithm (Kim et al., 2012), running multiple (where applicable) combinations of the decision variables, to find the unique one that is optimal. Appendix 5.19 shows a series of illustrations of using the *SimRunner* optimisation tool. In Appendix 5.19, illustration (a) shows the *SimRunner* initial screen. The grand total cost per week *variable* (*response category*) is selected as the optimisation *variable* or the *objective function*, and the *response statistic* is the *order-up-to-level* S, as shown in illustrations (b) and (c) respectively. Illustrations (c) and (d) depict the range of values for the *response statistic* and the model parameter values already determined through separate detailed analysis, respectively. The results of the experiments run by the *SimRunner* for different values of S and the confirmation that *SimRunner* has indeed found the optimum S for the specific $(R, S, T = R)$ policy, using *just-in-time* ordering $(T = 5\ weeks)$, are finally depicted in illustrations (e) and (f), respectively.

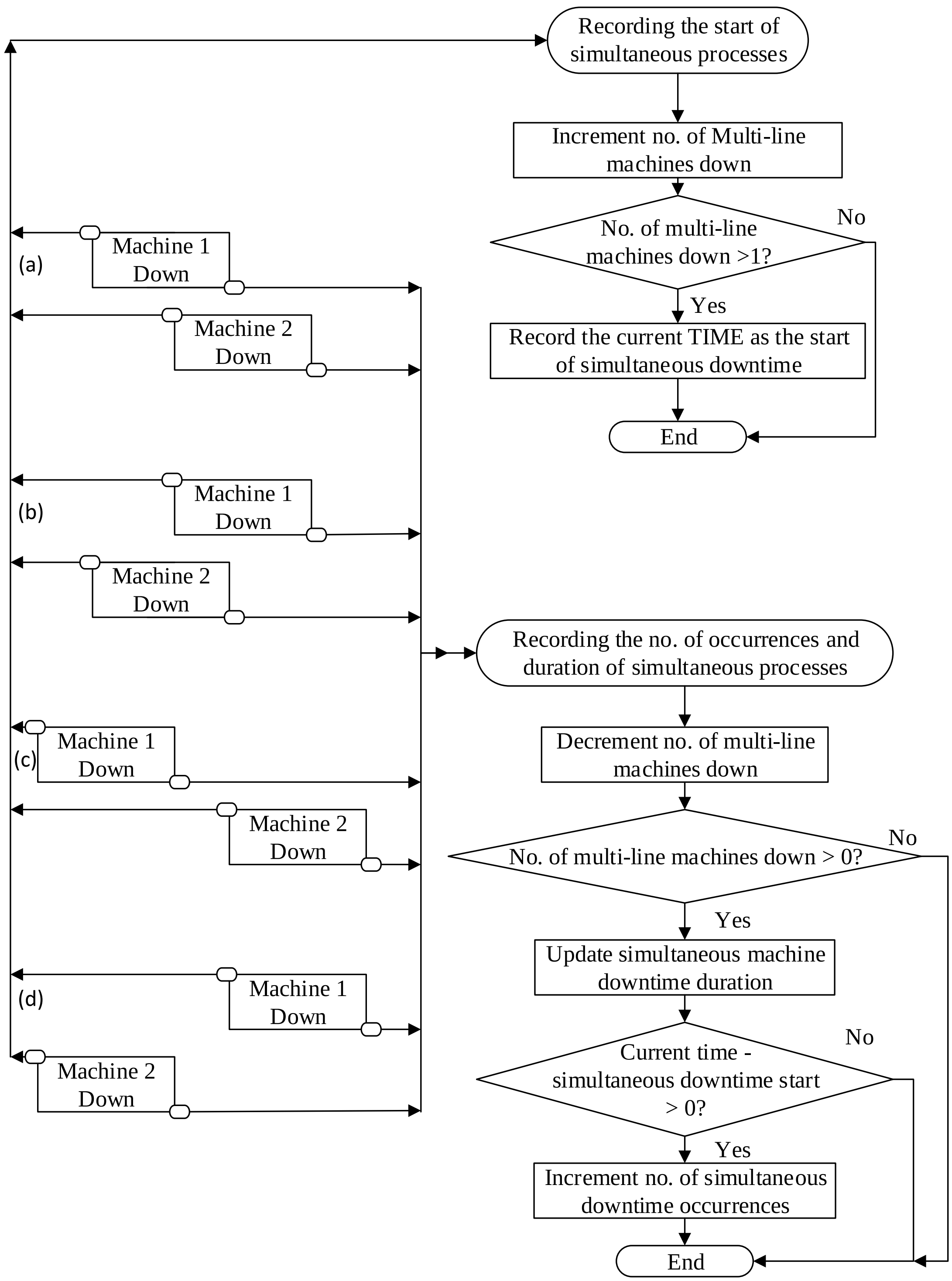


Figure 5.8. Flowchart,

depicting the process of capturing and recording simultaneous machine downtime.

5.5. Results analysis and discussion

5.5.1. Joint optimisation

Figure 5.9 displays the schematic diagram of ten joint inventory-maintenance policy variants, for which simulation models were developed. These policies are: $(R, S, T = R)$; $(R, S, T = 2R)$; $(R, S, T = 3R)$; $(R, S, T = 4R)$; and $(R, S, T = 0.5R)$ using *standard* ordering (Figure 5.9(a)), and the remaining same five models using *just-in-time* ordering (such that the delivery of parts coincides with maintenance interventions) (Figure 5.9(b)). For all variants considered, the joint optimisation policy contains the decision variables: inspection interval T; the review period R; and the order-up-to-level S. Here, values of the decision variables that minimise the cost-rate, $C(T)$, are sought. For all models, the preventive maintenance interval $T = kR$ for $k > 0$. Other values of k were investigated for $T = kR$ but found to be cost-sub-optimal for the range of parameter values used here. For all policies under *standard* ordering, the preventive maintenance coincides with one of the ordering events, and under *just-in-time* ordering, orders are placed so that the arrival of one of them coincides *just-in-time* with the preventive maintenance event.

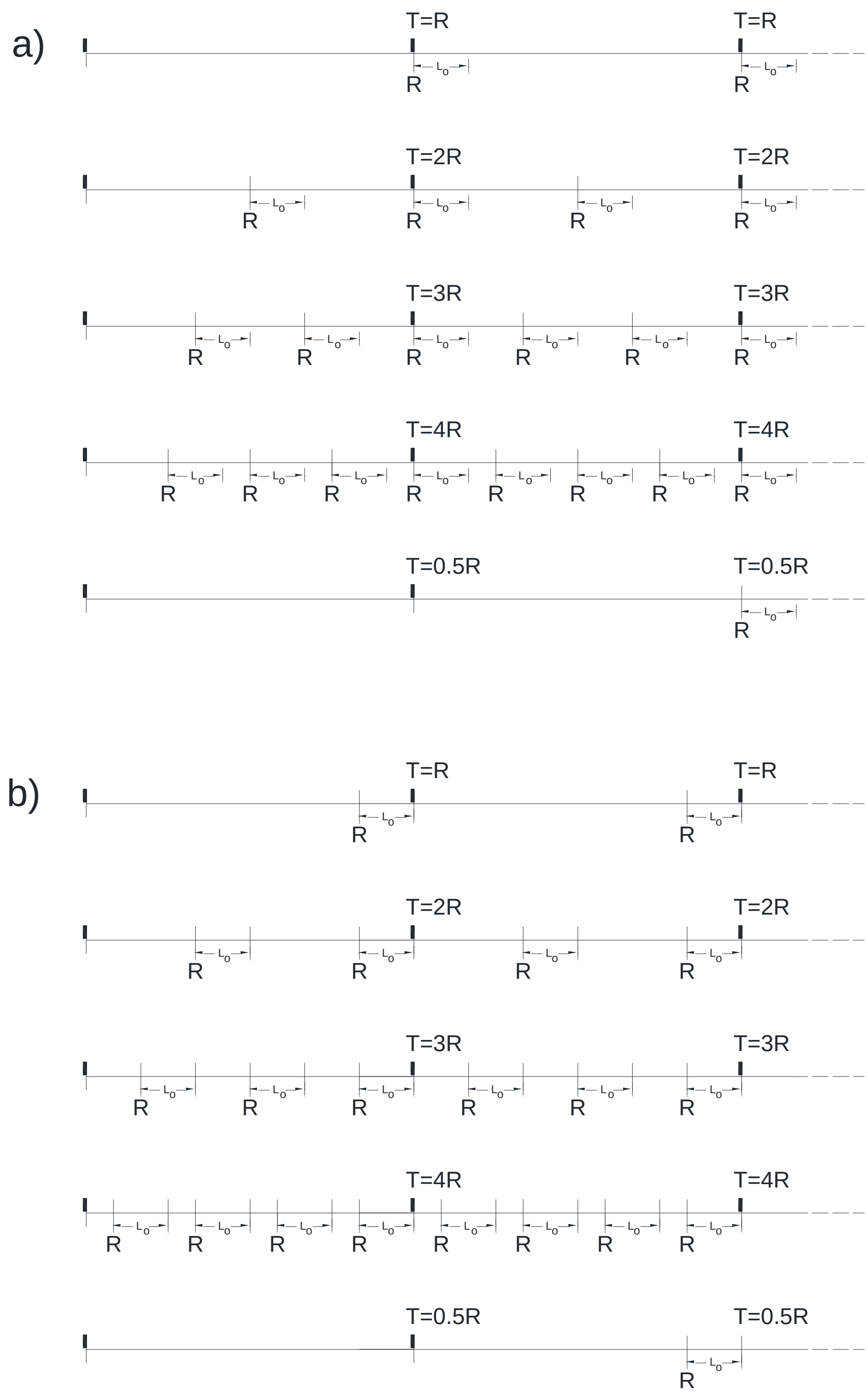


Figure 5.9. Joint maintenance-inventory policy variants considered under: a) *standard*; and b) *just-in-time* ordering.

Figure 5.10(a) and Table 5.1 illustrate the results for the best (lowest cost) four-out-of-ten policy variants that were depicted in Figure 5.9. The results demonstrate that the $(R, S, T = R)$ policy using *just-in-time* ordering, has the lowest total cost per unit time (cost-rate), thus maintaining bearings and ordering spares every 5 weeks in the plant. Figure 5.10(b) confirms that the *SimRunner* optimisation tool has indeed found the optimum S for the cost-optimal policy at $T = 5\ weeks$. Considering the results in Table 5.1 further indicates that the second lowest cost-rate policy, $(R, S, T = 0.5R)$, also uses *just-in-time* ordering, but maintaining bearings every 5 weeks and ordering spares every 10 weeks. The third *best* policy, $(R, S, T = R)$, uses the same frequency of maintenance and ordering spares as the cost-optimal policy but using *standard* ordering. Throughout the analysis, an arbitrary but convenient unit of time (one week) is used for the reporting of the results. The results generally suggest that it is not cost optimal to place multiple orders between preventive maintenance intervals.

The last three columns on the right side of Table 5.1 illustrate the percentage difference in cost of those particular policies compared to the cost-optimal policy. It is interesting to observe that, moving the preventive maintenance interval T by 1 week to either side of the optimum has bigger cost effect (+2.9% and +0.95% for $T = 4$ and $T = 6$, respectively) than changing the type of policy to the second (+0.19%), or third (+0.79%) best policies. This phenomenon is at the heart of the maintenance decision problem.

Under the cost-optimal policy, potentially more frequent orders are placed (every 5 weeks) compared to the second best policy, which will order spare parts every 10 weeks, thus minimising stock-outs and ultimately reducing cost. Thus, more orders might be placed, for $(R, S, T = R)$, or more stock might be held, for $(R, S, T = 0.5R)$, to reduce the possibility of stock-outs, depending on the relative sizes of the order cost and the holding cost. Clearly, the cost of inventory is traded-off with the cost of stock-outs, which lies at the heart of the inventory decision problem. However, there might be times that based on the current inventory position, the optimal cost will be achieved by ordering no spares, determined by the optimal order-up-to-level S. In the joint optimisation problem, where the inspection period is a multiple of the order period, more frequent ordering can, by implication, potentially reduce the frequency of stock-outs, and this certainly appears to be the case for $k = 1$ and $k = 0.5$ in this study.

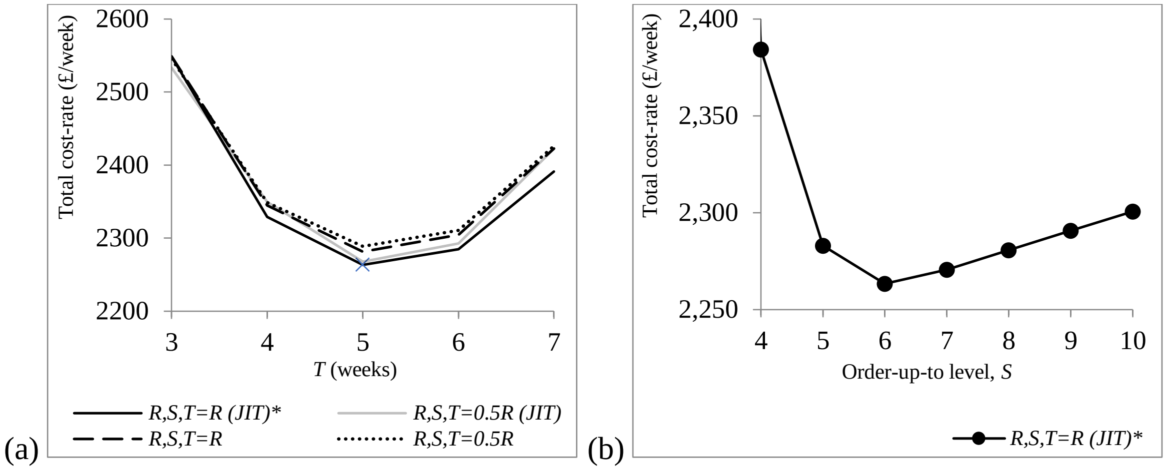


Figure 5.10. The effect on the joint optimisation cost:

(a) the best four-out-of-ten policy variants (optimal policy*; optimum interval^{X}); and

(b) order-up-to-level S for T=5 (weeks), for the cost-optimal policy.

Table 5.1. Comparison of the joint optimisation cost for the best four-out-of-ten policy variants, based on parameter values. Optimal cost for each policy variant. **Overall lowest-cost policy**.

	Cost-rate (best four-out-of-ten policy variants (types))				% difference, compared to the baseline*		
	(R,S,T=R)	*(R,S,T=0.5R)*	*(R,S,T=R)*	*(R,S,T=0.5R)*	*(R,S,T=0.5R)*	*(R,S,T=R)*	*(R,S,T=0.5R)*
Ordering>	*JIT**	*JIT*	*Standard*	*Standard*	*JIT*	*Standard*	*Standard*
T (weeks)	Cost/week	Cost/week	Cost/week	Cost/week	%	%	%
2	3,009.43	3,003.66	3,011.79	3,003.66	-0.19	0.08	-0.19
3	2,548.74	2,533.45	2,548.74	2,545.47	-0.60	0.00	-0.13
4	2,329.09	2,349.60	2,344.53	2,348.62	0.88	0.66	0.84
5	**2,263.35**	2,267.60	2,281.30	2,288.67	0.19	0.79	1.12
6	2,284.74	2,292.45	2,304.46	2,310.44	0.34	0.86	1.12
7	2,391.10	2,422.07	2,422.41	2,425.88	1.30	1.31	1.45
8	2,541.58	2,569.04	2,554.03	2,556.38	1.08	0.49	0.58
9	2,701.57	2,721.71	2,713.57	2,746.03	0.75	0.44	1.65
10	2,896.55	2,930.98	2,948.58	2,975.61	1.19	1.80	2.73

5.5.2. Insights into simulation results and characteristics of different policies

The results for the best four-out-of-ten policy variants (policy types) in Table 5.1 were further analysed in detail and are illustrated in Figures 5.11 to 5.15. The optimal inspection interval for all those policies is the same ~ 5 weeks.

Considering the mean number of defects removed at maintenance events, Figure 5.11(a) demonstrates that as inspection becomes infrequent, the number of defects identified and thus removed declines. With an arrival intensity of 0.125 defects per week, for both machines we would expect to find 25 defects in total every 100 weeks. Figure 5.11(a) illustrates that, on average, 24.26 of these defects are identified and removed if inspection is carried out at the optimal interval. Figure 5.11(b) supports this observation since, as inspection becomes infrequent, the failure rate rises (because defect identification/removal rate falls as illustrated in Figure 5.11(a)). Obviously, for each policy, the total number of defects identified/removed and the number of failures should accumulate to the number of defects arriving into the system (i.e. 25) as shown in Figure 5.11(c). The information in Figure 5.11(c) further illustrates that if inspection is carried out at the optimum interval, one failure at most, will be expected every 100 weeks. While failure demand is always unit-sized, demand at an inspection may be greater than unit-sized when more than one bearing is found to be defective.

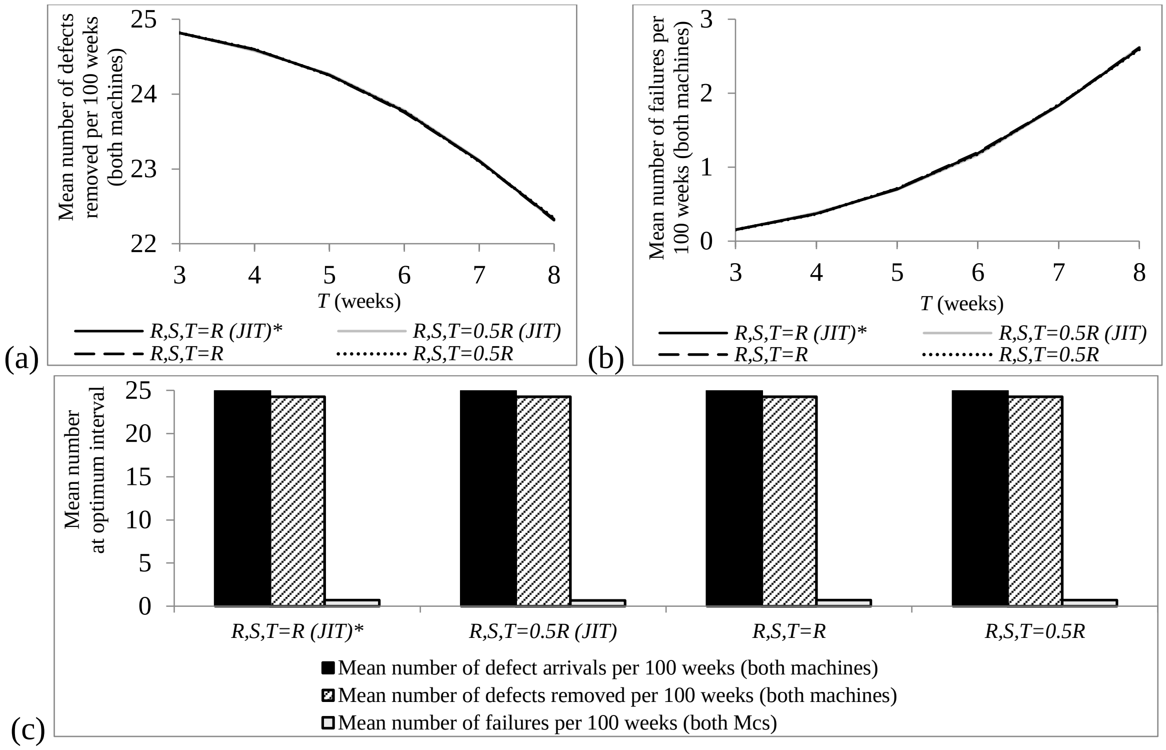


Figure 5.11. For the best four-out-of-ten joint policy variants (optimum policy*):

(a) defect removal rate;

(b) failure rate; and

(c) defect/failure statistics at optimum interval.

Comparing Figures 5.12(a) and 5.12(b) highlights that as inspection becomes infrequent, the percentage of positive inspections (i. e. defects found ≥ 1) rise (31%; 39%; 45%; 51%; 55% and 59% for inspection intervals 3 to 8, respectively). Figure 5.12(c) (2nd and 3rd bars) demonstrate that there is understandably a direct association between the number of positive inspections and the number of spares replaced (at the optimal interval, every 100 weeks, there are on average 18 positive inspections and 24 spares are ordered).

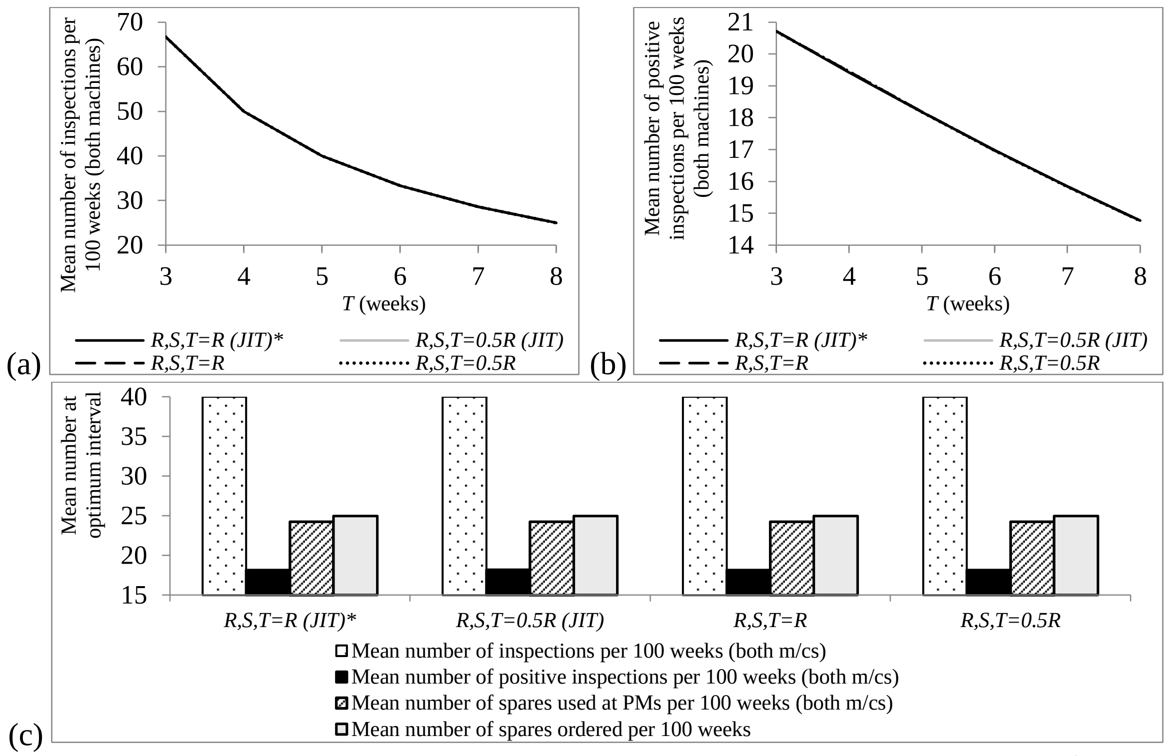


Figure 5.12. For the best four-out-of-ten joint policy variants (optimum policy*):

(a) inspection rate; (b) positive inspection rate; and

(c) inspection/spares statistics at optimum interval.

In Figure 5.13(a), variations in three different cost-rates are illustrated. Although the variations between policy variants are not significant, their impact influence the policy ranking. The PM replacement cost-rates are slightly higher for the policies using *just-in-time* ordering (0.05%) (superior policies), but the failure replacement cost-rates are on average 1.62% higher for the policies using *standard* ordering (inferior policies) since more failures are likely to occur due to the arrival of spare parts being out-of-sequence with maintenance. The inspection cost-rates are identical as expected.

Considering the ordering cost-rate, Figure 5.13(b) demonstrates that the cost for the $(T = R)$ policies are higher since they can potentially place more orders. This is supported by Figure 5.14(a) since the cost-rates for the $(T = R)$ policies are 55% higher at the optimum interval.

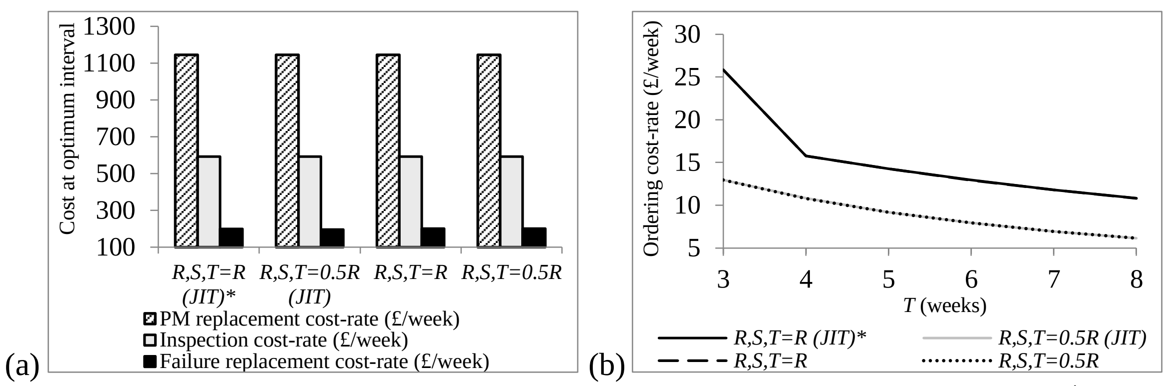


Figure 5.13. For the best four-out-of-ten joint policy variants (optimum policy*):
(a) maintenance cost-rates at optimum interval; and
(b) ordering cost-rate.

An interesting observation is that inventory costs seem to be traded-off. Thus considering Figures 5.14(a) and 5.14(b), it can be seen that for each policy type (*just-in-time* and *standard* ordering) where the ordering cost-rate is high, the holding cost-rate is low, and vice versa. The holding cost-rate is mainly influenced by the frequency of ordering and order-up-to-level S and seems to have a significant effect on the policy ranking. Whereas the difference between the joint optimisation cost-rates of the best and the second best policies is only £4.25 (Table 5.1), the difference between the holding cost-rates for the same policies is £3.90 (Figure 5.14(b)), which accounts for 92% of the cost difference and proves to be very significant.

Although the simultaneous machine downtime cost does not seem to be significant in term of the overall cost, Figure 5.14(c) shows this to be in line with the policy ranking; the lowest cost for the *best* policy and the highest cost for the most expensive policy. It is interesting to note that both spare holding cost-rates (Figure 5.14(b)) and simultaneous machine downtime cost-rates (Figure 5.14(c)) display similar trends. The implication is that the two policies which order spares less frequently ($(T = 0.5R)$ policies) and the policy that has its spares delivered out-of-sequence with maintenance ($(T = R)$ policy using *standard* ordering) are likely to have higher simultaneous machine downtime cost-rates, due to more likelihood of stock-outs. This observation certainly appears to be evident in Figure 5.14(c).

A number of points should be noted about the stock-out cost-rates shown in Figure 5.14(d). First, as expected, for $(T = 0.5R)$ policies, the cost-rates are much higher since the frequency of ordering spares is half as many as inspection and stock-outs are therefore more likely to occur. Second, the variation across the policies is relatively large, considering the overall cost-rate differences between policies. The third point is that generally the stock-out cost-rates are much lower than the failure cost-rates shown in Figure 5.13(a). Fourth, the $(R, S, T = R)$ policy using *standard* ordering, which may be perceived as a low risk policy since it has the lowest stock-out cost-rate, has a very large ordering cost-rate (Figure 5.14(a)). The fifth point is that the optimal policy has the second lowest stock-out cost-rate, which is much lower than the inferior policy (59% lower), making it a relatively low-risk policy. Finally, in general, stock-out cost-rates (Figure 5.14(d)) and ordering cost-rates (Figure 5.14(a)) display opposite trends - policies which have low ordering cost-rates tend to have high stock-out cost-rates, and vice versa.

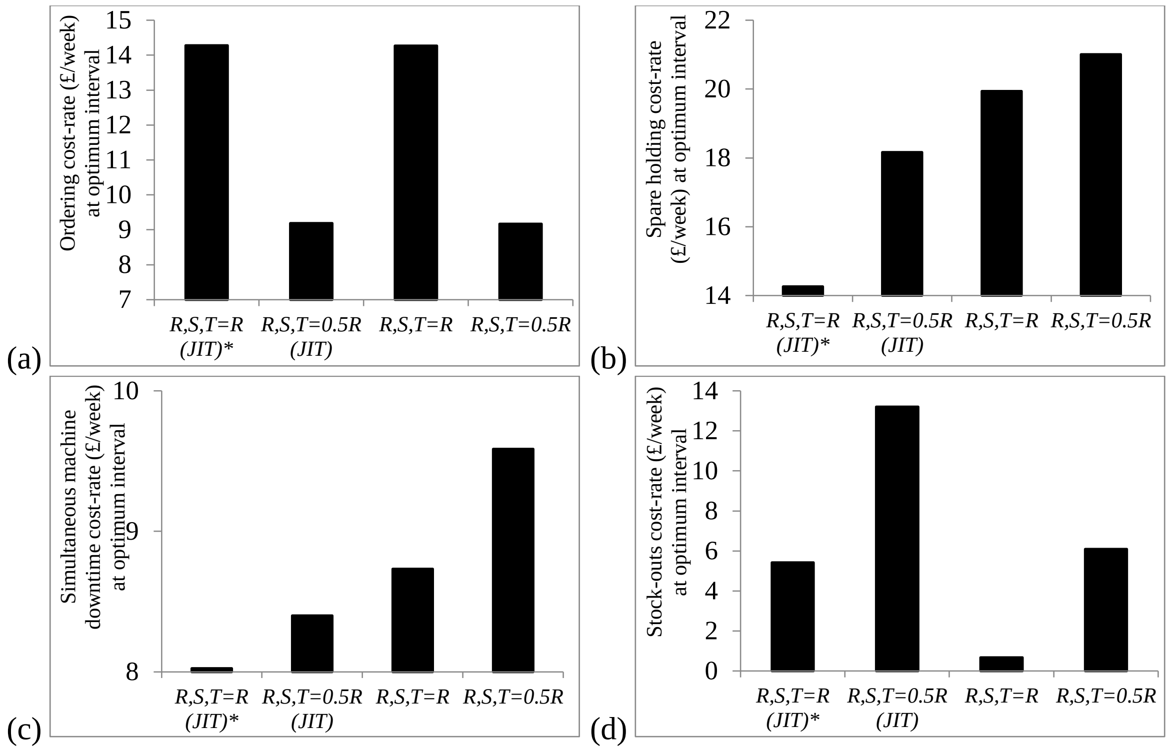


Figure 5.14. For the best four-out-of-ten joint policy variants at optimum interval (optimum policy*):
(a) ordering cost-rate;
(b) holding cost-rate;
(c) simultaneous machine downtime cost-rate;
and (d) stock-out cost-rate.

Where $(T = R)$ policies do appear to outperform $(T = 0.5R)$ policies is in the additional opportunities they offer for replenishments, even if they do not necessarily use them at every interval (Figure 5.15(a) vs 5.15(b)). Continuing the same discussion, as expected, the mean spares ordered per order for the $(T = R)$ policies are lower than the $(T = 0.5R)$ policies (Figure 5.15(c)) since the mean number of orders for the former policies are greater (potentially twice as many; Figure 5.15(b)), so there would be more opportunities (Figure 5.15(a)) for placing orders of lower spare quantities (Figure 5.15(c)). The final point is that the maximum spares ordered per order (Figure 5.15(d)) are generally lower for the $(T = R)$ policies since they will have twice as many opportunities to place orders compared to the $(T = 0.5R)$ policies. Therefore, there does not seem to be the need for placing orders of higher quantities to cover a longer period between review periods.

In summary, first and foremost, it is the *just-in-time* ordering (placing orders to arrive *just-in-time* to coincide with the preventive maintenance events) that has the greatest influence on the choice of policy. In addition, cost-rates are traded-off which also have an influence on the policy ranking including the spare holding cost-rate, the simultaneous machine downtime cost-rate, and the stock-out cost-rate. Thus, the different policies, at their optimal settings, place different demands on inventory.

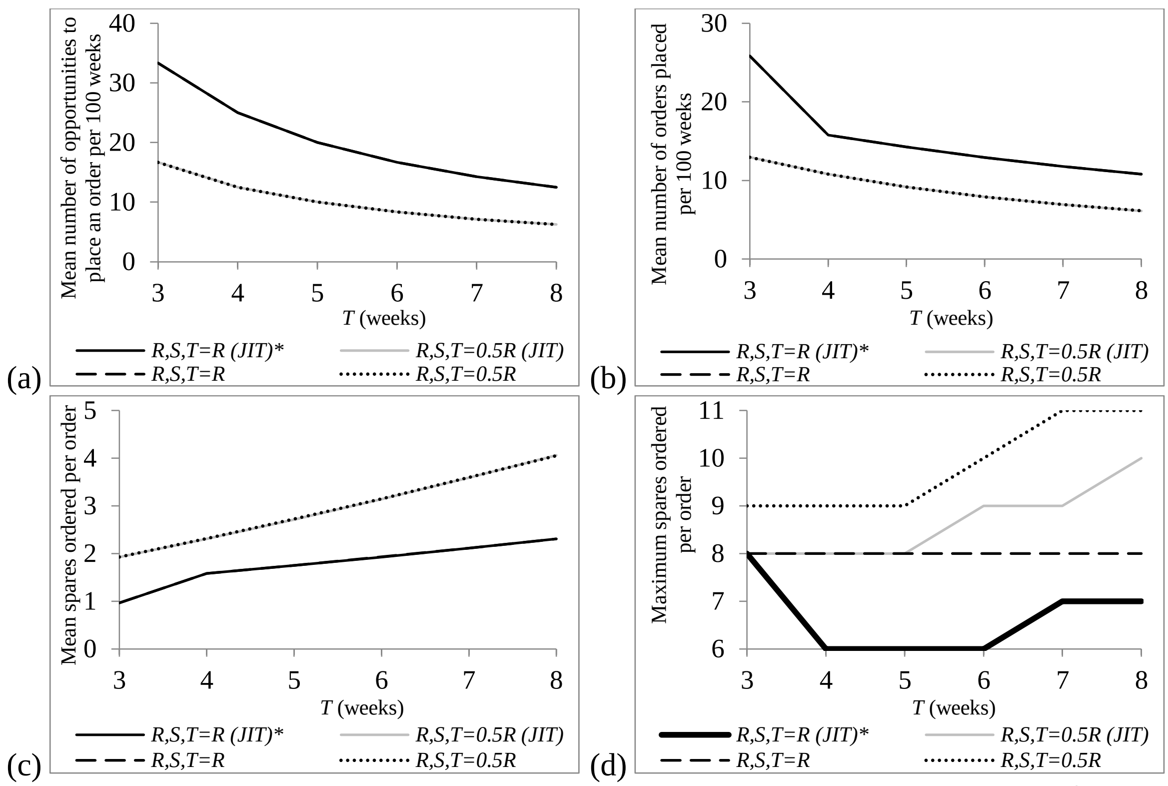


Figure 5.15. For the best four-out-of-ten joint policy variants (optimum policy*):

(a) order opportunity rate;

(b) order placing rate;

(c) mean order size; and

(d) maximum order size.

5.5.3. Sensitivity analysis of the optimal policy

The original values of eleven parameters were halved and doubled to study their effects on the joint optimisation cost of the cost-optimal policy, $(R, S, T = R)$ using *just-in-time* ordering. Table 5.2 classifies all parameters, in descending order of effects, showing the defect arrival rate and the cost of emergency shipment as the parameters with the most and least effects on the cost-optimal policy, respectively. The detailed discussion below should be considered in conjunction with the data in this table.

Table 5.2. Comparison of the effect of various parameters, in descending order, on the cost-optimal policy $(R, S, T = R\ (JIT))$.

		The effect of parameters and their values on the cost-optimal policy, in descending order										
		λ	$C_{d(ind)}$	α	C_i	C_u	C_h	$C_{d(sim)}$	L_o	C_o	L_{sh}	C_{sh}
		% of the baseline of the cost-optimal policy (optimum *T* = 5 weeks)										
Parameter value	Halved	62.5	73.8	124.0	91.2	93.4	99.2	99.8	99.8	99.7	99.9	100.0
	Doubled	170.2	150.9	86.2	115.7	113.3	103.0	103.1	101.3	100.6	100.3	100.0
	Halved	6	6	3	5	5	5	5	5	5	5	5
	Doubled	4	4	9	6	5	5	5	5	5	5	5
		Optimum *T* as a result of halving or doubling parameter values (original *T* = 5 weeks)										

For the defect arrival intensity (Figure 5.16(a)), the behaviour of the cost-rate is as expected since the cost-rate for 0.5λ and 2λ are at approximately 63% and 170% of the baseline, respectively. For the case of 2λ, inspection that is more frequent is expected to yield the optimal interval, and the behaviour is exactly as expected since the optimum interval is shortened by one week, and vice versa for 0.5λ.

Similarly, when the scale parameter, α of the Weibull *delay-time* distribution is doubled (Figure 5.16(b)), the mean *delay-time* increases and defects will take longer to develop into failures, and thus the cost-rate falls due to less failures as expected (86% of the baseline, and an optimum inspection interval of 9 weeks ~ almost double the original). However, when α is halved, defects develop into failures quicker, which increases the cost-rate, especially when inspection is infrequent. The cost-rate of the optimal inspection interval for 0.5α is 124% of the baseline with an optimum inspection interval of 3 weeks. In both cases, minimal effect is displayed when inspection is very frequent.

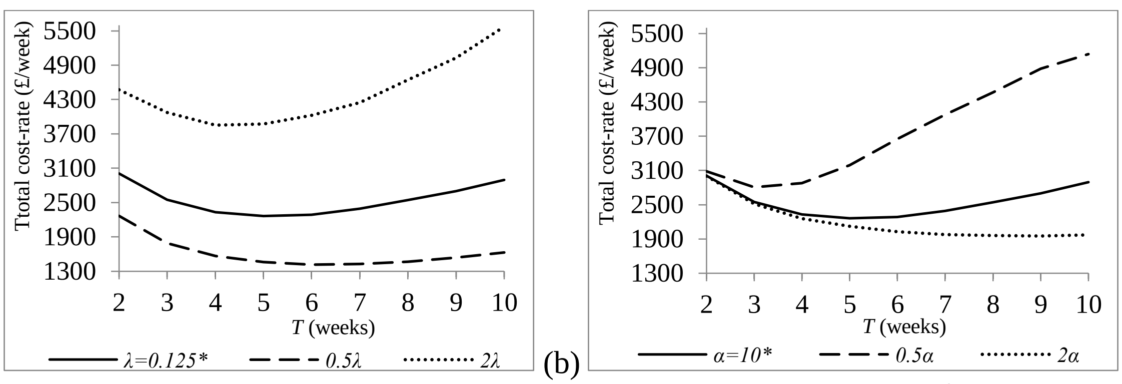
(a)


Figure 5.16. The effect of various parameters on the cost-optimal policy (*baseline): (a) defect arrival rate; and (b) failure *delay-time*.

Figure 5.17(a) displays the cost-rate of the optimal inspection interval for $0.5\mathrm{C}_{d(ind)}$ and $2\mathrm{C}_{d(ind)}$ at approximately 74% and 151% of the baseline, respectively. The optimal times between inspections also behave as expected (moving one interval to either side of the original, respectively), but with the greatest impact for the 100% increase in the cost of downtime and when inspection is infrequent. Overall, the greatest impact is evident when inspection is less frequent, as expected.

Moving the discussion to the cost of inspection, the cost-rate of the optimal inspection interval for $0.5\mathrm{C}_i$ and $2\mathrm{C}_i$ are at approximately 91% and 116% of the baseline respectively, as depicted in Figure 5.17(b). Overall, varying C_i has the greatest impact when inspection is frequent as expected, and particularly when the parameter value is doubled.

Finally, the cost-rate of the optimal inspection interval for $0.5\mathrm{C}_u$ and $2\mathrm{C}_u$ are approximately 93% and 113% of the baseline, respectively as depicted in Figure 5.17(c). The change in the unit cost does not seem to have a significant impact on the frequency of inspection.

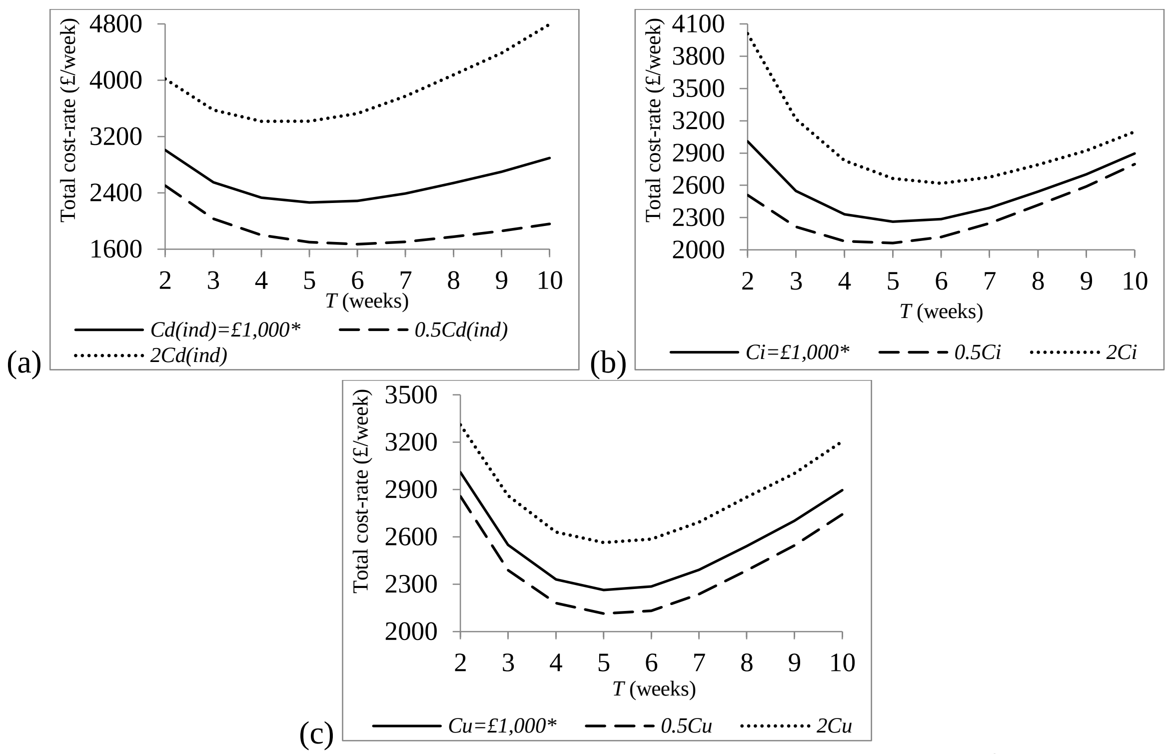


Figure 5.17. The effect of various parameters on the cost-optimal policy (*baseline): (a) individual machine downtime cost-rate; (b) inspection cost; and (c) unit cost.

Changing the values for the other parameters seem to have little effect on the cost-optimal policy, as seen in Table 5.2. Altering the holding cost has minimal effect when its value is halved but shows a moderate effect when the value is doubled. For the simultaneous machine downtime, the cost-rates for $0.5\mathrm{C}_{d(sim)}$ and $2\mathrm{C}_{d(sim)}$ are at 99.8% and 103.1% of the baseline. Compared to the high impact of the individual machine downtime cost-rate, it is not unusual to see a minimal effect for this parameter since simultaneous machine downtime is a rare event in this context. However, the $2C_{d(sim)}$ shows the greatest effect when inspection is infrequent and simultaneous machine downtime is more likely. The sensitivity of the $0.5C_o$ and $2C_o$ suggests that halving and doubling the order cost has minimal effect with 99.7% and 100.6% of the baseline. A change in the normal and emergency delivery lead-times will have very low effect on the overall cost-rate due to two reasons. First, since the optimal policy is one that uses *just-in-time* ordering, the effect of L_o is expected to be low anyway (99.8% and 101.3% of the baseline with the same optimal interval) since it matters little if the lead-time is large or small. Second, the trade-offs used in optimising the total cost-rate have already ensured that stock-outs are kept to a minimum, and thus the effect of L_{sh} is also expected to be very low (99.9% and 100.3% of the baseline with the same optimum interval). Therefore, the effect on the stock-out emergency delivery cost, C_{sh} is expected to be almost zero for every inspection interval.

5.6. Conclusions and further work

Several simulation models were developed for a multi-line production facility. The aim was to determine the joint optimisation of the planned maintenance inspection interval T, based on the delay-time concept, and the spare parts provisioning, considering various variants of the (R, S) inventory replenishment policy. The models were specifically developed for a complex system with multiple identical components in parallel machines. A paper machinery plant with two parallel machines provided the industrial context. However, the models are easily scalable for other multi-line industrial situations.

It is important to note that this is the first joint-optimisation study that addresses the maintenance of a multi-line system and considers a number of periodic review policy variants for the replenishment of spare parts. This optimisation is performed in order to eliminate, or at least minimise, the occurrence of simultaneous machine downtime in a system with parallel machines. The occurrence of simultaneous machine downtime in multi-line production settings may halt production, which will have a significant adverse effect on profitability or other performance measures. The main aim is twofold: (i) to find the joint cost-optimal policy among several policy variants; and (ii) to have insights into the characteristics of various policies in order to guide maintenance and inventory control practitioners about the suitability of policies for their particular contexts.

According to the studies in the literature, single-line studies are oversimplified and do not reflect the complexity and interactions of real systems in practice. Simulation is used in this study as a solution tool, which compared to mathematical modelling, has the advantage of being able to describe multivariate non-linear relations, and is the ideal tool for parallel-line production settings. However, since simulation is not an optimization technique, *SimRunner* (an optimization tool) was integrated with *ProModel* to find the global optimal policy.

For the ten joint maintenance-inventory policies considered in this study, it is found that:

- The $(R, S, T = R)$ policy, using *just-in-time* ordering is (i) the global cost-optimal policy; (ii) the lowest risk policy as it is associated with the lowest simultaneous machine downtime cost-rate, and a relatively low stock-out cost-rate; (iii) compensated by the reduction of potential machine downtime due to timely availability of spares; and finally (iv) associated with the lowest mean spares per order and the lowest maximum spares per order, thus reducing the holding cost.

- At the optimal interval, the two $(R, S, T = R)$ policies (using *just-in-time* and *standard* ordering) result in potentially placing more orders with fewer quantities every time, compared to the two $(R, S, T = 0.5R)$ policies, resulting in a higher ordering cost-rate but a lower holding cost-rate for the cost-optimal policy.

- It is not cost-effective to place multiple orders between preventive maintenance intervals.

- As inspection intervals get longer, the percentage of positive inspections increase from 31 to 59% for inspection intervals 3 to 8 (weeks), and is at 45% for the optimal interval. In addition, the ratio of the PM replacement cost-rate to the failure cost-rate is reduced since the ratio of the number of defects removed to the number of failures is also reduced. The results are similar for all policies.

- While the cost-rates are very similar across the four policies, the components of the cost-rates are quite diverse at different intervals because the trade-offs are different, and so the different policies place different demands on inventory.

- The effect of different values of parameters on the cost-optimal policy, $(R, S, T = R)$ using *just-in-time* ordering, give results that are broadly expected. Varying C_i and $\mathrm{C}_{d(ind)}$ have the greatest impact when inspection is frequent and infrequent, respectively. For the scale parameter of the *delay-time* distribution, minimum and maximum effects are displayed when inspection is very frequent and infrequent, respectively. The defect arrival rate and the cost of emergency shipment parameters have the most and least impact on the cost-optimal policy, respectively. When sensitivity analysis is broadly in line with expectations, it partly validates the simulation results, but at the same time increases the confidence for relying on results which are less obvious.

Extensions to the work presented in this study may be developed in several directions to model more realistic industrial situations. The future extensions might include: imperfect inspection; postponed replacement or repair of defective components; variable replenishment lead-times; manpower planning; modelling of dependent and/or non-identical multi-unit systems; and finally, joint-optimisation for non-identical multi-line parallel systems.

Chapter 6

Conclusions and future research

6.1. Conclusions

Almost all inspection models in the literature are concerned with single-line single-component systems or series systems with multiple components, with restrictions. Single-line studies that dominate the literature are oversimplified and do not reflect the complexity and interactions of real systems in practice (Alrabghi and Tiwari, 2015). To determine the optimal inspection interval for multi-line production systems, it may not be possible to use analytical models that require difficult mathematical formulations and analysis, except for limited situations and/or with simplifying assumptions. Thus in this PhD project, simulation is used as a solution tool, which compared to mathematical modelling, has the advantage to describe multivariate non-linear relations and is therefore the ideal tool for parallel-line production settings. Real-time decision making using simulation is very useful in dynamic situations where the condition of the system state is monitored. The use of simulation allows the models to make less simplifying assumptions than is usual with analytical models.

The work carried out in this thesis is presented in three principal Chapters, 3, 4 and 5. In Chapter 3, simulation is used to model and determine for the first time the optimal inspection policy for a number of multi-line production facility scenarios using the delay-time concept. In Chapter 4, a number of discrete-event simulation models are developed for the joint optimisation of maintenance inspection interval T, based on the delay-time concept, and the spare part inventory using simultaneously a selection of periodic and continuous review replenishment policies for the first time. The main aim is to find the most suitable and cost optimal policy for the industrial example of inspecting and replenishing bearings (critical spare part) for a paper-rolling plant. Whereas, the optimal inspection of a multi-line system was addressed in Chapter 3, and the joint optimisation of inspection maintenance and spare parts inventory for a single-line production setting was studied in Chapter 4, the third and final principal chapter, Chapter 5, combines the two important elements, thus simulating the joint modelling and simultaneous optimisation of preventive maintenance and the associated spare parts inventory for a multi-line production system.

This thesis makes three important contributions. The first contribution is the modelling and simulation of a number of parallel or multi-line production systems, using delay-time modelling. The second contribution is that the thesis considers a range of inventory replenishment policies in joint optimisation with maintenance planning and develops insights into the characteristics of each policy, which were not previously addressed in joint optimisation studies. The third and final contribution of this PhD project is the joint optimisation of preventive maintenance and spare parts inventory, specifically for a multi-line production system, which is not considered by the existing studies in the literature. In this respect, it can be argued that the overall contribution of this PhD thesis is in narrowing the gap between the theory and practice of managing multi-line systems.

In the first study (Chapter 3), initially a single-line facility is simulated in order to reproduce earlier analytical results. Then, a *modified* two-out-of-three parallel system is modelled and analysed to help address the issue of plant downtime under the *actual* operating conditions in the case study. Finally, two further model extensions are developed and analysed in order to consider whether modifications to either the operation of the system or the design of the system in the case study would be worthwhile. The latter three (out of four) models extend the study by Akbarov et al. (2008), in which the multi-line production facility is modelled as if it is a single line. The initial research findings of this PhD, documented in the first study, can therefore be summarised as:

- The simulation of the single-line system reproduces earlier analytical results.

- Consecutive inspection with prioritised failure repair lowers the cost (by 8.3%) and reduces the frequency of inspections (by 20%) compared to consecutive inspections.

- The standard two-out-of-three design configuration increases the cost (by 1.6% and 0.6%) compared to the *modified* two-out-of-three configuration operated by the management.

- The three parallel-line design configuration increases the frequency of inspections (by 25%) and increases the cost of maintenance (by 5.2%) for the consecutive inspection prioritising failure repair.

- The implications for this case study are substantial as the policy proposition suggests a cost reduction of 61.3% compared to the 'run-to-failure' policy.

- The scenarios and policies considered have economic and engineering implications for the management of the production line and that maintenance planning and execution first and foremost needs to be responsive to operational requirements. The study illuminates the practical problems that operations managers face in everyday real-world situations and the complexities that may exist in developing pragmatic solutions.

In the second study (Chapter 4), several simulation models were developed for a complex system with multiple identical bearings, in the context of a paper machinery plant. The planned maintenance inspection interval T, based on the delay-time concept, and the spare parts inventory policy, were jointly optimised. *SimRunner* (an optimisation tool), was integrated with the simulation models to find the global optimal policy. This is the first study in the literature that compares a number of periodic and continuous review replenishment policies, and analyses their efficacy when joined to the inspection policy. The following lists a summary of the conclusions drawn from the second study in this PhD:

- The $(R, S, T = 2R)$, is the cost-optimal policy, which is relatively low risk and is associated with the lowest stock-out cost-rate.

- The additional cost of more frequent ordering, and hence inspection, under the $(R, S, T = 2R)$ policy is compensated by the reduction of bearing failures.

- At the optimal interval, the $(R, s, Q, T = R)$ policy results in ordering more spares every time, compared to the $(R, S, T = R)$ policy, resulting in a lower ordering cost-rate but a higher holding cost-rate.

- The sensitivity analysis of different parameters on the optimum policy $(R, S, T = 2R)$ gives results that are broadly expected. Varying C_d and C_s has the greatest impact when inspection is infrequent and frequent, respectively. The optimal policy at its optimum settings is not sensitive to the order-up-to-level *S*.

- Whilst the cost-rates are similar across the range of policies, the components of the cost-rates are quite different because the policies' decision variables are different, and so the different policies, at their optimal settings, place different demands on inventory.

- Finally, the findings illustrate the characteristics of each policy so that engineers and practitioners may be guided about the suitability for their particular industrial contexts.

The third and final study (Chapter 5) to complete the work of this PhD developed several simulation models for a multi-line production facility to determine the joint optimisation of the planned maintenance inspection interval T, based on the delay-time concept, and the spare parts provisioning, considering several variants of the (R, S) inventory replenishment policy. The simulation models were specifically developed for a complex system with multiple identical components, in the context of a paper machinery plant with two parallel machines. However, the models are easily scalable for other industrial applications. It is important to note that this is the first optimisation study addressing the joint maintenance-inventory problem in a multi-line production system. For the joint policies considered in this final study, it is found that:

- The $(R, S, T = R)$ policy, using *just-in-time* ordering is the cost-optimal policy among several policy variants, which has the lowest simultaneous machine downtime cost-rate and a relatively low stock-out cost-rate. The policy is compensated by the reduction of potential machine downtime due to timely availability of spares, which in turn is due to *just-in-time* ordering, and also reduction in the holding cost.

- It is not cost-effective to place multiple orders between preventive maintenance intervals.

- At the optimal interval, the two $(R, S, T = R)$ policies (using *just-in-time* and *standard* ordering) result in potentially placing more orders with fewer quantities every time, compared to the two $(R, S, T = 0.5R)$ policies, resulting in a higher ordering cost-rate but a lower holding cost-rate for the cost-optimal policy.

- As inspection intervals get longer, the ratio of the PM replacement cost-rate to the failure cost-rate is reduced since the ratio of the number of defects removed to the number of failures is also reduced.

- While the cost-rates are very similar across the four policies, the components of the cost-rates are quite diverse at different intervals because the trade-offs are different, and so the different policies place different demands on inventory.

- The effect of different values of parameters indicate that the defect arrival rate and the cost of emergency shipment have the most and least impact on the cost-optimal policy, respectively.

6.2. Limitations

The simulation models developed in this thesis are based on the assumption of "perfect inspection", which requires the identification and removal of all defective components present in the system at the time of inspection. To model a more realistic industrial scenario will need access to reliable data in order to make the experimentation meaningful. Models are also based on the assumption of immediate replacement of all defective components, identified through inspection, provided spares are available. In real industrial situations, the replacement of some or all of the components may be delayed until the next replacement cycle, which may be more cost-effective, especially if spare parts are not immediately available. Constant replenishment lead-time and infinite manpower availability are also limitations which might not be fully justified in real-life situations. In the joint-optimisation models only standard inventory control policies are used for the replenishments of spare parts. The joint maintenance-inventory optimisation models for the multi-line parallel setting consider identical and independent components, which might not be the case in all industrial situations. The simulation models in this thesis are developed for specific industrial situations and production configurations. Unlike analytical models, simulation-based models require computation time for experimentation, which will inevitably take some time to produce results. Moreover, discrete-event simulation (by its nature), together with an optimisation tool, will not necessarily produce an exact optimum solution because the search space is not continuous.

6.3. Future research

Extensions to the work presented in this PhD thesis may be developed in several directions to address the limitations discussed in Section 6.2. The simulation models may be developed further by relaxing the assumption of "perfect inspection", which is rare in industrial situations. However, this will need good quality and reliable data to make the study and analysis realistic. Simulation is the ideal tool for extending delay-time modelling research. The development of the three-stage delay-time model is a step closer to reality since inspection might reveal more than just one defective state. With real-time condition monitoring, one will be able to identify the degree of defectiveness of each component and thus decide, for example, to delay or postpone the replacement until the next inspection interval, which may be more cost-effective provided the *delay-time* is long enough. This will certainly be beneficial for spare part management, if enough spares are not immediately available. If the item is in a minor defective state and the spare part is not available, one can wait and postpone replacement rather than rushing into an emergency replenishment. However, the extended model is more complicated than the basic delay-time model, and requires more information to enable the parameter estimation procedure. With the flexibility of simulation, there is no reason why a delay-time model of more than three-stages cannot be developed.

Models using variable replenishment lead-time may be developed, since in practice it is unlikely to know the exact duration of lead-time in advance. Manpower planning will also be essential if models are to be implemented in practice, provided the data is based on real industrial situations. The joint optimisation for dependent and/or non-identical units in multi-line parallel systems should be considered. The formulation of a spare ordering policy based on historic data, and *dynamic forecasting* to predict the demand would be very challenging.

The three principal chapters (Chapters 3, 4 and 5) in this PhD thesis are the start of more research studies, which will hopefully be completed and published in the future.

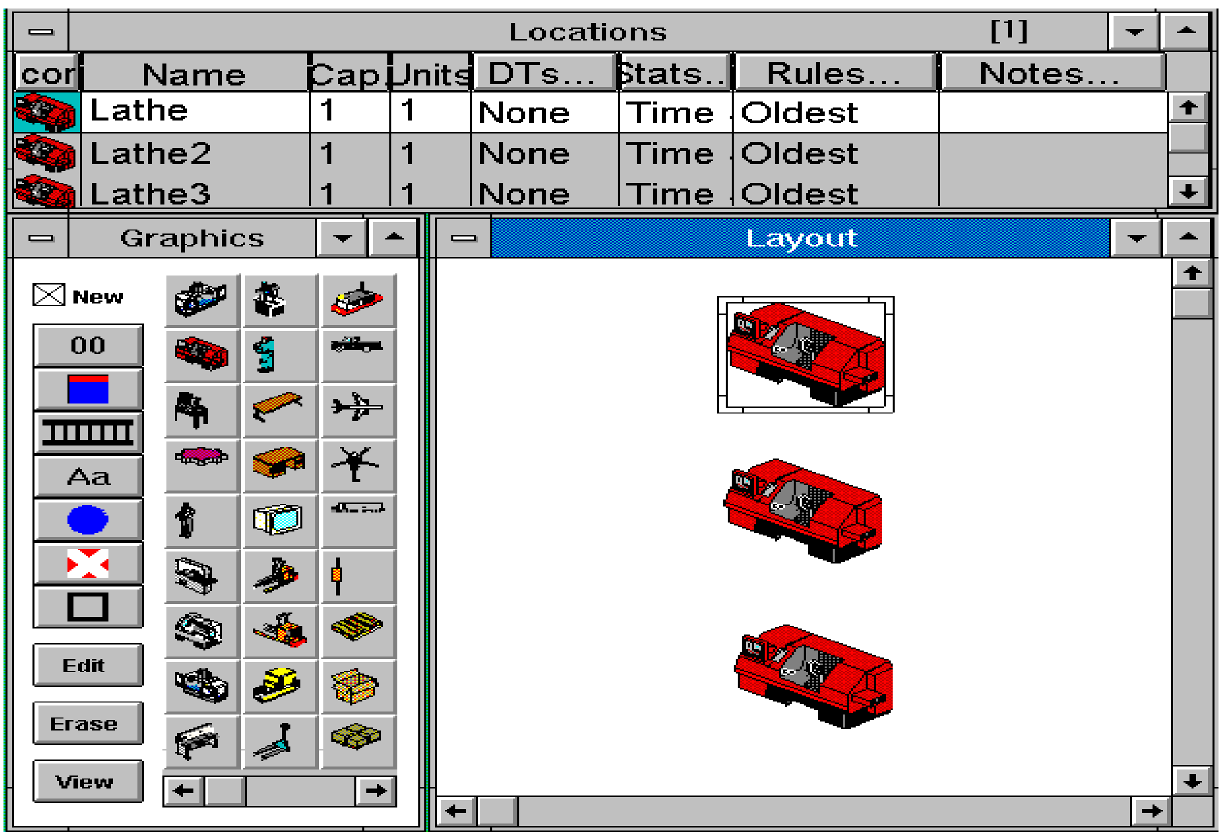
7. Appendices

Each appendix in this thesis is associated with a particular chapter. Appendix 1.1 is associated with Chapter 1. Chapter 2 does not have an appendix. Appendices 3.1 to 3.6 are associated with Chapter 3. Appendices 4.1 to 4.5 are associated with Chapter 4. Finally, Appendices 5.1 to 5.19 are associated with Chapter 5.

Appendix 1.1. Developing a simple model using the *ProModel* simulation tool.

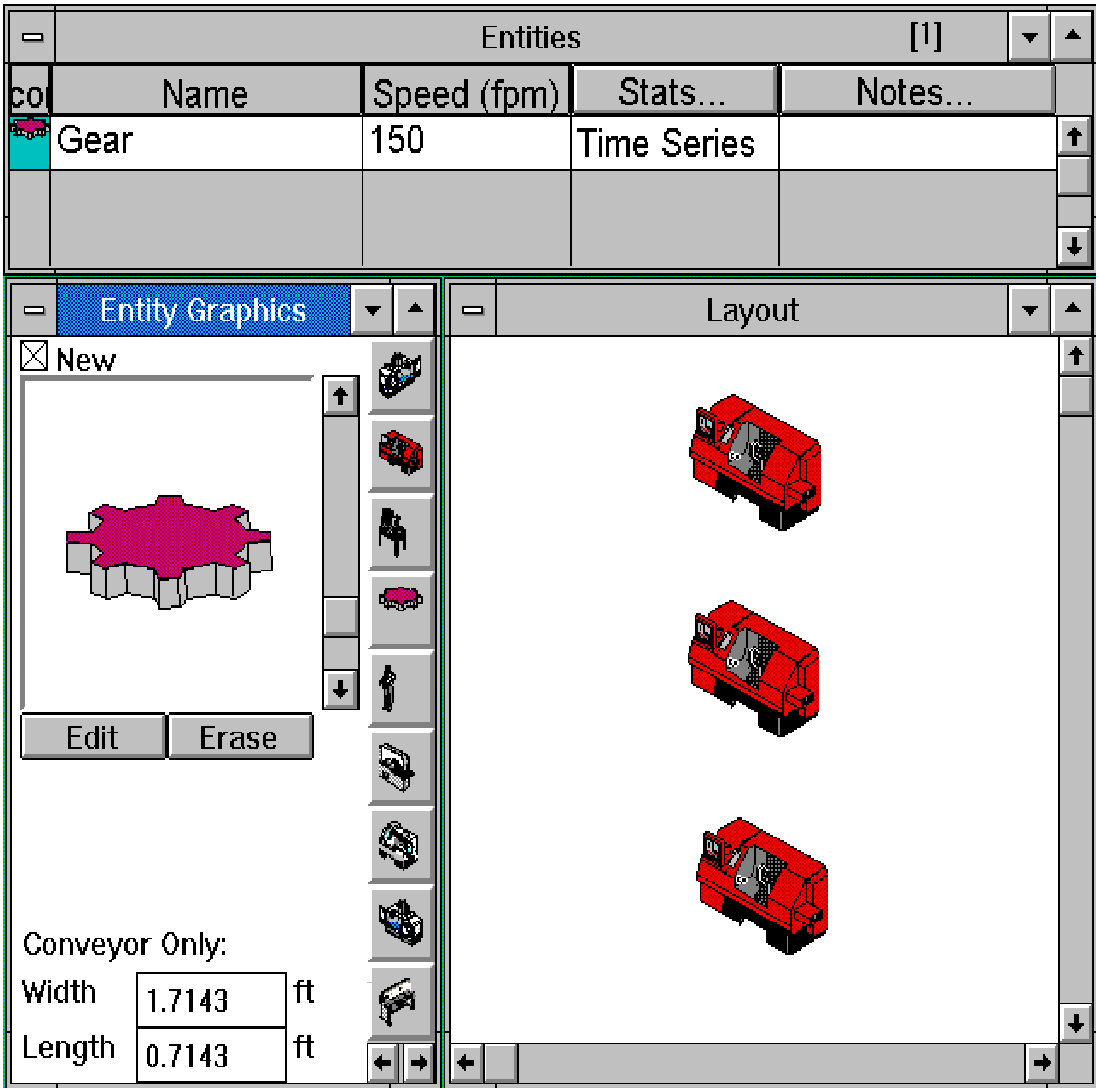
The development of the simplest model using the *ProModel* programming environment requires, at least, the use of *Locations*, *Entities*, *Arrivals*, and *Processing*, using the *LEAP* paradigm.

Locations, which may be single or multiple capacity, are generally fixed places in the system where entities are routed for processing, queuing, or making some decision about further routing. Examples include workstations, check-in-points, queues, storage areas (buffers), etc. *Locations* may have certain times that are available; may also have special input and output, such as, input based on highest priority, or output based on First-In-First-Out (FIFO). An example of three lathe machines used as *locations* from *ProModel* graphic library is shown in depiction (a) below.



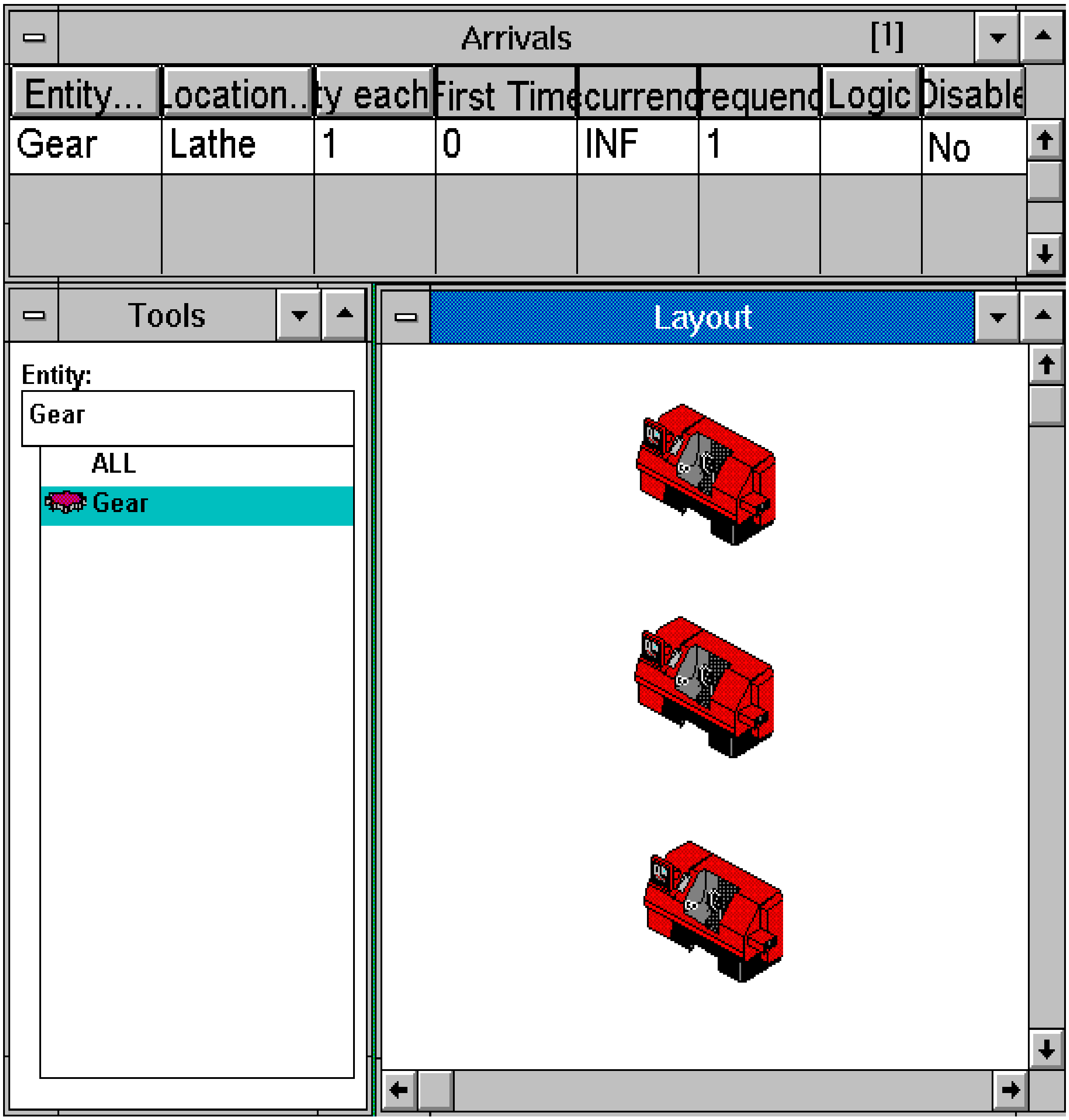
(a). *ProModel Locations* graphic library.

Entities are the objects processed in the model, that represent the inputs and outputs of the system; they enter into, flow through and depart from the system as complete objects, such as, parts, products, people, or even defect arrivals and failure occurrences. *Entities* may have special characteristics, such as, speed, size, etc.; may arrive from outside; or may also be created from within the system. An example of a 'gear' used as an *entity* from *ProModel* library of graphics is shown in depiction (b) below.



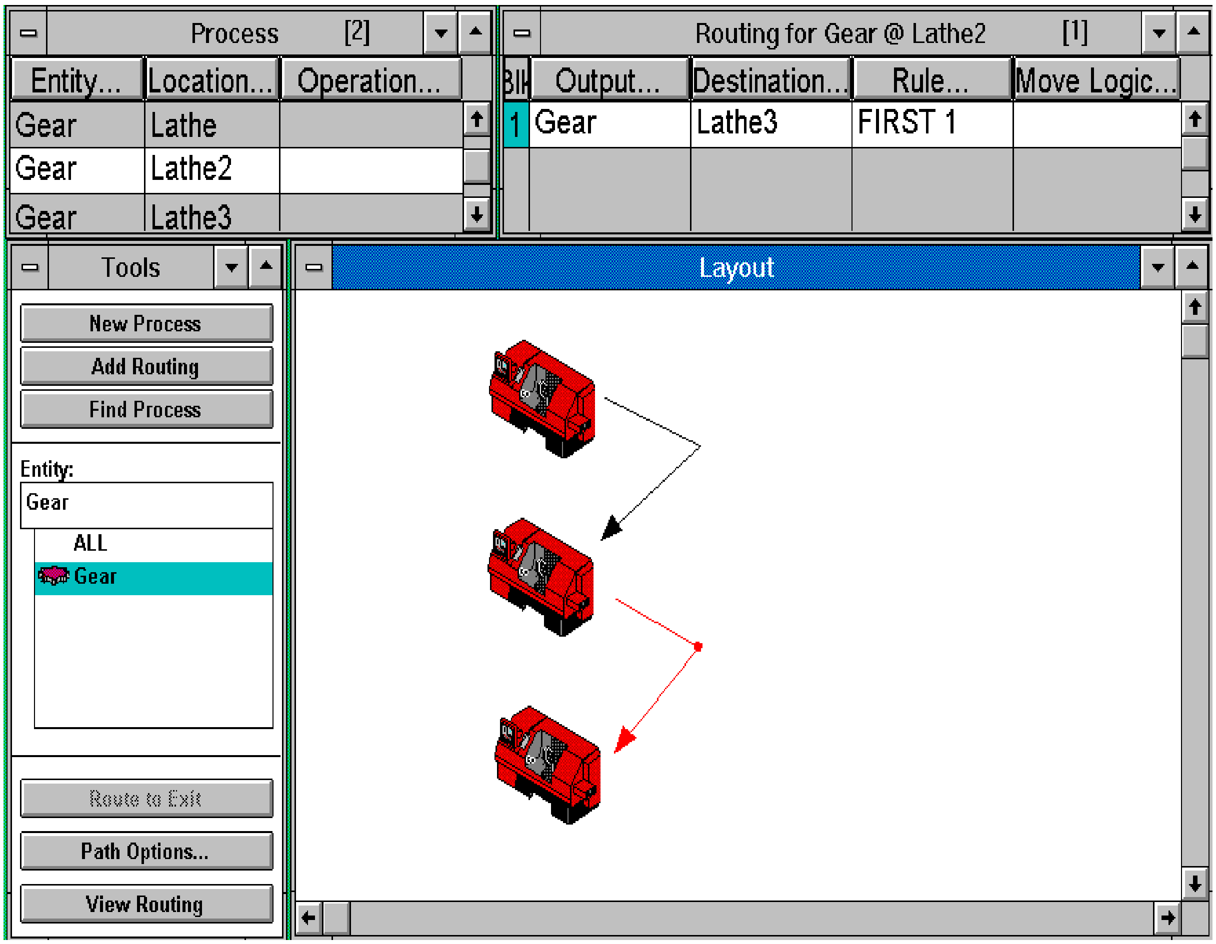
(b). *ProModel entities* graphic library.

Arrivals describe the precise pattern: timing; quantity; frequency; and *location* of *entities* entering into the system. Examples include: parts arriving to a manufacturing shop; customers arriving at a post office; passengers arriving at an airport; and defects arriving at a plant. An example of 'gears' arriving at the lathe machine, one at a time, from the start of simulation, every one minute, indefinitely, is shown in the top table in screen capture (c) below.



(c). ***ProModel arrivals***.

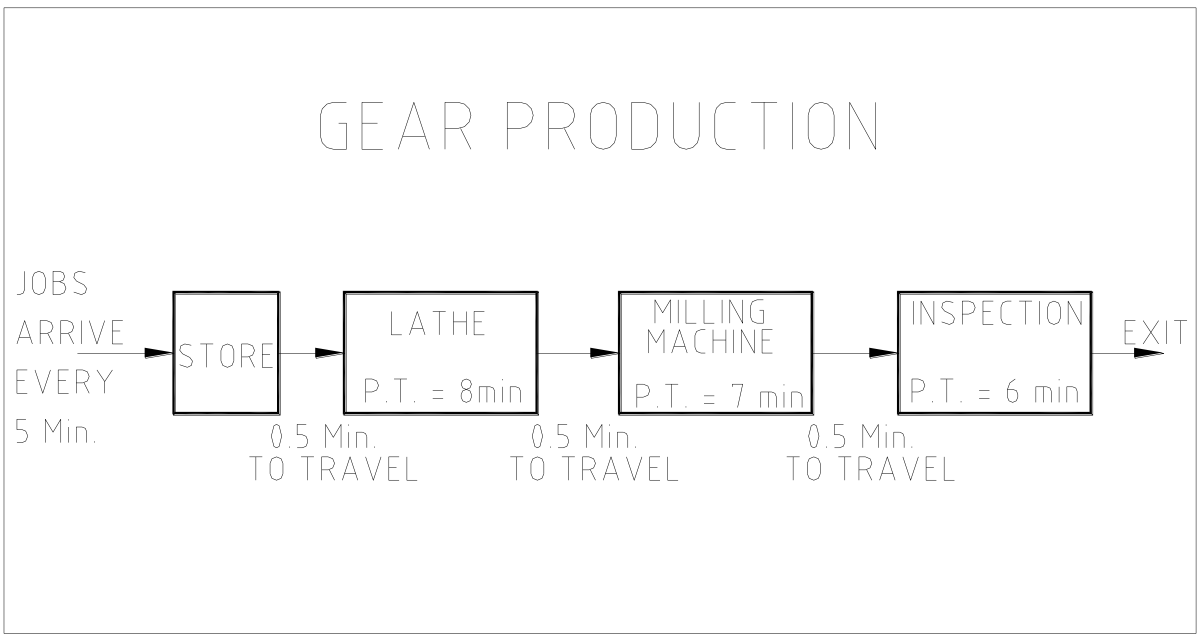
Finally, *Processing* defines the exact route that an *entity* follows, from entering into, to leaving the system. This includes any activity that happens at a *location*, such as, the required *operation(s)* that needs to be performed; the amount of time an *entity* spends at a *location*; the *resources* it needs to complete *processing*; and the selection of the *entity*'s next destination. An example of 'gears' being routed through different lathe machines is shown in the top two tables in screen capture (d) below.



(d). *ProModel processing*.

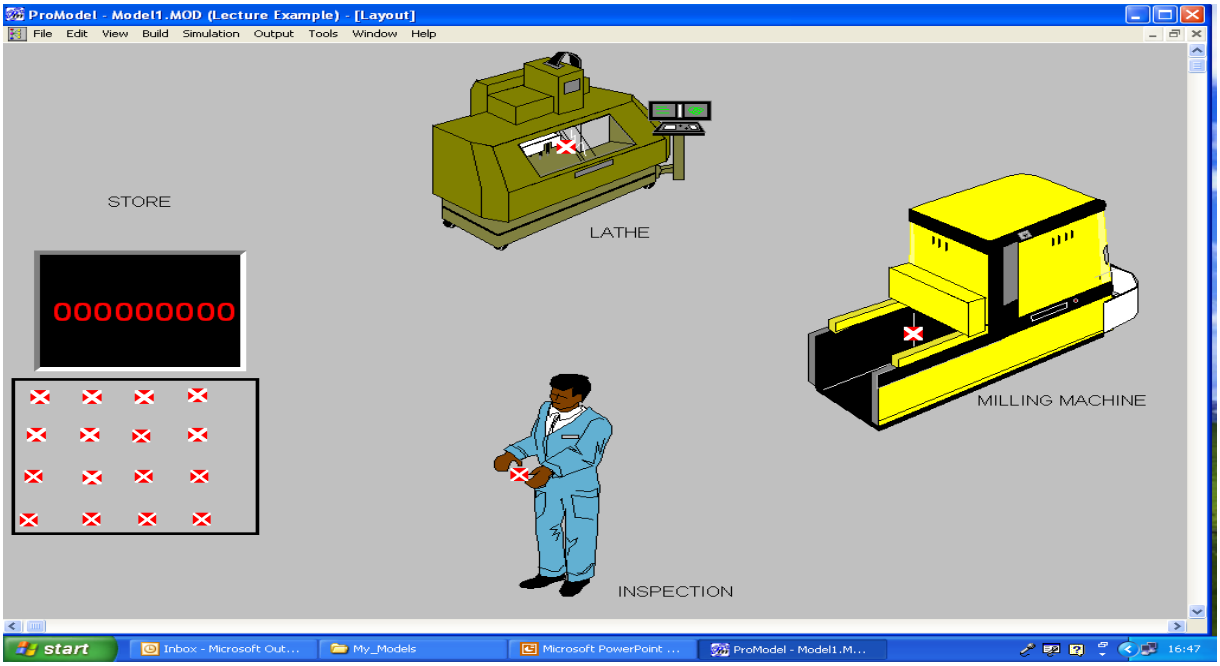
Although, the most simple model in this environment needs to have *LEAP* described, any further sophistication needed almost certainly will require the use of other *modules*, such as, *Path Networks*, *Resources*, *Variables*, *Attributes*, *Arrays*, *Subroutines*, *Shifts*, and the development of special *programming routines*.

To develop a simple model, imagine a gear production scenario as depicted in the screen capture (e) below. Round blank metals arrive into the store, one at a time, every five minutes, indefinitely. When production starts, blank metals are taken to the lathe machine to be machined for eight minutes. Once machined, turned parts go to the milling machine where they are milled for seven minutes. Gears from the milling machine are then taken to inspection, to be tested and measured which will take 6 minutes. Gears are finally dispatched as ready products. The travelling time from one machine to another, from the store all the way to exit takes 0.5 minute each. The objective of the simulation is to determine the production rate of gears in every eight-hour shift.

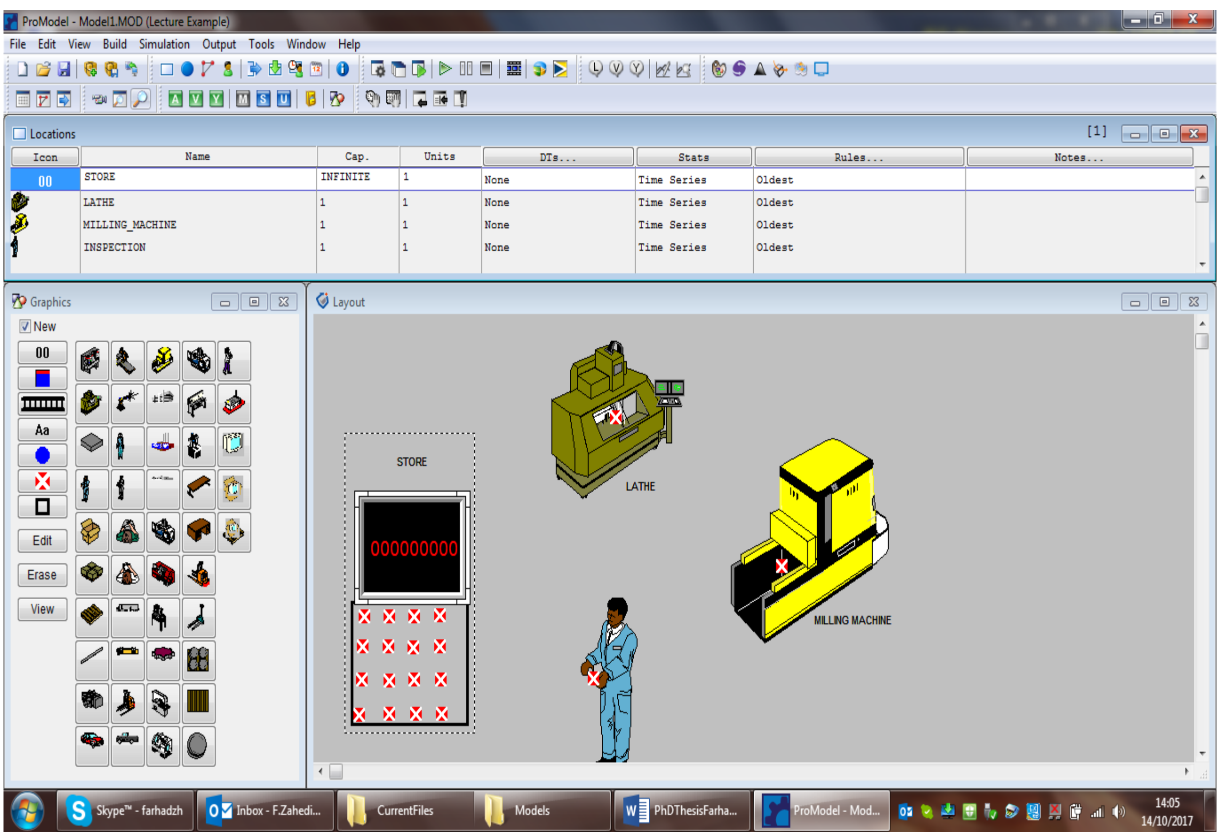


(e). Gear production scenario.

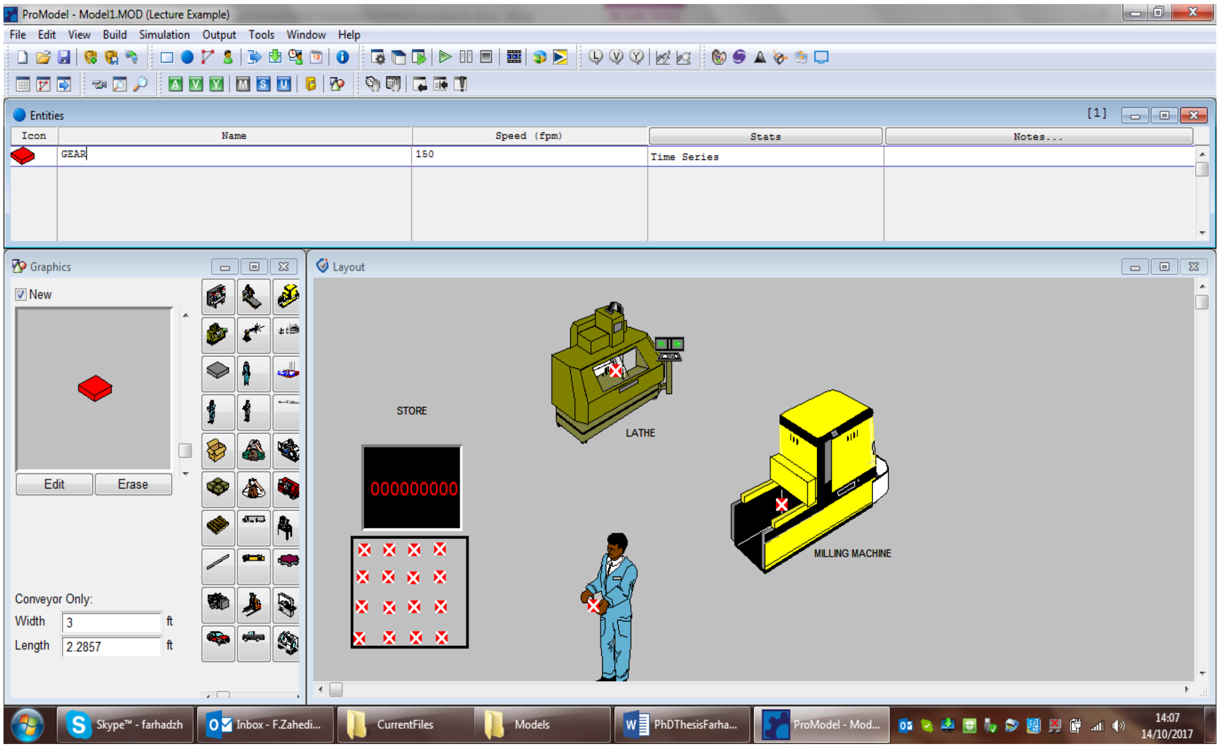
Following the procedure described above (shown in captions (a) to (d)) for developing *LEAP*, a simulation model for the gear production scenario is developed. The screen capture (f) below shows the layout of the simulation model for the gear production. Screen captures (g), (h), (i), and (j) depict the *Locations*, *Entities*, *Arrivals* and *Processing* respectively, which are developed for the gear production model. The completed model is then run for an eight-hour shift, for which the output results, depicted in screen capture (k), shows that 54 gears are produced.



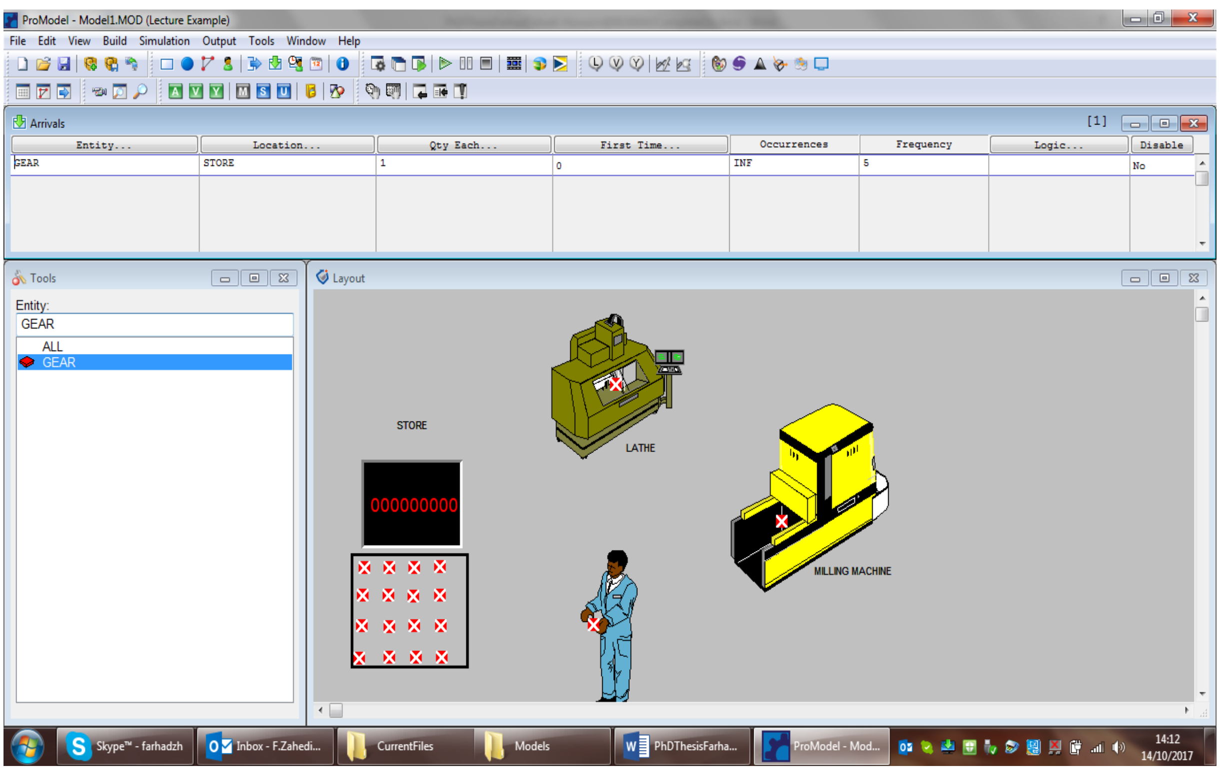
(f). *ProModel* layout for the gear production.



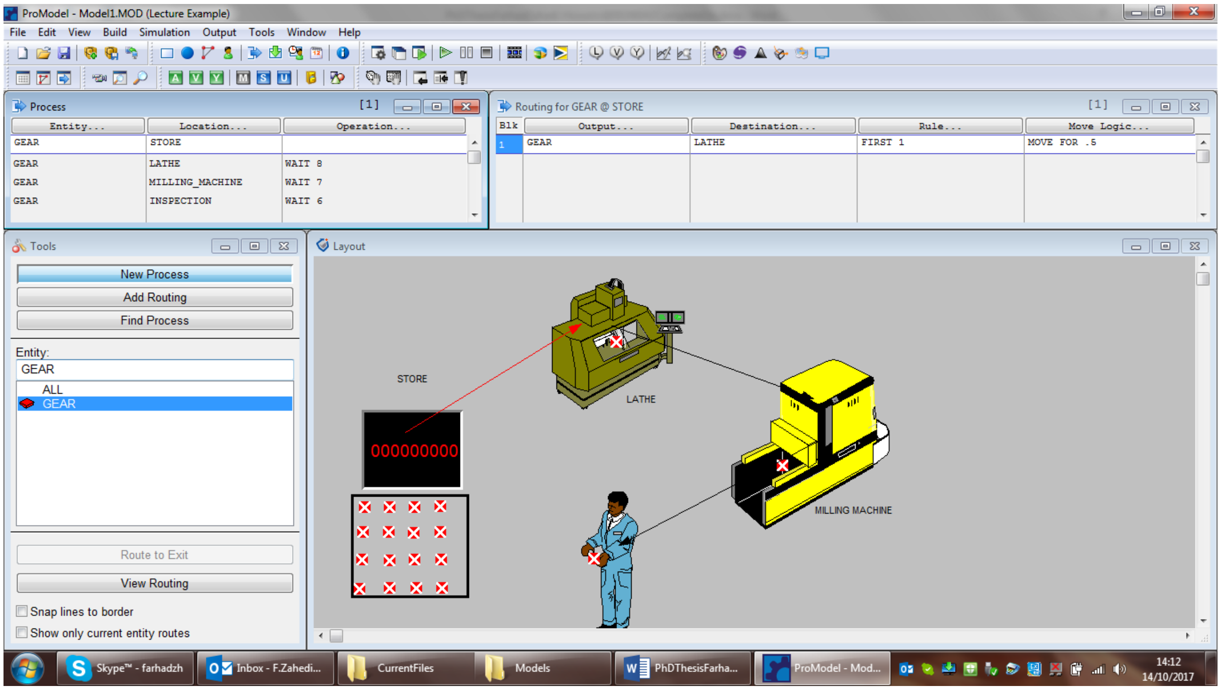
(g). *Locations* for the gear production model.



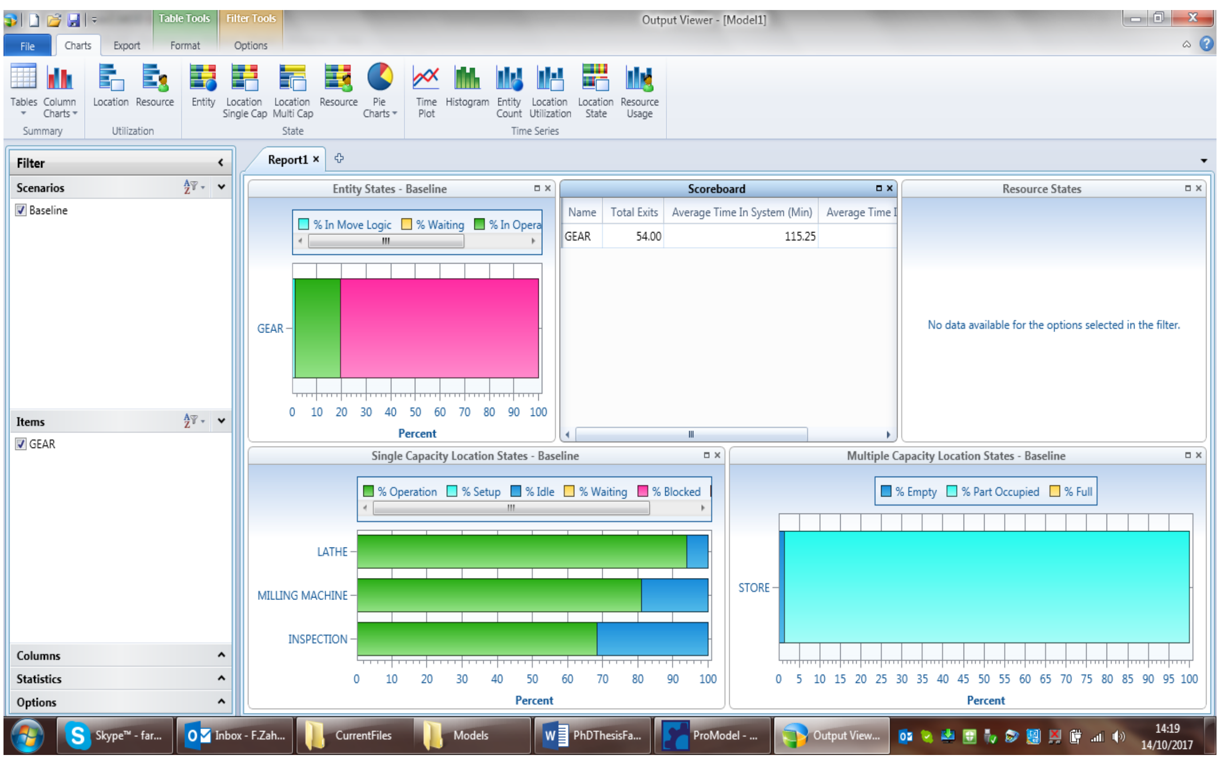
(h). *Entities* for the gear production model.



(i). *Arrivals* for the gear production model.



(j). *Processing* for the gear production model.



(k). Output results for the gear production model.

Appendix 3.1. A summary of the modelling framework specifically for the modified two-out-of-three parallel system with consecutive inspection prioritising failure repair:

(a) Planned maintenance algorithms.

A summary of the simulation algorithms for planned maintenance (i.e. machines going 'down' due to scheduled inspections) called 'Clock downtimes for Locations' is detailed below:

Start of algorithms

Repeat every PM interval time unit, starting from a simulation clock equal to PM interval time unit

Step 1: If status of M_i is 'down', wait until status is 'operational'

Step 2: Increment number of inspections, I

Step 3: Initiate the routine for removing any defects waiting in their *delay-time*

Step 4: Set status to 'on' for removing defects

Step 5: Set Downtime switch

Step 6: Set Arrays to hold defect and failing entities in 'time suspension' during inspection

Step 7: Delay by d_s

Step 8: Update total inspection duration by d_s

Step 9: Compute and update total cost, C_{mod}

Step 10: Set status to 'off'

Step 11: Increment total number of inspections I, for M_i

Step 12: Execute subroutine for calculating Downtime

Step 13: Decrement current total number of lines Down

End of algorithms

(b) Unplanned maintenance algorithms.

A summary of the simulation algorithms for unplanned maintenance (i.e. machines going ‘down’ due to component failure(s) named ‘Called downtimes for Locations’ is detailed below:

Start of algorithms

Step 1: Increment number of failures, F

Step 2: Set status to ‘on’ for removing defect

Step 3: Set Downtime switch

Step 4: Set Arrays to hold defect and failing entities in ‘time suspension’ during this process

Step 5: Delay by d_f

Step 6: Accumulate total failure duration by d_f

Step 7: Compute and update total cost, C_{mod}

Step 8: Set status to ‘off’

Step 9: Increment total number of failures F, for M_i

Step 10: Execute subroutine for calculating Downtime

Step 11: Decrement current total number of lines Down

End of algorithms

(c) Defect arrival algorithms.

Start of algorithms

Step 1: Set local variables

Step 2: Increment number of defects generated for M_i

Step 3: Generate a Random value

Step 4: Compute defect time lapse duration based on the time-to-defect distribution

Step 5: Record the time computed in the above step into an Array

Step 6: Delay defect by the time lapse duration

Step 7: In case there are concurrent inspections or failure stoppage processes in operation, execute the 'Prolonging' subroutine in order to 'suspend' the arrival of other defects and keep in 'suspension' the existing defects waiting in their '*delay-time*' evolving into failures. This is for implementing the 'suspension of component aging' concept

Step 8: If there is an ongoing machine downtime process, then

Step 8.1: Delay by d_f

Step 9: If there is a concurrent inspection process, then

Step 9.1: Compute the time lapse depending upon the time needed for each defect to be removed at inspection

Step 9.2: Delay by the computed time generated above

Step 10: Increment number of defect arrivals for M_i

End of algorithms

(d) Failure occurrence algorithms.

Start of algorithms

Step 1: Set local variables

Step 2: Increment number of failures evolved for M_i

Step 3: Generate a Random value

Step 4: Compute *delay-time* based on its distribution

Step 5: Record the time computed in the above step into an Array

Step 6: Delay failure by the *delay-time* duration

Step 7: In case there are concurrent inspections or failure stoppage processes in operation, execute the 'Prolonging' subroutine in order to 'suspend' the arrival of other defects and keep in 'suspension' the existing defects waiting in their '*delay-time*' evolving into failures. This is for implementing the 'suspension of component aging' concept

Step 8: If there is an ongoing machine downtime process, then

Step 8.1: Delay by d_f

Step 9: If there is a concurrent inspection process, then

Step 9.1: Compute the time lapse depending upon the time needed for each defect to be removed at inspection

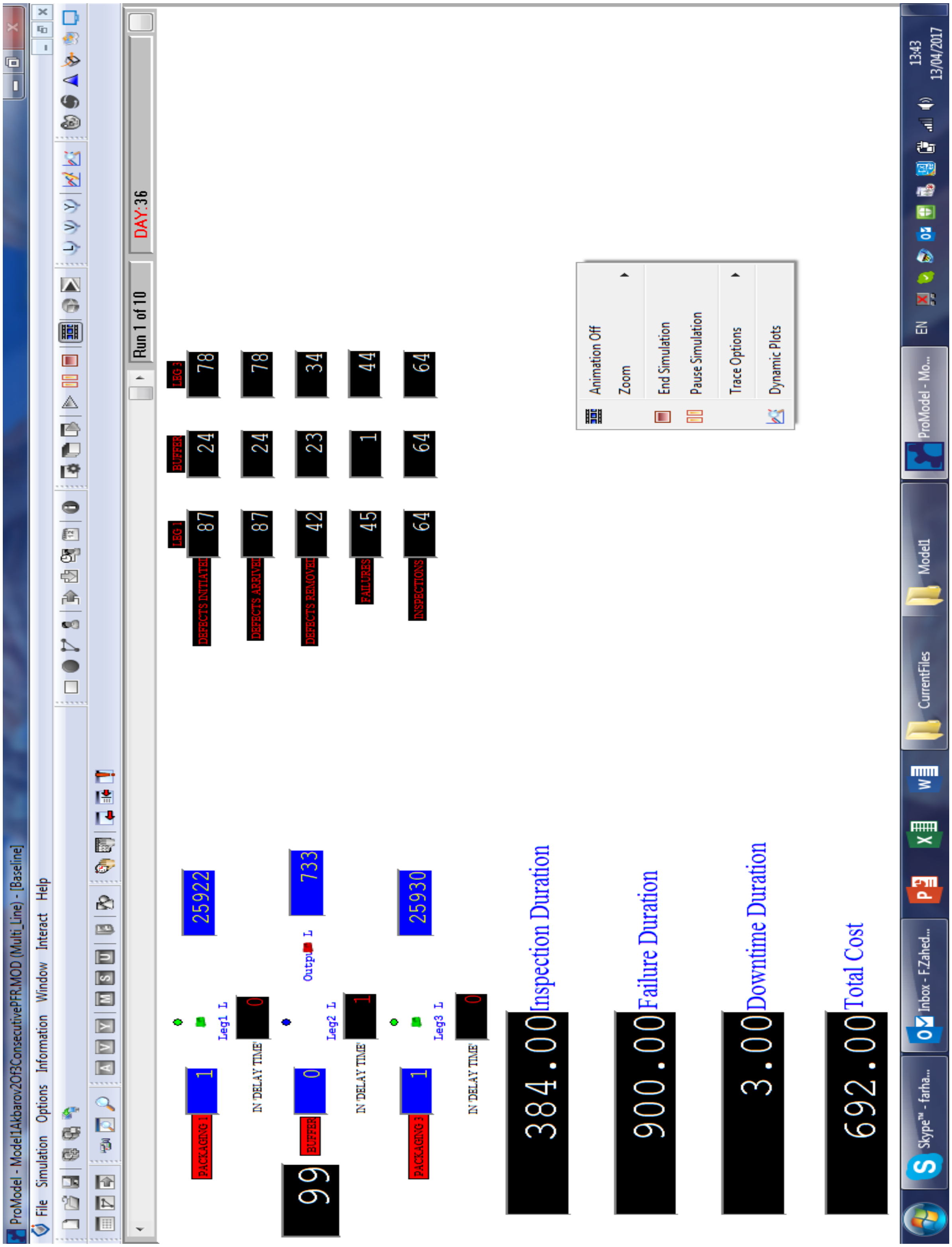
Step 9.2: Delay by the computed time generated above

Step 10: Increment number of failures for M_i

Step 11: Initiate machine Downtime

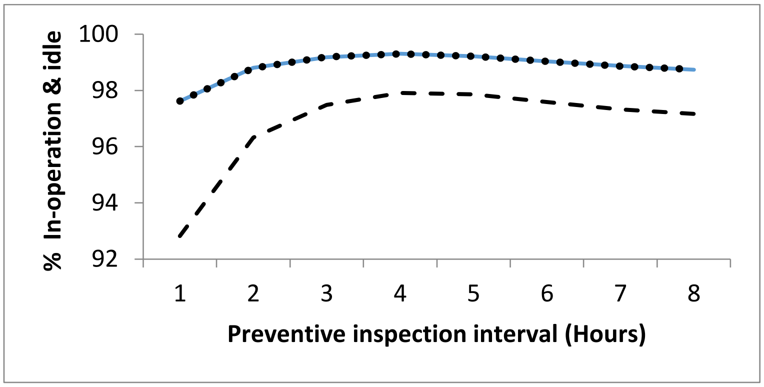
End of algorithms

Appendix 3.2. On-screen layout of simulation model.

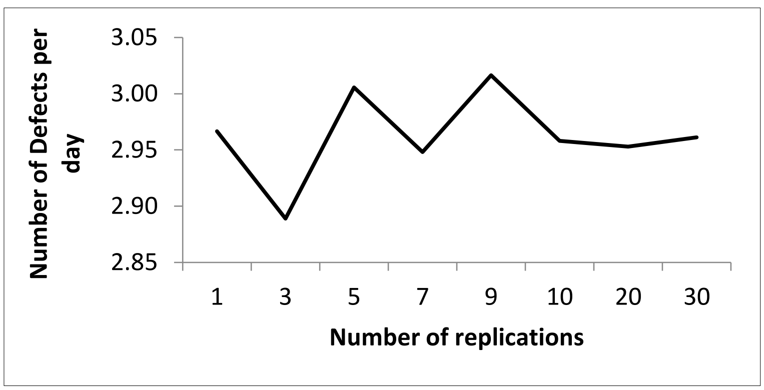


Appendix 3.3. Percentage time packing lines are either working or idle, under the modified 2-out-of-3 parallel system:

——— line 1 working; — — — line 2 idle; ●●● line 3 working.



Appendix 3.4. Average number of defects arriving at packing line 1.



Appendix 3.5. Detailed source data for the single-line packing system.

1,000 Simulation DAYS each							Per Day		
Inspection Interval (Hour)	No. of Defect Arrived	No. of failures	No. of inspections	No. of defects removed	Downtime (Minutes)	Downtime Per Day (Mins)	No. of failures	No. of inspections	No. of defects removed
1	2,965	0	23,809	2,965	47,618	47.6	0.0	23.8	3.0
2	2,984	2	12,048	2,982	24,112	24.1	0.0	12.0	3.0
3	2,992	25	8,000	2,966	16,204	16.2	0.0	8.0	3.0
4	2,995	146	5,988	2,849	13,230	13.2	0.1	6.0	2.8
5	2,992	448	4,808	2,544	13,671	13.7	0.4	4.8	2.5
6	2,990	821	4,000	2,169	15,765	15.8	0.8	4.0	2.2
7	2,983	1,138	3,425	1,845	17,837	17.8	1.1	3.4	1.8
8	2,977	1,333	3,003	1,643	18,982	19.0	1.3	3.0	1.6
9	2,977	1,542	2,666	1,435	20,404	20.4	1.5	2.7	1.4
10	2,974	1,676	2,398	1,297	21,292	21.3	1.7	2.4	1.3
11	2,971	1,793	2,184	1,178	22,043	22.0	1.8	2.2	1.2
12	2,970	1,893	2,000	1,077	22,689	22.7	1.9	2.0	1.1
13	2,970	1,971	1,845	998	23,175	23.2	2.0	1.8	1.0
14	2,965	2,039	1,715	926	23,624	23.6	2.0	1.7	0.9
15	2,963	2,093	1,600	870	23,911	23.9	2.1	1.6	0.9
16	2,962	2,158	1,499	804	24,387	24.4	2.2	1.5	0.8
17	2,962	2,194	1,412	768	24,572	24.6	2.2	1.4	0.8
18	2,961	2,237	1,333	724	24,875	24.9	2.2	1.3	0.7
19	2,960	2,285	1,263	675	25,238	25.2	2.3	1.3	0.7
20	2,960	2,323	1,200	637	25,472	25.5	2.3	1.2	0.6
21	2,959	2,354	1,142	605	25,681	25.7	2.4	1.1	0.6
22	2,959	2,352	1,090	606	25,574	25.6	2.4	1.1	0.6

Appendix 3.6. Detailed source data for the modified two-out-of-three parallel system.

	LEG1				LEG2				LEG3				SIMULTANEOUS INSPECTION		CONSECUTIVE INSPECTION	
Inspection Interval (Hour)	No. of Defect Arrived	No. of Failures	No. Of Inspections	No. of Defects Removed	No. of Defect Arrived	No. of Failures	No. Of Inspections	No. of Defects Removed	No. of Defect Arrived	No. of Failures	No. Of Inspections	No. of Defects Removed	Downtime per day (mins)	Cost per day (£) Simultaneous	Cost per day (£) Consecutive	Cost per day (£) Consecutive inspection, prioritising failure repair
1	2,986	0	23,809	2,986	917	0	23,809	917	2,960	0	23,809	2,960	47.62	865.06	71.43	71.43
2	2,966	2	12,048	2,965	938	0	12,048	938	2,998	1	12,048	2,997	24.10	437.76	36.21	36.16
3	3,004	30	8,000	2,973	930	0	8,000	930	2,981	26	8,000	2,955	16.00	290.95	24.62	24.25
4	2,985	149	5,988	2,836	934	0	5,988	934	3,007	146	5,988	2,861	11.98	219.09	**21.03**	19.51
5	2,994	448	4,807	2,545	942	0	4,807	942	2,994	443	4,807	2,551	9.64	179.61	22.06	**19.29**
6	2,984	825	4,000	2,159	940	1	4,000	939	3,006	826	4,000	2,179	8.08	154.97	25.27	21.74
7	2,979	1,124	3,424	1,855	962	2	3,424	960	2,987	1,131	3,424	1,856	6.98	137.93	26.61	23.73
8	3,002	1,352	3,003	1,651	966	3	3,003	963	2,959	1,343	3,003	1,616	6.17	125.27	27.80	25.73
9	2,984	1,554	2,667	1,431	965	9	2,667	957	2,978	1,526	2,667	1,452	5.55	116.01	29.04	27.12
10	2,974	1,680	2,399	1,295	984	16	2,399	968	2,969	1,691	2,399	1,278	5.05	108.23	30.10	27.52

Appendix 4.1. Survey questionnaire:

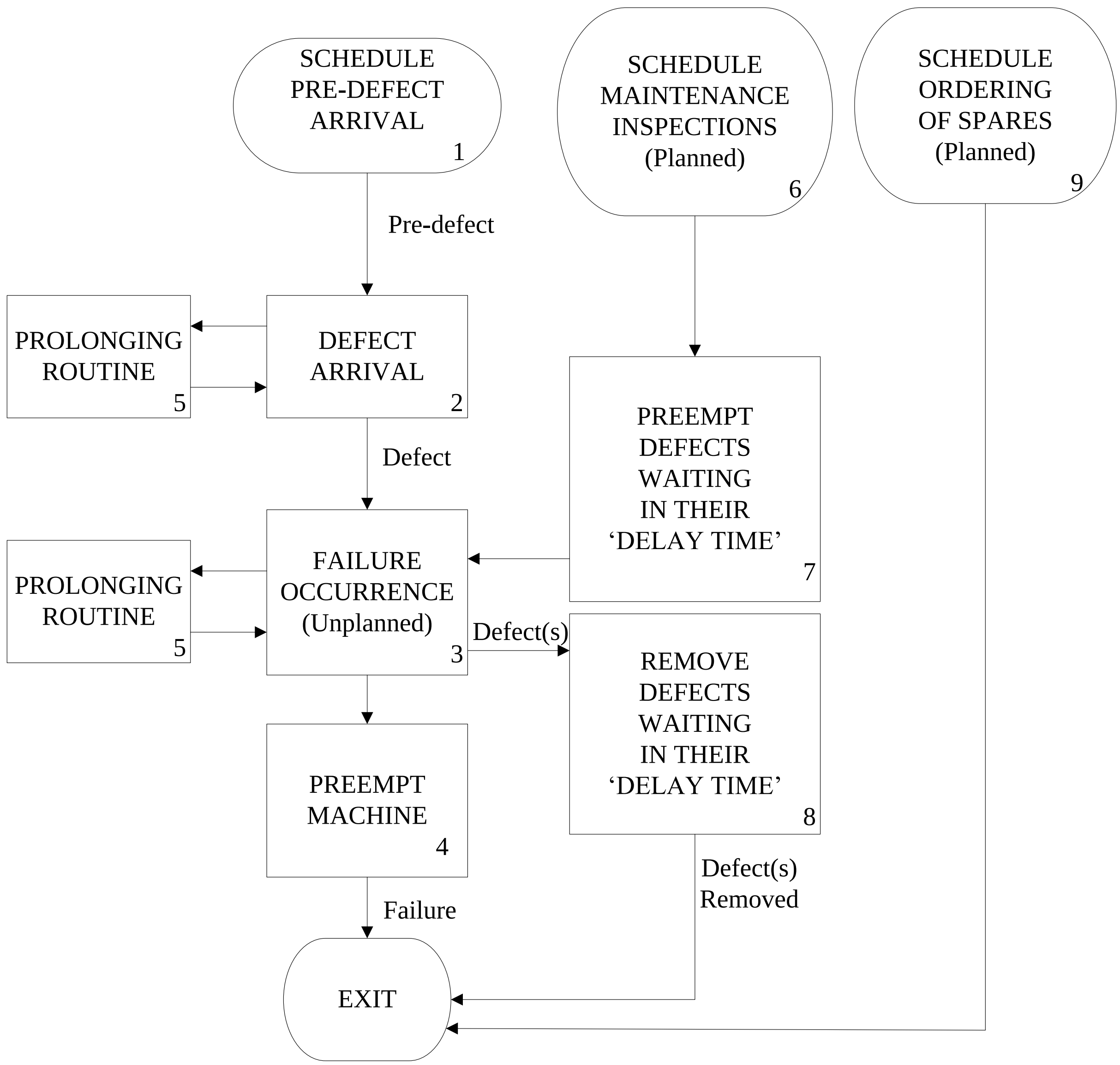
(a) Collecting data from maintenance experts and paper machine manufacturers.

Information	Explanation	Value
Unit cost of the critical spare part	What is the unit cost of the critical spare part (£)? Although the model can be expanded in the future, our current model assumes that there is one critical part that when it fails will result in the immediate downtime of the machinery/equipment.	
Inspection	How long does it take to carry out a Planned Maintenance inspection for the machinery/equipment, from start to finish, to identify if one or more units (instances) of the critical part need to be replaced (excluding the time it takes for replacing each critical part) (Sec/ Min/ Hour/ Day)?	
	How many operators/technicians (resources) are needed for carrying out the inspection activity, in order to determine if one or more critical spare part(s) need to be replaced?	
	What is the cost of using each operator/technician (£ per Hour/ Day)?	
Replacement	How long does it take to actually replace a single critical spare part after identifying, at inspection, that the part needs to be replaced (Sec/ Min/ Hour/ Day)?	
	How many operators/technicians are needed for the replacement of a unit (an instance) of the critical spare part?	
	If different from the data supplied in the 'inspection' section above, what is the cost of using each operator/technician (£ per Hour/ Day)?	
Failure replacement	How long does it take to replace a single unit (instance) of the critical spare part as a result of Corrective Maintenance, i.e. when the critical part fails (Sec/ Min/ Hour/ Day)?	
	How many operators/technicians are needed for the replacement of a single unit of the critical part?	
	If different from the data supplied in the 'inspection' section above, what is the cost of using each operator/technician (£ per Hour/ Day)?	
Downtime cost	How much does it cost the company (£ per Hour/Day) for the loss of production (or sales) while the machinery/equipment is down due to one of the following reasons? • the inspection process • the replacement of a critical part during Preventive maintenance • the replacement of a critical part during Corrective maintenance • waiting for shortages to arrive	
Ordering cost	How much does it cost to order the critical spare part during normal production times (not ordering shortages), excluding the unit cost of the part (£)?	
Holding cost	How much does it cost to hold a single unit of inventory of the critical spare part in stock (£ per unit per Day/Week/Month/Year)?	
Order Delivery lead time	How long does it take for an order of the critical spare part to arrive, from the time that an order is placed until it arrives (Day/Week)?	
	How variable is this lead time? Can this be estimated, or does it conform to a known statistical distribution. For example, conforming to a Normal distribution, with a mean lead time of 7 days and a standard deviation of 1 day!	
Shortage Shipment cost	How much extra (excluding the unit cost) does it cost to replenish a critical spare part in emergency (£)?	
Shortage Delivery Lead-time	How long does it take for a shortage (or emergency) order of the critical spare part to arrive, from the time that the order is placed until the time it arrives (Day/Week)?	
Faults (defects) distribution & parameters	Do you know how often faults related to the critical spare part occur? Does the pattern of occurrences follow a known statistical distribution? Alternatively, can you provide a time-series history of fault arrivals?	
Failure Distribution & parameters	Do you know if the time between failures of the critical part follow a known statistical distribution?	
Currently, how often is the machinery/equipment inspected in connection with this critical spare part?		
What is the current policy for replenishing the inventory of this critical spare part? Periodic or Continuous Review?		

(b) Survey response values.

Category	Questions	Response values
Unit cost of the critical spare part	What is the unit cost of the critical spare part (£)? Although the model can be expanded in the future, our current model assumes that there is one critical part that when it fails will result in the immediate downtime of the machinery/equipment.	£1,000-4,000.
Inspection	How long does it take to carry out a Planned Maintenance inspection for the machinery/equipment, from start to finish, to identify if one or more units (instances) of the critical part need to be replaced (excluding the time it takes for replacing each critical part) (Sec/ Min/ Hour/ Day)?	Depends and varies.
	How many operators/technicians (resources) are needed for carrying out the inspection activity, in order to determine if one or more critical spare part(s) need to be replaced?	
	What is the cost of using each operator/technician (£ per Hour/ Day)?	£60 per hour.
Replacement	How long does it take to actually replace a single critical spare part after identifying, at inspection, that the part needs to be replaced (Sec/ Min/ Hour/ Day)?	1-6 hours.
	How many operators/technicians are needed for the replacement of a unit (an instance) of the critical spare part?	3 technicians.
	If different from the data supplied in the 'inspection' section above, what is the cost of using each operator/technician (£ per Hour/ Day)?	£60 per hour.
Failure replacement	How long does it take to replace a single unit (instance) of the critical spare part as a result of Corrective Maintenance, i.e. when the critical part fails (Sec/ Min/ Hour/ Day)?	1-36 hours.
	How many operators/technicians are needed for the replacement of a single unit of the critical part?	3 technicians.
	If different from the data supplied in the 'inspection' section above, what is the cost of using each operator/technician (£ per Hour/ Day)?	£60 per hour.
Downtime cost	How much does it cost the company (£ per Hour/Day) for the loss of production (or sales) while the machinery/equipment is down due to one of the following reasons? • the inspection process • the replacement of a critical part during Preventive maintenance • the replacement of a critical part during Corrective maintenance • waiting for shortages to arrive	£1,000 per hour.
Ordering cost	How much does it cost to order the critical spare part during normal production times (not ordering shortages), excluding the unit cost of the part (£)?	£100.
Holding cost	How much does it cost to hold a single unit of inventory of the critical spare part in stock (£ per unit per Day/Week/Month/Year)?	1% of item cost per week.
Order Delivery lead time	How long does it take for an order of the critical spare part to arrive, from the time that an order is placed until it arrives (Day/Week)?	2-6 weeks.
	How variable is this lead time? Can this be estimated, or does it conform to a known statistical distribution. For example, conforming to a Normal distribution, with a mean lead time of 7 days and a standard deviation of 1 day!	
Shortage Shipment cost	How much extra (excluding the unit cost) does it cost to replenish a critical spare part in emergency (£)?	£500-1,200.
Shortage Delivery Lead-time	How long does it take for a shortage (or emergency) order of the critical spare part to arrive, from the time that the order is placed until the time it arrives (Day/Week)?	1-10 days.
Faults (defects) distribution & parameters	Do you know how often faults related to the critical spare part occur? Does the pattern of occurrences follow a known statistical distribution? Alternatively, can you provide a time-series history of fault arrivals?	Depends on quality.
Failure Distribution & parameters	Do you know if the time between failures of the critical part follow a known statistical distribution?	Varies.

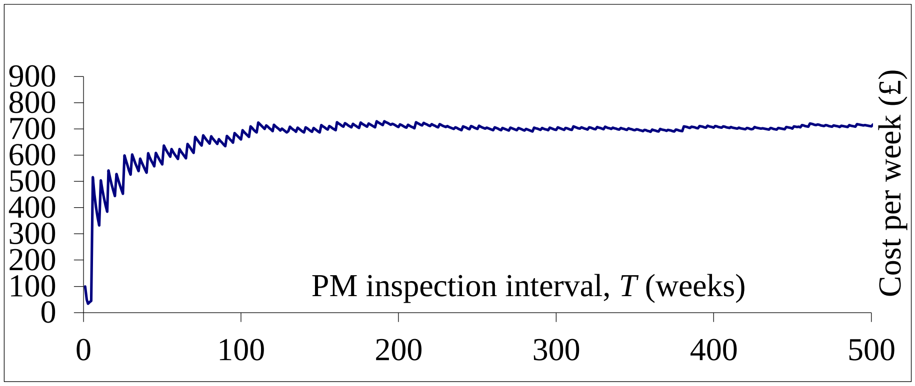
Appendix 4.2. Flowchart of the general simulation procedure, showing the flow of entities from one modelling routine to another.



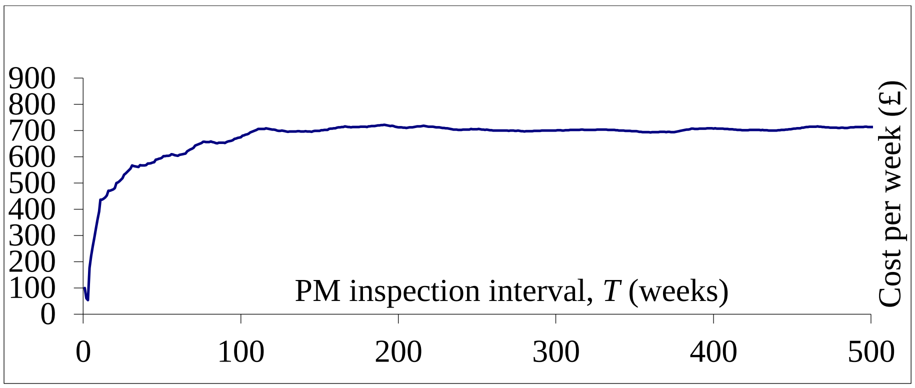
Appendix 4.3. Determination of the simulation *warm-up* period using:

(a) the Time Series method; (b) Welch's method; and

(c) sample analysis showing the number of *replications* needed to achieve a 95% confidence interval in the results.



(a)

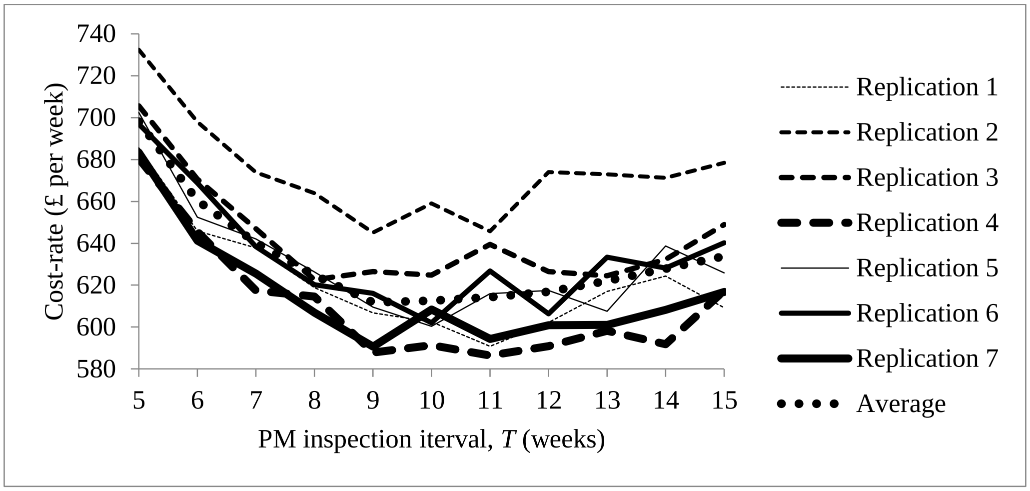


(b)

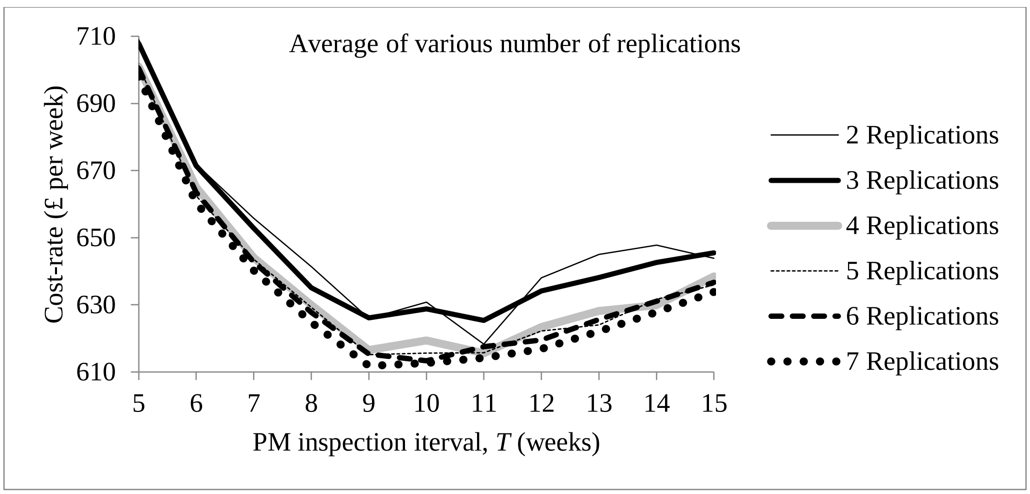
(c)

				Significance level	5.0%
				Confidence interval	
Replication	**Result**	**Cum. mean average**	**Standard deviation**	**Lower limit**	**Upper limit**
1	711.39	711.39	n/a	n/a	n/a
2	711.06	711.23	0.233	709.14	713.32
3	710.74	711.06	0.328	710.25	711.88
4	711.50	711.17	0.344	710.63	711.72
5	711.17	711.17	0.298	710.80	711.54
6	710.84	711.12	0.299	710.80	711.43
7	710.51	711.03	0.356	710.70	711.36
8	710.19	710.92	0.445	710.55	711.30
9	710.60	710.89	0.429	710.56	711.22
10	710.28	710.83	0.448	710.51	711.15

Appendix 4.4. A sample analysis for the determination of number of replications.

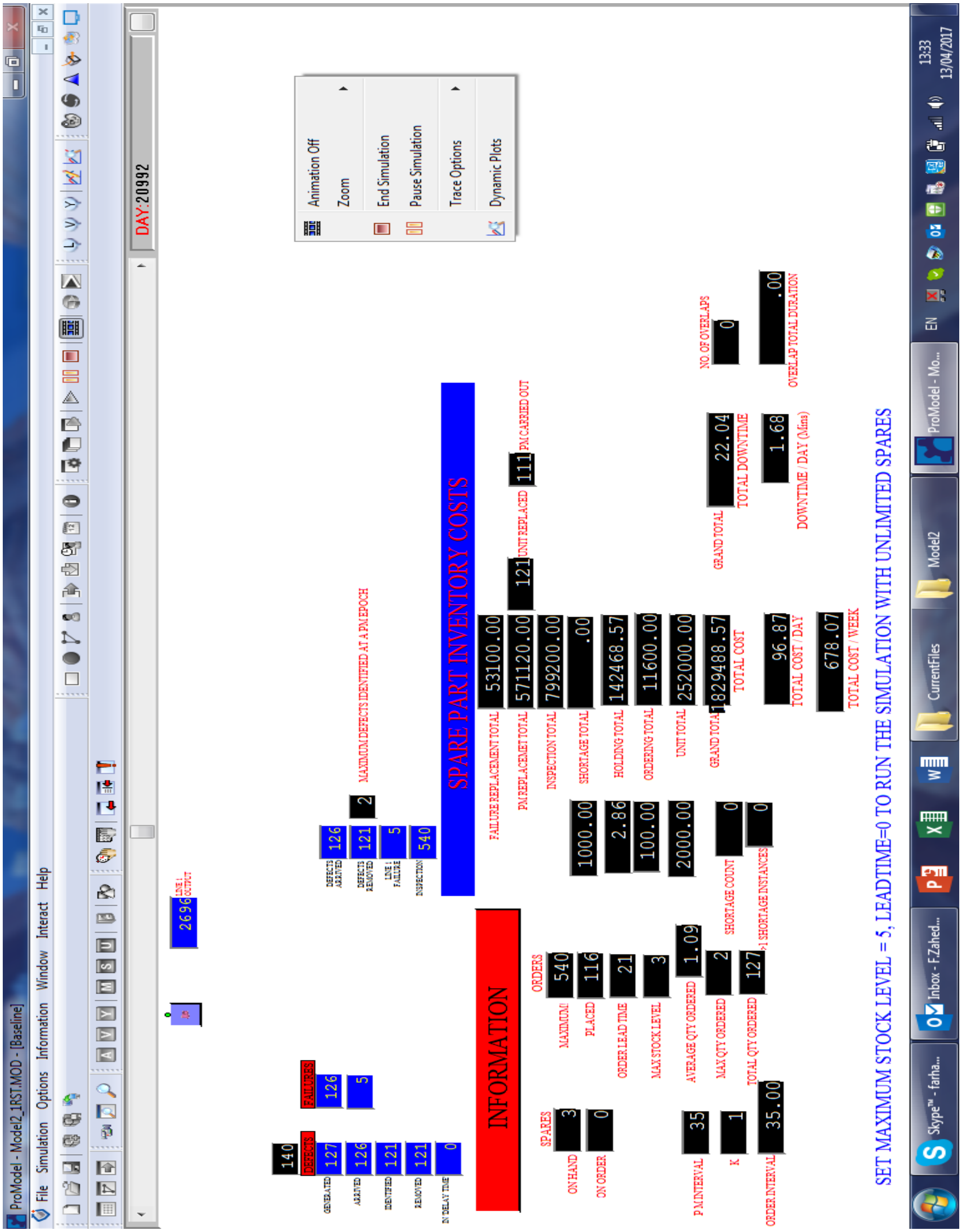


(a)

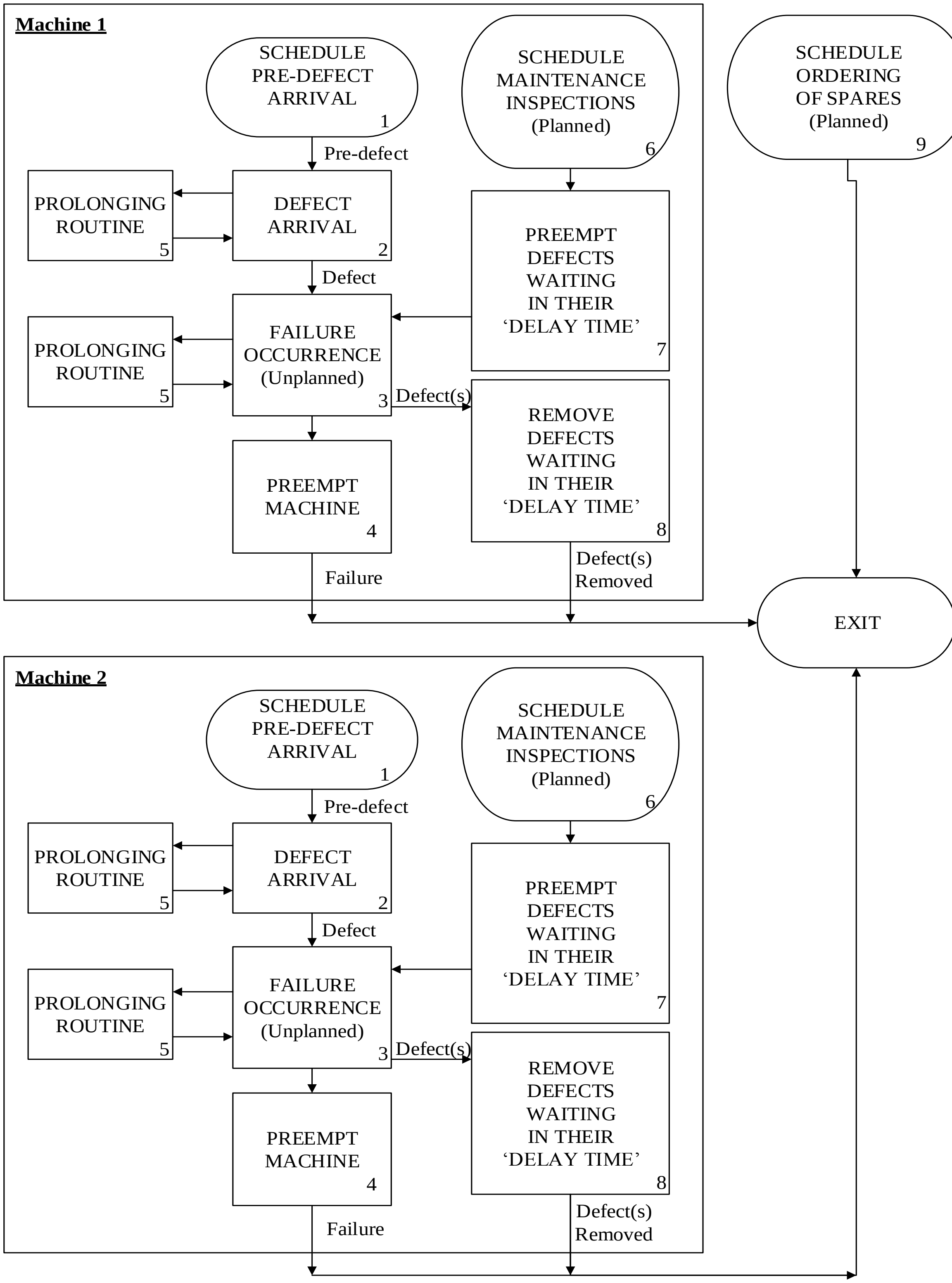


(b)

Appendix 4.5. On-screen layout of simulation model.



Appendix 5.1. Flowchart of the general simulation procedure, showing the flow of entities from one modelling routine to another.



Appendix 5.2. Details of the model: (a) *locations*; and (b) *entities*.

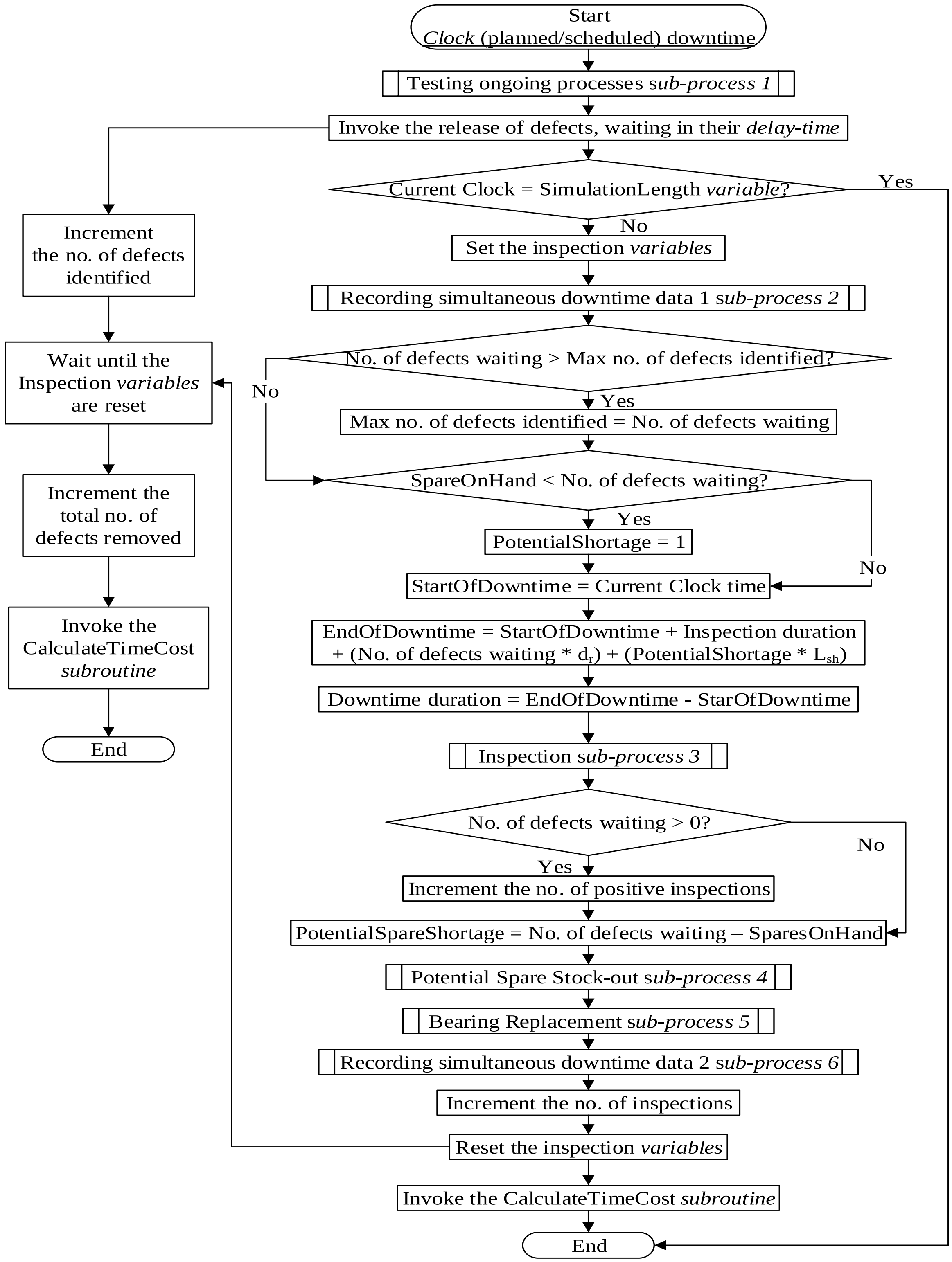
(a)

Location	**Capacity**	**Units**	**Downtimes**
MC1_MachineProcess	1	1	*Clock* – see Appendices 5.3 to 5.5.
MC1_DownstreamProcess	1	1	None
MC1_DefectArrival	Infinite	1	None
MC1_FailureOccurrence	Infinite	1	*Called* – see Appendices 5.6 & 5.7.
MC1_DefectDump	Infinite	1	None
MC1_Reset	Infinite	1	None
MC1_HoldingCost	Infinite	1	None
MC1_PlaceOrders	Infinite	1	None
MC2_MachineProcess	1	1	*Clock* – see Appendices 5.3 to 5.5.
MC2_DownstreamProcess	1	1	None
MC2_DefectArrival	Infinite	1	None
MC2_FailureOccurrence	Infinite	1	*Called* – see Appendices 5.6 & 5.7.
MC2_DefectDump	Infinite	1	None
MC2_Reset	Infinite	1	None
MC2_HoldingCost	Infinite	1	None
MC2_PlaceOrders	Infinite	1	None

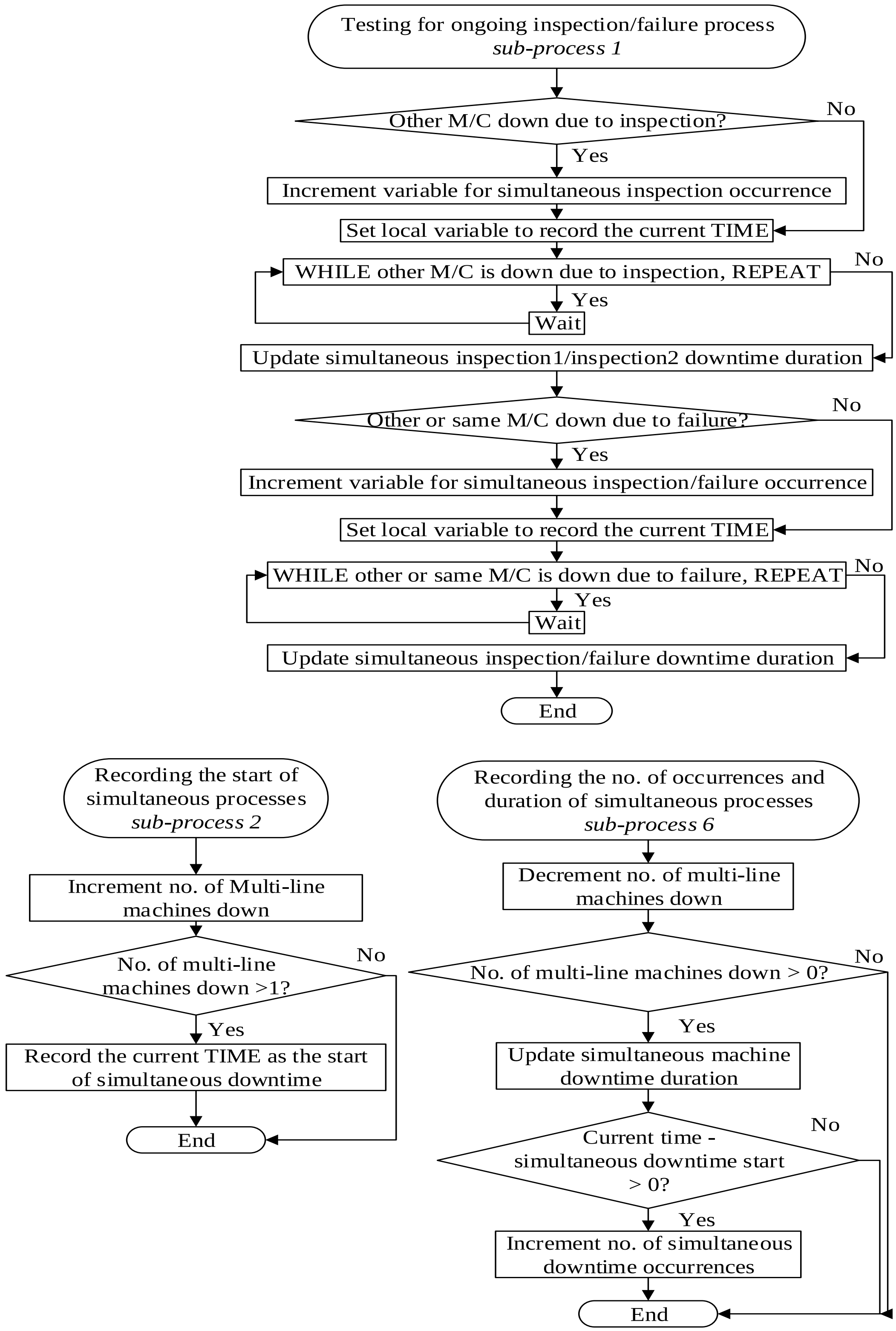
(b)

Entity	
Product2B	DefectIdentified
Product	Reset
ProductMade	HoldingCost
PreDefect	PlaceOrders
Defect	Calculate
Failure	

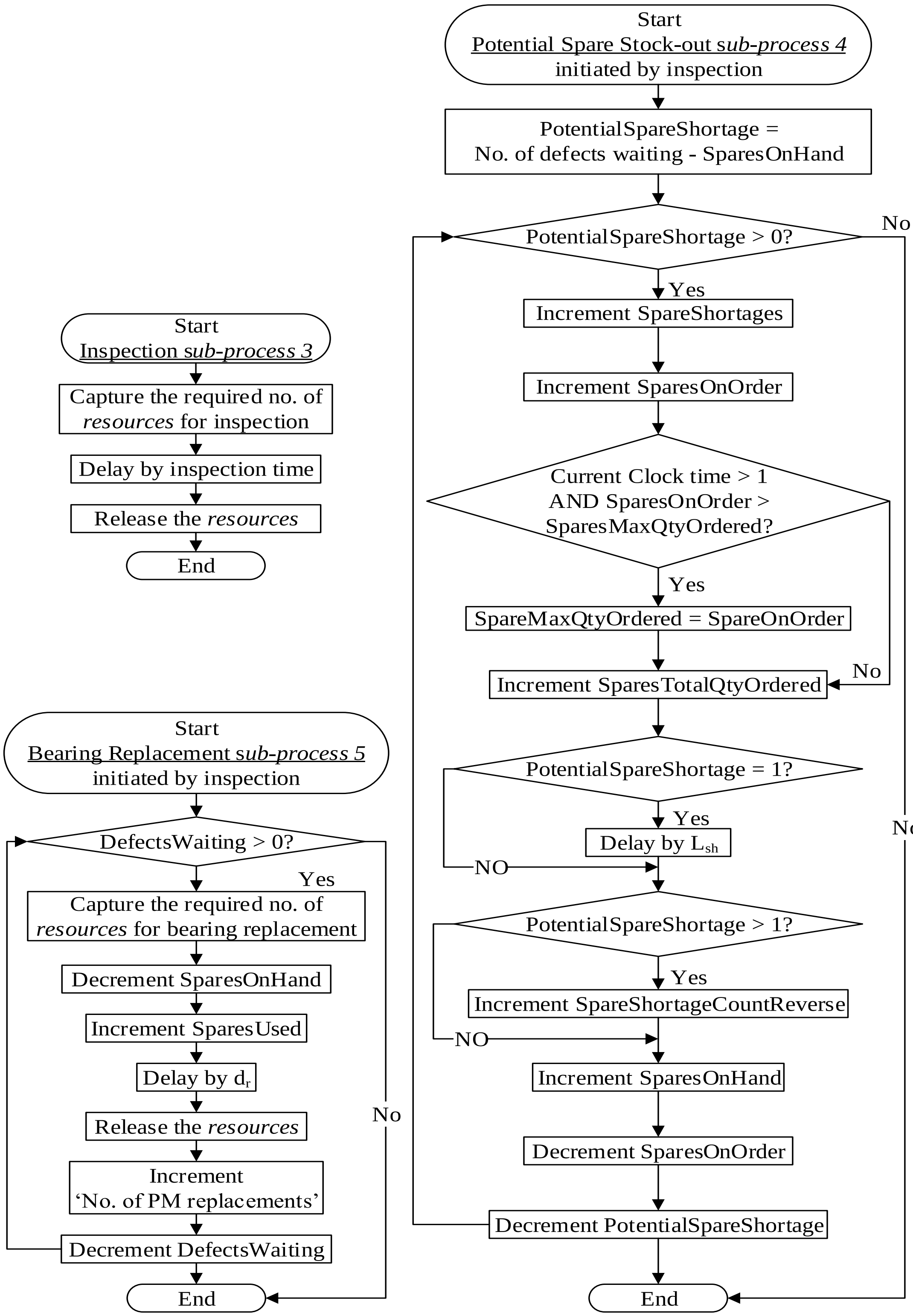
Appendix 5.3. Flowchart of the *Clock* (planned/scheduled) downtime routine for each machine.



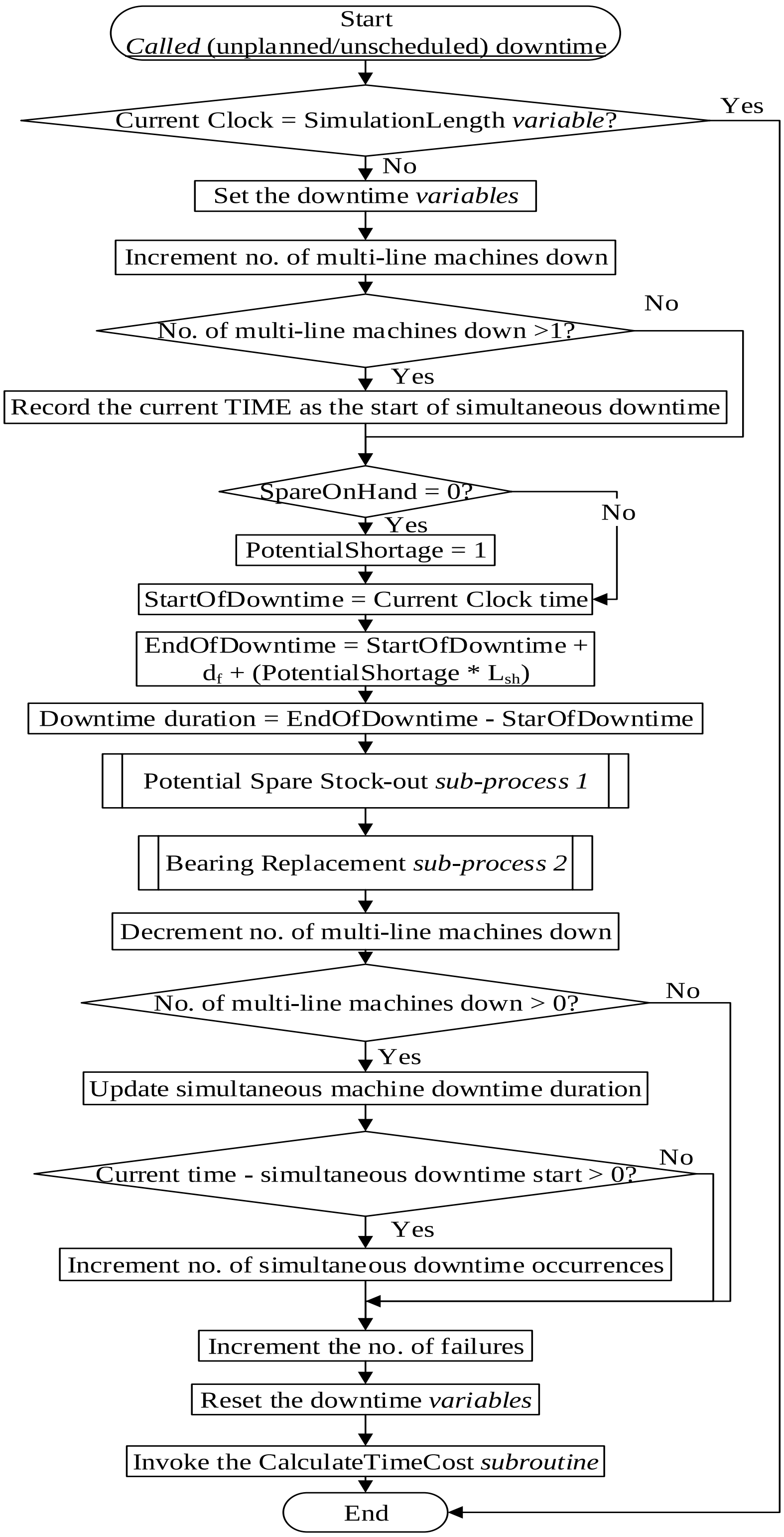
Appendix 5.4. Flowchart of the *Clock* (planned/scheduled) downtime *sub-processes 1, 2 & 6* routines for each machine.



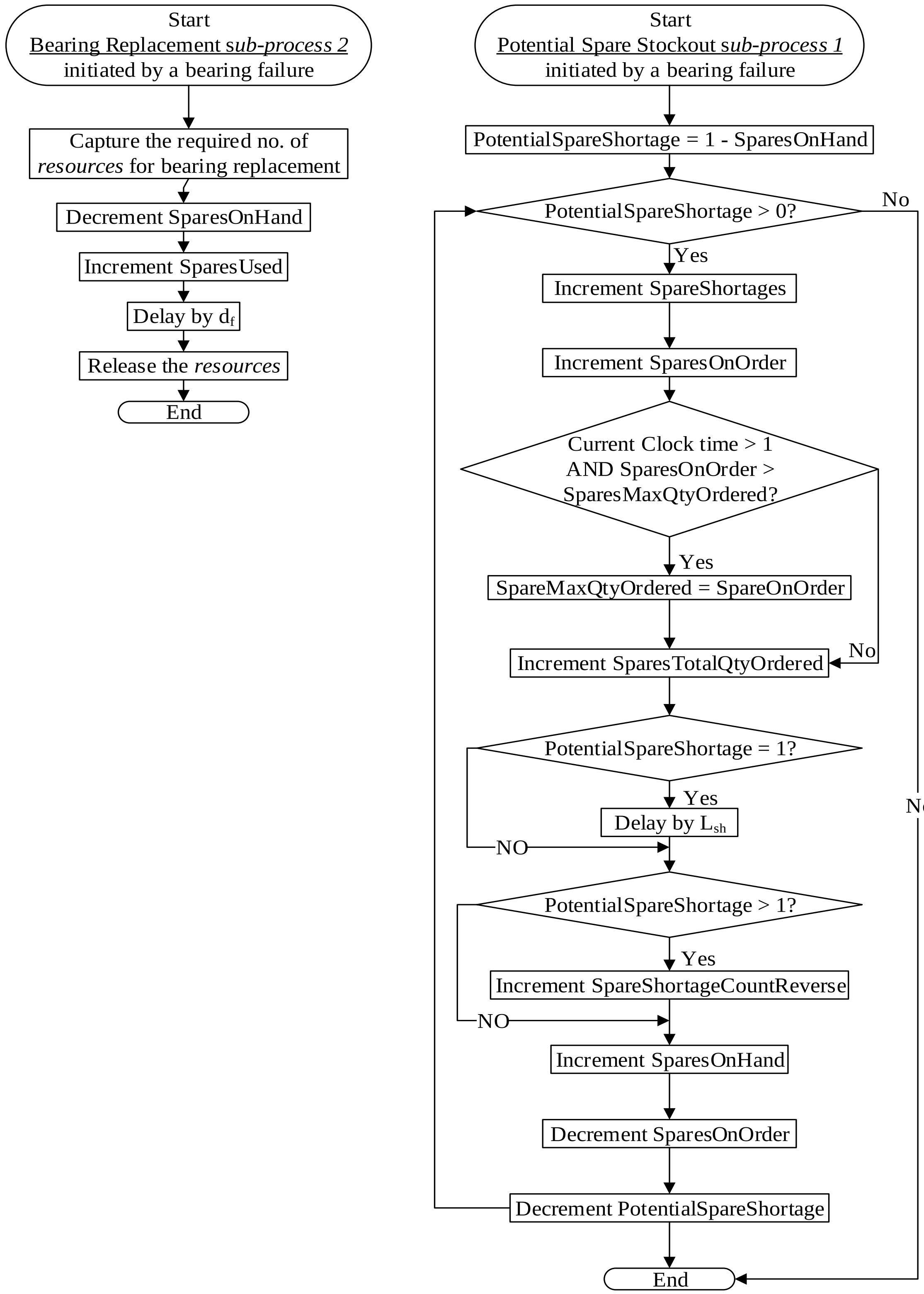
Appendix 5.5. Flowchart of the *Clock* (planned/scheduled) downtime *sub-processes 3, 4 & 5* routines for each machine.



Appendix 5.6. Flowchart of the *Called* (unplanned/unscheduled) downtime routine for each machine.



Appendix 5.7. Flowchart of the *Called* (unplanned/unscheduled) downtime *sub-processes* routine for each machine.



Appendix 5.8. Details of the model: (a) *resources*; (b) *processing*; and (c) *arrivals*.

(a)

Resource	**Number of units**
Machine	2
Technician	6

(b)

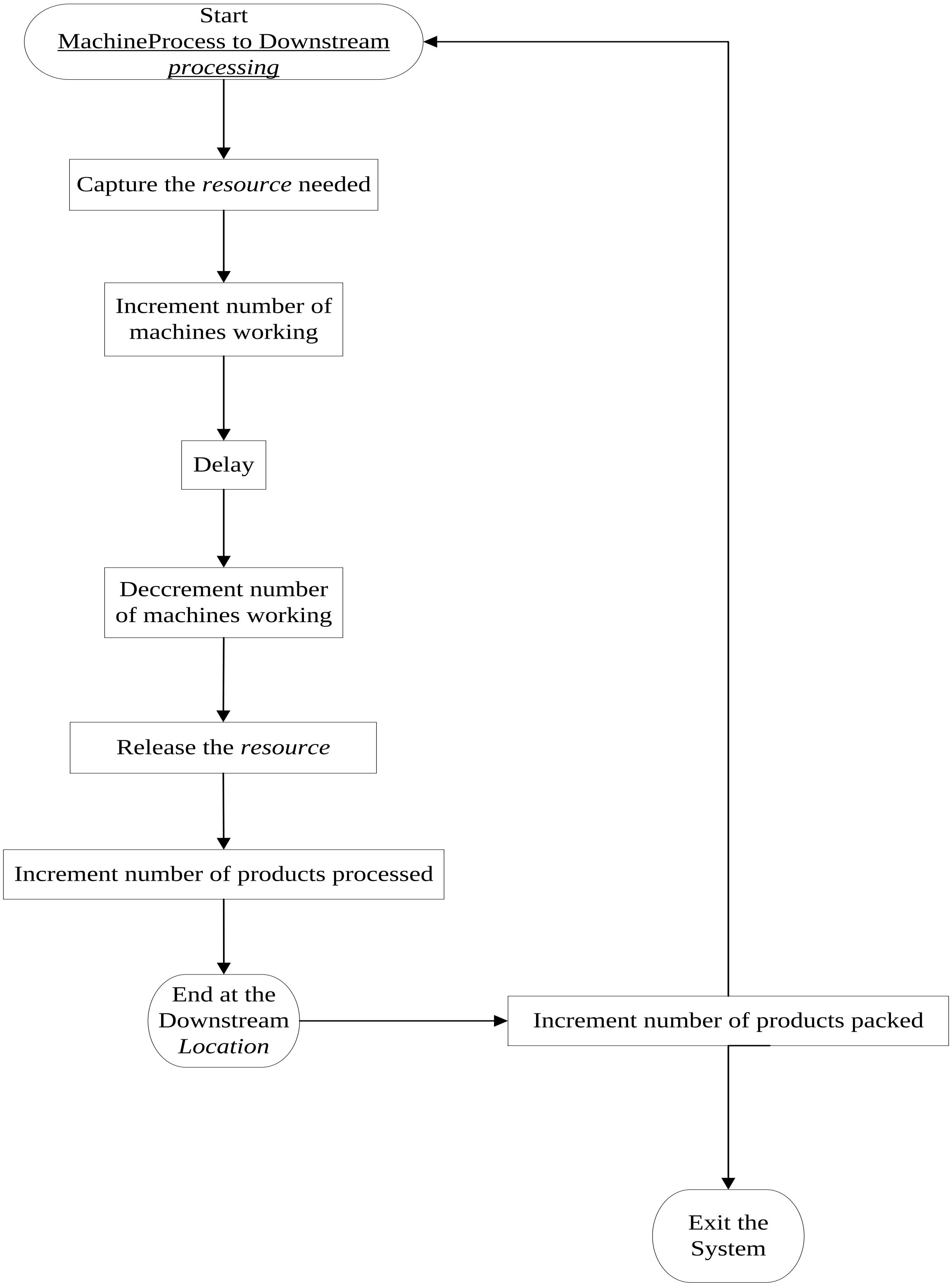
Processing				
Entity	***Location***	**Operation**	**Output**	**Destination**
Product	MC1 MachinePrss	Appendix 5.9	ProductMade	MC1 DownstreamPrss
All	MC1 DownstreamPrss	Appendix 5.9	ProductMade	Exit
PreDefect	MC1 DefectArrival	Appendix 5.10	Defect	MC1 FailureOccurrence
Defect	MC1 FailureOccurrence	Appendix 5.11	Failure	Exit
Defect	MC1 FailureOccurrence	Appendix 5.12	DefectIdentified	MC1 DefectDump
DefectIdentified	MC1 DefectDump	Appendix 5.12	DefectIdentified	Exit
Reset	MC1 Reset	Reset *variables*	Reset	Exit
PlaceOrders	MC1 PlaceOrders	Invoke OrderSpares	PlaceOrders	Exit
HoldingCost	MC1 HoldingCost	Calculate HoldingCost	HoldingCost	Exit
Product	MC2 MachinePrss	Appendix 5.9	ProductMade	MC2 DownstreamPrss
All	MC2 DownstreamPrss	Appendix 5.9	ProductMade	Exit
PreDefect	MC2 DefectArrival	Appendix 5.10	Defect	MC2 FailureOccurrence

Defect	MC2 FailureOccurrence	Appendix 5.11	Failure	Exit
Defect	MC2 FailureOccurrence	Appendix 5.12	DefectIdentified	MC2 DefectDump
DefectIdentified	MC2 DefectDump	Appendix 5.12	DefectIdentified	Exit
Reset	MC2 Reset	Reset *variables*	Reset	Exit
PlaceOrders	MC2 PlaceOrders	Invoke OrderSpares	PlaceOrders	Exit
HoldingCost	MC2 HoldingCost	Calculate HoldingCost	HoldingCost	Exit
Calculate	MCS Calculate	Calculate Grand Total	Calculate	Exit

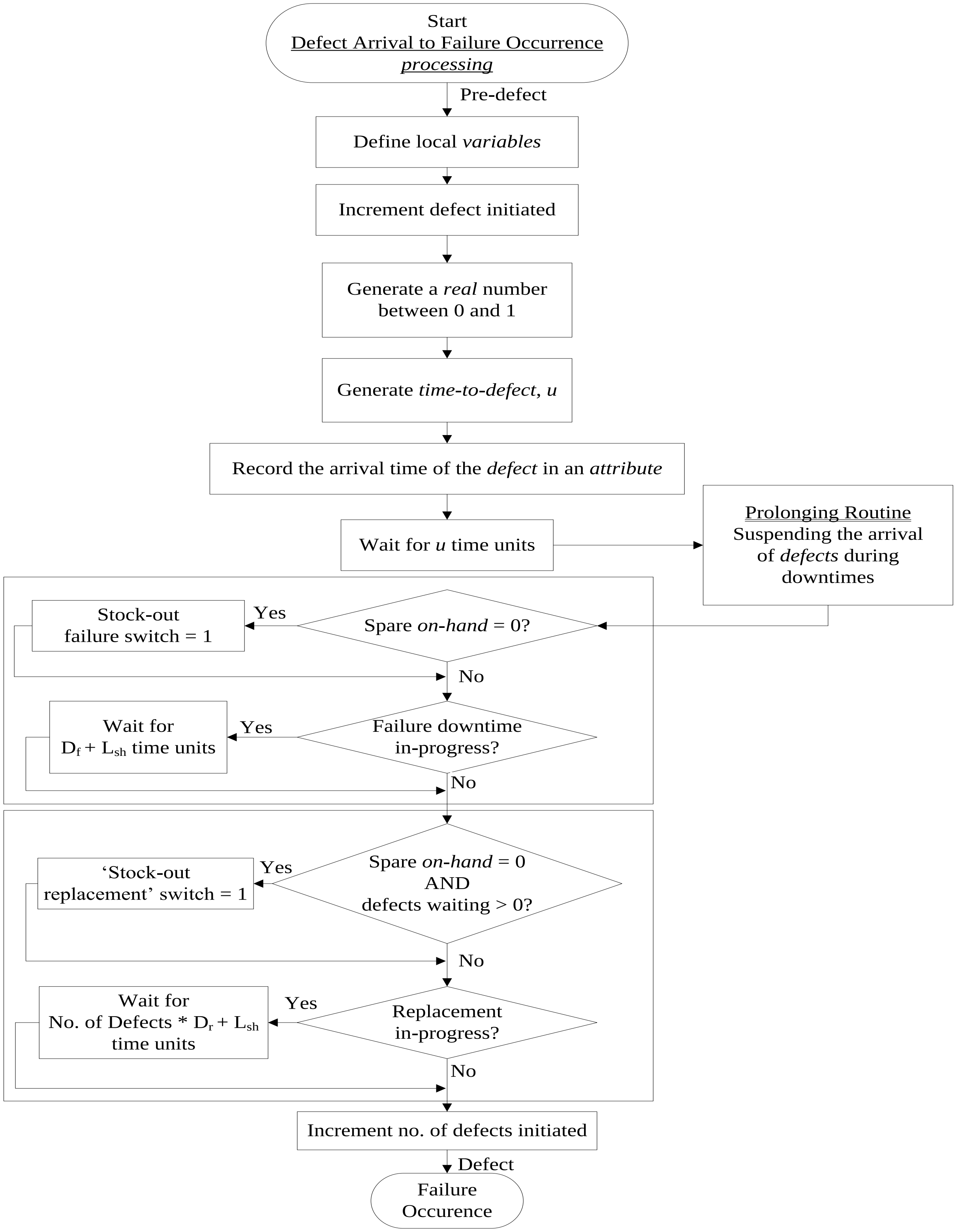
(c)

Arrival					
Entity	***Location***	**Qty**	**First Time**	**Occurrence**	**Frequency**
Product	MC1 Machine Prss	1	0	1	None
PreDefect	MC1 DefectArrival	1	0	1	None
Reset	MC1 Reset	1	WarmUpPeriod	1	None
HoldingCost	MC1 HoldingCost	1	1	Infinite	1
PlaceOrders	MC1 PlaceOrders	1	MC1OrderInterval-SpLeadTime	Infinite	MC1 OrderInterval
Product	MC2 Machine Prss	1	0	1	None
PreDefect	MC2 DefectArrival	1	0	1	None
Reset	MC2 Reset	1	WarmUpPeriod	1	None
Calculate	MCSCalculate	1	0	Infinite	7

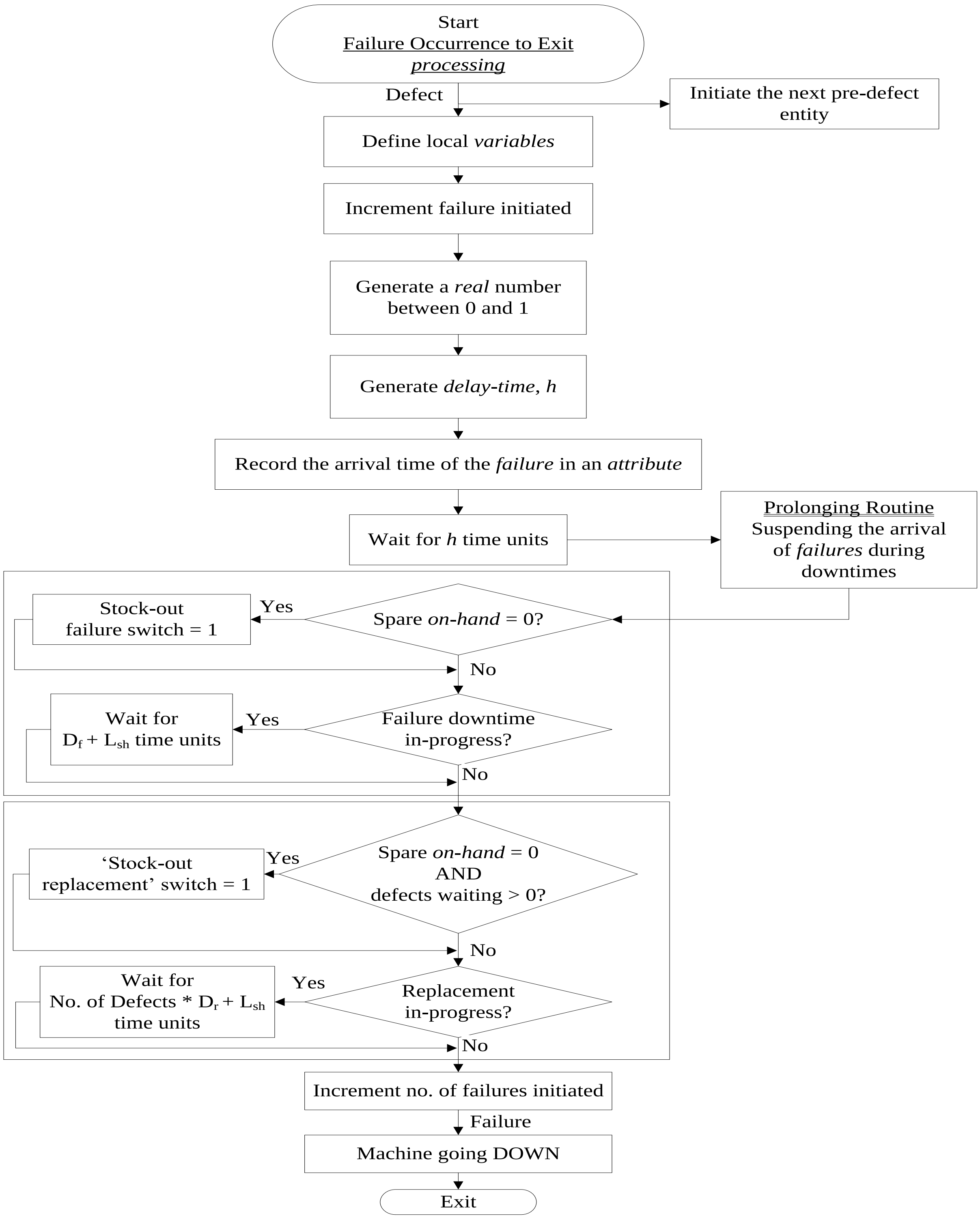
Appendix 5.9. Flowchart of the *Processing Operations* routine for the Machine Process to Downstream Process and Exit.



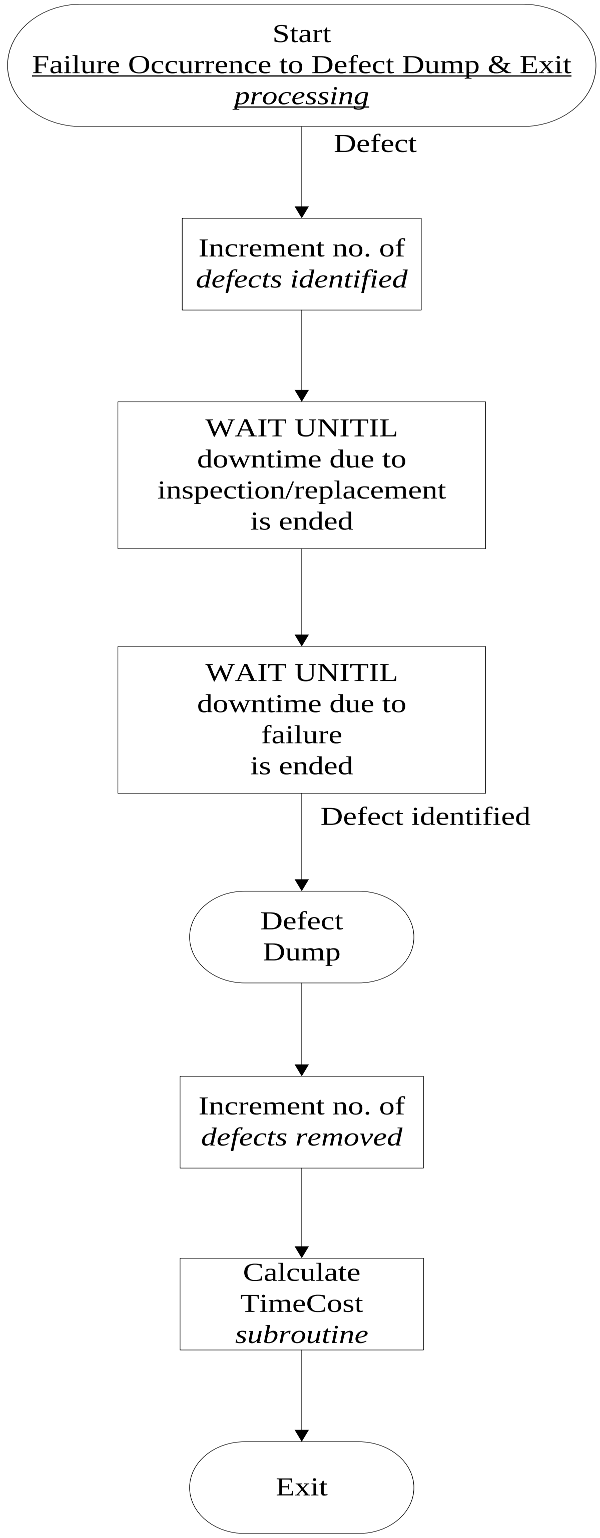
Appendix 5.10. Flowchart of the *Processing Operations* routine for the Defect Arrival to Failure Occurrence.



Appendix 5.11. Flowchart of the *Processing Operations* routine for the Failure Occurrence to Exit.



Appendix 5.12. Flowchart of the *Processing Operations* routine for the Failure Occurrence to Defect Dump & Exit.



Appendix 5.13. Details of the model: (a) *attributes*; (b) *macros*; and (c) *subroutines*.

(a)

Attribute	**Type**	**Classification**
MC1StartTime	Real	Entity
MC2StartTime	Real	Entity

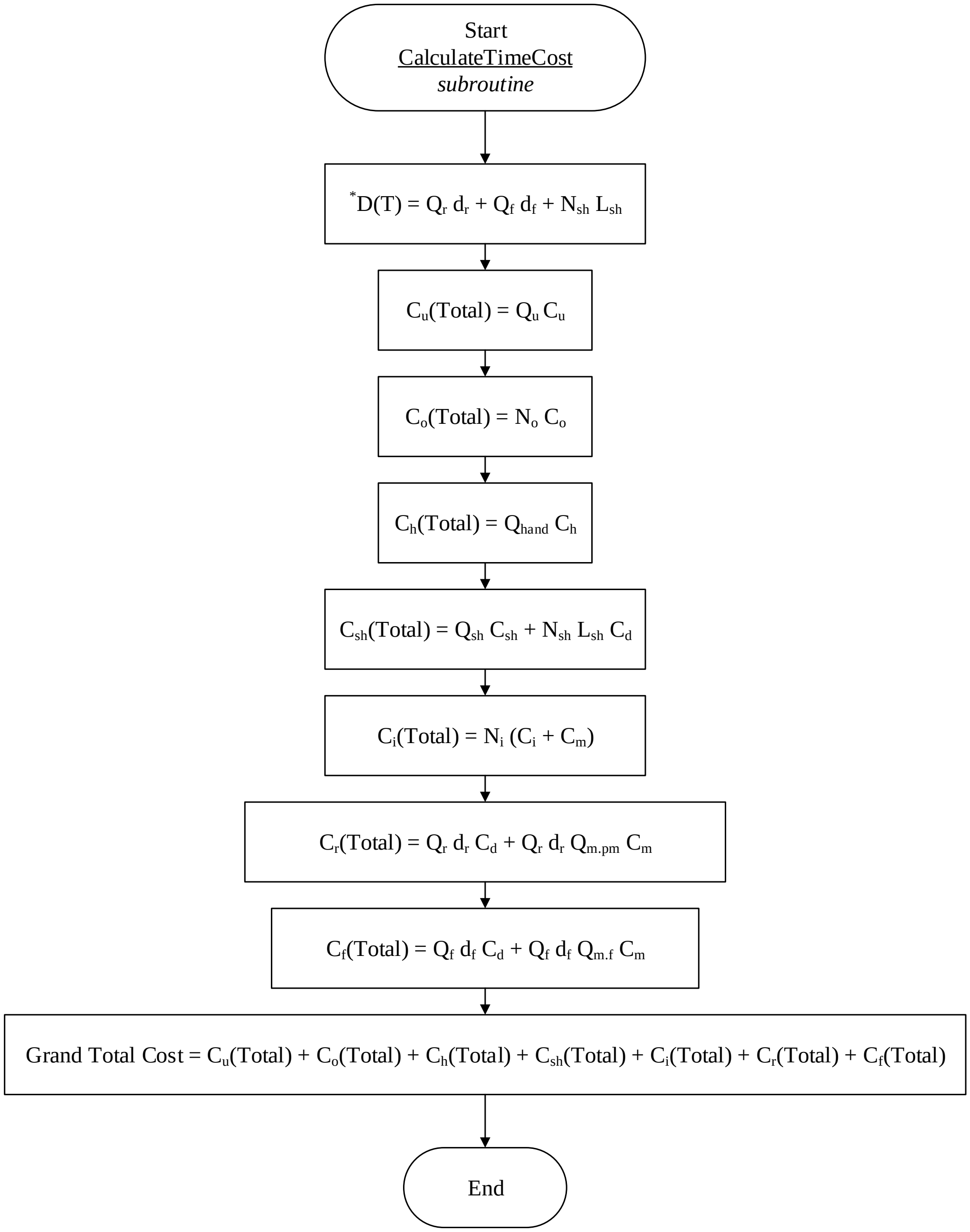
(b)

Macro
OrderIntervalInWeeks
MSL
Mc1DefectArrivalIntensityPerWeek
Mc2DefectArrivalIntensityPerWeek
SpareNormalLeadTimeInWeeks
SpareShortageLeadTimeInWeeks
K
WarmUpInWeeks
NumberOfReplications
SimulationLengthInWeeks

(c)

Subroutine	**Type**	**Parameters**	**Logic**
MC1CalculateTimeCost	None	None	See Appendix 5.14
MC2CalculateTimeCost	None	None	See Appendix 5.14
MCSOrderSpares	None	None	See Appendix 5.15
MCSCalculateTimeCost	None	None	See Appendix 5.16

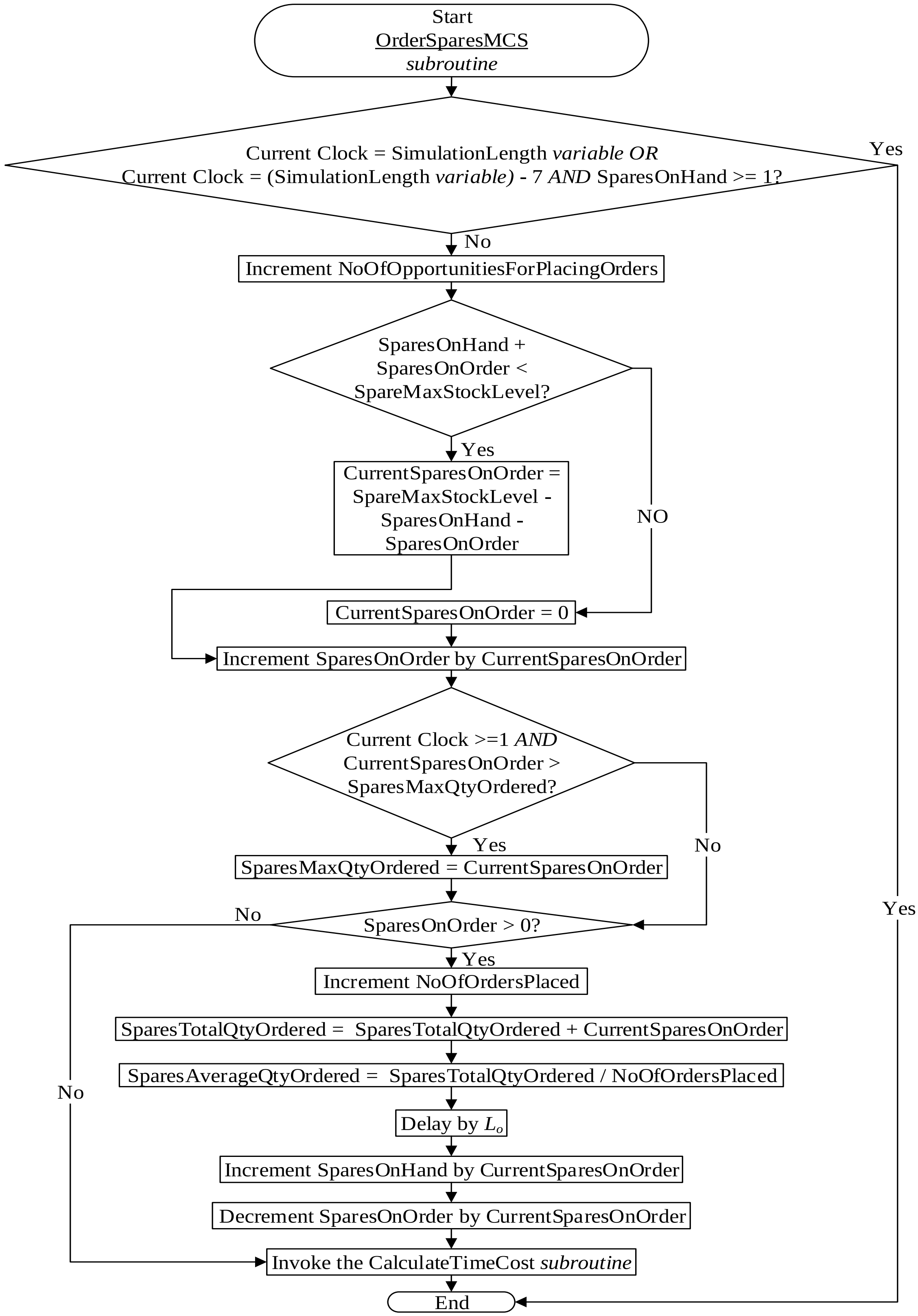
Appendix 5.14. Flowchart of the CalculateTimeCost *subroutine* (individual machines).



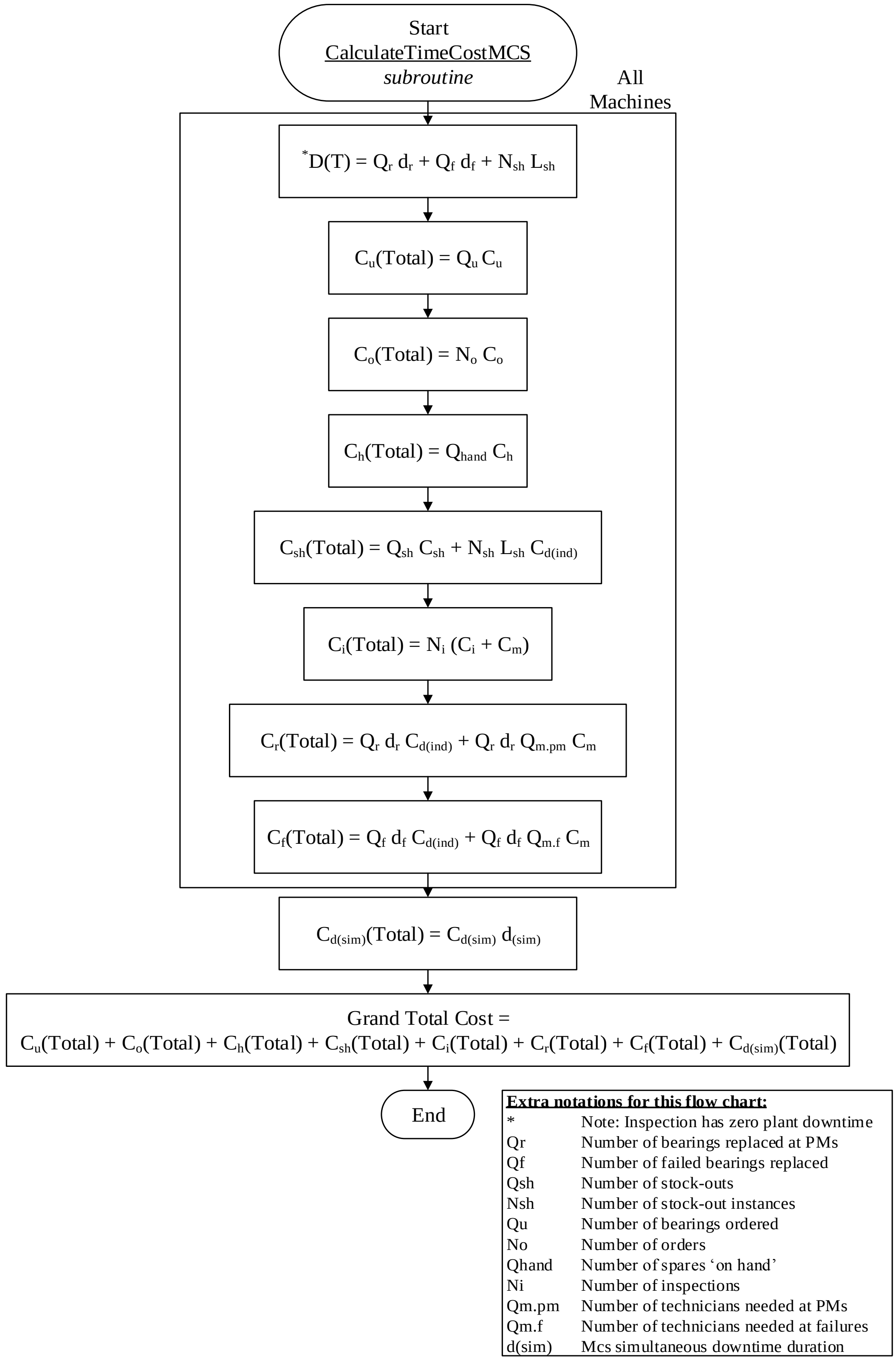
Extra notations for this flow chart:

*	Note: Inspection has zero plant downtime
Q_r	Number of bearings replaced at PMs
Q_f	Number of failed bearings replaced
Q_{sh}	Number of stock-outs
N_{sh}	Number of stock-out instances
Q_u	Number of bearings ordered
N_o	Number of orders
Q_{hand}	Number of spares ‘on hand’
N_i	Number of inspections
$Q_{m.pm}$	Number of technicians needed at PMs
$Q_{m.f}$	Number of technicians needed at failures

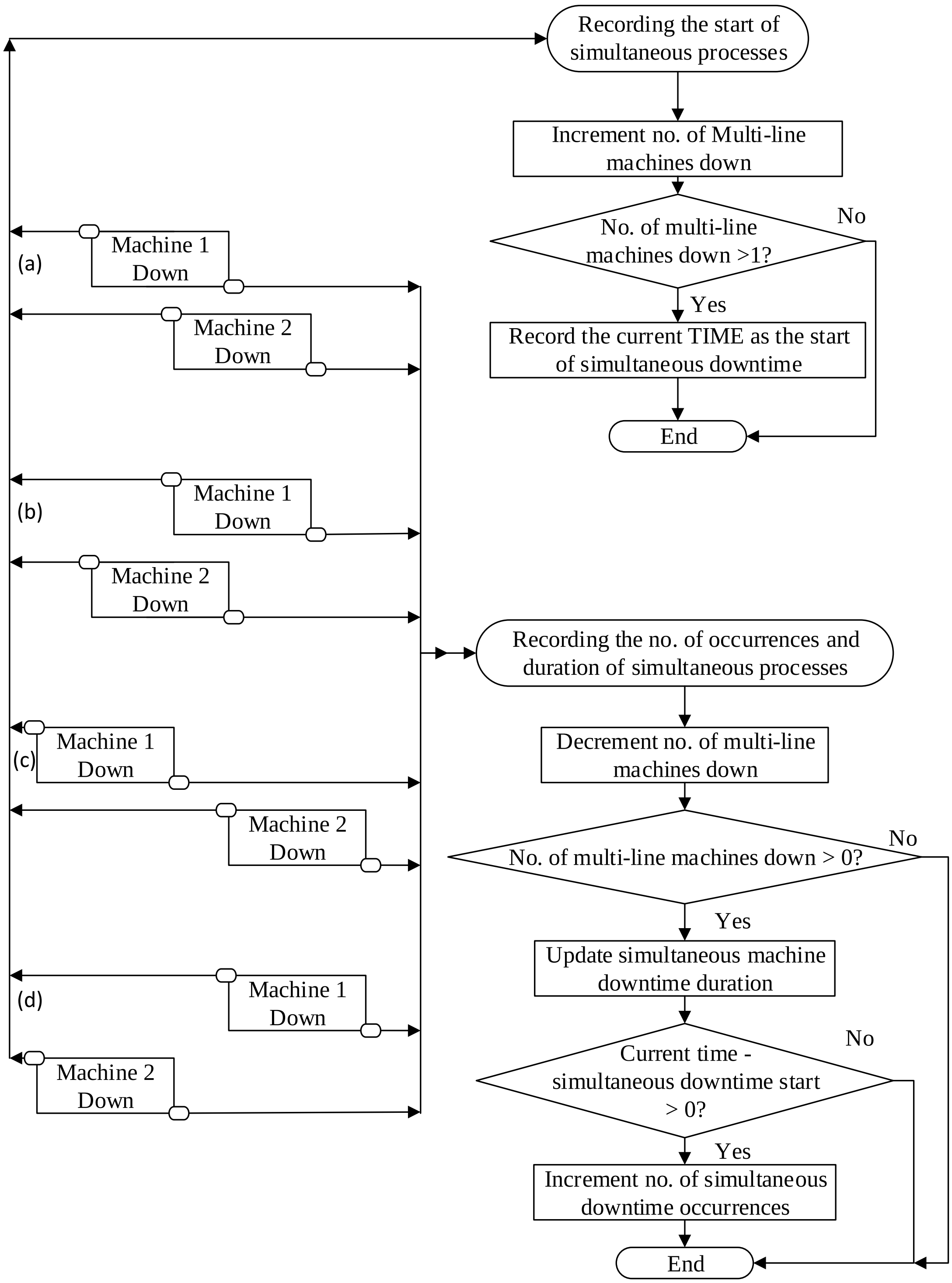
Appendix 5.15. Flowchart of the OrderSparesMCS *subroutine*.



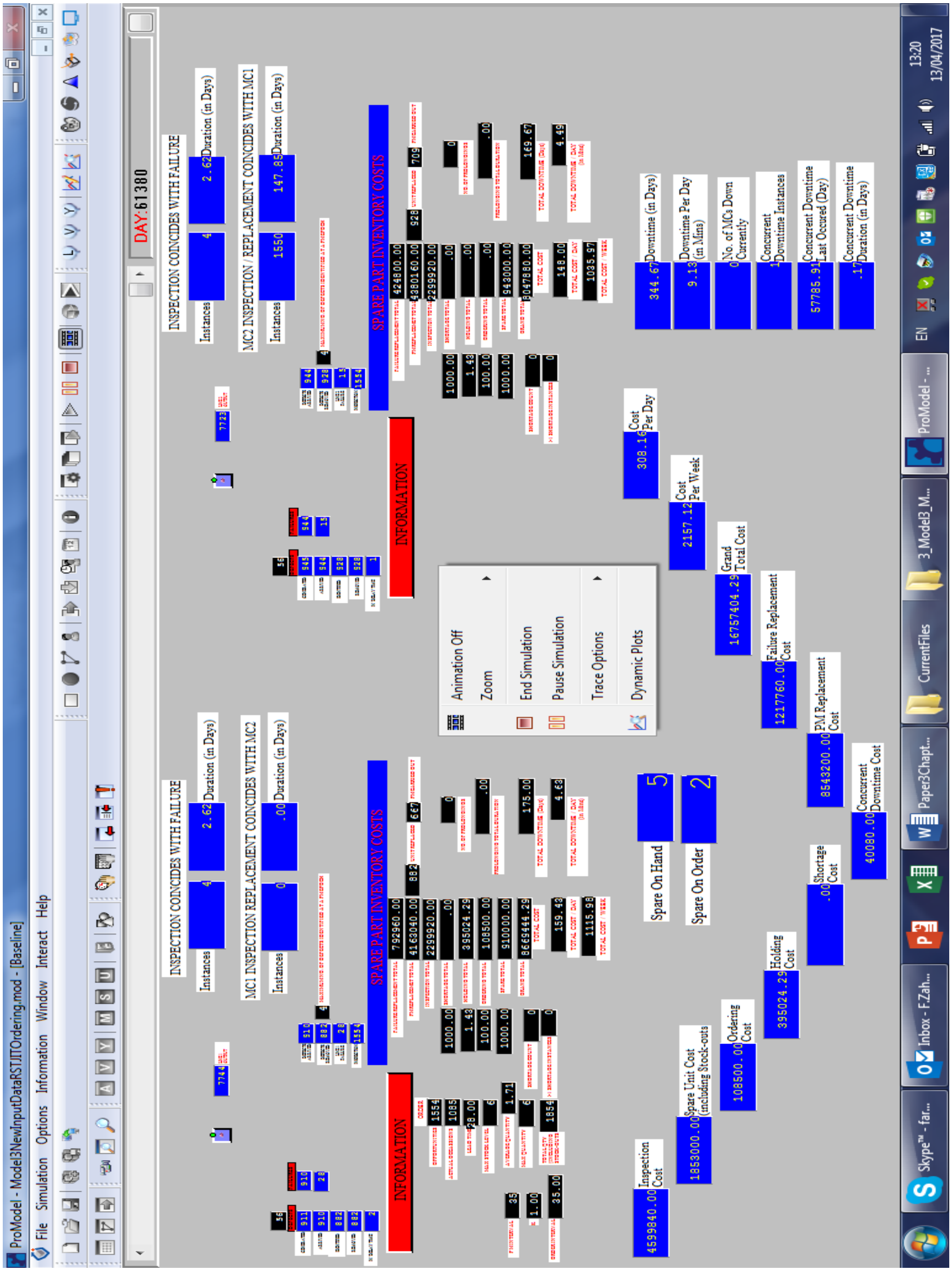
Appendix 5.16. Flowchart of the CalculateTimeCostMCS *subroutine* (multi-line).



Appendix 5.17. Flowchart, depicting the process of capturing and recording simultaneous machine downtime.

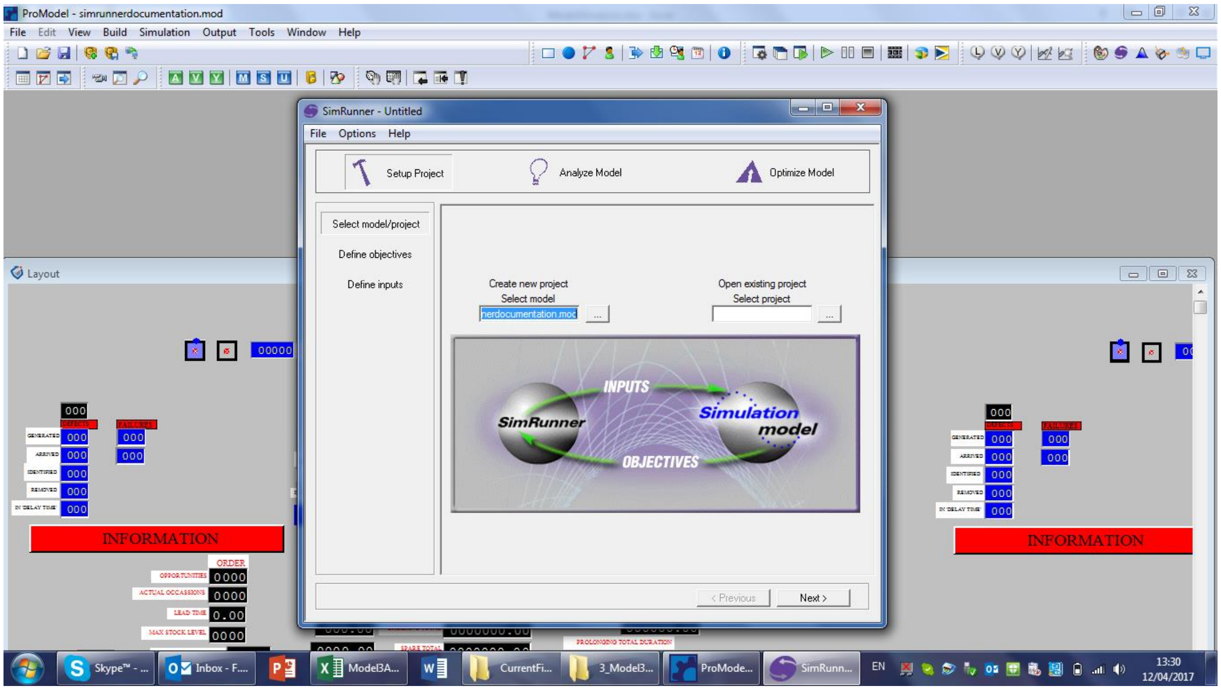


Appendix 5.18. On-screen layout of simulation model.

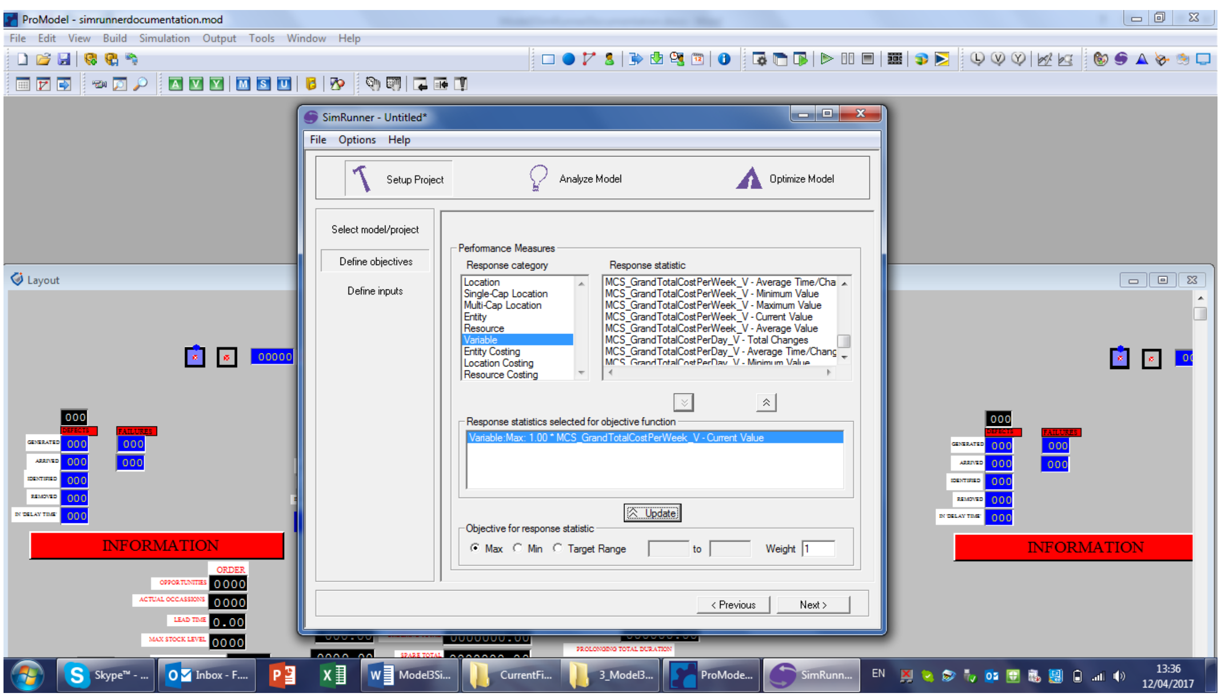


Appendix 5.19. *SimRunner* screen-captures:- inputting model parameter values for determining the cost-optimal policy for a given inspection interval:

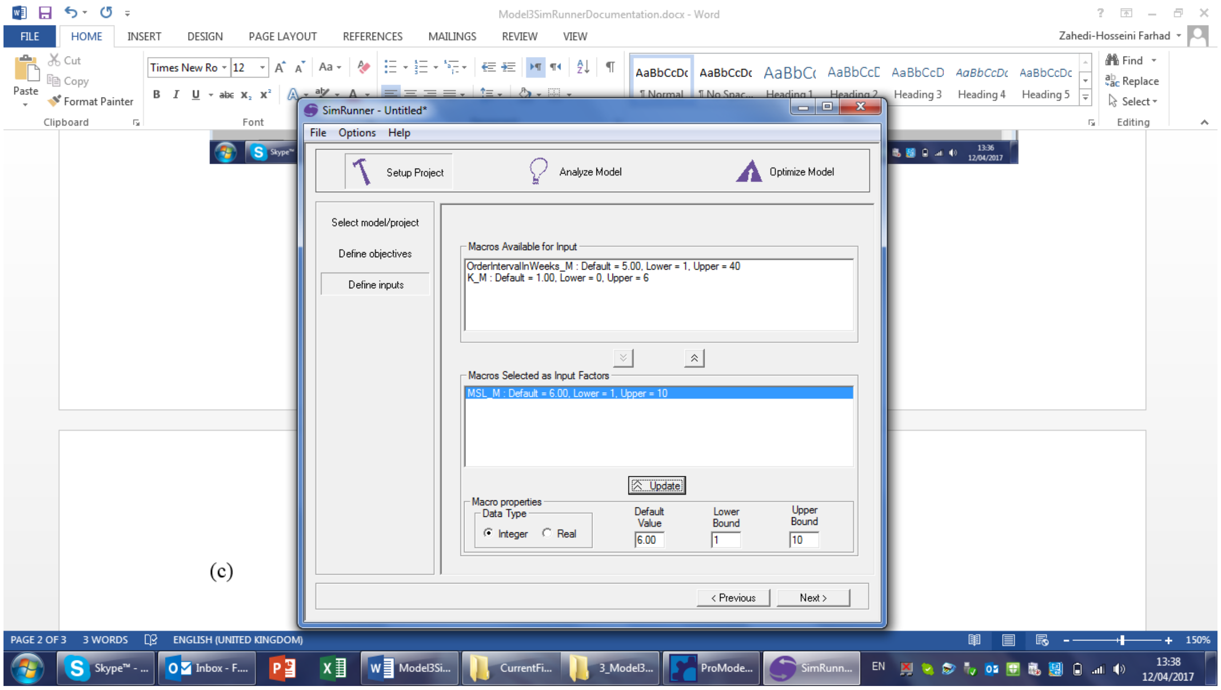
(a) The *SimRunner* optimisation tool - initial screen.



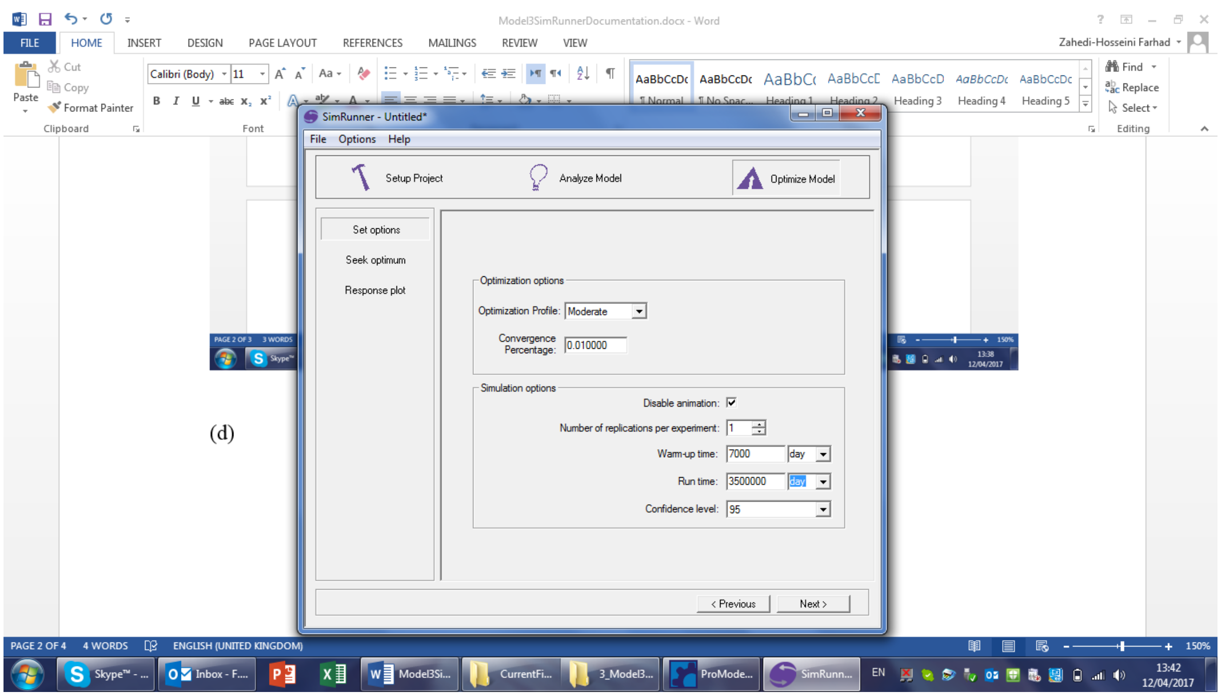
(b) The selection of the *response category* and the specific *response statistic*.



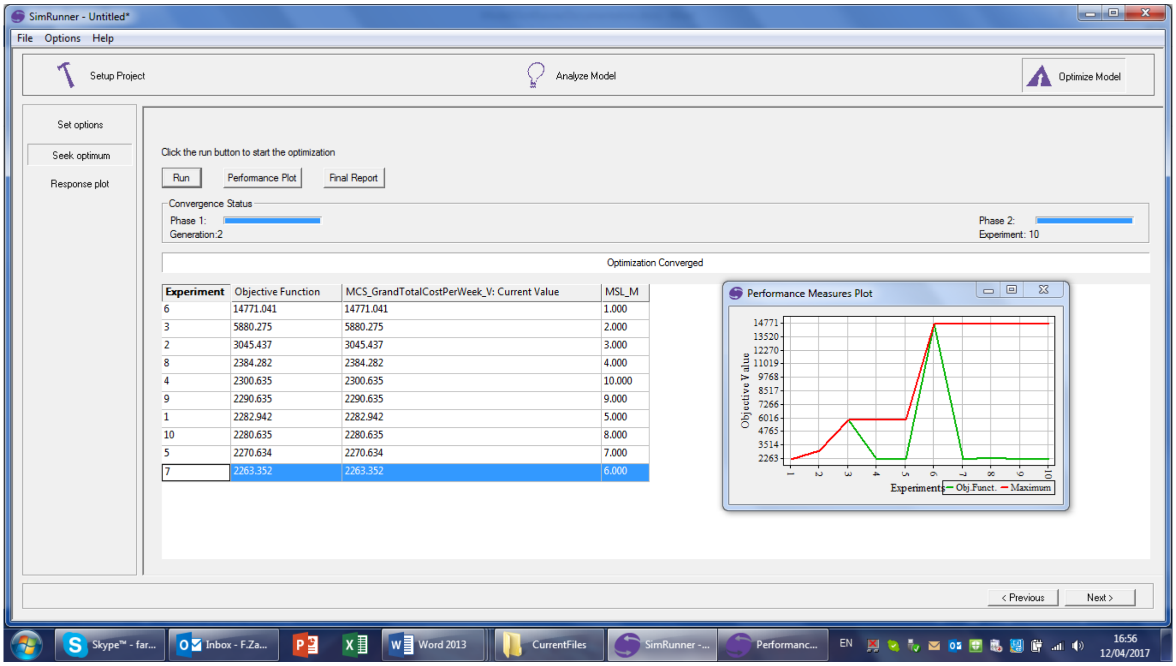
(c) Setting the range for the *response statistic*.



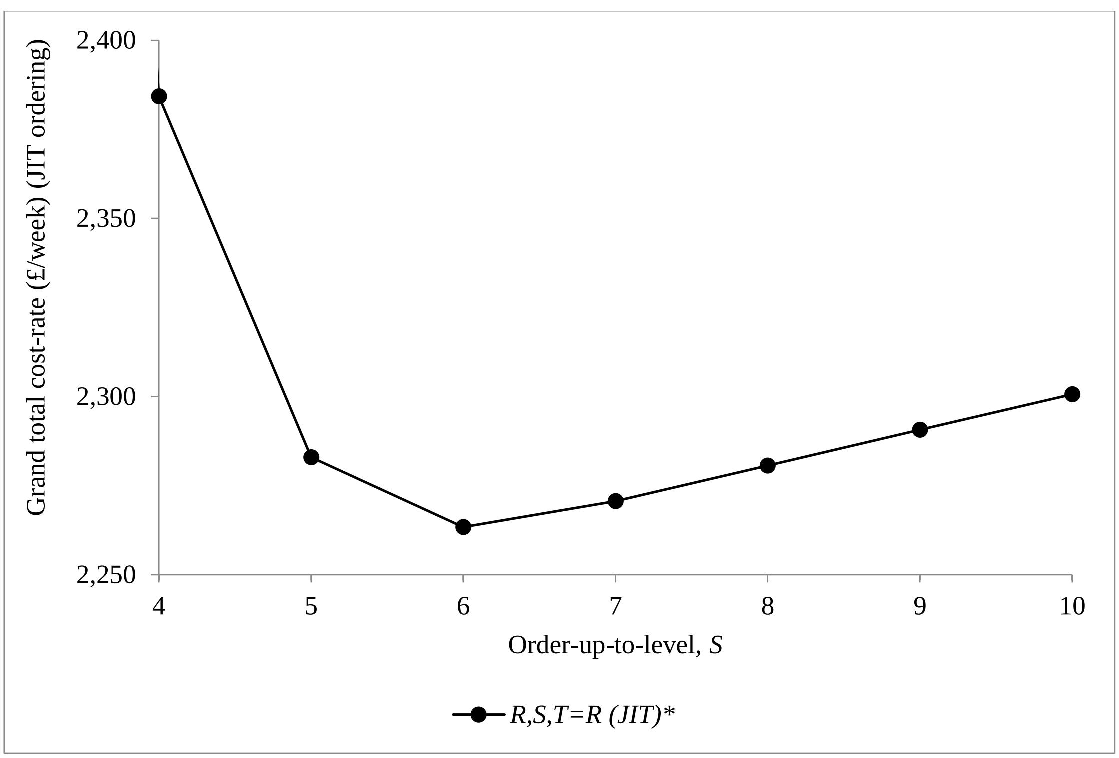
(d) Specifying the model parameters: *warm-up*; *no. of replications*; and *run time*.



(e) *SimRunner* experiments, confirming the cost-optimal policy is achieved with $S = 6$.



(f) The plot illustrating that the *SimRunner* has indeed found the optimum S for the specific $(R, S, T = R)$ policy using *just-in-time* ordering for $T = 5$.



8. References

Abdel-Hameed, M. (1995). “Inspection, maintenance and replacement models.” *Computers and Operations Research* 22: 435-441.

Abdollahzadeh, H., Atashgar, K., and Abbasi, M. (2016). “Multi-objective opportunistic maintenance optimization of a wind farm considering limited number of maintenance groups.” *Renewable Energy* 88: 247-261.

Akbarov, A., Wang, W., and Christer, A. H. (2008). “Problem identification in maintenance modelling: a case study.” *International Journal of Production Research* 46: 1031-1046.

Allaoui, H., and Artiba, A. (2004). “Integrating simulation and optimization to schedule a hybrid flow shop with maintenance constraints.” *Computers and Industrial Engineering* 47: 431-450.

Alrabghi, A., and Tiwari, A. (2015). “State of the art in simulation-based optimisation for maintenance systems.” *Computers & Industrial Engineering* 82: 167–182.

Alrabghi, A., and Tiwari, A. (2016). “A novel approach for modelling complex maintenance systems using discrete event simulation.” *Reliability Engineering and System Safety* 154: 160-170.

Alrabghi, A., Tiwari, A., and Savill, M. (2017). “Simulation-based optimisation of maintenance systems: Industrial case studies.” *Journal of Manufacturing Systems* 44: 191-206.

Alsyouf, I. (2006). “Measuring maintenance performance using a balanced scorecard approach.” *Journal of Quality in Maintenance Engineering* 12: 133–149.

Alsyouf, I. (2009). “Maintenance practices in Swedish industries: Survey results.” *International Journal of Production Economics* 121: 212–223.

Alsyouf, I., Shamsuzzaman, M., Abdelrahman, G. and Al-Taha, M. (2016). “Improving reliability of repairable systems using preventive maintenance and time-between-failures monitoring.” *European Journal of Industrial Engineering* 10: 596-617.

Arthur, N. (2005). "Optimization of vibration analysis inspection intervals for an offshore oil and gas water injection pumping system." *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* 219: 251-259.

Babishin, V., and Taghipour, S. (2016). "Optimal maintenance policy for multicomponent systems with periodic and opportunistic inspections and preventive replacements." *Applied Mathematical Modelling* 40: 10480-10505.

Baker, R. D., and Christer, A. H. (1994). "Review of delay-time OR modelling of engineering aspects of maintenance." *European Journal of Operational Research* 73: 407-422.

Baker, R. D., and Wang, W. (1992). "Estimating the delay-time distribution of faults in repairable machinery from failure data." *IMA Journal of Mathematics Applied in Business & Industry* 3: 259-281.

Baker, R. D, Wang, W. (1993). "Developing and testing the delay time model." *Journal of Operational Research Society* 44: 361-374.

Banks, J. (2010). *Discrete-event System Simulation*. Upper Saddle River: Pearson.

Banks, J., Carson, J.S., Nelson, B.L., and Nicol, D.M. (2013). *Discrete-event System Simulation*. Pearson: Harlow, UK.

Barlow, R., Hunter, L. (1960). "Optimum preventive maintenance policies." *Operations Research* 8: 90-100.

Barlow, R. E. and Proschan, F. (1965). *Mathematical Theory of Reliability*. Wiley, New York.

Berrade, M. D., Cavalcante, C. A. V., and Scarf, P. A. (2012). "Maintenance scheduling of a protection system subject to imperfect inspection and replacement." *European Journal of Operational Research* 218: 716-725.

Boschian, V., Rezg, N., and Chelbi, A. (2009). "Contribution of simulation to the optimization of maintenance strategies for a randomly failing production system." *European Journal of Operational Research* 197: 1142–1149.

Boulet, J. F., Gharbi, A., and Kenn, J. P. (2009). "Multiobjective optimization in an unreliable failure-prone manufacturing system." *Journal of Quality in Maintenance Engineering* 15: 397-411.

Boylan, J. E., and Syntetos, A. A., (2010). "Spare parts management: a review of forecasting research and extensions." *IMA Journal of Management Mathematics* 21: 227-237.

Brezavscek, A., and Hudoklin, A. (2003). "Joint optimization of block replacement and periodic-review spare-provisioning policy." *IEEE Transactions on Reliability* 52: 112-117.

Chelbi, A., and Ait-Kadi, D. (2001). "Spare provisioning strategy for preventively replaced systems subjected to random failure." *International Journal of Production Economics* 74: 183-189.

Cavory, G., Dupas, R., and Goncalves, G. (2001). "A genetic approach to the scheduling of preventive maintenance tasks on a single product manufacturing production line." *International Journal of Production Economics* 74: 135-146.

Chellappachetty, B., and Raju, R. (2015). "Analysing the effect of grouping subsystems for periodic maintenance inspection of equipment using delay time methodology to minimise the downtime per unit time." *Indian Journal of Science and Technology* 8: 849-858.

Chen, L., Ye, Z. S., and Xie, M. (2013). "Joint maintenance and spare parts provisioning policy for k-out-of-n systems." *Asia-Pacific Journal of Operational Research* 30: 1-21.

Cheng, R., Macal, C., Nelson, B., Rabe, M., Currie, C., Fowler, J., and Lee, L. H. (2016). "Simulation: The past 10 years and the next 10 years." *Proceedings of the Winter Simulation Conference. Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Cheng, Y. H. and Tsao, H. L. (2010). "Rolling stock maintenance strategy selection, spares parts' estimation, and replacements' interval calculation." *International Journal of Production Economics* 128: 404-412.

Cho, D. I., and Parlar, M. (1991). "A survey of maintenance models for multi-unit systems." *European Journal of Operational Research* 51: 1-23.

Cholasuke, C., Bhardwa, R., and Antony, J. (2004). "The status of maintenance management in UK manufacturing organisations: results from a pilot survey." *Journal of Quality in Maintenance Engineering* 10: 5-15.

Christer, A. H. (1976). "Innovative Decision Making." In *Proceedings of the NATO conference on the role and effectiveness of theories of decision in practice*. Bowen, K. C. and White, D. J. (Editors), Hodder and Stoughton, 368-377.

Christer, A. H. (1987). “Delay-time model of reliability of equipment subject to inspection monitoring.” *Journal of the Operational Research Society* 38: 329-334.

Christer, A. H. (1999). “Developments in delay time analysis for modelling plant maintenance.” *Journal of the Operational Research Society* 50: 1120-1137.

Christer, A. H., and Waller, W. M. (1984a). “Delay time Models of Industrial Inspection Maintenance Problems.” *Journal of the Operational Research Society* 35: 401-406.

Christer, A. H, Waller, W. M. (1984b). “Reducing production downtime using delay-time analysis.” *Journal of the Operational Research Society* 35: 499-512.

Christer, A. H., and Wang, W. (1995). “A delay time based maintenance model of a multi-component system.” *IMA Journal of Management Mathematics* 6: 205-222.

Christer, A. H., Wang, W., Baker, R. D., and Sharp, J. M. (1995). “Modelling maintenance practice of production plant using the delay-time concept.” *IMA Journal of Management Mathematics* 6: 67-83.

Collins, D. (2017). “What is L10 life and why does it matter?” 23 February 2017. Accessed 12 October 2017. http://www.linearmotiontips.com/what-is-l10-life-and-why-does-it-matter/

Crespo, A. M., Gupta, J. N. D., and Heguedas, A. S. (2003). “Maintenance policies for a production system with constrained production rate and buffer capacity.” *International Journal of Production Research* 41: 1909-1926.

Das, A. N., and Sarmah S. P. (2010). “Preventive replacement models: an overview and their application in process industries.” *European Journal of Industrial Engineering* 4: 280–307.

De Almeida, A. T. (2001). “Multicriteria decision making on maintenance, spares and contracts planning.” *European Journal of Operational Research* 129: 235-241.

De Almeida, A. T., Ferreira, R. J. P., and Cavalcante, C.A.V. (2015). “A review of the use of multicriteria and multi-objective models in maintenance and reliability.” *IMA Journal of Management Mathematics* 26: 249-271.

De Smidt-Destombes, K. S., Van Der Heigden, M. C. and Van Harten, A. (2007). “Availability of k-out-of-N systems under block replacement sharing limited spares and repair capacity.” *International Journal of Production Economics* 107: 404-421.

De Smidt-Destombes, K. S., Van Der Heijden, and M. C., Van Harten, A. (2009). "Joint optimisation of spare part inventory, maintenance frequency and repair capacity for k-out-of-N systems." *International Journal of Production Economics* 118: 260-268.

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

Dekker, R., and Scarf, P. A. (1998). "On the impact of optimisation models in maintenance decision making: the state of the art." *Reliability Engineering and System Safety* 60: 111-119.

Dias, L. M. S., Vieira, A. A. C., Pereira, G. A. B., and Oliveira, J. A. (2016). "Discrete simulation software ranking – a top list of the worldwide most popular and used tools." *Proceedings of the Winter Simulation Conference. Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Ding, S. H., Kamaruddin, S. (2015). "Maintenance policy optimization - literature review and directions." *International Journal of Advanced Manufacturing Technology* 76: 1263–1283.

Dinmohammadi, F., Alkali, B., Shafiee, M., Bérenguer, C., and Labib, A. (2016). "Risk evaluation of railway rolling stock failures using FMECA technique: A case study of passenger door system." *Urban Rail Transit* 2: 128-145.

Emovon, I., Norman, R. A., Murphy, A. J. (2016). "An integration of multi-criteria decision making techniques with a delay time model for determination of inspection intervals for marine machinery systems." *Applied Ocean Research* 59: 65-82.

Erguido, A., Marquez, A. C., Castellano, E., and Fernandez, J. F. G. (2017). "A dynamic opportunistic maintenance model to maximize energy-based availability while reducing the life cycle cost of wind farms." *Renewable Energy* 114: 843-856.

Flage, R. (2014). "A delay time model with imperfect and failure-inducing inspections." *Reliability Engineering and System Safety* 124: 1–12.

Folger, R., Rodes, J., and Novak, D. (2014a). "Bearing Killer: preventing common causes of bearing system damage – part 1." *Maintenance & Engineering* 14: 12-15.

Folger, R., Rodes, J., and Novak, D. (2014b). "Bearing Killer: preventing common causes of bearing system damage – part 2." *Maintenance & Engineering* 14: 8-10.

Gan, S., Zhang, Z., Zhou, Y., and Shi, J. (2015). “Joint optimization of maintenance, buffer, and spare parts for a production system.” *Applied Mathematical Modelling* 39: 6032-6042.

Geraerds, W. M. J. (1978). “Estimation of Cost of Maintenance Expenditure within the Netherlands.” Internal Report, Faculty of Technology Management, Eindhoven University of Technology, Netherlands.

Gharbi, A., and Kenné, J. P. (2005). “Maintenance scheduling and production control of multiple-machine manufacturing systems.” *Computers and Industrial Engineering* 48: 693-707.

Ghobbar, A. A., and Friend, C. H. (2003). “Evaluation of forecasting methods for intermittent parts demand in the field of aviation: A predictive model.” *Computers and Operations Research* 30: 2097-2114.

Gopalakrishnan, M., Laroque, C., and Skoogh, A. (2016). “Buffer utilisation based scheduling of maintenance activities by a shifting priority approach – a simulation study.” *Proceedings of the Winter Simulation Conference. Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Gupta, A., and Lawsirirat, C. (2006). “Strategically optimum maintenance of monitoring-enabled multi-component systems using continuous-time jump deterioration models.” *Journal of Quality in Maintenance Engineering*, 12: 306–329.

Harrell, C., Ghosh, B. K., and Bowden, R. O. (2011). *Simulation using ProModel*. *Third Edition*. McGraw Hill.

Hu, Q., Chakhar, S., Siraj, S., and Labib, A. (2017). “Spare parts classification in industrial manufacturing using the dominance-based rough set approach.” *European Journal of Operational Research* 262: 1136-1163.

Huang, R., Meng, L., Xi, L., and Liu, C. (2008). “Modeling and analyzing a joint optimization policy of block-replacement and spare inventory with random-leadtime.” *IEEE Transactions Reliability* 57: 113-124.

Ilgin, M. A., and Tunali, S. (2007). “Joint optimization of spare parts inventory and maintenance policies using genetic algorithms.” *International Journal of Advanced Manufacturing Technology* 34: 594-604.

Irawan, C. A., Ouelhadj, D., Jones, D., Stålhane, M., and Sperstad, I. B. (2017). "Optimisation of maintenance routing and scheduling for offshore wind farms." *European Journal of Operational Research* 256: 76-89.

Jacobs, W., Van Hooreweder, B., Boonen, R., Sas, P, Moens, D. (2016). "The influence of external dynamic loads on the lifetime of rolling element bearings: Experimental analysis of the lubricant film and surface wear." *Mechanical Systems and Signal Processing* 74: 144-164.

Jardine, A. K. S., and Hassounah, M. I. (1990). "An optimal vehicle-fleet inspection schedule." *Journal of the Operational Research Society* 33: 723-732.

Jiang, Y., Chen, M., and Zhou, D. (2015). "Joint optimization of preventive maintenance and inventory policies for multi-unit systems subject to deteriorating spare part inventory." *Journal of Manufacturing Systems* 35: 191-205.

Jones, B., Jenkinson, I., and Wang, J. (2009). "Methodology of using delay-time analysis for a manufacturing industry." *Reliability Engineering and System Safety* 94: 111-124.

Jones, B., Jenkinson, I., Yang, Z., and Wang, J. (2010). "The use of Bayesian network modelling for maintenance planning in a manufacturing industry." *Reliability Engineering and System Safety* 95: 267-277.

Kaio, N., and Osaki, S. (1989). "Comparison of Inspection Policies." *Journal of the Operational Research Society* 40: 499-503.

Kennedy, W. J., Patterson, J. W., and Fredendall, L. D. (2002). "An overview of recent literature on spare parts inventories." *International Journal of Production Economics* 76: 201-215.

Khalili, M. H., and Zahedi-Hosseini, F. (2013). "Modelling and simulation of a mattress production line using ProModel." In *Proceedings of the Winter Simulation Conference (WSC)*. Pasupathy R., Kim, S. H., Tolk, A., Hill, R., and Kuhl, M. E. (Editors). 8-11 December 2013, Washington, DC.

Kim, W. K., Yoon, K. P., Kim, Y., and Bronson, G. J. (2012). "Improving system performance for stochastic activity network: A simulation approach." *Computers & Industrial Engineering* 62: 1-12.

Komonen, K. (2002). "A cost model of industrial maintenance for profitability analysis and benchmarking." *International Journal of Production Economics* 79: 15-31.

Kuntz, P. A., Christie, R. D., and Venkata, S. S. (2001). “A reliability centred optimal visual inspection model for distribution feeders.” *IEEE Transactions on Power Delivery* 16: 718-723.

Law, A. M. (2015). *Simulation Modelling and Analysis*. *Fifth edition*. McGraw-Hill: New York.

Linneusson, G., Ng, A., and Aslam, T. (2016). “Investigating maintenance performance: A simulation study.” *7th Production Symposium, Sweden*.

Liu, X., Wang, W., Peng, R. (2015). “An integrated production and delay-time based preventive maintenance planning model for a multi-product production system.” *Journal of Maintenance and Reliability* 17: 215-221.

Lu, W. Y., Wang, W. (2011). “Modelling Preventive maintenance based on the delay time concept in the context of a case study.” *Journal of Maintenance and Reliability* 3: 4-10.

Lu, B. Zhou,X. and Li, Y. (2016). “Joint modelling of preventive maintenance and quality improvement for deteriorating single-machine manufacturing systems.” *Computers & Industrial Engineering* 91: 188-196.

Lung, B., Do, P., Levrat, E., and Voisin, A. (2016). “Opportunistic maintenance based on multi-dependent components of manufacturing system.” *CIRP Annals – Manufacturing Technology* 65: 401-404.

Luss, H. (1983). “An Inspection Policy Model for Production Facilities.” *Management Science* 29: 1102-1109.

Lynch, P., Adendorff, K., Yadavalli, V. S. S., and Adetunji, O. (2013). “Optimal spares and preventive maintenance frequencies for constrained industrial systems.” *Computers and Industrial Engineering* 65: 378–387.

Macal, C. M. (2016). “Everything you need to know about agent-based modelling and simulation.” *Journal of Simulation* 10: 144-156.

Maintenance & Engineering. (2017). “Improving reliability with an outsourced maintenance strategy.” *Maintenance & Engineering*: 16.

Marseguerra, M., Zio, E., and Podofillini, L. (2005). “Multiobjective spare part allocation by means of genetic algorithms and Monte Carlo simulation.” *Reliability Engineering and System Safety* 87: 325-335.

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

Molenaers, A., Baets, H., Pintelon, L., and Waeyenbergh, G. (2012). "Criticality classification of spare parts: A case study." *International Journal of Production Economics* 140: 570-578.

Muller, M. (2011). *Essentials of Inventory Management*. Second Edition. AMACOM.

Natarajan, A. R. (2016). "The design of a novel bearing manufacturing line based on simulation." *American Journal of Operations Research* 6: 425-435.

Nelson, B. L. (2016). "Some tactical problems in digital simulation for the next 10 years." *Journal of Simulation* 10: 2-11.

Nicolai, R. P., and Dekker, R. (2008). "Optimal maintenance of multi-component systems: a review." In *Complex System Maintenance Handbook*. Kobbacy, K. A. H., Murthy, D. N. P. (Editors). Springer, London, 263-286.

Nguyen, T., Deloux, E., and Dieulle, L. (2016). "Maintenance policies for a heterogeneous two-component system with stochastic and economic dependencies." *IFAC-PapersOnLine* 49-12: 787-792.

Okumura, S., Jardine, A. K. S., and Yamashina, H. (1996). "An inspection policy for a deteriorating single-unit system characterized by a delay-time model." *International Journal of Production Research* 34: 2441-2460.

Olde Keizer, M. C. A., Flapper, S. D. P., Teunter, R. H. (2017). "Condition-based maintenance policies for systems with multiple dependent components: A review." *European Journal of Operational Research* 261: 405-420.

Panagiotidou, S. (2014). "Joint optimization of spare parts ordering and maintenance policies for multiple identical items subject to silent failures." *European Journal of Operational Research* 235: 300-314.

Pham, H., and Wang, H. (1996). "Imperfect maintenance." *European Journal of Operational Research* 94: 428-438.

Pierskalla, W. P., and Voelker, J. A. (2006). "A survey of maintenance models: The control and surveillance of deteriorating systems." *Naval Research Logistics Quarterly* 23: 353-388.

Pietruczuk, A. J., and Werbińska-Wojciechowska, S. (2017). "Block inspection policy model with imperfect maintenance for single-unit systems." *Procedia Engineering* 187: 570-581.

Pillay, A., Wang, J., Wall, A. D., and Ruxton T. (2001a). "A maintenance study of fishing vessel equipment using delay-time analysis." *Journal of Quality in Maintenance Engineering* 7: 118-128.

Pillay, A., Wang, J., and Wall, A. D. (2001b). "Optimal inspection period for fishing vessel equipment: a cost and downtime model using delay time analysis." *Marine Technology* 38: 122-129.

Pophaley, M., and Vyas, R. K. (2010). "Plant maintenance management practices in automobile industries: a retrospective and literature review." *Journal of Industrial Engineering and Management* 3: 512-541.

ProModel. (2010). *SimRunner User Guide*. ProModel Corporation.

ProModel. (2016). *ProModel User Guide*. ProModel Corporation.

Rabe, M. and Dross, F. (2015). "A reinforcement learning approach for a decision support system for logistics networks". *Proceedings of the Winter Simulation Conference, Yilmaz, L., Chan, W.K.V., Moon, I., Roeder, T.M.K., Macal, C., and Rossetti, M. (Editors): 2020-2032, Piscataway: IEEE*.

Rezg, N., Chelbi, A., and Xie, X. (2005). "Modelling and optimizing a joint inventory control and preventive maintenance strategy for a randomly failing production unit: analytical and simulation approaches." *International Journal of Computer Integrated Manufacturing* 8: 225-235.

Robinson, S. (2004). *Simulation: The practice of model development and use*. Wiley.

Roux, O., Duvivier, D., Quesnel, G., Ramat, E. (2013). "Optimization of preventive maintenance through a combined maintenance-production simulation model." *International Journal of Production Economics* 143: 3-12.

Roux, O., Jamali, M. A., Kadi, D. A., and Châtelet, E. (2008). "Development of simulation and optimization platform to analyse maintenance policies performances for manufacturing systems." *International Journal of Computer Integrated Manufacturing* 21: 407-414.

Roy, R., Stark, R., Tracht, K., Takata, S., and Mori, M. (2016). "Continuous maintenance and the future – Foundations and technological challenges." *CIRP Annals – Manufacturing Technology* 65: 667-688.

Rozen, K., and Byrne, N. M. (2016). "Using simulation to improve semiconductor factory cycle time by segregation of preventive maintenance activities." *Proceedings of the Winter Simulation Conference. Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Samal, N. K., and Pratihar, D. K. (2015). "Joint optimization of preventive maintenance and spare parts inventory using genetic algorithms and particle swarm optimization algorithm." *International Journal of System Assurance Engineering and Management* 6: 248-258.

Santos, A. F., and Bispo, C. F. (2016). "Simulation based optimisation package for periodic review inventory control." *Proceedings of the Winter Simulation Conference, Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Sarker, B. R., and Faiz, T. I. (2016). "Minimizing maintenance cost for offshore wind turbines following multi-level opportunistic preventive strategy." *Renewable Energy* 85: 104-113.

Sarker, R., and Haque, A. (2000). "Optimization of maintenance and spare provisioning policy using simulation." *Applied Mathematical Modelling* 24: 751-760.

Scarf, P. A., and Cavalcante, C.A.V. (2012). "Modelling quality in replacement and inspection maintenance." *International Journal of Production Economics* 135: 372-381.

Scarf, P. A. (1997). "On the application of mathematical models in maintenance." *European Journal of Operational Research* 99: 493-506.

Shafiee, M., Finkelstein, M., and Bérenguer, C. (2015). "An opportunistic condition-based maintenance policy for offshore wind turbine blades subjected to degradation and environmental shocks." *Reliability Engineering and System Safety* 142: 463-471.

Shafiee, M., Patriksson, M., Chukova, S. (2016). "An optimal age–usage maintenance strategy containing a failure penalty for application to railway tracks." *Proceeding of Institution of Mechanical Engineers Part F: Journal of Rail and Rapid Transit* 230: 407-417.

Sharma, A., and Yadava, G. S. (2011). "A literature review and future perspectives on maintenance optimization." *Journal of Quality in Maintenance Engineering* 17: 5-25.

Silver, E. A., Pyke, D. F., and Thomas, D. J. (2016). *Inventory Management and Production Planning and Scheduling*. *Third Edition*. Wiley.

Smith, M. A. J. and Dekker, R. (1997). "Preventive maintenance in a 1 out of n system: The uptime, downtime and costs." *European Journal of operational research* 99: 565-583.

Tateyama, T., Tateno, T., and Shimizu, K. (2010). "Dynamic work planning by using simulation-based optimization in consideration of workers' skill and training." *Journal of Advanced Mechanical Design, Systems and Manufacturing* 4: 597–604.

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

Thomas, L. C., Gaver, D. P., and Jacobs, P. A. (1991). "Inspection models and their application." *IMA Journal of Management Mathematics* 3: 283-303.

Thörnblad, K., Strömberg, A-B., Patriksson, M. and Almgren, T. (2015). "Scheduling optimisation of a real flexible job shop including fixture availability and preventive maintenance." *European Journal of Industrial Engineering* 9: 126-145.

Van Horenbeek, A., Pintelon, L., and Muchiri, P. (2010). "Maintenance optimization models and criteria." *International Journal of Systems Assurance Engineering and Management* 1: 189–200.

Van Horenbeek, A., Buré, J., Cattrysse, D., Pintelon, L., and Van Steenwegen, P. (2013). "Joint maintenance and inventory optimization systems: A review." *International Journal of Production Economics* 143: 499-508.

Van Oosterom, C. D., Elwany, A. H., Çelebi, D., and Van Houtum, G. J. (2014). "Optimal policies for a delay time model with postponed replacement." *European Journal of Operational Research* 232: 186–197.

Vaughan, T. S. (2005). "Failure replacement and preventive maintenance spare parts ordering policy." *European Journal of Operations Research* 161: 183-190.

Volovoi, V. (2016a). “Simulation of maintenance processes in the big data era.” *Proceedings of the Winter Simulation Conference. Roeder, T. M. K., Szechtman, R., Zhou, E., Huschka, T., and Chick S. E. (Editors).*

Volovoi, V. (2016b). “Big data for reliability engineering: Threat and opportunity”. *IEEE Reliability Magazine*.

Wall, D. (2013). “Trunnion bearing replacement on a BOS Plant.” *Maintenance & Engineering: MaintenanceOnLine.Co.UK* (28/03/2013): 22.

Wang, H. Z. (2002). “A survey of maintenance policies of deteriorating systems.” *European Journal of Operational Research* 139: 469-489.

Wang, S., and Liu, M. (2015). “Multi-objective optimization of parallel machine scheduling integrated with multi-resources preventive maintenance planning.” *Journal of Manufacturing Systems* 37: 182-192.

Wang, W. (2008). “Delay Time Modelling.” In *Complex System Maintenance Handbook*. Kobbacy, K. A. H., Murthy, D. N. P. (Editors). Springer, London, 345-370.

Wang, W. (2011a). “A joint spare part and maintenance inspection optimisation model using the delay-time concept.” *Reliability Engineering and System Safety* 96: 1535-1541.

Wang, W. (2011b). “An inspection model based on a three-stage failure process.” *Reliability Engineering and System Safety* 96: 838-848.

Wang, W. (2012a). “An overview of the recent advances in delay-time-based maintenance modelling.” *Reliability Engineering and System Safety* 106: 165-178.

Wang, W. (2012b). “A stochastic model for joint spare parts inventory and planned maintenance optimisation.” *European Journal of Operational Research* 216: 127-139.

Wang, W., Christer, A. H. (1997). “A modelling procedure to optimise component safety inspection over a finite time horizon.” *Quality and Reliability Engineering International* 13: 217-224.

Wang, W., and Syntetos, A. A. (2011). “Spare parts demand: Linking forecasting to equipment maintenance.” *Transportation Research Part E* 47: 1194-1209.

Wang W., and Wang, H. (2015). “Preventive replacement for systems with condition monitoring and additional manual inspections.” *European Journal of Operational Research* 247: 459-471.

Williams, J. W. (2014). “Simulation attacks manufacturing challenges.” *Proceedings of the Winter Simulation Conference. Tolk, A., Diallo, S. Y., Ryzhov, I. O., Yilmaz, L., Buckley, S., Miller, J. A. (Editors)*.

Xiang, Y., Cassady, C. R., and Pohl, E. A. (2012). “Optimal maintenance policies for systems subject to a Markovian operating environment.” *Computers and Industrial Engineering* 62: 190-197.

Yang, L., Ma, X., Zhai, Q., Zhao, Y. (2016). “A delay time model for a mission-based system subject to periodic and random inspection and postponed replacement.” *Reliability Engineering and System Safety* 150: 96-104.

Yoo, Y. K., Kim, K. J., and Seo, J. (2001). “Optimal joint spare stocking and block replacement policy (cost modelling of spare stocking and block replacement).” *International Journal of Advanced Manufacturing Technology* 18: 906-909.

Zahedi-Hosseini, F., Scarf, P. A., and Syntetos, A. A. (2017). “Joint optimisation of inspection maintenance and spare parts provisioning: a comparative study of inventory policies using simulation and survey data.” *Reliability Engineering and System Safety* 168: 306-316. http://dx.doi.org/10.1016/j.ress.2017.03.007.

Zhang, N., Fouladirad, M., and Barros, A. (2016). “Maintenance of a two dependent component system: a case study.” *IFAC-PapersOnLine* 49-12: 793-798.

Zhang, C., Gao, W., Guo, S., Li, Y., and Yang, T. (2017). “Opportunistic maintenance for wind turbines considering imperfect, reliability-based maintenance.” *Renewable Energy* 103: 606-612.