AN EVALUATION OF THE REINSPECTION DECISION POLICIES FOR SOFTWARE CODE INSPECTIONS

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

BY

SERKAN NALBANT

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

JANUARY 2005

Approval of the Graduate School of Natural and Applied Sciences.

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Çağlar GÜVEN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Gülser KÖKSAL
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Yasemin SERİN (METU, IE)

Assoc. Prof. Dr. Gülser KÖKSAL (METU, IE)

Assoc. Prof. Dr. Tayyar ŞEN (METU, IE)

Assoc. Prof. Dr. Ali Hikmet DOĞRU (METU, CENG)

M.Sc. Levent TÜREL (MilSOFT Software Technologies)

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

Name, Last Name :

Signature :

ABSTRACT

AN EVALUATION OF THE REINSPECTION DECISION POLICIES FOR SOFTWARE CODE INSPECTIONS

Nalbant, Serkan

M.S., Department of Industrial Engineering

Supervisor: Assoc. Prof. Dr. Gülser KÖKSAL

January 2005, 197 pages

This study evaluates a number of software reinspection decision policies for software code inspections with the aim of revealing their effects regarding cost, schedule and quality related objectives of a software project.

Software inspection is an effective defect removal technique for software projects. After the initial inspection, a reinspection may be performed for decreasing the number of remaining defects further. Although, various reinspection decision methods are proposed in the literature, no study provides information on the results of employing different methods. In order to obtain insight about this unaddressed issue, this study compares the reinspection decision policies by finding out and analyzing their performance with respect to designated measures and preference profiles for cost, schedule, and quality perspectives in the context of a typical Software Capability Maturity Model Level 3 software organization. For this purpose, a Monte Carlo simulation model, which represents the process comprising initial code inspection, reinspection, testing and field use activities, is employed in the study together with the experiment designed in order to consider different circumstances under which the mentioned process operates.

The study recommends concluding the reinspection decision by comparing inspection effectiveness measure for major defects with respect to a moderately high threshold value (i.e. 75%). The study also reveals that applying default decisions of ‘Never Reinspect’ and ‘Always Reinspect’ do not exhibit the most appropriate outcomes regarding cost, schedule, and quality. Additionally, the study presents suggestions for further improving the cost, schedule, and quality of the software based on the analysis of the experiment factors.

Key Words: Software reinspection, software code inspection, decision making, Monte Carlo simulation, design of experiments.

ÖZ

YAZILIM KOD MUAYENELERİNE YÖNELİK YENİDEN MUAYENEYE KARAR VERME POLİTİKALARININ KARŞILAŞTIRILMASI

Nalbant, Serkan

Yüksek Lisans, Endüstri Mühendisliği Bölümü

Tez Yöneticisi: Doç. Dr. Gülser Köksal

Ocak 2005, 197 sayfa

Bu çalışma yazılım kod muayeneleri için ortaya konan belirli sayıdaki yeniden muayeneye karar verme politikalarını, bu politikaların yazılım projesinin maliyet, takvim, ve kaliteye ilişkin hedefleri üzerindeki etkilerini ortaya çıkarmak amacıyla değerlendirmektedir.

Yazılım muayeneleri yazılım projelerinde oluşan hataların ortadan kaldırılması için etkili bir tekniktir. İlk muayaneden sonra yapılabilen yeniden muayene sayesinde ise yazılım ürününde kalan hatalar daha da azaltılabilir. Literatürde yeniden muayeneye karar vermeye yönelik bir çok metot önerilmesine rağmen, farklı metotların kullanılması sonucu karşılaşılan sonuçları gösteren bir çalışma bulunmamaktadır. Bu çalışma, literatürde değinilmemiş olan bu konu hakkında fikir sağlamak üzere, Yazılım Yetenek Olgunluk Modeli Seviye 3'e göre yapılanmış bir yazılım organizasyonu kontekstinde, yeniden muayeneye karar verme politikalarının maliyet, takvim, ve kalite bakış açılarına ilişkin performanslarını, belirlenen ölçülere ve tercih profillerine göre ortaya çıkarıp analiz ederek karşılaştırmaktadır. Bu amaç doğrultusunda, çalışmada, ilk kod muayenesi, yeniden muayene, test ve yazılımın sahada kullanılması faaliyetlerini kapsayan süreci temsil etmek üzere bir Monte Carlo simülasyon modeli, bu süreç kapsamındaki farklı koşulların dikkate alınmasını sağlayan bir deney tasarımı ile beraber kullanılmaktadır.

Çalışma, yeniden muayene kararının majör hataların etkililik ölçüsünün kısmen yüksek bir eşik değerine göre karşılaştırılması sonucu alınmasını tavsiye etmektedir. Ayrıca, 'Hiçbir Zaman Yeniden Muayane Yapma' ve 'Her Zaman Yeniden Muayane Yap' sabit kararlarının uygulanmasının maliyet, takvim, ve kalite açılarından en uygun sonuçları vermediği de çalışma sonuçları tarafından ortaya konmaktadır. Buna ilaveten, çalışma, deneyde içerilen faktörlerin analiz edilmesi sonucunda, yazılım maliyet, takvim, ve kalitesinin daha da geliştirilmesi için çeşitli öneriler sunmaktadır.

Anahtar Kelimeler: Yazılım yeniden muayenesi, yazılım kod muayenesi, karar verme, Monte Carlo simülasyonu, deney tasarımı.

To My Family

ACKNOWLEDGEMENTS

I send my kindest regards and thanks to Mrs. Köksal for providing invaluable advices related to my research and supporting me endlessly while I am working on thesis study.

I would like to express my sincere thanks to my family for always trusting me and tolerating all my fancies in the stressful days.

I also thank to my company for excusing and supporting me during my Master of Science education.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Today, software is a basic component of many businesses. In some areas, it is even impossible to survive for an organization without the use of associated computer software (e.g. banking, telecommunications). Because of this fact along with the increasing competition, advances in technology and enhancing capabilities of software development organizations, the need for more and more sophisticated software systems is growing constantly. The realization of such systems requires successful completion of complex projects. This means that the software product is delivered on time, within budget and with high quality. The accomplishment of these aims demands effective software processes which underpins the software development activities.

The delivery of high quality software is achieved by eliminating the defects injected in various phases of the software development life cycle, such as requirements analysis, software design, coding. Software organizations employ numerous techniques to remove the defects from the software, thus preventing their propagation to the user. These techniques include automated analysis of the software code, peer reviews and execution testing (generally performed in different levels such as unit testing, integration testing, acceptance testing). Among these techniques, peer review is getting more and more popular due to increasing recognition of its effectiveness in removing the defects in software. Also, software development related standards and models mandate the implementation of peer reviews. In the scope of peer review a software artifact is examined by the related project personnel (other than the author) with the aim of pointing out to defects and improvements.

Software Inspection is a special type of peer review, where the examination of the software work product is conducted according to well-defined procedures with certain stages and by personnel who are trained on inspection procedures and techniques. The aim of conducting a software inspection is to improve the software project's performance in terms of cost, schedule, and quality. Additionally, software reinspections may be performed after the initial inspection is completed in order to further reduce the number of defects in the software work product, thus to obtain more of inspection benefits. However, this requires increasing the amount of project resources that are allocated for inspecting the product. Hence, the project management should make the decision to devote higher level of valuable resources for scrutinizing the software work product anew rationally, i.e. without arbitrariness and basing it on quantitative data. Although, various objective decision methods are proposed for concluding this important decision in addition to ad-hoc and historical data based decision methods, currently there is no guidance for the software engineering practitioners for selecting the appropriate objective reinspection decision method among the available ones. Hence, in this study this niche is addressed by evaluating the performance of different objective reinspection decision methods for code inspections conducted in the context of a Software Capability Maturity Model Level 3 organization. This evaluation is based on the comparisons of different policies (constituted from the available objective decision methods) with respect to the outcomes they depict at the end of the project lifecycle. Namely, for each considered reinspection policy the resulting cost, schedule delay and defect containment are revealed with the aim of determining the ranking of the policy for different preference profiles. These profiles refer to the different weights assigned to cost, schedule, defectiveness due to varying preferences shaped according to organizational policies, project structure, software type etc. In order to observe the related outcomes for various conditions underlying the study, an experiment is designed by designating the factors that affect software inspection. Then, the policies are evaluated by conducting Monte Carlo simulations that execute the model of the study for the determined experiment design. The usage of simulation technique enables observing the effects of various policies under various conditions as software code is flowing through the life cycle without conducting actual inspections, testing etc., i.e. by just expending computer processor time. The study also utilizes the simulation results for providing guidance about the effects of changing factor levels on cost, quality, and schedule measures.

The study is presented as follows. In Chapter 2, first, software inspection and the issue regarding software defects are described in detail. This is followed by the explanation and comparison of the techniques that enable the estimation of the defect population in a software work product during inspection. Then, Chapter 2 concludes by providing information on software reinspections and the techniques to predict the number of defects to be found during a probable reinspection. Chapter 3 initially puts forward the method to be followed in the study along with the underlying simulation model. Then, the factors that are considerable for the purposes of the study are introduced, and by taking into account these factors, the experiment design which enables the representation of the different circumstances related to the study is constituted. Chapter 3 continues by identifying the reinspection decision policies which are compared with each other during the study. Further, the defect estimation techniques required while applying the selected reinspection policies are nominated. Afterwards Chapter 3 presents the simulation results that show the ranking of different policies under various preference profiles regarding cost, schedule, and quality. The ANOVA studies regarding the effects considered factors on output measures of the study completes Chapter 3. The succeeding chapter, namely Chapter 4, provides the analysis regarding the reinspection policy rankings and ANOVA studies along with the discussion on the validity of the study results. Chapter 4 also lists the suggestions for the software organizations which enable them to improve their cost, schedule, and quality performance by manipulating the factors considered in the scope of the study. Finally, in Chapter 5, the study is concluded by summarizing the study, portraying the overall findings and describing the potential research opportunities that may be considered in the future.

CHAPTER 2

BACKGROUND AND LITERATURE SURVEY

2.1. Software Peer Reviews

Peer Review is a process, where colleagues of the author, who developed the software work product, examine the product with intend to find defects and identify improvement opportunities (Wiegers, 2002). Wiegers defines a defect (also known as bug or fault) as "a condition in a software work product that would cause the software to produce an unsatisfactory or unexpected result". In software industry, peer reviews are employed to detect the defects in various work products such as requirements specifications, design descriptions, source code, planning documentation, test case descriptions, process descriptions, etc.

The software process improvement models such as Software Capability Maturity Model (SW-CMM) (Paulk et al., 1993), Capability Maturity Model Integration (CMMI) (CMMI Product Team, 2001), and Software Process Improvement and Capability Determination (SPICE) (ISO, 1998) impose peer reviews as an effective verification practice. SW-CMM deserves special attention among these models, since it is probably the most widely used and recognized one by software organizations (although it is recently started to be replaced by its extension CMMI). Even only recent (between January 2000 and June 2004) SW-CMM assessments are considered, the number of organizations assessed with respect to SW-CMM adds up to 1,543 (Software Engineering Institute, 2004). SW-CMM enables software development organizations to improve their software processes. On the other hand, it helps the software acquiring organizations in assessing the quality of their contractors. For supporting the process improvement goal, SW-CMM puts forward a framework through which software organizations selects process improvement strategies after determining the current level of their process maturity and identifying the key factors that would lead to improvement. The underlying assumption of SW-CMM, as in all process models, is better processes lead to improved quality in the product. When a software organization is assessed with respect to SW-CMM, it is given one of the five available ratings. These ratings represent the software development maturity of the organization and called as 'maturity level'. For acquiring a certain maturity level a software organization must meet the requirements of this level and the requirements of the lower levels. Each maturity level is composed of a number of processes, which are called 'Key Process Area (KPA)', that correspond to these requirements. A KPA is defined as, "a cluster of related activities that, when performed collectively, achieve a set of goals considered important for establishing process capability" (Paulk et al., 1993). Hence, the software organizations must prove that it implements the related KPAs required by the aimed maturity level. One of the KPAs put forward by maturity level three is 'Peer Reviews'. The goal of Peer Review KPA is to remove and identify the defects in the software work products by performing planned peer review activities. In order to accomplish this the following practices shall be conducted; (i) a peer review policy is designated, (ii) resources are allocated to perform peer reviews, (iii) peer review participants are trained about peer reviews, (iv) planned peer reviews are carried out according to documented procedures by reviewers who have defined roles and by using checklists, (v) the actions identified during the peer review performed in order to remove the defects, (vi) quantitative data regarding the peer review are stored.(Paulk et al., 1993).

Actually, peer review is an umbrella term used to denote the different kinds of processes that enables manually examining the software work product for finding defects, i.e. there exists different types of peer reviews. 'IEEE Standard for Software Reviews' classifies peer reviews into three as (i) Technical Review, (ii) Walkthrough and (iii) Inspection (IEEE, 1998). The following definitions are put forward by this standard for these peer review types (IEEE, 1998):

(i) Technical Review: A systematic evaluation of a software product by a team of qualified personnel that examines the suitability of the software product for its intended use and identifies discrepancies from specifications and standards. Technical reviews may also provide recommendations of alternatives and examination of various alternatives.

(ii) Walkthrough: A static analysis technique in which a designer or programmer leads members of the development team and other interested parties through a software product, and the participants ask questions and make comments about possible errors, violation of development standards, and other problems.

(iii) Inspection: A visual examination of a software product to detect and identify software anomalies, including errors and deviations from standards and specifications. Inspections are peer examinations led by impartial facilitators who are trained in inspection techniques. Determination of remedial or investigative action for an anomaly is a mandatory element of a software inspection, although the solution should not be determined in the inspection meeting.

Among these peer review types, inspection deserves special attention because of the following reasons:

- Inspection is more effective than other peer review types in terms of defect removal (Wiegers, 2002, Gilb and Graham, 1993, Radice, 2002).
- Many software organizations which undertake initiatives (such as SW-CMM, CMMI, ISO 9000, Six Sigma) to improve quality of their processes and products employ inspection.
- Inspection is more rigorous and systematic than other peer review types (Wiegers, 2002, Radice, 2002, Gilb and Graham, 1993).

A more detailed discussion on the differences between software inspection and other peer review types is available in Wheeler et al. (1997).

2.2. Software Inspections

Software Inspections are introduced to software community by Michael Fagan as a result of his development efforts regarding the inspection methodology at IBM in the early 1970s (Fagan, 1976). The main goal of software inspection is to remove the defects in the software work product right after their injection. In this way inspection enables; (i) saving of rework cost and development time, which needs to be expended if the defects pass to later stages of software development life cycle and (ii) improves the quality of software product by enhancing its reliability, maintainability and availability. Software code is the work product for which software inspection is applied in most cases, i.e. applying software inspections is more common for software code vis-à-vis other work products such as requirements specifications, design descriptions etc (Laitenberger and DeBaud, 2000).

'IEEE Standard for Software Reviews' put forwards the scope of the software inspection by listing its objectives as follows (IEEE, 1998):

- To verify that software product;
  - satisfies its specifications,
  - satisfies specified quality attributes,
  - conforms to applicable regulations, standards, guidelines, plans, and procedures.
- To identify deviations from standards and specifications.
- To collect software engineering data (for example, anomaly and effort data) (optional)
- Uses the collected software engineering data to improve the inspection process itself and its supporting documentation (for example, checklists) (optional).

The original inspection process proposed by Fagan consists of the following five activities (Fagan, 1976):

- **Overview:** The inspection participants obtain and are informed about the work product to be inspected.
- **Preparation:** Participants individually examine the work product to develop an understanding on the work product and to generate the issues that they deem as defects.

- **Inspection Meeting:** The inspection team finalizes and documents the list of defects in the work product by carrying out necessary discussions on the issues found during preparation and raised during the inspection meeting while going over the work product.
- **Rework:** The defects are corrected by the owner of the work product.
- **Follow-Up:** Inspection Moderator verifies whether the defects are resolved appropriately. If more than 5% of the work product is affected during the rework, a reinspection is conducted.

After Fagan's initial work, software inspections are employed widely by software organizations as an effective quality control technique. Also, many variations of original inspection process emerge as a result of widespread application of inspections in the software industry due to different needs of the organizations and because of the attempts to design a more efficient inspection process. These variations incorporate new activities to inspection process such as Entry Condition Checking, Planning, Data Recording, Consolidation, Entry Condition Checking and Prevention Meeting. Furthermore, among these variations organization, number of activities, participant roles, number of participants, the work product type, the reading techniques employed during preparation differ. MacDonald and Miller (1997) describe main inspection methods with the aim of developing an inspection process definition language which represents various inspection methods by considering their commonalities and differences. In this study, they list the following as the main inspection methods along with the related references; (i) Fagan Inspection, (ii) Structured Walkthrough, (iii) Humphrey's Inspection, (iv) Gilb Inspection, (v) Asynchronous Inspection, (vi) Active Design Reviews, (vii) Phased Inspection, (viii) N-Fold Inspection.

Also, Laitenberger and DeBaud (2000) manifest the following activities, which characterize most inspection methods, with the aim of having a reference model while discussing the similarities and differences among inspection methods:

- **Planning:** The inspection is organized by selecting the participants, assigning roles (such as moderator, recorder, reader etc.) to participants, scheduling the inspection meeting, and distribution of the inspection material.

- **Overview:** The work product is explained to the participants in order to facilitate their inspection and understanding.
- **Defect Detection:** Work product is examined with the aim of finding defects. During this activity, reading techniques are employed for facilitating defect identification. Among various inspection methods, there is no consensus on whether this activity should be carried out individually, in groups or both.
- **Defect Collection:** The issues that are accepted as defects are consolidated and documented. Furthermore, the decision for a second inspection is made.
- **Defect Correction:** The author makes the editions regarding the defects accepted in collection activity.
- **Follow-Up:** The resolution of the defects accepted in collection activity is ensured by checking the reworked work product.

As evident from the above explanations, software inspection is a group activity conducted by a number of people assigned among the members of the project team. The group that is composed of people participating to the inspection is generally referred as inspection team. The two factors related to an inspection team are team roles and team size.

Inspection participants perform different roles, whose proper conduct is critical for the success of the inspection. There are various roles put forward by different inspection methods. These can be listed as; Organizer, Moderator, Inspector, Reader/Presenter, Author, Recorder, Collector (Laitenberger and Debaud, 2000). Among these the main roles are author and inspector. The others actually represent the additional duties performed by the inspectors, i.e. in the course of an inspection cycle a person generally performs more than one role. An inspector is responsible for finding the defects in the inspected artifact. Whilst, an author is responsible for answering the questions related to the artifact during inspection meeting and for correcting the defects identified. Usually, all team members are assumed as inspectors regardless of other roles they are assigned. There are two exceptions to this, namely, organizer, who plans inspection activities to be performed in the course of the project (generally the project manager), and author, who must not evaluate the work product for inspection purposes due to independency constraints.

The team size of a software inspection is determined by the number of inspectors plus the author (in some cases multiple people may attend as the author). The information regarding the average number of inspectors in an inspection team vary greatly throughout the literature. Actually, this is an expected outcome, since the appropriate value for ‘number of inspectors’ depends on many factors such as, the type of inspected artifact, the availability of resources, budget allocated to software inspections, the size of the artifact, and the criticality of the software developed (e.g. for a software product whose malfunction may cause losing of human life, the number of inspectors may be very high, since it is positively correlated with the number of defects found). This is also true for software inspections regarding software code. Wiegers (2002) states that two inspectors are usually sufficient while performing code inspection. Whilst, Radice (2002) suggests four as the maximum number of inspectors that should be allocated to code inspections. He also adds that the values lower than four can be equally effective with respect to aim of finding all available defects. Further, a study exploring the effect of varying values regarding the number of inspectors in code inspections considers 1, 2, and 4 as the inspector number (Porter et al., 1997). The results of this study provide evidence for the suggestion of Radice (2002), since it founds little difference between the effectiveness of code inspections performed with 2 and 4 inspectors. Whilst, conducting code inspections with one inspector is found to be the least effective of all.

There are many studies which mention success stories regarding software inspections. These studies report that inspection may detect and remove between 30% and 93% of the defects in the software (Laitenberger and DeBaud, 2000). In a study, Briand et al. (1998a) perform a simulation by using the published inspection data and find out 57% as the benchmark value for the ratio of defects eliminated by the inspection. They also report that code and design inspections save 39% and 44% of defect removal costs, respectively, vis-à-vis testing. Some studies also focus on the maintenance effort saved by applying software inspections. For example, Russell (1991) and Doolan (1992) state that each hour spent for inspections avoids a rework effort about 33 and 30 hours, respectively, during the maintenance phase. More comprehensive information on experiences with software inspection can be found in

Radice (2002), Wheeler et al. (1996), Gilb and Graham (1993) and Laitenberger and DeBaud (2000).

Besides these quantitative results, the benefits of software inspection can be listed qualitatively as below (Radice 2002, Wiegers, 2002, Gilb and Graham, 1993):

Software inspection;

1. Decreases the number of defects pass on to the testing and field use.
2. Reduces development cycle time.
3. Increases the probability of delivering the software on schedule.
4. Saves from testing, maintenance and support costs.
5. Reduces testing and debugging time.
6. Improves productivity (effort spent per unit code size).
7. Supports knowledge sharing and education.
8. Enhances teamwork and collaboration.
9. Provides early information on the quality of end product.

In order to obtain the above benefits, a software organization should invest in software inspections, which requires the expending of related costs. These include the start-up costs spent while deploying the software inspections throughout the organization (such as training, process definition, and adaptation costs), and implementation costs spent while actually carrying out the inspection steps. The latter one, in addition to indirect costs (such as overhead), is largely determined by the personnel effort used to perform software inspections. Since software inspection is a human-based activity, the studies reporting related cost values provide data in terms of effort spent per unit artifact size or effort spent per defect. A number of studies provide data on these costs for code inspections. Briand et al. (1998a), Laitenberger and Debaud (2000) and Radice (2002) provide good summaries of these studies. The ranges obtained from these summaries are listed in Table 1.

Table 1 Ranges of Published Code Inspection Cost Data

Cost Data Description	Range (in hours)
Individual preparation effort in code inspections per thousand lines	4.91-7.9
Meeting time in code inspections per thousand lines	3.32-4.4
Average effort to find and fix a defect in code inspections	0.2-2.7

Several studies publish the Return on Investment (ROI) values obtained by applying software inspections. In these studies, the return acquired from software inspections are expressed in terms of rework savings gathered due to early detection of defects. Grady and Van Slack (1994) report a ROI of 10.4 for Hewlett-Packard's inspection program, which resulted an estimated saving of 21.5 million dollars in 1993. By employing the data of a software organization, Mah (2001) reports the software inspection ROI values, which are 7 for code inspections and, 14 for design and requirements inspections. In addition to these, the results obtained from a more comprehensive study, namely National Software Quality Experiment (NSQE) (in which about 80 organizations participate by sharing their software quality related data), depict ROI values between 2 and 8 for the participating organizations (O'Neill, 2002). By employing NSQE data, O'Neill (2003) also figures out the code inspection ROI values for organizations that have varying degrees of process maturity. According to his results, an organization that implements structured software engineering (corresponds to SW-CMM Maturity Level 3) can obtain 6 as ROI value.

Other important information related to cost of performing software inspections comprises the rates regarding preparation and meeting steps of the inspection. Preparation rate is defined as the average quantity of material covered per labor hour of individual preparation (Wiegers, 2002). Whilst, meeting rate is the average quantity of material inspected per meeting hour (Wiegers, 2002). If the average values of these rates are known for an organization, a project manager can calculate the expected cost a software inspection based on the given size of the artifact that will be inspected. Although the published data about these rates varies for code inspections, some sources provide guidance on suitable preparation and meeting rates. Namely, Tervonen and Iisakka (1997), Radice (2002), and Wiegers (2002) consistently suggest 150-200, 100-200, 150-200 source lines of code (SLOC) per hour, respectively, for both rates.

The literature also focuses on the ways for increasing the effectiveness of defect detection phase. These are commonly referred as 'Reading Techniques'. A reading technique can be defined as a series of steps that provides direction to the inspector on the ways for checking the work product and facilitates his/her understanding. Although the description of different reading techniques is out of the scope of this study, the main reading techniques can be named with related references as; Ad-Hoc Reading (i.e. no explicit guidance is available for the inspector) (Doolan, 1992), Checklist Based Reading (Fagan, 1976), (Gilb and Graham, 1993), Scenario-based Reading (Basili, 1997), (Cheng and Jeffrey, 1996) , Defect-based Reading (Porter et al., 1995), Traceability-based Reading (Travassos et al., 1999), Perspective-based Reading (Basili et al., 1996), Reading by Stepwise Abstraction (Dyer, 1992), Usage-based Reading (Thelin et al., 2001). Among these the most widely used one is Checklist Based Reading. Laitenberger and DeBaud (2000) provides a good discussion on the different reading techniques by comparing their main characteristics, which are Application Context, Usability, Repeatability, Adaptability, Coverage, Overlap among Inspectors. Ad-hoc Reading and Checklist-based Reading are most commonly used reading techniques throughout the software industry (Freimut et al., 2001, Laitenberger and DeBaud, 2000). Ad-hoc Reading stands for examining the work product without employing any specific reading technique. However, regardless of how much widespread is a particular reading technique, a number of studies report on the experiments for comparing different reading techniques. Examples for such studies can be found in Thelin et al., (2003), Porter et al., (1995), Laitenberger et al., (2000), Laitenberger et al., (2001).

Testing is another activity which aims the detection and removal of the software defects in the software code. Testing is generally performed after the inspection is completed. During testing the software code is executed and evaluated with respect to test cases, which outlines the actions to be performed and expected outcomes. When any deviation occurs from the expected outcomes, this is treated as a defect.

Usually software development organizations conduct different levels of tests such as unit test, unit integration test, system test, which are related to different states of software code. Testing verifies the software code dynamically, i.e. as it is working, whilst inspection statically analyzes it. For instance, testing can not reveal the problems regarding the maintainability of the code, which may cause the expending of high amount of rework effort due to increased difficulty for identifying the location that should be modified. Similarly, it may be hard for an inspector to see the malfunctioning that will be encountered as code is run. However, there is no agreement in the literature regarding whether inspection and testing are mutually exclusive alternatives, i.e. if any defect can be identified by both inspection and testing or some defects can be only detectable only by one of two alternatives. For example, Gilb and Graham (1993) claim that although there are defects catchable by both testing and inspection, some defects are only detectable through testing and some are only detectable with the means of inspection. Consequently, according to them the two methods are complementary for each other. On the contrary, a study conducted by Laitenberger (1998) shows no evidence for the claim that states testing and inspection enable the finding of different defect classes. Besides, many studies that explore the savings gained by applying software inspections assume that both inspection and testing are capable of identifying a certain defect (examples can be found in (O'Neill, 2003, Radice, 2002, Gilb and Graham, 1993). Furthermore, a number of reports show that software inspections are more efficient than testing in terms of average effort spent to find a single defect. For example, a banking computer services firm's data is reported by Ackerman et al. (1989), where 2.2 hours are spent to remove a defect via code inspections on the average, whilst a value of 4.5 hours is observed during testing. Further, other illustrative reports provides the following values for the average effort (in terms of hours) required to eliminate a defect via code inspections and testing, respectively; 1 and 6 (Franz and Shih, 1994), 1.46 and 17 (Kelly et al., 1992), 1 and 6 (Weller, 1993). Hence, it can be concluded that removing a defect with the means of inspection is cheaper than removing it via testing. The main reason for this finding is the ease of locating and fixing a defect during code inspection vis-à-vis testing. Testing reveals symptom of a failure, so the project team should spent time to locate the problematic statements in the software code. However, since the software code is directly examined in the course of an inspection, the root cause of a defect is located when it is found by the inspector.

2.3. Software Code Defects and Relevant Studies

A code defect is a condition which causes the software code to deviate with respect to expectations. These expectations are determined by (i) the standards that the code must conform to, (ii) the design that the code must comply with, (iii) the requirements that code must meet in order to satisfy the needs of the users, and (iv) the results (obtained by running the code) that code must depict. Code defects are generally classified according to their severity as major and minor defects. Severity refers to the significance of the adverse effects that are caused by a defect. A major defect affects the proper execution with respect to requirements put forward for the software. Thus it represents a problem for the user, if it remains undetected until field use. On the other hand, a minor defect generally refers to format, writing or representation errors that does not impede/halt the execution, but it may still be problematic for the user (although a work around solution exists) or be important regarding the maintenance of the software.

As defects are identified, they need to be removed from the software by performing necessary corrections. Certainly, this requires the spending of additional rework costs. The amount of this cost usually escalates as the artifacts progress to later phases of software development life cycle. By considering the published data, Radice (2002) provides the summary of defect removal costs (usually given in terms of personnel effort as in most of software engineering studies), for various phases where defect is encountered, namely inspection, test and field use. Further, with the aim of performing return on investment and saving analysis regarding software inspections, he utilizes this information to figure out the relative cost values to fix software defects identified during inspection, test and field as 1,10, and 100, respectively (Radice, 2002). Also, NSQE study supplies data regarding the cost to repair a code defect (O'Neill, 2003). As mentioned before in the text this study is a comprehensive one, since it considers data obtained from about 80 organizations. By using the data obtained from NSQE, O'Neill (2003) reports that for an organization operating according to SW-CMM Maturity Level 3 practices, a major defect consumes an additional repair effort of 5-7 times during testing when compared to inspection, whilst a minor defect requires 3 times more effort. Further, he states that these ratios are also same when the repairing efforts during testing and field use are compared. Besides these quantitative costs, the defects found by the customer during usage also results in loss of goodwill for the software organization due to dissatisfaction of the customer.

The literature also provides guidance about the techniques for predicting the number of defects contained in the software code. These techniques enable the managers to assess the project progress, to plan the defect detection activities, to evaluate the quality of software product, and to carry out process improvement initiations. Most of the defect prediction techniques employ historical defect data. Namely, these techniques are; Empirical Defect Prediction (Humphrey, 1999), Orthogonal Defect Classification (Chillarege et al., 1992), Fault Proneness Evaluation (Selby and Basili, 1991), and Statistical Process Control (Florac and Carleton, 1999). Since they rely on the historical data, the application of such techniques requires data collected from the environment of the specific software development organization, which intends to use prediction techniques. So, these techniques do not help for obtaining benchmark information regarding the number of defects present in the software code at different phases of software development life cycle. However, another technique called Constructive Quality Model (COQUALMO) enables the prediction of the software defect level without the usage of historical data (Boehm et al., 2000). COQUALMO is actually an extension to well known Constructive Cost Model II (COCOMO II), which deals with the estimation of the cost, effort, and duration required to complete software projects. COQUALMO comprises of two submodels called Defect Introduction (DI) Model and Defect Removal (DR) Model.

DI model enables to predict the number of the defects injected into a software product given the size of the code. This is accomplished by adjusting the baseline defect values with the parameters regarding the environment of a software project. These parameters are specifically called Defect Introduction Drivers. DI submodel of COQUALMO puts forward 21 factors that determine the drivers. These factors are grouped into four categories as platform, product, personnel, and project as listed in Table 2 (Boehm et al., 2000).

Table 2 Defect Introduction Factors for COQUALMO Defect Introduction Submodel

Category	Defect Introduction Factor
Platform	Required Software Reliability (RELY) Data Base Size (DATA) Required Reusability (RUSE) Documentation Match to Life-Cycle Needs (DOCU) Product Complexity (CPLX)
Product	Execution Time Constraint (TIME) Main Storage Constraint (STOR) Platform Volatility (PVOL)
Personnel	Analyst Capability (ACAP) Programmer Capability (PCAP) Applications Experience (AEXP) Platform Experience (PEXP) Language and Tool Experience (LTEX) Personnel Continuity (PCON)
Project	Use of Software Tools (TOOL) Multisite Development (SITE) Required Development Schedule (SCED) Precedentedness (PREC) Architecture/Risk Resolution (RESL) Team Cohesion (TEAM) Process Maturity (PMAT)

The baseline number of code defects put forward by DI submodel is 33 per 1000 source lines of code. The total number of code defects can be computed by the following formula (Boehm et al., 2000).

$$D_{ICode} = D_{BCode} \cdot S \cdot \prod_{i=1}^{21} (DI - driver)_i \qquad \text{(Eq. 1)}$$

where,

D_{ICode}: Estimated number of code defects introduced

D_{BCode}: Baseline rate for code defect introduction per 1000 source lines of code (SLOC)

S: Size of the software code in kilo source lines of code (KSLOC)

$(DI\text{-}driver)_i$: Defect introduction driver corresponding to factor i

The value of a particular factor's defect introduction driver is designated by determining the rating corresponding to the as Very Low, Low, Nominal, High, Very High, and Extra High. If the rating of a specific factor is selected as Nominal, the value of the driver becomes 1. Whilst, a driver value less than 1 is found if the related factor's rating (different than nominal) affects the defect introduction positively (i.e. less defects are injected), and the driver value is greater than 1, otherwise. For instance, $(\text{DI-driver})_{\text{PCAP}}$ values are 0.76 and 1.32, for the cases where the rating of programmer capability factor is 'Very High' and 'Very Low', respectively. Consequently, if all factors are at their nominal level for a software project, this means that the predicted number of code defects introduced as a result of coding activity would be 33 for a 1 KSLOC of software code.

Defect Removal submodel of COQUALMO enables the prediction of the percentage of defects removed by applying certain defect removal activities, namely, automated analysis, people reviews, and execution testing and tools. For each of these activities, six different defect removal levels (ratings) are designated as Very Low, Low, Nominal, High, Very High, and Extra High. The description of the profiles that constitute the defect removal levels are given in Table 3 (Boehm et al., 2000).

Table 3 Defect Removal Profiles for COQUALMO Defect Removal Submodel

Rating	Automated Analysis	People Reviews	Execution Testing and Tools
Very Low	Simple compiler syntax checking.	No people review.	No testing.
Low	Basic compiler capabilities for static module-level code analysis, syntax, type-checking.	Ad-hoc informal walkthroughs. Minimal preparation, no Follow-up.	Ad-hoc testing and debugging. Basic text-based debugger.
No-minal	Some compiler extensions for static module and inter-module level code analysis, syntax, type checking. Basic requirements and design consistency, traceability checking.	Well-defined sequence of preparation, review, minimal follow-up. Informal review roles and procedures.	Basic unit test, integration test, system test process. Basic test data management, problem tracking support. Test criteria based on checklists.
High	Intermediate-level module and inter-module code syntax and semantic analysis. Simple requirements/ design view consistency checking.	Formal review roles and procedures applied to all products using basic checklists, follow up.	Well-defined test sequence tailored to organization (acceptance / alpha / beta / flight etc.) test. Basic test coverage tools, test support system. Basic test process management.
Very High	More elaborate requirements/design view consistency checking. Basic distributed-processing and temporal analysis, model checking, symbolic execution.	Formal review roles and procedures applied to all product artifacts & changes (formal change control boards). Basic review checklists, root cause analysis. Use of historical data on inspection rate, preparation rate, fault density.	More advanced test tools, test data preparation, basic test oracle support, distributed monitoring and analysis, assertion checking. Metrics-based test process management.
Extra High	Formalized specification and verification. Advanced distributed processing and temporal analysis, model checking, symbolic execution.	Formal review roles and procedures for fixes, change control. Extensive review checklists, root cause analysis. Continuous review process improvement. User/Customer involvement, Statistical Process Control.	Highly advanced tools for test oracles, distributed monitoring and analysis, assertion checking. Integration of automated analysis and test tools. Model-based test process management.

For each of the above levels a defect removal percentage is assigned. By using the following formula and determining the inherent defect removal ratings, a software project can predict the number of remaining defects in the software code when it is deployed to the customer (Boehm et al., 2000).

$$D_{RCode} = D_{ICode} \cdot \prod_{i=1}^{3} (1 - DRF_i) \qquad \text{(Eq. 2)}$$

Where,

D_{RCode}: Estimated number of residual code defects

D_{ICode}: Estimated number of code defects introduced

i: index for defect removal activities. i is equal to 1,2,3 for automated analysis, people reviews, and execution testing and tools, respectively.

DRF_i: Code defect removal fraction corresponding to defect removal activity i.

COQUALMO puts forward the DRF values in Table 4, for various defect removal ratings and activities (Boehm et al., 2000). For example, if all defect removal ratings are at their nominal level and if the number of introduced defects is 33, then the number of residual code defects is predicted as, 33×(1-0.2) ×(1-0.48) ×(1-0.58) ≅ 6 per KSLOC.

Table 4 Code Defect Removal Fractions for COQUALMO Defect Removal Submodel

Rating	**Automated Analysis**	**People Reviews**	**Execution Testing and Tools**
Very Low	0.00	0.00	0.00
Low	0.10	0.30	0.38
Nominal	0.20	0.48	0.58
High	0.30	0.60	0.69
Very High	0.48	0.73	0.78
Extra High	0.55	0.83	0.88

2.4. Defect Content Estimation Techniques for Software Inspection

Defect Content Estimation Techniques (DCET) enable the estimation of the number of defects contained in the software work product that is subject to inspection. This information can be employed by the inspection team to estimate the number of remaining defects in the work product by subtracting the number of defects found in the inspection from the estimate for total number of defects contained. The number of remaining defects is important to make informed decisions about performing a second inspection activity (i.e. reinspection), where the inspection team repeats the inspection process with the aim of reducing the number of defects to a more suitable level before the work product is passed to the next phase of the development life cycle. The DCETs in the literature can be classified as objective techniques and subjective techniques. Objective techniques are further categorized as curve fitting models and capture-recapture models.

The main property of subjective techniques, which favors them with respect to objective techniques, is the ease of obtaining the estimate. Estimating the number of defects present in an inspected work product with subjective techniques is simpler because it does not require any data collection other than asking the guess of an individual inspector or a group of inspectors. However, its dependence on the human judgement (i.e. on the knowledge and capability of the person(s) from which the estimate is requested) is the main drawback of the subjective techniques, although the initial study on using subjective estimates, claims that these techniques can perform satisfactorily in terms of accuracy (El Amam et al., 2000). On the other hand, whilst objective techniques do not depend on personal opinion, they are more costly than subjective ones, since they require the rigorous collection of high amount of data. This data shall be sophisticated enough to depict the defects that are found by a particular inspector and the inspectors that catch a particular defect. The following sub-sections describe objective and subjective DCETs in more detail.

2.4.1 Curve Fitting Models

The idea of fitting curves to the defect data collected during inspection with the aim of defect content estimation is originated by Wohlin and Runeson (1998). They propose two methods based on sorting and plotting the defect data gathering during inspection; Detection Profile Method (DPM) and Cumulative Method. In DPM, first for each defect the number of inspectors that found a particular defect is found out. Then a plot, where the defect index is located in x-axis and number of inspectors identified the defect is located in y-axis, is created. After, sorting the defect indexes according to decreasing value for number of inspectors, an exponential curve is fitted to the scattered data. Lastly, the exponential curve is utilized to produce the estimate for total number of defects. In particular, this estimate is the largest integer in the x-axis for which the y-axis value of the exponential curve equals or greater than 0.5. Similarly, Cumulative Method plots the cumulative number of inspectors in the y-axis starting with the defect that is captured by the highest number of inspectors, and adding the inspector number corresponding to other defects in the decreasing order (e.g. if defect a, b and c are found by 5, 3 and 8 inspectors respectively, the plot depicts three bars from left-to-right with values 8, 13, 16). Again, after fitting an exponential function to this plot, the defect estimate can be produced by employing some reliability models. Although this original study does not report one of the two method as superior than the other, the succeeding studies focus on DPM by replicating the original procedure or proposing and conducting its variations (for example fitting linear, quadratic or other types of exponential functions to the plotted data) (Thelin and Runeson, 2000, Briand et al., 1998b). However, in these studies the performance of the tested alternatives is similar or inferior from the initially proposed DPM (Freimut et al., 2001).

2.4.2 Capture-Recapture Models

The capture-recapture models are adapted to software engineering from the biology field which developed them to estimate the size of animal populations. In order to produce this estimate, biologists settle to an area where the animal population lives. Then, they start to capture the animals, whose population size to be revealed, by conducting trapping occasions (i.e. the different days in which the capturing is performed). When an animal is captured in an occasion, it is marked, and released back to its habitat. If a marked animal is caught again in another trapping occasion, it is said that the animal is recaptured and the animal’s tag is noted. The information collected in this way (i.e. as a result of the completion of all trapping occasions) is then used to make population size estimations based on statistical inference.

Biostatisticians proposed different models for open and closed populations to estimate the population size. An open population's size changes from one trapping occasion to another due to birth, death, migration etc., whilst the size of a closed population is assumed to be constant between trapping occasions. For estimating the fault content in an inspected software artifact, the closed population capture-recapture models are appropriate, since the defects in the artifact are fixed. So, for illustrating how the capture-recapture data is employed to estimate the size of an animal population, an estimator for closed populations should be used. For such a demonstration, consider the following: Suppose that, in a Capture-Recapture study which is composed of two trapping occasions, the following data is produced; $\mathbf{n_1}$: the number of animals captured in the first occasion, $\mathbf{n_2}$: the number of animals captured in the second occasion, $\mathbf{n_{12}}$: the number of animals captured in both occasions. By assuming the percentage of the recaptured animals in the second occasion is equal to the percentage of animals captured in the first occasion with respect to the population size (say **N**), the estimation for the population size can be generated as follows:

$$\frac{n_1}{\hat{N}} = \frac{n_{12}}{n_2} \Rightarrow \hat{N} = \frac{n_1 \times n_2}{n_{12}} \qquad \text{(Eq. 3)}$$

In biology and wildlife research, the above formulation is one of the basic estimators to estimate the population size and it is known as Lincoln-Peterson Estimator (Seber, 1982). However, this estimator is not applicable for capture-recapture studies with more than two occasions.

In biological studies the closed population capture-recapture models can be classified with respect to their assumptions about catchability of animals (Otis et al., 1978). These models consider the following three factors as the sources which result in variations regarding the catchability of an animal (i.e. the probability that a particular animal will be captured):

- **Time Response:** The catchabilityof an animal differs according to the trapping occasion. For example, in cold days most of the animals prefer not to go out of their homes. So, the number of animals captured in such a day is expected to be lower when compared to a tepid day.

- **Heterogeneity:** Different animals possess different capture probabilities. For instance, old animals, which are less mobile, are captured less often vis-à-vis young animals.
- **Behavioral Response:** The catchability of an animal changes, when it is captured. Hence, a previously marked animal will avoid the trap or will become attracted with it based upon its past experience (e.g. the animal tends to be captured again, if during its first encounter with the trap, it obtained food without getting hurt).

The eight closed population capture-recapture models constituted with different assumptions regarding the source of the variation (i.e. a certain combination of the above factors) in capturing probability of the animals are listed in Table 5 (Otis et al., 1978). As can be understood by examining this table, the models are named by incorporating the corresponding letter according to the sources of variation considered while building the model. Namely, letters **t**, **h** and **b** are used to denote the time response, heterogeneity and behavioral response, respectively. For each of these models at least one population size estimator is proposed by the researchers in the biology field, see for example (Otis et al., 1978).

Table 5 Closed Population Capture-Recapture Models & Their Sources of Variation

Model Name	Sources of Variation
Model M0	None, i.e. capture probabilities are same regardless of any factor.
Model Mt	Capture Probabilities vary with time.
Model Mb	Capture Probabilities vary by behavioral response to capture.
Model Mh	Capture Probabilities vary by individual animal.
Model Mtb	Capture Probabilities vary by time and behavioral response to capture.
Model Mth	Capture Probabilities vary by time and individual animal.
Model Mhb	Capture Probabilities vary by individual animal and behavioral response to capture.
Model Mtbh	Capture Probabilities vary by behavioral response to capture, time, and individual animal.

Regardless of the sources of variation taken into account, the following assumptions are valid for all of the above models (Otis et al., 1978).

1. The population is closed, i.e. animals enter or leave the population due to death, birth or migration reasons.
2. Animals do not lose their marks during the experiment.
3. All marks are correctly noted and recorded at each trapping occasion.

Actually, in addition to the above ones, a fourth assumption is manifested in the original work by Otis et al. (1978), as “each animal has a constant and equal probability of capture on each trapping occasion”. Otis et al. (1978) state that one of the objectives of their work is to relax this assumption by putting forward the models with behavioral response. For other models, i.e. for models M0, Mt, Mh, and Mth all of the four assumptions are applicable.

The first adaptation of estimators based on capture-recapture models to the field of software engineering was made by Mills (1972), who applied the Lincoln-Peterson estimator with the aim of obtaining an estimate regarding the number of defects remaining after testing. Later, Eick et al. (1992) applied the capture-recapture techniques to estimate the number of defects remaining after a software inspection. The application of the capture-recapture models for software inspection is based on considering the defects as animals and the inspectors as trapping occasions. More clearly, the defects in the inspected artifact represent the population, and from this population the inspector obtains a sample as he finds the defects in the artifact. If the same defect is identified by two or more inspector, this defect is referred as recaptured. By using this analogy, it is possible to estimate the number of defects in the inspected artifact. This is achieved by collecting data on the particular defects identified by each inspector during software inspection, and using the capture-recapture estimators to calculate the estimate for total number of defects. Then, by using this estimate and the actual number of defects found in the inspection, the inspection team or project manager can estimate the number of remaining defects in the artifact. Certainly, since the estimation is based on the degree of commonality regarding the defects found by different inspectors, in order to employ capture-recapture models for a particular software inspection, the number of inspectors participate in the inspection shall be at least two. A good description for the rationale regarding the usage of capture-recapture models to estimate the number of defects in a software work product is provided by Petersson et al. (2004) as follows:

> The overlap among the faults that the reviewers found is used as a basis for the estimation. The smaller overlap among the reviewers, the more faults are assumed to remain, and the larger overlap, the fewer faults are assumed to remain. The extreme cases are the following: either, all reviewers have found exactly the same faults, which means that there are probably very few faults left, or none of the reviewers has found a fault that another reviewer has found, which indicates that there are probably many faults left.

The assumptions regarding the closed capture-recapture models (given before in the text) are translated to software inspection context by Miller (1999) as follows:

1. Closed population => The artifact is not revised once it is delivered for inspection; and a particular inspector finds exactly same defects if he is given the same artifact twice, i.e. inspector performance is constant.
2. Animals do not lose their marks => Inspectors do not publish the defects they found to other inspectors.
3. The marks are correctly recorded => Inspectors correctly record and document the defects they identify.
4. Equal catchability => All inspectors are continuously provided with identical information (such as inspected artifact, inspection aids, standards against which the artifact is evaluated etc.).

Furthermore, the three sources of variations that underpin the capture-recapture models can be considered in the context of software inspection as given below (Freimut, 1997):

- **Time Response:** As mentioned previously, the time response is utilized to model the varying capturing probability in different trapping occasions. Since the individual inspectors participating in the inspection is mapped to these trapping occasions, the time response explains the variations among different inspectors. In software inspection setting, the defect finding ability of the inspectors is viewed as the difference. Certainly, this detection ability is related to the factors such as expertise and education level of the inspector along with his familiarity with the work product.

- **Heterogeneity:** By applying the similar reasoning used for the catchability of the animals, this source of variation is considered as the factor that indicates the varying detectability of the defects in the inspected artifact. The defects in a software work product are not equally responsive to the reading activity of the inspection, i.e. more effort is needed to identify certain defects. Even a large amount of effort is spent, finding some of the defects may be impossible during the inspection (which are later noticed during testing or field use). This is in line with the results of Vander Wiel and Votta (1993), who reported that developers generally classify defects as 'easy to detect' and 'hard to detect'.
- **Behavioral Response:** Some ideas have been put forward for this source of variation in software inspection. According to these approaches, the capture-recapture models incorporating behavioral response can be used to adjust the detection probability of the defects that are pointed out by many inspectors (Briand et al., 2000), or taking into account the behavioral response might be useful in a situation where one inspector passes the inspected artifact to another with the markings that indicate the defects he found (Freimut, 1997). However, despite these attempts, none of the behavior based capture-recapture models have been employed for software inspection context due to the assumption of independence among inspectors and unreasonableness of ordering the inspectors as trapping occasions ordered while using the related models for animal populations.

Therefore, based on the non-applicability of behavioral response, the remaining four closed population capture-recapture models have been used throughout the literature for estimating the defect content of an inspected software artifact. These models are listed in Table 6 together with their underlying assumptions regarding software inspection and corresponding estimators. Thelin et al. (2002) defined an estimator as; "A formula used to predict the number of faults remaining in an artifact".

In order to model the detection probability of a particular defect, one can adapt a notation $\mathbf{p_{ij}}$, which denotes the probability that defect i is found by inspector j. In Table 6, also, the equality expressing $\mathbf{p_{ij}}$ value according to the assumptions of the corresponding model is provided (Freimut, 1997) (where, p_i: the probability that defect i being detected by any inspector, and p_j: the probability that inspector j detects any defect).

Table 6 Capture-Recapture Models and Estimators Used for Software Inspection

Model	Assumptions	Estimators	p_{ij}
M0	All defects have equal detection probability. All inspectors have equal detection ability.	• Maximum Likelihood Estimator (Otis et al., 1978)	$p_{ij} = p$
Mt	All defects have equal detection probability. Inspectors may have different detection abilities.	• Maximum Likelihood Estimator (Otis et al., 1978) • Chao's Time Estimator (Chao, 1989)	$p_{ij} = p_i$
Mh	Defects may have different detection probabilities. All inspectors have equal detection ability.	• Jack-knife Estimator (Burnham and Overton, 1978) • Chao's Heterogeneity Estimator (Chao, 1987)	$p_{ij} = p_j$
Mth	Defects may have different detection probabilities. Inspectors may have different detection abilities.	• Chao's Heterogeneity-Time Estimator (Chao et al., 1992)	$p_{ij} = p_i x p_j$

As can be seen from Table 6, for each model, there is at least one estimator, which can be used to calculate a point estimate and corresponding confidence interval. The usage of any capture-recapture estimator for software inspection defect content estimation requires the collection of raw data in the following form.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & x_{ij} & \dots \\ x_{D1} & x_{D2} & \dots & x_{Dk} \end{bmatrix}$$

where $x_{ij} = 1$ if inspector i detected defect j.

Succeeding paragraphs provide the formulation relevant to capture-recapture estimators. The following notation is consistently used throughout these sections (Freimut, 1997).

k : number of inspectors

n_j : number of defects detected by inspector j

$n.$: the sum of n_j values for all inspectors, i.e. $\sum_{j=1}^{k} n_j$

N : total number of defects in the inspected document

$\hat{N}$: estimated total number of defects in the inspected document

D : number of distinct defects found during inspection

f_j : number of defects found by exactly j inspectors. (Hence, f_0 denotes the number of defects identified by none of the inspectors, actually this is equivalent to the number of remaining defects which is subject to the estimation procedure)

Z_j : the number of defects found only by inspector j

1. **Maximum Likelihood Estimator for Model M0 (MLE-M0):** The derivation of this estimator includes the employment of a log-likelihood function based on multi nominal distribution. The estimator ($\hat{N}$) is the N value that maximizes the following equation (Otis et al., 1978):

$$\left[\ln\left(\frac{N!}{(N-D)!} \right) + (n.)\ln(n.) + (kN - n.)\ln(kN - n.) - kN\ln(kN) \right] \qquad \text{(Eq. 4)}$$

Where, $N \in \aleph$ and $\aleph$ is the set defined as $\aleph := \{D, D+1, D+2, \ldots\}$. Hence, a search over $\aleph$, provides the estimated value regarding MLE-M0.

2. **Maximum Likelihood Estimator for Model Mt (MLE-Mt):** This estimator is the extension of MLE-M0 by incorporating time response into the related formulations. The estimator ($\hat{N}$) is the N value that maximizes the following equation (Otis et al., 1978):

$$\left[\ln\left(\frac{N!}{(N-D)!}\right)+\sum_{j=1}^{k} n_j \ln(n_j)+\sum_{j=1}^{k}(N-n_j)\ln(N-n_j)-kN\ln(kN)\right] \quad \text{(Eq. 5)}$$

Where, $N \in \aleph$ and $\aleph$ is the set defined as $\aleph := \{D, D+1, D+2, \ldots\}$. Hence, a search over $\aleph$, provides the estimated value regarding MLE-Mt.

3. Chao's Time Estimator (ChaoMt): This estimator is proposed by Chao by claiming to overcome the situations where MLE estimator (for Model Mt) overestimates in the case of sparse data. The estimator is deduced by using the expected values for f_1 and f_2 with aim of calculating f_0. The underlying formula is presented below (Chao, 1989):

$$\hat{N}_{tCh} = D + \frac{f_1^2 - \sum_{j=1}^{k} Z_j^2}{2(f_2+1)} \quad \text{(Eq. 6)}$$

4. Chao's Heterogeneity Estimator (ChaoMh): This estimator is proposed by Chao by claiming to overcome the situations where Jack-knife estimator's (for Model Mh) low performance in the case the animals are caught mostly once or twice. The corresponding formula is the following (Chao, 1987):

$$\hat{N}_{hCh} = D + \frac{f_1^2}{2f_2} \quad \text{(Eq. 7)}$$

5. **Jack-knife Estimator (JKMh):** This estimator is actually composed of sub-estimators which employ the capture frequency information (i.e. f_i). The sub-estimators are referred to as first order jack-knife estimator, second order jack-knife estimator, and so forth. A procedure for selecting the appropriate sub-estimator is put forward by Burnham and Overton (1978). This procedure starting with the first order jack-knife estimator computes the estimator and compares it with the next order estimator by the means of hypothesis testing, i.e. the zero difference between the values outputted by two estimators is the null hypothesis. If the null hypothesis is not rejected the lower order estimator is employed to estimate the defect content. Otherwise, the hypothesis testing is repeated with second and third order estimators. The procedure continues until the procedure stops due to failure of rejecting the null

hypothesis. The first five order sub-estimators of Jack-knife are given below (Burnham and Overton, 1978).

$$\hat{N}_{Jk1} = D + \left(\frac{k-1}{k}\right) f_1 \quad \text{(Eq. 8)}$$

$$\hat{N}_{Jk2} = D + \left(\frac{2k-3}{k}\right) f_1 - \left(\frac{(k-2)^2}{k(k-1)}\right) f_2 \quad \text{(Eq. 9)}$$

$$\hat{N}_{Jk3} = D + \left(\frac{3k-6}{k}\right) f_1 - \left(\frac{3k^2-15k+19}{k(k-1)}\right) f_2 + \left(\frac{(k-3)^3}{k(k-1)(k-2)}\right) f_3 \quad \text{(Eq. 10)}$$

$$\hat{N}_{Jk4} = D + \left(\frac{4k-10}{k}\right) f_1 - \left(\frac{6k^2-36k+55}{k(k-1)}\right) f_2 + \left(\frac{4k^3-42k^2+148t-175}{k(k-1)(k-2)}\right) f_3 - \left(\frac{(k-4)^4}{k(k-1)(k-2)(k-3)}\right) f_4 \quad \text{(Eq. 11)}$$

$$\hat{N}_{Jk5} = D + \left(\frac{5k-15}{k}\right) f_1 - \left(\frac{10k^2-70+125}{k(k-1)}\right) f_2 + \left(\frac{10k^3-120k^2+485k-660}{k(k-1)(k-2)}\right) f_3 - \left(\frac{(k-4)^5-(k-5)^5}{k(k-1)(k-2)(k-3)}\right) f_4 + \left(\frac{(k-5)^5}{k(k-1)(k-2)(k-3)(k-4)}\right) f_5 \quad \text{(Eq. 12)}$$

Actually, it is possible to derive higher order estimators, however it is seen that these estimators do not improve the estimation. Additionally, Miller (1999) claims that the selection algorithm of Burnham and Overton (1978) is not suitable in software inspection setting due to different losses encountered in software engineering and biological domains when the estimation deviates from the true value. Because of this underlying reason, he proposes to treat the sub-estimators of Jack-knife estimator as separate estimators.

6. Chao's Heterogeneity-Time Estimator (ChaoMth): This estimator is the only one which enables to estimate when both time response and heterogeneity exist. By using the concept of sample coverage, Chao et al. (1992) put forward three versions of the estimator, but without mentioning the differences among them. So it is wise to treat them as separate estimators. The formulas underlying these three estimators are given below:

$$\hat{C}_1 = 1 - \frac{f_1}{\sum_{j=1}^{k} jf_j} \quad \text{(Eq. 13)}$$

$$\hat{C}_2 = 1 - \frac{f_1 - 2\frac{f_2}{k-1}}{\sum_{j=1}^{k} jf_j} \quad \text{(Eq. 14)}$$

$$\hat{C}_3 = 1 - \frac{f_1 - 2\frac{f_2}{k-1} + 6\frac{f_3}{(k-1)(k-2)}}{\sum_{j=1}^{k} jf_j} \quad \text{(Eq. 15)}$$

$$\hat{\gamma}_i^2 = \max\left[\frac{D/\hat{C}_i \sum_{j=1}^{k} j(j-1)f_j}{2\sum\sum_{i<j} n_i n_j} - 1, 0\right], \quad \hat{N}_i = \frac{D}{\hat{C}_i} + \frac{f_1}{\hat{C}_i}\hat{\gamma}_i^2 \quad , \text{i}= 1, 2, 3 \quad \text{(Eq. 16)}$$

2.4.3 Subjective Defect Content Estimation Techniques

The defect content estimation techniques using subjective information, i.e. personal opinion, are less common in the literature when compared to the techniques aiming the generation the defect estimate with objective means. Subjective techniques rely on the knowledge and experience of the inspection participants, and their feelings with the current inspection, while producing the defect estimate. El Amam et al. (2000) propose asking the effectiveness value (i.e. percentage of defects found) to the individual inspectors and using this information to find out an estimate regarding the total number of defects in the artifact. They claim that, if the actual number of defects found by a particular inspector is divided by the estimated effectiveness value obtained from the same inspector, the total number of defects in the artifact can be obtained. Then, they conduct an experiment to test this proposal by using professional developers, and report that subjective estimation can be employed as an alternative to objective techniques since a median relative error of zero is observed.

Other approaches related to utilizing the individuals' perceptions for defect content estimation are provided by Biffl (2000). Biffl (2000) introduces the following three models for coming up with team estimates after obtaining individual opinions for the total number of defects present in the artifact: Largest Interval (LI), Weighted Average of Individual Estimates (WAE), and Weighted Average of Individual Offsets (WAO). Largest Interval model calculates the team estimate by averaging the maximum and minimum values that inspectors provide. WAE model reveals the weighted average of the individual weights, i.e. the sum of multiplications regarding individual weights and estimates is divided by total weight. Whilst, WAO model calculates the weighted average of the differences between the individual estimate and the actual defects found by the same inspector during the inspection. Further, Biffl (2000) puts forward three different approaches to represent the weights required for WAO and WAE models above, i.e. the contributions of the individual estimates to the team estimate. Later, these models are extended again by Biffl (2003) to handle the cases when more than one inspection is carried out with the aim of taking into account the defect data from the previous inspection.

2.5. Evaluations of Different Software Inspection DCETs

2.5.1 Evaluation Measures for Software Inspection DCETs

As it has been for any kind of evaluation, evaluation of DCETs also requires the designation of objective measures which enable the comparison of DCETs with respect to different aspects such as accuracy and variability. The related measures put forward in the literature, for evaluating DCETs, are outlined below.

- **Relative Error (RE) (Briand et al., 2000, Thelin et al., 2002, El Amam and Laitenberger, 2001):** This measure provides insight about the accuracy of a DCET by calculating the normalized bias between the estimated and actual number of defects contained in the artifact. As the median or mean of the relative error corresponding to a number of estimations performed with a particular DCET is found, the overestimation/underestimation tendency of the DCET can be observed. Whilst, the variance or inter quartile range of RE inform about the variability of the DCET regarding its accuracy. The underlying formula of this measure is given below. The optimum value for RE measure is zero. However, in general, an overall RE value between ± 20% range of zero is acceptable (Briand et al., 2000).

$$RE = \frac{\hat{N} - N}{N} \quad \text{(Eq. 17)}$$

- **Decision Accuracy (DA) (El Amam et al., 2000, El Amam and Laitenberger, 2001):** This measure calculates the proportion of correct reinspection decisions that a DCET proposed. This requires the designation of reinspection decision criteria and, the storing of the decision concluded according to estimated and actual number of defects contained in the artifact. DA measure enables to observe the capability of a DCET for guiding the practitioners in making correct decisions. The formula of DA is provided below. As it should be evident, the optimum value for DA measure is 1.

$$DA = \frac{m_0 + m_1}{M} \quad \text{(Eq. 18)}$$

where,

M: Total number of reinspection decision instances

m_0: Number of instances that DCET correctly proposed 'do not reinspect' alternative

m_1: Number of instances that DCET correctly proposed 'reinspect' alternative

- **Relative Decision Accuracy (RDA) (El Amam et al., 2000, El Amam and Laitenberger, 2001):** This measure is similar to DA, however it excludes the cases where the DCET correctly proposes the default decision of passing the artifact to the next phase, i.e. without performing a reinspection. This measure provides information about how much a DCET is successful beyond making the default decision. The formula of RDA is depicted below. The optimum value for RDA measure is zero.

$$RDA = DA - \frac{m_0}{m_0 + m_0^{'}} \quad \text{(Eq. 19)}$$

where,

$m_0^{'}$: Number of instances that DCET proposed 'do not reinspect' alternative, although the correct decision was the opposite.

- **Root Mean Square Error (RMSE) (Thelin et al., 2002):** This measure is used to evaluate the bias and the variability of a DCET at the same time. If two DCETs are compared solely with respect to their RMSE values, the one with lower RMSE is deemed better. The formula of RMSE measure is provided below.

$$RMSE = \sqrt{Var\left[\hat{N}\right] + \left(E\left[\hat{N}\right] - N\right)^2} \qquad \text{(Eq. 20)}$$

- **Failure Rate (FR) (Briand et al., 1998b, Briand et al., 2000):** Some of the DCETs fail to estimate in certain situations. For example, Chao's Heterogeneity Estimator fail due to division by zero error, if none of the defects is captured by exactly two inspectors. Hence, FR measure accounts for the failure frequency of a DCET. If a DCET can not provide an estimation most of the time, it precludes the aims regarding the proper estimation of the number of defects. So, a failure rate of zero is the most preferable value for FR measure. The underlying formula of failure rate measure is given below.

$$FR = \frac{Number\ of\ cases\ that\ the\ DCET\ fails}{Total\ number\ of\ cases\ where\ the\ DCET\ used} \qquad \text{(Eq. 21)}$$

2.5.2 Findings of the Studies that Evaluate Software Inspection DCETs

A considerable amount of studies in the literature report the evaluations conducted for revealing the differences among Software Inspection Defect Content Estimation Techniques (DCET). The aim of these studies is to find out the most appropriate technique regarding the proper estimation of the number of defects contained in the inspected artifact. Further, these studies account for various circumstances that can be encountered in software development environments.

A recent survey study provides a good summary regarding the usage of Capture-Recapture estimators in software inspection defect estimation along with other related methods and issues (Petersson et al., 2004). By taking into account 15 studies that report evaluation results performed for different estimators, Petersson et al. (2004) finds out that; (i) Most estimators tend to underestimate, (ii) Jack-knife estimator is the best estimator for using in software inspections, and (iii) The recommended minimum number of inspectors, which makes the usage of Jack-knife estimator more appropriate, is four, (iv) The studies that evaluate Capture-Recapture estimators for the cases where two inspectors participates, do not show consensus about the best estimator, and (v) The related studies show that usage of different reading techniques does not affect the estimators' performance, so Capture-Recapture estimators can be employed with any reading technique. Furthermore, Petersson et al. (2004) also report that DPM is the most appropriate curve-fitting method, although it is generally found to be inferior than Jack-knife estimator. The findings of this study resulted from a comprehensive exploration of the literature. Consequently, the details of the studies included by Petersson et al. (2004) while constituting the above findings is not repeated here, unless they dwell upon the specific subjects related to this study. Among the studies that evaluate various Capture-Recapture estimators, Miller's (1999) study deserves special attention. Miller (1999) claims that the sub-estimator selection procedure of Burnham and Overton (1978) for Jack-knife estimator is inappropriate in the context of software engineering. So, he evaluates various estimators by considering the Jack-knife sub-estimators as separate estimators. His evaluations favor the usage of Jack-knife estimator order 1. However, in this study Miller (1999) does not include the estimations obtained by using Burnham and Overton's (1978) selection procedure (generally referred as full Jack-knife estimator). Thelin et al. (2002) later replicates Miller's (1999) proposal by including also the full Jack-knife estimator and DPM. In this replicated study, Jack-knife estimator order 2 is found as the most appropriate estimator. Since also full Jack-knife estimator is considered in the study, this finding supports the original claim of Miller (1999).

Several studies in the literature focus on the subjective DCETs by inspecting their performance in estimating the defect content of an inspected artifact. The idea for using of subjective estimates in order to predict the number of defects present in the software work product is originated by El Amam et al (2000). Their original study does not include any comparison of the subjective estimations with estimations obtained by any objective DCET. Fortunately, Biffl (2000) provides guidance about the performance of subjective DCETs proposed by himself. Biffl (2000) reports that when subjective DCETs are considered alone, Largest Interval (LI) model performs better than others. Additionally, he shows that subjective models tend to underestimate. Further, the results of the study favor subjective DCETs over objective DCETs, although Jack-knife estimator is even considered (which exhibits the best performance among objective techniques also in this study). By employing the same requirements inspection experiment underpinning this study, Biffl (2003) provides insight also for the usage of DCETs for estimating the number of major defects. According to results of this study; (i) all DCETs tend to underestimate major defects, (ii) objective DCETs depict similar accuracy for major defects defects, but Chao's estimators for models Mh and Mth consistenly show good results, (iii) The performance of objective and subjective DCETs are comparable and (iv) The accuracy of DCETs are not affected significantly by the usage of different reading techniques.

2.6. Reinspections

After a software inspection is completed, i.e. the identified defects are resolved and verified; the project manager should decide either to pass the corresponding work product to the next phase in the software life cycle or to inspect it again for finding the defects missed in the first inspection. If the latter is the case, it is referred as a reinspection. To be more specific, a reinspection is inspecting the work product anew with the aim of reducing the number of defects further to a more suitable level. In reinspection, the parameters that can be changed with respect to the first inspection are as follows; (i) inspection process, (ii) number and content of the inspectors, and (iii) reading technique (Biffl et al., 2001). In practice, a considerable amount of software development organizations pass the work product to the next phase even giving no consideration to reinspection, i.e. the default decision is to baseline the work product after it is verified with the inspection and to take it as an input to the succeeding activity (Radice, 2002). However, performing reinspections is becoming more common in the software industry. Software Engineering Institute's (SEI) Capability Maturity Model Integration (CMMI) (CMMI Product Team, 2001), recommends the devising of reinspections by putting forward related subpractices in verification process area. This evidence shows the current commonality of reinspections in the industry and also, the prospects for the increasing implementation throughout the industry in the near future.

By performing a reinspection the software development project can obtain more of the potential benefits offered by software inspections (see Section 2.2). However, repeating of the inspection also requires additional resources and time to invest in the inspection activity. Hence, as for all decisions, reinspection decision also comprises a trade-off situation: to devote the valuable effort of software developers to a second inspection cycle or to use the effort for other planned activities. In the literature, many methods are proposed to make the reinspection decision. Some of these methods put forward ad-hoc criteria to conclude the decision. For example, the repeating of the inspection can be based on the belief of inspection leader or inspection team regarding the inspected work product defect level (Gilb and Graham, 1993, Strauss and Ebenau, 1993). However, such an approach results in arbitrariness and subjectivity while coming up with the decision for reinspection, which is not a desired situation for making good decisions with respect to corresponding objectives (Briand et al., 2000, Radice, 2002). Employing the historical information stands as a second approach for finding out whether the reinspection is needed. According to this approach, the results of the current inspection are compared to historical norms or related benchmarks. For example, the defect density (number of defects found in the first inspection divided by the size of the inspected artifact) of the inspected work product is evaluated against the upper and lower limits deduced from historical inspection data (Radice, 2002). If the defect density is above the upper limit, the document is deemed as inferior in terms of quality. Whilst, the application of the inspection is concluded to be poor, if the defect density is below lower limit. Consequently, in both cases a reinspection is justified. Although, this quantitative approach is employed by many software organizations (especially which implements statistical process control practices), the following shortcomings are supplied in the literature for explaining its inappropriateness.

- The historical or benchmark data may not be available, or obtaining it is too costly (Briand et al., 2000, Biffl, 2000).
- The historical or benchmark data may not be appropriate to use for the current project, since the circumstances of the projects, from which the data is obtained, differs with respect to current one (Biffl, 2000).
- Inspectors may tend to find a passing number of defects (Briand et al., 2000).
- The low quality work products may pass to next phase without reinspection and high quality work products may be needlessly reinspected (Biffl, 2000, El Amam and Laitenberger, 2001, Briand et al., 2000, El Amam et al., 2000).

Consequently, in the literature, using the estimated number of defects in the inspected artifact is proposed as a third and appropriate approach for concluding reinspection decision objectively (Miller, 1999, Briand et al., 2000, Biffl and Halling, 2001, Petersson et al., 2004). This approach requires the employment of Defect Content Estimation Techniques (described in Section 2.4) along with the related data from the initial inspection. We can call this approach as *deciding the reinspection with objective reinspection decision methods*, since it is purified from peoples' opinion and it does not include the bias of historical or benchmark data. In the literature, different objective methods are proposed regarding reinspection decision, which are listed and described in the paragraphs below. From now on, 'reinspection decision method' term will be used throughout the text for referring briefly to *objective reinspection decision methods*.

I. Deciding based on the effectiveness of the current inspection (El Amam and Laitenberger, 2001, Thelin and Runeson, 2000).

Criterion: Reinspect, if the effectiveness of the first inspection is smaller than a predetermined inspection effectiveness threshold.

The effectiveness of an inspection is the percentage of total defects found during the inspection. This method evaluates the quality of the inspection process with respect to defect finding performance and suggests a reinspection if it is not satisfactory. According to the method, the artifact should be reinspected if the following inequality holds.

$$\frac{D}{\hat{N}} < e \qquad \text{(Eq. 22)}$$

where,

D: The number of distinct defects found during inspection

$\hat{N}$: Estimated total number of defects in the inspected document

e: Threshold value for inspection effectiveness

As evident from the above formulation, this method requires the estimate for the total number of defects in the document, which can be obtained by using appropriate defect content estimation technique.

II. Deciding based on the defect density after the current inspection (El Amam et al., 2000).

Criterion: Reinspect, if the defect density of the artifact corrected as a result of the initial inspection is greater than or equal to a given defect density threshold.

The defect density is the number of defects contained in the artifact per unit size. Hence, this method concludes the reinspection decision by considering a minimum quality level that should be attained by the artifact subject to inspection. According to the method, the artifact should be reinspected if the following inequality holds.

$$\frac{\hat{N}-D}{S} \geq d \qquad \text{(Eq. 23)}$$

where,

S: The size of the artifact (in source lines of code or number of pages)

d: Defect density threshold

If the both sides of the above inequality is multiplied with artifact size, the method becomes deciding with respect to number of remaining defects. This method, again, needs the use of DCETs.

III.Deciding based on the defect density after the reinspection (Biffl and Halling, 2001).

Criterion: Reinspect, if the defect density is above the corresponding threshold and the reinspection will bring it to the acceptable level.

If the defect density is high after the first inspection with respect to the predetermined threshold, the potential of the reinspection for decreasing the defect containment to the suitable value is evaluated via this model. In order to reinspect the artifact this method requires the following inequalities to hold at the same time.

$$\frac{\hat{N}-D}{S} \geq d \qquad \text{(Eq. 24)}$$

$$\frac{\hat{N}-D-\hat{D}_2}{S} < d \qquad \text{(Eq. 25)}$$

where,

$\hat{D}_2$: Estimate for the number of defects that would be identified during reinspection.

As can be seen from the above formulation, this method needs utilization of the suitable defect detection capability estimation and defect content estimation techniques together. By multiplying both sides of both inequalities with artifact size, the method can be converted to one considering the number of defects as the issue on which reinspection criterion is based.

IV. Deciding based on the effectiveness after the reinspection (Biffl and Gutjahr, 2002).

Criterion: Reinspect, if the effectiveness of the current inspection is below the effectiveness threshold and the reinspection will enable the exceeding of this threshold.

This method suggests reinspecting if it is worth to attain the required level of inspection quality, i.e. effectiveness, if it is not accomplished with the initial inspection. According to the method, the reinspection decision should be made if the following two inequalities are satisfied.

$$\frac{D}{\hat{N}} < e \qquad \text{(Eq. 26)}$$

$$\frac{D + \hat{D}_2}{\hat{N}} \geq e \qquad \text{(Eq. 27)}$$

Using this method means that the reinspection is not beneficial for the project as long as it would not decrease the percentage of the remaining defects below the required level.

V. Deciding based on the net benefit obtained from the reinspection (Biffl and Halling, 2001).

Criterion: Reinspect, if the net benefit (i.e. benefits minus costs) is above the predetermined benefit threshold.

This method proposes to quantify the benefits acquired by conducting a second inspection and compare them with the corresponding costs. The benefit of a reinspection is the saved rework effort regarding later phases of the software development life cycle, since if a defect remains undetected during inspection the cost to detect and correct it later is much higher than the one during inspection phase. If this method is employed, a reinspection is conducted as long as the following inequality is satisfied. Note that, the left side of the inequality corresponds to the net benefit value. Further, as in most software engineering studies the benefit and cost values are expressed in terms of staff effort in man-hours.

$$\frac{\hat{D}_2}{D}\left(M \cdot s_{mj} + (D-M) \cdot s_{mn}\right) - C_R > b \qquad \text{(Eq. 28)}$$

where,

M: The number of major defects found during inspection

s_{mj} : Average effort saved per major defect found during inspection

s_{mn} : Average effort saved per minor defect found during inspection

C_R: Effort spent to conduct the reinspection

b: Threshold for net benefit

2.7. Defect Detection Capability Estimation Techniques

Defect Detection Capability Estimation Techniques (DDCET) are constituted to predict number of defects to be found during reinspection. Hence, by using these techniques, an inspection team can conclude the reinspection decision based on the estimated gain that will be obtained if reinspection is carried out. DDCETs are relatively novel according to various DCETs. Several papers deal with this subject. The three DDCETs proposed in the literature are (i) Optimistic Linear Model (OLM), (ii) Improved Linear Model (ILM), and (iii) Reliability Growth Model (RGM). Actually, the first two models are heuristics structured with the assumption of linear relationships between first and second inspection (i.e. reinspection). Further, all three models assume that in the second inspection cycle, the same inspection process and the same inspectors are employed as in the first inspection. When these three techniques are compared, the usage of ILM is recommended (Biffl and Gutjahr, 2002) or ILM is deemed comparable with RGM (Biffl and Halling, 2001). The brief description of each DDCET is provided below.

1. **Optimistic Linear Model (OLM):** This model assumes that the efficiency in the first inspection is conserved during reinspection. Inspection efficiency is defined as the number of defects found per unit effort spent. Hence, the predicted number of defects to be found by reinspection team can be expressed as follows (Biffl and Halling, 2001):

$$\hat{D}_2 = \left(\frac{D_1}{E_1}\right) \times E_2 \qquad \text{(Eq. 29)}$$

where,

$\hat{D}_2$: Estimate for the number of defects that would be identified during reinspection.

D_1: Number of defects detected during the first inspection.

E_1, E_2: The effort for examining the work product in the first and second inspection, respectively.

2. **Improved Linear Model (ILM):** This model is similar to OLM, but it discounts the number of defects to be found in reinspection by the estimated percentage of remaining defects after the first inspection. The aim of this is to account for the decreased efficiency during second inspection cycle due to decreased number of defects available for detection. This approach is concordant with the results of the study conducted by Biffl et al. (2001), which reports that reinspections provide lower net benefit when compared to initial inspection. The corresponding formula to this model is given below, in which –again- the efficiency values are utilized (Biffl and Halling, 2001):

$$\hat{D}_2 = \left(\frac{D_1}{E_1}\right) \times E_2 \times \left(1 - \frac{D_1}{\hat{N}_0}\right) \qquad \text{(Eq. 30)}$$

where, $\hat{N}_0$ is the estimate for the total number of defects in the work product at the start of inspections and generated by using appropriate DCET with the data from the first inspection.

3. **Reliability Growth Model (RGM):** This model adapts the reliability approaches for inspection process. In software engineering, these approaches are frequently used to predict the number of defects to be found in testing. Reliability Growth models are

based on the assumption of decreasing rate of defect detection. Biffl and Gutjahr (2002) propose to use Jelinski-Moranda reliability model in software inspections. According to this method, during the first inspection cycle the time when a particular defect is detected needs to be stored. Then, this data is utilized together with the reliability model structured for a parametrized stochastic process, i.e. second inspection process, to estimate the mean time between defect detection events during the second inspection cycle. Finally, the mean time is employed to predict the number of defects that will be found in the reinspection given the duration of the reinspection cycle. This information can also be used to estimate the total number of defects in the artifact, if unlimited duration for reinspection is considered. Biffl and Halling (2001) demonstrate the usage of RGM for software inspections with the plot given in Figure 1. The drawback of the approach explained above is the need for collecting time-stamped data in the initial inspection, which is not available in most software inspection implementations.

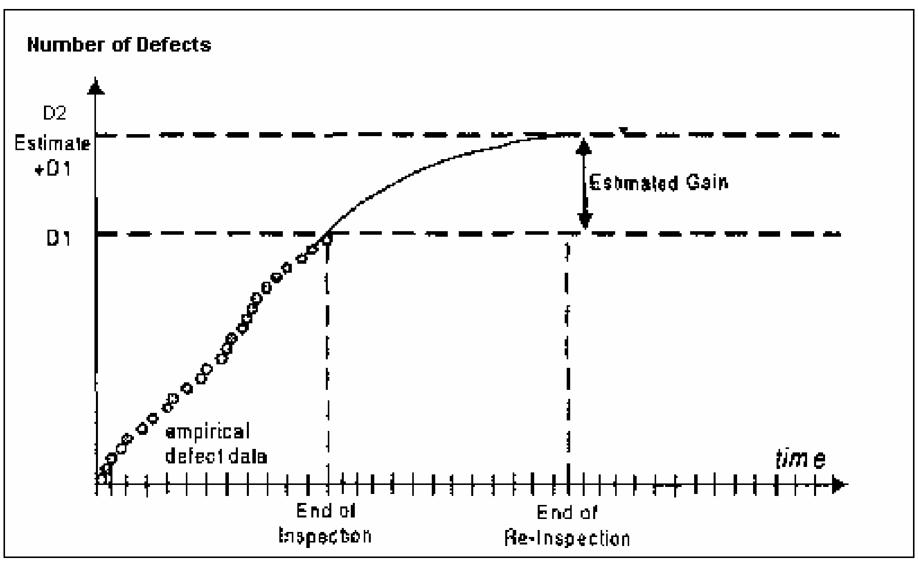

Figure 1 Reliability Growth Model for Software Inspection (Biffl and Halling, 2001)

CHAPTER 3

DESCRIPTION AND CONDUCT OF THE STUDY

3.1. Description of the Method and the Simulation Model Applied in the Study

As depicted in Chapter 2, the literature investigates various issues related to software inspection in order to find answers to following questions: How should software inspection process be designed for making it efficient and effective?, How many inspectors should be allocated for obtaining most benefit from the software inspection with no significant cost increase?, How much benefit is acquired from performing software inspections with respect to return on investment, rework cost savings, percentage of defects found, and decrease in the costs of other defect removal activities such as testing, What is the optimum preparation and meeting rates for software inspection?, How much cost is incurred during software inspection?, How should the software artifact be examined during inspection, i.e. which reading technique is more appropriate for maximizing the effectiveness of the software inspection?, How much does it cost to correct (i.e. rework cost) a defect during software inspection phase and subsequent phases?, How can the number of defects contained in the software work product be estimated appropriately by employing the results of the software inspection? How does reinspection differs from the initial inspection in terms of benefits and other factors?, How should the number of defects to be identified during a potential reinspection be predicted?. In addition to these, the literature also provides different methods for concluding the reinspection decision by objectively guiding the decision maker based on the data from the initial inspection. The reinspection decision is a good example of various technical and managerial decisions that are made during a typical software development project. Decision making in the context of software engineering is a complex, important and difficult task, due to complexity, human intensity, uncertainty, and existence of multiple objectives inherent in software development process (Rus et al., 2003). The three main objectives, which are generally pursued in software development projects, are related to cost, schedule and quality. More clearly, software project managers, in general, aim to assure a certain quality level, to reduce development costs and to keep the project within the schedule. Consequently, there is a constant striving in software engineering literature and software development industry to identify the ways for developing cheaper, faster, and better software. For this reason, any software engineering related decision should be evaluated with respect to its effects on the achievement of these objectives (Rus et al, 2003). Since, the final results regarding cost, quality and schedule objectives are revealed at the end of software development life cycle, the mentioned evaluation certainly requires the analysis of the outputs observed after completion of the software project for three objectives. When this necessity is considered, it is concluded that the effects of software reinspection decisions should also be investigated for their effects on software project objectives. However, although a number of reinspection decision methods proposed in the literature, no study exists (to author's knowledge) that evaluates and compares the different reinspection decision methods by revealing their impact on cost, quality and schedule of the software development project. Hence, this study aims to address this niche by following the method described extensively in subsequent paragraphs and sections.

In this study, the reinspection decision that needs to be made in a SW-CMM (Paulk, et al., 1993) Maturity Level 3 organization's project after the completion of a code inspection is considered. The requirements put forward by SW-CMM for peer reviews corresponds to software inspection, when they are examined according to characteristics of different peer review types, which are outlined in Section 2.1. In this study the software code inspections are considered, since, as stated in Section 2.2, software inspections are more common for removing the defects in software code. Hence, by using the underlying implementation data available through the literature, providing insight to reinspection decision for code inspections is more valuable. Further, SW-CMM Maturity Level 3 provides a good context for the study as explained in the following sentences. Maturity Level 3 (also called as 'Defined Level') enables the "establishment of an infrastructure that institutionalizes effective software engineering and management processes across all projects" (Paulk, et al., 1993). By this way, all projects (in a maturity level 3 organization) operate according to procedures which are tailored from the standard processes put forward in the organizational level. This brings consistency among the process implementations performed by various projects. Hence, when this fact is considered together with the peer review KPA's rules that each maturity level 3 organization shall comply with, designating the context of the study as SW-CMM enables the applicability of the potential results to many organizations. More clearly, the variability among the different software inspection implementations is minimized, since the considered software organizations conduct peer reviews according to peer review rules of SW-CMM. This is also true for other processes such as testing, requirements analysis, software design etc. From this perspective, considering higher and lower maturity levelled organizations in this study is not reasonable. In organization with lower maturity levels of SW-CMM the software inspection may be in place, although not required by SW-CMM. However, since the process standardization is not implemented and the SW-CMM rules are not ensured, the variability of the software inspection implementations are expected to be high for such organizations. This is fact is also valid for other processes. Clearly, this hinders the generalization of the study results. Whilst, for organizations which have maturity levels 4 or 5, the employed technologies and followed project execution practices may vary greatly due to continuous process improvement and quantitative project management concepts in place. Again, this represents a situation that may be detrimental while generalizing the study results. Consequently, to sum up, considering maturity level 3 of a widely employed process improvement model, namely SW-CMM, makes sense for the purposes of this study.

For evaluating the effects of the decision made at the point where the initial inspection is completed, a Monte-Carlo simulation model to represent the activities that the software code goes through before and after the decision is constituted. The cycle that starts with the initial inspection and ends with the field use (i.e. the related operational environment the software is used in) is taken into account as the scope of this model. This process is depicted in Figure 2. According to this process, a software code piece goes through the inspection. During inspection, the code follows the general inspection steps, namely planning, overview, preparation, meeting, rework (correction), and follow-up. The process assumes that no new insertion new defects are inserted during correction step. After the inspection is completed, the decision is made whether to reinspect the artifact (i.e. software code) or continue with testing. Based on this decision the code is passed to testing or it is verified with inspection once more. So, the code enters to testing phase after the defects are eliminated via one or two inspection cycles (the second being the reinspection). During testing, some of the remaining defects in the code are removed by executing the pre-designed test cases and identifying the failing functions. Then, the code is deployed to the field with the residual defects, i.e. the defects which could not be found by means of inspection or testing. Eventually, these defects are revealed by the users sooner or later and corrected by the software organization.

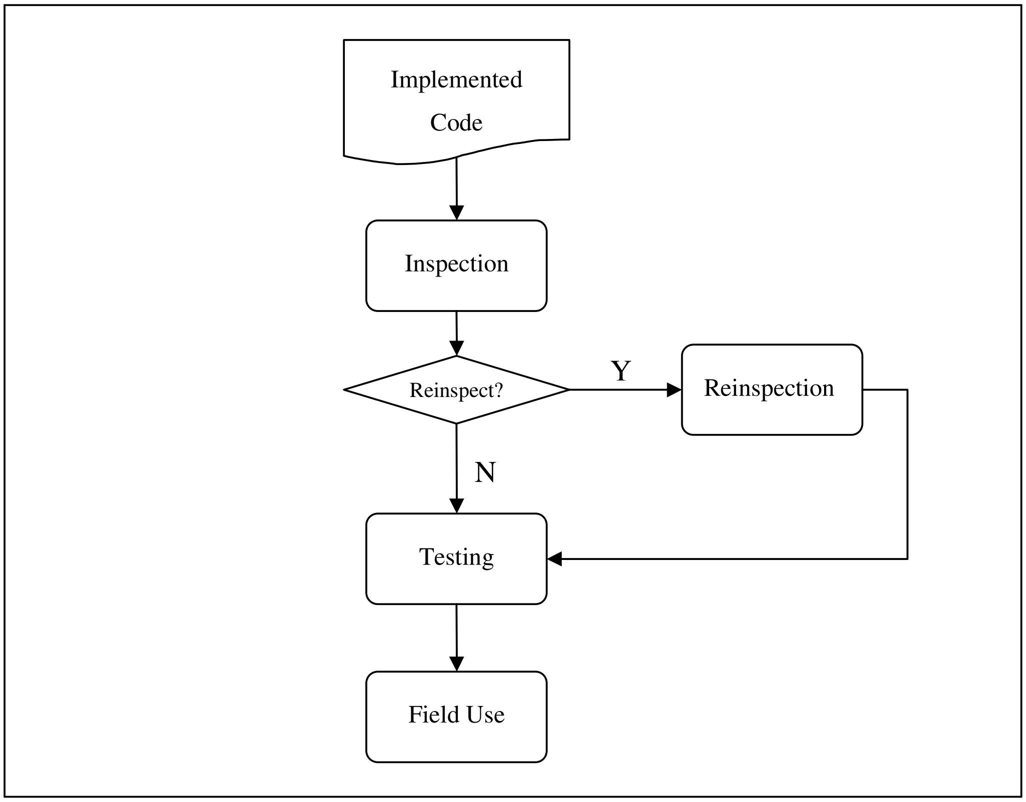

Figure 2 The Inspection-Reinspection Process Flow

As clear from the above explanations, since all of the included activities aim to capture the defects in the software code, the defects should be considered as the main entity. Hence, the simulation model represents the different points that can detect the defects as they flow through the software life cycle.

As described before, the aim of the study is to evaluate different methods for code reinspection decision. In line with this aim, the study tries to find out the resulting cost, defect level encountered by the users and the schedule delay with respect to planned end date of the project. Among these, the cost and defect level require the information regarding how many defects are caught in each of the relevant detection activity, i.e. inspection, reinspection, testing or field use. Whilst, the schedule delay is caused by performing the reinspection, which is an unanticipated activity in the project plans. Consequently, structuring the model based on the events the defects come across is reasonable. In order to achieve this, the following information needs to be known regarding the defects:

1. The activity where a particular defect is detected
2. Number and content of the defects present in the code at the beginning of the inspection
3. The cost to correct a particular defect when it is detected
4. The cost of conducting the reinspection

3.1.1 Modeling the Defect Detection

For evaluating the different reinspection decision methods Monte Carlo simulation is utilized in this study. This requires the modeling of the defect detection events for the activities in the model. Actually, if a defect is detected during field use that means it can not be identified by means of inspection or testing, thus propagated to the user. So, the consideration of the detection events regarding other activities (namely inspection, reinspection–if performed-, testing) is sufficient. For these activities, the related detection models are described in the succeeding paragraphs.

- **Modeling the detection event for inspection activity:** By recalling the background information provided in Section 2.4.2, a defect's detection is related to two factors which are defect's detectability and inspector's (defect detection) ability. Both of these factors can be expressed in probability terms. More clearly, the detectability of a defect refers the probability that it is found by any of the inspectors and the capability of the inspector refers to the probability that he identifies any defect. Thus, the multiplication of these two probabilities reveals the probability of detection for a particular defect by a specific inspector. Furthermore, since more than one inspector participate in inspection, a defect remains undetected if it is not found by any of the inspectors, i.e. the mentioned defect detection probability is realized for none of the inspectors. Additionally, here the detected defects refer to the defects which are really defects, i.e. it is assumed that no false positives exist and if an inspector claims a proper situation as a defect, this is identified during meeting phase, so it is not counted as a defect. Thus, the inspector defect detection event does not possess Type II error. As evident, the capabilities of different inspectors in an inspection session vary. Besides, as mentioned in the text before, the defects can be classified into two as ‘easy’ and ‘difficult’ according the their degree of detectability, i.e. easy and difficult defects have high and low probabilities of being captured, respectively. Consequently, the following formulation can be constructed to represent the event of detection during inspection activity in the context of Monte Carlo simulation.

$$p_{ij} = p_i\, q_j \qquad \text{(Eq. 31)}$$

$$p_i = \begin{cases} p_1, \text{if defect i is difficult to find} \\ p_2, \text{if defect i is easy to find} \end{cases} \qquad \text{(Eq. 32)}$$

$$x_{ij} = \begin{cases} 1, \text{if } \mathrm{r_{ij}} \le \mathrm{p_{ij}}, \text{i.e. if defect i is captured by inspector j} \\ 0, \text{otherwise} \end{cases} \qquad \text{(Eq. 33)}$$

where,

p_i: the probability that a defect i is detected by any inspector

p_1: the probability that a difficult defect is detected by any inspector

p_2: the probability that an easy defect is detected by any inspector

q_j: the probability that inspector j detects any defect

p_{ij}: the probability that a defect i is detected by inspector j

r_{ij}: A Uniform(0,1) random number

By employing the above notation the following measures can be derived regarding the results of the inspection activity.

$$X_i = \begin{cases} 1, \text{if } \sum_{j=1}^{k} x_{ij} \ge 1, \text{i.e if defect i is detected during inspection} \\ 0, \text{otherwise} \end{cases} \qquad \text{(Eq. 34)}$$

$$D = \sum_{i=1}^{N} X_i \qquad \text{(Eq. 35)}$$

where,

X_i: The indicator if defect i is detected during inspection or not k: Number of inspectors

D: Number of defects found during inspection

N: Total number of defects present in the inspected code

- **Modeling the detection event for reinspection activity:** The modeling of the detection event during reinspection is the same as the one described for inspection activity above. Hence, the corresponding formulation can be directly provided as below without repeating the details. Remember that, the reinspection is performed for the remaining defects if the related decision method suggests reinspection. Recall that, during a reinspection it is possible to change the inspection process, reading technique, and number and content of inspectors. In this study, only the number and content of inspectors is assumed to be subject to change. However, this does not cause any change in the inspector's probability of catching a defect, since even different inspectors are employed during reinspection, they are selected from the same inspector pool which provides inspectors with similar characteristics due to common training given in SW-CMM Level 3 context. Consequently, it is assumed that the inspectors in reinspection team possess the same capability with the inspectors participated in the initial inspection.

$$p'_{ij} = p_i \, q'_j \qquad \text{(Eq. 36)}$$

$$p_i = \begin{cases} p_1, \text{if defect i is difficult to find} \\ p_2, \text{if defect i is easy to find} \end{cases} \qquad \text{(Eq. 37)}$$

$$x'_{ij} = \begin{cases} 1, \text{if reinspection is performed, and } r'_{ij} \le p'_{ij}, \text{ and } X_i = 0, \\ \quad \text{i.e. if defect i is captured by reinspector j} \\ 0, \text{otherwise} \end{cases} \qquad \text{(Eq. 38)}$$

where,

p_i: the probability that a defect i is detected by any reinspector

p_1: the probability that a difficult defect is detected by any inspector

p_2: the probability that an easy defect is detected by any inspector

q'_j: the probability that reinspector j detects any defect (actually, it is equal with q_j)

p'_{ij}: the probability that a defect i is detected by reinspector j

r'_{ij}: A Uniform(0,1) random number

By employing the above notation the following measures can be derived regarding the results of the reinspection activity.

$$X'_i = \begin{cases} 1, \text{if } \sum_{j=1}^{k'} x'_{ij} \geq 1, \text{i.e if defect i is detected during reinspection} \\ 0, \text{otherwise} \end{cases} \quad \text{(Eq. 39)}$$

$$D' = \sum_{i=1}^{N} X'_i \quad \text{(Eq. 40)}$$

where,

X'_i: The indicator if defect i is detected during reinspection or not

k': Number of reinspectors

D': Number of defects found during reinspection

- **Modeling the detection event for testing activity:** In order to represent the test detection event, simply the fraction of defects that will be found during testing is considered. Hence, it is assumed that all defects reaching to testing activity are equally likely to be removed from the software. The formulation underpinning the defect detection event for testing is depicted as follows.

p_t = defect removal percentage of testing

$$y_i = \begin{cases} 1, \text{if } r_i \leq p_t \text{ and } X_i = 0 \text{ and } X'_i = 0, \text{i.e. if defect i is found in testing} \\ 0, \text{otherwise} \end{cases} \quad \text{(Eq. 41)}$$

where,

p_t: the probability that any defect propagating to testing is detected during testing

r_i: A Uniform(0,1) random number

Thus, the number of defects removed during testing can be stated as follows.

$$T = \sum_{i=1}^{N} y_i \quad \text{(Eq. 42)}$$

Finally, the number of defects that are encountered by the users during field use is expressed as; $F = N - D - D' - T$.

3.1.2 Modeling the Defect Injection

The number (N) and content of the defects at the beginning of the initial inspection has to be known for being able to execute the Monte Carlo simulation. The COQUALMO model is employed for determining the number of defects, since it provides the most comprehensive and generic information with respect to other available data (Boehm et al., 2000). As explained in Section 2.3, COQUALMO enables to estimate the number of defects contained in for requirements, design and code artifacts injected during software requirements analysis, software design and software coding activities, respectively. For unit-sized software code (i.e. 1 kilo source lines of code), the Defect Introduction submodel of COQUALMO states that 33 defects is present at the end of the coding activity on the average, if all defect introduction drivers are at their nominal levels. A nominal level of process maturity driver is assumed in this study, since, a SW-CMM Maturity Level 3 organization is taken into account. Further, all of the remaining drivers (such as programmer capability, required software reliability, platform volatility, team cohesion etc.) are assumed at their nominal levels, for being able to model the environment of a typical software project. The defect removal submodel of COQUALMO states that the nominal defect removal percentage for code defects eliminated through automated code analysis is 0.2 (See Table 4). So, 80% of the defects reaches to inspection activity, meaning that $33 \times 0.8 \cong 26$ defects are waiting to be removed via inspection. The underlying reason for using a nominal level of automated analysis capability is the variability of the tools used for these purposes by various software projects due to different technologies utilized. Hence, there is no information regarding the automated analysis profile of SW-CMM Maturity Level 3 software organizations. Using 26 as the average number of defects present in a 1000 SLOC sized software code is also supported by O'Neill (2003), who reports a defect insertion rate of 20 to 30 for Software CMM Level 3 companies.

With the content of the defects, the severity (as major or minor) and the degree of inspection detectability (as difficult or easy) of each defect are meant. So, the severity and detectability for each of the 26 defects should be designated as they are inputted to the simulation model. Let p_m and p_d be the probabilities regarding a certain defect being major and difficult, respectively. Then, the formulations for setting the severity and detectability of a particular defect can be stated as follows.

$$\text{severity of defect i} = \begin{cases} \text{major, if } r_m \le p_m \\ \text{minor, otherwise} \end{cases} \quad \text{(Eq. 43)}$$

$$\text{detectability degree of defect i} = \begin{cases} \text{difficult, if } r_d \le p_d \\ \text{easy, otherwise} \end{cases} \quad \text{(Eq. 44)}$$

Where, r_m and r_d are Uniform(0,1) random numbers used to simulate the defect severity and detectability, respectively.

3.1.3 Modeling the Cost of Defect Correction

In order to calculate the resulting cost at the end of the software life cycle described in this study, the cost of correcting a defect, i.e. rework cost (which also includes the cost of troubleshooting) needs to be computed, besides the labor cost of a reinspection. As explained in Section 2.3, in software engineering studies the costs are generally represented in terms of effort, i.e. man-hours. Further, again by recalling from Section 2.3, the rework costs are different for major and minor defects. In this study, the cost values supplied by National Software Quality Experiment (NSQE) (O'Neill, 2003) are employed due to its comprehensive nature with respect to other studies that report cost values from specific environments and experiments. According to NSQE the costs that are faced by a typical SW-CMM Maturity Level 3 organization are as given in Table 7 in terms of required effort hours. Here, for analysis purposes, NSQE study assumes an average cost of 1 effort hour to correct (repair) a defect (major or minor) during software inspection. This value is used by O'Neill (2003) to calculate Return on Investment value regarding software inspections for different software development environments. Consequently, in this study the same value is utilized, since no separate data can be found in the literature regarding the cost to correct a defect.

Table 7 NSQE Defect Correction Costs in terms of Man-hours

Activity	Cost to Correct Minor Defect	Cost to Correct Major Defect
Inspection	1	1
Testing	4	Between 6 and 8
Field Use	16	6 to 8 times the cost in testing

Throughout this study, the testing and field use costs for major defects are assumed to be uniformly distributed between the values outlined in Table 7, since results obtained from NSQE do not provide any information regarding the distribution.

3.1.4 Modeling the Labor Cost of Reinspection

The two other cost consuming activities of inspection process in addition to the defect correction, are preparation and meeting. The cost of other activities (such as overview, follow-up and consolidation) are not even reported by the related studies because of three possible reasons; (i) the inclusion of the corresponding costs in preparation and meeting costs, (ii) not applying the related activity since it is optional or not embraced in the selected inspection method, (iii) neglecting the corresponding costs. So, in this study the costs regarding the meeting and preparation activities are considered to model the extra labor cost that the project will bear due to conduct of the reinspection. Besides, the indirect costs (such as facility overhead or general & administration) are not included in the cost of reinspection since the related data is unavailable and subject to high variation for different organizations. Furthermore, these costs are actually irrelevant since they are expended by the organization in any case, i.e. whether reinspection is carried out or not. By reviewing the preparation rate and meeting rate values provided in Section 2.2 and by following a conservative approach, it is reasonable to assume a rate of 200 Source Lines of Code (SLOC) per hour for both values in the context of a SW-CMM Maturity Level 3 organization. Since, the code size subject to the study is 1000 SLOC, having these rates results in 5 hours of preparation and 5 hours of meeting effort for each inspector participating to the reinspection. In addition to this, since the author of the inspected artifact is assumed to be among the reinspection participants in order to discuss and answer issues, an additional 5 hours is spent for meeting. Thus, the labor cost of reinspection (C_R) can be expressed as; (Number of Reinspectors × 10 + 5) man-hours.

3.1.5 Defining the Output Measures of the Simulation Model

As mentioned before in the text, the three objectives, for which software projects strive to improve their performance, are cost, schedule and quality. For being able to evaluate the cost, schedule delay and defect containment that come out at the end of the software life cycle (defined by the simulation model) as a result of applying various reinspection decision methods, the related measures need to be designated. The selected measures and the underpinning reasons are described in the paragraphs below. In this study, these measures are computed without considering the specific software development in which the values of some measures may escalate. For example, a defect that propagates to the user for life critical software has much more important consequences in terms of quality. However, the effect of software type is accounted while evaluating the reinspection decision policies by determining a number of preference profiles for output measures, which actually correspond to different importances assigned to cost, quality and schedule according to software type developed (See Section 3.6 and Table 25).

- **Total Cost (TC):** When a reinspection is conducted, the project faces the corresponding labor cost and correction cost. However, at the same time the correction costs in testing and field use are decreased due the defects found in reinspection, i.e. the rework cost is saved. So, this impact of reinspection should be taken into account in the study. This measure provides the cost performance observed as a result of following a certain reinspection decision method and refers to the costs incurred throughout the phases embraced by the simulation model. In the study, costs of initial inspection and testing are irrelevant along with indirect costs, since all of these costs are assumed to be independent of the conduct of reinspection, i.e. the related expenditures are the same regardless of reinspection decision. For example, the testing effort is constant as stabilized in the project schedule and it is not altered according to results of previous inspection activities. So, the defect correction costs and reinspection cost (if conducted) are taken into account to calculate the total cost. Consequently, in the context of the study, the total cost for a single simulation replication is calculated as follows.

$$TC = C_R + c_R^{mj} D'_{mj} + c_R^{mn} D'_{mn} + c_T^{mj} T_{mj} + c_T^{mn} T_{mn} + c_F^{mj} F_{mj} + c_F^{mn} F_{mn} \quad \text{(Eq. 45)}$$

where,

TC: Total cost incurred during the software life cycle

C_R: The labor cost of conducting the reinspection (excluding the related correction cost)

$c_R^{mj}, c_T^{mj}, c_F^{mj}$: The cost to correct a major defect found during reinspection, testing and field use, respectively.

$c_R^{mn}, c_T^{mn}, c_F^{mn}$: The cost to correct a minor defect found during reinspection, testing and field use, respectively.

D'_{mj}, D'_{mn}: The number of major and minor defects, respectively, identified during reinspection.

T_{mj}, T_{mn}: The number of major and minor defects, respectively, identified during testing.

F_{mj}, F_{mn}: The number of major and minor defects, respectively, identified during field use.

In order to evaluate the reinspection decision methods, the total cost values (TC) resulting for all replications are averaged.

- **Schedule Delay (SD):** As a reinspection is carried out, the project's completion time may be affected adversely, since the reinspection may cause the late start of other activities which are on the critical path. Hence, this impact of the reinspection should be considered while analyzing the effects of different reinspection decision methods. Actually, finding out the delay in the project schedule –if any- when a reinspection is performed, requires the availability of the entire project schedule. However, this information is not covered in the model used in this study. Because of this reason, another way should be found to investigate the delay caused by the reinspection. In order to achieve this, the percentage of replications in which reinspection is conducted, is selected as the required measure.

- **Major Field Defects (MjFD):** It is expected that the number of defects transferred to the field use (i.e. found by the user) decreases as the software code is reinspected. Hence, this is another impact that should be analyzed regarding reinspections. The number of defects detected during field use provides the indication of software quality as seen by the user. However, since minor defects do not generally hinder the users' business, the number of major defects propagating to the field use is a more appropriate measure to investigate the effect of reinspections regarding software quality. So, in the study the average (among all replications) number of major defects that users encounter is employed as the third measure.

3.1.6 Outline of the Study

Throughout the study a number of steps are carried out in line with the aim of evaluating various reinspection decision methods in the context of a SW-CMM Maturity Level 3 organization by using the simulation model described above along with the measures that underpin the evaluation task. Firstly, the factors that possibly impact the results of the study are put forward. Then, the levels of these factors are deduced for manifesting the varying circumstances inherent in the simulation model. By using these levels, an experiment is designed which enables the enumeration of different circumstances that are encountered. Since, the main objects in the study are the reinspection decision methods, the next step is to expose the reinspection policies (regarding code inspections) that are going to be compared in the course of the study. This in turn requires the selection of a defect content estimation technique (DCET), which serves satisfactorily in all conditions, for estimating the number of remaining defects in the inspected artifact. Further, since some of the considered reinspection decision methods are based on the estimated number of defects to be found during reinspection, a suitable defect detection capability estimation technique (DDCET) needs to be chosen. Next, the total cost, schedule delay, and major field defects measures are found out for all of the selected reinspection policies by executing the pre-specified simulation model in all the cases covered by the designed experiment. Then, the results from the simulation are evaluated by employing the performance measures regarding cost, schedule and quality. During this evaluation, different weight combinations assigned to cost, schedule and quality dimensions are considered with the aim of proposing appropriate reinspection decision policies for the software projects with different objectives and circumstances. Finally, the effects of the factors considered in the study on the output measures are investigated by performing analysis of variance (ANOVA).

3.2. Determining the Factors and Experimental Layout of the Study

A number of factors, which underpin the study and the simulation model, are described in Section 3.1. The varying values of these factors certainly affect the results regarding the evaluations for different reinspection decision methods. Hence, the factors and their levels that determine the various conditions where simulation model will be executed should be put forward. By reviewing the discussions in the previous section the following factors and corresponding levels come into play in the context of a SW-CMM Maturity Level 3 organization.

- **The probability that a given defect is major (p_m):** As mentioned before, the severity of a defect (as major or minor) specifies the related correction costs. Note that, the complement of p_m value (i.e. $1-p_m$) denotes the probability that a defect is minor. In the course of software projects major defects are less likely to be encountered than minor defects. Hence, it is reasonable to select the levels of p_m as smaller than 0.5. These levels are selected as 0.1, 0.25, and 0.4 which correspond to the low, medium, and high probabilities, respectively.

- **The probability that a given defect is difficult (p_d):** The detectability degree of a defect determines its possibility for being captured by the inspectors in inspection or reinspection. So, as a defect is introduced in the simulation model its detectability should be assigned as difficult or easy. p_d value enables to handle this task, i.e. according to this probability, if it occurs, the defect is treated as a difficult one. Otherwise, the defect at hand is an easy one. Consequently, in this study three levels are proposed for p_d as 0.2, 0.5, and 0.8. Actually, these levels also represent the inspection difficulty of the artifact as low, medium, and high.

- **Inspection detection probabilities of difficult and easy defects (p_1, p_2):** The simulation model should take as one of its inputs the probability of detection during inspection and reinspection for both easy and difficult defects. As it should be clear from the previous explanations, the probability that a defect is captured by any inspector is lower for difficult defects than easy defects, i.e. $p_1 \leq p_2$. In this study, the possible values for detection probabilities of easy and difficult defects should be considered along with the proximity of their probabilities. Hence, in order to achieve this, the following (p1, p2) pairs are designated: (0.1, 0.9), (0.25, 0.75), and (0.4, 0.6). Note that these levels correspond to extreme, moderate, and low difference in detection probabilities, respectively.

- **Number of inspectors (k):** In general, as the number of inspectors increases, the number of defects captured during inspection also increases. So, inspection team size is another factor that should be considered throughout the study. In Section 2.2, the range of inspector team size is stated as 2 to 4 for code inspection. So, by using this information, the corresponding levels are selected as 2, 3, and 4.

- **Number of reinspectors (k'):** By following the same reasoning that is applied for number of inspectors above, the levels regarding the number of reinspectors are designated as 2,3, and 4. Certainly, this factor is relevant if 'reinspect' decision is made by the reinspection decision method. If this is the case, the project manager determines the inspection team that will conduct the reinspection. Hence, the reinspection may be carried out with the same, higher or lower number of inspectors. For example, it is possible to have 2 inspectors during inspection and 4 for reinspection, which represents probably a situation that project manager wants to ensure the quality of the inspected artifact by devoting additional resources for reinspection.

- **Capability of a particular inspector (q_j):** For assigning the capturing probability of a specific defect during inspection or reinspection, the capabilities of the related inspectors should also be known, besides the defect detectability. As explained in the previous sections of the text, the capability of an inspector is represented with the probability that s/he detects any defect (which is independent of severity and detectability of the defect). In order to determine the levels corresponding to this probability, the Code People Review defect removal fraction values proposed by Defect Removal submodel of COQUALMO are employed. Very low and low people review ratings in COQUALMO correspond to no peer reviews and ad-hoc peer reviews, respectively. The peer review implementation put forward by SW-CMM Maturity Level 3 requires documented procedure, training, defined roles, and use of checklists. Hence, it is reasonable to include the high and very high ratings of COQUALMO. Also, the nominal rating should be included in order to account for the situations where the related peer review procedures are followed improperly (causing a decrease in peer review effectiveness with respect to expectations). Furthermore, extra high rating is not included, because it refers to the practices (such as root cause analysis and statistical process control) which are not embraced in many inspection methods. As a result, the inspector capability levels to be used in this study are chosen as 0.48, 0.6, and 0.73 (See Table 4 in Section 2.3). Actually, in the study these are employed as the capability of an average inspector. It is assumed that there is not much variation among the capabilities of different inspectors, since they obtain same training for inspection process and other processes. However, still some degree of variation is present due to different background and experience of the inspectors. Hence, in order to account for this issue, in this study, the capability of a specific inspector is assumed to be normally distributed around the related capability level with standard deviation 0.05 (assuming a deviation of ± 0.15 around the average that covers about 99% of the population), e.g. if the inspector capability level at hand is 0.6, during a single replication of the Monte Carlo simulation, the probability representing the capability of an inspector will be sampled from the distribution N (0.6, 0.0025).

- **Testing detection probability of a defect (p_t):** If a defect remains undetected after inspection activities, it has still chance to be identified via testing before it is transferred to the user. The model proposed in this study requires the determination of the probability that a particular defect is detected during testing. As mentioned before in the text, this probability is taken as constant for all defects. The levels of the testing detection probability are selected again by referring to COQUALMO model. The Defect Removal submodel of COQUALMO provides different ratings for eliminating defects via execution testing and tools. In this study, the nominal and very high ratings are taken into account as two cases which represent the content of the testing activities employed by SW-CMM Maturity Level 3 organizations. Namely, nominal rating corresponds to the basic activities regarding different testing types as unit testing, integration testing and system (acceptance) testing. Whilst, very high rating refers to the addition of advanced testing tools and techniques while applying the practices given by nominal case. The high rating is not included in the study since the corresponding defect removal fraction value does not differ much from one higher and lower ratings, i.e. from nominal and very high ratings. This also enables the saving from the number of cases enumerated in the study due to assumption that small difference in defect removal fraction value does not affect the results significantly. Besides, very low and low ratings are not considered, since they relate to cases where no testing and ad-hoc testing are applied, respectively. And extra high rating is not included, since it describes very extreme cases regarding testing capability. Finally, the two levels selected for p_t value are 0.58 and 0.78 (See Table 4 in Section 2.3).

As a result of the above discussions the factors and the levels to be considered throughout the study while executing the replications of Monte Carlo simulation can be outlined as given in Table 8 below. In this table, the levels are numbered according the increasing levels of factors.

Table 8 Factors and Corresponding Levels Considered in the Study

Factor	Related Activity	Levels
p_m	Coding	1) 0.1 2) 0.25 3) 0.4
p_d	Coding	1) 0.2 2) 0.5 3) 0.8
(p_1,p_2)	Inspection	1) (0.1,0.9) 2) (0.25,0.75) 3) (0.4,0.6)
k	Inspection	1) 2 2) 3 3) 4
$k^{'}$	Reinspection	1) 2 2) 3 3) 4
q_j	Inspection and Reinspection	1) N(0.48, 0.0025) 2) N(0.60, 0.0025) 3) N(0.73, 0.0025)
p_t	Testing	1) 0.58 2) 0.78

In order to represent different conditions resulting from the varying levels of the above factors, an experiment needs to be designed. For this purpose, the Taguchi design (orthogonal array, fractional factorial design) shown in Table 9 is constructed by using the Minitab Statistical Software. The replications to be used for various analysis purposes in the succeeding parts of the study employ the treatments given in this table. Namely, the related treatments are repeatedly simulated for each of the 54 treatments in order to obtain data for different situations comprised in the study context. (The details of the simulation study are explained in Section 3.6.)

Table 9 Experimental Layout Employed in the Study

Treatment	p_t	p_d	p_m	q_j	(p_1,p_2)	**k**	k'
1	1	1	1	1	1	1	1
2	1	1	1	2	1	1	1
3	1	1	1	3	1	1	1
4	1	2	2	1	2	1	2
5	1	2	2	2	2	1	2
6	1	2	2	3	2	1	2
7	1	3	3	1	3	1	3
8	1	3	3	2	3	1	3
9	1	3	3	3	3	1	3
10	1	1	2	1	2	2	1
11	1	1	2	2	2	2	1
12	1	1	2	3	2	2	1
13	1	2	3	1	3	2	2
14	1	2	3	2	3	2	2
15	1	2	3	3	3	2	2
16	1	3	1	1	1	2	3
17	1	3	1	2	1	2	3
18	1	3	1	3	1	2	3
19	1	2	1	1	3	3	1
20	1	2	1	2	3	3	1
21	1	2	1	3	3	3	1
22	1	3	2	1	1	3	2
23	1	3	2	2	1	3	2
24	1	3	2	3	1	3	2
25	1	1	3	1	2	3	3
26	1	1	3	2	2	3	3
27	1	1	3	3	2	3	3
28	2	3	3	1	2	1	1
29	2	3	3	2	2	1	1

Table 9 Experimental Layout Employed in the Study (cont.)

Treatment	p_t	p_d	p_m	q_j	(p_1,p_2)	**k**	k'
30	2	3	3	3	2	1	1
31	2	1	1	1	3	1	2
32	2	1	1	2	3	1	2
33	2	1	1	3	3	1	2
34	2	2	2	1	1	1	3
35	2	2	2	2	1	1	3
36	2	2	2	3	1	1	3
37	2	2	3	1	1	2	1
38	2	2	3	2	1	2	1
39	2	2	3	3	1	2	1
40	2	3	1	1	2	2	2
41	2	3	1	2	2	2	2
42	2	3	1	3	2	2	2
43	2	1	2	1	3	2	3
44	2	1	2	2	3	2	3
45	2	1	2	3	3	2	3
46	2	3	2	1	3	3	1
47	2	3	2	2	3	3	1
48	2	3	2	3	3	3	1
49	2	1	3	1	1	3	2
50	2	1	3	2	1	3	2
51	2	1	3	3	1	3	2
52	2	2	1	1	2	3	3
53	2	2	1	2	2	3	3
54	2	2	1	3	2	3	3

3.3. Determining the Reinspection Policies to be Evaluated

Recall that, the aim of this study is to compare various reinspection decision methods by observing their effects on cost, schedule and quality related objectives of a software project. In Section 2.6, five reinspection decision methods revealed from the literature are described. By employing these methods, the reinspection policies to be evaluated in this study are determined. In the context of the study, a reinspection policy refers to the rule used to decide whether to reinspect the code or not. Determination of the policies requires the designation of different values for the thresholds inherent in the outlined reinspection decision methods. Besides adapting all of the available methods for all defects (i.e. without differentiating major and minor defects), the methods I through IV are also adapted for major defects. The consideration of this kind of policies enables the study to account for the perspective which focuses on the removal of major defects, since they are deemed as much more critical than minor defects. Furthermore, in addition to the policies determined by adapting the reinspection decision methods, also 'Never Reinspect' and 'Always Reinspect' policies are included in the analysis. In order to increase the validity of the study, the thresholds regarding the policies are selected with the aim of coming up with a representative set of policies. Hence, while determining the policies tight threshold values are included as well as loose and moderate ones. However, in practice, i.e. for software development projects in real life, these thresholds are generally determined by considering the historical data generated from previous projects of the organization. The resulting policies are provided in Table 10. Note that, the thresholds for policies based on defect density include 'number of defects' in the description column, because in the study the code size is fixed in all cases as 1 KSLOC (recall that when defect density value is multiplied by code size it provides the number of defects).

Table 10 Reinspection Policies Considered in the Study

Policy	Threshold Value	Reinspection decision method	Threshold Description	Considered Defect Severity Class
1	0.25	I	Minimum Effectiveness After Inspection	All
2	0.5			
3	0.6			
4	0.75			
5	0.9			
6	0.25	I	Minimum Effectiveness After Inspection	Major
7	0.5			
8	0.6			
9	0.75			
10	0.9			
11	3	II	Upper Bound for the Allowable Number of Defects After Inspection	All
12	6			
13	9			
14	12			
15	15			
16	1	II	Upper Bound for the Allowable Number of Defects After Inspection	Major
17	2			
18	3			
19	4			
20	5			
21	2	III	Upper Bound for the Allowable Number of Defects After Reinspection	All
22	4			
23	6			
24	8			
25	10			

Table 10 Reinspection Policies Considered in the Study (cont.)

Policy	**Threshold Value**	**Reinspection decision method**	**Threshold Description**	**Considered Defect Severity Class**
26	1	III	Upper Bound for the Allowable Number of Defects After Reinspection	Major
27	2			
28	3			
29	4			
30	5			
31	0.25	IV	Minimum Effectiveness After Reinspection	All
32	0.5			
33	0.6			
34	0.75			
35	0.9			
36	0.25	IV	Minimum Effectiveness After Reinspection	Major
37	0.5			
38	0.6			
39	0.75			
40	0.9			
41	0	V	Minimum Net Benefit	All
42	25			
43	50			
44	75			
45	100			
46	N/A	Never Reinspect	N/A	N/A
47	N/A	Always Reinspect	N/A	N/A

In the study, it is possible to encounter some occurrences where during inspection the number of identified defects is zero. This situation is especially probable for major defects, since their number is low vis-à-vis all defects. Hence, in these cases the above policies are adapted to result in 'do not reinspect decision', due to (i) unavailability of the information that enables the estimation of number of defects contained in the artifact, and (ii) in practice this probably is deemed as inexistence of considerable amount of defects in the artifact. Furthermore, the above policy formulations include the estimated values, which are obtained by applying appropriate DCETs and DDCETs. These estimates may result in non-integer values. For such situations, the rounding of the resulting value is not carried out, by assuming it is actually an approximation/average for the number of defects contained in the artifact and for the number of defects that will be identified during reinspection.

3.4. Selecting the Defect Content Estimation Technique to be Employed

Most of the reinspection decision methods described in Section 2.6 are based on the estimate for number of defects present in the inspected artifact ($\hat{N}$). Consequently, the policies to be evaluated in this study (See Section 3.3) also require the calculation of this estimate for concluding reinspection decision. Hence, a defect content estimation technique (DCET), which performs satisfactorily in the underlying context, should be determined for accomplishing the purposes of this study. While selecting an appropriate DCET among the various DCETs outlined in Section 2.4, the techniques that rely on the subjective estimations of software practitioners are not taken into account, since they necessitate the people, from whom the estimates will be obtained. This shows that, such techniques are not applicable in a simulation study like this one. Furthermore, the objective curve-fitting models are also not preferable for the simulation purposes. Because, making the defect estimation by using one of the available curve-fitting models requires extensive work for determining the line that best represents the inspection data at hand. Certainly, to conduct that fitting operation for each replication of the simulation is infeasible with respect to cost (measured in terms of processing time). Furthermore, the literature implies that Capture-Recapture estimators perform better than curve-fitting techniques. As a result, in this study, the Capture-Recapture estimators are considered, since they are also suitable and practical for the conditions of the study.

Recalling from Section 2.4, there exist six different estimators provided by the CR models that are applicable to software inspections. Again as mentioned in Section 2.4, some of these estimators, namely Chao's Heterogeneity-Time and Jack-knife estimators, have sub-estimators. In this study these sub-estimators are treated as separate estimators while selecting the best CR estimator to compute the value of $\hat{N}$. For Chao's Heterogeneity-Time estimator, the unavailability of the guidance for selecting the proper sub-estimators is the main reason for this fact. Whilst, for Jack-knife estimator the claims of Miller (1999) regarding the inappropriateness of using the selection procedure proposed by Burnham and Overton (1978) for software engineering practices, and the results supplied by Thelin et al. (2002), are accepted as the evidence for employing the sub-estimators along with their practicality for simulation purposes. Consequently, this results in 12 estimators considered in this study. These estimators are listed in Table 11. Actually, some of these estimators are not candidates for selection at certain levels designated for number of inspectors and reinspectors in the designed experiment (see Section 3.2). By definition, they are not applicable in the cases where the number of inspectors is equal to 2 (estimators 7, 10, 11, 12), 3 (estimators 11, 12), or 4 (estimator 12).

Table 11 Capture-Recapture Estimators Considered in the Study

Estimator Number	Estimator Name
1	Maximum Likelihood Estimator for Model M0 (MLE-M0)
2	Maximum Likelihood Estimator for Model Mt (MLE-Mt)
3	Chao’s Time Estimator (ChaoMt)
4	Chao’s Heterogeneity Estimator (ChaoMh)
5	Chao’s Heterogeneity-Time Estimator (ChaoMth) Order 1
6	Chao’s Heterogeneity-Time Estimator (ChaoMth) Order 2
7	Chao’s Heterogeneity-Time Estimator (ChaoMth) Order 3
8	Jack-knife Estimator (JKMh) Order 1
9	Jack-knife Estimator (JKMh) Order 2
10	Jack-knife Estimator (JKMh) Order 3
11	Jack-knife Estimator (JKMh) Order 4
12	Jack-knife Estimator (JKMh) Order 5

In order to differentiate among the remaining 8 estimators, different DCET evaluation measures available through the literature are utilized (see Section 2.5.1). Namely, Failure Rate (FR), Relative Error (RE), Decision Accuracy (DA), Relative Decision Accuracy (RDA), and Root Mean Square Error (RMSE) measures are computed by replicating the underlying simulation model 1000 times for each estimator and for each treatment. Certainly, increasing the number of replications above 1000 would make the analysis results more reliable. However, this is not feasible because of the limited processing time. So, the reasonability for number of replications is considered by computing the estimated standard error values regarding 1000 replications for the related measures, namely, RE, FR and DA. For RDA and RMSE measures, estimated standard error is not revealed, since RDA is actually based on DA measure and RMSE, by definition, includes the variability among 1000 replications. The highest estimated standard error values encountered through all treatments for RE, FR and DA measures are 0.0212, 0.0158, and, 0.0158, respectively. These values show that the number of replications employed for determining the appropriate estimator is reasonable. Because the estimated standard errors are small enough (compared to the average of all replications) to decrease the variability among different replications to an acceptable level, while obtaining the estimator's performance for each treatment. During the estimator comparison, using the designed experiment, which represents the different software inspection circumstances, enhances further the validity of the results with respect to variability. The succeeding paragraphs present the results of the analysis performed for each measure.

The values for FR measure are revealed by calculating, for a given replication, the percentage of replications, in which the related estimator fails to produce an estimate. This generally occurs when 'division by zero' is resulted due to estimator formulation. Table 12 supplies the FR statistics of each estimator for 54 treatments of the designed experiment. Further, the related boxplots are depicted in Figure 3. As can be seen from seen from Table 12, a number of estimators always estimate the number of defects present in the artifact, whilst others fail 4-5% of the time. Although, the mean values for different estimators seem not to differ significantly, the corresponding boxplots reveal that high failure rates are encountered for several estimators, namely for estimators 1, 2, 4, 5, and 6 in a number of treatments. When the estimators' FR profile for defects with major severity is examined with the aim of inspecting their performance related to policies based on major defects, the situation is even worse for these estimators. As evident from Figure 4 and Table 13, for major estimators they systematically result in failure, since the number of major defects is low (due to the experimental design) with respect to total number of defects. So, these estimators are deemed as inappropriate for using them throughout the study, since the estimator to be selected should produce estimates most of the time for all treatments and such estimators are already available. Hence, as a result of these facts, estimators 1, 2, 4, 5, and 6 are removed from the evaluations, which leaves estimators 3, 8, and 9 as the available options regarding the best DCET for the study.

Table 12 Failure Rate Statistics for the Estimators Regarding All Defects

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
1	0.0379	0.002	0.075	0	0.356	0	0.038
2	0.03783	0.0035	0.07266	0	0.326	0	0.04
3	0	0	0	0	0	0	0
4	0.0487	0.0095	0.0776	0	0.361	0.0008	0.078
5	0.0397	0.0035	0.0794	0	0.377	0	0.0395
6	0.0371	0.0025	0.0753	0	0.38	0	0.0305
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0

Table 13 Failure Rate Statistics for the Estimators Regarding Major Defects

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
1	0.3453	0.3292	0.2508	0	0.8218	0.0913	0.5742
2	0.3441	0.3326	0.2518	0	0.8164	0.0935	0.5466
3	0	0	0	0	0	0	0
4	0.3989	0.3943	0.245	0.027	0.8058	0.1445	0.5933
5	0.3408	0.3294	0.249	0	0.8105	0.0987	0.5543
6	0.344	0.3423	0.2484	0	0.8008	0.1085	0.5669
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0

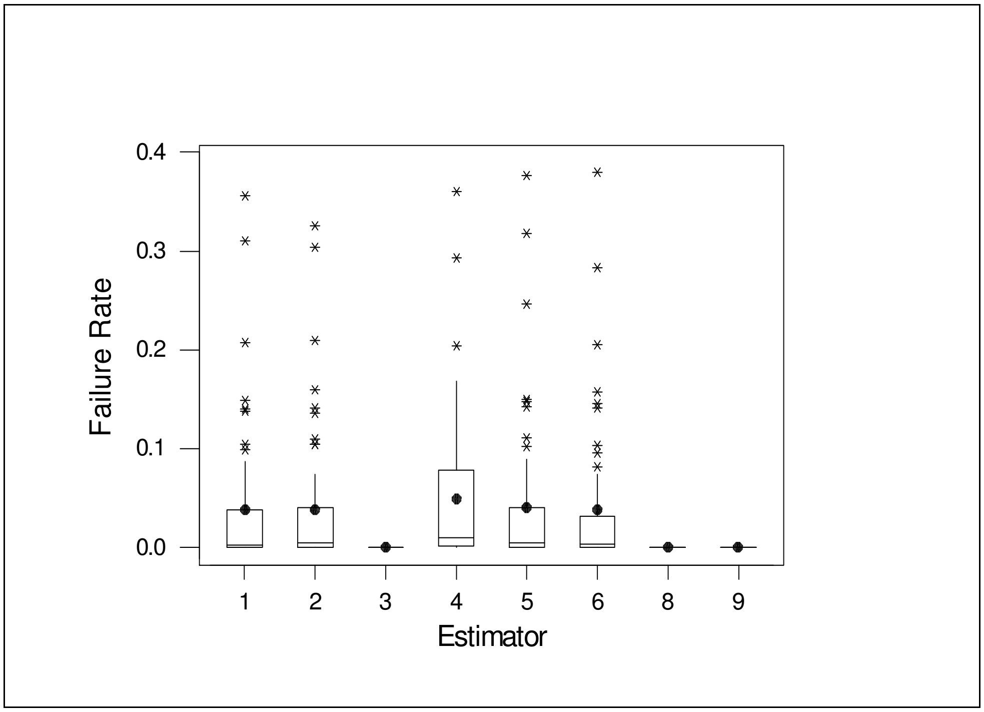

Figure 3 Boxplots for Estimator Failure Rate Regarding All Defects

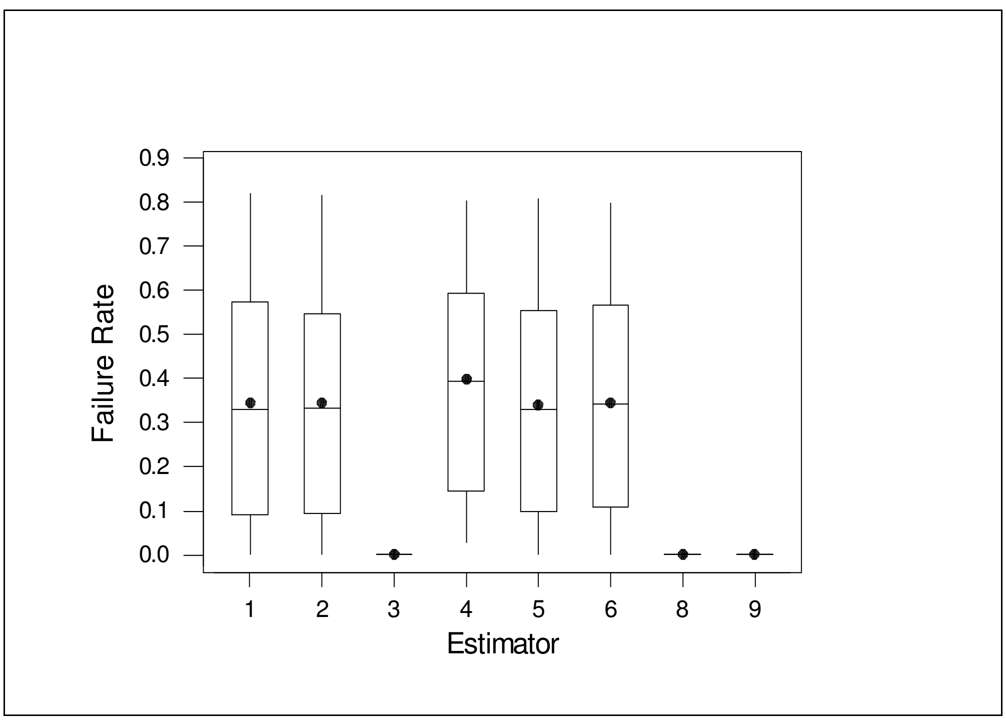

Figure 4 Boxplots for Estimator Failure Rate Regarding Major Defects

For the remaining three estimators, DA and RDA measures are considered by computing their average values through all treatments for each reinspection policy listed in Section 3.3. Actually, the average of the mean DA and mean RDA for each treatment, which are computed by taking into account 1000 replications corresponding to the treatment, are calculated. For each estimator, the DA and RDA statistics generated by taking into account the related average values of 45 reinspection policies, are given in Table 14 and Table 15, respectively. Also, the corresponding boxplots are shown in Figure 5 and Figure 6. Even without conducting hypothesis testing, these tables and figures obviously reveal no evidence supporting the fact, that one of the three estimators is better than others with respect to enabling more accurate decisions. So, based on these two measures, none of the remaining estimators can be eliminated. Further, by recalling that the best values for DA and RDA measures are 1 and 0, respectively, it can be concluded that, for the purposes of this study all of the three estimators perform satisfactorily on the average with respect to decision accuracy.

Table 14 Decision Accuracy Statistics for the Estimators

Estimator	**Mean**	**Median**	**Std Dev.**	**Min**	**Max**	**1st Quartile**	**3rd Quartile**
3	0.6763	0.6550	0.1367	0.4374	0.9622	0.5824	0.7471
8	0.6764	0.6673	0.1669	0.3054	0.9645	0.5704	0.7795
9	0.6707	0.6660	0.1579	0.3051	0.9647	0.5605	0.7620

Table 15 Relative Decision Accuracy Statistics for the Estimators

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
3	0.0378	-0.0002	0.1964	-0.3388	0.7649	-0.0285	0.0866
8	0.0337	-0.0011	0.2666	-0.5252	0.9047	-0.0634	0.0132
9	0.0288	-0.0234	0.2639	-0.3923	0.8946	-0.1143	0.0151

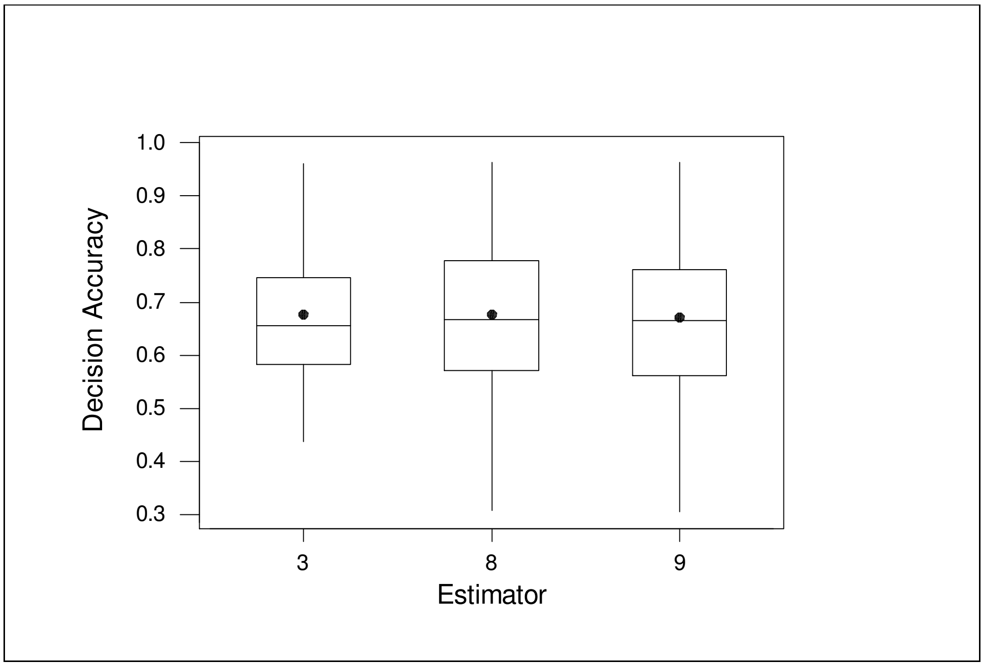

Figure 5 Boxplots for Estimator Decision Accuracy

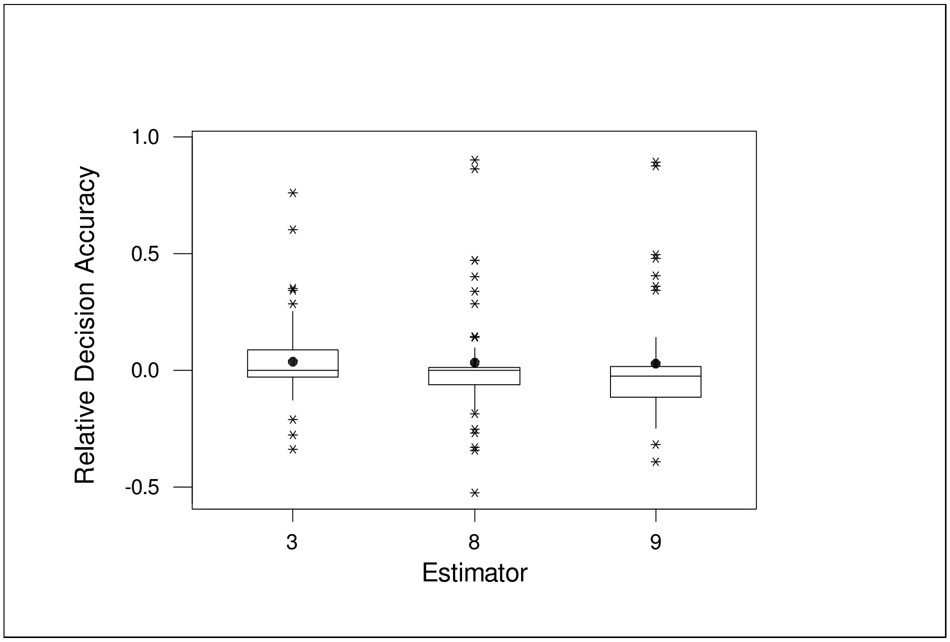

Figure 6 Boxplots for Estimator Relative Decision Accuracy

In order to compare the performance of the estimators regarding the accuracy of the estimations they produce, RE and RMSE measures are employed for each. Since, some of the selected policies depend on only major defect containment; the related evaluations are also performed for major defects, in addition to evaluations conducted for all defects (i.e. regardless of the defect severity as major or minor). By calculating RMSE value through all replications of a treatment, the variance and bias of an estimator are accounted at the same time. If we deem each 1000 replications as a sample, this enables to observe in-sample results for the accuracy of the estimations with respect to the true value. As can be seen from Table 16, Figure 7 and Figure 9, RMSE values for major defects do not exhibit distinction among the estimators 3 and 9. However, estimator 8 seems to be better than the other two (recall that low values are more favorable for this measure), although not justified via pairwise comparisons of Fisher's Test (given in Figure 7). Further, regarding all defects, RMSE performance of estimator 3 seems deficient vis-à-vis estimators 8 and 9, based on results depicted in Table 17, Figure 8 and Figure 10. However, same exhibitions show that when only RMSE performance regarding all defects is considered, the tie between estimators 8 and 9 is not broken. Consequently, as a result of the above discussion, it can be concluded that RMSE measure favors estimators 8 and 9 over estimator 3, but estimators 8 and 9 do not outperform each other, although estimator 8's RMSE performance seems better. Estimators 8 and 9 are even deemed as equivalent with respect to RMSE values computed regardless of defect severity.

Table 16 Root Mean Square Error Statistics for the Estimators through Major Defects

Estimator	**Mean**	**Median**	**Std Dev.**	**Min**	**Max**	**1st Quartile**	**3rd Quartile**
3	3.422	3.503	1.37	1.638	6.796	2.06	4.293
8	3.007	2.852	1.141	1.61	6.508	1.887	3.584
9	3.409	3.532	1.175	1.591	6.267	2.284	4.131

```
Analysis of Variance for RMSE
Source     DF        SS        MS        F        P
Estimato    2      5.99      3.00     1.97    0.143
Error     159    241.59      1.52
Total     161    247.58
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
3          54     3.422     1.370                 (----------*----------)
8          54     3.007     1.141  (----------*----------)
9          54     3.409     1.175                 (----------*----------)
                                   -+---------+---------+---------+-----
Pooled StDev =    1.233           2.70      3.00      3.30      3.60

Fisher's pairwise comparisons

    Family error rate = 0.122
Individual error rate = 0.0500

Critical value = 1.975

Intervals for (column level mean) - (row level mean)

                  3           8

       8      -0.054
               0.883

       9      -0.455      -0.870

               0.482       0.067
```

Figure 7 MINITAB Output for One-way ANOVA and Fisher's Test Results (with 95 % confidence) of Root Mean Square Error through Major Defects

Table 17 Root Mean Square Error Statistics for the Estimators through All Defects

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
3	8.949	9.293	3.434	2.826	16.012	5.422	11.483
8	7.315	6.36	4.01	2.979	15.945	3.761	10.689
9	7.593	6.665	3.304	3.492	15.249	4.856	10.01

```
Analysis of Variance for RMSE
Source     DF        SS        MS        F        P
Estimato    2      82.6      41.3     3.19    0.044
Error     159    2056.0      12.9
Total     161    2138.6
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+--------
3          54     8.949     3.434                   (--------*--------)
8          54     7.315     4.010  (---------*---------)
9          54     7.593     3.304     (---------*---------)
                                   -------+---------+---------+--------
Pooled StDev =    3.596                 7.0       8.0       9.0

Fisher's pairwise comparisons

    Family error rate = 0.122
Individual error rate = 0.0500

Critical value = 1.975

Intervals for (column level mean) - (row level mean)

                  3           8

       8       0.268
               3.001

       9      -0.010      -1.645

               2.723       1.089
```

Figure 8 MINITAB Output for One-way ANOVA and Fisher’s Test Results (with 95 % confidence) of Root Mean Square Error through All Defects

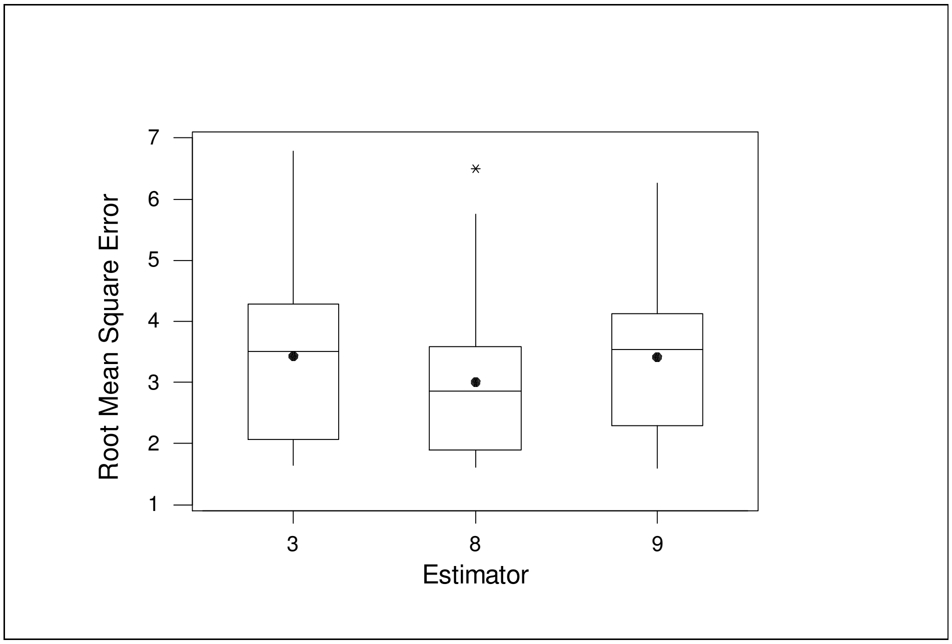

Figure 9 Boxplots for Estimator Root Mean Square Error through Major Defects

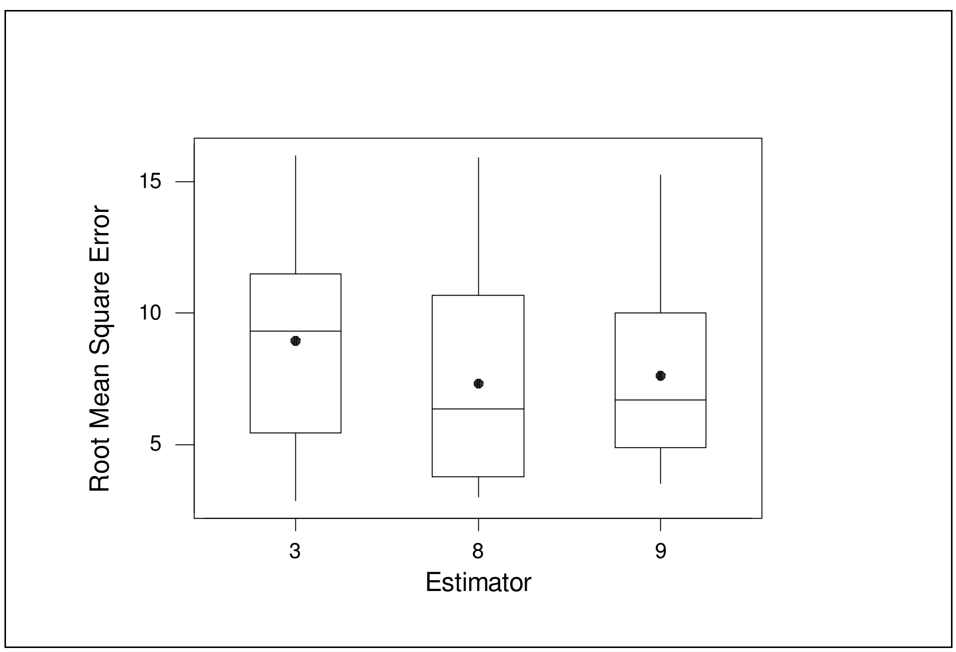

Figure 10 Boxplots for Estimator Root Mean Square Error through All Defects

Relative Error (RE) measure is further used as a final criterion to compare the three estimators. While finding out the values for RE measure, the mean value among 1000 replications is computed for each treatment. The related RE statistics are summarized in Table 18 and Table 19, for all defects and major defects, respectively. Additionally, the boxplots showing the distribution of RE values among 54 treatments are plotted in Figure 11 and Figure 12, for all defects and major defects, respectively. As clearly seen from this analysis, all three estimators tend to underestimate. This result conforms to the findings of the previous studies in the literature, as explained in Section 2.5.2. Additionally, as mentioned again in Section 2.5.1, the RE value of an estimator is considered acceptable if it lies in the range ± 20% on the average. Hence, by considering this advice along with its RMSE measure performance, estimator 3 should not be selected as the DCET to be used in this study based on the mean and median RE values regarding major defects. Further, when the RE statistics of estimator 8 is examined, its mean and median RE values are found out unsatisfactory according to ± 20% criterion, since they all lie on the boundary of -0.2. Whilst, estimator 9 performs well in terms of its accuracy (especially in the case of major defects), i.e. corresponding mean and median results are in the required range and close to the optimum value, which is zero. Consequently, since also its RMSE performance is adequate, estimator 9, namely Jack-knife (JKMh) estimator order 2, seems as the most appropriate Capture-Recapture estimator in the scope of this study. This finding supports the general result in the literature regarding the superiority of Jack-knife estimator in software inspection context. Further, when the subestimators are considered, the outcome of this study is in line with the work of Thelin et al. (2002), in which, also, second order estimator of Jack-knife is suggested.

Table 18 Relative Error Statistics for the Estimators through All Defects

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
3	-0.1894	-0.1394	0.166	-0.5785	-0.0094	-0.335	-0.0544
8	-0.2059	-0.1913	0.2054	-0.5989	0.0932	-0.3836	-0.027
9	-0.1534	-0.1255	0.2307	-0.5708	0.1859	-0.3523	0.0647

Table 19 Relative Error Statistics for the Estimators through Major Defects

Estimator	Mean	Median	Std Dev.	Min	Max	1st Quartile	3rd Quartile
3	-0.3026	-0.2969	0.1856	-0.7154	-0.0408	-0.4546	-0.1323
8	-0.2051	-0.1893	0.2057	-0.5962	0.0921	-0.399	-0.0206
9	-0.0794	-0.109	0.289	-0.5602	0.4029	-0.3529	0.2015

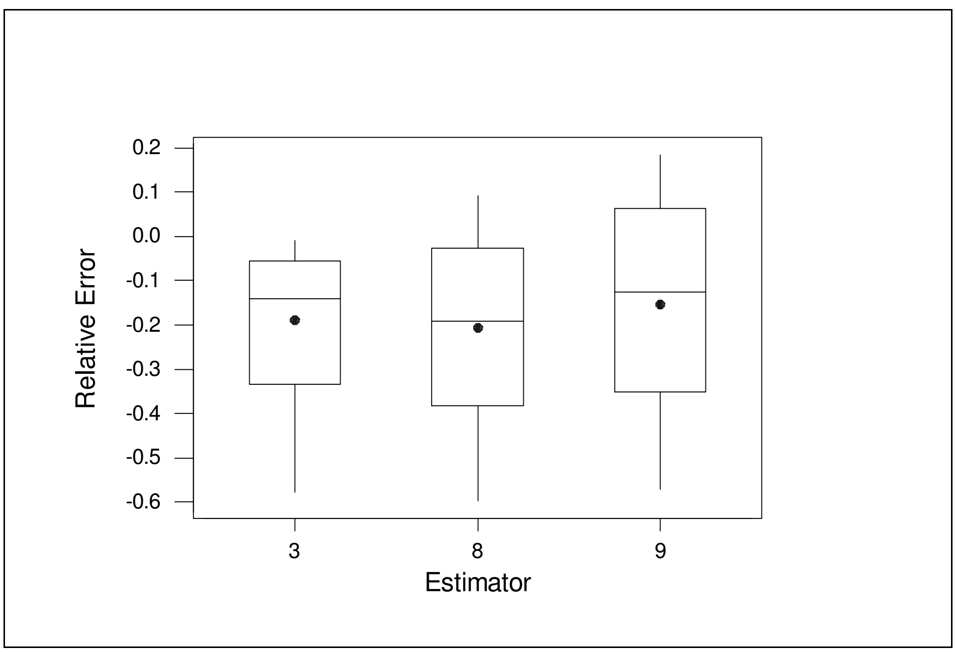

Figure 11 Boxplots for Estimator Relative Error through All Defects

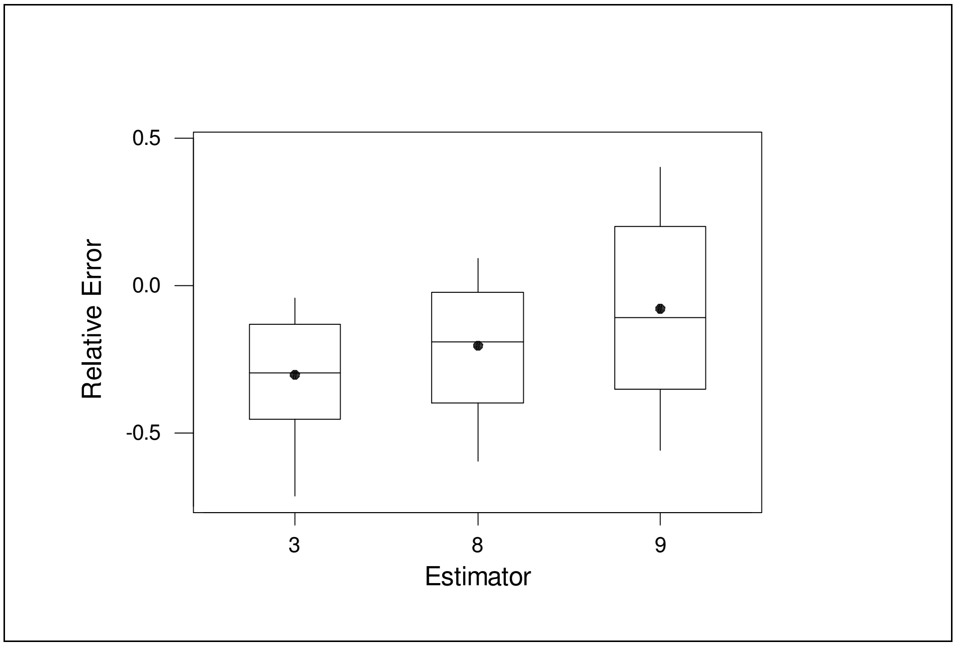

Figure 12 Boxplots for Estimator Relative Error through Major Defects

The simulated data employed during the analysis in this section are provided in Appendix A for FR, in Appendix B for RE, and RMSE measures, and in Appendix C for DA and RDA measures.

3.5. Selecting the Defect Detection Capability Estimation Technique to be Employed

Several policies outlined in Section 3.3, namely the ones that are based on methods III IV and V, require the calculation of the estimated number of defects to be found in a probable reinspection. As mentioned in Section 2.7, Defect Detection Capability Estimation Techniques (DDCET) make this possible. Hence, in order to obtain the results of such policies a DDCET should be designated which serves satisfactorily for the purposes of the study. This means that the approach of determining different DDCETs for various situations is not followed, because the practicality of the proposed policy is one of the main concerns of the study. Therefore, since such an approach complicates the application of the policies in practice, a single DDCET is considered throughout the study. Among the available three DDCETS outlined in Section 2.7, the one based on reliability growth model is not adapted in the study, due to 5 main reasons; (i) the study of Biffl and Gutjahr recommend another DDCET, namely ILM (Biffl and Gutjahr, 2002), (ii) it requires time-stamped data regarding each defect detection event of inspectors, which is not available through the designed simulation model, (iii) available methods for software inspection and their practical applications do not include the collection of this kind of data, (iv) infeasibility of fitting a reliability curve to the data (even it is available) for each replication of the simulation study, and (v) costly and impractical implementation of this approach in practice. The remaining two techniques, namely Optimistic Linear Model (OLM) and Improved Linear Model (ILM), are similar except the latter one accounts for the decreased efficiency in reinspection, which is the case in general as showed in the study of Biffl et al. (2001). Further, ILM is found superior than OLM in the literature (Biffl and Gutjahr, 2002, Biffl and Halling, 2001). Consequently, ILM is deemed as the most appropriate technique for this study, in order to reveal estimated number of defects to be eliminated if reinspection is performed.

3.6. Evaluating the Reinspection Decision Policies

The main aim of this study is to provide guidance on the effects of employing different reinspection policies for code inspections. In order to achieve this, the output measures outlined in Section 3.1.5 are revealed by running the simulation model described in Section 3.1 for each reinspection policy listed in Section 3.3 through all treatments of the designed experiment in Section 3.2. Further, while calculating the estimated values for number of defects in the inspected code and number of defects to be found during reinspection, the estimators designated in sections 3.4 and 3.5, respectively, are employed. The simulation model is executed 500 times for each treatment and for each reinspection policy. In order to make the results comparable, the same random number seed is initialized at the start of each 500 replications. The process describes in the previous sentences is demonstrated in Table 20. The data resulted from the running of all these replications are available in Appendix D. Increasing the number of replications further is deemed unsuitable due to processing time constraints. In order to check the appropriateness for selected number of replications, the estimated standard error values of three output measures are computed. The highest estimated standard error values encountered through all policies and treatments for Total Cost, Schedule Delay and Major Field Defects measures are 3.687, 0.0223, and, 0.0692, respectively. These values show that the number of replications used for evaluating the reinspection decision policies is acceptable, since the estimated standard errors are small enough (compared to the average of all replications) to decrease the variability among different replications to a reasonable level. During the policy evaluation, using the designed experiment, which represents the different software inspection circumstances, improves further the validity of the results with respect to variability.

Table 20 Layout for Evaluating Reinspection Decision Policies

Policies	Treatments	Replications	p_m	p_d	p_1,p_2	k	k'	q_j	p_t	R	TC	SD	MjFD
			Factors								Output Measures		
1	1	Replications	p_m	p_d	p_1,p_2	k	k'	q_j	p_t	R	TC	SD	MjFD
		1	1	1	1	1	1	1	1	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		500	1	1	1	1	1	1	1	:	:	:	:
		Output Measure Values Averaged over Replications									$\overline{TC}^1_1$	$\overline{SD}^1_1$	$\overline{MjFD}^1_1$
	:	:	:	:	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:	:	:	:	:
	54	Replications	p_m	p_d	p_1,p_2	k	k'	q_j	p_t	R	TC	SD	MjFD
		1	2	2	1	3	2	3	3	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		500	2	2	1	3	2	3	3	:	:	:	:
		Output Measure Values Averaged over Replications									$\overline{TC}^1_{54}$	$\overline{SD}^1_{54}$	$\overline{MjFD}^1_{54}$
		Output Measure Values Averaged over Treatments									$\overline{\overline{TC}}_1$	$\overline{\overline{SD}}_1$	$\overline{\overline{MjFD}}_1$
:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:
47	Treatments												
	1	Replications	p_m	p_d	p_1,p_2	k	k'	q_j	p_t	R	TC	SD	MjFD
		1	1	1	1	1	1	1	1	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		500	1	1	1	1	1	1	1	:	:	:	:
		Output Measure Values Averaged over Replications									$\overline{TC}^{47}_1$	$\overline{SD}^{47}_1$	$\overline{MjFD}^{47}_1$
	:	:	:	:	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:	:	:	:	:
	54	Replications	p_m	p_d	p_1,p_2	k	k'	q_j	p_t	R	TC	SD	MjFD
		1	2	2	1	3	2	3	3	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
		500	2	2	1	3	2	3	3	:	:	:	:
		Output Measure Values Averaged over Replications									$\overline{TC}^{47}_{54}$	$\overline{SD}^{47}_{54}$	$\overline{MjFD}^{47}_{54}$
		Output Measure Values Averaged over Treatments									$\overline{\overline{TC}}_{47}$	$\overline{\overline{SD}}_{47}$	$\overline{\overline{MjFD}}_{47}$

where, R: 1 if reinspection is performed, 0, otherwise.

$\overline{TC}^j_i$, $\overline{SD}^j_i$, $\overline{MjFD}^j_i$: Average Total Cost, Schedule Delay, Major Field Defects over 500 replications , respectively, for treatment j of policy i.

$\overline{\overline{TC}}_i$, $\overline{\overline{SD}}_i$, $\overline{\overline{MjFD}}_i$: Average Total Cost, Schedule Delay, Major Field Defects over 54 treatments , respectively, for policy i.

As mentioned before, three measures that are designated in this study for evaluation purposes are Total Cost, Schedule Delay and Major Field Defects. Table 21 depicts, for each reinspection policy, the mean values of these measures averaged over 54 treatments, which are obtained from the simulation study. As the results are examined, it is seen that, although the conduct of a reinpection requires extra cost, as the number of reinspections increase the cost and defect measures tend to decrease. Hence, it is possible to have less cost and less major defects by allocating more time to reinspections, which in turn increases the risk of being late for accomplishing project milestones. Furthermore, as the values in Table 21 are studied more rigorously, it is deduced that some of the policies show equivalent performance regarding cost, schedule and quality. For example, both of policy 1 and 6 never propose to perform a reinspection, thus they end up with same cost and defect values. Truly, when the detailed results given in Appendix D are examined, it is seen that these policies results in exactly the same values for each treatment regarding the number of reinspections performed. Hence, it is wise to retain one policy from such policy groups and to remove the others from the subsequent analysis. The equivalent policy groups are; (1,6,31,36,46), (2,32), and (7,37). In line with the decision for eliminating the redundant policies, the policies 1,6,31,36,32,37 are not further included in the analysis work, but will be considered while discussing the results of the study. Therefore, this remains 41 reinspection policies available for further consideration.

Table 21 Mean Output Measure Values for Reinspection Policies

Reinspection Policy	Mean Total Cost	Mean Schedule Delay	Mean Major Field Defects
1	116.798048	0	0.812370368
2	116.7724958	0.005740741	0.810814814
3	115.7054587	0.162555554	0.764666667
4	105.716796	0.755185184	0.506777776
5	104.7543723	0.970407409	0.443222226
6	116.798048	0	0.812370368
7	116.8169714	0.054259259	0.801111109
8	115.7619471	0.196703704	0.754148146
9	107.8830846	0.587185185	0.563407407
10	105.7882993	0.776518517	0.49137037
11	105.7635882	0.911296296	0.469962963
12	113.4406781	0.540703706	0.641407407
13	115.7654827	0.265962963	0.732962961
14	116.7832228	0.101259258	0.786592594
15	116.9437742	0.020481481	0.809185183
16	107.7358969	0.633962965	0.539074075
17	111.9711735	0.360555555	0.649925928
18	114.8449318	0.186925925	0.730074071
19	115.8786667	0.094407408	0.769518522
20	116.3694679	0.046555556	0.791592596
21	106.9019087	0.506629627	0.578925931
22	108.8620646	0.569666671	0.58148148
23	113.7690238	0.428481484	0.664407412
24	115.4504815	0.290481483	0.71781482
25	116.1648702	0.181999999	0.756000001
26	108.4767602	0.48088889	0.599259256
27	111.7367286	0.354407407	0.658629627
28	114.153731	0.24614815	0.715111107

Table 21 Mean Output Measure Values for Reinspection Policies (cont.)

Reinspection Policy	**Mean Total Cost**	**Mean Schedule Delay**	**Mean Major Field Defects**
29	115.0771243	0.173111112	0.748111113
30	115.5498474	0.131444445	0.768481484
31	116.798048	0	0.812370368
32	116.7724958	0.005740741	0.810814814
33	115.7125367	0.162	0.76474074
34	106.0156293	0.67611111	0.523000003
35	107.3144223	0.533666666	0.571518516
36	116.798048	0	0.812370368
37	116.8169714	0.054259259	0.801111109
38	115.847503	0.17337037	0.756999998
39	108.0554617	0.478222223	0.587555554
40	108.509004	0.37674074	0.615
41	106.1201175	0.826555558	0.495555554
42	110.7139721	0.407740743	0.62714815
43	114.4025569	0.151481481	0.73137037
44	116.3389104	0.056666667	0.786296297
45	116.7985377	0.014888889	0.807111116
46	116.798048	0	0.812370368
47	105.1280341	1	0.436777778

In order to be able to compare output measures resulted for each reinspection policy the average squared total cost, schedule, and major field defect values are employed. The reason for using squared values is to account for both the variability and average over all treatments, since $E(X^2) = [E(X)]^2 + \text{Variance}(X)$, where X is any of the output measures calculated for any reinspection policy. Hence, the average of all 54 squared values obtained for an output measure provides the overall performance of the reinspection policy (regarding all the conditions put forward by the experimental design) with respect to this particular measure. The resulting mean squared values are listed in Table 22 for three output measures.

Table 22 Mean Squared Output Measure Values for Reinspection Policies

Reinspection Policy	Mean Squared Total Cost	Mean Squared Schedule Delay	Mean Squared Major Field Defects
2	16553.98829	0.000406444	0.999372584
3	16266.11611	0.068060961	0.907877918
4	13250.14333	0.631229775	0.395909255
5	12957.5875	0.949252522	0.327686451
7	16535.68954	0.014109704	0.979244876
8	16262.62936	0.072797704	0.894562357
9	13783.46999	0.374804148	0.475079851
10	13200.09919	0.632750072	0.378645702
11	13264.37093	0.843644667	0.370930444
12	15541.4098	0.391066742	0.713909187
13	16243.59498	0.150827481	0.885277765
14	16501.5413	0.028385777	0.962294369
15	16579.19716	0.001701556	0.999253023
16	13668.77735	0.445804225	0.440940814
17	14967.59866	0.191509555	0.646302376
18	15924.99456	0.075129556	0.830155846
19	16280.40206	0.02757237	0.917446754
20	16429.28299	0.008453556	0.95999179
21	13499.63813	0.407193996	0.521987491
22	14064.66165	0.392312081	0.530739257
23	15602.55936	0.254584224	0.734568235
24	16170.04461	0.159995186	0.859102305
25	16339.33067	0.074566666	0.91447733
26	13853.0968	0.287927852	0.538304437
27	14895.47595	0.146690888	0.660125252
28	15746.58048	0.081191557	0.80870162
29	16072.84262	0.047047111	0.886079341
30	16216.95876	0.030616667	0.926005861

Table 22 Mean Squared Output Measure Values for Reinspection Policies (cont.)

Reinspection Policy	**Mean Squared Total Cost**	**Mean Squared Schedule Delay**	**Mean Squared Major Field Defects**
33	16267.35959	0.067423407	0.907933325
34	13310.31405	0.525972813	0.419008823
35	13587.65251	0.440337405	0.525749845
38	16277.74862	0.058670444	0.896230282
39	13837.91261	0.273862815	0.513870811
40	13902.1977	0.239897333	0.570814742
41	13370.91196	0.721094892	0.414722364
42	14611.37483	0.246365411	0.619755189
43	15769.64245	0.082913333	0.823505989
44	16384.4794	0.018949778	0.946466224
45	16544.53541	0.001727852	0.993274977
46	16562.1141	0	1.004051691
47	13003.91995	1	0.322600074

Actually, the evaluation of the reinspection policies with respect to cost, schedule, and quality measures requires multi criteria analysis. However, since the data for three measures are in different scales, using directly the current values, while making the required comparisons, is not appropriate. For overcoming this obstacle, the related values are standardized, thus they can be incorporated to the formulations that find out the overall performance of a policy. This is accomplished by applying the following formula to each mean squared output measure value of each policy.

$$z_i = \frac{x_i - \overline{X}}{s_X} \qquad \text{(Eq. 46)}$$

where (the below definitions correspond to any of the three output measures),

z_i : Standardized mean squared output value of reinspection policy i

x_i : Mean squared output value of reinspection policy i

$\overline{X}$: Average of mean squared output values over all reinspection policies

s_X : Standard deviation of mean squared output values over all reinspection policies

$\overline{X}$ and s_X values calculated for each output measure are given in Table 23. Furthermore, standardized mean squared output values corresponding to each reinspection policy are available via Table 24.

Table 23 Average and Standard Deviation of Mean Squared Values for Three Output Measures

Output Measure	**Total Cost**	**Schedule Delay**	**Major Field Defects**
$\overline{X}$	15124.44501	0.255288988	0.717633093
s_X	1333.56839	0.275563435	0.22962187

Table 24 Standardized Mean Squared Output Measure Values for Reinspection Policies

Reinspection Policy	**Standardized Total Cost**	**Standardized Schedule Delay**	**Standardized Major Field Defects**
2	1.071968482	-0.924950527	1.226971501
3	0.856102405	-0.679437123	0.828513522
4	-1.405478487	1.364262233	-1.40110277
5	-1.624856681	2.518344043	-1.698212115
7	1.058246838	-0.875222379	1.139315617
8	0.853487801	-0.662247822	0.770524442
9	-1.00555399	0.433711967	-1.056315943
10	-1.443004971	1.369779283	-1.4762853
11	-1.394809661	2.135100687	-1.509885137
12	0.312668471	0.49272776	-0.016217558
13	0.839214533	-0.37908334	0.730090177
14	1.032640173	-0.823415529	1.065496395
15	1.090871804	-0.920250662	1.226450814
16	-1.091558311	0.691366172	-1.204990964
17	-0.117614029	-0.231451004	-0.310644265
18	0.600306333	-0.653785699	0.490034998
19	0.86681498	-0.826367324	0.870185666
20	0.978455994	-0.895748135	1.055468697
21	-1.218390366	0.551252411	-0.852033831
22	-0.794697419	0.497247004	-0.813920016
23	0.358522556	-0.002557539	0.073752304
24	0.784061481	-0.345814395	0.616096419
25	0.911003647	-0.655828387	0.857253869
26	-0.953343091	0.118444101	-0.780973764
27	-0.171696524	-0.394094741	-0.250445836
28	0.46651936	-0.631787129	0.396602147
29	0.711172831	-0.755694881	0.733581029

Table 24 Standardized Mean Squared Output Measure Values for Reinspection Policies (cont.)

Reinspection Policy	Standardized Total Cost	Standardized Schedule Delay	Standardized Major Field Defects
30	0.81924089	-0.81531979	0.907460459
33	0.857034848	-0.681750762	0.828754819
34	-1.360358399	0.982292243	-1.300504478
35	-1.152391217	0.671527474	-0.835648834
38	0.864825247	-0.713514637	0.777788233
39	-0.964729225	0.067403089	-0.88738186
40	-0.916523904	-0.055855218	-0.639391844
41	-1.314917973	1.690376317	-1.319171946
42	-0.38473481	-0.032383025	-0.426256891
43	0.483812791	-0.625538929	0.461074964
44	0.944859224	-0.857658095	0.996565049
45	1.064880068	-0.920155234	1.200416509
46	1.078061767	-0.926425485	1.247348943
47	-1.590113468	2.702503011	-1.720363218

Different reinspection policies show varying results with respect to three output measures, namely total cost, schedule delay, and major field defects. For each of these measures, the lower values are better, i.e. the aim should be to minimize each of them. However, decreasing the values for total cost and major field defects requires increasing the number of reinspections performed, thus the schedule delay measure. Hence, for being able to compare reinspection policies, the three values are combined into one value, the minimization of which will represent the concurrent minimization of three measures. In order to accomplish this, weights are assigned to each of the output measures. These weights are then multiplied with the standardized mean squared value of an output measure. Finally, the summation of weighted standardized values is carried out through all measures for obtaining the aggregate standardized mean squared value. The corresponding formulation is provided below.

$$z_i^A = w_C z_i^C + w_S z_i^S + w_D z_i^D \qquad \text{(Eq. 47)}$$

where,

z_i^A: Aggregate standardized mean squared value of reinspection policy i.

w_C, w_S, w_D: The weights corresponding to total cost, schedule delay, and major field defects measures, respectively.

z_i^C, z_i^S, z_i^D: Standardized mean squared value of reinspection policy i for total cost, schedule delay, and major field defects measures, respectively.

Actually, the weights in the above formulation correspond to the different preferences regarding three output measures, thus also for cost, schedule, and quality perspectives underlying the inspected code and the software project. Briefly, the weight of an output measure implies the importance assigned to the corresponding perspective. As mentioned before in the text, these preferences are shaped according to organizational policies, project structure, software type, etc. Therefore, by varying the weights of each output measure, it is possible to reveal the performance of reinspection policies for different preference profiles. Hence, by this way, the study enables the exposing of the suitable reinspection policy for a particular preference profile. For determining the profiles to be considered in this study, the following profile patterns are identified for three output measures.

- Each measure is equally important.
- One of the measures is extremely important when compared to others.
- The importance of the three measures can be sorted as high, medium, and low.
- One of the measures is more important than two other equally important measures.
- One of the measures is less important than two other equally important measures.

By distributing a total weight of 1 to three measures according to these patterns, 16 different preference profiles depicted in Table 25 are designated for this study. The software organizations producing space shuttle software can be given as an example to profile 8 which represents a software development environment which can not bear even a single defect. Further, profile 5 illustrates an organization for which time-to-market is critical, thus no risk for late project completion is acceptable. Profile 2 corresponds to a software organization where the upper management is strictly against exceeding of project budget. Finally, a software development project that is executed according to a contract which comprises penalties for late completion and for user encountered defects is a good example for profile 7.

Table 25 Preference (Weight) Profiles for Output Measures Considered in the Study

Profile	$\mathbf{w_C}$	$\mathbf{w_S}$	$\mathbf{w_D}$
1	0.333	0.333	0.333
2	0.8	0.1	0.1
3	0.6	0.3	0.1
4	0.6	0.1	0.3
5	0.1	0.8	0.1
6	0.3	0.6	0.1
7	0.1	0.6	0.3
8	0.1	0.1	0.8
9	0.3	0.1	0.6
10	0.1	0.3	0.6
11	0.4	0.4	0.2
12	0.4	0.2	0.4
13	0.2	0.4	0.4
14	0.5	0.25	0.25
15	0.25	0.5	0.25
16	0.25	0.25	0.5

For each of the above profiles the reinspection policies' performance is revealed by computing aggregate standardized mean squared value. Detailed results corresponding to these calculations are available in Appendix E. Furthermore, the ranking of each policy in the context of each preference profile is given in Table 26. In this table a rank of 1 corresponds to the policy with lowest aggregate standardized value (z_i^A). When the ranks provided in the table are examined, policy 39 is seen as the best policy in the case of equal preference with respect to cost, schedule, and quality, which, in practice, can be deemed as the default profile. Further, if a policy's rank summation through 16 profiles is put forward as an indicator which shows the policy's overall performance, again policy 39 is the most suitable policy since its 'sum of ranks' value is the lowest. A more detailed discussion of the results regarding reinspection policy rankings is provided in Section 4.1.

Table 26 Ranks of Reinspection Policies for each Preference Profile

Profile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Sum of Ranks**
Policy																	
2	39	39	39	39	2	18	19	39	39	39	37	39	39	39	30	39	535
3	30	30	30	30	17	25	24	29	29	30	30	30	29	31	26	31	451
4	9	4	8	4	37	36	36	5	4	6	10	3	10	3	17	3	195
5	15	1	13	1	40	40	40	1	1	10	18	4	18	11	40	7	260
7	37	37	38	37	8	21	20	37	37	37	36	37	37	37	29	37	522
8	28	27	28	26	18	26	18	27	27	27	27	27	27	28	24	27	412
9	3	11	7	10	29	13	11	9	9	5	4	7	4	7	5	5	139
10	7	3	6	3	36	35	35	3	3	3	9	1	8	2	16	1	171
11	16	6	14	6	39	39	39	4	5	14	16	8	16	14	39	11	286
12	24	20	24	19	34	37	37	19	19	23	40	20	31	23	38	21	429
13	35	32	35	29	23	33	34	26	26	32	41	32	36	34	37	32	517
14	36	36	36	36	11	23	25	36	36	36	33	36	35	36	31	36	518
15	40	41	41	41	5	20	21	40	40	40	39	40	40	41	34	40	563
16	6	10	10	8	32	28	26	8	8	1	7	6	5	6	7	4	172
17	17	18	18	18	21	5	5	17	17	18	14	18	14	18	8	18	244
18	22	23	21	23	13	10	9	23	23	22	20	23	20	22	15	23	312
19	25	28	26	32	7	12	12	31	31	28	25	29	24	26	19	28	383
20	34	35	32	35	1	14	14	35	35	35	31	35	32	35	23	35	461
21	8	8	2	9	30	16	27	11	10	9	5	10	7	5	6	9	172
22	11	15	12	15	31	30	29	14	14	11	11	15	12	15	12	15	262
23	21	19	22	20	28	34	32	20	20	19	23	19	22	20	28	19	366
24	31	25	33	25	24	32	31	24	24	26	35	25	34	29	35	25	458
25	33	33	34	33	20	27	28	32	33	33	32	33	33	32	32	33	501
26	4	13	5	13	27	4	4	13	13	7	3	12	3	8	3	8	140
27	14	17	17	17	19	3	1	18	18	17	12	17	11	17	4	17	219
28	19	21	19	21	12	6	6	21	21	20	17	21	17	19	13	20	273
29	23	24	23	24	10	9	10	25	25	24	22	24	23	24	18	24	332
30	26	26	25	28	9	11	13	33	32	31	24	28	26	25	20	29	386
33	29	31	29	31	16	24	22	30	30	29	29	31	28	30	25	30	444

Table 26 Ranks of Reinspection Policies for each Preference Profile (cont.)

Profile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Sum of Ranks**
Policy																	
34	2	5	1	5	35	31	33	6	6	2	6	2	6	1	10	2	153
35	10	9	9	11	33	29	30	12	12	13	8	13	9	10	11	14	233
38	27	29	27	27	14	22	16	28	28	25	26	26	25	27	21	26	394
39	1	12	3	12	25	2	2	10	11	4	1	9	1	4	1	6	**104**
40	5	14	4	14	22	1	3	15	15	8	2	14	2	9	2	13	143
41	12	7	11	7	38	38	38	7	7	12	15	11	15	12	36	12	278
42	13	16	16	16	26	8	8	16	16	16	13	16	13	16	9	16	234
43	20	22	20	22	15	7	7	22	22	21	19	22	19	21	14	22	295
44	32	34	31	34	6	15	15	34	34	34	28	34	30	33	22	34	450
45	38	38	37	38	3	17	17	38	38	38	34	38	38	38	27	38	515
46	41	40	40	40	4	19	23	41	41	41	38	41	41	40	33	41	564
47	18	2	15	2	41	41	41	2	2	15	21	5	21	13	41	10	290

3.7. Studying the Effects of the Factors on Output Measures

In Section 3.2, seven factors, varying levels of which are considered throughout the study, are identified. Then these are employed to develop an experiment design which consists of 54 treatments. Certainly, as the levels corresponding to these factors are altered, the values of the three output measures are influenced. Because these measures indicate the outcomes encountered at the end of the software development life-cycle, which depends on the activities and parameters throughout the life-cycle. In this section, the results of Analysis of Variance (ANOVA) conducted for each output measure is reported, in order to acquire knowledge about the effects of different factor levels on total cost, schedule delay, and major field defects. While conducting the ANOVA study, the experiment described in Section 3.2 is employed. Actually, as it should be evident from the previous explanations, the simulation study corresponding to each reinspection policy is performed for all 54 treatments of the experiment. The ANOVA is carried out for the treatments related to the reinspection policy 39, since it is the most preferable policy when all designated preference profiles are considered. For each ANOVA, the main factor and two factor interaction effects are included in the analysis. Besides, 95% is employed as the significance level of the tests. Subsequent paragraphs summarize these analyses. Whilst, Appendix F includes normal probability and residual plots for each ANOVA study, along with the Tukey comparisons of factor levels.

For Total Cost measure, the concluded ANOVA table is given in Table 27. Furthermore, Figure 13 and Figure 14 provide corresponding Main Effects Plot and Interaction Plots for significant interactions, respectively. From these exhibitions, it is seen that the total cost faced in the context of the study is significantly affected by all of the seven factors except number of reinspectors. More specifically, if the probability regarding a defect's being major or difficult is increased, the total cost also increases. Whilst the remaining four factors result in a decrease in total cost as their levels are increased. Furthermore, three significant two-factor interactions (depicted in Figure 14) obviously show that the detection capability of testing activity determines how most of other factors influence the total cost measure. For instance, having one more inspector when the testing detection probability is high, lowers the cost value much more than the decrease gained when testing detection probability is low.

Table 27 Total Cost Measure's ANOVA Table for Policy 39

Source	**DF**	**Seq SS**	**Adj SS**	**Adj MS**	**F**	**p**
Testing detection probability of a defect (p_t)	1	9586.3	9586.3	9586.3	239.09	0.00
Probability that a given defect is difficult (p_d)	2	44386.9	21440	10720	267.37	0.00
Probability that a given defect is major (p_m)	2	2423.7	3273.8	1636.9	40.83	0.00
Inspector capability of a particular inspector (q_j)	2	10367	10367	5183.5	129.28	0.00
Inspection detection probabilities of difficult and easy defects (p1, p2)	2	22902.3	15421.0	7710.5	192.31	0.00
Number of inspectors (k)	2	8947	8947	4473.5	111.58	0.00
p_t*p_d	2	8002.3	8002.3	4001.2	99.79	0.00
p_t*(p1, p2)	2	3308.8	3308.8	1654.4	41.26	0.00
p_t*k	2	5376.5	5376.5	2688.2	67.05	0.00
Error	36	1443.4	1443.4	40.1		
Total	53	116744.2				

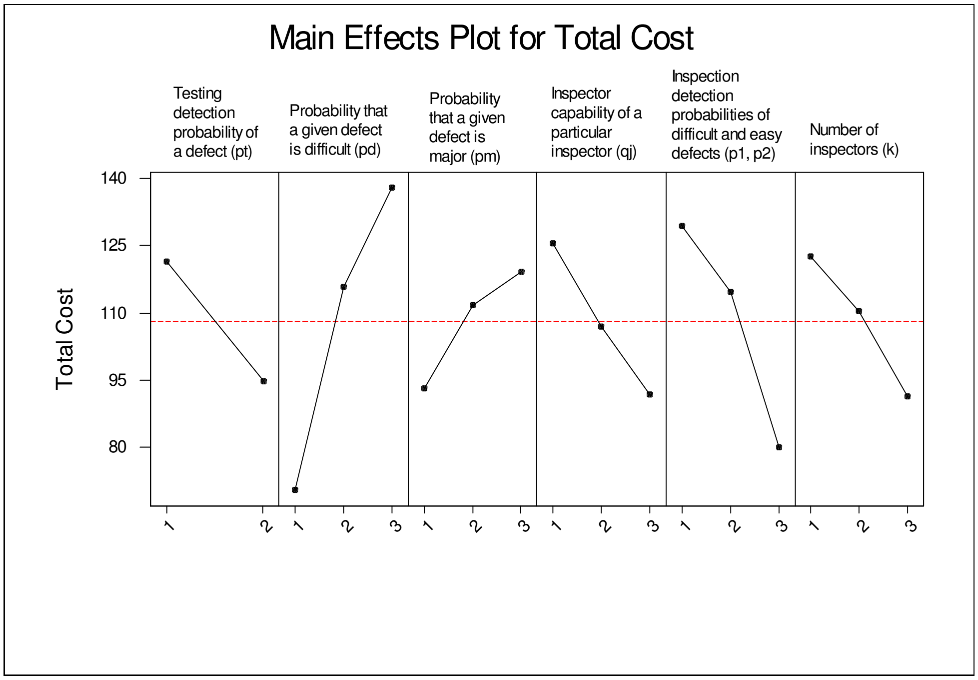

Figure 13 Total Cost Measure’s Main Effects Plot for Policy 39

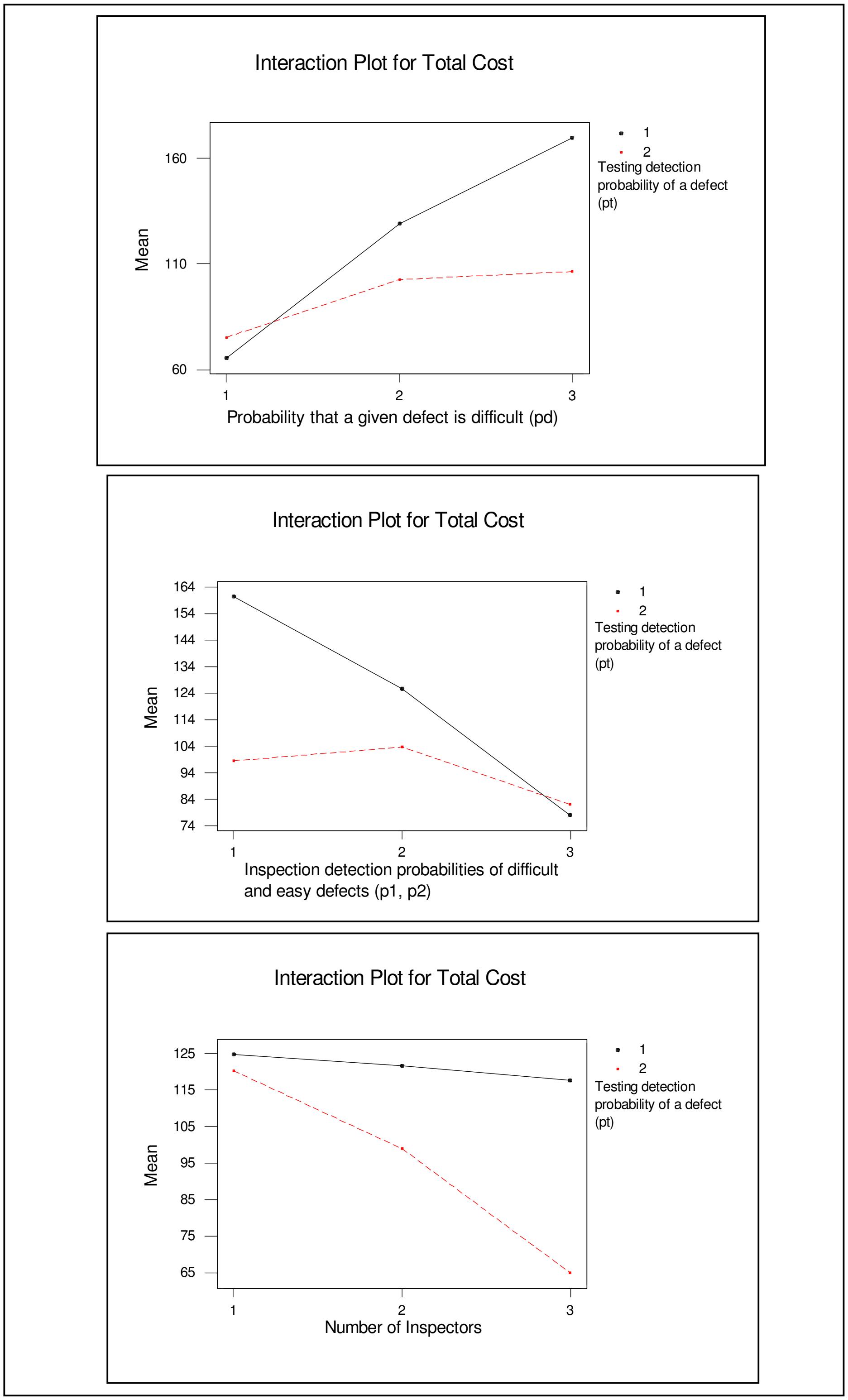

Figure 14 Total Cost Measure's Interaction Plots for Policy 39

The Schedule Delay measure's final ANOVA table, and Main Effects Plot are depicted in Table 28, and Figure 15, respectively. Note that, ANOVA results in no significant two-factor interactions. Furthermore, two factors, namely testing detection probability of a defect and probability that a given defect is difficult, do not have any effect on schedule delay. Regarding the remaining five factors, increasing the levels for probability that a given defect is major, and inspection detection probabilities cause an increase in schedule delay measure. Whilst, high values of inspector number and capability exhibit a decreasing effect. Although the effect of number of reinspectors factor is not same for identical for its different levels, nevertheless it can be concluded that it has a non-decreasing effect on schedule delay measure as number of reinspectors is increased.

Table 28 Schedule Delay Measure's ANOVA Table for Policy 39

Source	**DF**	**Seq SS**	**Adj SS**	**Adj MS**	**F**	**p**
Probability that a given defect is major (p_m)	2	0.42059	0.42059	0.2103	26.77	0.000
Inspector capability of a particular inspector (q_j)	2	0.09003	0.09003	0.04502	5.73	0.006
Inspection detection probabilities of difficult and easy defects (p1, p2)	2	0.37166	0.37166	0.18583	23.66	0.000
Number of inspectors (k)	2	1.02692	1.02692	0.51346	65.37	0.000
Number of reinspectors (k')	2	0.19203	0.19203	0.09601	12.22	0.000
Error	43	0.33775	0.33775	0.00785		
Total	53	2.43898				

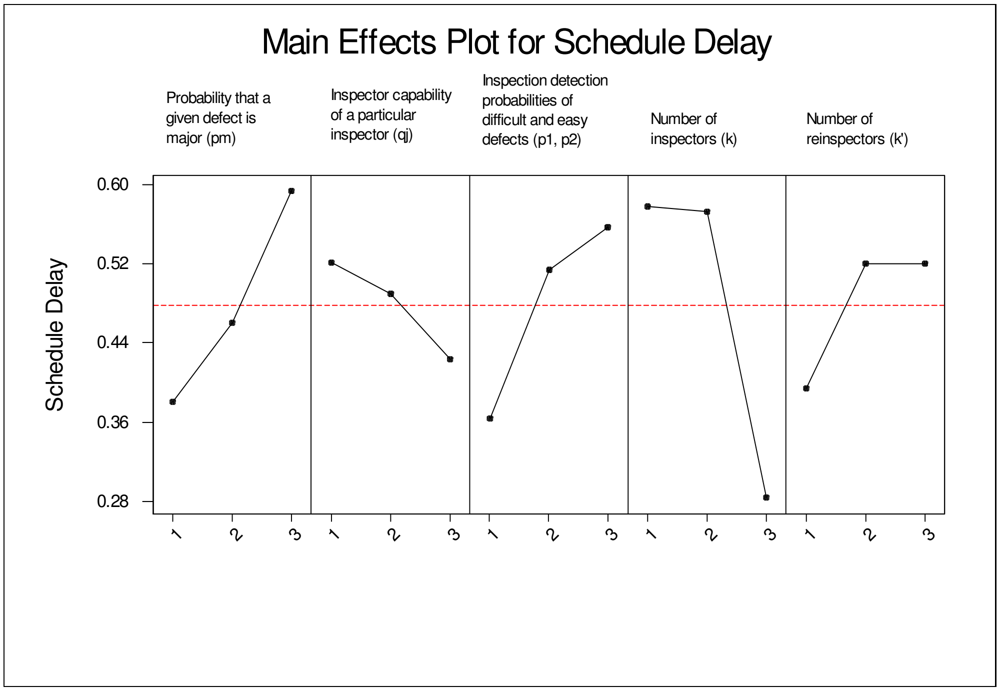

Figure 15 Schedule Delay Measure's Main Effects Plot for Policy 39

The Major Field Defects measure's final ANOVA table, Main Effects Plot, and Interaction Plots are depicted in Table 29, Figure 16, and Figure 17, respectively. These are obtained after applying square root transformation on 54 major field defects values. A brief examination of these results reveals that all seven factors affect the number of major defects found by the users. Among these, testing detection probability, inspector capability, inspection detection probabilities, number of inspectors, and number of reinspectors, on the average, decrease the major field defects as their corresponding levels are increased. Whilst, if a defect's being difficult or major probability is increased, a higher value of major field defects is observed. Furthermore, five different two-factor interactions seem significant for major field defects measure. For example, at the third highest level of probability that a given defect is difficult, major field defects increase if the testing detection probability is at its low level, and the major field defects decrease otherwise.

Additionally, when the ANOVAs for total cost and major field defects are examined together, it is seen that they result in exactly same significant factors and two-factor interactions.

Table 29 Major Field Defects Measure's ANOVA Table for Policy 39

Source	**DF**	**Seq SS**	**Adj SS**	**Adj MS**	**F**	**P**
Testing detection probability of a defect (p_t)	1	0.46622	0.46622	0.46622	838.18	0.00
Probability that a given defect is difficult (p_d)	2	0.8625	0.43569	0.21785	391.65	0.00
Probability that a given defect is major (p_m)	2	1.13449	0.70673	0.35336	635.29	0.00
Inspector capability of a particular inspector (q_j)	2	0.15687	0.15687	0.07843	141.01	0.00
Inspection detection probabilities of difficult and easy defects (p1, p2)	2	0.59992	0.21321	0.1066	191.66	0.00
Number of inspectors (k)	2	0.12503	0.12503	0.06251	112.39	0.00
Number of reinspectors (k')	2	0.03224	0.17043	0.08522	153.2	0.00
p_t*p_d	2	0.20614	0.20614	0.10307	185.3	0.00
p_t*q_j	2	0.00871	0.00871	0.00435	7.83	0.002
p_t*(p1, p2)	2	0.02381	0.02381	0.0119	21.4	0.000
p_t*k	2	0.07857	0.07857	0.03929	70.63	0.000
q_j *(p1, p2)	4	0.01936	0.01936	0.00484	8.7	0.000
Error	28	0.01557	0.01557	0.00056		
Total	53	3.72941				

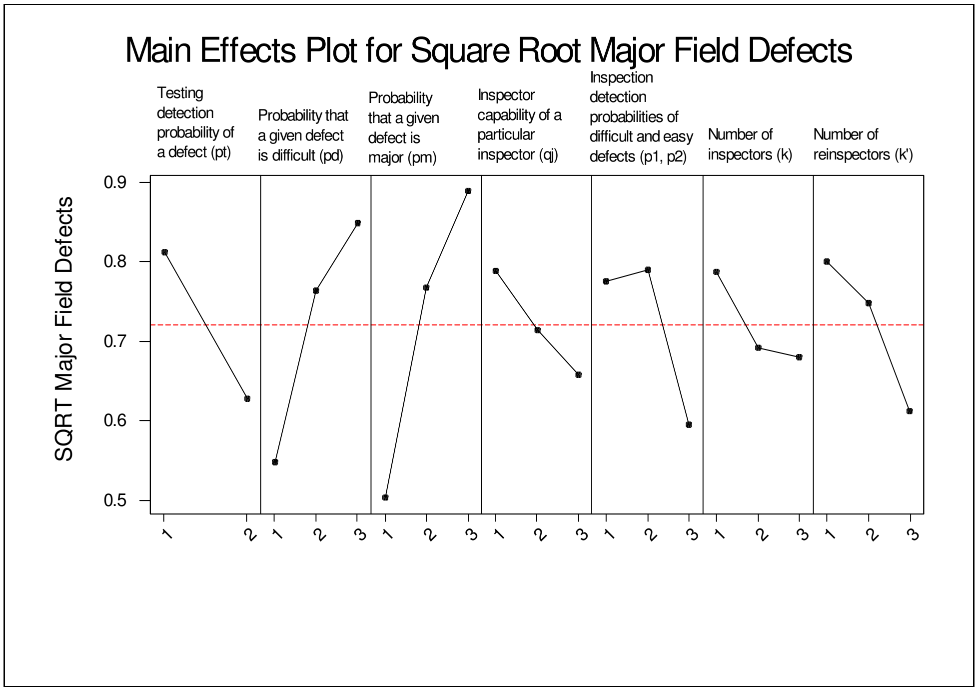

Figure 16 Major Field Defects Measure's Main Effects Plot for Policy 39

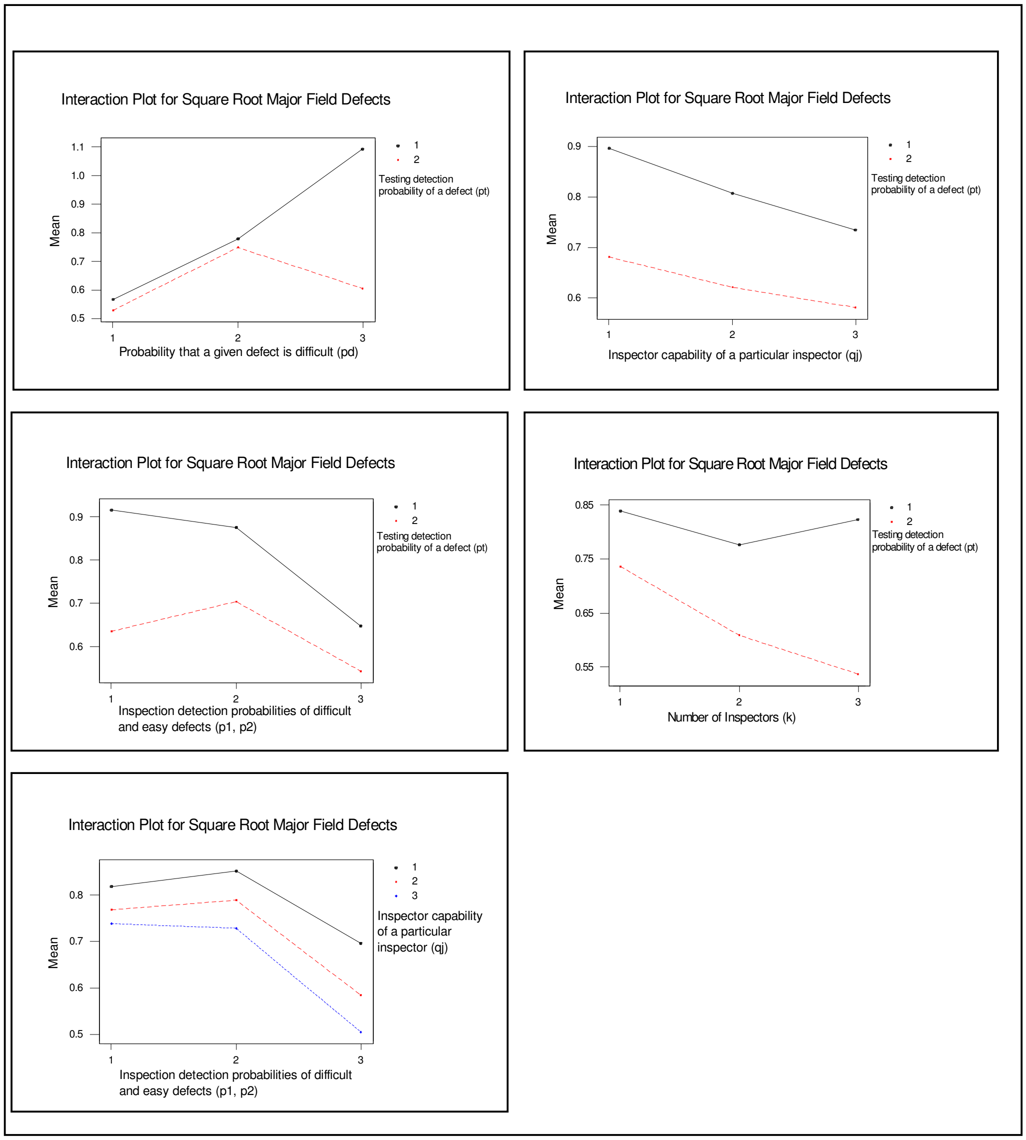

Figure 17 Major Field Defects Measure's Interaction Plots for Policy 39

CHAPTER 4

DISCUSSION

4.1. Analyzing the Results Regarding Suitable Policies for Different Preference Profiles

The rankings given in Table 26 reveal many important inferences about reinspection policies. First, 8 reinspection policies become the first policy in at least one of the preference profiles. The common property of these 8 policies is their aggressive threshold values, i.e they put forward effectiveness or defect density values that are difficult to attain as thresholds. For instance, the most preferable policy, namely policy 39, requires the removal of 75% of the defects after reinspection is performed. The only exception to this finding is policy 20, which is ranked first for preference profile 5. Actually, this is a normal outcome, since profile 5 favors the policies that cause low schedule delay values without giving much importance to cost and defectiveness perspectives. Consequently, with a reinspection percentage of 4.6%, policy 20 outperforms other policies with low schedule delay (namely 2, 7, 15, 19, 44, 45, and 46) by resulting acceptable values also for total cost and major field defects measures. Another important finding through Table 26 is; the reinspection policies based on inspection effectiveness metric are generally more successful in satisfying the expectations put forward by different preference profiles. Since, except for three of preference profiles (namely, profiles 5, 7, and 10), the superior policies are the ones which are generated from reinspection decision methods I and IV. Actually, for three profiles that favor reinspection policies related to other decision methods, the second best alternatives (policies) are again based on inspection effectiveness, and they are outperformed due to very small differences in aggregate standardized mean squared value. Besides, the study results emphasize none of the policies underpinned by net benefit criterion, i.e. reinspection decision method V. Thus, the estimated effectiveness after initial or second inspection should be employed while making the reinspection decision. Additionally, the majority of the superior policies employ the information regarding major defects while concluding the reinspection decision. This is due to important cost and quality related consequences of major defects.

As explained in the previous sections, 'Never Reinspect' is the default reinspection decision in most of the software development environments. Actually, in most of the cases, the code is directly passed to testing, even without considering the 'reinspect' alternative. Hence, policy 46 deserves special attention, since it always suggests continuing to next phase without conducting a reinspection. The ranks corresponding to policy 46 reveal that this policy may be beneficial only if the schedule performance of the project is deemed as extremely critical. Consequently, always skipping reinspection, is not a wise practice in general. This is also evident from the high total cost and major field defects values encountered when policy 46 is applied. Also, policy 47 that corresponds to 'Always Reinspect' strategy is ranked first for none of preference profiles, although it is in second place for some of the profiles which favor measures regarding cost and number of defects (namely, profiles 2, 4, 8, and 9). In all these cases, policy 47 is outperformed by policy 5, which results in a slightly less number of reinspections. Hence, it is possible to obtain same cost and defect performance by relaxing 'Always Reinspect' strategy. Further, it is concluded that applying a straightforward policy that makes the conduct of the second inspection a common practice is not the solution, since policy 47 never succeeds to become the first policy. When the performance of the policies that represent the policies excluded from the analysis (since they are equivalent) is examined, it is seen that none of them (namely policy 2, 7, and 46) is superior in any of the preference profiles. So, this shows no necessity for dwelling upon the performance of removed policies.

It should be clear from the discussions up to this point; this study shows that the appropriate reinspection policy to be employed for reinspection decision making, differs according to the preference structure of the software organization/project regarding cost, schedule, and quality. The questionnaire given in Table 30 is constituted for providing guidance to practitioners on selecting the policy that best fits to the organizational needs. Further, this questionnaire also summarizes the results obtained by evaluating various reinspection policies under different preference profiles. In the questionnaire, the questions are designed according to importance assigned to three perspectives, namely, cost, schedule, and quality, which correspond to the output measures taken into account throughout this study, i.e. total cost, schedule delay, and major field defects, respectively. Also, note that, for simplicity reasons regarding the usage of the questionnaire, some of the preference profiles' the second or third best reinspection policy included in the questionnaire, since its aggregate standardized value is very close to the one corresponding to the best policy.

Table 30 Questionnaire for Determining the Appropriate Reinspection Policy

No	**Question**	**Answer**	**Action**	**Related Profile(s)**
1	Are all perspectives equally important?	Yes	Apply Policy 39	1
		No	Answer Question 2	-
2	Are there any two criteria that are equally important?	Yes	Answer Question 3	-
		No	Answer Question 4	-
3	Are there any perspective that is important than schedule?	Yes	Apply Policy 10	12,14,16
		No	Apply Policy 39	11,13,15
4	Is schedule perspective is extremely important?	Yes	Apply Policy 20	5
		No	Answer Question 5	-
5	Is the importance given to schedule perspective low, while the cost or defect perspectives are assigned the highest priority?	Yes	Apply Policy 5	2,4,8,9
		No	Answer Question 6	-
6	Is the importance of defect perspective is high?	Yes	Apply Policy 16	10
		No	Apply Policy 39	3,6,7

Policy 39, which requires the estimated inspection effectiveness after reinspection to be at least 75%, is the best reinspection policy with respect to total rank value. Further, it is the policy with highest rank if the default weights are employed for three perspectives. In addition to this, the questionnaire above proposes to employ this policy for 7 out of 16 preference profiles. Consequently, policy 39 can be deemed as the policy which is overall acceptable and performs satisfactorily in most of the considered situations. So, it is concluded that while making the reinspection decision, one should examine the percentage of major defects captured when the reinspection is completed. Further, this percentage level should be moderately high as 75%, i.e. increasing the related threshold to a higher value as in the case of policy 40 –with 90% threshold- does not result in better performance in terms of aggregate standardized mean squared value. Recall that, since policy 39 is based on reinspection decision method IV, it does not propose reinspection, if the required effectiveness determined by the threshold will not be attained when reinspection is carried out. This implies that a reinspection should be performed if it provides the related benefits with respect to effectiveness objective. Finally, superiority of policy 39 implies that the reinspection decision should be based on the data regarding major defects. Actually, this is an expected outcome, since generally major defects are deemed as more critical and more costly than minor defects.

As a result, this study suggests using policy 39 if one policy that serves well for majority of software organizations is requested. In order to scrutinize the effects of using policy 39 when actually another policy is more appropriate, a sensitivity analysis is performed. For the preference profiles where other policies are ranked first, this analysis reveals the gain or loss with respect to three output measures, if policy 39 is applied instead of the corresponding most appropriate policy. The results of this analysis are depicted in Table 31. In this table, the positive difference percentage values correspond to situations where the employment of policy 39 causes to a loss with respect to value of the best policy. By selecting the 20% as the threshold value above which the related deviation with respect to corresponding best value is not acceptable, the following findings can be reached.

- If the schedule perspective's importance is high (especially if it is extremely important), policy 39 should not be applied (See profiles 5, 6, and 7). Because in this case the schedule delay value is affected adversely. So, the policy ranked first should be followed as usual.

- When policy 5 is the most appropriate policy, replacing it with policy 39 should be avoided, since this causes a high increase in the major field defects.
- Regarding total cost, the usage of policy 39 never results a significant deviation vis-à-vis the original policy. However, since this observation is due to close mean total cost values of different policies, it can be more suitable to put forward a scalar threshold. For instance, if expending 3 hours more is deemed critical for an organization, policy 39 should not be an alternative for policy 5.

Furthermore, if the specific values of the experiment factors are known for a software development environment for which the results of this study will be used, the output measure values of reinspection decision policies may be employed by just considering the treatment corresponding to these specific factor levels. Thus, by this way, the best policy can be identified for the subject environment.

Table 31 Sensitivity Analysis Regarding the Usage of Policy 39

	Preference Profile	2	3	4	5
Output Measure	**Policy Ranked First**	**5**	**34**	**5**	**20**
Mean Total Cost	**Policy 39**	108.0555	108.0555	108.0555	108.0555
	Other Policy	104.7544	106.0156	104.7544	116.3695
	Difference Percentage (%)	3.151266	1.924087	3.151266	-7.14449
Mean Schedule Delay	**Policy 39**	0.478222	0.478222	0.478222	0.478222
	Other Policy	0.970407	0.676111	0.970407	0.046556
	Difference Percentage (%)	-50.7194	-29.2687	-50.7194	927.2076
Mean Major Field Defects	**Policy 39**	0.587556	0.587556	0.587556	0.587556
	Other Policy	0.443222	0.523	0.443222	0.791593
	Difference Percentage (%)	32.56455	12.34332	32.56455	-25.7755
	Preference Profile	**6**	**7**	**8**	**9**
Output Measure	**Policy Ranked First**	**40**	**27**	**5**	**5**
Mean Total Cost	**Policy 39**	108.0555	108.0555	108.0555	108.0555
	Other Policy	108.509	111.7367	104.7544	104.7544
	Difference Percentage (%)	-0.41798	-3.29459	3.151266	3.151266
Mean Schedule Delay	**Policy 39**	0.478222	0.478222	0.478222	0.478222
	Other Policy	0.376741	0.354407	0.970407	0.970407
	Difference Percentage (%)	26.93669	34.93573	-50.7194	-50.7194
Mean Major Field Defects	**Policy 39**	0.587556	0.587556	0.587556	0.587556
	Other Policy	0.615	0.65863	0.443222	0.443222
	Difference Percentage (%)	-4.46251	-10.7912	32.56455	32.56455

Table 31 Sensitivity Analysis Regarding the Usage of Policy 39 (cont.)

	Preference Profile	10	12	14	16
Output Measure	**Policy Ranked First**	**16**	**10**	**34**	**10**
Mean Total Cost	**Policy 39**	108.0555	108.0555	108.0555	108.0555
	Other Policy	107.7359	105.7883	106.0156	105.7883
	Difference Percentage (%)	0.296619	2.143113	1.924087	2.143113
Mean Schedule Delay	**Policy 39**	0.478222	0.478222	0.478222	0.478222
	Other Policy	0.633963	0.776519	0.676111	0.776519
	Difference Percentage (%)	-24.5662	-38.4146	-29.2687	-38.4146
Mean Major Field Defects	**Policy 39**	0.587556	0.587556	0.587556	0.587556
	Other Policy	0.539074	0.49137	0.523	0.49137
	Difference Percentage (%)	8.993473	19.57488	12.34332	19.57488

4.2. Reviewing the Results of Tolerance Analysis

Results of the tolerance analysis given in Section 3.7, provide valuable insights about how a software organization can improve its cost, schedule, and quality performance. The following are the suggestions revealed by this study regarding the most suitable factor levels and related actions that can be taken by the organization.

- The effectiveness of testing activity should be enhanced as much as possible, since as the capability of identifying the defects during testing is increased, the number of major defects propagating to the field decreases along with the total cost incurred for defect correction. This can be accomplished by employing more advanced testing tools, more talented testers, more sophisticated testing techniques, and managing the testing process quantitatively. However, the schedule performance (as it understood in this study) does not get better, because the testing activity's parameters seem irrelevant to the number of reinspections performed.
- Improving the capability of the inspectors should be one of the main goals of the organization, since this influences all of the performance indicators positively, i.e. as more capable inspectors participate in inspections and reinspections, significant reductions are observed in cost, schedule, and quality related measures. Using more experienced employees as inspectors, deploying intelligent tools that help inspectors while scrutinizing the software code, improving the quality of the inspection training, and lessons learned meetings about inspection, can be listed among the possible initiatives for improving the inspector capability.
- Although adding more inspectors means increasing the labor costs devoted to the inspection activity, this in general enables better performance in terms of cost, schedule, and quality. For example, in this study, the total cost, schedule delay, and major field defects measures are brought to their lowest values (when number of inspectors factor considered alone) by assigning 4 inspectors. If the number of inspectors is increased, more defects are found during the inspection, thus (i) more correction cost is saved regarding later phases, (ii) less defects pass to the next phase, hence also to the user, and (iii) the selected reinspection policy less frequently proposes the conduct of the reinspection.
- When the detection probabilities of difficult and easy defects are about 0.5 and close to each other, the cost and quality performances improve, whilst schedule

performance worsens. An organization may decrease the variance of the coding activity (that is where the defects are injected) in order to have uniform defect detection probabilities during inspection, if the cost and defect aspects are deemed critical.

- As the intensity of the defects that are difficult to find during inspection increases, the total cost and the number of major defects reaching to the users also increase. So, the number of difficult defects should be as low as possible for having better cost and quality performance. However, no evident action is available for altering the difficulty level of a defect, since this is related to the defect injection. Nevertheless, performing cause analysis meetings after the defects are captured in the later phases, in order to develop inspection mechanisms that will enable the removal of more difficult defects may be a solution.
- Similar to difficulty aspect of the defects, as the number of major defects gets higher, cost, schedule, and quality indicators are exposed to adverse effects. Consequently, the software project is in better position as less major defects are generated during coding activity. However, unfortunately, one can not manipulate the severity level of a defect, since it depends on the circumstances where the defect is injected.
- If the number of reinspectors factor is increased, this influences major field defects positively, but schedule delay measure negatively, i.e. the number of reinspections increases as the number of reinspectors is changed from 2 to 3. Furthermore, no effect for total cost measure is observed regarding number of reinspectors. Consequently a trade-off situation occurs between schedule delay and major field defects measures, because as more resources are allocated to the reinspection, the estimated number of defects to be identified during reinspection increases, which in turn enhances the quality related benefits, but at the same time increases the number of reinspections that is performed (which means an increase in schedule delay measure). So, if the importance of the schedule perspective is low when compared to quality perspective, the number of reinspectors should be retained at its highest level as long as the timing and budget constraints of the project are not violated.

4.3. Validity of the Study Results

Since the simulation techniques are utilized throughout the study the main threat to the validity comes from the ability of the simulation model and the related factors to represent the environment encountered in real software development projects. In order to avoid the related shortcomings that may expose due to the discrepancy between the study context and the real life, the general approach of determining each component in the study by employing underlying studies and results in the literature is followed. In line with this approach, the undertakings described in the subsequent paragraphs are performed in the study.

The study considers the environment of a typical SW-CMM Maturity Level 3 software development project. Maturity Level 3 of SW-CMM Model requires the conduct of software inspections according to a standard procedure which is constituted with respect to the rules put forward by Peer Review Key Process Area of the model. So, by employing SW-CMM Maturity Level 3 context, the variability among different organizations and projects while conducting software inspections is minimized. The study results are not valid for an organization which has higher or lower maturity than SW-CMM Maturity Level 3. Actually, this study can not be replicated with the aim of making the same analysis for lower maturity organizations, since these organizations' processes are not standardized (i.e. different projects may execute according to different procedures). For higher maturity organizations, again the coding, inspection, testing practices generally differ, since quantitative project management and continuous improvement paradigms are in place. This results in varying degree for software development (high) technology and managing the project according to quantitative analysis among the organizations. So, for such organizations (lower or higher maturity) the study should be repeated by taking into account the specifics regarding defect numbers, rework costs, inspector capability.

While putting forward the reinspection decision policies, all the different objective reinspection decision methods available in the literature are utilized. For observing the results obtained from different threshold levels, each reinspection decision method is adapted several times by varying the related threshold values. Also, the related reinspection decision methods are duplicated for major defects. These make the study confident with respect to the coverage of possible reinspection policies.

The COQUALMO model, which provides the underlying information to predict the number of defects injected into and removed from the software for a software project, is utilized for selecting the various values considered in the study. First, contrary to some simulation studies in the literature where the number of defects present in the inspected artifact is set arbitrarily (See for example, (El Amam and Laitenberger, 2000)), this study determines the defect content of the software code by employing the Defect Introduction submodel of COQUALMO. Further, the inspector's and testing activity's defect detection probabilities are designated according to COQUALMO Defect Removal submodel. While adapting COQUALMO's defect removal submodel the ratings regarding various defect detection activities are selected by taking into account the circumstances of a SW-CMM Level 3 software development project. However, the ratings for 21 factors of Defect Introduction submodel are assumed at their nominal level, since these rating values are dependent to the specific properties and conditions of a project which can not be defined in the SW-CMM model (except Process Maturity factor, whose nominal value corresponds to maturity level 3 of SW-CMM). So, because of this reason a typical project (i.e. a project with nominal rankings) is the subject of the study. In order to increase the reliability of the results the study may be replicated by using the specific ratings corresponding to the project for which the appropriate reinspection policy is desired to be identified. In order to cover the different conditions that may be faced during the software development life cycle, the study takes into account all the relevant factors that affect the number and content of the defects flowing from coding activity to the field use. Namely, seven factors are designated for the purposes of this study. However, since the decision that is dwelled upon in this study is related to the software inspection, the issues regarding this activity are included with more detail with respect to testing. Consequently, the testing activity is treated as a black-box which results in either the detection of a defect or allows the defects to pass to the user. Further, for determining these conditions an experiment is designed by changing the levels of the seven selected factors. The levels of the factors are determined by either using the related guidance from the literature (such as COQUALMO, the information regarding number of inspectors for code inspections) or selecting a representative set of levels (for instance, for the probability of generating a difficult defect the values corresponding to high, medium, and low probabilities are designated). Another parameter, which needs attention regarding the validity, is size of the inspected code. Because, the defect introduction and removal information supplied by COQUALMO is given for 1000 SLOC of code, in this study the inspected code is assumed to be 1000 SLOC also. This assumption seems reasonable, since there is no guidance in the literature for the typical code size that is subject to inspection. Actually, in practice generally the software unit, whose scope is determined during software design phase, is the inspected artifact. So, the size of the unit may vary greatly. Consequently, future work may address the replication of study with different code sizes in order to reveal the effects regarding changing code size.

The study embraced all the typical activities performed in a software project after the completion of coding phase. By this way, the effects of the reinspection decision on the end-project performance indicators are revealed. Namely, the three aspects; schedule, cost, and quality are examined for various reinspection policies. These are the main issue for whom quantitative objectives are designated and against which software project's actual performance is monitored. Additionally, 16 different preference profiles, which are believed to represent all priority structures (for cost, quality, and schedule) that may be encountered in real software projects, are comprised in the study, thus the outcomes of applying different reinspection policies can be observed for various software development project types with respect to varying degree of importance assigned to three perspectives.

As mentioned previously in the text, all evaluated reinspection policies depend on the estimation of defect number in the inspected artifact. Hence, the selection of the DCET to be used in the study is important for having accurate estimates. With the aim of identifying the appropriate DCET, the DCETs that are available for use in the context of a simulation study are evaluated. This enables the minimization of the adverse consequences that may be resulted from using a DCET which provides inaccurate and misleading estimates. At the end, the main finding regarding the most appropriate technique is validated in the scope and context of the study. Furthermore, the DDCET needed to estimate the number o defects to be found during reinspection is also determined by employing the related guidance available through literature.

However, the related DCETs (namely, subjective and curve-fitting techniques) and DDCETs (namely, reliability growth model) can not be considered in the study, since their incorporation into the reinspection policies is not possible due to feasibility reasons inherent in simulation techniques. So, the question whether these techniques result in different outcomes for the reinspection policies, remains unanswered.

Since one of the output measures considered in the study is related to cost, the determination of cost related parameters also depicts importance. First, the labor cost of performing a reinspection is determined by using the suggested inspection rates in the literature. Further, the study employs the results of National Software Quality Experiment (NSQE) for obtaining the rework costs (in terms of effort) to be spent when a defect is found. Since, NSQE results are extracted from a considerable number of projects, the usage of the cost values provided by NSQE enables the study to represent the real software projects cost structure. Further, as mentioned in the text, one hour is assumed to be the average correction effort when a defect is caught during software inspection phase. At first sight this assumption seems unrealistic. However, if it is changed to another value, for example if it is halved, the total cost values corresponding to all policies will also be halved. Thus, their relative cost values will not change. Consequently, validity of the study remains unthreatened, since it is thought that using unit correction cost as one does not cause any significant change during the standardization of total cost values and also it is suggested as a reasonable value by O'Neill while he is analyzing the software inspection ROI values related to NSQE (O'Neill, 2003). Anyway, when an organization aims to reveal the most appropriate reinspection policy for its environment, it can replicate the study by using the actual value for the average correction cost during inspection in order to obtain the exact total cost values.

CHAPTER 5

CONCLUSION

Software Inspection is a valuable technique for software development organizations which aim to minimize the number of defects propagating to subsequent phases of the software development life cycle, and eventually to the user. By defining a structured procedure that aims the identification and removal of defects just after they are injected into the software work product, software inspection also enables the software organizations to save from the higher rework costs emerging if the defect is captured later. A software development organization can gain more of these and other benefits of software inspection by conducting reinspections when appropriate. Since, performing an additional inspection necessitates the expending of extra resources; it is generally unreasonable to repeat every inspection conducted. Consequently, the reinspection should be carried out if designated criteria are satisfied. The decision that allows selecting among 'Reinspect' and 'Do Not Reinspect' options is called 'Reinspection Decision'. The literature proposes a number of objective methods for concluding reinspection decision. However, the literature does not provide guidance on the appropriateness of different objective reinspection decision methods. With this realization, this study evaluates various reinspection policies for code inspection in the context of a SW-CMM Level 3 organization. These policies are generated by adapting the reinspection decision methods available through literature. The evaluation is conducted by revealing the outcomes obtained at the end of a project due to employment of different reinspection policies. Namely, for each reinspection policy considered, total cost, schedule delay, and major defects found by the users are computed as output measures by employing Monte Carlo simulation. These values are in turn used to extract the ranking of the policies under different preference profiles regarding cost, quality, and schedule. In the study, various factors affecting the software inspection activity and succeeding reinspection decision are taken into account for constituting an experiment which represent circumstances under software projects are executed. This makes possible the analysis of the changes in output measures due to varying levels of the considered factors along with the increased validity of the results reported by the study.

By providing the effects of various reinspection decision policies on project's cost, schedule and quality related objectives; this study provides valuable insights which will hopefully guide software practitioners in determining a reinspection decision policy. First, the reinspection decision policy, which examines the expected percentage of major defects found when the probable reinspection is conducted and compares it against a moderately high threshold value (i.e. 75%), is suggested by the results of the study. Further, it is seen that using inspection effectiveness measure while making the reinspection decision is much more reasonable when compared to the cases where the defect density and net benefit measures are utilized. The study also reveals that applying default decisions of 'Never Reinspect' and 'Always Reinspect' do not exhibit the most appropriate outcomes regarding cost, schedule, and quality. Another guidance that can be obtained from the study is; regardless of the reinspection decision method in charge, the corresponding threshold should be set aggressively, i.e. by employing threshold values close to the maximum or minimum possible. The study additionally provides information about the actions that may be performed by practitioners for improving software project's performance with respect to cost, quality, and schedule. Accordingly, in general, the related initiatives should be taken to increase testing effectiveness and inspector capability. Further, it is seen that increasing the number of inspectors participating to initial inspection and to reinspection promises better performance, although the corresponding labor costs are raised.

Mainly due to time constraints, the study does not have opportunity for dwelling upon some further issues which may provide additional insights regarding software reinspection decision policies. Nevertheless, the author believes that the study signifies a good initial step towards understanding the software life cycle consequences of applying various reinspection decision methods available in the literature. The following lists some areas that can be addressed by prospective research studies aimed on the same subject:

- The study can be replicated for inspections conducted on other software work products such as requirements specifications, design descriptions, test procedures.
- The study can be repeated by employing data from a real software engineering environment.
- The adaptation of the study for higher and lower maturity organizations (vis-à-vis a Software CMM Maturity Level 3 organization) can be performed.
- A more comprehensive simulation model, which also considers the testing process in more detail, can be constituted.
- The definition of total cost measure can be enhanced by quantifying and incorporating the loss of goodwill due to dissatisfaction of the users.
- The preference profiles for cost, quality, and schedule can be constituted by requesting input from software practitioners.
- More advanced multi-objective decision making techniques can be employed while comparing performance of reinspection decision policies with respect to three output measures.
- The size of the inspected artifact and the correction effort during inspection phase can be varied by generating it from an underlying distribution or by selecting different sizes as the levels of a new factor that will be incorporated into experimental design.
- The benefits and feasibility of using software inspection defect estimation techniques, which can not be included in this study, can be explored more extensively.
- Software inspection experiments that employ real software artifacts and inspectors can be designed in order to scrutinize the outcomes of using different reinspection decision policies.

- The study may be repeated by adjusting the reinspection decision policy thresholds based on the data coming from field experience.

REFERENCES

Ackerman, A.F., Buchwald, L.S., Lewski, F.H., 1989. “Software Inspections: An Effective Verification Process”. IEEE Software, Vol. 6, No. 3, pp. 31-36.

Basili, V.R., 1997. “Evolving and Packaging Reading Technologies”. Journal of Systems and Software, Vol. 38, No. 1, pp. 3-12.

Basili, V.R., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Sorumgard, S., Zelkowitz, M.V., 1996. “The Empirical Investigation of Perspective-Based Reading”. Empirical Software Engineering: An International Journal, Vol. 1, No. 2, pp. 133-164.

Biffl, S., 2000. “Using Inspection Data for Defect Estimation”. IEEE Software, Vol. 17, No. 6, pp. 36-43.

Biffl, S., 2003. “Evaluating Defect Estimation Models with Major Defects”. Journal of Systems and Software, Vol. 65, No. 1, pp. 13-29.

Biffl, S., Freimut, B., Laitenberger, O., 2001. “Investigating the Cost-Effectiveness of Reinspections in Software Development”. Proceedings of the Twenty-Third International Conference on Software Engineering, pp. 155-164.

Biffl, S., Gutjahr, W.J., 2002. “Using a Reliability Growth Model to Control Software Inspection”. Empirical Software Engineering: An International Journal, Vol. 7, No. 3, pp. 257-284.

Biffl, S., Halling, M., 2001. “Investigating Reinspection Decision Accuracy Regarding Product-Quality and Cost-Benefit Estimates”. Proceedings of Twenty-Fifth Annual International Computer Software and Applications Conference, pp. 87-96.

Boehm, B.W., Horowitz, E., Madachy, R., Reifer, D., Clark, B.K., Steece, B., Brown, A.W., Chulani, S., Abts, C., 2000. *Software Cost Estimation with Cocomo II*. Prentice Hall PTR.

Briand, L., El Amam, K., Freimut, B., 1998b. "A Comparison and Integration of Capture-Recapture Models and the Detection Profile Method". Proceedings of Ninth International Symposium on Software Reliability Engineering, pp. 32-41.

Briand, L., El Amam, K., Freimut, B., Laitenberger, O., 2000. "A Comprehensive Evaluation of Capture-Recapture Models for Estimating Software Defect Content". IEEE Transactions on Software Engineering, Vol. 26, No. 6, pp. 518-540.

Briand, L., El Amam, K., Laitenberger, O., Fussbroich, T., 1998a. "Using Simulation to Build Inspection Efficiency Benchmarks for Development Projects". Proceedings of Ninth International Symposium on Software Reliability Engineering, pp. 340-349.

Burnham, K.P., Overton, W.S., 1978. "Estimation of the Size of a Closed Population when Capture Probabilities Vary among Animals". Biometrica, Vol. 65, No. 3, pp. 625-633.

Chao, A., 1987. "Estimating Population Size for Capture-Recapture Data with Unequal Catchability". Biometrics, Vol. 43, No. 4, pp. 783-791.

Chao, A., 1989. "Estimating Population Size for Sparse Data in Capture-Recapture Experiments". Biometrics, Vol. 45, No. 2, pp. 427-438.

Chao, A., Lee, S.M., Jeng, S.L., 1992. "Estimating Population Size for Capture-Recapture Data When Capture Probabilities Vary by Time and Individual Animal". Biometrics, Vol. 48, No. 1, pp. 201-216.

Cheng, B., Jeffery, R., 1996. "Comparing Inspection Strategies for Software Requirement Specifications". Proceedings of the 1996 Australian Software Engineering Conference, pp. 203-211.

Chillarege, R., Bhandari, I.S., Chaar, J.K., Halliday, M.J., Moebus, D.S., Ray, B.K., Wong, M.Y, 1992. "Orthogonal Defect Classification-A Concept for In-Process Measurements". IEEE Transactions on Software Engineering, Vol. 18, No. 11, pp. 943-956.

CMMI Product Team, 2001. *CMMI^{SM} for Systems Engineering, Software Engineering, and Integrated Product and Process Development (CMMI-SE/SW/IPPD, V1.1), Continuous Representation.* Carnegie Mellon University.

Doolan, E.P., 1992. "Experience with Fagan's Inspection Method". Software-Practice and Experience, Vol. 22, No. 2, pp. 173-182.

Dyer, M., 1992. *The Cleanroom Approach to Quality Software Development*. John Wiley & Sons.

Eick, S.G., Loader, C.R., Long, M.D., Votta, L.G., Vander Wiel, S.A., 1992. "Estimating Software Fault Content before Coding". Proceedings of Fourteenth International Conference on Software Engineering, pp. 59-65.

El Amam, K., Laitenberger, O., 2001. "Evaluating Capture-Recapture Models with Two Inspectors". IEEE Transactions on Software Engineering, Vol. 27, No. 9, pp. 851-864.

El Amam, K., Laitenberger, O., Harbich, T., 2000. "The Application of Subjective Estimates of Effectiveness to Controlling Software Inspections". Journal of Systems and Software, Vol. 54, No. 2, pp. 119-136.

Fagan, M.E., 1976. "Design and Code Inspections to Reduce Errors in Program Development". IBM Systems Journal, Vol. 15, No. 3, pp. 182-211.

Florac, W.A., Carleton, A.D., 1999. *Measuring the Software Process: Statistical Process Control for Software Process Improvement*. Addison-Wesley Professional.

Franz, L.A., Shih, J.C., 1994. "Estimating the Value of Inspections and Early Testing for Software Projects". Hewlett-Packard Journal, Vol. 45, No. 6, pp. 60-67.

Freimut, B., 1997. "Capture-Recapture Models to Estimate Software Fault Content". Diploma Thesis, University of Kaiserslautern, Germany.

Freimut, B., Laitenberger, O., Biffl, S., 2001. "Investigating the Impact of Reading Techniques on the Accuracy of Different Defect Content Estimation Techniques". Proceedings of Seventh International Software Metrics Symposium, pp. 51-62.

Gilb, T., Graham, D., 1993. *Software Inspection.* Addison-Wesley Longman Limited.

Grady, R.B., Van Slack, T., 1994. “Key Lessons in Achieving Widespread Inspection Use. IEEE Software, Vol. 11, No. 4, pp. 46-57.

Humphrey, W.S., 1999. *Introduction to the Team Software Process*. Addison-Wesley Publishing Company.

IEEE STD 1028-1997, 1998. “IEEE Standard for Software Reviews”.

ISO/IEC TR 15504-2: 1998(E), 1998. “Information Technology - Software process assessment – Part 2: A reference model for processes and process capability”.

Kelly, J.C., Sherif, J.S., Hops, J., 1992. “An analysis of Defect Densities Found During Software Inspections”. Journal of Systems and Software, Vol. 17, No. 2, pp. 111-117.

Laitenberger, O., 1998. “Studying the Effects of Code Inspection and Structural Testing on Software Quality”. Proceedings of Ninth International Symposium on Software Reliability Engineering, pp. 237-246.

Laitenberger, O., Atkinson, C., Schlich, M., El Amam, K., 2000. “An experimental comparison of reading techniques for defect detection in UML design documents”. Journal of Systems and Software, Vol. 53, No. 2, pp. 183-204.

Laitenberger, O., DeBaud, J., 2000. “An Encompassing Life Cycle Centric Survey of Software Inspection”. Journal of Systems and Software, Vol. 50, No. 1, pp. 5-31.

Laitenberger, O., El Emam, K., Harbich, T.G., 2001. “An Internally Replicated Quasi-Experimental Comparison of Checklist and Perspective-Based Reading of Code Documents”. IEEE Transactions on Software Engineering, Vol. 27, No. 5, pp. 387-421.

Macdonald, F., Miller, J., 1997. “A Software Inspection Process Definition Language and Prototype Support Tool”. Software Testing, Verification and Reliability, Vol. 7, No. 2, pp. 99-128.

Mah, M., 2001. “Defect Metrics, Inspections, and Testing: Part II”. IT Metrics Strategies, Vol. 7, No. 2, pp. 6-16.

Miller, J., 1999. “Estimating the Number of Remaining Defects after Inspection”. Software Testing, Verification, and Reliability, Vol. 9, No. 3, pp. 167-189.

Mills, H.D., 1972. “On the statistical validation of computer programs”. Technical Report no. FSC-72-6015, IBM Federal Systems Division.

O’Neill, D., 2002. “Peer Reviews”. In: *Encyclopedia of Software Engineering* (2nd edition). John Wiley & Sons.

O’Neill, D., 2003. “Determining Return on Investment Using Software Inspections”. Crosstalk: The Journal of Defense Software Engineering, Vol. 16, No. 3, pp. 16-21.

Otis, D.L., Burnham, K.P., White, G.C., Anderson, D.R., 1978. “Statistical Inference from Capture Data on Closed Animal Populations”. Wildlife Monographs, Vol. 62, pp. 1–135.

Paulk, M.C., Curtis, B., Chrissis, M.B., Weber, C.V., 1993. *Capability Maturity Model for Software, Version 1.1 (CMU/SEI-93-TR-024, ADA 263403).* Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University.

Petersson, H., Thelin, T., Runeson, P., Wohlin, C., 2004. “Capture–Recapture in Software Inspections after 10 Years Research—Theory, Evaluation and Application”. Journal of Systems and Software, Vol. 72, No. 2, pp. 249-264.

Porter, A.A. Votta, L.G., Jr. Basili, V.R., 1995. “Comparing Detection Methods for Software Requirements Inspections: A Replicated Experiment. IEEE Transactions on Software Engineering, Vol. 21, No. 6, pp. 563-575.

Porter, A.A., Siy, H.P., Toman, C.A., Votta, L.G., 1997. “An experiment to Assess the Cost-Benefits of Code Inspections in Large Scale Software Development”. IEEE Transactions on Software Engineering, Vol. 23, No. 6, pp. 329-346.

Radice, R.A., 2002. *High Quality Low Cost Software Inspections.* Paradoxicon Publishing.

Rus, I., Halling, M., Biffl, S., 2003. “Supporting Decision-Making in Software Engineering with Process Simulation and Empirical Studies”. International Journal of Software Engineering and Knowledge Engineering (IJSEKE), Vol. 13, No. 5, pp. 531-545.

Russell, G.W., 1991. “Experience with Inspection in Ultralarge-Scale Developments”. IEEE Software, Vol. 8, No. 1, pp. 25-31.

Seber, G.A.F, 1982. *The Estimation of Animal Abundance and Related Parameters,* 2nd Edition. Charles Griffin and Co.

Selby, R., Basili, V., 1991. "Analyzing Error-prone System Structure". IEEE Transactions on Software Engineering, Vol. 17, No. 2, pp. 141-152.

Software Engineering Institute, 2004. *Software CMM Process Maturity Profile: 2004 Mid-Year Update*. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University. (URL: http://www.sei.cmu.edu/sema/pdf/SW-CMM/2004aug.pdf, Last Accessed on: 02.01.2005).

Strauss, S.H., Ebenau, R.G., 1993. *Software Inspection Process.* Mcgraw Hill Systems Design and Implementation Series.

Tervonen, I., Iisakka, J., 1997. "Monitoring Software Inspections with Prescriptive Metrics". Software QA, Vol. 4, No. 3, pp. 16-23.

Thelin, T., Petersson, H., Runeson, P., 2002. "Confidence Intervals for Capture-Recapture Estimations in Software Inspections". Information and Software Technology, Vol. 44, No. 12, pp. 683-702.

Thelin, T., Runeson, P., 2000. "Fault Content Estimations Using Extended Curve Fitting Models and Model Selection". Proceedings of the Fourth International Conference on Empirical Assessment and Evaluation in Software Engineering.

Thelin, T., Runeson, P., Regnell, B., 2001. "Usage-Based Reading - An Experiment to Guide Reviewers with Use Cases". Information and Software Technology, Vol. 43, No. 15, pp. 925-938.

Thelin, T., Runeson, P., Wohlin, C., 2003. "An Experimental Comparison of Usage-based and Checklist-based Reading". IEEE Transactions on Software Engineering, Vol. 29, No. 8, pp. 687- 704.

Travassos, G., Shull, F., Fredericks, M., Basili, V.R., 1999. "Detecting Defects in Object-Oriented Designs: Using Reading Techniques to Increase Software Quality". Proceedings of Object Oriented Programming Conference (OOPSLA'99), pp. 47-56.

Vander Wiel, S.A., Votta, L.G., 1993. “Assessing Software Designs Using Capture–Recapture Methods”. IEEE Transactions on Software Engineering, Vol. 19, No. 11, pp. 1045–1054.

Weller, E.F., 1993. “Lessons from Three Years of Inspection Data”. IEEE Software, Vol. 10, No. 5, pp. 38-45.

Wheeler, D.A., Brykczynski, B., Meeson, R.N., 1996. *Software Inspection: An industry Best Practice*. IEEE Computer Society Press.

Wheeler, D.A., Brykczynski, B., Meeson, R.N., 1997. “Software Peer Reviews”. In: *Software Engineering Project Management* (2nd edition). John Wiley & Sons- IEEE Computer Society Press.

Wiegers, K.E., 2002. *Peer Reviews in Software: A Practical Guide.* Addison-Wesley Information Technology Series.

Wohlin, C., Runeson, P., 1998. “Defect Content Estimations from Review Data”. Proceedings of the Twentieth International Conference on Software Engineering, pp.400-409.

APPENDICES

APPENDIX A

SIMULATED DATA FOR FAILURE RATE MEASURE OF ESTIMATORS

Table 32 Failure Rate Data of Considered Estimators for All Defects

Estimator	Treatment	Failure Rate
1	1	0.017
1	2	0.001
1	3	0
1	4	0.149
1	5	0.056
1	6	0.011
1	7	0.311
1	8	0.138
1	9	0.056
1	10	0.002
1	11	0
1	12	0
1	13	0.025
1	14	0
1	15	0
1	16	0.104
1	17	0.05
1	18	0.01
1	19	0
1	20	0
1	21	0
1	22	0.033
1	23	0.005
1	24	0.002
1	25	0
1	26	0
1	27	0
1	28	0.356
1	29	0.207
1	30	0.087
1	31	0.14
1	32	0.034
1	33	0.004
1	34	0.098
1	35	0.021

Estimator	Treatment	Failure Rate
5	1	0.015
5	2	0.002
5	3	0
5	4	0.147
5	5	0.053
5	6	0.014
5	7	0.318
5	8	0.15
5	9	0.059
5	10	0.002
5	11	0
5	12	0
5	13	0.011
5	14	0.003
5	15	0
5	16	0.111
5	17	0.035
5	18	0.011
5	19	0.002
5	20	0
5	21	0
5	22	0.041
5	23	0.012
5	24	0.002
5	25	0
5	26	0
5	27	0
5	28	0.377
5	29	0.246
5	30	0.089
5	31	0.142
5	32	0.039
5	33	0.005
5	34	0.087
5	35	0.018

Table 32 Failure Rate Data of Considered Estimators for All Defects (cont.)

Estimator	Treatment	Failure Rate
1	36	0.001
1	37	0.002
1	38	0
1	39	0
1	40	0.087
1	41	0.022
1	42	0.003
1	43	0.013
1	44	0
1	45	0
1	46	0.001
1	47	0
1	48	0
1	49	0
1	50	0
1	51	0
1	52	0.001
1	53	0
1	54	0
2	1	0.024
2	2	0
2	3	0
2	4	0.141
2	5	0.049
2	6	0.011
2	7	0.304
2	8	0.135
2	9	0.052
2	10	0.003
2	11	0
2	12	0
2	13	0.017
2	14	0.002
2	15	0
2	16	0.109
2	17	0.037
2	18	0.015
2	19	0
2	20	0
2	21	0
2	22	0.03
2	23	0.01
2	24	0.003
2	25	0
2	26	0
2	27	0
2	28	0.326
2	29	0.209
2	30	0.104
2	31	0.159
2	32	0.05
2	33	0.007
2	34	0.104
2	35	0.019
2	36	0
2	37	0.007
2	38	0
2	39	0
2	40	0.073
2	41	0.027
2	42	0.008
2	43	0.004
2	44	0
2	45	0

Estimator	Treatment	Failure Rate
5	36	0.001
5	37	0.004
5	38	0
5	39	0
5	40	0.102
5	41	0.031
5	42	0.005
5	43	0.008
5	44	0
5	45	0
5	46	0.004
5	47	0
5	48	0
5	49	0
5	50	0
5	51	0
5	52	0
5	53	0
5	54	0
6	1	0.015
6	2	0.002
6	3	0
6	4	0.145
6	5	0.049
6	6	0.005
6	7	0.284
6	8	0.157
6	9	0.058
6	10	0.002
6	11	0
6	12	0
6	13	0.023
6	14	0.003
6	15	0
6	16	0.103
6	17	0.024
6	18	0.014
6	19	0.001
6	20	0
6	21	0
6	22	0.021
6	23	0.008
6	24	0.004
6	25	0
6	26	0
6	27	0
6	28	0.38
6	29	0.205
6	30	0.095
6	31	0.141
6	32	0.044
6	33	0.009
6	34	0.074
6	35	0.018
6	36	0.001
6	37	0.003
6	38	0
6	39	0
6	40	0.081
6	41	0.026
6	42	0.002
6	43	0.005
6	44	0
6	45	0

Table 32 Failure Rate Data of Considered Estimators for All Defects (cont.)

Estimator	Treatment	Failure Rate
2	46	0.004
2	47	0
2	48	0
2	49	0
2	50	0
2	51	0
2	52	0
2	53	0
2	54	0
3	1	0
3	2	0
3	3	0
3	4	0
3	5	0
3	6	0
3	7	0
3	8	0
3	9	0
3	10	0
3	11	0
3	12	0
3	13	0
3	14	0
3	15	0
3	16	0
3	17	0
3	18	0
3	19	0
3	20	0
3	21	0
3	22	0
3	23	0
3	24	0
3	25	0
3	26	0
3	27	0
3	28	0
3	29	0
3	30	0
3	31	0
3	32	0
3	33	0
3	34	0
3	35	0
3	36	0
3	37	0
3	38	0
3	39	0
3	40	0
3	41	0
3	42	0
3	43	0
3	44	0
3	45	0
3	46	0
3	47	0
3	48	0
3	49	0
3	50	0
3	51	0
3	52	0
3	53	0
3	54	0
4	1	0.01

Estimator	Treatment	Failure Rate
6	46	0.002
6	47	0.001
6	48	0
6	49	0
6	50	0
6	51	0
6	52	0
6	53	0
6	54	0
8	1	0
8	2	0
8	3	0
8	4	0
8	5	0
8	6	0
8	7	0
8	8	0
8	9	0
8	10	0
8	11	0
8	12	0
8	13	0
8	14	0
8	15	0
8	16	0
8	17	0
8	18	0
8	19	0
8	20	0
8	21	0
8	22	0
8	23	0
8	24	0
8	25	0
8	26	0
8	27	0
8	28	0
8	29	0
8	30	0
8	31	0
8	32	0
8	33	0
8	34	0
8	35	0
8	36	0
8	37	0
8	38	0
8	39	0
8	40	0
8	41	0
8	42	0
8	43	0
8	44	0
8	45	0
8	46	0
8	47	0
8	48	0
8	49	0
8	50	0
8	51	0
8	52	0
8	53	0
8	54	0
9	1	0

Table 32 Failure Rate Data of Considered Estimators for All Defects (cont.)

Estimator	Treatment	Failure Rate
4	2	0.003
4	3	0
4	4	0.158
4	5	0.046
4	6	0.011
4	7	0.293
4	8	0.149
4	9	0.061
4	10	0.002
4	11	0
4	12	0
4	13	0.033
4	14	0.002
4	15	0
4	16	0.168
4	17	0.09
4	18	0.074
4	19	0.003
4	20	0
4	21	0
4	22	0.116
4	23	0.11
4	24	0.099
4	25	0
4	26	0.001
4	27	0
4	28	0.361
4	29	0.204
4	30	0.107
4	31	0.141
4	32	0.045
4	33	0.009
4	34	0.073
4	35	0.013
4	36	0.002
4	37	0.015
4	38	0.002
4	39	0.001
4	40	0.126
4	41	0.043
4	42	0.016
4	43	0.013
4	44	0.001
4	45	0
4	46	0.014
4	47	0.004
4	48	0
4	49	0
4	50	0.001
4	51	0
4	52	0.005
4	53	0
4	54	0.003

Estimator	Treatment	Failure Rate
9	2	0
9	3	0
9	4	0
9	5	0
9	6	0
9	7	0
9	8	0
9	9	0
9	10	0
9	11	0
9	12	0
9	13	0
9	14	0
9	15	0
9	16	0
9	17	0
9	18	0
9	19	0
9	20	0
9	21	0
9	22	0
9	23	0
9	24	0
9	25	0
9	26	0
9	27	0
9	28	0
9	29	0
9	30	0
9	31	0
9	32	0
9	33	0
9	34	0
9	35	0
9	36	0
9	37	0
9	38	0
9	39	0
9	40	0
9	41	0
9	42	0
9	43	0
9	44	0
9	45	0
9	46	0
9	47	0
9	48	0
9	49	0
9	50	0
9	51	0
9	52	0
9	53	0
9	54	0

Table 33 Failure Rate Data of Considered Estimators for Major Defects

Estimator	Treatment	Failure Rate
1	1	0.663113
1	2	0.535676
1	3	0.366273
1	4	0.601602
1	5	0.477
1	6	0.328
1	7	0.61
1	8	0.488
1	9	0.348
1	10	0.202202
1	11	0.092
1	12	0.051
1	13	0.208
1	14	0.077
1	15	0.024
1	16	0.781148
1	17	0.717063
1	18	0.618844
1	19	0.501587
1	20	0.367412
1	21	0.232238
1	22	0.434434
1	23	0.33033
1	24	0.253253
1	25	0.017
1	26	0.006
1	27	0.001
1	28	0.673
1	29	0.526
1	30	0.389
1	31	0.821845
1	32	0.723769
1	33	0.629274
1	34	0.56513
1	35	0.355
1	36	0.234
1	37	0.134
1	38	0.039
1	39	0.014
1	40	0.771368
1	41	0.688749
1	42	0.601704
1	43	0.315631
1	44	0.178
1	45	0.089
1	46	0.266
1	47	0.154
1	48	0.079
1	49	0.006
1	50	0
1	51	0
1	52	0.487779
1	53	0.319654
1	54	0.255365
2	1	0.659188
2	2	0.535865
2	3	0.355603
2	4	0.631632
2	5	0.456456

Estimator	Treatment	Failure Rate
5	1	0.659597
5	2	0.490967
5	3	0.316348
5	4	0.608609
5	5	0.495
5	6	0.327
5	7	0.635
5	8	0.461
5	9	0.332
5	10	0.202405
5	11	0.087
5	12	0.046092
5	13	0.196
5	14	0.098
5	15	0.023
5	16	0.769892
5	17	0.729412
5	18	0.640043
5	19	0.46044
5	20	0.340812
5	21	0.214665
5	22	0.417
5	23	0.332
5	24	0.259
5	25	0.013
5	26	0.003
5	27	0
5	28	0.659
5	29	0.539
5	30	0.387
5	31	0.810458
5	32	0.734672
5	33	0.612834
5	34	0.530531
5	35	0.369
5	36	0.248248
5	37	0.121
5	38	0.041
5	39	0.02
5	40	0.763158
5	41	0.690811
5	42	0.600212
5	43	0.32032
5	44	0.174524
5	45	0.099
5	46	0.281
5	47	0.144144
5	48	0.071071
5	49	0.009
5	50	0.001
5	51	0
5	52	0.4375
5	53	0.331887
5	54	0.24866
6	1	0.618182
6	2	0.52234
6	3	0.370095
6	4	0.63
6	5	0.495

Table 33 Failure Rate Data of Considered Estimators for Major Defects (cont.)

Estimator	Treatment	Failure Rate
2	6	0.351
2	7	0.613
2	8	0.5
2	9	0.328
2	10	0.204
2	11	0.087087
2	12	0.032032
2	13	0.201
2	14	0.095
2	15	0.045
2	16	0.816435
2	17	0.715344
2	18	0.623391
2	19	0.469762
2	20	0.361588
2	21	0.226115
2	22	0.457457
2	23	0.313
2	24	0.273
2	25	0.018
2	26	0.003
2	27	0
2	28	0.689
2	29	0.527
2	30	0.415
2	31	0.801927
2	32	0.72043
2	33	0.61242
2	34	0.532533
2	35	0.37
2	36	0.235
2	37	0.127
2	38	0.032
2	39	0.009
2	40	0.786022
2	41	0.678919
2	42	0.578834
2	43	0.307307
2	44	0.157
2	45	0.089089
2	46	0.259
2	47	0.149
2	48	0.081
2	49	0.004
2	50	0
2	51	0
2	52	0.478587
2	53	0.337272
2	54	0.23395
3	1	0
3	2	0
3	3	0
3	4	0
3	5	0
3	6	0
3	7	0
3	8	0
3	9	0
3	10	0
3	11	0
3	12	0
3	13	0
3	14	0
3	15	0

Estimator	Treatment	Failure Rate
6	6	0.336673
6	7	0.611
6	8	0.487
6	9	0.321
6	10	0.196
6	11	0.098098
6	12	0.042042
6	13	0.204
6	14	0.112
6	15	0.021
6	16	0.794243
6	17	0.714898
6	18	0.646432
6	19	0.487674
6	20	0.364316
6	21	0.225772
6	22	0.428
6	23	0.348
6	24	0.273273
6	25	0.009
6	26	0.001
6	27	0.001
6	28	0.678
6	29	0.53
6	30	0.379
6	31	0.800843
6	32	0.720213
6	33	0.592119
6	34	0.572573
6	35	0.377
6	36	0.233233
6	37	0.121
6	38	0.044
6	39	0.019
6	40	0.763713
6	41	0.66525
6	42	0.565032
6	43	0.3
6	44	0.166
6	45	0.089089
6	46	0.272
6	47	0.152
6	48	0.075075
6	49	0.006
6	50	0.002
6	51	0
6	52	0.483279
6	53	0.348712
6	54	0.26327
8	1	0
8	2	0
8	3	0
8	4	0
8	5	0
8	6	0
8	7	0
8	8	0
8	9	0
8	10	0
8	11	0
8	12	0
8	13	0
8	14	0
8	15	0

Table 33 Failure Rate Data of Considered Estimators for Major Defects (cont.)

Estimator	Treatment	Failure Rate
3	16	0
3	17	0
3	18	0
3	19	0
3	20	0
3	21	0
3	22	0
3	23	0
3	24	0
3	25	0
3	26	0
3	27	0
3	28	0
3	29	0
3	30	0
3	31	0
3	32	0
3	33	0
3	34	0
3	35	0
3	36	0
3	37	0
3	38	0
3	39	0
3	40	0
3	41	0
3	42	0
3	43	0
3	44	0
3	45	0
3	46	0
3	47	0
3	48	0
3	49	0
3	50	0
3	51	0
3	52	0
3	53	0
3	54	0
4	1	0.646932
4	2	0.506329
4	3	0.357374
4	4	0.597
4	5	0.471
4	6	0.326326
4	7	0.621
4	8	0.489
4	9	0.335
4	10	0.263
4	11	0.134
4	12	0.111111
4	13	0.251
4	14	0.134
4	15	0.05
4	16	0.805794
4	17	0.778378
4	18	0.745474
4	19	0.592077
4	20	0.454352
4	21	0.4037
4	22	0.582164
4	23	0.552
4	24	0.577
4	25	0.045
8	16	0
8	17	0
8	18	0
8	19	0
8	20	0
8	21	0
8	22	0
8	23	0
8	24	0
8	25	0
8	26	0
8	27	0
8	28	0
8	29	0
8	30	0
8	31	0
8	32	0
8	33	0
8	34	0
8	35	0
8	36	0
8	37	0
8	38	0
8	39	0
8	40	0
8	41	0
8	42	0
8	43	0
8	44	0
8	45	0
8	46	0
8	47	0
8	48	0
8	49	0
8	50	0
8	51	0
8	52	0
8	53	0
8	54	0
9	1	0
9	2	0
9	3	0
9	4	0
9	5	0
9	6	0
9	7	0
9	8	0
9	9	0
9	10	0
9	11	0
9	12	0
9	13	0
9	14	0
9	15	0
9	16	0
9	17	0
9	18	0
9	19	0
9	20	0
9	21	0
9	22	0
9	23	0
9	24	0
9	25	0

Table 33 Failure Rate Data of Considered Estimators for Major Defects (cont.)

Estimator	Treatment	Failure Rate	Estimator	Treatment	Failure Rate
4	26	0.032	9	26	0
4	27	0.027	9	27	0
4	28	0.67	9	28	0
4	29	0.562	9	29	0
4	30	0.385	9	30	0
4	31	0.804747	9	31	0
4	32	0.719828	9	32	0
4	33	0.619808	9	33	0
4	34	0.54	9	34	0
4	35	0.38038	9	35	0
4	36	0.229229	9	36	0
4	37	0.176	9	37	0
4	38	0.106	9	38	0
4	39	0.081	9	39	0
4	40	0.805223	9	40	0
4	41	0.708021	9	41	0
4	42	0.676344	9	42	0
4	43	0.364364	9	43	0
4	44	0.244244	9	44	0
4	45	0.133	9	45	0
4	46	0.341	9	46	0
4	47	0.216	9	47	0
4	48	0.148	9	48	0
4	49	0.044	9	49	0
4	50	0.047	9	50	0
4	51	0.066	9	51	0
4	52	0.586538	9	52	0
4	53	0.51746	9	53	0
4	54	0.480423	9	54	0

APPENDIX B

SIMULATED DATA FOR RELATIVE ERROR AND ROOT MEAN SQUARE ERROR MEASURES OF ESTIMATORS

Table 34 Relative Error and Root Mean Square Error Data of Estimators 3, 8, and 9 for All Defects

Estimator	Treatment	Mean Relative Error	RMSE
3	1	-0.170518	9.3303
3	2	-0.158997	6.5324
3	3	-0.160115	5.4528
3	4	-0.262659	11.4591
3	5	-0.219821	10.319
3	6	-0.207428	8.2442
3	7	-0.219408	12.1363
3	8	-0.132839	11.3554
3	9	-0.059723	10.7285
3	10	-0.078872	7.523
3	11	-0.070376	4.9831
3	12	-0.07252	3.6758
3	13	-0.053781	10.0168
3	14	-0.028502	7.6615
3	15	-0.023221	6.1682
3	16	-0.578454	16.0122
3	17	-0.557886	15.4028
3	18	-0.519636	14.4799
3	19	-0.019421	8.7882
3	20	-0.033512	5.2772

Estimator	Treatment	Mean Relative Error	RMSE
8	28	-0.568038	15.1856
8	29	-0.491981	13.2964
8	30	-0.400981	11.0302
8	31	-0.342692	9.673
8	32	-0.216058	6.7825
8	33	-0.112673	4.6262
8	34	-0.470712	12.7772
8	35	-0.389692	10.7333
8	36	-0.344635	9.5645
8	37	-0.344885	9.7044
8	38	-0.297577	8.5403
8	39	-0.277	7.9549
8	40	-0.381538	10.6747
8	41	-0.303167	8.8469
8	42	-0.213462	6.7758
8	43	-0.102192	4.7029
8	44	-0.006923	3.6764
8	45	0.066077	3.7772
8	46	-0.082173	4.6732
8	47	0.013654	3.7139

Table 34 Relative Error and Root Mean Square Error Data of Estimators 3, 8, and 9 for All Defects (cont.)

Estimator	Treatment	Mean Relative Error	RMSE
3	21	-0.022307	3.9528
3	22	-0.509716	14.7552
3	23	-0.485719	13.8878
3	24	-0.408147	12.7712
3	25	-0.076513	4.8248
3	26	-0.058543	3.8445
3	27	-0.054581	2.8257
3	28	-0.417466	13.8397
3	29	-0.329295	12.1063
3	30	-0.258733	11.5549
3	31	-0.108693	11.7101
3	32	-0.060656	9.6867
3	33	-0.026251	8.1724
3	34	-0.416569	12.6753
3	35	-0.391158	11.4017
3	36	-0.384049	10.6809
3	37	-0.38141	10.9974
3	38	-0.362761	10.3582
3	39	-0.352108	9.7126
3	40	-0.243902	11.7328
3	41	-0.198002	10.4163
3	42	-0.174224	9.255
3	43	-0.016401	9.7813
3	44	-0.011704	7.1579
3	45	-0.011737	4.8389
3	46	-0.042875	8.8952
3	47	-0.010262	7.0636
3	48	-0.009441	4.7775
3	49	-0.1404	5.3287
3	50	-0.140415	4.5379
3	51	-0.128868	3.9568
3	52	-0.138423	8.3336
3	53	-0.119246	6.9722
3	54	-0.107923	4.9115
8	1	-0.235058	7.1914
8	2	-0.134712	4.919
8	3	-0.071904	3.5621
8	4	-0.429519	11.7611
8	5	-0.321962	9.1387
8	6	-0.232058	7.0437
8	7	-0.4595	12.5402
8	8	-0.355558	9.9862

Estimator	Treatment	Mean Relative Error	RMSE
8	48	0.075635	3.9415
8	49	-0.031548	3.3393
8	50	-0.036365	2.979
8	51	-0.060163	2.9808
8	52	-0.111788	4.8495
8	53	-0.055654	3.9857
8	54	-0.013548	3.2308
9	1	-0.227481	6.9994
9	2	-0.132788	4.857
9	3	-0.068346	3.5562
9	4	-0.430173	11.7783
9	5	-0.3325	9.4153
9	6	-0.232538	7.0627
9	7	-0.460365	12.4823
9	8	-0.352231	9.9313
9	9	-0.247288	7.4601
9	10	0.030397	4.6456
9	11	0.093821	4.8536
9	12	0.094571	4.4123
9	13	-0.055615	4.9397
9	14	0.034103	4.6119
9	15	0.117782	5.0598
9	16	-0.554045	14.9652
9	17	-0.504994	13.7356
9	18	-0.470603	12.885
9	19	0.102974	5.6401
9	20	0.167064	6.5555
9	21	0.175683	6.2477
9	22	-0.462955	12.8562
9	23	-0.410029	11.5547
9	24	-0.352583	10.2459
9	25	0.116673	5.5763
9	26	0.112542	5.1326
9	27	0.074147	4.1439
9	28	-0.570846	15.2486
9	29	-0.490269	13.2087
9	30	-0.412692	11.2751
9	31	-0.336827	9.5977
9	32	-0.223442	6.9306
9	33	-0.108712	4.5271
9	34	-0.464327	12.6026
9	35	-0.393635	10.8127

Table 34 Relative Error and Root Mean Square Error Data of Estimators 3, 8, and 9 for All Defects (cont.)

Estimator	Treatment	Mean Relative Error	RMSE
8	9	-0.252077	7.4994
8	10	-0.077359	4.2392
8	11	0.000526	3.3149
8	12	0.041577	3.1104
8	13	-0.16909	5.9452
8	14	-0.063615	4.2594
8	15	0.017628	3.3406
8	16	-0.598897	15.9451
8	17	-0.555231	14.8327
8	18	-0.512269	13.77
8	19	-0.034375	3.9107
8	20	0.047769	3.7935
8	21	0.093221	3.8546
8	22	-0.5275	14.207
8	23	-0.478212	12.9619
8	24	-0.442885	12.1274
8	25	0.018962	3.461
8	26	0.0545	3.2742
8	27	0.061019	3.0056

Estimator	Treatment	Mean Relative Error	RMSE
9	36	-0.334558	9.3528
9	37	-0.274603	8.3673
9	38	-0.247526	7.5633
9	39	-0.248962	7.4787
9	40	-0.306179	9.2252
9	41	-0.207391	7.1974
9	42	-0.118218	5.4446
9	43	0.009218	4.5535
9	44	0.094346	4.9519
9	45	0.146423	5.5602
9	46	0.061503	5.4737
9	47	0.140663	6.0734
9	48	0.185923	6.7739
9	49	0.014785	4.5011
9	50	-0.027417	3.8234
9	51	-0.073875	3.4916
9	52	-0.010728	4.9792
9	53	0.032356	4.8022
9	54	0.056808	4.6029

Table 35 Relative Error and Root Mean Square Error Data of Estimators 3, 8, and 9 for Major Defects

Estimator	Treatment	Mean Relative Error	RMSE	Estimator	Treatment	Mean Relative Error	RMSE
3	1	-0.371652	1.8769	8	28	-0.579143	6.50812
3	2	-0.305583	1.71324	8	29	-0.493189	5.7683
3	3	-0.213303	1.63794	8	30	-0.422036	5.17821
3	4	-0.459541	3.98666	8	31	-0.331273	1.7666
3	5	-0.357722	3.9011	8	32	-0.209901	1.73656
3	6	-0.289081	3.54186	8	33	-0.105339	1.74316
3	7	-0.429792	5.96946	8	34	-0.478948	3.70829
3	8	-0.304648	5.42674	8	35	-0.410373	3.43375
3	9	-0.20323	5.31631	8	36	-0.344776	3.15978
3	10	-0.167783	3.47843	8	37	-0.32064	4.4127
3	11	-0.11043	3.05165	8	38	-0.294231	4.15108
3	12	-0.089469	2.75356	8	39	-0.278764	4.01223
3	13	-0.153295	5.17997	8	40	-0.395225	1.94278
3	14	-0.085988	4.94615	8	41	-0.298798	1.75411
3	15	-0.040815	4.39081	8	42	-0.204504	1.84108
3	16	-0.715372	2.16315	8	43	-0.106331	2.75233
3	17	-0.644706	2.12417	8	44	-0.010829	2.69168
3	18	-0.632593	2.08685	8	45	0.056689	2.79454
3	19	-0.254854	1.92293	8	46	-0.089387	2.82987
3	20	-0.152016	1.76798	8	47	0.016888	2.84163
3	21	-0.110692	1.74465	8	48	0.071735	2.9521
3	22	-0.612539	4.49137	8	49	-0.023796	3.18731
3	23	-0.564111	4.26103	8	50	-0.039018	2.9798
3	24	-0.528091	4.07621	8	51	-0.055021	2.80334
3	25	-0.090726	3.62797	8	52	-0.111696	1.91047
3	26	-0.07112	3.37683	8	53	-0.053292	1.74014
3	27	-0.049415	2.94447	8	54	-0.000109	1.79426
3	28	-0.569576	6.79597	9	1	-0.215067	1.66836
3	29	-0.479694	6.23228	9	2	-0.113929	1.59122
3	30	-0.371151	5.64374	9	3	-0.063683	1.63364
3	31	-0.48136	2.06801	9	4	-0.426805	3.54738
3	32	-0.36209	1.91565	9	5	-0.317762	3.10385
3	33	-0.275979	1.88058	9	6	-0.229155	2.83215
3	34	-0.523265	4.15182	9	7	-0.450233	5.37261
3	35	-0.452992	3.74887	9	8	-0.365498	4.70206
3	36	-0.41076	3.59561	9	9	-0.248084	3.98969
3	37	-0.39882	5.2391	9	10	0.076533	3.29953
3	38	-0.37999	4.88541	9	11	0.17735	3.50717
3	39	-0.360901	4.67138	9	12	0.227854	3.57845
3	40	-0.561306	2.03639	9	13	-0.004453	3.76151
3	41	-0.470088	2.06981	9	14	0.109305	4.0781
3	42	-0.395571	1.87854	9	15	0.208726	4.46434
3	43	-0.178127	3.77466	9	16	-0.540142	2.10564
3	44	-0.087611	3.52801	9	17	-0.4867	2.10837
3	45	-0.04914	3.2395	9	18	-0.411653	1.96703
3	46	-0.157254	4.13051	9	19	0.250362	2.56139
3	47	-0.09218	3.54221	9	20	0.355274	2.63076
3	48	-0.078026	3.06038	9	21	0.402876	2.60624
3	49	-0.159092	3.59744	9	22	-0.399369	3.75961
3	50	-0.132354	3.18844	9	23	-0.350811	3.54133
3	51	-0.131993	2.8763	9	24	-0.297533	3.40323
3	52	-0.329027	1.69103	9	25	0.30314	5.29435
3	53	-0.246853	1.81373	9	26	0.313848	5.18042
3	54	-0.197498	1.77416	9	27	0.291319	4.84373
8	1	-0.279741	1.75989	9	28	-0.560163	6.26718
8	2	-0.103625	1.6105	9	29	-0.483006	5.60066
8	3	-0.087408	1.66665	9	30	-0.409094	5.07144

Table 35 Relative Error and Root Mean Square Error Data of Estimators 3, 8, and 9 for Major Defects (cont.)

Estimator	Treatment	Mean Relative Error	RMSE
8	4	-0.41274	3.48887
8	5	-0.337709	3.20836
8	6	-0.237246	2.83033
8	7	-0.453145	5.4265
8	8	-0.342344	4.62343
8	9	-0.252142	4.05683
8	10	-0.080385	2.8626
8	11	-0.004173	2.73266
8	12	0.052062	2.68858
8	13	-0.17403	3.66891
8	14	-0.058227	3.29225
8	15	0.019683	3.43844
8	16	-0.596197	2.06203
8	17	-0.536831	1.92789
8	18	-0.484301	1.85555
8	19	-0.033315	1.8364
8	20	0.048417	1.8977
8	21	0.092077	1.79104
8	22	-0.546019	4.12637
8	23	-0.476784	3.77352
8	24	-0.425822	3.55525
8	25	0.016226	3.42253
8	26	0.062133	3.21983
8	27	0.069007	3.18367

Estimator	Treatment	Mean Relative Error	RMSE
9	31	-0.359288	1.83279
9	32	-0.196578	1.72374
9	33	-0.104072	1.66057
9	34	-0.466796	3.73684
9	35	-0.390737	3.39453
9	36	-0.338808	3.11203
9	37	-0.230104	4.08717
9	38	-0.193155	3.86544
9	39	-0.171926	3.80532
9	40	-0.281108	2.01524
9	41	-0.160776	2.04061
9	42	-0.059939	2.11841
9	43	0.049987	3.30714
9	44	0.169637	3.57356
9	45	0.238918	3.80246
9	46	0.196148	3.76614
9	47	0.296478	4.2613
9	48	0.336124	4.44001
9	49	0.210535	4.78601
9	50	0.180644	4.30411
9	51	0.078981	3.52252
9	52	0.134317	2.30539
9	53	0.199079	2.22026
9	54	0.230994	2.30637

APPENDIX C

SIMULATED DATA FOR DECISION ACCURACY AND RELATIVE DECISION ACCURACY MEASURES OF ESTIMATORS

Table 36 Decision Accuracy and Relative Decision Accuracy Data of Estimators 3, 8, and 9

Estimator	Policy No	Treat-ment	Mean DA	Mean RDA
3	1	54	0.957167	-0.005626
3	2	54	0.688185	0.010018
3	3	54	0.576852	0.090521
3	4	54	0.566556	0.343406
3	5	54	0.788926	0.764902
3	6	54	0.953185	-0.000055
3	7	54	0.737593	-0.011328
3	8	54	0.587889	0.005104
3	9	54	0.500148	0.082608
3	10	54	0.521296	0.252945
3	11	54	0.628852	0.606697
3	12	54	0.437407	0.286318
3	13	54	0.4445	0.102229
3	14	54	0.553074	0.011606
3	15	54	0.699704	-0.029748
3	16	54	0.486704	0.240585
3	17	54	0.515907	0.094648
3	18	54	0.614556	0.030548
3	19	54	0.712204	0.001912
3	20	54	0.793	-0.014142
3	21	54	0.552278	-0.338758
3	22	54	0.602519	-0.127417
3	23	54	0.625593	-0.010177
3	24	54	0.635019	0.010963
3	25	54	0.655037	0.007883
3	26	54	0.603722	-0.211128
3	27	54	0.642685	-0.122354
3	28	54	0.690463	-0.099646
3	29	54	0.729611	-0.099962
8	24	54	0.578722	-0.018962
8	25	54	0.627889	-0.018511
8	26	54	0.488889	-0.330175
8	27	54	0.588574	-0.177923
8	28	54	0.666352	-0.117696
8	29	54	0.727519	-0.101417
8	30	54	0.777704	-0.09451
8	31	54	0.9645	0
8	32	54	0.754796	0
8	33	54	0.67163	-0.001341
8	34	54	0.521611	-0.099777
8	35	54	0.432463	-0.343603
8	36	54	0.958852	0
8	37	54	0.826648	0
8	38	54	0.728926	-0.032382
8	39	54	0.563796	-0.186406
8	40	54	0.576944	-0.25191
8	41	54	0.667259	0.471895
8	42	54	0.588519	0.095842
8	43	54	0.722944	-0.011061
8	44	54	0.866019	-0.009596
8	45	54	0.938204	-0.00107
9	1	54	0.962111	0
9	2	54	0.680111	0.000387
9	3	54	0.515333	0.028887
9	4	54	0.715481	0.496356
9	5	54	0.951667	0.894639
9	6	54	0.95287	0
9	7	54	0.738981	-0.023442

Table 36 Decision Accuracy and Relative Decision Accuracy Data of Estimators 3, 8, and 9 (cont.)

Estimator	Policy No	Treat-ment	Mean DA	Mean RDA		Estimator	Policy No	Treat-ment	Mean DA	Mean RDA
3	30	54	0.767815	-0.103723		9	8	54	0.573185	-0.045346
3	31	54	0.959426	-0.005257		9	9	54	0.625481	0.137719
3	32	54	0.747574	-0.009013		9	10	54	0.775778	0.344272
3	33	54	0.685574	0.009863		9	11	54	0.894389	0.878607
3	34	54	0.599111	-0.025734		9	12	54	0.479444	0.362224
3	35	54	0.527074	-0.277925		9	13	54	0.305074	-0.007115
3	36	54	0.962185	-0.000165		9	14	54	0.458111	-0.076906
3	37	54	0.811148	-0.014147		9	15	54	0.717241	-0.02305
3	38	54	0.746611	-0.012096		9	16	54	0.699037	0.408655
3	39	54	0.6995	-0.040958		9	17	54	0.564574	0.141463
3	40	54	0.701204	-0.105153		9	18	54	0.578167	0.001351
3	41	54	0.6085	0.352205		9	19	54	0.669833	-0.037754
3	42	54	0.623741	0.107317		9	20	54	0.769611	-0.035717
3	43	54	0.743815	0.003864		9	21	54	0.484481	-0.392259
3	44	54	0.844667	-0.024694		9	22	54	0.505889	-0.247645
3	45	54	0.906685	-0.027255		9	23	54	0.616944	-0.020979
8	1	54	0.963796	0		9	24	54	0.575185	-0.035285
8	2	54	0.677111	0		9	25	54	0.5565	-0.072134
8	3	54	0.487352	0.001223		9	26	54	0.495185	-0.318481
8	4	54	0.687889	0.472516		9	27	54	0.586685	-0.179482
8	5	54	0.962296	0.904695		9	28	54	0.640926	-0.1418
8	6	54	0.955074	0		9	29	54	0.6955	-0.133699
8	7	54	0.756481	0		9	30	54	0.745926	-0.125156
8	8	54	0.59363	-0.004681		9	31	54	0.964741	0
8	9	54	0.616519	0.141538		9	32	54	0.754481	-0.003664
8	10	54	0.781222	0.340649		9	33	54	0.611556	-0.064772
8	11	54	0.885463	0.867327		9	34	54	0.5	-0.130329
8	12	54	0.396537	0.287256		9	35	54	0.502778	-0.229021
8	13	54	0.305352	-0.026731		9	36	54	0.960944	0
8	14	54	0.547704	-0.003234		9	37	54	0.790907	-0.042199
8	15	54	0.738667	-0.000017		9	38	54	0.666037	-0.103427
8	16	54	0.666315	0.402154		9	39	54	0.58137	-0.178256
8	17	54	0.557889	0.145438		9	40	54	0.6165	-0.21425
8	18	54	0.583204	0.006201		9	41	54	0.686111	0.483154
8	19	54	0.697278	-0.009765		9	42	54	0.553759	0.075204
8	20	54	0.801907	-0.004887		9	43	54	0.703481	-0.023506
8	21	54	0.390426	-0.52516		9	44	54	0.836519	-0.036001
8	22	54	0.487593	-0.268267		9	45	54	0.921796	-0.013809
8	23	54	0.661519	0.020159						

APPENDIX D

SIMULATED OUTPUT MEASURES DATA OF REINSPECTION DECISION POLICIES THROUGH ALL TREATMENTS

Table 37 Output Measures Data of Reinspection Decision Policies

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
1	1	121.365	0.000	0.466
1	2	97.158	0.000	0.368
1	3	75.930	0.000	0.296
1	4	194.929	0.000	1.544
1	5	169.795	0.000	1.360
1	6	145.021	0.000	1.148
1	7	239.830	0.000	2.554
1	8	207.280	0.000	2.202
1	9	176.905	0.000	1.866
1	10	112.110	0.000	0.878
1	11	85.320	0.000	0.692
1	12	61.474	0.000	0.484
1	13	169.818	0.000	1.772
1	14	129.510	0.000	1.332
1	15	96.765	0.000	1.000
1	16	200.128	0.000	0.826
1	17	187.998	0.000	0.764
1	18	177.161	0.000	0.720
1	19	93.681	0.000	0.360
1	20	67.371	0.000	0.252
1	21	45.410	0.000	0.166

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
24	28	187.662	0.000	1.558
24	29	171.032	0.002	1.422
24	30	153.993	0.004	1.274
24	31	103.461	0.026	0.270
24	32	85.926	0.070	0.220
24	33	69.053	0.106	0.174
24	34	143.955	0.002	0.862
24	35	128.685	0.002	0.772
24	36	115.671	0.004	0.702
24	37	141.202	0.250	1.116
24	38	125.736	0.160	1.016
24	39	114.879	0.084	0.952
24	40	117.029	0.472	0.268
24	41	102.031	0.540	0.208
24	42	89.537	0.602	0.166
24	43	81.407	0.832	0.192
24	44	66.947	0.812	0.132
24	45	54.873	0.730	0.098
24	46	86.641	0.362	0.412
24	47	65.374	0.378	0.266
24	48	49.091	0.402	0.186

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
1	22	227.052	0.000	1.836
1	23	212.298	0.000	1.722
1	24	198.816	0.000	1.600
1	25	95.760	0.000	0.982
1	26	68.846	0.000	0.728
1	27	49.195	0.000	0.528
1	28	187.662	0.000	1.558
1	29	171.019	0.000	1.422
1	30	154.051	0.000	1.276
1	31	103.188	0.000	0.270
1	32	85.471	0.000	0.224
1	33	68.599	0.000	0.182
1	34	143.955	0.000	0.862
1	35	128.613	0.000	0.772
1	36	115.546	0.000	0.702
1	37	140.295	0.000	1.168
1	38	124.804	0.000	1.046
1	39	114.017	0.000	0.964
1	40	115.420	0.000	0.318
1	41	101.659	0.000	0.274
1	42	88.452	0.000	0.240
1	43	87.545	0.000	0.472
1	44	65.862	0.000	0.362
1	45	47.402	0.000	0.272
1	46	89.673	0.000	0.498
1	47	67.247	0.000	0.352
1	48	48.987	0.000	0.254
1	49	66.426	0.000	0.546
1	50	53.009	0.000	0.442
1	51	45.896	0.000	0.388
1	52	75.198	0.000	0.218
1	53	60.950	0.000	0.182
1	54	49.225	0.000	0.158
2	1	121.365	0.000	0.466
2	2	97.158	0.000	0.368
2	3	75.930	0.000	0.296
2	4	194.929	0.000	1.544
2	5	169.795	0.000	1.360
2	6	145.021	0.000	1.148
2	7	239.830	0.000	2.554
2	8	207.280	0.000	2.202
2	9	176.905	0.000	1.866
2	10	112.110	0.000	0.878
24	49	72.139	0.360	0.482
24	50	55.879	0.102	0.438
24	51	46.188	0.010	0.388
24	52	85.454	0.656	0.166
24	53	71.392	0.582	0.132
24	54	59.948	0.460	0.126
25	1	121.365	0.000	0.466
25	2	97.158	0.000	0.368
25	3	75.974	0.002	0.296
25	4	194.929	0.000	1.544
25	5	169.795	0.000	1.360
25	6	145.021	0.000	1.148
25	7	239.830	0.000	2.554
25	8	207.280	0.000	2.202
25	9	176.905	0.000	1.866
25	10	103.842	0.484	0.700
25	11	80.573	0.350	0.546
25	12	61.071	0.182	0.434
25	13	148.331	0.494	1.330
25	14	108.936	0.610	0.874
25	15	84.100	0.534	0.652
25	16	200.445	0.016	0.824
25	17	188.122	0.014	0.762
25	18	177.420	0.014	0.718
25	19	89.527	0.424	0.286
25	20	65.892	0.388	0.206
25	21	46.434	0.330	0.140
25	22	227.577	0.070	1.822
25	23	212.821	0.086	1.704
25	24	199.529	0.102	1.580
25	25	94.317	0.426	0.752
25	26	71.369	0.234	0.636
25	27	51.059	0.074	0.516
25	28	187.662	0.000	1.558
25	29	171.019	0.000	1.422
25	30	154.051	0.000	1.276
25	31	103.188	0.000	0.270
25	32	85.471	0.000	0.224
25	33	68.639	0.002	0.182
25	34	143.955	0.000	0.862
25	35	128.613	0.000	0.772
25	36	115.546	0.000	0.702
25	37	140.657	0.080	1.156

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
2	11	85.320	0.000	0.692
2	12	61.474	0.000	0.484
2	13	169.818	0.000	1.772
2	14	129.510	0.000	1.332
2	15	96.765	0.000	1.000
2	16	200.128	0.000	0.826
2	17	187.998	0.000	0.764
2	18	177.161	0.000	0.720
2	19	93.337	0.024	0.356
2	20	67.355	0.004	0.252
2	21	45.410	0.000	0.166
2	22	227.066	0.108	1.802
2	23	212.183	0.044	1.706
2	24	198.051	0.030	1.580
2	25	95.814	0.002	0.982
2	26	68.846	0.000	0.728
2	27	49.195	0.000	0.528
2	28	187.662	0.000	1.558
2	29	171.019	0.000	1.422
2	30	154.051	0.000	1.276
2	31	103.188	0.000	0.270
2	32	85.471	0.000	0.224
2	33	68.599	0.000	0.182
2	34	143.955	0.000	0.862
2	35	128.613	0.000	0.772
2	36	115.546	0.000	0.702
2	37	140.295	0.000	1.168
2	38	124.804	0.000	1.046
2	39	114.017	0.000	0.964
2	40	115.420	0.000	0.318
2	41	101.659	0.000	0.274
2	42	88.452	0.000	0.240
2	43	87.545	0.000	0.472
2	44	65.862	0.000	0.362
2	45	47.402	0.000	0.272
2	46	89.514	0.082	0.492
2	47	67.290	0.008	0.352
2	48	48.987	0.000	0.254
2	49	66.426	0.000	0.546
2	50	53.009	0.000	0.442
2	51	45.896	0.000	0.388
2	52	75.107	0.008	0.214
2	53	60.950	0.000	0.182

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
25	38	125.129	0.052	1.038
25	39	114.168	0.020	0.960
25	40	115.989	0.152	0.296
25	41	102.464	0.252	0.236
25	42	89.490	0.302	0.200
25	43	84.722	0.590	0.274
25	44	68.026	0.598	0.208
25	45	53.559	0.466	0.170
25	46	88.175	0.400	0.422
25	47	67.303	0.448	0.288
25	48	50.382	0.424	0.186
25	49	69.904	0.180	0.526
25	50	54.030	0.032	0.442
25	51	45.966	0.002	0.388
25	52	82.348	0.428	0.178
25	53	67.545	0.326	0.158
25	54	55.280	0.240	0.144
26	1	109.374	0.518	0.350
26	2	90.657	0.468	0.290
26	3	74.295	0.404	0.242
26	4	160.207	0.712	1.020
26	5	133.175	0.764	0.816
26	6	112.508	0.778	0.634
26	7	160.994	0.880	1.208
26	8	124.961	0.928	0.804
26	9	99.596	0.952	0.520
26	10	106.926	0.186	0.796
26	11	82.348	0.216	0.622
26	12	60.914	0.268	0.418
26	13	157.784	0.196	1.550
26	14	119.147	0.224	1.134
26	15	88.914	0.272	0.818
26	16	202.989	0.812	0.642
26	17	188.709	0.770	0.578
26	18	175.711	0.716	0.530
26	19	89.349	0.406	0.286
26	20	65.606	0.330	0.200
26	21	45.810	0.268	0.134
26	22	224.610	0.336	1.692
26	23	210.530	0.310	1.606
26	24	196.165	0.296	1.488
26	25	91.890	0.324	0.794
26	26	69.157	0.414	0.520

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
2	54	49.225	0.000	0.158
3	1	121.365	0.000	0.466
3	2	97.158	0.000	0.368
3	3	75.930	0.000	0.296
3	4	194.929	0.000	1.544
3	5	169.795	0.000	1.360
3	6	145.021	0.000	1.148
3	7	239.830	0.000	2.554
3	8	207.280	0.000	2.202
3	9	176.905	0.000	1.866
3	10	107.671	0.224	0.798
3	11	84.395	0.044	0.668
3	12	61.422	0.006	0.482
3	13	140.220	0.576	1.208
3	14	113.916	0.356	1.024
3	15	93.235	0.148	0.904
3	16	202.071	0.480	0.718
3	17	188.243	0.300	0.704
3	18	176.406	0.198	0.670
3	19	89.090	0.524	0.284
3	20	66.431	0.216	0.230
3	21	45.610	0.050	0.160
3	22	228.032	0.506	1.680
3	23	211.936	0.414	1.588
3	24	196.631	0.338	1.468
3	25	95.453	0.124	0.924
3	26	68.992	0.020	0.722
3	27	49.112	0.002	0.526
3	28	187.662	0.000	1.558
3	29	171.019	0.000	1.422
3	30	154.051	0.000	1.276
3	31	103.188	0.000	0.270
3	32	85.471	0.000	0.224
3	33	68.599	0.000	0.182
3	34	143.955	0.000	0.862
3	35	128.613	0.000	0.772
3	36	115.546	0.000	0.702
3	37	140.237	0.166	1.132
3	38	124.775	0.038	1.038
3	39	114.093	0.004	0.964
3	40	115.412	0.676	0.248
3	41	100.581	0.510	0.216
3	42	88.158	0.350	0.204
26	27	56.412	0.428	0.386
26	28	171.376	0.768	1.228
26	29	152.169	0.792	1.074
26	30	132.457	0.822	0.892
26	31	93.041	0.544	0.174
26	32	77.452	0.528	0.140
26	33	63.143	0.534	0.108
26	34	142.786	0.662	0.644
26	35	131.226	0.650	0.590
26	36	120.679	0.588	0.552
26	37	140.127	0.182	1.124
26	38	125.408	0.206	1.002
26	39	114.564	0.284	0.896
26	40	114.135	0.668	0.232
26	41	99.727	0.624	0.188
26	42	87.489	0.560	0.164
26	43	81.826	0.624	0.266
26	44	65.720	0.710	0.158
26	45	54.610	0.718	0.110
26	46	88.857	0.084	0.478
26	47	67.052	0.068	0.338
26	48	49.527	0.102	0.244
26	49	69.486	0.206	0.514
26	50	58.853	0.264	0.412
26	51	51.678	0.270	0.346
26	52	81.131	0.490	0.166
26	53	67.672	0.438	0.130
26	54	56.815	0.406	0.112
27	1	114.289	0.236	0.396
27	2	93.111	0.198	0.316
27	3	74.175	0.166	0.258
27	4	184.554	0.226	1.388
27	5	158.445	0.254	1.184
27	6	136.081	0.260	1.002
27	7	203.608	0.468	1.926
27	8	159.060	0.598	1.344
27	9	124.068	0.702	0.912
27	10	105.874	0.324	0.754
27	11	80.794	0.312	0.570
27	12	59.843	0.282	0.408
27	13	142.325	0.482	1.252
27	14	105.683	0.548	0.854
27	15	78.048	0.618	0.544

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
3	43	83.555	0.484	0.308
3	44	66.083	0.204	0.310
3	45	48.323	0.052	0.268
3	46	86.719	0.654	0.372
3	47	66.711	0.390	0.292
3	48	49.324	0.148	0.234
3	49	66.965	0.028	0.542
3	50	53.009	0.000	0.442
3	51	45.896	0.000	0.388
3	52	79.905	0.358	0.182
3	53	63.373	0.150	0.170
3	54	49.794	0.040	0.152
4	1	105.806	0.700	0.320
4	2	91.554	0.402	0.288
4	3	75.310	0.128	0.272
4	4	149.980	0.866	0.874
4	5	130.585	0.768	0.762
4	6	119.429	0.606	0.738
4	7	152.706	0.964	1.064
4	8	124.908	0.916	0.794
4	9	105.172	0.868	0.600
4	10	92.633	0.944	0.510
4	11	72.220	0.834	0.372
4	12	58.707	0.534	0.334
4	13	117.434	0.990	0.790
4	14	88.573	0.978	0.504
4	15	70.836	0.918	0.360
4	16	203.929	0.906	0.622
4	17	189.516	0.866	0.556
4	18	176.151	0.800	0.522
4	19	84.007	0.974	0.206
4	20	62.974	0.890	0.130
4	21	46.747	0.710	0.094
4	22	228.526	0.906	1.560
4	23	211.173	0.874	1.420
4	24	194.496	0.842	1.276
4	25	87.633	0.832	0.490
4	26	70.255	0.530	0.462
4	27	53.342	0.230	0.466
4	28	168.576	0.910	1.174
4	29	149.040	0.886	1.004
4	30	131.517	0.828	0.870
4	31	86.112	0.932	0.114

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
27	16	202.028	0.560	0.690
27	17	188.133	0.508	0.636
27	18	175.564	0.482	0.578
27	19	89.344	0.334	0.294
27	20	65.479	0.284	0.214
27	21	45.412	0.240	0.140
27	22	227.016	0.322	1.742
27	23	211.287	0.310	1.610
27	24	197.822	0.324	1.492
27	25	90.443	0.516	0.674
27	26	71.161	0.514	0.496
27	27	56.088	0.462	0.362
27	28	183.253	0.310	1.458
27	29	164.060	0.408	1.282
27	30	144.108	0.472	1.090
27	31	96.920	0.278	0.214
27	32	80.684	0.220	0.176
27	33	65.508	0.204	0.142
27	34	143.431	0.174	0.796
27	35	129.282	0.142	0.728
27	36	116.463	0.128	0.666
27	37	140.065	0.284	1.102
27	38	125.731	0.314	0.980
27	39	116.516	0.266	0.930
27	40	114.396	0.506	0.248
27	41	99.721	0.432	0.206
27	42	86.642	0.400	0.176
27	43	83.001	0.600	0.278
27	44	66.222	0.624	0.186
27	45	53.210	0.562	0.148
27	46	88.749	0.200	0.466
27	47	66.579	0.188	0.316
27	48	49.269	0.178	0.228
27	49	71.074	0.318	0.480
27	50	59.554	0.292	0.408
27	51	50.868	0.212	0.362
27	52	79.842	0.350	0.184
27	53	65.286	0.280	0.156
27	54	53.613	0.266	0.124
28	1	114.983	0.216	0.406
28	2	93.228	0.174	0.320
28	3	74.205	0.144	0.260
28	4	190.256	0.078	1.468

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
4	32	71.597	0.858	0.082
4	33	61.158	0.714	0.090
4	34	142.272	0.700	0.638
4	35	129.891	0.486	0.626
4	36	117.705	0.238	0.650
4	37	140.502	0.858	0.972
4	38	127.343	0.668	0.914
4	39	116.504	0.356	0.900
4	40	115.080	0.984	0.206
4	41	100.264	0.976	0.160
4	42	86.230	0.972	0.108
4	43	78.260	0.976	0.130
4	44	64.817	0.954	0.074
4	45	55.166	0.842	0.076
4	46	84.662	0.984	0.290
4	47	65.271	0.954	0.178
4	48	50.550	0.868	0.116
4	49	74.738	0.508	0.464
4	50	57.483	0.178	0.428
4	51	46.350	0.016	0.388
4	52	86.915	0.922	0.120
4	53	74.238	0.812	0.102
4	54	61.863	0.624	0.106
5	1	98.770	1.000	0.266
5	2	83.439	1.000	0.204
5	3	72.611	0.990	0.162
5	4	142.878	1.000	0.774
5	5	120.287	1.000	0.618
5	6	101.253	1.000	0.448
5	7	148.617	1.000	0.998
5	8	116.961	1.000	0.662
5	9	94.832	1.000	0.432
5	10	90.949	1.000	0.482
5	11	70.712	0.994	0.326
5	12	56.774	0.966	0.224
5	13	116.704	1.000	0.776
5	14	87.964	1.000	0.488
5	15	68.324	0.996	0.300
5	16	204.354	0.984	0.606
5	17	189.646	0.984	0.536
5	18	175.792	0.974	0.482
5	19	83.722	1.000	0.200
5	20	62.098	0.996	0.112

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
28	5	166.630	0.056	1.308
28	6	143.505	0.046	1.120
28	7	230.527	0.138	2.388
28	8	193.030	0.208	1.938
28	9	159.555	0.280	1.548
28	10	106.927	0.260	0.780
28	11	81.475	0.226	0.598
28	12	60.629	0.192	0.426
28	13	142.404	0.568	1.248
28	14	106.225	0.614	0.830
28	15	78.599	0.656	0.534
28	16	201.895	0.490	0.710
28	17	188.156	0.442	0.650
28	18	175.295	0.414	0.592
28	19	90.522	0.208	0.312
28	20	66.481	0.164	0.226
28	21	45.493	0.136	0.148
28	22	227.480	0.220	1.778
28	23	211.888	0.176	1.662
28	24	197.660	0.164	1.538
28	25	90.837	0.534	0.670
28	26	70.989	0.442	0.518
28	27	54.837	0.332	0.430
28	28	186.490	0.088	1.534
28	29	169.363	0.110	1.382
28	30	151.633	0.122	1.224
28	31	97.220	0.264	0.220
28	32	81.352	0.200	0.188
28	33	66.033	0.172	0.152
28	34	143.057	0.082	0.818
28	35	128.572	0.054	0.748
28	36	115.456	0.040	0.682
28	37	140.944	0.256	1.112
28	38	126.205	0.222	1.004
28	39	115.193	0.152	0.940
28	40	113.958	0.396	0.254
28	41	99.417	0.314	0.214
28	42	86.747	0.272	0.190
28	43	86.702	0.360	0.378
28	44	66.797	0.352	0.266
28	45	50.689	0.304	0.200
28	46	88.365	0.240	0.446
28	47	66.588	0.262	0.300

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
5	21	47.343	0.980	0.068
5	22	228.388	0.982	1.534
5	23	210.746	0.976	1.386
5	24	193.732	0.984	1.234
5	25	85.181	0.994	0.376
5	26	72.171	0.948	0.274
5	27	61.220	0.782	0.248
5	28	166.690	1.000	1.138
5	29	145.863	1.000	0.954
5	30	127.180	1.000	0.788
5	31	84.858	1.000	0.104
5	32	69.081	1.000	0.058
5	33	57.379	1.000	0.044
5	34	140.894	1.000	0.544
5	35	130.881	0.996	0.488
5	36	122.006	0.988	0.440
5	37	140.216	0.980	0.938
5	38	128.124	0.970	0.850
5	39	118.622	0.914	0.786
5	40	114.919	1.000	0.204
5	41	99.961	1.000	0.154
5	42	86.425	1.000	0.108
5	43	77.885	1.000	0.124
5	44	64.772	1.000	0.066
5	45	56.325	0.994	0.038
5	46	84.415	1.000	0.284
5	47	65.480	0.998	0.176
5	48	51.231	0.992	0.106
5	49	80.361	0.930	0.374
5	50	69.171	0.728	0.360
5	51	56.239	0.442	0.334
5	52	87.863	0.996	0.116
5	53	76.567	0.988	0.082
5	54	67.860	0.956	0.060
6	1	121.365	0.000	0.466
6	2	97.158	0.000	0.368
6	3	75.930	0.000	0.296
6	4	194.929	0.000	1.544
6	5	169.795	0.000	1.360
6	6	145.021	0.000	1.148
6	7	239.830	0.000	2.554
6	8	207.280	0.000	2.202
6	9	176.905	0.000	1.866

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
28	48	49.628	0.246	0.212
28	49	71.281	0.316	0.480
28	50	57.927	0.226	0.410
28	51	48.167	0.092	0.378
28	52	77.583	0.230	0.184
28	53	63.430	0.188	0.156
28	54	51.796	0.154	0.138
29	1	114.983	0.216	0.406
29	2	93.434	0.172	0.324
29	3	74.205	0.144	0.260
29	4	190.927	0.058	1.480
29	5	167.128	0.036	1.316
29	6	143.868	0.020	1.130
29	7	236.999	0.036	2.504
29	8	204.466	0.046	2.150
29	9	173.396	0.070	1.794
29	10	109.598	0.132	0.828
29	11	83.250	0.124	0.638
29	12	61.192	0.068	0.464
29	13	148.982	0.440	1.374
29	14	112.951	0.484	0.968
29	15	85.116	0.474	0.686
29	16	201.859	0.486	0.710
29	17	187.916	0.430	0.650
29	18	175.493	0.398	0.596
29	19	90.981	0.164	0.314
29	20	66.322	0.126	0.230
29	21	45.044	0.104	0.146
29	22	226.137	0.132	1.786
29	23	211.523	0.100	1.684
29	24	198.382	0.098	1.570
29	25	92.142	0.436	0.728
29	26	70.245	0.316	0.568
29	27	52.566	0.188	0.474
29	28	187.133	0.040	1.550
29	29	170.818	0.024	1.416
29	30	153.717	0.018	1.266
29	31	97.220	0.264	0.220
29	32	81.352	0.200	0.188
29	33	66.305	0.170	0.156
29	34	143.127	0.072	0.822
29	35	128.481	0.042	0.752
29	36	115.371	0.030	0.686

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
6	10	112.110	0.000	0.878
6	11	85.320	0.000	0.692
6	12	61.474	0.000	0.484
6	13	169.818	0.000	1.772
6	14	129.510	0.000	1.332
6	15	96.765	0.000	1.000
6	16	200.128	0.000	0.826
6	17	187.998	0.000	0.764
6	18	177.161	0.000	0.720
6	19	93.681	0.000	0.360
6	20	67.371	0.000	0.252
6	21	45.410	0.000	0.166
6	22	227.052	0.000	1.836
6	23	212.298	0.000	1.722
6	24	198.816	0.000	1.600
6	25	95.760	0.000	0.982
6	26	68.846	0.000	0.728
6	27	49.195	0.000	0.528
6	28	187.662	0.000	1.558
6	29	171.019	0.000	1.422
6	30	154.051	0.000	1.276
6	31	103.188	0.000	0.270
6	32	85.471	0.000	0.224
6	33	68.599	0.000	0.182
6	34	143.955	0.000	0.862
6	35	128.613	0.000	0.772
6	36	115.546	0.000	0.702
6	37	140.295	0.000	1.168
6	38	124.804	0.000	1.046
6	39	114.017	0.000	0.964
6	40	115.420	0.000	0.318
6	41	101.659	0.000	0.274
6	42	88.452	0.000	0.240
6	43	87.545	0.000	0.472
6	44	65.862	0.000	0.362
6	45	47.402	0.000	0.272
6	46	89.673	0.000	0.498
6	47	67.247	0.000	0.352
6	48	48.987	0.000	0.254
6	49	66.426	0.000	0.546
6	50	53.009	0.000	0.442
6	51	45.896	0.000	0.388
6	52	75.198	0.000	0.218
29	37	140.456	0.158	1.128
29	38	125.384	0.116	1.026
29	39	114.275	0.066	0.948
29	40	114.022	0.352	0.262
29	41	99.559	0.284	0.222
29	42	86.721	0.230	0.194
29	43	87.181	0.196	0.418
29	44	65.638	0.148	0.308
29	45	48.248	0.142	0.226
29	46	89.057	0.132	0.466
29	47	67.253	0.160	0.326
29	48	49.826	0.152	0.236
29	49	70.786	0.244	0.504
29	50	55.330	0.098	0.430
29	51	46.599	0.032	0.382
29	52	77.235	0.194	0.190
29	53	62.812	0.156	0.160
29	54	51.153	0.130	0.138
30	1	114.983	0.216	0.406
30	2	93.434	0.172	0.324
30	3	74.205	0.144	0.260
30	4	190.877	0.054	1.480
30	5	167.078	0.032	1.316
30	6	143.804	0.018	1.130
30	7	237.878	0.018	2.520
30	8	206.724	0.008	2.190
30	9	176.102	0.008	1.850
30	10	111.114	0.048	0.858
30	11	84.557	0.042	0.672
30	12	61.393	0.024	0.478
30	13	157.996	0.270	1.542
30	14	119.945	0.298	1.118
30	15	90.537	0.284	0.824
30	16	201.793	0.482	0.710
30	17	187.916	0.430	0.650
30	18	175.421	0.396	0.596
30	19	90.895	0.162	0.314
30	20	66.109	0.118	0.226
30	21	45.016	0.106	0.146
30	22	226.272	0.124	1.790
30	23	211.504	0.092	1.686
30	24	198.353	0.082	1.572
30	25	94.646	0.256	0.846

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
6	53	60.950	0.000	0.182
6	54	49.225	0.000	0.158
7	1	121.365	0.000	0.466
7	2	97.158	0.000	0.368
7	3	75.930	0.000	0.296
7	4	194.929	0.000	1.544
7	5	169.795	0.000	1.360
7	6	145.021	0.000	1.148
7	7	239.830	0.000	2.554
7	8	207.280	0.000	2.202
7	9	176.905	0.000	1.866
7	10	112.110	0.000	0.878
7	11	85.320	0.000	0.692
7	12	61.474	0.000	0.484
7	13	169.818	0.000	1.772
7	14	129.510	0.000	1.332
7	15	96.765	0.000	1.000
7	16	200.128	0.000	0.826
7	17	187.998	0.000	0.764
7	18	177.161	0.000	0.720
7	19	90.588	0.362	0.320
7	20	66.100	0.280	0.218
7	21	45.814	0.190	0.152
7	22	226.340	0.340	1.704
7	23	210.970	0.276	1.618
7	24	196.363	0.234	1.494
7	25	94.937	0.042	0.950
7	26	68.814	0.010	0.726
7	27	49.219	0.002	0.528
7	28	187.662	0.000	1.558
7	29	171.019	0.000	1.422
7	30	154.051	0.000	1.276
7	31	103.188	0.000	0.270
7	32	85.471	0.000	0.224
7	33	68.599	0.000	0.182
7	34	143.955	0.000	0.862
7	35	128.613	0.000	0.772
7	36	115.546	0.000	0.702
7	37	140.295	0.000	1.168
7	38	124.804	0.000	1.046
7	39	114.017	0.000	0.964
7	40	115.420	0.000	0.318
7	41	101.659	0.000	0.274

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
30	26	69.615	0.180	0.634
30	27	50.898	0.072	0.514
30	28	187.168	0.032	1.552
30	29	170.853	0.016	1.418
30	30	154.006	0.004	1.276
30	31	97.220	0.264	0.220
30	32	81.352	0.200	0.188
30	33	66.305	0.170	0.156
30	34	143.127	0.070	0.822
30	35	128.481	0.042	0.752
30	36	115.371	0.030	0.686
30	37	140.224	0.078	1.146
30	38	124.751	0.036	1.036
30	39	114.140	0.016	0.960
30	40	113.966	0.348	0.262
30	41	99.517	0.278	0.222
30	42	86.920	0.220	0.198
30	43	87.450	0.090	0.450
30	44	66.140	0.062	0.348
30	45	47.916	0.064	0.254
30	46	89.281	0.130	0.476
30	47	67.567	0.094	0.344
30	48	49.005	0.072	0.238
30	49	69.182	0.144	0.528
30	50	53.633	0.030	0.436
30	51	46.024	0.004	0.388
30	52	77.121	0.190	0.190
30	53	62.773	0.150	0.162
30	54	51.135	0.128	0.138
31	1	121.365	0.000	0.466
31	2	97.158	0.000	0.368
31	3	75.930	0.000	0.296
31	4	194.929	0.000	1.544
31	5	169.795	0.000	1.360
31	6	145.021	0.000	1.148
31	7	239.830	0.000	2.554
31	8	207.280	0.000	2.202
31	9	176.905	0.000	1.866
31	10	112.110	0.000	0.878
31	11	85.320	0.000	0.692
31	12	61.474	0.000	0.484
31	13	169.818	0.000	1.772
31	14	129.510	0.000	1.332

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
7	42	88.452	0.000	0.240
7	43	87.545	0.000	0.472
7	44	65.862	0.000	0.362
7	45	47.402	0.000	0.272
7	46	89.053	0.266	0.458
7	47	67.411	0.152	0.330
7	48	48.916	0.074	0.240
7	49	66.539	0.006	0.546
7	50	53.073	0.002	0.442
7	51	45.966	0.002	0.388
7	52	78.974	0.298	0.190
7	53	64.394	0.230	0.156
7	54	52.591	0.164	0.144
8	1	121.365	0.000	0.466
8	2	97.158	0.000	0.368
8	3	75.930	0.000	0.296
8	4	194.929	0.000	1.544
8	5	169.795	0.000	1.360
8	6	145.021	0.000	1.148
8	7	239.830	0.000	2.554
8	8	207.280	0.000	2.202
8	9	176.905	0.000	1.866
8	10	105.036	0.350	0.746
8	11	81.844	0.214	0.600
8	12	60.746	0.106	0.448
8	13	143.466	0.512	1.280
8	14	113.770	0.372	1.008
8	15	90.627	0.204	0.850
8	16	201.247	0.312	0.760
8	17	188.550	0.310	0.698
8	18	177.089	0.266	0.668
8	19	90.222	0.456	0.308
8	20	65.666	0.386	0.206
8	21	45.853	0.272	0.144
8	22	227.315	0.482	1.680
8	23	211.475	0.416	1.578
8	24	196.421	0.398	1.444
8	25	93.199	0.258	0.824
8	26	68.557	0.108	0.668
8	27	49.552	0.024	0.520
8	28	187.662	0.000	1.558
8	29	171.019	0.000	1.422
8	30	154.051	0.000	1.276
31	15	96.765	0.000	1.000
31	16	200.128	0.000	0.826
31	17	187.998	0.000	0.764
31	18	177.161	0.000	0.720
31	19	93.681	0.000	0.360
31	20	67.371	0.000	0.252
31	21	45.410	0.000	0.166
31	22	227.052	0.000	1.836
31	23	212.298	0.000	1.722
31	24	198.816	0.000	1.600
31	25	95.760	0.000	0.982
31	26	68.846	0.000	0.728
31	27	49.195	0.000	0.528
31	28	187.662	0.000	1.558
31	29	171.019	0.000	1.422
31	30	154.051	0.000	1.276
31	31	103.188	0.000	0.270
31	32	85.471	0.000	0.224
31	33	68.599	0.000	0.182
31	34	143.955	0.000	0.862
31	35	128.613	0.000	0.772
31	36	115.546	0.000	0.702
31	37	140.295	0.000	1.168
31	38	124.804	0.000	1.046
31	39	114.017	0.000	0.964
31	40	115.420	0.000	0.318
31	41	101.659	0.000	0.274
31	42	88.452	0.000	0.240
31	43	87.545	0.000	0.472
31	44	65.862	0.000	0.362
31	45	47.402	0.000	0.272
31	46	89.673	0.000	0.498
31	47	67.247	0.000	0.352
31	48	48.987	0.000	0.254
31	49	66.426	0.000	0.546
31	50	53.009	0.000	0.442
31	51	45.896	0.000	0.388
31	52	75.198	0.000	0.218
31	53	60.950	0.000	0.182
31	54	49.225	0.000	0.158
32	1	121.365	0.000	0.466
32	2	97.158	0.000	0.368
32	3	75.930	0.000	0.296

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
8	31	103.188	0.000	0.270
8	32	85.471	0.000	0.224
8	33	68.599	0.000	0.182
8	34	143.955	0.000	0.862
8	35	128.613	0.000	0.772
8	36	115.546	0.000	0.702
8	37	139.152	0.276	1.080
8	38	124.678	0.130	1.010
8	39	114.437	0.048	0.958
8	40	115.870	0.434	0.276
8	41	102.353	0.424	0.236
8	42	88.491	0.408	0.196
8	43	82.857	0.472	0.302
8	44	64.896	0.334	0.254
8	45	49.351	0.226	0.218
8	46	87.486	0.574	0.388
8	47	66.720	0.468	0.270
8	48	50.030	0.340	0.204
8	49	68.409	0.126	0.524
8	50	53.452	0.020	0.440
8	51	46.164	0.008	0.388
8	52	80.507	0.382	0.186
8	53	65.543	0.284	0.154
8	54	53.800	0.222	0.138
9	1	111.395	0.494	0.376
9	2	91.652	0.408	0.296
9	3	75.626	0.286	0.266
9	4	159.502	0.668	1.014
9	5	139.673	0.610	0.898
9	6	122.027	0.516	0.774
9	7	169.475	0.792	1.344
9	8	137.632	0.782	0.996
9	9	115.764	0.744	0.788
9	10	94.277	0.798	0.534
9	11	75.401	0.646	0.444
9	12	59.282	0.512	0.346
9	13	121.350	0.914	0.860
9	14	92.603	0.884	0.578
9	15	72.733	0.806	0.406
9	16	201.843	0.368	0.754
9	17	189.375	0.390	0.684
9	18	177.643	0.372	0.648
9	19	88.861	0.630	0.280

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
32	4	194.929	0.000	1.544
32	5	169.795	0.000	1.360
32	6	145.021	0.000	1.148
32	7	239.830	0.000	2.554
32	8	207.280	0.000	2.202
32	9	176.905	0.000	1.866
32	10	112.110	0.000	0.878
32	11	85.320	0.000	0.692
32	12	61.474	0.000	0.484
32	13	169.818	0.000	1.772
32	14	129.510	0.000	1.332
32	15	96.765	0.000	1.000
32	16	200.128	0.000	0.826
32	17	187.998	0.000	0.764
32	18	177.161	0.000	0.720
32	19	93.337	0.024	0.356
32	20	67.355	0.004	0.252
32	21	45.410	0.000	0.166
32	22	227.066	0.108	1.802
32	23	212.183	0.044	1.706
32	24	198.051	0.030	1.580
32	25	95.814	0.002	0.982
32	26	68.846	0.000	0.728
32	27	49.195	0.000	0.528
32	28	187.662	0.000	1.558
32	29	171.019	0.000	1.422
32	30	154.051	0.000	1.276
32	31	103.188	0.000	0.270
32	32	85.471	0.000	0.224
32	33	68.599	0.000	0.182
32	34	143.955	0.000	0.862
32	35	128.613	0.000	0.772
32	36	115.546	0.000	0.702
32	37	140.295	0.000	1.168
32	38	124.804	0.000	1.046
32	39	114.017	0.000	0.964
32	40	115.420	0.000	0.318
32	41	101.659	0.000	0.274
32	42	88.452	0.000	0.240
32	43	87.545	0.000	0.472
32	44	65.862	0.000	0.362
32	45	47.402	0.000	0.272
32	46	89.514	0.082	0.492

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
9	20	65.519	0.580	0.190
9	21	46.512	0.480	0.124
9	22	228.190	0.662	1.636
9	23	212.041	0.650	1.514
9	24	196.643	0.654	1.384
9	25	88.218	0.726	0.554
9	26	70.514	0.514	0.474
9	27	54.168	0.326	0.422
9	28	171.200	0.774	1.216
9	29	153.473	0.736	1.086
9	30	136.634	0.670	0.956
9	31	93.554	0.560	0.178
9	32	76.485	0.554	0.122
9	33	63.001	0.500	0.102
9	34	142.789	0.554	0.702
9	35	129.462	0.456	0.638
9	36	116.553	0.306	0.604
9	37	139.716	0.722	0.988
9	38	126.815	0.574	0.928
9	39	116.437	0.390	0.892
9	40	116.909	0.550	0.274
9	41	103.080	0.572	0.224
9	42	88.931	0.572	0.180
9	43	79.517	0.854	0.170
9	44	65.398	0.802	0.122
9	45	53.251	0.676	0.110
9	46	85.341	0.828	0.320
9	47	65.744	0.794	0.214
9	48	50.760	0.686	0.150
9	49	74.226	0.510	0.450
9	50	59.142	0.284	0.404
9	51	48.769	0.118	0.374
9	52	82.926	0.546	0.174
9	53	69.541	0.488	0.136
9	54	58.113	0.420	0.126
10	1	107.652	0.682	0.344
10	2	88.595	0.682	0.266
10	3	75.024	0.612	0.234
10	4	149.016	0.902	0.868
10	5	124.731	0.932	0.692
10	6	105.061	0.940	0.510
10	7	151.963	0.968	1.054
10	8	117.843	0.990	0.680
32	47	67.290	0.008	0.352
32	48	48.987	0.000	0.254
32	49	66.426	0.000	0.546
32	50	53.009	0.000	0.442
32	51	45.896	0.000	0.388
32	52	75.107	0.008	0.214
32	53	60.950	0.000	0.182
32	54	49.225	0.000	0.158
33	1	121.365	0.000	0.466
33	2	97.158	0.000	0.368
33	3	75.930	0.000	0.296
33	4	194.929	0.000	1.544
33	5	169.795	0.000	1.360
33	6	145.021	0.000	1.148
33	7	239.830	0.000	2.554
33	8	207.280	0.000	2.202
33	9	176.905	0.000	1.866
33	10	107.671	0.224	0.798
33	11	84.395	0.044	0.668
33	12	61.422	0.006	0.482
33	13	140.220	0.576	1.208
33	14	113.916	0.356	1.024
33	15	93.235	0.148	0.904
33	16	202.071	0.480	0.718
33	17	188.243	0.300	0.704
33	18	176.406	0.198	0.670
33	19	89.181	0.518	0.284
33	20	66.418	0.214	0.230
33	21	45.610	0.050	0.160
33	22	228.032	0.506	1.680
33	23	211.936	0.414	1.588
33	24	196.631	0.338	1.468
33	25	95.453	0.124	0.924
33	26	68.992	0.020	0.722
33	27	49.112	0.002	0.526
33	28	187.662	0.000	1.558
33	29	171.019	0.000	1.422
33	30	154.051	0.000	1.276
33	31	103.188	0.000	0.270
33	32	85.471	0.000	0.224
33	33	68.599	0.000	0.182
33	34	143.955	0.000	0.862
33	35	128.613	0.000	0.772

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
10	9	95.293	0.994	0.440
10	10	92.717	0.920	0.508
10	11	72.283	0.886	0.362
10	12	57.782	0.830	0.268
10	13	117.656	0.986	0.792
10	14	88.141	0.982	0.490
10	15	68.791	0.972	0.312
10	16	201.939	0.376	0.754
10	17	189.399	0.402	0.684
10	18	177.570	0.390	0.644
10	19	88.611	0.668	0.276
10	20	65.324	0.650	0.180
10	21	47.118	0.576	0.118
10	22	228.334	0.736	1.614
10	23	212.075	0.752	1.480
10	24	196.536	0.754	1.354
10	25	85.677	0.924	0.420
10	26	70.610	0.828	0.314
10	27	59.875	0.670	0.306
10	28	167.994	0.938	1.156
10	29	146.797	0.962	0.968
10	30	127.989	0.976	0.800
10	31	91.877	0.662	0.160
10	32	74.545	0.714	0.100
10	33	60.977	0.744	0.074
10	34	142.109	0.854	0.612
10	35	131.303	0.860	0.540
10	36	121.739	0.822	0.494
10	37	140.007	0.912	0.950
10	38	127.674	0.860	0.872
10	39	117.846	0.784	0.808
10	40	116.888	0.566	0.268
10	41	103.293	0.598	0.220
10	42	89.136	0.616	0.172
10	43	78.503	0.942	0.140
10	44	64.842	0.938	0.078
10	45	55.860	0.922	0.058
10	46	85.035	0.926	0.302
10	47	65.630	0.904	0.194
10	48	51.332	0.872	0.130
10	49	78.142	0.782	0.394
10	50	66.858	0.642	0.354
10	51	56.227	0.456	0.326

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
33	36	115.546	0.000	0.702
33	37	140.237	0.166	1.132
33	38	124.775	0.038	1.038
33	39	114.093	0.004	0.964
33	40	115.412	0.676	0.248
33	41	100.581	0.510	0.216
33	42	88.158	0.350	0.204
33	43	83.555	0.484	0.308
33	44	66.083	0.204	0.310
33	45	48.323	0.052	0.268
33	46	87.018	0.634	0.376
33	47	66.717	0.388	0.292
33	48	49.324	0.148	0.234
33	49	66.965	0.028	0.542
33	50	53.009	0.000	0.442
33	51	45.896	0.000	0.388
33	52	79.905	0.358	0.182
33	53	63.373	0.150	0.170
33	54	49.794	0.040	0.152
34	1	105.806	0.700	0.320
34	2	91.554	0.402	0.288
34	3	75.310	0.128	0.272
34	4	149.980	0.866	0.874
34	5	130.585	0.768	0.762
34	6	119.429	0.606	0.738
34	7	152.706	0.964	1.064
34	8	124.908	0.916	0.794
34	9	105.172	0.868	0.600
34	10	95.507	0.790	0.560
34	11	72.754	0.810	0.384
34	12	58.707	0.534	0.334
34	13	117.434	0.990	0.790
34	14	88.573	0.978	0.504
34	15	70.836	0.918	0.360
34	16	203.929	0.906	0.622
34	17	189.516	0.866	0.556
34	18	176.151	0.800	0.522
34	19	90.923	0.254	0.320
34	20	64.550	0.490	0.180
34	21	45.911	0.526	0.108
34	22	227.858	0.526	1.674
34	23	211.661	0.606	1.508
34	24	196.413	0.630	1.366

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
10	52	83.847	0.594	0.168
10	53	70.930	0.572	0.126
10	54	59.572	0.510	0.106
11	1	99.392	0.986	0.270
11	2	83.834	0.980	0.208
11	3	72.815	0.966	0.166
11	4	149.101	0.904	0.856
11	5	122.191	0.972	0.642
11	6	102.871	0.972	0.470
11	7	160.207	0.908	1.176
11	8	120.130	0.978	0.710
11	9	96.513	0.986	0.456
11	10	91.049	0.996	0.484
11	11	70.672	0.990	0.326
11	12	56.855	0.954	0.228
11	13	117.430	0.996	0.788
11	14	87.964	1.000	0.488
11	15	68.693	0.990	0.306
11	16	205.016	0.770	0.668
11	17	190.964	0.764	0.608
11	18	176.920	0.776	0.538
11	19	83.834	0.996	0.202
11	20	62.066	0.988	0.112
11	21	47.277	0.966	0.070
11	22	229.231	0.852	1.586
11	23	211.868	0.870	1.430
11	24	195.406	0.872	1.284
11	25	85.840	0.980	0.394
11	26	71.963	0.912	0.288
11	27	61.213	0.738	0.276
11	28	174.911	0.710	1.274
11	29	151.365	0.860	1.042
11	30	130.544	0.930	0.844
11	31	85.733	0.978	0.114
11	32	69.151	0.996	0.058
11	33	57.411	0.998	0.044
11	34	143.332	0.776	0.626
11	35	131.309	0.842	0.542
11	36	122.313	0.806	0.510
11	37	140.469	0.928	0.954
11	38	127.642	0.874	0.866
11	39	118.781	0.736	0.832
11	40	115.285	0.974	0.210

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
34	25	87.580	0.830	0.490
34	26	70.255	0.530	0.462
34	27	53.342	0.230	0.466
34	28	168.576	0.910	1.174
34	29	149.040	0.886	1.004
34	30	131.517	0.828	0.870
34	31	86.112	0.932	0.114
34	32	71.597	0.858	0.082
34	33	61.158	0.714	0.090
34	34	142.272	0.700	0.638
34	35	129.891	0.486	0.626
34	36	117.705	0.238	0.650
34	37	140.507	0.732	0.998
34	38	127.406	0.648	0.920
34	39	116.428	0.352	0.900
34	40	115.080	0.984	0.206
34	41	100.264	0.976	0.160
34	42	86.230	0.972	0.108
34	43	78.260	0.976	0.130
34	44	64.817	0.954	0.074
34	45	55.166	0.842	0.076
34	46	88.733	0.174	0.454
34	47	65.577	0.346	0.270
34	48	49.445	0.520	0.166
34	49	74.772	0.506	0.466
34	50	57.483	0.178	0.428
34	51	46.350	0.016	0.388
34	52	87.006	0.914	0.124
34	53	74.238	0.812	0.102
34	54	61.863	0.624	0.106
35	1	99.651	0.962	0.276
35	2	83.577	0.992	0.206
35	3	72.611	0.990	0.162
35	4	142.878	1.000	0.774
35	5	120.287	1.000	0.618
35	6	101.253	1.000	0.448
35	7	148.617	1.000	0.998
35	8	116.961	1.000	0.662
35	9	94.832	1.000	0.432
35	10	111.487	0.016	0.868
35	11	84.094	0.074	0.662
35	12	60.633	0.204	0.432
35	13	166.059	0.070	1.704

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
11	41	100.251	0.984	0.156
11	42	86.685	0.996	0.110
11	43	77.914	0.998	0.124
11	44	64.772	1.000	0.066
11	45	56.325	0.994	0.038
11	46	84.428	0.996	0.284
11	47	65.418	0.990	0.176
11	48	51.195	0.986	0.106
11	49	80.005	0.884	0.388
11	50	68.117	0.664	0.374
11	51	55.065	0.390	0.340
11	52	87.683	0.978	0.116
11	53	76.417	0.958	0.086
11	54	67.397	0.922	0.068
12	1	116.748	0.374	0.406
12	2	94.514	0.402	0.312
12	3	76.746	0.278	0.266
12	4	189.405	0.160	1.434
12	5	159.767	0.256	1.176
12	6	135.781	0.316	0.972
12	7	232.123	0.136	2.400
12	8	189.244	0.284	1.844
12	9	150.599	0.442	1.346
12	10	94.498	0.908	0.536
12	11	72.469	0.864	0.364
12	12	58.748	0.728	0.290
12	13	122.098	0.934	0.868
12	14	90.494	0.950	0.534
12	15	71.593	0.930	0.360
12	16	202.489	0.254	0.774
12	17	189.648	0.272	0.704
12	18	178.218	0.266	0.662
12	19	84.630	0.952	0.214
12	20	63.264	0.916	0.132
12	21	47.469	0.828	0.082
12	22	229.487	0.408	1.736
12	23	213.648	0.432	1.594
12	24	197.600	0.466	1.432
12	25	87.874	0.876	0.464
12	26	71.677	0.670	0.410
12	27	56.772	0.392	0.418
12	28	187.642	0.032	1.554
12	29	169.571	0.080	1.388

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
35	14	123.058	0.160	1.198
35	15	87.333	0.306	0.772
35	16	202.541	0.638	0.674
35	17	189.779	0.798	0.572
35	18	176.399	0.870	0.510
35	19	93.561	0.008	0.358
35	20	67.259	0.028	0.248
35	21	45.722	0.058	0.160
35	22	226.882	0.072	1.810
35	23	211.729	0.090	1.690
35	24	198.086	0.112	1.568
35	25	90.338	0.410	0.714
35	26	71.245	0.706	0.394
35	27	60.240	0.738	0.256
35	28	174.493	0.668	1.288
35	29	151.223	0.828	1.046
35	30	129.719	0.908	0.826
35	31	84.858	1.000	0.104
35	32	69.081	1.000	0.058
35	33	57.379	1.000	0.044
35	34	140.894	1.000	0.544
35	35	130.881	0.996	0.488
35	36	122.006	0.988	0.440
35	37	140.176	0.064	1.152
35	38	125.391	0.154	1.016
35	39	114.869	0.352	0.886
35	40	114.837	0.050	0.306
35	41	101.711	0.086	0.266
35	42	87.833	0.156	0.212
35	43	79.635	0.726	0.208
35	44	64.475	0.936	0.082
35	45	56.110	0.984	0.038
35	46	89.659	0.004	0.498
35	47	67.367	0.012	0.352
35	48	48.999	0.022	0.252
35	49	71.020	0.358	0.464
35	50	63.271	0.492	0.378
35	51	55.632	0.418	0.336
35	52	77.792	0.246	0.184
35	53	67.712	0.442	0.138
35	54	60.844	0.626	0.090
36	1	121.365	0.000	0.466
36	2	97.158	0.000	0.368

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
12	30	151.601	0.138	1.222	36	3	75.930	0.000	0.296
12	31	101.295	0.310	0.236	36	4	194.929	0.000	1.544
12	32	80.791	0.510	0.146	36	5	169.795	0.000	1.360
12	33	64.472	0.656	0.098	36	6	145.021	0.000	1.148
12	34	144.484	0.054	0.850	36	7	239.830	0.000	2.554
12	35	129.714	0.068	0.762	36	8	207.280	0.000	2.202
12	36	116.657	0.062	0.690	36	9	176.905	0.000	1.866
12	37	141.201	0.592	1.040	36	10	112.110	0.000	0.878
12	38	126.981	0.458	0.964	36	11	85.320	0.000	0.692
12	39	116.415	0.254	0.928	36	12	61.474	0.000	0.484
12	40	116.792	0.748	0.234	36	13	169.818	0.000	1.772
12	41	101.410	0.822	0.180	36	14	129.510	0.000	1.332
12	42	87.857	0.846	0.138	36	15	96.765	0.000	1.000
12	43	78.834	0.946	0.142	36	16	200.128	0.000	0.826
12	44	65.640	0.944	0.084	36	17	187.998	0.000	0.764
12	45	56.597	0.924	0.058	36	18	177.161	0.000	0.720
12	46	85.111	0.954	0.302	36	19	93.681	0.000	0.360
12	47	65.212	0.946	0.178	36	20	67.371	0.000	0.252
12	48	50.985	0.918	0.110	36	21	45.410	0.000	0.166
12	49	76.511	0.576	0.460	36	22	227.052	0.000	1.836
12	50	60.309	0.278	0.426	36	23	212.298	0.000	1.722
12	51	47.437	0.054	0.386	36	24	198.816	0.000	1.600
12	52	86.977	0.858	0.134	36	25	95.760	0.000	0.982
12	53	74.480	0.798	0.104	36	26	68.846	0.000	0.728
12	54	63.217	0.678	0.092	36	27	49.195	0.000	0.528
13	1	121.388	0.004	0.466	36	28	187.662	0.000	1.558
13	2	97.229	0.006	0.368	36	29	171.019	0.000	1.422
13	3	76.018	0.004	0.296	36	30	154.051	0.000	1.276
13	4	194.929	0.000	1.544	36	31	103.188	0.000	0.270
13	5	169.829	0.002	1.360	36	32	85.471	0.000	0.224
13	6	145.061	0.002	1.148	36	33	68.599	0.000	0.182
13	7	239.830	0.000	2.554	36	34	143.955	0.000	0.862
13	8	207.280	0.000	2.202	36	35	128.613	0.000	0.772
13	9	177.041	0.006	1.866	36	36	115.546	0.000	0.702
13	10	101.619	0.638	0.662	36	37	140.295	0.000	1.168
13	11	78.243	0.494	0.494	36	38	124.804	0.000	1.046
13	12	60.327	0.316	0.392	36	39	114.017	0.000	0.964
13	13	138.669	0.662	1.160	36	40	115.420	0.000	0.318
13	14	102.314	0.728	0.758	36	41	101.659	0.000	0.274
13	15	80.392	0.668	0.558	36	42	88.452	0.000	0.240
13	16	200.748	0.032	0.824	36	43	87.545	0.000	0.472
13	17	188.082	0.042	0.748	36	44	65.862	0.000	0.362
13	18	177.509	0.036	0.714	36	45	47.402	0.000	0.272

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
13	19	86.892	0.772	0.238	36	46	89.673	0.000	0.498
13	20	65.830	0.662	0.174	36	47	67.247	0.000	0.352
13	21	46.908	0.532	0.118	36	48	48.987	0.000	0.254
13	22	228.391	0.124	1.814	36	49	66.426	0.000	0.546
13	23	213.298	0.150	1.688	36	50	53.009	0.000	0.442
13	24	198.984	0.166	1.546	36	51	45.896	0.000	0.388
13	25	91.721	0.556	0.660	36	52	75.198	0.000	0.218
13	26	70.764	0.306	0.582	36	53	60.950	0.000	0.182
13	27	52.426	0.116	0.516	36	54	49.225	0.000	0.158
13	28	187.662	0.000	1.558	37	1	121.365	0.000	0.466
13	29	171.019	0.000	1.422	37	2	97.158	0.000	0.368
13	30	154.083	0.002	1.276	37	3	75.930	0.000	0.296
13	31	103.186	0.002	0.270	37	4	194.929	0.000	1.544
13	32	85.545	0.004	0.224	37	5	169.795	0.000	1.360
13	33	68.894	0.020	0.182	37	6	145.021	0.000	1.148
13	34	143.955	0.000	0.862	37	7	239.830	0.000	2.554
13	35	128.613	0.000	0.772	37	8	207.280	0.000	2.202
13	36	115.546	0.000	0.702	37	9	176.905	0.000	1.866
13	37	140.526	0.134	1.138	37	10	112.110	0.000	0.878
13	38	125.750	0.110	1.032	37	11	85.320	0.000	0.692
13	39	114.370	0.040	0.956	37	12	61.474	0.000	0.484
13	40	116.205	0.300	0.278	37	13	169.818	0.000	1.772
13	41	102.455	0.382	0.226	37	14	129.510	0.000	1.332
13	42	89.803	0.448	0.184	37	15	96.765	0.000	1.000
13	43	82.539	0.732	0.218	37	16	200.128	0.000	0.826
13	44	67.349	0.710	0.166	37	17	187.998	0.000	0.764
13	45	54.623	0.604	0.132	37	18	177.161	0.000	0.720
13	46	87.001	0.776	0.354	37	19	90.588	0.362	0.320
13	47	67.481	0.758	0.240	37	20	66.100	0.280	0.218
13	48	51.116	0.642	0.162	37	21	45.814	0.190	0.152
13	49	70.797	0.248	0.510	37	22	226.340	0.340	1.704
13	50	54.403	0.048	0.440	37	23	210.970	0.276	1.618
13	51	45.966	0.002	0.388	37	24	196.363	0.234	1.494
13	52	83.945	0.552	0.168	37	25	94.937	0.042	0.950
13	53	69.067	0.466	0.136	37	26	68.814	0.010	0.726
13	54	57.716	0.358	0.134	37	27	49.219	0.002	0.528
14	1	121.365	0.000	0.466	37	28	187.662	0.000	1.558
14	2	97.158	0.000	0.368	37	29	171.019	0.000	1.422
14	3	75.930	0.000	0.296	37	30	154.051	0.000	1.276
14	4	194.929	0.000	1.544	37	31	103.188	0.000	0.270
14	5	169.795	0.000	1.360	37	32	85.471	0.000	0.224
14	6	145.021	0.000	1.148	37	33	68.599	0.000	0.182
14	7	239.830	0.000	2.554	37	34	143.955	0.000	0.862

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
14	8	207.280	0.000	2.202	37	35	128.613	0.000	0.772
14	9	176.905	0.000	1.866	37	36	115.546	0.000	0.702
14	10	109.701	0.214	0.812	37	37	140.295	0.000	1.168
14	11	83.945	0.142	0.638	37	38	124.804	0.000	1.046
14	12	61.877	0.048	0.476	37	39	114.017	0.000	0.964
14	13	162.267	0.226	1.598	37	40	115.420	0.000	0.318
14	14	120.298	0.302	1.114	37	41	101.659	0.000	0.274
14	15	91.550	0.290	0.826	37	42	88.452	0.000	0.240
14	16	200.128	0.000	0.826	37	43	87.545	0.000	0.472
14	17	188.021	0.002	0.764	37	44	65.862	0.000	0.362
14	18	177.161	0.000	0.720	37	45	47.402	0.000	0.272
14	19	91.902	0.436	0.312	37	46	89.053	0.266	0.458
14	20	67.141	0.332	0.220	37	47	67.411	0.152	0.330
14	21	47.207	0.234	0.152	37	48	48.916	0.074	0.240
14	22	227.132	0.016	1.832	37	49	66.539	0.006	0.546
14	23	212.741	0.026	1.720	37	50	53.073	0.002	0.442
14	24	199.035	0.014	1.600	37	51	45.966	0.002	0.388
14	25	96.527	0.232	0.886	37	52	78.974	0.298	0.190
14	26	69.877	0.098	0.690	37	53	64.394	0.230	0.156
14	27	49.640	0.018	0.526	37	54	52.591	0.164	0.144
14	28	187.662	0.000	1.558	38	1	121.365	0.000	0.466
14	29	171.019	0.000	1.422	38	2	97.158	0.000	0.368
14	30	154.051	0.000	1.276	38	3	75.930	0.000	0.296
14	31	103.188	0.000	0.270	38	4	194.929	0.000	1.544
14	32	85.471	0.000	0.224	38	5	169.795	0.000	1.360
14	33	68.599	0.000	0.182	38	6	145.021	0.000	1.148
14	34	143.955	0.000	0.862	38	7	239.830	0.000	2.554
14	35	128.613	0.000	0.772	38	8	207.280	0.000	2.202
14	36	115.546	0.000	0.702	38	9	176.905	0.000	1.866
14	37	140.341	0.016	1.166	38	10	105.036	0.350	0.746
14	38	124.908	0.014	1.044	38	11	81.844	0.214	0.600
14	39	114.018	0.006	0.962	38	12	60.746	0.106	0.448
14	40	115.727	0.060	0.312	38	13	143.466	0.512	1.280
14	41	101.820	0.076	0.258	38	14	113.770	0.372	1.008
14	42	88.914	0.102	0.224	38	15	90.627	0.204	0.850
14	43	88.008	0.300	0.390	38	16	201.247	0.312	0.760
14	44	68.826	0.308	0.304	38	17	188.550	0.310	0.698
14	45	50.229	0.202	0.218	38	18	177.089	0.266	0.668
14	46	89.728	0.462	0.432	38	19	93.315	0.094	0.348
14	47	68.055	0.406	0.296	38	20	66.750	0.108	0.236
14	48	50.867	0.326	0.210	38	21	45.449	0.082	0.158
14	49	67.505	0.060	0.532	38	22	227.315	0.482	1.680
14	50	53.399	0.012	0.442	38	23	211.475	0.416	1.578

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
14	51	45.966	0.002	0.388
14	52	78.568	0.210	0.196
14	53	64.701	0.168	0.168
14	54	52.248	0.108	0.150
15	1	121.365	0.000	0.466
15	2	97.158	0.000	0.368
15	3	75.930	0.000	0.296
15	4	194.929	0.000	1.544
15	5	169.795	0.000	1.360
15	6	145.021	0.000	1.148
15	7	239.830	0.000	2.554
15	8	207.280	0.000	2.202
15	9	176.905	0.000	1.866
15	10	112.063	0.028	0.872
15	11	85.278	0.014	0.686
15	12	61.540	0.006	0.484
15	13	169.380	0.024	1.754
15	14	128.987	0.032	1.314
15	15	96.808	0.048	0.982
15	16	200.128	0.000	0.826
15	17	187.998	0.000	0.764
15	18	177.161	0.000	0.720
15	19	93.639	0.134	0.348
15	20	67.588	0.106	0.244
15	21	45.850	0.046	0.164
15	22	227.097	0.002	1.836
15	23	212.338	0.002	1.722
15	24	198.816	0.000	1.600
15	25	96.542	0.046	0.970
15	26	69.128	0.016	0.724
15	27	49.285	0.002	0.528
15	28	187.662	0.000	1.558
15	29	171.019	0.000	1.422
15	30	154.051	0.000	1.276
15	31	103.188	0.000	0.270
15	32	85.471	0.000	0.224
15	33	68.599	0.000	0.182
15	34	143.955	0.000	0.862
15	35	128.613	0.000	0.772
15	36	115.546	0.000	0.702
15	37	140.369	0.004	1.168
15	38	124.804	0.000	1.046
15	39	114.017	0.000	0.964

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
38	24	196.421	0.398	1.444
38	25	93.199	0.258	0.824
38	26	68.557	0.108	0.668
38	27	49.552	0.024	0.520
38	28	187.662	0.000	1.558
38	29	171.019	0.000	1.422
38	30	154.051	0.000	1.276
38	31	103.188	0.000	0.270
38	32	85.471	0.000	0.224
38	33	68.599	0.000	0.182
38	34	143.955	0.000	0.862
38	35	128.613	0.000	0.772
38	36	115.546	0.000	0.702
38	37	139.152	0.276	1.080
38	38	124.678	0.130	1.010
38	39	114.437	0.048	0.958
38	40	115.870	0.434	0.276
38	41	102.353	0.424	0.236
38	42	88.491	0.408	0.196
38	43	82.857	0.472	0.302
38	44	64.896	0.334	0.254
38	45	49.351	0.226	0.218
38	46	88.261	0.338	0.424
38	47	66.684	0.340	0.290
38	48	50.139	0.274	0.218
38	49	68.409	0.126	0.524
38	50	53.452	0.020	0.440
38	51	46.164	0.008	0.388
38	52	80.507	0.382	0.186
38	53	65.543	0.284	0.154
38	54	53.800	0.222	0.138
39	1	111.395	0.494	0.376
39	2	91.652	0.408	0.296
39	3	75.626	0.286	0.266
39	4	159.502	0.668	1.014
39	5	139.673	0.610	0.898
39	6	122.027	0.516	0.774
39	7	169.475	0.792	1.344
39	8	137.632	0.782	0.996
39	9	115.764	0.744	0.788
39	10	101.188	0.454	0.662
39	11	78.732	0.438	0.534
39	12	59.946	0.410	0.380

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
15	40	115.382	0.002	0.316
15	41	101.646	0.010	0.272
15	42	88.522	0.008	0.240
15	43	88.017	0.038	0.466
15	44	66.422	0.050	0.348
15	45	48.459	0.036	0.270
15	46	90.308	0.154	0.480
15	47	67.755	0.124	0.336
15	48	49.785	0.072	0.250
15	49	66.566	0.004	0.546
15	50	53.009	0.000	0.442
15	51	45.896	0.000	0.388
15	52	76.318	0.044	0.216
15	53	61.982	0.036	0.180
15	54	49.764	0.018	0.158
16	1	115.756	0.302	0.410
16	2	94.175	0.298	0.330
16	3	76.020	0.260	0.278
16	4	164.259	0.658	1.084
16	5	135.893	0.732	0.860
16	6	113.725	0.760	0.652
16	7	162.864	0.864	1.240
16	8	125.411	0.924	0.814
16	9	99.596	0.952	0.520
16	10	95.409	0.814	0.558
16	11	74.413	0.774	0.410
16	12	58.733	0.710	0.298
16	13	120.076	0.946	0.836
16	14	90.855	0.940	0.544
16	15	69.951	0.944	0.334
16	16	201.426	0.338	0.758
16	17	188.958	0.350	0.692
16	18	177.325	0.338	0.650
16	19	89.298	0.520	0.290
16	20	65.341	0.470	0.194
16	21	46.215	0.390	0.134
16	22	227.447	0.564	1.662
16	23	211.820	0.582	1.536
16	24	196.487	0.634	1.388
16	25	86.930	0.896	0.454
16	26	70.874	0.806	0.330
16	27	60.175	0.658	0.318
16	28	171.432	0.784	1.220
39	13	121.350	0.914	0.860
39	14	92.603	0.884	0.578
39	15	72.733	0.806	0.406
39	16	201.843	0.368	0.754
39	17	189.375	0.390	0.684
39	18	177.643	0.372	0.648
39	19	92.779	0.142	0.342
39	20	67.229	0.156	0.240
39	21	45.714	0.160	0.148
39	22	228.008	0.238	1.780
39	23	212.952	0.294	1.640
39	24	198.670	0.298	1.520
39	25	89.041	0.684	0.586
39	26	70.546	0.504	0.476
39	27	54.144	0.324	0.422
39	28	171.200	0.774	1.216
39	29	153.473	0.736	1.086
39	30	136.634	0.670	0.956
39	31	93.554	0.560	0.178
39	32	76.485	0.554	0.122
39	33	63.001	0.500	0.102
39	34	142.789	0.554	0.702
39	35	129.462	0.456	0.638
39	36	116.553	0.306	0.604
39	37	140.670	0.454	1.070
39	38	127.013	0.456	0.962
39	39	116.042	0.344	0.898
39	40	116.909	0.550	0.274
39	41	103.080	0.572	0.224
39	42	88.931	0.572	0.180
39	43	79.517	0.854	0.170
39	44	65.398	0.802	0.122
39	45	53.251	0.676	0.110
39	46	87.656	0.202	0.438
39	47	65.878	0.236	0.304
39	48	49.295	0.270	0.210
39	49	73.384	0.440	0.466
39	50	58.944	0.274	0.406
39	51	48.641	0.114	0.374
39	52	79.149	0.248	0.202
39	53	66.097	0.258	0.162
39	54	54.747	0.256	0.140
40	1	116.944	0.222	0.428

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
16	29	151.256	0.842	1.050	40	2	93.720	0.306	0.334
16	30	130.882	0.888	0.854	40	3	75.438	0.350	0.262
16	31	99.009	0.280	0.224	40	4	149.016	0.902	0.868
16	32	81.572	0.328	0.176	40	5	124.731	0.932	0.692
16	33	65.437	0.364	0.134	40	6	105.061	0.940	0.510
16	34	143.613	0.592	0.684	40	7	151.963	0.968	1.054
16	35	131.359	0.608	0.610	40	8	117.843	0.990	0.680
16	36	120.854	0.558	0.568	40	9	95.293	0.994	0.440
16	37	140.057	0.808	0.972	40	10	111.090	0.084	0.860
16	38	127.289	0.762	0.888	40	11	83.909	0.128	0.650
16	39	117.601	0.662	0.838	40	12	60.661	0.190	0.438
16	40	116.002	0.476	0.274	40	13	160.861	0.156	1.606
16	41	102.179	0.504	0.224	40	14	117.759	0.262	1.100
16	42	88.751	0.508	0.184	40	15	87.515	0.332	0.782
16	43	80.743	0.832	0.200	40	16	200.819	0.064	0.820
16	44	65.418	0.844	0.110	40	17	188.907	0.096	0.750
16	45	55.549	0.826	0.080	40	18	177.643	0.124	0.696
16	46	85.932	0.794	0.332	40	19	93.465	0.018	0.358
16	47	66.038	0.792	0.214	40	20	67.235	0.036	0.246
16	48	51.358	0.768	0.146	40	21	45.689	0.050	0.164
16	49	77.779	0.742	0.406	40	22	227.196	0.074	1.814
16	50	66.345	0.594	0.372	40	23	212.291	0.098	1.690
16	51	55.401	0.424	0.328	40	24	198.626	0.092	1.572
16	52	81.626	0.460	0.180	40	25	90.724	0.420	0.724
16	53	68.018	0.404	0.142	40	26	70.075	0.502	0.488
16	54	56.807	0.366	0.126	40	27	57.797	0.520	0.354
17	1	120.671	0.020	0.456	40	28	182.167	0.312	1.446
17	2	96.835	0.026	0.360	40	29	160.316	0.450	1.218
17	3	75.900	0.022	0.294	40	30	138.296	0.594	0.982
17	4	188.607	0.172	1.452	40	31	91.877	0.662	0.160
17	5	161.162	0.222	1.228	40	32	74.545	0.714	0.100
17	6	137.298	0.242	1.020	40	33	60.977	0.744	0.074
17	7	205.478	0.452	1.958	40	34	142.109	0.854	0.612
17	8	159.510	0.594	1.354	40	35	131.303	0.860	0.540
17	9	124.068	0.702	0.912	40	36	121.739	0.822	0.494
17	10	101.268	0.580	0.658	40	37	140.053	0.096	1.142
17	11	77.482	0.530	0.480	40	38	125.176	0.146	1.018
17	12	59.289	0.428	0.366	40	39	114.924	0.244	0.912
17	13	126.626	0.842	0.952	40	40	116.414	0.096	0.312
17	14	94.962	0.870	0.616	40	41	102.987	0.108	0.270
17	15	72.392	0.860	0.386	40	42	89.014	0.124	0.226
17	16	200.362	0.078	0.806	40	43	81.988	0.580	0.272
17	17	188.214	0.078	0.750	40	44	65.139	0.688	0.156

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
17	18	177.305	0.086	0.702	40	45	54.366	0.764	0.092
17	19	91.293	0.254	0.326	40	46	89.291	0.058	0.484
17	20	66.294	0.242	0.228	40	47	67.152	0.056	0.342
17	21	45.596	0.198	0.152	40	48	49.360	0.092	0.242
17	22	229.039	0.334	1.764	40	49	70.078	0.252	0.496
17	23	212.677	0.344	1.618	40	50	60.045	0.332	0.394
17	24	198.202	0.394	1.474	40	51	52.754	0.310	0.344
17	25	88.729	0.784	0.524	40	52	77.839	0.156	0.204
17	26	71.253	0.660	0.412	40	53	64.655	0.200	0.164
17	27	57.122	0.508	0.354	40	54	52.650	0.180	0.134
17	28	183.748	0.278	1.464	41	1	101.192	0.948	0.282
17	29	164.161	0.394	1.284	41	2	84.109	0.972	0.208
17	30	144.154	0.468	1.090	41	3	73.129	0.952	0.170
17	31	102.888	0.014	0.264	41	4	150.501	0.888	0.874
17	32	84.803	0.020	0.212	41	5	123.646	0.956	0.664
17	33	67.803	0.034	0.168	41	6	102.852	0.970	0.470
17	34	144.259	0.104	0.836	41	7	167.009	0.850	1.300
17	35	129.414	0.100	0.748	41	8	123.416	0.946	0.768
17	36	116.638	0.098	0.682	41	9	97.944	0.976	0.482
17	37	140.414	0.580	1.028	41	10	91.527	0.990	0.490
17	38	126.788	0.522	0.940	41	11	71.315	0.978	0.338
17	39	116.964	0.380	0.902	41	12	57.047	0.932	0.228
17	40	115.906	0.162	0.304	41	13	116.862	0.996	0.778
17	41	101.921	0.164	0.258	41	14	87.964	1.000	0.488
17	42	88.190	0.188	0.218	41	15	68.491	0.994	0.302
17	43	83.050	0.608	0.274	41	16	202.515	0.324	0.762
17	44	66.242	0.624	0.184	41	17	190.733	0.384	0.690
17	45	53.343	0.570	0.148	41	18	178.896	0.422	0.638
17	46	87.662	0.658	0.380	41	19	85.663	0.910	0.226
17	47	66.484	0.662	0.238	41	20	63.369	0.884	0.132
17	48	50.643	0.586	0.166	41	21	47.547	0.796	0.086
17	49	75.604	0.590	0.432	41	22	229.622	0.432	1.732
17	50	61.826	0.390	0.396	41	23	214.488	0.490	1.596
17	51	51.680	0.240	0.360	41	24	198.221	0.532	1.430
17	52	78.209	0.202	0.202	41	25	86.319	0.974	0.404
17	53	63.943	0.160	0.172	41	26	72.248	0.908	0.298
17	54	52.069	0.152	0.144	41	27	60.767	0.720	0.280
18	1	121.365	0.000	0.466	41	28	171.721	0.812	1.224
18	2	96.951	0.002	0.364	41	29	149.766	0.904	1.018
18	3	75.930	0.000	0.296	41	30	128.602	0.970	0.810
18	4	194.308	0.024	1.532	41	31	87.287	0.914	0.118
18	5	169.347	0.024	1.352	41	32	69.708	0.982	0.060
18	6	144.722	0.028	1.138	41	33	57.659	0.988	0.044

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
18	7	232.397	0.122	2.420
18	8	193.480	0.204	1.948
18	9	159.555	0.280	1.548
18	10	106.455	0.304	0.764
18	11	81.248	0.262	0.584
18	12	60.703	0.210	0.422
18	13	138.536	0.674	1.172
18	14	103.644	0.698	0.776
18	15	77.496	0.712	0.500
18	16	200.230	0.008	0.826
18	17	188.237	0.012	0.764
18	18	177.035	0.018	0.716
18	19	93.029	0.072	0.350
18	20	67.208	0.070	0.244
18	21	45.360	0.054	0.160
18	22	228.402	0.132	1.816
18	23	212.946	0.124	1.694
18	24	198.514	0.140	1.554
18	25	90.331	0.630	0.612
18	26	71.139	0.476	0.504
18	27	55.037	0.340	0.428
18	28	186.984	0.056	1.540
18	29	169.464	0.096	1.384
18	30	151.678	0.118	1.224
18	31	103.188	0.000	0.270
18	32	85.471	0.000	0.224
18	33	68.328	0.002	0.178
18	34	143.885	0.012	0.858
18	35	128.705	0.012	0.768
18	36	115.631	0.010	0.698
18	37	140.826	0.356	1.084
18	38	126.239	0.276	0.992
18	39	115.347	0.170	0.936
18	40	115.411	0.048	0.310
18	41	101.566	0.038	0.266
18	42	88.278	0.052	0.232
18	43	86.716	0.352	0.378
18	44	66.764	0.346	0.266
18	45	50.665	0.302	0.200
18	46	88.445	0.464	0.410
18	47	67.146	0.478	0.274
18	48	50.369	0.410	0.192
18	49	72.976	0.420	0.462

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
41	34	144.489	0.558	0.698
41	35	133.046	0.656	0.620
41	36	121.541	0.584	0.574
41	37	140.860	0.908	0.966
41	38	127.778	0.872	0.866
41	39	118.777	0.754	0.828
41	40	116.524	0.798	0.234
41	41	101.613	0.896	0.176
41	42	87.641	0.938	0.132
41	43	77.969	0.992	0.124
41	44	64.842	0.996	0.066
41	45	56.324	0.992	0.038
41	46	85.185	0.960	0.302
41	47	65.427	0.962	0.182
41	48	51.306	0.952	0.110
41	49	79.254	0.846	0.392
41	50	66.706	0.592	0.382
41	51	53.058	0.304	0.354
41	52	87.215	0.842	0.140
41	53	74.555	0.808	0.100
41	54	64.239	0.730	0.086
42	1	118.029	0.228	0.424
42	2	95.554	0.272	0.332
42	3	76.500	0.234	0.272
42	4	176.146	0.454	1.254
42	5	144.509	0.604	0.964
42	6	122.052	0.656	0.758
42	7	205.036	0.468	1.936
42	8	153.192	0.680	1.232
42	9	116.354	0.808	0.770
42	10	101.607	0.670	0.646
42	11	77.317	0.614	0.456
42	12	59.529	0.490	0.346
42	13	124.253	0.924	0.910
42	14	90.494	0.958	0.540
42	15	70.278	0.962	0.334
42	16	200.792	0.060	0.816
42	17	188.657	0.062	0.752
42	18	177.629	0.092	0.700
42	19	94.029	0.066	0.356
42	20	67.666	0.066	0.250
42	21	45.778	0.058	0.158
42	22	227.444	0.040	1.830

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
18	50	58.390	0.244	0.408
18	51	48.295	0.096	0.378
18	52	75.717	0.050	0.208
18	53	61.649	0.040	0.176
18	54	49.886	0.026	0.158
19	1	121.365	0.000	0.466
19	2	97.158	0.000	0.368
19	3	75.930	0.000	0.296
19	4	194.979	0.004	1.544
19	5	169.845	0.004	1.360
19	6	145.085	0.002	1.148
19	7	238.869	0.020	2.536
19	8	204.916	0.042	2.160
19	9	173.396	0.070	1.794
19	10	109.918	0.126	0.830
19	11	83.515	0.126	0.640
19	12	61.265	0.070	0.464
19	13	148.435	0.460	1.362
19	14	112.054	0.498	0.950
19	15	85.097	0.486	0.682
19	16	200.194	0.004	0.826
19	17	187.998	0.000	0.764
19	18	177.233	0.002	0.720
19	19	93.586	0.016	0.356
19	20	67.121	0.018	0.248
19	21	45.042	0.014	0.160
19	22	227.080	0.030	1.828
19	23	212.560	0.034	1.718
19	24	199.128	0.048	1.592
19	25	92.341	0.466	0.716
19	26	70.320	0.326	0.566
19	27	52.566	0.188	0.474
19	28	187.627	0.008	1.556
19	29	170.920	0.010	1.418
19	30	153.763	0.014	1.266
19	31	103.188	0.000	0.270
19	32	85.471	0.000	0.224
19	33	68.599	0.000	0.182
19	34	143.955	0.002	0.862
19	35	128.613	0.000	0.772
19	36	115.546	0.000	0.702
19	37	140.373	0.178	1.120
19	38	125.359	0.122	1.024
42	23	212.605	0.046	1.714
42	24	198.715	0.072	1.574
42	25	88.057	0.870	0.470
42	26	71.885	0.698	0.394
42	27	57.191	0.448	0.386
42	28	185.134	0.222	1.490
42	29	163.855	0.394	1.268
42	30	143.627	0.502	1.070
42	31	99.521	0.386	0.220
42	32	79.977	0.552	0.142
42	33	63.823	0.678	0.098
42	34	144.993	0.152	0.824
42	35	130.255	0.164	0.736
42	36	117.978	0.168	0.670
42	37	140.805	0.340	1.090
42	38	126.436	0.294	0.992
42	39	115.769	0.182	0.938
42	40	116.740	0.216	0.296
42	41	103.137	0.320	0.236
42	42	89.860	0.400	0.188
42	43	80.084	0.934	0.160
42	44	65.689	0.960	0.084
42	45	56.274	0.938	0.048
42	46	90.279	0.306	0.460
42	47	67.830	0.360	0.296
42	48	50.657	0.320	0.210
42	49	74.540	0.476	0.470
42	50	58.067	0.192	0.430
42	51	47.170	0.040	0.388
42	52	81.107	0.332	0.190
42	53	67.718	0.324	0.158
42	54	55.933	0.266	0.140
43	1	121.397	0.002	0.466
43	2	97.142	0.002	0.368
43	3	75.980	0.002	0.296
43	4	191.516	0.114	1.478
43	5	163.992	0.170	1.262
43	6	139.896	0.206	1.054
43	7	230.185	0.154	2.370
43	8	186.874	0.304	1.810
43	9	153.029	0.408	1.408
43	10	111.409	0.068	0.858
43	11	84.761	0.064	0.668

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
19	39	114.244	0.070	0.946	43	12	61.443	0.040	0.474
19	40	115.476	0.004	0.318	43	13	141.823	0.644	1.216
19	41	101.709	0.008	0.274	43	14	100.709	0.780	0.716
19	42	88.253	0.010	0.236	43	15	75.515	0.822	0.442
19	43	87.195	0.188	0.418	43	16	200.158	0.002	0.826
19	44	65.605	0.142	0.308	43	17	188.044	0.008	0.762
19	45	48.224	0.140	0.226	43	18	177.275	0.010	0.718
19	46	89.622	0.242	0.458	43	19	93.681	0.000	0.360
19	47	67.776	0.250	0.320	43	20	67.371	0.000	0.252
19	48	50.003	0.216	0.226	43	21	45.410	0.000	0.166
19	49	71.075	0.280	0.494	43	22	227.062	0.002	1.836
19	50	55.509	0.104	0.430	43	23	212.348	0.004	1.722
19	51	46.669	0.034	0.382	43	24	198.826	0.002	1.600
19	52	75.403	0.012	0.216	43	25	92.657	0.588	0.660
19	53	61.031	0.008	0.180	43	26	72.077	0.388	0.558
19	54	49.243	0.002	0.158	43	27	54.524	0.218	0.490
20	1	121.365	0.000	0.466	43	28	187.595	0.006	1.556
20	2	97.158	0.000	0.368	43	29	170.650	0.038	1.410
20	3	75.930	0.000	0.296	43	30	153.161	0.052	1.256
20	4	194.929	0.000	1.544	43	31	103.048	0.032	0.266
20	5	169.795	0.000	1.360	43	32	84.812	0.092	0.206
20	6	145.021	0.000	1.148	43	33	68.147	0.146	0.156
20	7	239.748	0.002	2.552	43	34	144.108	0.016	0.860
20	8	207.174	0.004	2.200	43	35	128.836	0.018	0.768
20	9	176.102	0.008	1.850	43	36	115.775	0.018	0.698
20	10	111.492	0.038	0.862	43	37	140.443	0.028	1.164
20	11	84.832	0.034	0.676	43	38	125.120	0.024	1.044
20	12	61.511	0.022	0.480	43	39	114.089	0.014	0.960
20	13	157.918	0.272	1.540	43	40	115.581	0.016	0.316
20	14	119.945	0.298	1.118	43	41	101.864	0.026	0.272
20	15	90.537	0.284	0.824	43	42	88.668	0.044	0.234
20	16	200.128	0.000	0.826	43	43	83.696	0.752	0.240
20	17	187.998	0.000	0.764	43	44	67.507	0.796	0.140
20	18	177.161	0.000	0.720	43	45	55.995	0.772	0.094
20	19	93.496	0.010	0.356	43	46	89.779	0.012	0.496
20	20	67.094	0.008	0.248	43	47	67.373	0.006	0.352
20	21	45.064	0.012	0.160	43	48	49.031	0.002	0.254
20	22	227.123	0.018	1.832	43	49	69.060	0.130	0.532
20	23	212.449	0.022	1.720	43	50	53.850	0.028	0.442
20	24	199.035	0.030	1.594	43	51	45.966	0.002	0.388
20	25	94.831	0.262	0.846	43	52	76.341	0.042	0.218
20	26	69.663	0.182	0.634	43	53	61.831	0.034	0.178
20	27	50.898	0.072	0.514	43	54	50.307	0.032	0.158

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
20	28	187.662	0.000	1.558
20	29	170.955	0.002	1.420
20	30	154.051	0.000	1.276
20	31	103.188	0.000	0.270
20	32	85.471	0.000	0.224
20	33	68.599	0.000	0.182
20	34	143.955	0.000	0.862
20	35	128.613	0.000	0.772
20	36	115.546	0.000	0.702
20	37	140.254	0.080	1.146
20	38	124.833	0.040	1.036
20	39	114.171	0.018	0.960
20	40	115.420	0.000	0.318
20	41	101.667	0.002	0.274
20	42	88.452	0.000	0.240
20	43	87.464	0.082	0.450
20	44	66.107	0.056	0.348
20	45	47.892	0.062	0.254
20	46	89.717	0.160	0.474
20	47	67.935	0.134	0.342
20	48	49.206	0.100	0.238
20	49	69.233	0.156	0.522
20	50	53.633	0.030	0.436
20	51	46.024	0.004	0.388
20	52	75.289	0.008	0.216
20	53	60.992	0.002	0.182
20	54	49.225	0.000	0.158
21	1	100.217	0.890	0.282
21	2	83.992	0.918	0.218
21	3	72.273	0.968	0.166
21	4	143.559	0.992	0.786
21	5	120.686	0.992	0.622
21	6	101.447	0.998	0.450
21	7	150.837	0.988	1.034
21	8	117.670	0.996	0.674
21	9	94.984	0.998	0.434
21	10	111.664	0.010	0.870
21	11	84.548	0.048	0.676
21	12	60.483	0.086	0.456
21	13	163.874	0.078	1.666
21	14	124.864	0.102	1.240
21	15	91.738	0.134	0.892
21	16	203.038	0.788	0.644

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
44	1	121.365	0.000	0.466
44	2	97.158	0.000	0.368
44	3	75.930	0.000	0.296
44	4	194.840	0.014	1.540
44	5	169.770	0.020	1.354
44	6	144.699	0.024	1.134
44	7	238.599	0.026	2.530
44	8	203.274	0.066	2.124
44	9	172.203	0.118	1.766
44	10	112.110	0.000	0.878
44	11	85.320	0.000	0.692
44	12	61.474	0.000	0.484
44	13	160.788	0.252	1.574
44	14	117.571	0.412	1.046
44	15	89.259	0.426	0.760
44	16	200.128	0.000	0.826
44	17	187.998	0.000	0.764
44	18	177.239	0.002	0.720
44	19	93.681	0.000	0.360
44	20	67.371	0.000	0.252
44	21	45.410	0.000	0.166
44	22	227.052	0.000	1.836
44	23	212.298	0.000	1.722
44	24	198.816	0.000	1.600
44	25	96.972	0.238	0.884
44	26	70.983	0.160	0.670
44	27	50.583	0.048	0.524
44	28	187.662	0.000	1.558
44	29	171.019	0.000	1.422
44	30	153.993	0.004	1.274
44	31	103.188	0.000	0.270
44	32	85.509	0.004	0.224
44	33	68.599	0.000	0.182
44	34	143.955	0.002	0.862
44	35	128.631	0.002	0.772
44	36	115.617	0.002	0.702
44	37	140.295	0.000	1.168
44	38	124.804	0.000	1.046
44	39	114.017	0.000	0.964
44	40	115.420	0.000	0.318
44	41	101.667	0.002	0.274
44	42	88.452	0.000	0.240
44	43	87.528	0.366	0.378

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
21	17	189.334	0.830	0.572
21	18	175.945	0.850	0.512
21	19	93.555	0.008	0.358
21	20	67.319	0.018	0.250
21	21	45.348	0.032	0.160
21	22	226.207	0.160	1.772
21	23	211.668	0.182	1.664
21	24	196.506	0.206	1.508
21	25	91.706	0.224	0.822
21	26	69.903	0.480	0.498
21	27	57.865	0.626	0.284
21	28	168.480	0.906	1.166
21	29	147.499	0.940	0.980
21	30	128.632	0.940	0.814
21	31	85.036	0.996	0.106
21	32	69.081	1.000	0.058
21	33	57.379	1.000	0.044
21	34	141.595	0.954	0.564
21	35	131.056	0.968	0.496
21	36	122.348	0.960	0.452
21	37	140.200	0.102	1.144
21	38	124.806	0.198	0.998
21	39	114.790	0.348	0.884
21	40	114.841	0.166	0.300
21	41	100.997	0.148	0.256
21	42	87.306	0.170	0.210
21	43	81.083	0.482	0.302
21	44	63.052	0.650	0.148
21	45	54.289	0.818	0.086
21	46	89.578	0.010	0.496
21	47	67.316	0.012	0.352
21	48	48.917	0.018	0.252
21	49	69.024	0.220	0.490
21	50	61.182	0.412	0.386
21	51	55.534	0.424	0.334
21	52	76.597	0.176	0.194
21	53	64.280	0.288	0.144
21	54	56.576	0.450	0.096
22	1	103.798	0.874	0.302
22	2	85.353	0.898	0.214
22	3	73.822	0.832	0.186
22	4	162.939	0.688	1.038
22	5	132.889	0.816	0.788

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
44	44	69.164	0.432	0.266
44	45	52.959	0.428	0.168
44	46	89.673	0.000	0.498
44	47	67.247	0.000	0.352
44	48	48.987	0.000	0.254
44	49	66.592	0.008	0.544
44	50	53.079	0.002	0.442
44	51	45.896	0.000	0.388
44	52	75.198	0.000	0.218
44	53	61.034	0.002	0.182
44	54	49.225	0.000	0.158
45	1	121.365	0.000	0.466
45	2	97.158	0.000	0.368
45	3	75.930	0.000	0.296
45	4	194.933	0.002	1.544
45	5	169.799	0.002	1.360
45	6	145.021	0.000	1.148
45	7	239.766	0.002	2.552
45	8	207.197	0.004	2.200
45	9	176.147	0.016	1.852
45	10	112.110	0.000	0.878
45	11	85.320	0.000	0.692
45	12	61.474	0.000	0.484
45	13	168.154	0.058	1.730
45	14	127.236	0.082	1.276
45	15	95.302	0.126	0.938
45	16	200.128	0.000	0.826
45	17	187.998	0.000	0.764
45	18	177.161	0.000	0.720
45	19	93.681	0.000	0.360
45	20	67.371	0.000	0.252
45	21	45.410	0.000	0.166
45	22	227.052	0.000	1.836
45	23	212.298	0.000	1.722
45	24	198.816	0.000	1.600
45	25	96.284	0.046	0.968
45	26	69.479	0.026	0.722
45	27	49.356	0.004	0.528
45	28	187.662	0.000	1.558
45	29	171.019	0.000	1.422
45	30	154.051	0.000	1.276
45	31	103.188	0.000	0.270
45	32	85.471	0.000	0.224

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
22	6	108.550	0.884	0.552
22	7	182.477	0.698	1.556
22	8	132.170	0.866	0.898
22	9	101.997	0.938	0.540
22	10	106.777	0.192	0.798
22	11	82.092	0.274	0.616
22	12	59.688	0.440	0.382
22	13	143.966	0.454	1.312
22	14	108.403	0.464	0.924
22	15	77.838	0.582	0.564
22	16	203.802	0.578	0.710
22	17	190.536	0.602	0.650
22	18	178.233	0.598	0.588
22	19	92.820	0.044	0.344
22	20	66.285	0.072	0.232
22	21	45.346	0.150	0.156
22	22	228.060	0.454	1.704
22	23	210.927	0.458	1.550
22	24	196.836	0.440	1.436
22	25	86.759	0.640	0.562
22	26	70.685	0.726	0.372
22	27	58.823	0.630	0.302
22	28	181.929	0.386	1.422
22	29	159.480	0.564	1.186
22	30	138.392	0.674	0.976
22	31	88.737	0.850	0.124
22	32	70.502	0.954	0.064
22	33	58.032	0.972	0.046
22	34	144.198	0.462	0.714
22	35	132.670	0.578	0.634
22	36	120.291	0.508	0.584
22	37	139.609	0.464	1.054
22	38	126.903	0.528	0.942
22	39	117.190	0.486	0.876
22	40	115.102	0.630	0.254
22	41	100.033	0.632	0.206
22	42	86.122	0.652	0.152
22	43	78.072	0.944	0.138
22	44	64.733	0.970	0.076
22	45	56.293	0.978	0.042
22	46	89.077	0.038	0.482
22	47	67.095	0.050	0.342
22	48	49.160	0.068	0.248

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
45	33	68.599	0.000	0.182
45	34	143.955	0.000	0.862
45	35	128.613	0.000	0.772
45	36	115.546	0.000	0.702
45	37	140.295	0.000	1.168
45	38	124.804	0.000	1.046
45	39	114.017	0.000	0.964
45	40	115.420	0.000	0.318
45	41	101.659	0.000	0.274
45	42	88.452	0.000	0.240
45	43	88.272	0.124	0.444
45	44	67.500	0.166	0.328
45	45	50.042	0.146	0.248
45	46	89.673	0.000	0.498
45	47	67.247	0.000	0.352
45	48	48.987	0.000	0.254
45	49	66.426	0.000	0.546
45	50	53.009	0.000	0.442
45	51	45.896	0.000	0.388
45	52	75.198	0.000	0.218
45	53	60.950	0.000	0.182
45	54	49.225	0.000	0.158
46	1	121.365	0.000	0.466
46	2	97.158	0.000	0.368
46	3	75.930	0.000	0.296
46	4	194.929	0.000	1.544
46	5	169.795	0.000	1.360
46	6	145.021	0.000	1.148
46	7	239.830	0.000	2.554
46	8	207.280	0.000	2.202
46	9	176.905	0.000	1.866
46	10	112.110	0.000	0.878
46	11	85.320	0.000	0.692
46	12	61.474	0.000	0.484
46	13	169.818	0.000	1.772
46	14	129.510	0.000	1.332
46	15	96.765	0.000	1.000
46	16	200.128	0.000	0.826
46	17	187.998	0.000	0.764
46	18	177.161	0.000	0.720
46	19	93.681	0.000	0.360
46	20	67.371	0.000	0.252
46	21	45.410	0.000	0.166

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
22	49	73.496	0.470	0.462
22	50	63.393	0.462	0.388
22	51	51.226	0.222	0.364
22	52	81.013	0.544	0.154
22	53	71.107	0.662	0.112
22	54	62.824	0.722	0.082
23	1	116.748	0.374	0.406
23	2	94.514	0.402	0.312
23	3	76.746	0.278	0.266
23	4	189.405	0.160	1.434
23	5	159.767	0.256	1.176
23	6	135.781	0.316	0.972
23	7	232.123	0.136	2.400
23	8	189.244	0.284	1.844
23	9	150.599	0.442	1.346
23	10	102.415	0.462	0.708
23	11	76.162	0.614	0.480
23	12	59.424	0.626	0.326
23	13	125.654	0.832	0.950
23	14	92.460	0.866	0.584
23	15	72.160	0.862	0.390
23	16	202.489	0.254	0.774
23	17	189.648	0.272	0.704
23	18	178.218	0.266	0.662
23	19	91.555	0.174	0.338
23	20	65.010	0.260	0.210
23	21	46.175	0.322	0.128
23	22	229.013	0.334	1.752
23	23	213.070	0.352	1.608
23	24	197.234	0.386	1.450
23	25	87.619	0.804	0.492
23	26	71.353	0.650	0.414
23	27	56.772	0.392	0.418
23	28	187.642	0.032	1.554
23	29	169.571	0.080	1.388
23	30	151.601	0.138	1.222
23	31	101.295	0.310	0.236
23	32	80.791	0.510	0.146
23	33	64.472	0.656	0.098
23	34	144.484	0.054	0.850
23	35	129.714	0.068	0.762
23	36	116.657	0.062	0.690
23	37	140.961	0.504	1.056
46	22	227.052	0.000	1.836
46	23	212.298	0.000	1.722
46	24	198.816	0.000	1.600
46	25	95.760	0.000	0.982
46	26	68.846	0.000	0.728
46	27	49.195	0.000	0.528
46	28	187.662	0.000	1.558
46	29	171.019	0.000	1.422
46	30	154.051	0.000	1.276
46	31	103.188	0.000	0.270
46	32	85.471	0.000	0.224
46	33	68.599	0.000	0.182
46	34	143.955	0.000	0.862
46	35	128.613	0.000	0.772
46	36	115.546	0.000	0.702
46	37	140.295	0.000	1.168
46	38	124.804	0.000	1.046
46	39	114.017	0.000	0.964
46	40	115.420	0.000	0.318
46	41	101.659	0.000	0.274
46	42	88.452	0.000	0.240
46	43	87.545	0.000	0.472
46	44	65.862	0.000	0.362
46	45	47.402	0.000	0.272
46	46	89.673	0.000	0.498
46	47	67.247	0.000	0.352
46	48	48.987	0.000	0.254
46	49	66.426	0.000	0.546
46	50	53.009	0.000	0.442
46	51	45.896	0.000	0.388
46	52	75.198	0.000	0.218
46	53	60.950	0.000	0.182
46	54	49.225	0.000	0.158
47	1	98.770	1.000	0.266
47	2	83.439	1.000	0.204
47	3	72.558	1.000	0.162
47	4	142.878	1.000	0.774
47	5	120.287	1.000	0.618
47	6	101.253	1.000	0.448
47	7	148.617	1.000	0.998
47	8	116.961	1.000	0.662
47	9	94.832	1.000	0.432
47	10	90.949	1.000	0.482

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects	Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
23	38	126.676	0.416	0.970	47	11	70.553	1.000	0.324
23	39	116.294	0.242	0.930	47	12	56.426	1.000	0.218
23	40	116.674	0.714	0.238	47	13	116.704	1.000	0.776
23	41	101.477	0.776	0.192	47	14	87.964	1.000	0.488
23	42	87.643	0.816	0.140	47	15	68.174	1.000	0.296
23	43	78.750	0.944	0.142	47	16	204.320	1.000	0.602
23	44	65.658	0.942	0.086	47	17	189.635	1.000	0.534
23	45	56.597	0.924	0.058	47	18	175.500	1.000	0.472
23	46	87.886	0.170	0.448	47	19	83.722	1.000	0.200
23	47	65.258	0.180	0.292	47	20	62.168	1.000	0.112
23	48	48.801	0.278	0.200	47	21	47.310	1.000	0.066
23	49	75.111	0.492	0.476	47	22	228.192	1.000	1.526
23	50	59.989	0.268	0.426	47	23	210.472	1.000	1.374
23	51	47.367	0.052	0.386	47	24	193.759	1.000	1.232
23	52	85.063	0.740	0.146	47	25	85.204	1.000	0.374
23	53	73.065	0.738	0.110	47	26	72.085	1.000	0.252
23	54	62.672	0.656	0.092	47	27	63.281	1.000	0.162
24	1	121.226	0.030	0.462	47	28	166.690	1.000	1.138
24	2	97.244	0.030	0.366	47	29	145.863	1.000	0.954
24	3	76.306	0.022	0.296	47	30	127.180	1.000	0.788
24	4	194.918	0.008	1.542	47	31	84.858	1.000	0.104
24	5	169.604	0.014	1.354	47	32	69.081	1.000	0.058
24	6	144.656	0.026	1.138	47	33	57.379	1.000	0.044
24	7	239.712	0.004	2.550	47	34	140.894	1.000	0.544
24	8	206.944	0.016	2.190	47	35	130.877	1.000	0.488
24	9	175.986	0.044	1.842	47	36	122.104	1.000	0.438
24	10	98.273	0.662	0.612	47	37	140.200	1.000	0.934
24	11	76.071	0.582	0.452	47	38	128.018	1.000	0.840
24	12	59.430	0.438	0.352	47	39	119.312	1.000	0.774
24	13	131.285	0.788	1.034	47	40	114.919	1.000	0.204
24	14	97.972	0.812	0.680	47	41	99.961	1.000	0.154
24	15	77.053	0.784	0.478	47	42	86.425	1.000	0.108
24	16	201.028	0.072	0.814	47	43	77.885	1.000	0.124
24	17	188.547	0.082	0.744	47	44	64.772	1.000	0.066
24	18	177.836	0.080	0.702	47	45	56.264	1.000	0.034
24	19	88.883	0.360	0.292	47	46	84.415	1.000	0.284
24	20	64.915	0.420	0.188	47	47	65.518	1.000	0.176
24	21	45.374	0.394	0.118	47	48	51.193	1.000	0.104
24	22	228.367	0.200	1.788	47	49	80.774	1.000	0.350
24	23	213.271	0.218	1.660	47	50	73.810	1.000	0.302
24	24	198.534	0.256	1.512	47	51	69.228	1.000	0.270
24	25	90.479	0.684	0.592	47	52	87.971	1.000	0.116
24	26	71.149	0.446	0.514	47	53	76.867	1.000	0.082

Table 37 Output Measures Data of Reinspection Decision Policies (cont.)

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
24	27	54.081	0.204	0.492

Policy No	Treat-ment	Total Cost	Schedule Delay	Major Field Defects
47	54	68.443	1.000	0.054

APPENDIX E

REINSPECTION DECISION POLICIES' AGGREGATE STANDARDIZED MEAN SQUARED VALUES FOR ALL PREFERENCE PROFILES

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
	Profile 1					Profile 2				
2	0.357	-0.308	0.409	0.458	39	0.858	-0.092	0.123	0.888	39
3	0.285	-0.226	0.276	0.335	30	0.685	-0.068	0.083	0.700	30
4	-0.468	0.455	-0.467	-0.481	9	-1.124	0.136	-0.140	-1.128	4
5	-0.542	0.839	-0.566	-0.268	15	-1.300	0.252	-0.170	-1.218	1
7	0.353	-0.292	0.380	0.441	37	0.847	-0.088	0.114	0.873	37
8	0.284	-0.221	0.257	0.321	28	0.683	-0.066	0.077	0.694	27
9	-0.335	0.145	-0.352	-0.543	3	-0.804	0.043	-0.106	-0.867	11
10	-0.481	0.457	-0.492	-0.517	7	-1.154	0.137	-0.148	-1.165	3
11	-0.465	0.712	-0.503	-0.257	16	-1.116	0.214	-0.151	-1.053	6
12	0.104	0.164	-0.005	0.263	24	0.250	0.049	-0.002	0.298	20
13	0.280	-0.126	0.243	0.397	35	0.671	-0.038	0.073	0.706	32
14	0.344	-0.274	0.355	0.425	36	0.826	-0.082	0.107	0.850	36
15	0.364	-0.307	0.409	0.466	40	0.873	-0.092	0.123	0.903	41
16	-0.364	0.230	-0.402	-0.535	6	-0.873	0.069	-0.120	-0.925	10
17	-0.039	-0.077	-0.104	-0.220	17	-0.094	-0.023	-0.031	-0.148	18
18	0.200	-0.218	0.163	0.146	22	0.480	-0.065	0.049	0.464	23
19	0.289	-0.275	0.290	0.304	25	0.693	-0.083	0.087	0.698	28
20	0.326	-0.299	0.352	0.379	34	0.783	-0.090	0.106	0.799	35
21	-0.406	0.184	-0.284	-0.506	8	-0.975	0.055	-0.085	-1.005	8
22	-0.265	0.166	-0.271	-0.370	11	-0.636	0.050	-0.081	-0.667	15
23	0.120	-0.001	0.025	0.143	21	0.287	0.000	0.007	0.294	19
24	0.261	-0.115	0.205	0.351	31	0.627	-0.035	0.062	0.654	25
25	0.304	-0.219	0.286	0.371	33	0.729	-0.066	0.086	0.749	33

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
26	-0.318	0.039	-0.260	-0.539	4	-0.763	0.012	-0.078	-0.829	13
27	-0.057	-0.131	-0.083	-0.272	14	-0.137	-0.039	-0.025	-0.202	17
28	0.156	-0.211	0.132	0.077	19	0.373	-0.063	0.040	0.350	21
29	0.237	-0.252	0.245	0.230	23	0.569	-0.076	0.073	0.567	24
30	0.273	-0.272	0.302	0.304	26	0.655	-0.082	0.091	0.665	26
33	0.286	-0.227	0.276	0.335	29	0.686	-0.068	0.083	0.700	31
34	-0.453	0.327	-0.434	-0.560	2	-1.088	0.098	-0.130	-1.120	5
35	-0.384	0.224	-0.279	-0.439	10	-0.922	0.067	-0.084	-0.938	9
38	0.288	-0.238	0.259	0.310	27	0.692	-0.071	0.078	0.698	29
39	-0.322	0.022	-0.296	-0.595	1	-0.772	0.007	-0.089	-0.854	12
40	-0.306	-0.019	-0.213	-0.537	5	-0.733	-0.006	-0.064	-0.803	14
41	-0.438	0.563	-0.440	-0.315	12	-1.052	0.169	-0.132	-1.015	7
42	-0.128	-0.011	-0.142	-0.281	13	-0.308	-0.003	-0.043	-0.354	16
43	0.161	-0.209	0.154	0.106	20	0.387	-0.063	0.046	0.371	22
44	0.315	-0.286	0.332	0.361	32	0.756	-0.086	0.100	0.770	34
45	0.355	-0.307	0.400	0.448	38	0.852	-0.092	0.120	0.880	38
46	0.359	-0.309	0.416	0.466	41	0.862	-0.093	0.125	0.895	40
47	-0.530	0.901	-0.573	-0.203	18	-1.272	0.270	-0.172	-1.174	2
	Profile 3					**Profile 4**				
2	0.643	-0.277	0.123	0.488	39	0.643	-0.092	0.368	0.919	39
3	0.514	-0.204	0.083	0.393	30	0.514	-0.068	0.249	0.694	30
4	-0.843	0.409	-0.140	-0.574	8	-0.843	0.136	-0.420	-1.127	4
5	-0.975	0.756	-0.170	-0.389	13	-0.975	0.252	-0.509	-1.233	1
7	0.635	-0.263	0.114	0.486	38	0.635	-0.088	0.342	0.889	37
8	0.512	-0.199	0.077	0.390	28	0.512	-0.066	0.231	0.677	26
9	-0.603	0.130	-0.106	-0.579	7	-0.603	0.043	-0.317	-0.877	10
10	-0.866	0.411	-0.148	-0.602	6	-0.866	0.137	-0.443	-1.172	3
11	-0.837	0.641	-0.151	-0.347	14	-0.837	0.214	-0.453	-1.076	6
12	0.188	0.148	-0.002	0.334	24	0.188	0.049	-0.005	0.232	19
13	0.504	-0.114	0.073	0.463	35	0.504	-0.038	0.219	0.685	29
14	0.620	-0.247	0.107	0.479	36	0.620	-0.082	0.320	0.857	36
15	0.655	-0.276	0.123	0.501	41	0.655	-0.092	0.368	0.930	41
16	-0.655	0.207	-0.120	-0.568	10	-0.655	0.069	-0.361	-0.947	8
17	-0.071	-0.069	-0.031	-0.171	18	-0.071	-0.023	-0.093	-0.187	18
18	0.360	-0.196	0.049	0.213	21	0.360	-0.065	0.147	0.442	23
19	0.520	-0.248	0.087	0.359	26	0.520	-0.083	0.261	0.699	32
20	0.587	-0.269	0.106	0.424	32	0.587	-0.090	0.317	0.814	35
21	-0.731	0.165	-0.085	-0.651	2	-0.731	0.055	-0.256	-0.932	9
22	-0.477	0.149	-0.081	-0.409	12	-0.477	0.050	-0.244	-0.671	15
23	0.215	-0.001	0.007	0.222	22	0.215	0.000	0.022	0.237	20
24	0.470	-0.104	0.062	0.428	33	0.470	-0.035	0.185	0.621	25
25	0.547	-0.197	0.086	0.436	34	0.547	-0.066	0.257	0.738	33
26	-0.572	0.036	-0.078	-0.615	5	-0.572	0.012	-0.234	-0.794	13
27	-0.103	-0.118	-0.025	-0.246	17	-0.103	-0.039	-0.075	-0.218	17
28	0.280	-0.190	0.040	0.130	19	0.280	-0.063	0.119	0.336	21
29	0.427	-0.227	0.073	0.273	23	0.427	-0.076	0.220	0.571	24
30	0.492	-0.245	0.091	0.338	25	0.492	-0.082	0.272	0.682	28
33	0.514	-0.205	0.083	0.393	29	0.514	-0.068	0.249	0.695	31
34	-0.816	0.295	-0.130	-0.652	1	-0.816	0.098	-0.390	-1.108	5

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
35	-0.691	0.201	-0.084	-0.574	9	-0.691	0.067	-0.251	-0.875	11
38	0.519	-0.214	0.078	0.383	27	0.519	-0.071	0.233	0.681	27
39	-0.579	0.020	-0.089	-0.647	3	-0.579	0.007	-0.266	-0.838	12
40	-0.550	-0.017	-0.064	-0.631	4	-0.550	-0.006	-0.192	-0.747	14
41	-0.789	0.507	-0.132	-0.414	11	-0.789	0.169	-0.396	-1.016	7
42	-0.231	-0.010	-0.043	-0.283	16	-0.231	-0.003	-0.128	-0.362	16
43	0.290	-0.188	0.046	0.149	20	0.290	-0.063	0.138	0.366	22
44	0.567	-0.257	0.100	0.409	31	0.567	-0.086	0.299	0.780	34
45	0.639	-0.276	0.120	0.483	37	0.639	-0.092	0.360	0.907	38
46	0.647	-0.278	0.125	0.494	40	0.647	-0.093	0.374	0.928	40
47	-0.954	0.811	-0.172	-0.315	15	-0.954	0.270	-0.516	-1.200	2
	Profile 5					**Profile 6**				
2	0.107	-0.740	0.123	-0.510	2	0.322	-0.555	0.123	-0.111	18
3	0.086	-0.544	0.083	-0.375	17	0.257	-0.408	0.083	-0.068	25
4	-0.141	1.091	-0.140	0.811	37	-0.422	0.819	-0.140	0.257	36
5	-0.162	2.015	-0.170	1.682	40	-0.487	1.511	-0.170	0.854	40
7	0.106	-0.700	0.114	-0.480	8	0.317	-0.525	0.114	-0.094	21
8	0.085	-0.530	0.077	-0.367	18	0.256	-0.397	0.077	-0.064	26
9	-0.101	0.347	-0.106	0.141	29	-0.302	0.260	-0.106	-0.147	13
10	-0.144	1.096	-0.148	0.804	36	-0.433	0.822	-0.148	0.241	35
11	-0.139	1.708	-0.151	1.418	39	-0.418	1.281	-0.151	0.712	39
12	0.031	0.394	-0.002	0.424	34	0.094	0.296	-0.002	0.388	37
13	0.084	-0.303	0.073	-0.146	23	0.252	-0.227	0.073	0.097	33
14	0.103	-0.659	0.107	-0.449	11	0.310	-0.494	0.107	-0.078	23
15	0.109	-0.736	0.123	-0.504	5	0.327	-0.552	0.123	-0.102	20
16	-0.109	0.553	-0.120	0.323	32	-0.327	0.415	-0.120	-0.033	28
17	-0.012	-0.185	-0.031	-0.228	21	-0.035	-0.139	-0.031	-0.205	5
18	0.060	-0.523	0.049	-0.414	13	0.180	-0.392	0.049	-0.163	10
19	0.087	-0.661	0.087	-0.487	7	0.260	-0.496	0.087	-0.149	12
20	0.098	-0.717	0.106	-0.513	1	0.294	-0.537	0.106	-0.138	14
21	-0.122	0.441	-0.085	0.234	30	-0.366	0.331	-0.085	-0.120	16
22	-0.079	0.398	-0.081	0.237	31	-0.238	0.298	-0.081	-0.021	30
23	0.036	-0.002	0.007	0.041	28	0.108	-0.002	0.007	0.113	34
24	0.078	-0.277	0.062	-0.137	24	0.235	-0.207	0.062	0.089	32
25	0.091	-0.525	0.086	-0.348	20	0.273	-0.393	0.086	-0.034	27
26	-0.095	0.095	-0.078	-0.079	27	-0.286	0.071	-0.078	-0.293	4
27	-0.017	-0.315	-0.025	-0.357	19	-0.052	-0.236	-0.025	-0.313	3
28	0.047	-0.505	0.040	-0.419	12	0.140	-0.379	0.040	-0.199	6
29	0.071	-0.605	0.073	-0.460	10	0.213	-0.453	0.073	-0.167	9
30	0.082	-0.652	0.091	-0.480	9	0.246	-0.489	0.091	-0.153	11
33	0.086	-0.545	0.083	-0.377	16	0.257	-0.409	0.083	-0.069	24
34	-0.136	0.786	-0.130	0.520	35	-0.408	0.589	-0.130	0.051	31
35	-0.115	0.537	-0.084	0.338	33	-0.346	0.403	-0.084	-0.026	29
38	0.086	-0.571	0.078	-0.407	14	0.259	-0.428	0.078	-0.091	22
39	-0.096	0.054	-0.089	-0.131	25	-0.289	0.040	-0.089	-0.338	2
40	-0.092	-0.045	-0.064	-0.200	22	-0.275	-0.034	-0.064	-0.372	1
41	-0.131	1.352	-0.132	1.089	38	-0.394	1.014	-0.132	0.488	38
42	-0.038	-0.026	-0.043	-0.107	26	-0.115	-0.019	-0.043	-0.177	8
43	0.048	-0.500	0.046	-0.406	15	0.145	-0.375	0.046	-0.184	7

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
44	0.094	-0.686	0.100	-0.492	6	0.283	-0.515	0.100	-0.131	15
45	0.106	-0.736	0.120	-0.510	3	0.319	-0.552	0.120	-0.113	17
46	0.108	-0.741	0.125	-0.509	4	0.323	-0.556	0.125	-0.108	19
47	-0.159	2.162	-0.172	1.831	41	-0.477	1.622	-0.172	0.972	41
	Profile 7					**Profile 8**				
2	0.107	-0.555	0.368	-0.080	19	0.107	-0.092	0.982	0.996	39
3	0.086	-0.408	0.249	-0.073	24	0.086	-0.068	0.663	0.680	29
4	-0.141	0.819	-0.420	0.258	36	-0.141	0.136	-1.121	-1.125	5
5	-0.162	1.511	-0.509	0.839	40	-0.162	0.252	-1.359	-1.269	1
7	0.106	-0.525	0.342	-0.078	20	0.106	-0.088	0.911	0.930	37
8	0.085	-0.397	0.231	-0.081	18	0.085	-0.066	0.616	0.636	27
9	-0.101	0.260	-0.317	-0.157	11	-0.101	0.043	-0.845	-0.902	9
10	-0.144	0.822	-0.443	0.235	35	-0.144	0.137	-1.181	-1.188	3
11	-0.139	1.281	-0.453	0.689	39	-0.139	0.214	-1.208	-1.134	4
12	0.031	0.296	-0.005	0.322	37	0.031	0.049	-0.013	0.068	19
13	0.084	-0.227	0.219	0.075	34	0.084	-0.038	0.584	0.630	26
14	0.103	-0.494	0.320	-0.071	25	0.103	-0.082	0.852	0.873	36
15	0.109	-0.552	0.368	-0.075	21	0.109	-0.092	0.981	0.998	40
16	-0.109	0.415	-0.361	-0.056	26	-0.109	0.069	-0.964	-1.004	8
17	-0.012	-0.139	-0.093	-0.244	5	-0.012	-0.023	-0.249	-0.283	17
18	0.060	-0.392	0.147	-0.185	9	0.060	-0.065	0.392	0.387	23
19	0.087	-0.496	0.261	-0.148	12	0.087	-0.083	0.696	0.700	31
20	0.098	-0.537	0.317	-0.123	14	0.098	-0.090	0.844	0.853	35
21	-0.122	0.331	-0.256	-0.047	27	-0.122	0.055	-0.682	-0.748	11
22	-0.079	0.298	-0.244	-0.025	29	-0.079	0.050	-0.651	-0.681	14
23	0.036	-0.002	0.022	0.056	32	0.036	0.000	0.059	0.095	20
24	0.078	-0.207	0.185	0.056	31	0.078	-0.035	0.493	0.537	24
25	0.091	-0.393	0.257	-0.045	28	0.091	-0.066	0.686	0.711	32
26	-0.095	0.071	-0.234	-0.259	4	-0.095	0.012	-0.625	-0.708	13
27	-0.017	-0.236	-0.075	-0.329	1	-0.017	-0.039	-0.200	-0.257	18
28	0.047	-0.379	0.119	-0.213	6	0.047	-0.063	0.317	0.301	21
29	0.071	-0.453	0.220	-0.162	10	0.071	-0.076	0.587	0.582	25
30	0.082	-0.489	0.272	-0.135	13	0.082	-0.082	0.726	0.726	33
33	0.086	-0.409	0.249	-0.075	22	0.086	-0.068	0.663	0.681	30
34	-0.136	0.589	-0.390	0.063	33	-0.136	0.098	-1.040	-1.078	6
35	-0.115	0.403	-0.251	0.037	30	-0.115	0.067	-0.669	-0.717	12
38	0.086	-0.428	0.233	-0.108	16	0.086	-0.071	0.622	0.637	28
39	-0.096	0.040	-0.266	-0.322	2	-0.096	0.007	-0.710	-0.800	10
40	-0.092	-0.034	-0.192	-0.317	3	-0.092	-0.006	-0.512	-0.609	15
41	-0.131	1.014	-0.396	0.487	38	-0.131	0.169	-1.055	-1.018	7
42	-0.038	-0.019	-0.128	-0.186	8	-0.038	-0.003	-0.341	-0.383	16
43	0.048	-0.375	0.138	-0.189	7	0.048	-0.063	0.369	0.355	22
44	0.094	-0.515	0.299	-0.121	15	0.094	-0.086	0.797	0.806	34
45	0.106	-0.552	0.360	-0.085	17	0.106	-0.092	0.960	0.975	38
46	0.108	-0.556	0.374	-0.074	23	0.108	-0.093	0.998	1.013	41
47	-0.159	1.622	-0.516	0.946	41	-0.159	0.270	-1.376	-1.265	2
	Profile 9					**Profile 10**				
2	0.322	-0.092	0.736	0.965	39	0.107	-0.277	0.736	0.566	39
3	0.257	-0.068	0.497	0.686	29	0.086	-0.204	0.497	0.379	30

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
4	-0.422	0.136	-0.841	-1.126	4	-0.141	0.409	-0.841	-0.572	6
5	-0.487	0.252	-1.019	-1.255	1	-0.162	0.756	-1.019	-0.426	10
7	0.317	-0.088	0.684	0.914	37	0.106	-0.263	0.684	0.527	37
8	0.256	-0.066	0.462	0.652	27	0.085	-0.199	0.462	0.349	27
9	-0.302	0.043	-0.634	-0.892	9	-0.101	0.130	-0.634	-0.604	5
10	-0.433	0.137	-0.886	-1.182	3	-0.144	0.411	-0.886	-0.619	3
11	-0.418	0.214	-0.906	-1.111	5	-0.139	0.641	-0.906	-0.405	14
12	0.094	0.049	-0.010	0.133	19	0.031	0.148	-0.010	0.169	23
13	0.252	-0.038	0.438	0.652	26	0.084	-0.114	0.438	0.408	32
14	0.310	-0.082	0.639	0.867	36	0.103	-0.247	0.639	0.496	36
15	0.327	-0.092	0.736	0.971	40	0.109	-0.276	0.736	0.569	40
16	-0.327	0.069	-0.723	-0.981	8	-0.109	0.207	-0.723	-0.625	1
17	-0.035	-0.023	-0.186	-0.245	17	-0.012	-0.069	-0.186	-0.268	18
18	0.180	-0.065	0.294	0.409	23	0.060	-0.196	0.294	0.158	22
19	0.260	-0.083	0.522	0.700	31	0.087	-0.248	0.522	0.361	28
20	0.294	-0.090	0.633	0.837	35	0.098	-0.269	0.633	0.462	35
21	-0.366	0.055	-0.511	-0.822	10	-0.122	0.165	-0.511	-0.468	9
22	-0.238	0.050	-0.488	-0.677	14	-0.079	0.149	-0.488	-0.419	11
23	0.108	0.000	0.044	0.152	20	0.036	-0.001	0.044	0.079	19
24	0.235	-0.035	0.370	0.570	24	0.078	-0.104	0.370	0.344	26
25	0.273	-0.066	0.514	0.722	33	0.091	-0.197	0.514	0.409	33
26	-0.286	0.012	-0.469	-0.743	13	-0.095	0.036	-0.469	-0.528	7
27	-0.052	-0.039	-0.150	-0.241	18	-0.017	-0.118	-0.150	-0.286	17
28	0.140	-0.063	0.238	0.315	21	0.047	-0.190	0.238	0.095	20
29	0.213	-0.076	0.440	0.578	25	0.071	-0.227	0.440	0.285	24
30	0.246	-0.082	0.544	0.709	32	0.082	-0.245	0.544	0.382	31
33	0.257	-0.068	0.497	0.686	30	0.086	-0.205	0.497	0.378	29
34	-0.408	0.098	-0.780	-1.090	6	-0.136	0.295	-0.780	-0.622	2
35	-0.346	0.067	-0.501	-0.780	12	-0.115	0.201	-0.501	-0.415	13
38	0.259	-0.071	0.467	0.655	28	0.086	-0.214	0.467	0.339	25
39	-0.289	0.007	-0.532	-0.815	11	-0.096	0.020	-0.532	-0.609	4
40	-0.275	-0.006	-0.384	-0.664	15	-0.092	-0.017	-0.384	-0.492	8
41	-0.394	0.169	-0.792	-1.017	7	-0.131	0.507	-0.792	-0.416	12
42	-0.115	-0.003	-0.256	-0.374	16	-0.038	-0.010	-0.256	-0.304	16
43	0.145	-0.063	0.277	0.359	22	0.048	-0.188	0.277	0.137	21
44	0.283	-0.086	0.598	0.796	34	0.094	-0.257	0.598	0.435	34
45	0.319	-0.092	0.720	0.948	38	0.106	-0.276	0.720	0.551	38
46	0.323	-0.093	0.748	0.979	41	0.108	-0.278	0.748	0.578	41
47	-0.477	0.270	-1.032	-1.239	2	-0.159	0.811	-1.032	-0.380	15
	Profile 11					**Profile 12**				
2	0.429	-0.370	0.245	0.304	37	0.429	-0.185	0.491	0.735	39
3	0.342	-0.272	0.166	0.236	30	0.342	-0.136	0.331	0.538	30
4	-0.562	0.546	-0.280	-0.297	10	-0.562	0.273	-0.560	-0.850	3
5	-0.650	1.007	-0.340	0.018	18	-0.650	0.504	-0.679	-0.826	4
7	0.423	-0.350	0.228	0.301	36	0.423	-0.175	0.456	0.704	37
8	0.341	-0.265	0.154	0.231	27	0.341	-0.132	0.308	0.517	27
9	-0.402	0.173	-0.211	-0.440	4	-0.402	0.087	-0.423	-0.738	7
10	-0.577	0.548	-0.295	-0.325	9	-0.577	0.274	-0.591	-0.894	1
11	-0.558	0.854	-0.302	-0.006	16	-0.558	0.427	-0.604	-0.735	8

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
12	0.125	0.197	-0.003	0.319	40	0.125	0.099	-0.006	0.217	20
13	0.336	-0.152	0.146	0.330	41	0.336	-0.076	0.292	0.552	32
14	0.413	-0.329	0.213	0.297	33	0.413	-0.165	0.426	0.675	36
15	0.436	-0.368	0.245	0.314	39	0.436	-0.184	0.491	0.743	40
16	-0.437	0.277	-0.241	-0.401	7	-0.437	0.138	-0.482	-0.780	6
17	-0.047	-0.093	-0.062	-0.202	14	-0.047	-0.046	-0.124	-0.218	18
18	0.240	-0.262	0.098	0.077	20	0.240	-0.131	0.196	0.305	23
19	0.347	-0.331	0.174	0.190	25	0.347	-0.165	0.348	0.530	29
20	0.391	-0.358	0.211	0.244	31	0.391	-0.179	0.422	0.634	35
21	-0.487	0.221	-0.170	-0.437	5	-0.487	0.110	-0.341	-0.718	10
22	-0.318	0.199	-0.163	-0.282	11	-0.318	0.099	-0.326	-0.544	15
23	0.143	-0.001	0.015	0.157	23	0.143	-0.001	0.030	0.172	19
24	0.314	-0.138	0.123	0.299	35	0.314	-0.069	0.246	0.491	25
25	0.364	-0.262	0.171	0.274	32	0.364	-0.131	0.343	0.576	33
26	-0.381	0.047	-0.156	-0.490	3	-0.381	0.024	-0.312	-0.670	12
27	-0.069	-0.158	-0.050	-0.276	12	-0.069	-0.079	-0.100	-0.248	17
28	0.187	-0.253	0.079	0.013	17	0.187	-0.126	0.159	0.219	21
29	0.284	-0.302	0.147	0.129	22	0.284	-0.151	0.293	0.427	24
30	0.328	-0.326	0.181	0.183	24	0.328	-0.163	0.363	0.528	28
33	0.343	-0.273	0.166	0.236	29	0.343	-0.136	0.332	0.538	31
34	-0.544	0.393	-0.260	-0.411	6	-0.544	0.196	-0.520	-0.868	2
35	-0.461	0.269	-0.167	-0.359	8	-0.461	0.134	-0.334	-0.661	13
38	0.346	-0.285	0.156	0.216	26	0.346	-0.143	0.311	0.514	26
39	-0.386	0.027	-0.177	-0.536	1	-0.386	0.013	-0.355	-0.727	9
40	-0.367	-0.022	-0.128	-0.517	2	-0.367	-0.011	-0.256	-0.634	14
41	-0.526	0.676	-0.264	-0.114	15	-0.526	0.338	-0.528	-0.716	11
42	-0.154	-0.013	-0.085	-0.252	13	-0.154	-0.006	-0.171	-0.331	16
43	0.194	-0.250	0.092	0.036	19	0.194	-0.125	0.184	0.253	22
44	0.378	-0.343	0.199	0.234	28	0.378	-0.172	0.399	0.605	34
45	0.426	-0.368	0.240	0.298	34	0.426	-0.184	0.480	0.722	38
46	0.431	-0.371	0.249	0.310	38	0.431	-0.185	0.499	0.745	41
47	-0.636	1.081	-0.344	0.101	21	-0.636	0.541	-0.688	-0.784	5
	Profile 13					**Profile 14**				
2	0.214	-0.370	0.491	0.335	39	0.536	-0.231	0.307	0.611	39
3	0.171	-0.272	0.331	0.231	29	0.428	-0.170	0.207	0.465	31
4	-0.281	0.546	-0.560	-0.296	10	-0.703	0.341	-0.350	-0.712	3
5	-0.325	1.007	-0.679	0.003	18	-0.812	0.630	-0.425	-0.607	11
7	0.212	-0.350	0.456	0.317	37	0.529	-0.219	0.285	0.595	37
8	0.171	-0.265	0.308	0.214	27	0.427	-0.166	0.193	0.454	28
9	-0.201	0.173	-0.423	-0.450	4	-0.503	0.108	-0.264	-0.658	7
10	-0.289	0.548	-0.591	-0.331	8	-0.722	0.342	-0.369	-0.748	2
11	-0.279	0.854	-0.604	-0.029	16	-0.697	0.534	-0.377	-0.541	14
12	0.063	0.197	-0.006	0.253	31	0.156	0.123	-0.004	0.275	23
13	0.168	-0.152	0.292	0.308	36	0.420	-0.095	0.183	0.507	34
14	0.207	-0.329	0.426	0.303	35	0.516	-0.206	0.266	0.577	36
15	0.218	-0.368	0.491	0.341	40	0.545	-0.230	0.307	0.622	41
16	-0.218	0.277	-0.482	-0.424	5	-0.546	0.173	-0.301	-0.674	6
17	-0.024	-0.093	-0.124	-0.240	14	-0.059	-0.058	-0.078	-0.194	18
18	0.120	-0.262	0.196	0.055	20	0.300	-0.163	0.123	0.259	22

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
19	0.173	-0.331	0.348	0.191	24	0.433	-0.207	0.218	0.444	26
20	0.196	-0.358	0.422	0.260	32	0.489	-0.224	0.264	0.529	35
21	-0.244	0.221	-0.341	-0.364	7	-0.609	0.138	-0.213	-0.684	5
22	-0.159	0.199	-0.326	-0.286	12	-0.397	0.124	-0.203	-0.477	15
23	0.072	-0.001	0.030	0.100	22	0.179	-0.001	0.018	0.197	20
24	0.157	-0.138	0.246	0.265	34	0.392	-0.086	0.154	0.460	29
25	0.182	-0.262	0.343	0.263	33	0.456	-0.164	0.214	0.506	32
26	-0.191	0.047	-0.312	-0.456	3	-0.477	0.030	-0.195	-0.642	8
27	-0.034	-0.158	-0.100	-0.292	11	-0.086	-0.099	-0.063	-0.247	17
28	0.093	-0.253	0.159	-0.001	17	0.233	-0.158	0.099	0.174	19
29	0.142	-0.302	0.293	0.133	23	0.356	-0.189	0.183	0.350	24
30	0.164	-0.326	0.363	0.201	26	0.410	-0.204	0.227	0.433	25
33	0.171	-0.273	0.332	0.230	28	0.429	-0.170	0.207	0.465	30
34	-0.272	0.393	-0.520	-0.399	6	-0.680	0.246	-0.325	-0.760	1
35	-0.230	0.269	-0.334	-0.296	9	-0.576	0.168	-0.209	-0.617	10
38	0.173	-0.285	0.311	0.199	25	0.432	-0.178	0.194	0.448	27
39	-0.193	0.027	-0.355	-0.521	1	-0.482	0.017	-0.222	-0.687	4
40	-0.183	-0.022	-0.256	-0.461	2	-0.458	-0.014	-0.160	-0.632	9
41	-0.263	0.676	-0.528	-0.115	15	-0.657	0.423	-0.330	-0.565	12
42	-0.077	-0.013	-0.171	-0.260	13	-0.192	-0.008	-0.107	-0.307	16
43	0.097	-0.250	0.184	0.031	19	0.242	-0.156	0.115	0.201	21
44	0.189	-0.343	0.399	0.245	30	0.472	-0.214	0.249	0.507	33
45	0.213	-0.368	0.480	0.325	38	0.532	-0.230	0.300	0.603	38
46	0.216	-0.371	0.499	0.344	41	0.539	-0.232	0.312	0.619	40
47	-0.318	1.081	-0.688	0.075	21	-0.795	0.676	-0.430	-0.550	13
	Profile 15					**Profile 16**				
2	0.268	-0.462	0.307	0.112	30	0.268	-0.231	0.613	0.650	39
3	0.214	-0.340	0.207	0.081	26	0.214	-0.170	0.414	0.458	31
4	-0.351	0.682	-0.350	-0.020	17	-0.351	0.341	-0.701	-0.711	3
5	-0.406	1.259	-0.425	0.428	40	-0.406	0.630	-0.849	-0.626	7
7	0.265	-0.438	0.285	0.112	29	0.265	-0.219	0.570	0.615	37
8	0.213	-0.331	0.193	0.075	24	0.213	-0.166	0.385	0.433	27
9	-0.251	0.217	-0.264	-0.299	5	-0.251	0.108	-0.528	-0.671	5
10	-0.361	0.685	-0.369	-0.045	16	-0.361	0.342	-0.738	-0.756	1
11	-0.349	1.068	-0.377	0.341	39	-0.349	0.534	-0.755	-0.570	11
12	0.078	0.246	-0.004	0.320	38	0.078	0.123	-0.008	0.193	21
13	0.210	-0.190	0.183	0.203	37	0.210	-0.095	0.365	0.480	32
14	0.258	-0.412	0.266	0.113	31	0.258	-0.206	0.533	0.585	36
15	0.273	-0.460	0.307	0.119	34	0.273	-0.230	0.613	0.656	40
16	-0.273	0.346	-0.301	-0.228	7	-0.273	0.173	-0.602	-0.703	4
17	-0.029	-0.116	-0.078	-0.223	8	-0.029	-0.058	-0.155	-0.243	18
18	0.150	-0.327	0.123	-0.054	15	0.150	-0.163	0.245	0.232	23
19	0.217	-0.413	0.218	0.021	19	0.217	-0.207	0.435	0.445	28
20	0.245	-0.448	0.264	0.061	23	0.245	-0.224	0.528	0.548	35
21	-0.305	0.276	-0.213	-0.242	6	-0.305	0.138	-0.426	-0.593	9
22	-0.199	0.249	-0.203	-0.154	12	-0.199	0.124	-0.407	-0.481	15
23	0.090	-0.001	0.018	0.107	28	0.090	-0.001	0.037	0.126	19
24	0.196	-0.173	0.154	0.177	35	0.196	-0.086	0.308	0.418	25
25	0.228	-0.328	0.214	0.114	32	0.228	-0.164	0.429	0.492	33

Table 38 Reinspection Decision Policies' Aggregate Standardized Mean Squared Values for All Preference Profiles (cont.)

Policy	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank	$w_C z_i^C$	$w_S z_i^S$	$w_D z_i^D$	z_i^A	Rank
26	-0.238	0.059	-0.195	-0.374	3	-0.238	0.030	-0.390	-0.599	8
27	-0.043	-0.197	-0.063	-0.303	4	-0.043	-0.099	-0.125	-0.267	17
28	0.117	-0.316	0.099	-0.100	13	0.117	-0.158	0.198	0.157	20
29	0.178	-0.378	0.183	-0.017	18	0.178	-0.189	0.367	0.356	24
30	0.205	-0.408	0.227	0.024	20	0.205	-0.204	0.454	0.455	29
33	0.214	-0.341	0.207	0.081	25	0.214	-0.170	0.414	0.458	30
34	-0.340	0.491	-0.325	-0.174	10	-0.340	0.246	-0.650	-0.745	2
35	-0.288	0.336	-0.209	-0.161	11	-0.288	0.168	-0.418	-0.538	14
38	0.216	-0.357	0.194	0.054	21	0.216	-0.178	0.389	0.427	26
39	-0.241	0.034	-0.222	-0.429	1	-0.241	0.017	-0.444	-0.668	6
40	-0.229	-0.028	-0.160	-0.417	2	-0.229	-0.014	-0.320	-0.563	13
41	-0.329	0.845	-0.330	0.187	36	-0.329	0.423	-0.660	-0.566	12
42	-0.096	-0.016	-0.107	-0.219	9	-0.096	-0.008	-0.213	-0.317	16
43	0.121	-0.313	0.115	-0.077	14	0.121	-0.156	0.231	0.195	22
44	0.236	-0.429	0.249	0.057	22	0.236	-0.214	0.498	0.520	34
45	0.266	-0.460	0.300	0.106	27	0.266	-0.230	0.600	0.636	38
46	0.270	-0.463	0.312	0.118	33	0.270	-0.232	0.624	0.662	41
47	-0.398	1.351	-0.430	0.524	41	-0.398	0.676	-0.860	-0.582	10

APPENDIX F

ANALYSIS OF VARIANCE DETAILS REGARDING OUTPUT MEASURES OF RECOMMENDED POLICY

I. Normal Plot of Residuals, Residuals versus Fits Plot and Tukey Comparisons of Final Analysis of Variance Performed for Policy 39 Total Cost Measures

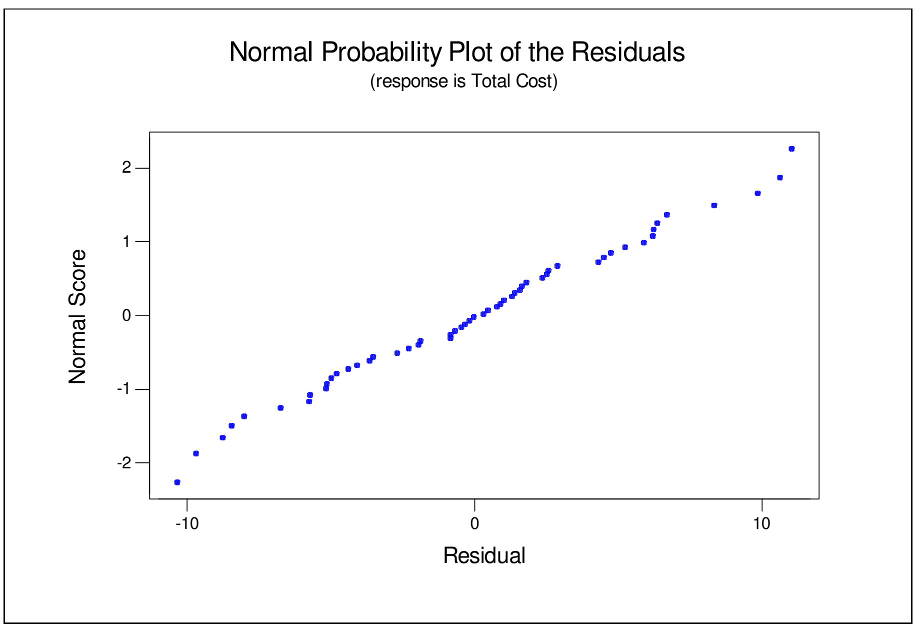

Figure 18 Normal Probability Plot for Policy 39 Total Cost Measures

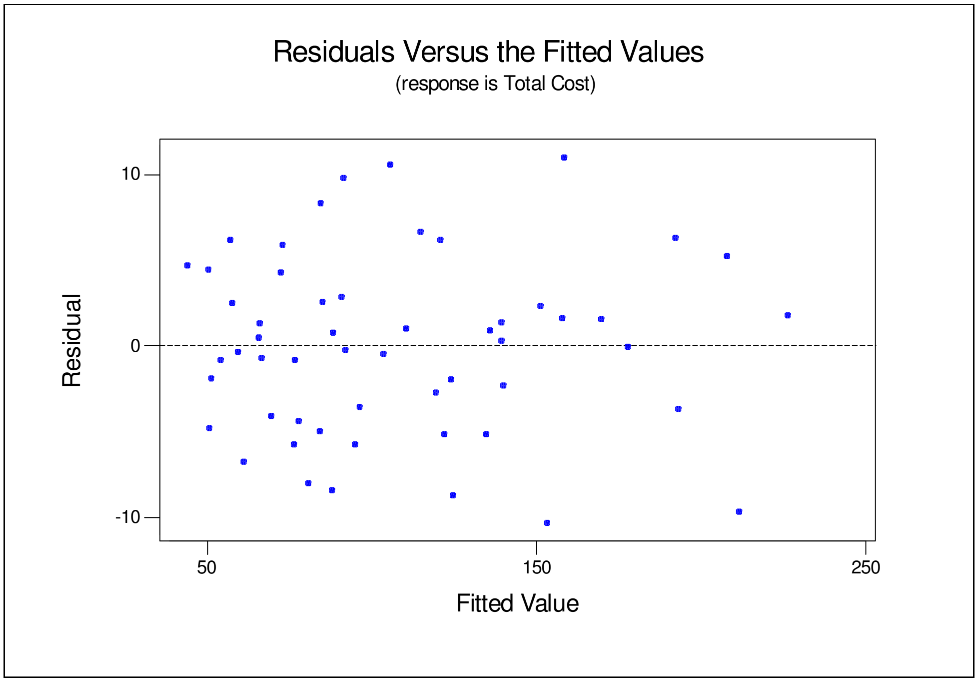

Figure 19 Residual versus Fits Plot for Policy 39 Total Cost Measures

Explanations of notations used in MINITAB Tukey Comparison Outputs in this Appendix;

pt: Testing detection probability of a defect

pd: Probability that a given defect is difficult

pm: Probability that a given defect is major

pj: Inspector capability of a particular inspector (pj)

p1, p2: Inspection detection probabilities of difficult and easy defects (p1, p2)

k: Number of inspectors

k`: Number of reinspectors

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Total Cost

pt = 1 subtracted from:
Level  Difference       SE of             Adjusted
pt       of Means  Difference   T-Value    P-Value
2          -26.65       1.723    -15.46     0.0000

pd = 1 subtracted from:
Level  Difference       SE of             Adjusted
pd       of Means  Difference   T-Value    P-Value
2           45.44       2.985     15.22     0.0000
3           67.72       2.985     22.69     0.0000
pd = 2 subtracted from:
Level  Difference       SE of             Adjusted
pd       of Means  Difference   T-Value    P-Value
3           22.28       2.985     7.463     0.0000

pm = 1 subtracted from:
Level  Difference       SE of             Adjusted
pm       of Means  Difference   T-Value    P-Value
2           18.69       2.985     6.260     0.0000
3           26.19       2.985     8.773     0.0000
pm = 2 subtracted from:
Level  Difference       SE of             Adjusted
pm       of Means  Difference   T-Value    P-Value
3           7.502       2.985     2.513     0.0428

qj = 1 subtracted from:
Level  Difference       SE of             Adjusted
qj       of Means  Difference   T-Value    P-Value
2          -18.51       2.111     -8.77     0.0000
3          -33.89       2.111    -16.06     0.0000
qj = 2 subtracted from:
Level  Difference       SE of             Adjusted
qj       of Means  Difference   T-Value    P-Value
3          -15.38       2.111    -7.287     0.0000

p1,p2 = 1 subtracted from:
Level     Difference       SE of             Adjusted
p1,p2       of Means  Difference   T-Value    P-Value
2             -14.76       2.585     -5.71     0.0000
3             -49.38       2.585    -19.10     0.0000
p1,p2 = 2 subtracted from:
Level     Difference       SE of             Adjusted
p1,p2       of Means  Difference   T-Value    P-Value
3             -34.62       2.585    -13.39     0.0000

k = 1 subtracted from:
Level Difference       SE of             Adjusted
k       of Means  Difference   T-Value    P-Value
2         -12.20       2.111     -5.78     0.0000
3         -31.28       2.111    -14.82     0.0000
k = 2 subtracted from:
Level Difference       SE of             Adjusted
k       of Means  Difference   T-Value    P-Value
3         -19.08       2.111    -9.037     0.0000
```

Figure 20 MINITAB Output for Significant Factors' Tukey Comparisons of Policy 39 Total Cost Measures

II. Normal Plot of Residuals, Residuals versus Fits Plot and Tukey Comparisons of Final Analysis of Variance Performed for Policy 39 Schedule Delay Measures

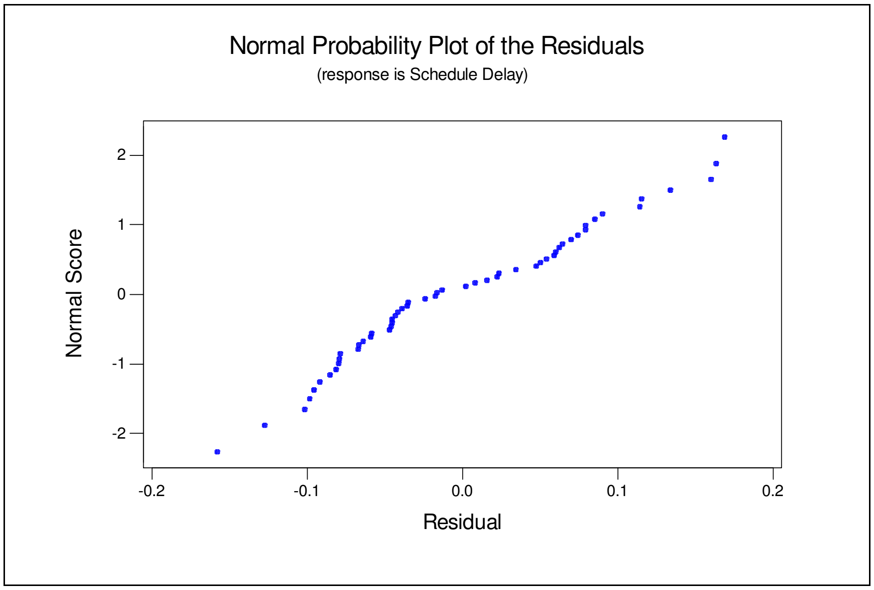

Figure 21 Normal Probability Plot for Policy 39 Schedule Delay Measures

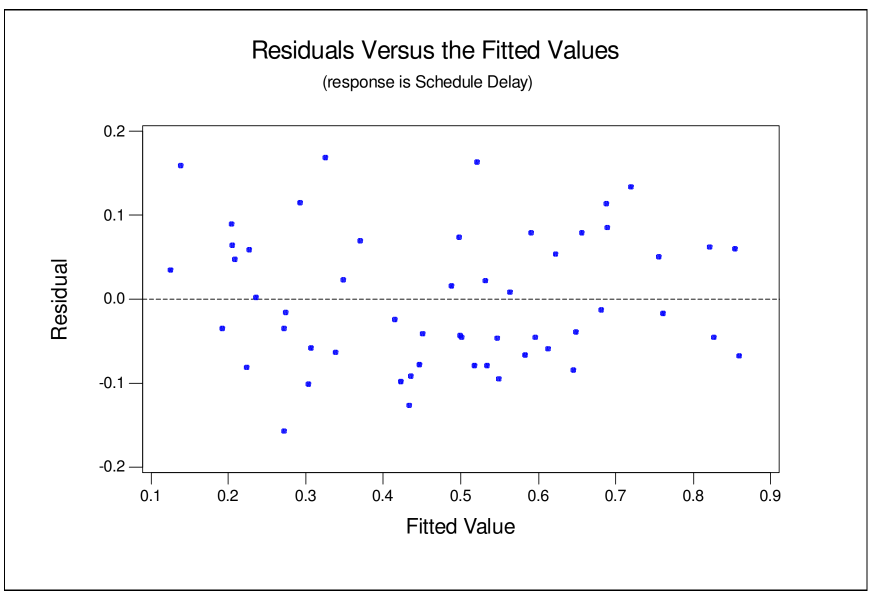

Figure 22 Residual versus Fits Plot for Policy 39 Schedule Delay Measures

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Schedule Delay

pm = 1 subtracted from:
Level  Difference       SE of             Adjusted
pm        of Means  Difference   T-Value    P-Value
2          0.07978     0.02954     2.700     0.0262
3          0.21389     0.02954     7.240     0.0000
pm = 2 subtracted from:
Level  Difference       SE of             Adjusted
pm        of Means  Difference   T-Value    P-Value
3           0.1341     0.02954     4.540     0.0001

qj = 1 subtracted from:
Level  Difference       SE of             Adjusted
qj        of Means  Difference   T-Value    P-Value
2         -0.03222     0.02954    -1.091     0.5249
3         -0.09811     0.02954    -3.321     0.0051
qj = 2 subtracted from:
Level  Difference       SE of             Adjusted
qj        of Means  Difference   T-Value    P-Value
3         -0.06589     0.02954    -2.230     0.0774

p1,p2 = 1 subtracted from:
Level     Difference       SE of             Adjusted
p1,p2       of Means  Difference   T-Value    P-Value
2             0.1499     0.02954     5.074     0.0000
3             0.1938     0.02954     6.559     0.0000
p1,p2 = 2 subtracted from:
Level     Difference       SE of             Adjusted
p1,p2       of Means  Difference   T-Value    P-Value
3            0.04389     0.02954     1.486     0.3078

k = 1 subtracted from:
Level Difference       SE of             Adjusted
k       of Means  Difference   T-Value    P-Value
2        -0.0052     0.02954    -0.177     0.9829
3        -0.2951     0.02954    -9.989     0.0000
k = 2 subtracted from:
Level Difference       SE of             Adjusted
k       of Means  Difference   T-Value    P-Value
3        -0.2899     0.02954    -9.813     0.0000

k` = 1 subtracted from:
Level  Difference       SE of             Adjusted
k`        of Means  Difference   T-Value    P-Value
2           0.1263     0.02954     4.276     0.0003
3           0.1267     0.02954     4.288     0.0003
k` = 2 subtracted from:
Level  Difference       SE of             Adjusted
k`        of Means  Difference   T-Value    P-Value
3         0.000333     0.02954   0.01128     0.9999
```

Figure 23 MINITAB Output for Significant Factors' Tukey Comparisons of Policy 39 Schedule Delay Measures

III. Normal Plot of Residuals, Residuals versus Fits Plot and Tukey Comparisons of Final Analysis of Variance Performed for Policy 39 Major Field Defects Measures

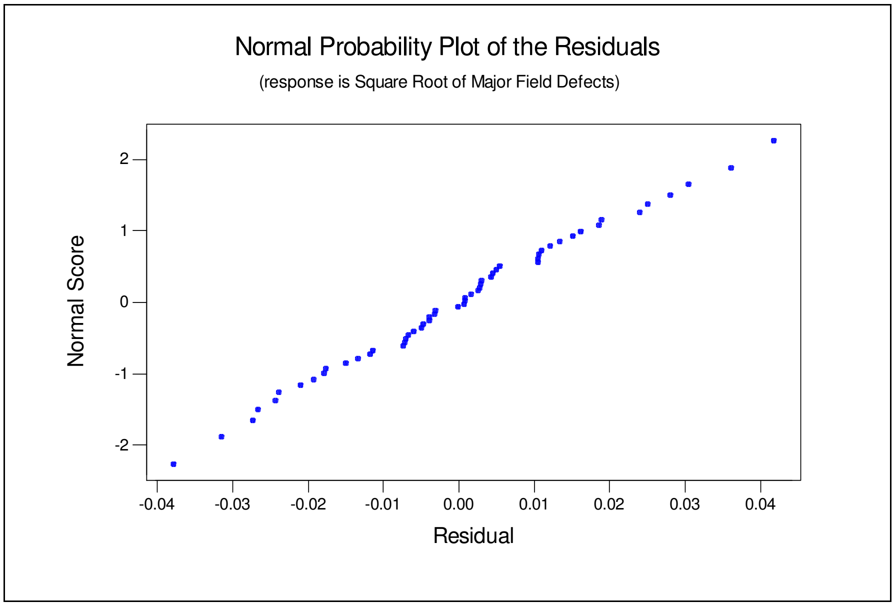

Figure 24 Normal Probability Plot for Policy 39 Major Field Defects Measures

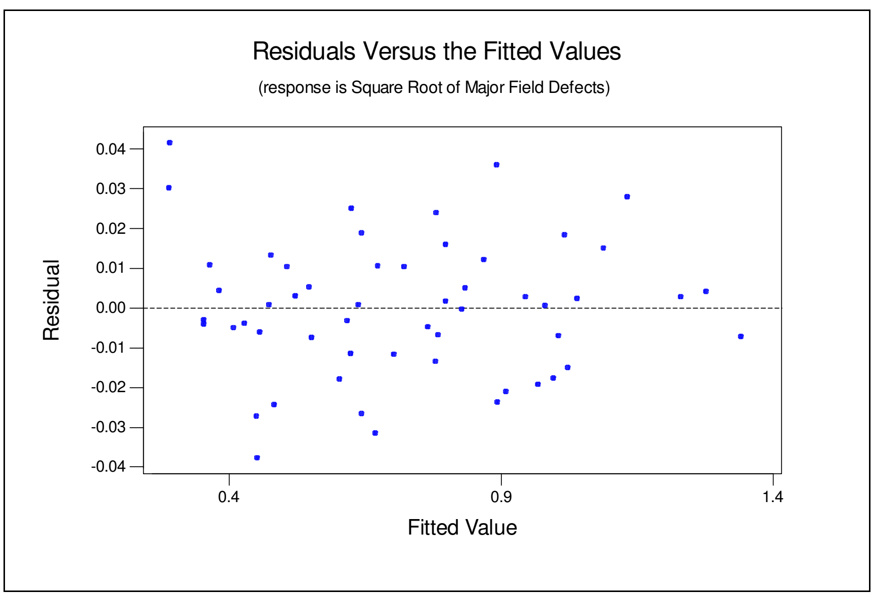

Figure 25 Residual versus Fits Plot for Policy 39 Major Field Defects Measures

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable SQRT Major Field Defects

pt = 1 subtracted from:
Level  Difference       SE of              Adjusted
pt       of Means  Difference    T-Value    P-Value
2         -0.1858    0.006419     -28.95     0.0000
pd = 1 subtracted from:
Level  Difference       SE of              Adjusted
pd       of Means  Difference    T-Value    P-Value
2          0.2167     0.01112      19.49     0.0000
3          0.3017     0.01112      27.14     0.0000
pd = 2 subtracted from:
Level  Difference       SE of              Adjusted
pd       of Means  Difference    T-Value    P-Value
3         0.08504     0.01112      7.649     0.0000
pm = 1 subtracted from:
Level  Difference       SE of              Adjusted
pm       of Means  Difference    T-Value    P-Value
2          0.2652     0.01112      23.85     0.0000
3          0.3876     0.01112      34.87     0.0000
pm = 2 subtracted from:
Level  Difference       SE of              Adjusted
pm       of Means  Difference    T-Value    P-Value
3          0.1224     0.01112      11.01     0.0000
qj = 1 subtracted from:
Level  Difference       SE of              Adjusted
qj       of Means  Difference    T-Value    P-Value
2         -0.0748    0.007861      -9.52     0.0000
3         -0.1316    0.007861     -16.74     0.0000
qj = 2 subtracted from:
Level  Difference       SE of              Adjusted
qj       of Means  Difference    T-Value    P-Value
3        -0.05677    0.007861     -7.222     0.0000
p1,p2 = 1 subtracted from:
Level     Difference       SE of              Adjusted
p1,p2       of Means  Difference    T-Value    P-Value
2             0.0147     0.01112       1.32     0.3943
3            -0.1807     0.01112     -16.25     0.0000
p1,p2 = 2 subtracted from:
Level     Difference       SE of              Adjusted
p1,p2       of Means  Difference    T-Value    P-Value
3            -0.1954     0.01112     -17.58     0.0000
k = 1 subtracted from:
Level Difference       SE of              Adjusted
k       of Means  Difference    T-Value    P-Value
2        -0.0953    0.007861     -12.13     0.0000
3        -0.1077    0.007861     -13.70     0.0000
k = 2 subtracted from:
Level Difference       SE of              Adjusted
k       of Means  Difference    T-Value    P-Value
3       -0.01234    0.007861     -1.569     0.2754
k` = 1 subtracted from:
Level  Difference       SE of              Adjusted
k`       of Means  Difference    T-Value    P-Value
2         -0.0518     0.01112      -4.66     0.0002
3         -0.1884     0.01112     -16.94     0.0000
k` = 2 subtracted from:
Level  Difference       SE of              Adjusted
k`       of Means  Difference    T-Value    P-Value
3         -0.1365     0.01112     -12.28     0.0000
```

Figure 26 MINITAB Output for Significant Factors' Tukey Comparisons of Policy 39 Major Field Defects Measures