

**BAYESIAN STATISTICS AND PRODUCTION RELIABILITY
ASSESSMENTS FOR MINING OPERATIONS**

by

Gaurav Kumar Sharma
B.E., Thapar University (India), 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate Studies

(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2008

© Gaurav Kumar Sharma, 2008

Abstract

This thesis presents a novel application of structural reliability concepts to assess the reliability of mining operations. “Limit-states” are defined to obtain the probability that the total productivity – measured in production time or economic gain – exceeds user-selected thresholds. Focus is on the impact of equipment downtime and other non-operating instances on the productivity and the economic costs of the operation. A comprehensive set of data gathered at a real-world mining facility is utilized to calibrate the probabilistic models. In particular, the utilization of Bayesian inference facilitates the inclusion of data – and updating of the production probabilities – as they become available. The thesis includes a detailed description of the Bayesian approach, as well as the limit-state-based reliability methodology. A comprehensive numerical example demonstrates the methodology and the usefulness of the probabilistic results.

Table of Contents

<i>Abstract</i>	<i>ii</i>
<i>Table of Contents</i>	<i>iii</i>
<i>List of Tables</i>	<i>iv</i>
<i>List of Figures</i>	<i>v</i>
<i>List of Symbols</i>	<i>vii</i>
<i>List of Abbreviations</i>	<i>viii</i>
<i>Acknowledgements</i>	<i>ix</i>
1 Introduction	1
1.1. Literature Review	4
2 Data Analysis For Mining Equipment	7
2.1. The Nature of the Raw Data	8
2.2. Modeling of Equipment States.....	13
2.3. Treatment of Uncertainty: A Bayesian Approach.....	16
2.4. The Predictive Analysis	20
3 Reliability Analysis	45
3.1. Reliability Formulation	46
3.2. Results and Applications	53
3.2.1. Production Time.....	54
3.2.2. Economic and Unscheduled Repair Costs	57
4 Discussion and Conclusions	76
<i>References</i>	<i>77</i>
APPENDIX A	80
A.1 Random Variable	81
A.2 Cumulative Distribution Function	82
A.3 Probability Mass Function	83
A.4 Probability Density Function	84
A.5 Types of Uncertainties	85
A.6 Basic Stochastic Processes.....	86
A.6.1 The Bernoulli Sequence:.....	86
A.6.2 The Poisson Process - Random occurrence model	89
APPENDIX B	90

List of Tables

Table 1: Classification of equipment states	24
Table 2: Type of components for repairs.	25
Table 3: Number of repairs observed in a year for the four equipments	25
Table 4: Time between successive scheduled repair events (in days)	26
Table 5: Statistics for time between successive UR sub-states.....	26
Table 6: Probability distributions for duration in sub-state for various equipments.	27
Table 7: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6161.	28
Table 8: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6162.	29
Table 9: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6163.	30
Table 10: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment 6164. 31	
Table 11: Bayesian updated mean occurrence rate for the random non-operating sub-states of equipments.	32
Table 12: Costs associated with different states of functioning of equipment.	61
Table 13: Empirical Productivity and Economic Cost Data.....	61

List of Figures

Figure 1: Excerpts from a typical equipment dispatch report.....	33
Figure 2: Relative frequency histogram, exponential and lognormal PDF for the time between successive UR sub-states	34
Figure 3: Cumulative frequency histogram, Exponential and Lognormal CDF for the time ..	35
Figure 4: Relative frequency histograms for UR sub-state duration	36
Figure 5: Cumulative frequency plot for the UR sub-state duration	37
Figure 6: Lognormal fit for the duration (time in state - TIS) of UR sub-state	38
Figure 7: An example of a time-window with several state transitions.....	40
Figure 8: Poisson probability mass function using a point estimate of rate of occurrences of UR sub-state.....	41
Figure 9: Number of Poisson occurrences with predictive and point estimation of sub-state occurrence rate for equipment 6161.	42
Figure 10: Predictive-CDF plot for the failure of equipment number 6161	43
Figure 11: Reliability plot for equipment 6161, Occurrence rate ~ 0.4	44
Figure 12: Methodology of calculating the total time in non-operation, $T_{NO}(\mathbf{x})$ for an equipment.....	62
Figure 13: Flowchart illustrating an algorithm for sampling time spent in non-operating state	63
Figure 14: Sampling results - PDF plot for the total production time (units: truck-days) of the four HEMM equipments in half-year.	64

Figure 15: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in half-year.	65
Figure 16: PDF plot for the total production time (units: truck-days) of the four HEMM equipments in one-year.	66
Figure 17: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in one-year.	67
Figure 18: PDF plot for the total production time (units: truck-days) of the four HEMM equipments in two-years.	68
Figure 19: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in two-years.	69
Figure 20: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in half-year.	70
Figure 21: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in half-year.	71
Figure 22: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in one-year.	72
Figure 23: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in one-year.	73
Figure 24: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in two years.	74
Figure 25: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in two years.	75

List of Symbols

λ_{UR}	Mean occurrence state rate of the unscheduled repair sub-state
T	Total Observation Period
$T_{NO}(x)$	Vector of random variables for time in non-operation
$T_O(x)$	Vector of random variables for time in operation
x	Name of a state
t_x	Time of Occurrence of state, x
d_x	Duration of state, x
λ_x	Mean occurrence state rate of state, x
N_x	Number of Poisson occurrences of state, x
k	Parameter 1 of Gamma distribution
ν	Parameter 2 of Gamma distribution
μ_λ	Mean of occurrence rate
δ_λ	Coefficient of variation of occurrence rate
T_{bf}	Time-between-failures
T_θ	Time threshold
$g(x)$	Expression for limit-state function
N	Total number of sub-states
C	Normalizing constant
C_{OP}	Operating cost of equipment per unit time
C_{UR}	Unscheduled repair cost of equipment per unit time
C_{SR}	Scheduled repair cost of equipment per unit time
C_E	Economic (Opportunity) cost of equipment per unit time
$C(x)$	Vector of random variables for economic and repair cost
C_θ	Cost threshold
D_x	Total time spent in random state x in the observation period
D	Total time spent in deterministic state in the observation period
C	Normalizing Constant
E	An event
$F(\lambda)$	Probability density function for occurrence rate, λ
$f'(\lambda)$	Prior distribution for λ
$F''(\lambda)$	Posterior distribution for λ
$P(A B)$	Conditional probability that event A will occur given that event B has occurred
$L(\lambda)$	Likelihood function for λ

List of Abbreviations

PDF	Probability Density Function
CDF	Cumulative Distribution Function
PMF	Probability Mass Function
HEMM	Heavy Earth Moving Machinery
UR	Unscheduled Repair

Acknowledgements

I would like to acknowledge the assistance, support and encouragement that have been provided by the following people and institution, without whom, this research would not have been possible.

- Dr. Terje Haukaas
- Dr. R. A. Hall
- Highland Copper Valley Mine
- The University of British Columbia
- Natural Sciences and Engineering Research Council of Canada, Collaborative Research and Development Grants

I owe particular thanks to Dr. Haukaas, whose enthusiasm and dedication towards his work and students will always inspire me in my career ahead.

Special thanks are also owed to my parents and my wife for supporting me in my education.

1 Introduction

The objective of this thesis is to combine statistical analysis of data recorded at a real mining site with modern reliability methods to predict the future productivity of a mining operation. Focus is on the impact of equipment downtime and other non-operating instances on the productivity and the economic costs of the operation. The probability of equipment being non-operative is influenced by a number of uncertain factors. These include local weather, geological conditions, human factors, uncertain degree of machine deterioration, and inherent uncertainty in material and equipment quality. For example, equipment may become intermittently non-operative due to poor weather conditions such as snow, rain, low visibility, and excessive humidity. Similarly, varying geological conditions such as difficult mountainous terrain, rocks and boulders inflict higher-than-normal equipment deterioration; thus increasing the likelihood of breakdown. Human factors such as delays, non-availability of operator, inadequate equipment handling, and comprehensive equipment usage also may result in a higher number of non-operative instances, leading to low operational reliability.

In today's modern mining setups, real time data gathering systems are installed on equipment in order to monitor the various states of equipment operation. The data, upon analysis, serves to provide necessary feedback for strategic decision-making. For example, the data supports decisions concerning operational processes, production output, cost management, etc. The methods presented in this thesis provide novel tools for analyzing the equipment production data and for reliability assessment of production.

The case under consideration is the hauling of ore using heavy earth moving machinery (HEMM). This equipment is increasingly deployed because of their large hauling

capacity. This yields time-saving during loading compared to using smaller equipment, and carries reduced operator cost. However, the risk of deploying HEMM equipment is larger because of the high consequence of failure of one HEMM. It is necessary that downtimes and idle-time (time in non-operation) of such equipments are minimal because loss of production results in high economic costs. Importantly, the overall cost of an earth moving mining operation is significantly affected by the reliability of the HEMM equipments used in the operation. Lower reliability of the equipment alters productivity and drives the maintenance and operating costs; thus affecting profit margins. This motivates the focus on productivity and cost-based limit-states in the subsequent reliability analysis.

A comprehensive set of observations gathered at the Highland Valley Copper mine (Priyadarshini et al. 2005) are processed into a database and subjected to Bayesian statistical analysis. The observations include the occurrence and duration characteristics of the various states of functioning of a set of HEMM equipments. The database classifies the observations according to the stochastic properties of the states. The preliminary analysis of data is done using probabilistic modeling of occurrence and duration patterns of the different states of equipment functioning. Homogeneous Poisson process is adopted as the random occurrence model. Bayesian inference is used for addressing the uncertainty in occurrence rates of the equipment states. The utilization of Bayesian inference facilitates the inclusion of data – and updating of the equipment state occurrence rates – as they become available. The thesis includes a detailed description of the Bayesian approach. The results of the Bayesian analysis include statistics for the occurrence of various states of equipment functioning, which subsequently enter into a reliability formulation.

The reliability analysis comprises innovative utilization of concepts of structural reliability analysis to assess the production probabilities of the mining operation. Structural reliability problems, such as design code calibration and assessment of failure probabilities of structures are solved using a limit-state-based methodology. A limit-state function is a mathematical expression to represent an event of interest. The limit-state function is designated by $g(\mathbf{x})$, where \mathbf{x} is a vector of random variables. Negative and positive values of g signify the failure state and the safe state, respectively. The reliability of system is determined by first computing the failure probabilities of system components. Then the various failure modes (limit-states) of the components are determined. The reliability of the components is computed by integrating the joint probabilities of load and resistance over the failure domain, which is defined by $g(\mathbf{x}) \leq 0$ (Melchers 1999). Typically, the system is either a series system of parallel sub-systems (sub-parallel systems known as “cut sets”) or parallel system of series system (known as “link sets”). The overall system reliability is obtained from the failure probabilities of system components constituting the cut-sets.

In this thesis, limit-state-based reliability formulations are developed to study the impact of equipment downtime and other non-operating instances on the production times and the economic costs (loss of opportunity) of operation. In particular, limit-state functions are defined to obtain the probability that the total productivity – measured in production time or economic gain – exceeds user-selected thresholds. It is noted that the presented reliability formulations also provide a novel way of addressing the system reliability problem. In order to assess production reliability of a fleet of equipment, the contribution to production from the equipment is calculated by considering equipment downtime and idle-time in contrast to

considering the explicit failure or non-failure combinatorial of the equipment. This circumvents the traditional system reliability problem.

1.1. Literature Review

Today, large-sized mining equipment are increasingly being used for ore-hauling process. This is primarily to increase productivity and reduce operational costs of the process. The impact of sudden failures of such large and complex equipments on the cost and production is significant. Mitigation of this necessitates improvement in the assessment techniques of equipment reliability. In the following, the state-of-the-art in the field of reliability assessment of mine equipment productivity is discussed.

Reliability is defined as the probability that a unit will perform its intended function until a specified point in time under encountered use conditions. Importantly, the environment in which a product operates is a critical factor in evaluating a product's reliability (Meeker and Escobar 2004, p.1). Availability is the probability that the unit is operating properly when its use is intended, i.e., it is not undergoing repair action, etc. Maintainability is the ability of the unit to be retained in, or restored to a specified state when the maintenance action is performed on the unit. The three concepts are interlinked and form the primary standard in reliability-availability-maintainability (RAM) based approaches for acquiring quality products (Department of Defense 2005).

Specific studies have been performed to analyze the impact of heavy ore-hauling equipment on mine production. One such method described by Yuriy and Vayenas (2007) involves reliability assessment of load-haul-dump using genetic algorithms and discrete event simulation models. A typical mine development cycle is emulated for analyzing the effect of

load-haul-dump equipment failures on production throughput, mechanical availability and equipment utilization. Gupta et al. (1999) describe a method for analyzing productivity of surface mining systems. According to the method, a production process remains in only one of two states: working or non-working. Transition from one state to the other is defined as a stochastic process and the state probabilities of the process are determined using the Markov modeling technique.

The production attained from an equipment is dependent upon the reliability of equipment, which is dependent on factors including the operating environment of equipment, equipment's maintenance quality, its handling, etc. Vagenes and Nuziale (2001) present a reliability assessment model based on genetic algorithms, which considers the dependence of equipment reliability on equipment's age, operating environment, number and quality of repairs for mining equipment. Impact of these factors is incorporated in a genetic algorithm to assess equipment reliability and predict future time-between-failures.

An important basis for reliability computations is the collection and analysis of appropriate data. Methodologies for gathering and analyzing reliability data in mining environment are well developed (Hall and Daneshmend 2003). Meeker and Escobar (2004) describe the use of Bayesian techniques for analyzing failure data. The analysis involves combining prior information with available field data for quantifying the uncertainty. Apeland et al. (2003) combine a Bayesian approach and Monte-Carlo simulation techniques to deal with component failures.

The concept of delay-time considers a component to be present in one of the three states, i.e., non-defective, defective state and failure state. Failure propagation involves changes of state between non-defective to defective state and then into failure state. Delay-

time is the time spent by the equipment from defective to failure state, if it is not inspected. Apeland et al. (2003) describe approaches for modeling inspection times, time to defect and delay-times using fully subjective (Bayesian) approach and Monte-Carlo simulation techniques.

The impact of equipment downtime on the cost of operation is an area of great interest for mine planning engineers. Vorster and Garza (2008) present a detailed methodology for quantifying and categorizing the different costs associated with equipment downtime and lack of availability. The cost impacts of failure of an equipment on the productivity and cost effectiveness of other equipments and cost impact of capital assets becoming idle due to equipment downtime are two of the consequential cost quantities addressed by the authors. In particular, significant literature relating to reliability assessment of mining equipments is available when compared to the reliability assessment of production from a mine process.

The remainder of this thesis is organized into two major sections. The first section comprises data analysis. The second section investigates plausible reliability formulations to study the impact of equipment non-operation on the economic costs and production time. This section also discusses the results and application areas of the analyzed data and the reliability formulations.

2 Data Analysis For Mining Equipment

The first part of this thesis consists primarily of analyzing data relating to the various states of functioning of mining equipment used in hauling of ore. In particular, the equipment in consideration is a fleet of HEMM – Caterpillar Mining Truck-793®, which operate in an open-pit mine of the Highland Valley Copper Inc.©. For identification, the equipments have been assigned with unique numbers, 6161, 6162, 6163, and 6164. The preliminary data used in the presented analysis was initially gathered by Priyadarshini et al. (2005) at the Highland Copper Valley mine for determining the equipment failure and repair patterns.

In the following, the data has been analyzed to obtain probabilistic measures for the various states of functioning of mining equipment. The probabilistic measures include the number of occurrences of each state and the duration spent by equipment in each instance of state occurrence. Subsequently, the probabilistic information obtained from the analyzed data will serve as input to develop a cost and productivity-based reliability model. The reliability model would assist the decision making process in mining operations.

Section 2.1 describes the nature of preliminary data; section 2.2 describes the approach taken for classification of data and modeling of equipment states. Section 2.3 presents the key concepts used for probabilistic modeling of data along with the various practical challenges in modeling such a system. Section 2.4 includes a detailed description of Bayesian analysis of data for determining statistics for the rates of occurrence of different states of equipment functioning along with the initial results. Finally, section 2.5 explores the predictive analysis for interpreting analyzed data and applying it to draw useful inferences.

2.1. *The Nature of the Raw Data*

The data that are available for this study essentially consist of an annual dispatch report (time log) of the operational status of equipments recorded at the mining site. The data comprises status changes of equipment recorded continually in the form of date, time, duration, and the nature of occurrence. The status of equipment functioning is categorized on the basis of various types of events at the mine site such as, tie-down, standby, repair, non-operating event, etc. This categorization has been adopted during the initial data gathering performed at the Highland Valley Copper mine (Priyadarshini et al. 2005).

Figure 1 shows excerpts from a typical equipment dispatch report. The first row contains explanations of the different columns. The first five columns in the figure contain the equipment number, date, time, duration of an equipment status, and name of the equipment status. The remaining columns contain a code for the equipment status, category of equipment functioning corresponding to the status, reason for non-operation of the equipment, and comments for the equipment category. The equipment status changes several times within a short span of time, with each status lasting for certain duration. For example, operating status of the equipment is followed by non-operating instances resulting in a delay, stand-by, operating delay, or a repair status. Hereinafter, in this thesis, ‘equipment status’ is referred as ‘equipment state.’

Significant amount of data is available for analysis. There are approximately 65,000 instances of state changes for each equipment within a year. The data gathering methodology adopted by Priyadarshini et al. (2005) provides the flexibility of analyzing the occurrence and duration characteristics of different sub-states of the equipment. In order to model the stochastic characteristics of the various states of functioning of the equipment, the annual

equipment dispatch report generated at the Highland Copper Valley mine has been processed into a comprehensive database. This database is included in the CD-ROM that is attached to this thesis as Appendix B. The database includes occurrence and duration patterns of equipment states observed over one year. In particular, the database contains the number of observations of each sub-state, the time between successive sub-state occurrences, and the durations of sub-states for each of the four equipment. The database is obtained by processing the annual dispatch reports for the equipments.

Table 1 provides a complete classification of equipment states derived from the database. According to the classification, the state of equipment functioning is broadly categorized into operating and non-operating state. The non-operative state is sub-categorized into various mutually exclusive sub-states, with each sub-state getting triggered due to similar type of events. The advantage of such a discrete classification approach is the separation of sources of uncertainty associated with respective sub-states. It is observed in Table 1 that the sub-states are grouped under delay, operating delay, standby, scheduled repair and unscheduled repair (UR). For example, delay in operation of equipment (a non-operating state) resulting from a blasting event in the mine is classified as sub-state blast-delay under the category operating delay. The fourth column in Table 1 lists the occurrence patterns of different sub-states. Accordingly, the occurrence of a sub-state is either random or deterministic in nature. The occurrence pattern of an equipment sub-state is based on the nature of event that results in the equipment sub-state. For example, sudden equipment failure is a random event, resulting in the random non-operating UR sub-state. On the contrary, a daily lunch-break event results in a deterministic non-operating delay equipment sub-state.

In the following, the equipment repair data from the dispatch report is processed as an example for the overall data analysis approach. The equipment (HEMM) in consideration is a large and complex piece of machinery and includes a number of components, some of which are detailed in Table 2. One or more of these components malfunction or fail to cause entire equipment breakdown and necessitating the repair. Therefore, it is difficult to assess the exact time and cause of a component failure. In this study, the failure of one or more components that resulted in a repair action is considered as a failure event. The details of the time and durations of the repairs are obtained from the dispatch report, which includes two types of events:

- Repairs performed on the equipment according to a scheduled maintenance program;
- Repairs necessitated as a result of sudden failure of one or more component of the equipment.

The latter event is random in nature and has been assumed to be a failure event in the statistical studies. In other words, URs caused due to a sudden failure of one or more components is a failure event. It is noted that during a failure event, more than one component may have been repaired. Table 3 lists the number of repairs observed in a year for the four trucks. The first column of the table lists the equipment number, the total number of unscheduled repairs, and the total number of scheduled repairs observed over one year. It is observed that equipment 6163 and 6164 needed more repairs than 6161 and 6162 in the observation period. The likely reasons for this variation include equipment age, handling, amount of usage, the quality of manufacturing, etc. In the next few paragraphs, the occurrence and duration characteristics of the UR sub-state are discussed.

Table 4 lists the time (in unit of days) between successive scheduled repairs for the four equipments. This data is obtained from the annual dispatch reports for the equipments. It is clear from the table that successive scheduled repair events occur every 30-40 days according to a preventative maintenance program. Therefore, the occurrence of scheduled repair is deterministic in nature.

The time between UR sub-states is calculated by taking the difference between the start times of successive UR events. In order to determine the probabilistic characteristics of the occurrences of UR sub-states, the distribution of time between unscheduled repairs of the equipments is studied. The nature of this distribution will aid in determining an appropriate random occurrence model for equipment failures. For this, first the preliminary data is analyzed in order to obtain the distribution for time between UR sub-states. To provide an idea of the nature of the data, a few preliminary statistical measures for the time between UR sub-states are provided in Table 5. For equipment 6161, the mean of the time between UR sub-states is 6.3 days, with a standard deviation of 9.3 days, and a coefficient of variation of 148 %. Coefficient of variation for equipments 6162, 6163, and 6164 is close to 100%, indicating high uncertainty in sub-state occurrence data. Reducing the uncertainty in time between successive UR sub-states is one of the chief goals of the Bayesian methods described later on in this section. Notably, as more data relating to UR sub-state occurrence observations becomes available, the coefficient of variation reduces significantly, resulting in less uncertainty.

Figure 2 shows a comparison between the frequency histogram, exponential probability density function (PDF) and lognormal PDF of the time between successive UR sub-states for the four equipments. Specifically, the bars in Figure 2 show the normalized

frequency histogram. The dashed and dotted lines are the exponential and lognormal PDFs with the same measures. It is observed from the figure that time between UR sub-states is mostly 1-3 days, i.e., a significant number of successive UR sub-state observations occur within every 1-3 days. Furthermore, a comparison of frequency histogram with the exponential and lognormal distribution reveals that exponential distribution is a reasonable fit to the raw data.

Figure 3 shows the associated cumulative distribution functions (CDFs). The solid line is the cumulative frequency; the dashed line is an exponential CDF with the aforementioned mean and standard deviation. The dotted line is a lognormal CDF with the same measures. It is observed that the lognormal and exponential distributions are reasonable fits to the raw data. Notably, the slope of CDF points towards the uncertainty, or the spread of data. A CDF with a steep slope indicates that the data is not highly dispersed (less uncertainty), whereas, CDF with flatter slope points towards highly dispersed data (more uncertainty). From Figure 3 it is clear that the time between UR sub-state data is less dispersed for equipment 6161 than the other equipments. The Figure 2 also corroborates this observation.

The duration characteristics of UR sub-state are discussed next. The time duration of equipments in the UR sub-state is highly variable and uncertain. According to the database, during a year, the total time spent in the UR sub-state by equipments 6161, 6162, 6263, and 6614 is 220, 227.17, 380.5, and 234.5 hours respectively. The maximum time durations of the equipments in the UR sub-state are 42.3, 44.7, 58.9, and 107.2 hours respectively. Figures 4 and 5 show the spread of the UR sub-state duration data in the form of relative frequency

histogram and cumulative frequency. It is observed from Figure 5 that for almost 80% of the UR instances, 10 hours or less are spent by the equipments in the repair shop.

The duration data for UR sub-state is lognormally distributed. Accordingly, Figure 6 shows the lognormal distribution obtained for the data for the four equipments. These distribution have been obtained using the Crystal Ball[®] software. Notably, the coefficient of variation for the duration in UR sub-state is close to 200%.

2.2. *Modeling of Equipment States*

A fundamental task in the present study is to estimate the time spent by equipment in non-operating states during an observation period. Probabilistic estimates of the time in non-operation can be obtained from the occurrence and duration characteristics of equipment states. The occurrence and duration patterns of many equipment states are random in nature, due to the uncertainty associated with the event that causes the state. Conversely, there are deterministic non-operating states due to events such as lunch, coffee break, which occur according to a pre-defined schedule.

When an equipment transitions into a random state, the state lasts for a random duration. For example, the occurrence of the UR sub-state and corresponding repair duration are random. Figure 7 illustrates a time window in which a set of equipment state transitions occur within a week. It is observed that the equipment undergoes scheduled repair for preventative maintenance on day two, which lasts for ten hours and forty minutes, and then operates for one and a half hour. Thereafter, a number of state transitions occur within the next four-hour period, which is shown by zooming-in on the time axis. Within the zoomed-in time window, the equipment operates intermittently with interruptions occurring due to

transition into various non-operating states. The non-operating states occur due to events including lunch, blast delay, repair, crusher red lights on, short move, coffee break, and operator shift change respectively. Hereinafter we denote the time of occurrence as t_x , where x is the name of the state. Similarly, we denote the duration of the occurrence by d_x , where x is the name of the state.

In the formulation developed in this thesis, a key quantity is the total time spent by an equipment in a sub-state. For a given observation period, this duration is denoted D_x , where x is the name of the sub-state. In a given observation period, D_x for a random sub-state is uncertain. Therefore, in the proposed model, the number of occurrences, N_x , and corresponding duration, d_x , of the random sub-state in the observation period are treated as random variables. Given the realizations of these parameters for a random sub-state x , the total duration in this sub-state, D_x , is evaluated by the expression

$$D_x = \sum_{i=1}^{N_x} d_{x(i)} \quad (1)$$

where N_x is the random number of occurrences of sub-state x within the given observation time and $d_{x(i)}$ is the random duration of occurrence number i . For deterministic sub-states, the occurrence and duration parameters are pre-defined and the total time spent in the sub-state is obtained by multiplying the total number of occurrences with the duration.

To model the uncertainty in number of occurrences N of random sub-state x , a probabilistic occurrence model is used. The homogeneous Poisson process is an appealing choice due to its simplicity. However, the implicit assumptions of independence between occurrences and time-invariant statistical properties must be carefully examined. The Poisson process is characterized by one parameter; the mean occurrence rate, λ . With knowledge of this parameter, derived measures such as the probability distributions for number of sub-

states in a time interval (the Poisson distribution) and time between sub-states (the exponential distribution) become available. The probability of observing N occurrences in observation period T is given by the Poisson probability mass function (PMF)

$$p(N) = \frac{(\lambda T)^N e^{-\lambda T}}{N!} \quad (2)$$

The obvious approach to obtain a point estimate of λ is to count the number of sub-states within the time period and divide it by the total time. For example the mean occurrence rate of unscheduled repairs, λ_{UR} for equipments are essentially point estimates of the occurrence rate. Figure 8 shows the PMF according to Eq. (2) for the point estimation of λ_{UR} for the four equipment. The approach of point estimation of λ has several weaknesses, including its inability to obtain a measure of uncertainty in the estimate of λ . This is remedied in the Bayesian approach suggested by Benjamin (1968). The Bayesian analysis is carried out in the next section to obtain probability distributions for λ for a number of sub-states.

In this thesis, the probability distribution for the duration parameter d_x is not obtained by Bayesian updating. Instead, probability distributions for d_x are selected by means of a simple best-fit technique. For example, duration data for UR suggests that the lognormal distribution is appropriate for all equipments under consideration. Similarly, the durations of power outage is exponentially distributed. Likewise, appropriate standard-type probability distributions for the durations of other sub-states of equipment are chosen. For this purpose, a standard software is utilized to fit probability distributions to d_x for all the sub-states. Table 6 lists the best fitting distributions to d_x for the sub-states of equipment. The first row contains explanations of the different columns for the equipments. The subsequent rows contain the results from the fitting. For each equipment, the first, second and third columns respectively

provide the distribution type, the mean duration in unit of minutes, and standard deviation of the sub-states. For many of the sub-states, a particular distribution type fits to the duration data of all the equipment in consideration. The exponential distribution fits well to the duration data corresponding to sub-states with very large variation in the magnitude of duration. For example, the durations of typical power outages varies from 20 minutes to 2000 minutes. Examples of other such sub-states include slippery/poor visibility, maintenance foreman in vehicle, and no-dump available. Furthermore, it is observed that the duration data corresponding to sub-states; no-operator available and UR has very high standard deviation. This is expected because of the highly uncertain nature of such events in practice.

2.3. *Treatment of Uncertainty: A Bayesian Approach*

The objective of this section is to analyze the occurrence rate, λ , of the sub-states of mining equipment. In the Bayesian approach, λ is itself considered to be a random variable. Importantly, the probability distribution for this random variable is repeatedly updated with information about occurrences of sub-state. Such updating is performed on the database gathered from the mining operation studied in this project. Before presenting the results, the methodology is briefly reviewed. The basic version of Bayes' theorem for the two events A and B reads

$$P(A|B) = \frac{P(B|A)}{P(B)} P(A) \quad (3)$$

where $P(A|B)$ denotes the conditional probability that event A will occur *given* that event B has occurred. The significance of Eq. (3) is its ability to update the probability of A in light of the occurrence of event B . This is reflected in the Bayesian terminology, in which $P(A)$ is

denoted the *prior* probability, while $P(A|B)$ is the *posterior* probability, $P(B|A)$ is the likelihood function, and $P(B)$ is a normalizing constant.

Bayes' theorem applied to the PDF, $f(\lambda)$ of a random variable λ reads

$$f(\lambda|E) = \frac{P(E|\lambda)}{P(E)} f(\lambda) \quad (4)$$

where E is the event that influences the PDF of λ . It is more common to write Eq. (4) in the form

$$f''(\lambda) = c \cdot L(\lambda) \cdot f'(\lambda) \quad (5)$$

where $f''(\lambda)$ is the posterior PDF for the occurrence rate, c is a normalizing constant that ensures a valid posterior PDF, $L(\lambda)$ is the likelihood function that denotes the probability of making the observation, and $f'(\lambda)$ is the prior PDF. Clearly, the objective of Eq. (5) is to update the PDF of λ in light of an observation whose information is implicitly incorporated in the likelihood function. From the above it is also clear that the likelihood function represents the probability of observing the observation given λ . Specifically, the probability of observing N occurrences in a time interval T is, according to the Poisson process given by the Poisson probability mass function explained in Eq. (2).

Consequently, the likelihood of observing, say, 3 equipment sub-states within a period of 40 days, given λ , is proportional to $\lambda^3 e^{-40\lambda}$ (it is not necessary to retain the entire expression in Eq. (1), but only the terms that contain λ , because the normalizing constant c will ultimately ensure a valid PDF).

The choice of prior PDF is part of the art of Bayesian updating. Typically, subjective judgment enters into this choice. However, it is emphasized that the prior PDF has little

impact on the final posterior PDF when reasonable amount of data is available (Box and Tiao, 1992), which is the case in this study. A convenient choice for the distribution type of the prior PDF is a *conjugate prior* (Benjamin and Cornell, 1970). This implies that the posterior PDF will be of the same type as the prior PDF. For this reason, the Gamma distribution is selected as the prior PDF for the occurrence rate, namely

$$f'(\lambda_x) = \frac{\nu(\nu \cdot \lambda)^{k-1} e^{-\nu \lambda}}{\Gamma(k)} \quad (6)$$

where $\Gamma(\cdot)$ is the Gamma function, and ν and k are distribution parameters. That the posterior distribution $f''(\lambda)$ will also have the Gamma distribution type is verified by recognizing that Eq. (2) is the likelihood function in this case. Considering only the terms of Eq. (2) that depend on λ and then substituting these terms and Eq. (6) into Eq. (5) one obtains

$$f''(\lambda) = c \cdot \lambda^N e^{-\lambda T} \cdot \frac{\nu(\nu \cdot \lambda)^{k-1} e^{-\nu \lambda}}{\Gamma(k)} = \dots = \frac{\nu(\nu \cdot \lambda)^{k+N-1} e^{-(\nu+T)\lambda}}{\Gamma(k+N)} \quad (7)$$

The terms non-dependent of λ in Eq. (2) form a constant quantity, which is taken care of by the normalizing constant c . It is observed that $f''(\lambda)$ is similar to the Gamma distribution in Eq. (6), with the only difference that k is replaced by $k+N$, and that ν is replaced by $\nu+T$. This implies that the Gamma distribution indeed is a conjugate prior. This has the appealing consequence that the PDF for λ is readily updated given a new occurrence by the algebraic updating rule

$$\begin{aligned} k'' &= k' + N \\ \nu'' &= \nu' + T \end{aligned} \quad (8)$$

where double-prime signifies the posterior distribution parameters in $f''(\lambda)$, while single-prime signifies the prior parameters.

In summary, the prior distribution $f'(\lambda)$ indicates our state of knowledge about λ before a particular sub-state observation is made for an equipment. The Bayesian updating rule in Eq. (5) combines the prior information contained in $f'(\lambda)$ with new observations contained in $L(\lambda)$. The posterior distribution $f''(\lambda)$ represents the updated state of information about the mean sub-state occurrence rate λ . In passing, it is also noted that the mean, standard deviation, and coefficient of variation are available from the Gamma distribution for λ : Mean $\mu_\lambda = k/v$, standard deviation $\sigma_\lambda = \sqrt{k}/v$ and coefficient of variation $\delta_\lambda = 1/k$.

Bayesian analysis as described above is carried out for all random sub-states. In the following, it is demonstrated for the occurrence of unscheduled repairs (sudden breakdown of the equipment). The time between UR events is recorded in units of days. A significant number of UR occurrences were recorded; typically in the order of 60 sub-states over a time period of 12 months. The rate of unscheduled repairs is denoted λ_{UR} . Based on judgment, the prior distribution for λ_{UR} was chosen to have a mean equal to 0.167 occurrences per day with a large degree of uncertainty; a 100% coefficient of variation was selected. This corresponds to $v'_{UR} = 6$ and $k'_{UR} = 1$. It is reemphasized that this initial assumption usually has little impact on the final distribution $f'(\lambda_{UR})$ for the occurrence rate.

As an example of the application of the methodology outlined in the previous section, Tables 7, 8, 9, and 10 are provided. The tables present the Bayesian procedure for updating the occurrence rates of UR sub-states for equipments 6161, 6162, 6163, and 6164. The first row in each table contains explanations of the different columns and the second row contains the initial assumption stated in the previous paragraph. The subsequent rows contain the results from the Bayesian updating according to Eq. (8). Notably, as more and more information on UR sub-states are incorporated, the coefficient of variation of the λ_{UR}

distribution is reduced. As expected, this indicates that the incorporation of information reduces the uncertainty in λ_{UR} . The tables only shows very few of the UR occurrence observations that are included in the database, for illustration purposes. The last row of each table contains the graphs of the prior and posterior Gamma probability distributions $f''(\lambda_x)$ which are plotted over a range of λ_{UR} values. The plot contains six PDFs; the five first PDFs obtained during the Bayesian updating and the last, which is obtained after all the UR data, is incorporated. It is observed that the initial PDFs have substantial dispersion, while the last PDF is significantly narrower, thus indicating less statistical uncertainty. It is also observed that the final PDF has similar characteristics for each of the considered equipment.

Table 11 provides results of Bayesian analysis for all random non-operating sub-states of equipments 6161, 6162, 6163, and 6164. In Table 11, the first row contains explanations of the different columns for the equipments. The subsequent rows contain the results from the Bayesian updating. For each equipment, the first, second and third columns provide the number of observations of different sub-states within a year, the Bayesian updated mean occurrence rate μ_λ in units of days, and coefficient of variation of λ of the sub-states. Significantly lower value of the coefficient of variation of λ is observed for the sub-state with high number of observations within a time period. Contrarily, a high value of coefficient of variation is associated with sub-states such as power outage, no-dump available, which have very few data sets (observations).

2.4. *The Predictive Analysis*

In the previous section the statistical uncertainties associated with the estimation of λ have been dealt using the Bayesian approach. Accordingly, the posterior distribution $f''(\lambda)$

provides us now with an updated and more informed estimate of λ . The posterior distribution of occurrence rates of sub-states, $f''(\lambda)$ can be used to estimate the probabilities of number of future sub-state occurrences N . The predictive distribution of N in time T is obtained from the PMF (Benjamin, 1970)

$$\tilde{p}(N) = \int_0^{\infty} \frac{(\lambda \cdot T)^N e^{-(\lambda \cdot T)}}{N!} f''(\lambda) d\lambda \quad (9)$$

After substituting for $f''(\lambda)$ in the above equation, the solution of the above integral is given by the following expression

$$\tilde{p}(N) = \left(\frac{T}{\nu'' + T} \right)^N \cdot \left(\frac{\nu''}{\nu'' + T} \right)^{k''} \cdot \frac{\Gamma(k'' + N)}{\Gamma(k'') \cdot N!} \quad (10)$$

The PMF of N based on the point estimation of λ , which is considered equal to the posterior mean $\bar{\lambda}$, is given by

$$p(N | \bar{\lambda}) = \frac{(\bar{\lambda} \cdot T)^N e^{-(\bar{\lambda} \cdot T)}}{N!} \quad (11)$$

Figure 9 shows a comparison of PMFs for the number of UR sub-state occurrences of equipment 6161, obtained using Eq. (10) and (11). In Figure 9, the solid bar represents the predictive distribution of N and the transparent bars represent the PMF obtained from point estimation of λ . It is observed, that although the means for the two cases, are nearly identical, the standard deviation based on predictive distribution is much larger. This indicates that the statistical uncertainty in the parameter λ is incorporated in predictive distribution, making the distribution more informative.

The data analysis can be applied to obtain various probabilistic measures relating to the equipment sub-states. For example, consider the UR sub-state of the equipment as equipment failure. Then, for a failure occurrence rate of λ , the probability of equipment failure within a time period T is given by the exponential CDF of its time-between-failures, Tbf , as follows

$$P(Tbf \leq T | \lambda) = 1 - e^{-\lambda \cdot T} \quad (12)$$

Eq. (12) represents a conditional distribution that depends upon λ . Using the total probability theorem, we combine Eqs. (7) and (12) to obtain the predictive probability distribution of equipment failure, P_f by the expression

$$P_f = \int_{-\infty}^{\infty} P(Tbf \leq T | \lambda) f''(\lambda) d\lambda = \int_{-\infty}^{\infty} (1 - e^{-\lambda \cdot T}) \frac{\nu'' (\nu'' \cdot \lambda)^{k''-1} e^{-\lambda \cdot \nu''}}{\Gamma(k'')} d\lambda \quad (13)$$

For equipment 6161 with parameter values $\nu'' = 358.385$ days and $k'' = 57$, the predictive CDF plot of equipment failure is shown in

Figure 10. According to the plot, the probability that the equipment will fail at least once in two days or less is approximately 0.27, and the probability that the equipment will fail at least once in four days or less is approximately 0.47.

Alternatively, the probability of failure of an equipment within a day, ($T = 1$), can also be calculated using Eq. (13). The probability of failure in a day for the four equipments is 0.1482, 0.1479, 0.1763, and 0.1836 respectively. Equipment 6164 has the highest probability of failure in a day followed by 6163, 6161, and 6162 respectively.

The reliability of an equipment is its probability of non-failure. Accordingly, the reliability function, $R(T)$, is given by the following expression

$$R(T) = 1 - P(Tbf \leq T | \lambda) = 1 - [1 - e^{-\lambda T}] \quad (14)$$

For the observed range of values for time-between-failure of equipment 6161, which is obtained from the database, and for the occurrence rate of 0.4, the reliability curve is plotted in Figure 11. According to the plot, the probability that the equipment will function without failure for 5 days is approximately 0.2. Similar type of reliability plots are used in mining engineering in order to draw inferences about the performance of equipments. For example, Hall (2000) presents a plot of reliability function corresponding to the failure distribution, which is fitted using the software Weibull++TM, for a data set obtained from a fleet of drills. Reliability of the drill (component) is the probability of the drill successfully operating to a given footage.

Table 1: Classification of equipment states

S.no	State	Sub-state	Occurrence	Category
1	Operating state	Equipment in Full Operation	Random	<i>Operating</i>
2	Non-Operating states	Delay-First Coffee Break	Deterministic	<i>Delay</i>
3		Delay-Second Coffee Break	Deterministic	
4		Delay-First Lunch Break	Deterministic	
5		Delay-Second Lunch Break	Deterministic	
6		Delay-Power Outage	Random	
7		Delay-Slippery/ Poor Visibility	Random	
8		Operating Delay-Blast Delays	Random	<i>Operating Delay</i>
9		Operating Delay –Clean-up	Random	
10		Operating Delay -Cleaning Cab	Random	
11		Operating Delay –Crusher Down	Random	
12		Operating Delay –Crusher Red Light-on	Random	
13		Operating Delay –Crusher-trucks Queued	Random	
14		Operations Delay	Random	
15		Operating Delay –Service	Random	
16		Operating Delay -Short move	Random	
17		Operating Delay -Shovel Down	Random	
18		Operating Delay -Washroom Usage	Random	
19		Standby- Maintenance Foreman	Random	<i>Standby</i>
20		Standby- No Dump Available	Random	
21		Standby- No Operator Available	Random	
22		Standby- No Shovel Available	Random	
23		Standby-Operator on Board	Random	
24		Standby-Miscellaneous	Deterministic	
25		Scheduled Repair	Deterministic	
26		Unscheduled Repair (UR)	Random	<i>Repair</i>

Table 2: Type of components for repairs.

S.no.	Components	S.no.	Components
1	24/12 volt system	15	General Mechanical
2	Air System	16	Hoist System
3	Alternator	17	Hydraulic System
4	Axle housing repairs	18	Lube System
5	Brakes	19	Main frame Carby
6	Bucket	20	Propel Transmission
7	Cab or house	21	Steering System
8	Cooling System	22	Steering/Brakes
9	Crowd Transmission	23	Suspension
10	Differential	24	Tires
11	Dump body	25	Torque Converter and Pumps
12	Engine	26	Transmission and Pumps
13	Frame cracks	27	Trip Machinery
14	General Electrical		

Table 3: Number of repairs observed in a year for the four equipments

Equipment number	Total number of unscheduled repairs in one year	Total number of scheduled repairs in one year
6161	56	8
6162	54	9
6163	69	10
6164	70	9

Table 4: Time between successive scheduled repair events (in days)

Time between successive scheduled repair events (in days)			
Equipment 6161	Equipment 6162	Equipment 6163	Equipment 6164
34.0	36.01	33.95	35.00
37.0	35.03	30.00	35.00
32.0	36.96	37.00	32.00
35.0	34.00	31.00	34.00
36.0	35.00	34.00	36.00
36.0	33.99	38.00	30.00
32.0	37.00	36.00	36.00
37.0	33.00	33.00	34.00
-	34.00	33.00	33.00
-	-	46.25	-

Table 5: Statistics for time between successive UR sub-states

Equipment number	Mean (days)	Standard Deviation (days)	Coefficient of Variation
6161	6.29	9.34	148%
6162	6.30	5.64	89%
6163	5.11	5.32	104%
6164	4.96	4.54	92%

Table 6: Probability distributions for duration in sub-state for various equipments.

S.no.	Random Non-Operating State	Equipment 6161			Equipment 6162			Equipment 6163			Equipment 6164		
		Distribution Type	Mean Duration (minutes)	Standard Deviation (minutes)	Distribution Type	Mean Duration (minutes)	Standard Deviation (minutes)	Distribution Type	Mean Duration (minutes)	Standard Deviation (minutes)	Distribution Type	Mean Duration (minutes)	Standard Deviation (minutes)
1	Power Outage	Exponential	35.07	-	Exponential	38.9	-	Exponential	20.4	-	Exponential	29.7	-
2	Slippery/ poor visibility	Exponential	44.04	-	Exponential	20.4	-	Exponential	98.7	-	Exponential	82.2	-
3	Blast Delays	Lognormal	9.80	9.07	Lognormal	7.7	4.3	Gamma	10.8	7.7	Exponential	11.7	-
4	Clean-up	Lognormal	6.44	8.24	Gamma	13.4	4.4	Lognormal	6.9	8.17	Lognormal	7.7	8.1
5	Cleaning Cab	Lognormal	6.03	4.21	Lognormal	17.0	19.0	Gamma	7.1	4.5	Lognormal	13.7	13.3
6	Crusher Down	Lognormal	25.76	26.08	Lognormal	9.97	10.95	Lognormal	15.4	19.5	Lognormal	14.2	18.28
7	Crusher Red-light On	Gamma	3.16	2.29	Gamma	3.4	2.6	Gamma	3.9	3.0	Gamma	3.6	3.1
8	Crusher-trucks queued	Gamma	6.34	5.51	Lognormal	5.7	5.8	Lognormal	6.9	6.5	Lognormal	5.8	5.4
9	Operations Delay	Exponential	9.93	-	Gamma	9.1	11.7	Exponential	12.9	-	Exponential	11.0	-
10	Service	Gamma	8.29	4.10	Gamma	8.1	3.8	Gamma	8.6	4.2	Gamma	9.1	4.6
11	Short move	Lognormal	6.92	5.08	Gamma	7.5	5.1	Gamma	7.3	6.6	Gamma	6.7	5.3
12	Shovel Down	Exponential	13.34	-	Gamma	10.6	9.3	Gamma	20.8	18.8	Exponential	14.0	-
13	Washroom	Gamma	11.61	4.39	Gamma	10.8	4.6	Gamma	11.8	4.0	Lognormal	8.7	6.8
14	Maintenance foreman	Exponential	40.47	-	Exponential	176.2	-	Exponential	96.1	-	Exponential	109.5	-
15	No Dump available	Exponential	1.57	-	Exponential	1.5	-	Exponential	0.4	-	-	-	-
16	No Operator Available	Lognormal	121.8	251.6	Lognormal	91.24	195.2	Lognormal	95.1	250.73	Lognormal	106.75	197.63
17	No shovel available	Lognormal	8.08	11.18	Lognormal	3.07	8.12	Lognormal	5.38	12.28	Lognormal	5.86	8.18
18	Operator on Board	Lognormal	13.6	21.24	Lognormal	15.4	21.36	Lognormal	31.13	58.40	Lognormal	19.61	39.1
19	Unscheduled Repair	Lognormal	230.6	468	Lognormal	247.8	534.5	Lognormal	321.54	618.8	Lognormal	198.17	771.8

Table 7: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6161.

Number of occurrences, N_{UR}	Time between subsequent occurrences, t	$k_{UR}'' = k_{UR}' + N_{UR}$	$\nu_{UR}'' = \nu_{UR}' + t$	Mean of λ_{UR}'' $\mu_{\lambda_{UR}}''$	Coefficient of variation, $\delta_{\lambda_{UR}}''$
0	0.000	6.000	1	0.1667	1.0000
1	2.225	8.225	2	0.2432	0.7071
2	12.534	20.760	3	0.1445	0.5774
3	0.216	20.975	4	0.1907	0.5000
4	0.842	21.817	5	0.2292	0.4472
6	0.496	33.503	7	0.2089	0.3780
7	2.057	35.560	8	0.2250	0.3536
8	2.638	38.198	9	0.2356	0.3333
9	0.776	38.974	10	0.2566	0.3162
10	8.167	47.141	11	0.2333	0.3015
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
53	0.918	350.601	54	0.1540	0.1361
54	3.818	354.419	55	0.1552	0.1348
55	1.130	355.549	56	0.1575	0.1336
56	2.836	358.385	57	0.1590	0.1325

Bayesian Updating: Gamma PDF plots for the occurrence rates of "Unscheduled Repair" sub-state for equipment 6161

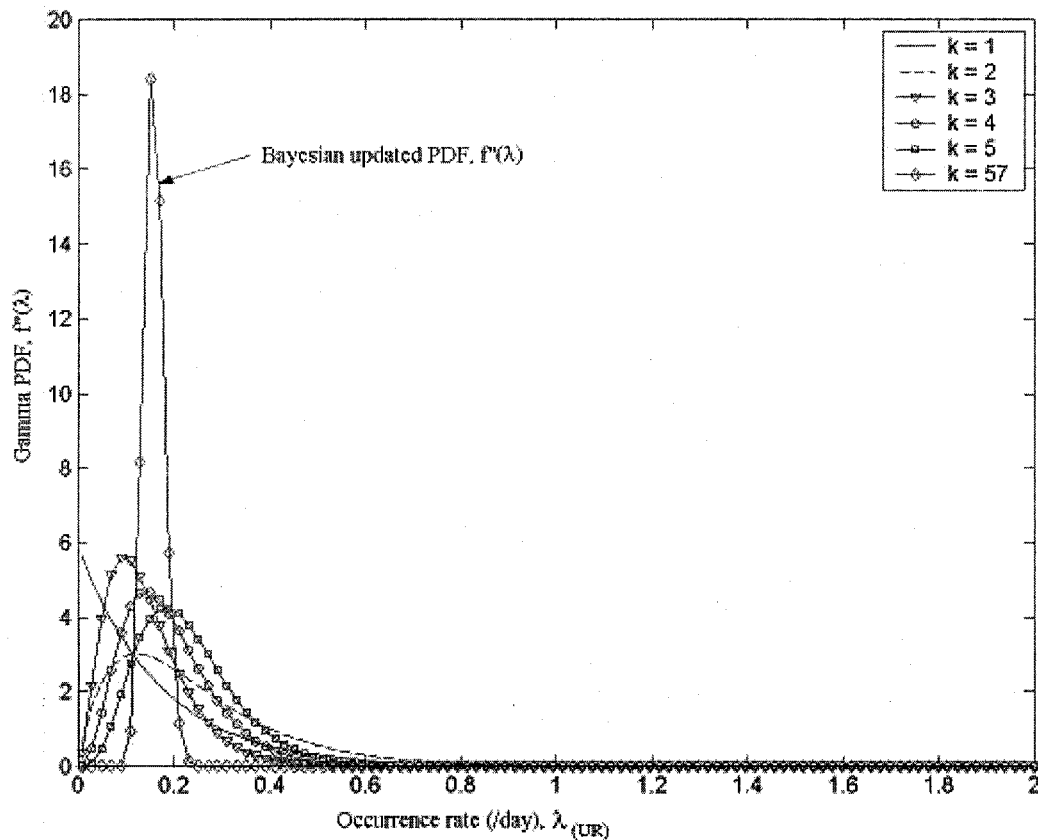


Table 8: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6162.

Number of occurrences, N_{UR}	Time between subsequent occurrences, t	$k_{UR}'' = k_{UR}' + N_{UR}$	$\nu_{UR}'' = \nu_{UR}' + t$	Mean of λ_{UR}'' $\mu_{\lambda_{UR}}''$	Coefficient of variation, $\delta_{\lambda_{UR}}''$
0	0.000	7.000	1	0.1429	1.0000
1	0.168	7.168	2	0.2790	0.7071
2	1.104	8.272	3	0.3627	0.5774
3	20.524	28.796	4	0.1389	0.5000
4	15.062	43.858	5	0.1140	0.4472
5	12.423	56.281	6	0.1066	0.4082
6	1.269	57.550	7	0.1216	0.3780
7	3.222	60.772	8	0.1316	0.3536
8	2.212	62.984	9	0.1429	0.3333
9	0.693	63.677	10	0.1570	0.3162
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
52	3.321	339.801	53	0.1530	0.1387
53	4.494	344.294	54	0.1539	0.1374
54	1.106	345.400	55	0.1563	0.1361
55	2.080	347.481	56	0.1583	0.1348

Bayesian Updating: Gamma PDF plots for the occurrence rates of "Unscheduled Repair" sub-state for equipment 6162

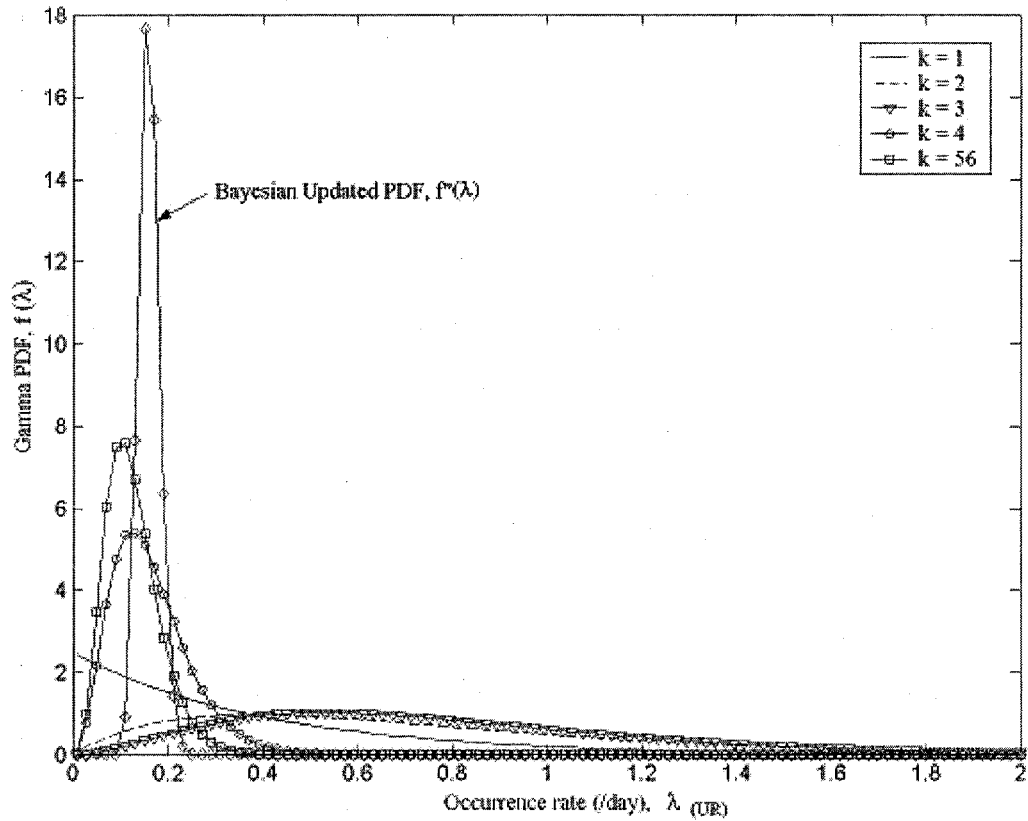


Table 9: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment no. 6163.

Number of occurrences, N_{UR}	Time between subsequent occurrences, t	$k_{UR}'' = k_{UR}' + N_{UR}$	$\nu_{UR}'' = \nu_{UR}' + t$	Mean of λ_{UR}'' $\mu_{\lambda_{UR}''}$	Coefficient of variation, $\delta_{\lambda_{UR}''}$
0	0.000	7.000	1	0.1428	1
1	3.258	10.258	2	0.1949	0.7071
2	1.544	11.801	3	0.2542	0.5773
3	0.263	12.064	4	0.3315	0.5
4	1.681	13.745	5	0.3637	0.4472
5	7.285	21.030	6	0.2853	0.4082
6	7.715	28.745	7	0.2435	0.3779
7	1.325	30.071	8	0.2660	0.3535
8	7.921	37.991	9	0.2368	0.3333
9	1.263	39.254	10	0.2547	0.3162
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
67	3.859	362.503	68	0.1875	0.1212
68	0.214	362.717	69	0.1902	0.1203
69	0.310	363.027	70	0.1928	0.1195
70	1.778	364.805	71	0.1946	0.1186

Bayesian Updating: Gamma PDF plots for the occurrence rates of "Unscheduled Repair" sub-state for equipment 6163

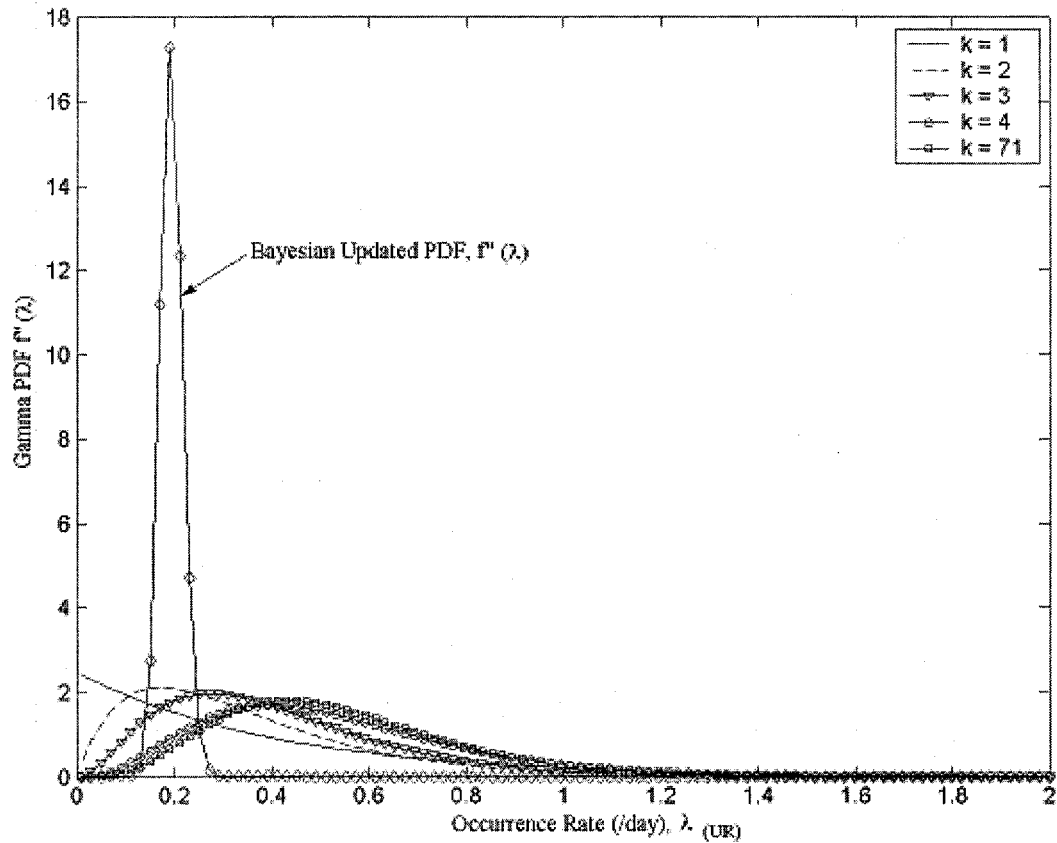


Table 10: Bayesian procedure to obtain updated occurrence rate (λ_{UR}) for equipment 6164.

Number of occurrences, N_{UR}	Time between subsequent occurrences, t	$k_{UR}'' = k_{UR}' + N_{UR}$	$\nu_{UR}'' = \nu_{UR}' + t$	Mean of λ_{UR}'' $\mu_{\lambda_{UR}}''$	Coefficient of variation, $\delta_{\lambda_{UR}}''$
0	0.000	7.000	1	0.1429	1.0000
1	5.251	12.251	2	0.1633	0.7071
2	4.895	17.146	3	0.1750	0.5774
3	11.211	28.357	4	0.1411	0.5000
4	0.385	28.742	5	0.1740	0.4472
5	4.004	32.746	6	0.1832	0.4082
6	9.117	41.864	7	0.1672	0.3780
7	12.022	53.886	8	0.1485	0.3536
8	0.130	54.016	9	0.1666	0.3333
9	1.735	55.750	10	0.1794	0.3162
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
68	0.141	353.502	69	0.1952	0.1204
69	0.047	353.549	70	0.1980	0.1195
70	0.505	354.054	71	0.2005	0.1187
71	0.505	354.559	72	0.2031	0.1179

Bayesian Updating: Gamma PDF plots for the occurrence rates of "Unscheduled Repair" sub-state for equipment 6164

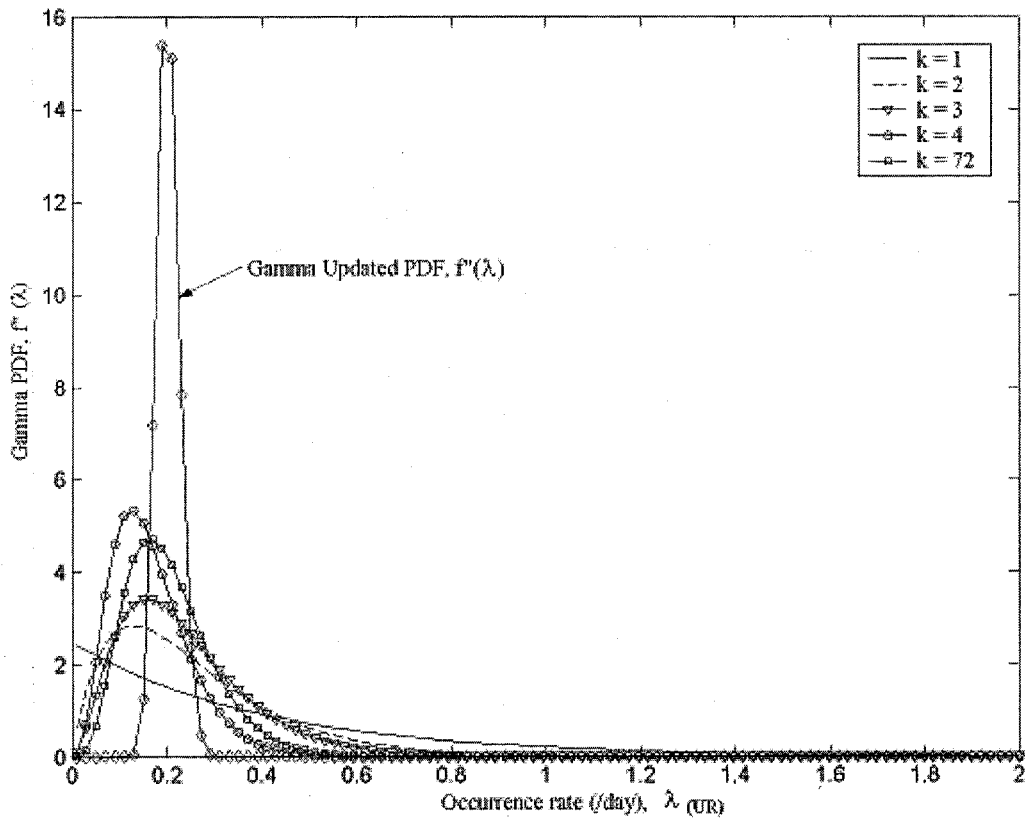


Table 11: Bayesian updated mean occurrence rate for the random non-operating sub-states of equipments.

S.no.	Random Non-Operating Sub-states	Equipment 6161			Equipment 6162			Equipment 6163			Equipment 6164		
		Number of observations in one year	Bayesian Updated Mean Occurrence Rate per	COV, % (δ_i)	Number of observations in one year	Bayesian Updated Mean Occurrence Rate per	COV, % (δ_i)	Number of observations in one year	Bayesian Updated Mean Occurrence Rate per	COV, % (δ_i)	Number of observations in one year	Bayesian Updated Mean Occurrence Rate per	COV, % (δ_i)
1	Power Outage	6	0.020	37.8	9	0.024	31.6	12	0.031	27.7	8	0.022	33.3
2	Slippery, p. visibility	9	0.038	31.6	6	0.033	37.8	15	0.060	25.0	17	0.056	23.6
3	Blast Delays	72	0.213	11.7	69	0.192	12.0	29	0.086	18.3	42	0.116	15.2
4	Clean-up	35	0.095	16.7	35	0.108	16.7	15	0.050	25.0	22	0.073	20.9
5	Cleaning Cab	159	0.437	7.9	316	0.867	5.6	140	0.385	8.4	144	0.405	8.3
6	Crusher Down	21	0.065	21.3	30	0.098	18.0	25	0.074	7.6	19	0.064	22.4
7	Crusher Red-light On	5218	14.337	1.4	4961	13.630	1.4	2555	7.006	2.0	3043	8.345	1.8
8	Cr. trucks queued	243	0.669	6.4	141	0.420	8.4	37	0.110	16.2	82	0.275	11.0
9	Operations Delay	145	0.403	8.3	204	0.568	7.0	172	0.484	7.6	176	0.505	7.5
10	Service	316	0.876	5.6	303	0.832	5.7	369	1.015	5.2	304	0.837	5.7
11	Short move	208	0.570	6.9	172	0.471	7.6	129	0.362	8.8	233	0.655	6.5
12	Shovel Down	76	0.252	11.4	81	0.256	11.0	93	0.281	10.3	66	0.190	12.2
13	Washroom	167	0.466	7.7	209	0.579	6.9	80	0.229	11.1	70	0.200	11.9
14	Maintenance foreman	7	0.022	35.4	6	0.026	37.8	13	0.052	26.7	10	0.030	30.2
15	No Dump available	8	0.039	33.3	12	0.133	27.7	4	0.024	44.7	-	-	-
16	No operator available	277	0.761	6.0	288	0.797	5.9	279	0.770	6.0	313	0.863	5.6
17	No shovel available	43	0.119	15.1	41	0.116	15.4	53	0.154	13.6	36	0.122	16.4
18	Operator on Board	16	0.047	24.3	14	0.043	25.8	25	0.075	19.6	23	0.069	20.4
19	Unscheduled Repair	56	0.159	13.2	54	0.158	13.5	70	0.195	11.9	70	0.203	11.8

Figure 1: Excerpts from a typical equipment dispatch report

16-FEB-06 -- Highland Valley Copper DISPATCH Report System -- 14:14:02									
Truck status Summary by Equipment									
1-JUL-04 Night Shift to 30-JUN-05 Day Shift									
Equipment	Date	Time	Duration	Status	Code	Category	Reason	Comments	

Status Changes for Trucks									
6124	22-SEP-04	08:00:00	11:11:24	Tiedown					
		19:14:28	144:19:08	Standby	1	Standby	NO OPERATOR		
Subtotal			155:30:32						
6125	16-SEP-04	08:00:00	1:50:08	Tiedown					
		09:50:08	0:05:50	Standby	1	Standby	NO OPERATOR		
		09:55:58	12:33:05	Standby					
		22:29:03	03:37:22	Standby	1	Standby	NO OPERATOR		
	20-SEP-04	20:06:25	0:10:49	Ready		Operating			
		20:17:14	0:14:24	Delay	11	Operating	De SHOVEL DOWN		
		20:31:38	1:28:08	Ready		Operating			
		21:59:46	0:11:03	Delay	200	Delay	COFFEE		FIRST COFFEE
		22:10:49	0:16:51	Ready		Operating			
		22:27:40	7:48:14	Standby	1	Standby	NO OPERATOR		
	21-SEP-04	06:15:54	1:44:06	Down	101	Repair	ENGINE		
		08:00:00	696:00:00	Standby					
	20-OCT-04	08:00:00	9:25:06	Tiedown		Standby			
		17:23:06	15:37:53	Standby	1	Standby	NO OPERATOR		
	21 OCT 04	09:02:59	1:57:00	Down	128	Repair	TRANSMISSION AND PUMPS		TRANS HOSE
		11:00:00	2:06:10	Standby	1	Standby	NO OPERATOR		
		13:06:10	0:11:52	Standby	1	Standby	NO OPERATOR		
		13:18:02	1:42:52	Ready		Operating			
		15:00:54	0:11:16	Delay	200	Delay	COFFEE		SECOND COFFE
		15:12:10	1:02:13	Ready		Operating			
		16:14:23	0:24:44	Delay	110	Operating	De OPERATIONS DELAY		
		16:39:07	0:22:44	Ready		Operating			
		17:01:51	0:30:30	Delay	201	Delay	LUNCH		SECOND LUNCH
		17:32:21	2:01:59	Ready		Operating			
		19:34:20	0:06:00	Delay	900	Standby	SHIFTCHANGE		TIEDOWN
		19:40:20	61:04:57	Standby	1	Standby	NO OPERATOR		
	24-OCT-04	08:45:17	23:14:43	Standby	100	Standby	MAINTENANCE FOREMAN		BRK PROBLEM
	25-OCT-04	08:00:00	1:00:00	Down	103	Repair	HOIST/HYD SYSTEM		PM REPAIRS
		09:00:00	239:00:00	Down	105	Repair	BRKES		
	04-NOV-04	08:00:00	8:58:00	Tiedown		Standby			
		16:58:00	15:00:37	Standby	1	Standby	NO OPERATOR		
	05-NOV-04	07:58:37	2:01:29	Ready		Operating			
		10:00:06	0:10:07	Delay	200	Delay	COFFEE		FIRST COFFEE
		10:10:13	2:21:34	Ready		Operating			
		12:31:47	0:30:04	Delay	201	Delay	LUNCH		FIRST LUNCH
		13:01:51	1:57:43	Ready		Operating			
		14:59:34	0:10:04	Delay	200	Delay	COFFEE		SECOND COFFEE
		15:09:38	1:51:28	Ready		Operating			
		17:01:06	0:30:36	Delay	201	Delay	LUNCH		SECOND LUNCH
		17:31:42	2:00:01	Ready		Operating			
		19:31:43	0:28:17	Delay	900	Standby	SHIFTCHANGE		TIEDOWN

Figure 2: Relative frequency histogram, exponential and lognormal PDF for the time between successive UR sub-states

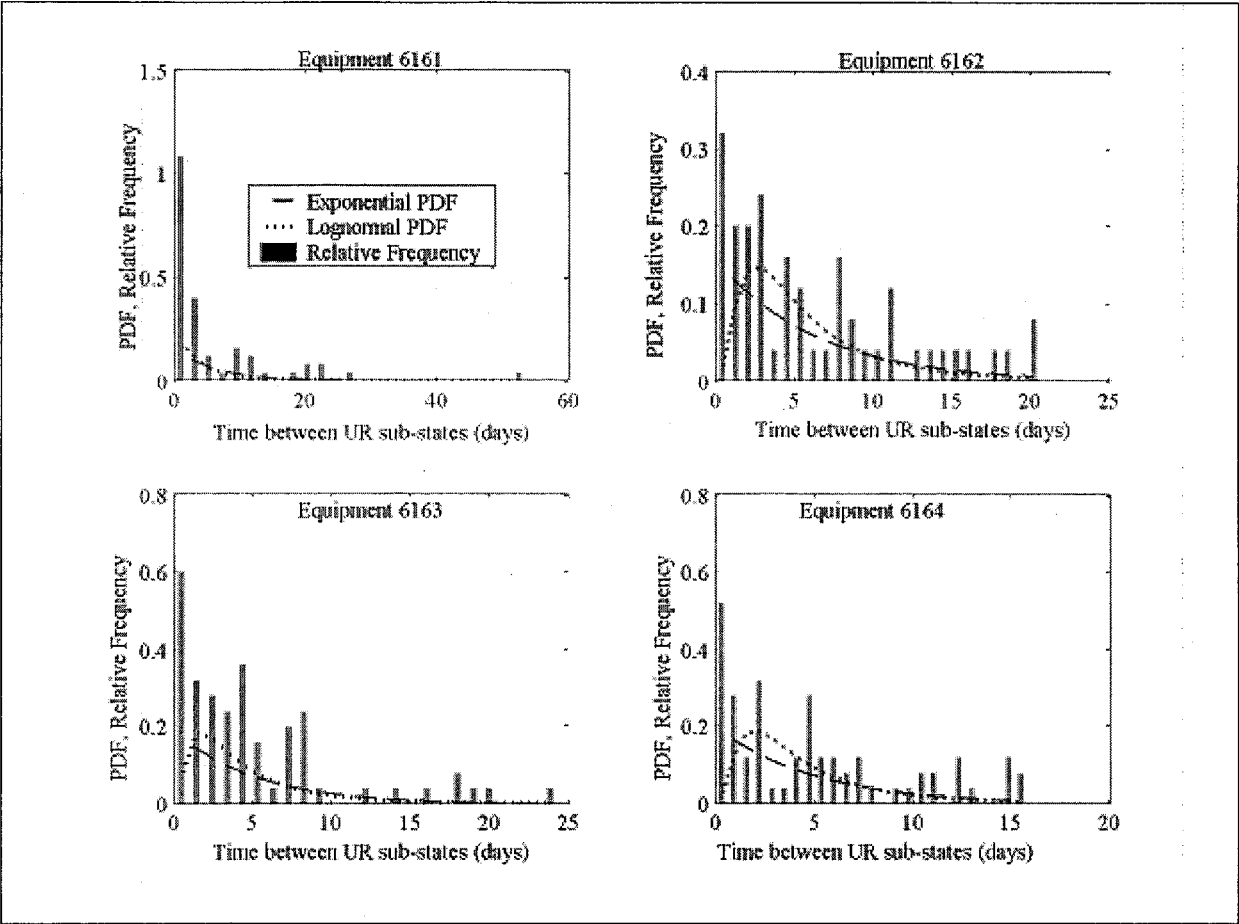


Figure 3: Cumulative frequency histogram, Exponential and Lognormal CDF for the time between successive UR sub-states

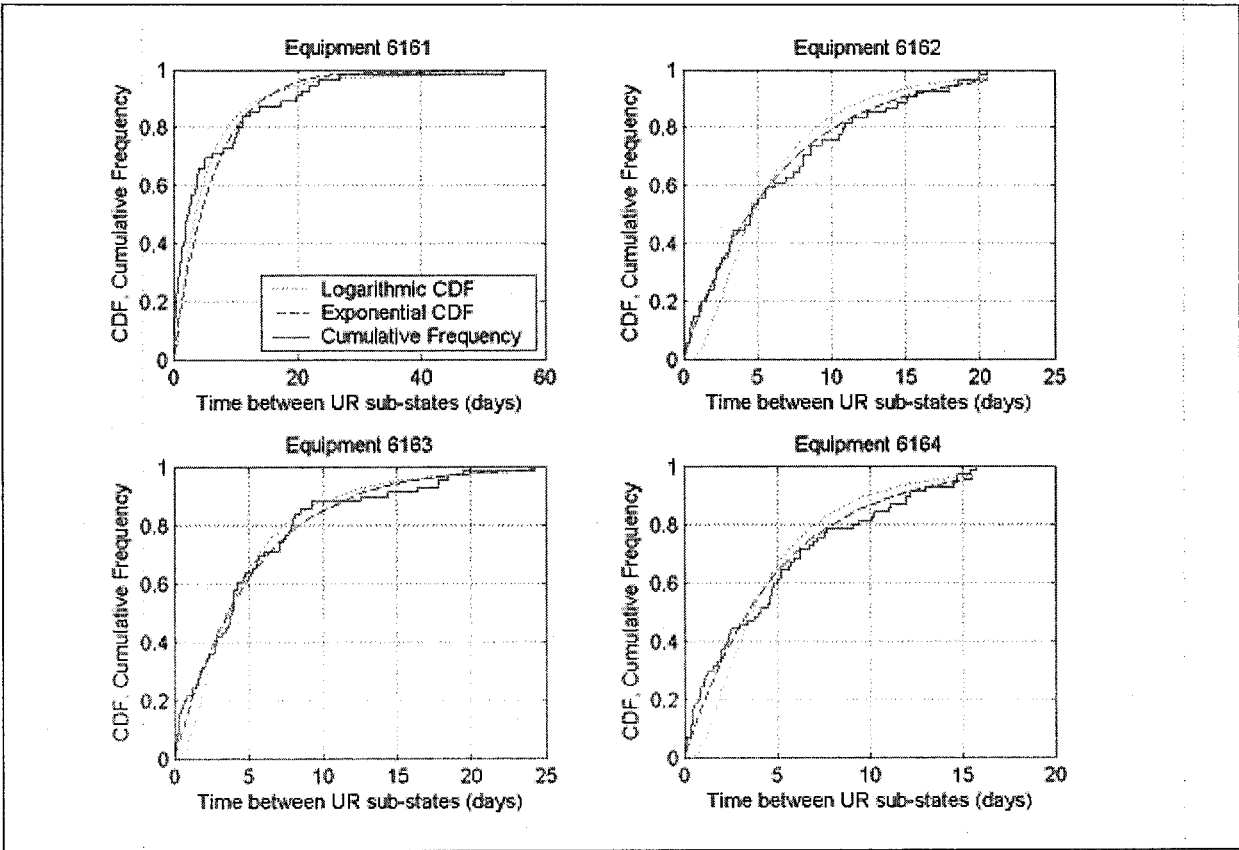


Figure 4: Relative frequency histograms for UR sub-state duration

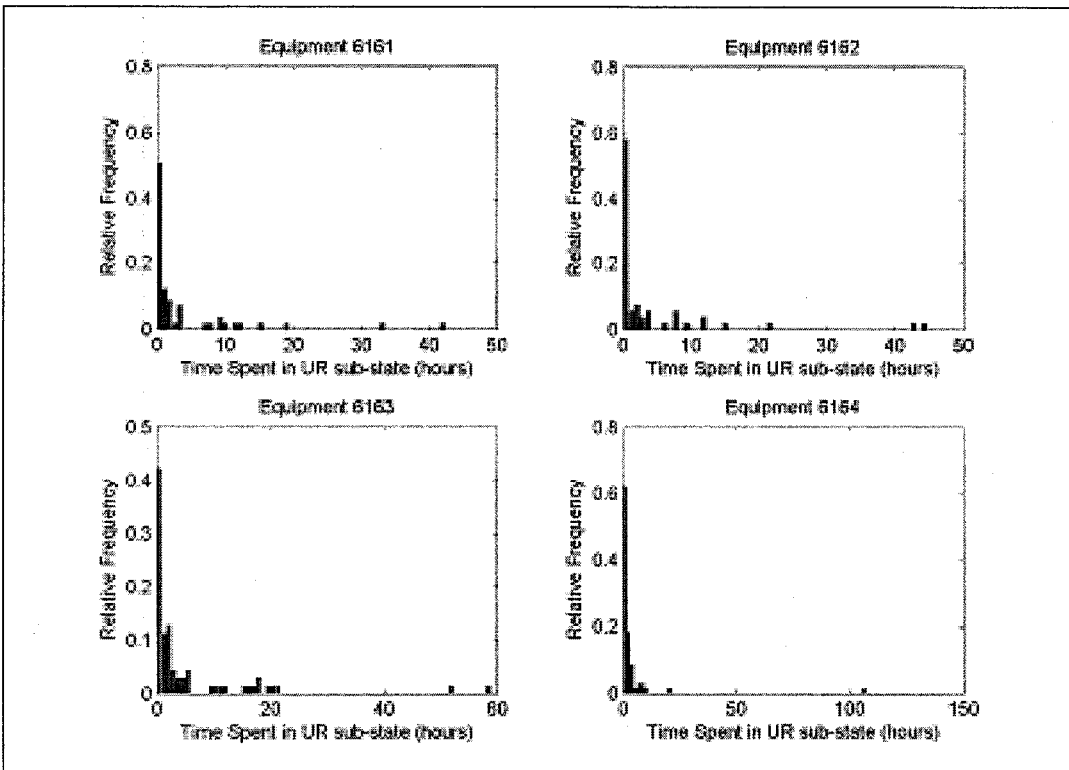


Figure 5: Cumulative frequency plot for the UR sub-state duration

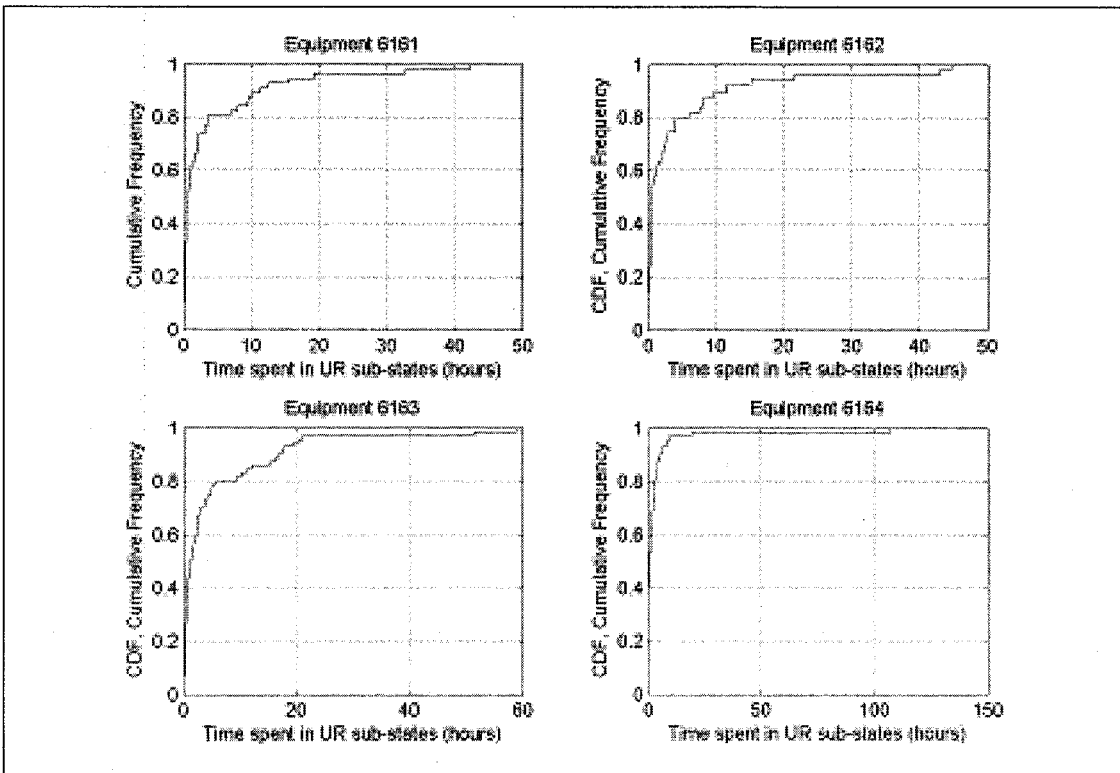
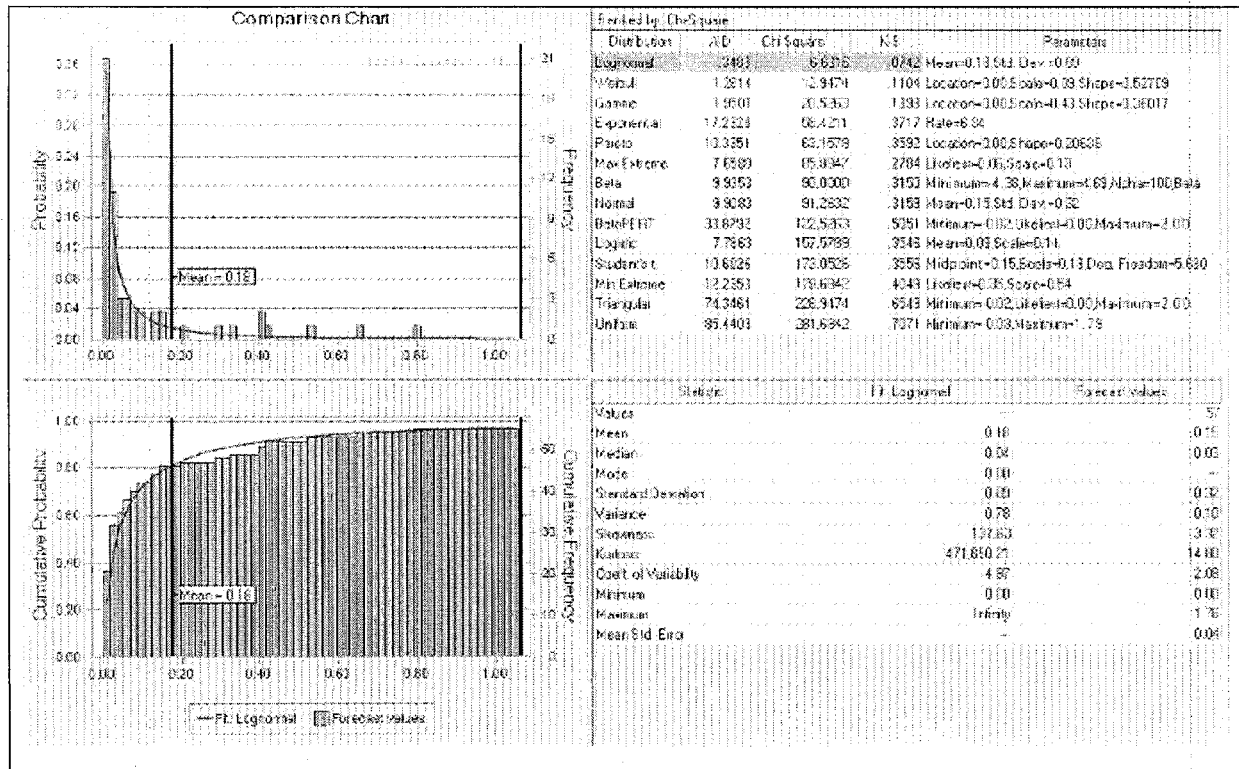
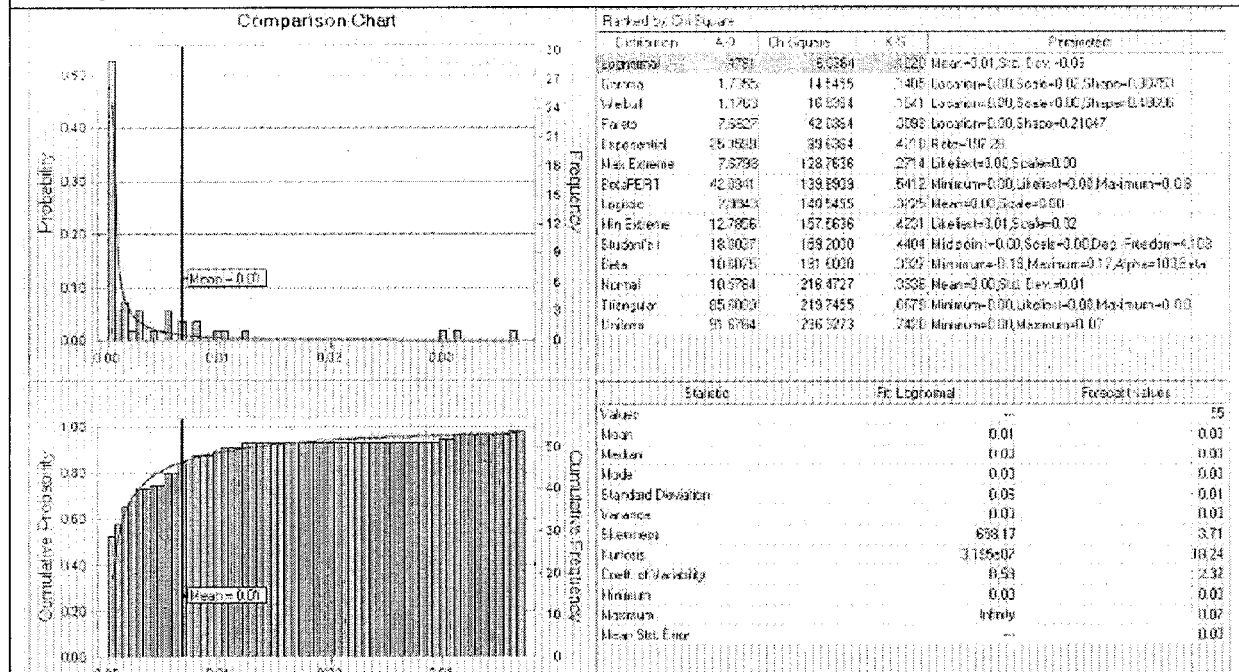


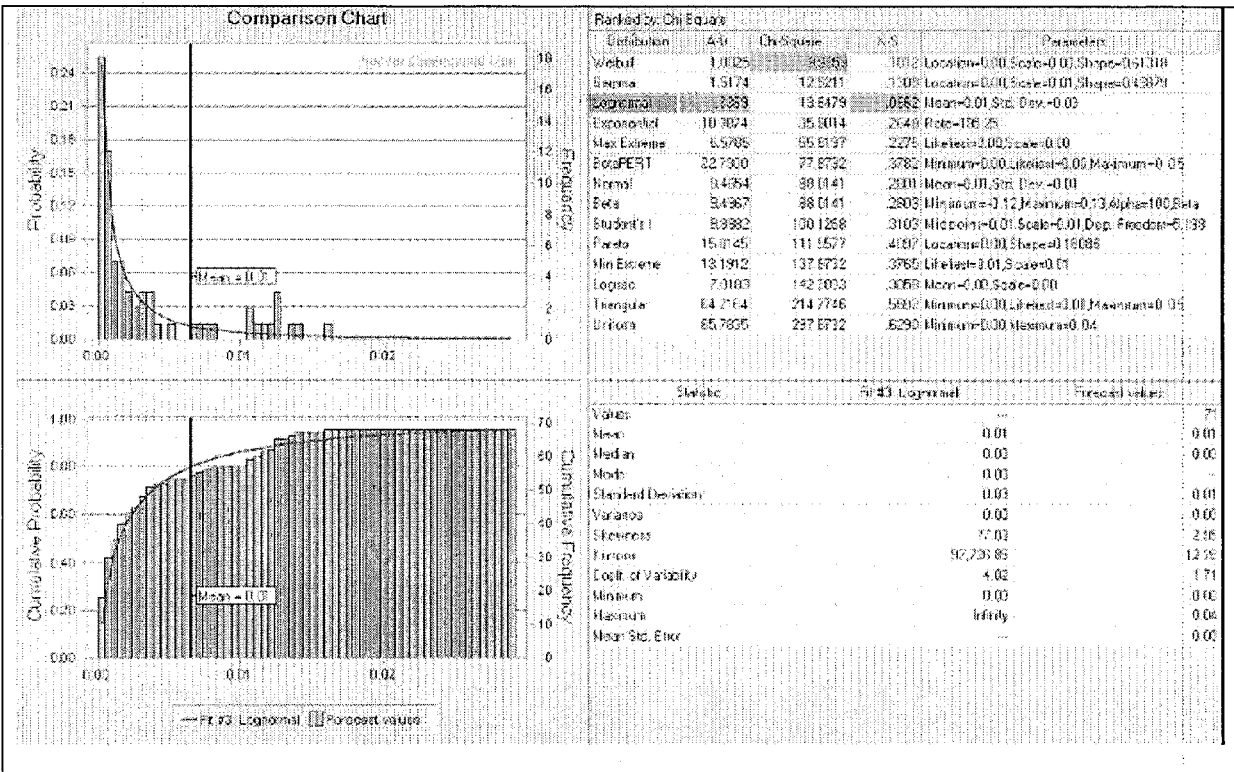
Figure 6: Lognormal fit for the duration (time in state - TIS) of UR sub-state



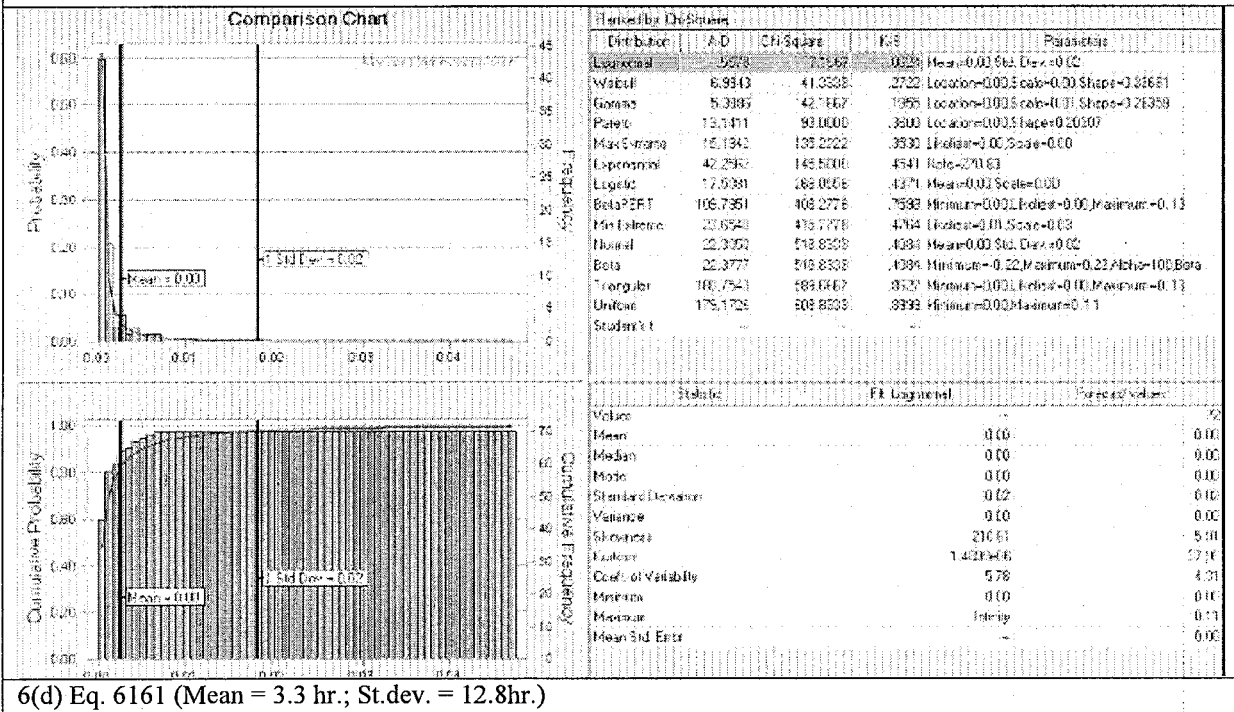
6(a) Eq. 6161 (Mean = 3.8 hr.; St.dev. = 7.8 hr.)



6(b) Eq. 6162 (Mean = hr.; St.dev. = hr.)



6(c) Eq. 6163 (Mean = 5.3 hr.; St.dev. = 10.3 hr.)



6(d) Eq. 6161 (Mean = 3.3 hr.; St.dev. = 12.8hr.)

Figure 7: An example of a time-window with several state transitions

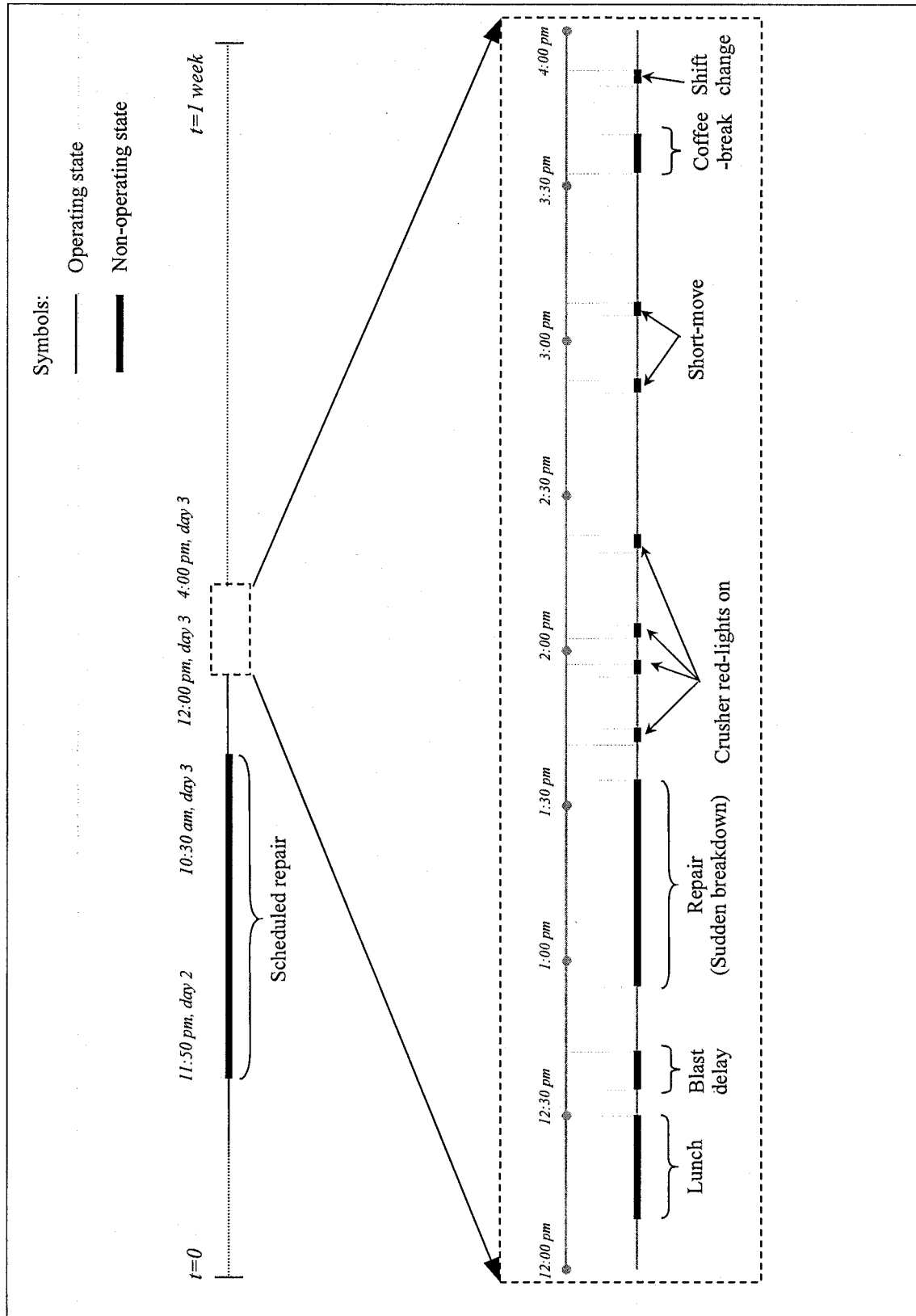


Figure 8: Poisson probability mass function using a point estimate of rate of occurrences of UR sub-state

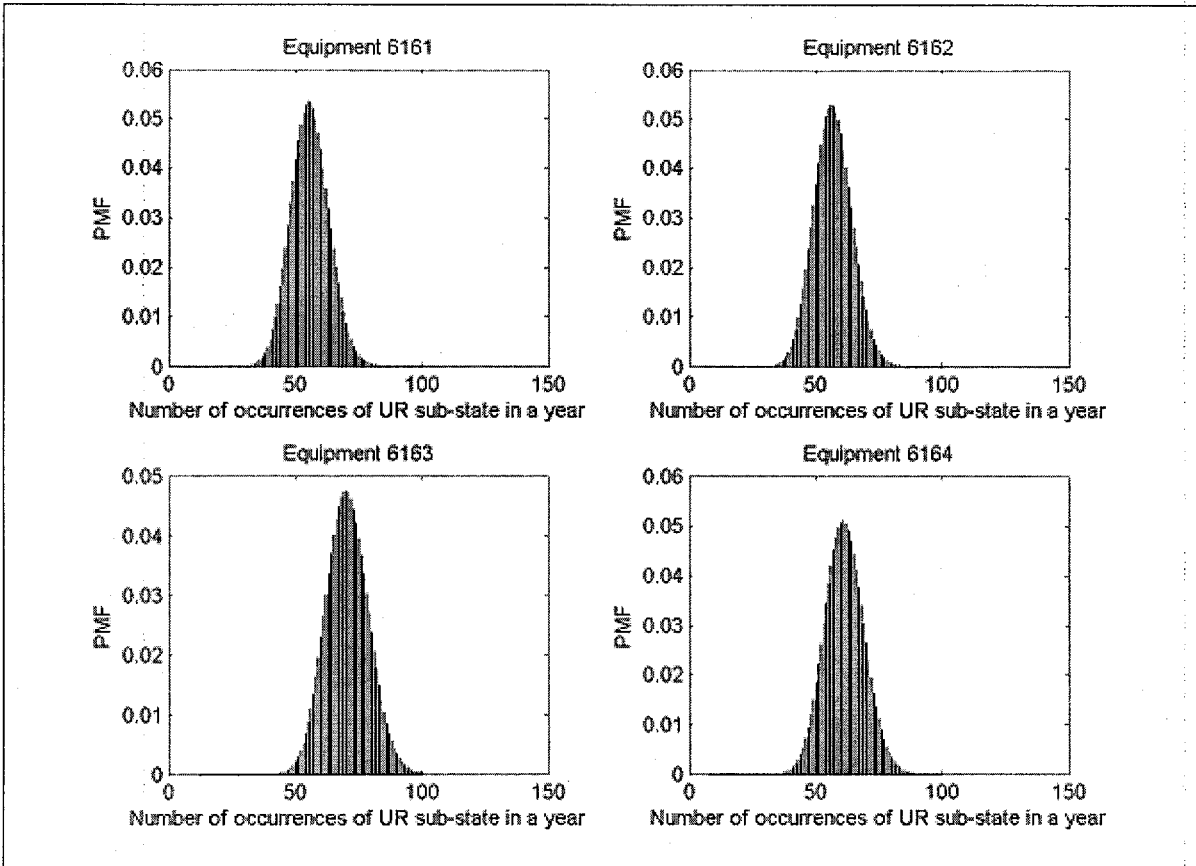


Figure 9: Number of Poisson occurrences with predictive and point estimation of sub-state occurrence rate for equipment 6161.

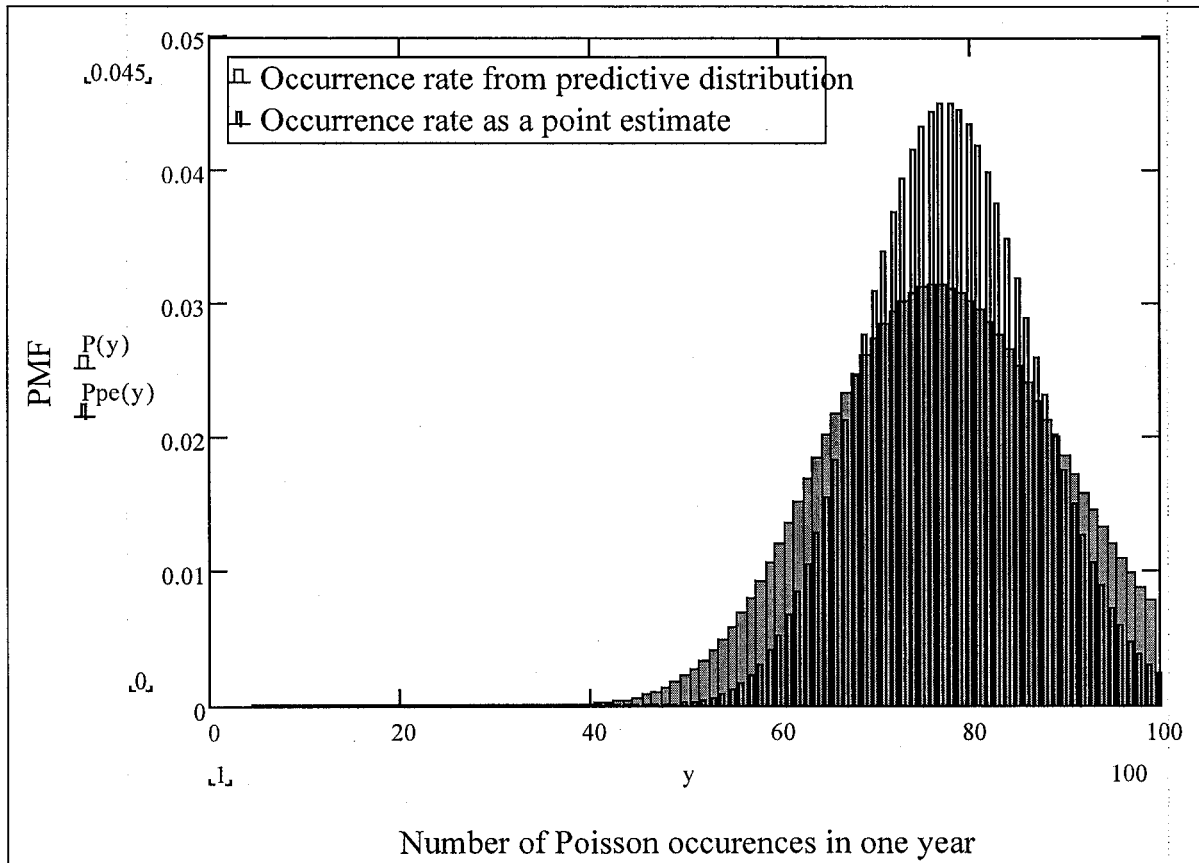


Figure 10: Predictive-CDF plot for the failure of equipment number 6161

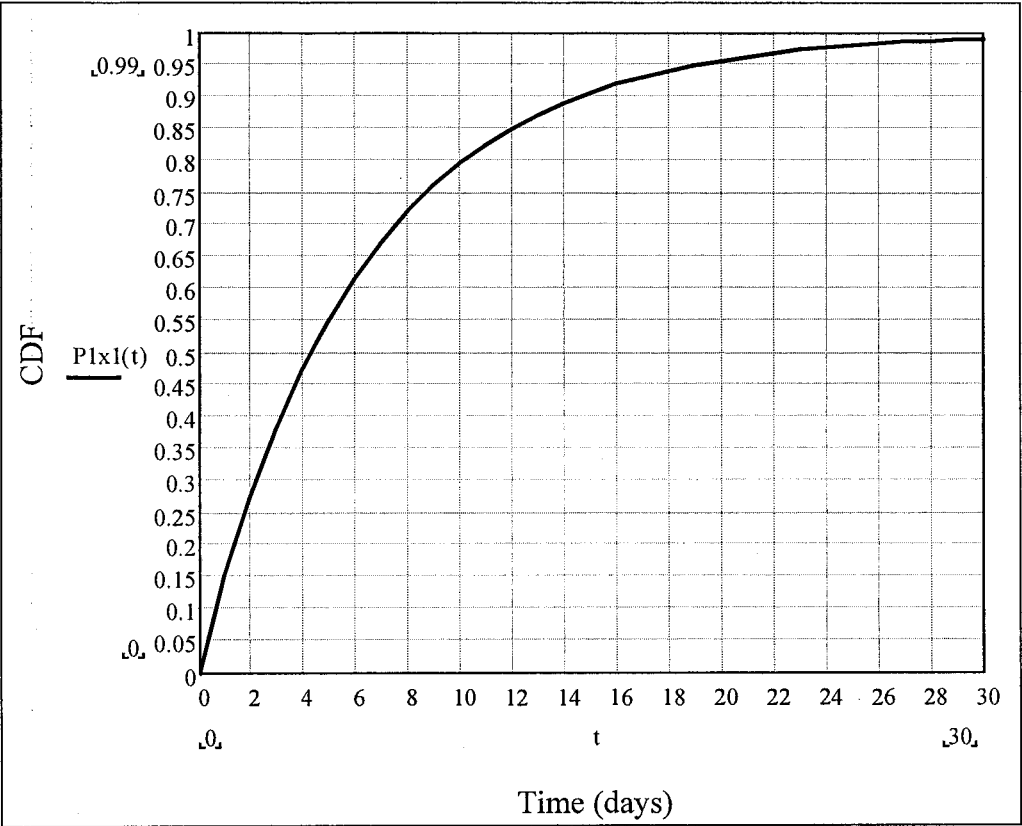
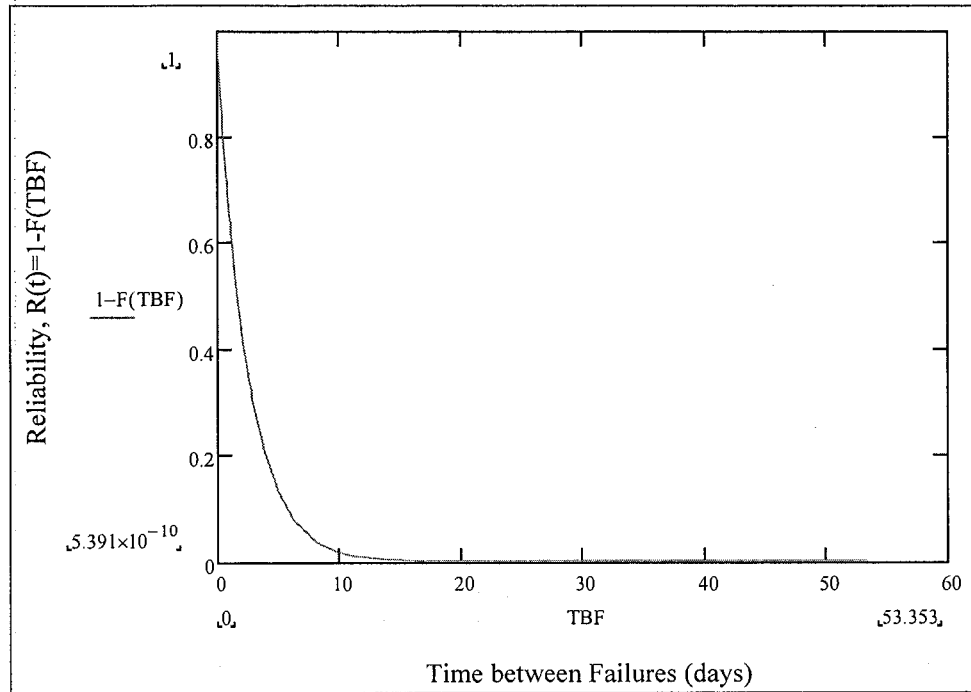


Figure 11: Reliability plot for equipment 6161, Occurrence rate ~ 0.4



3 Reliability Analysis

In this chapter, reliability formulations are developed for assessing the productivity and economic costs of the ore-hauling operation. The focus of the study is the impact of equipment downtime and other non-operating instances on the productivity and the economic costs of the operation. The productivity is measured in terms of total equipment time in operation, i.e., its production time. The data analysis carried out in the previous chapter provides probabilistic measures of the time spent by equipment in its various non-operating states. This serves as an input for the reliability formulations developed in this chapter. A novel aspect of the presented analysis is the utilization of structural reliability concepts for assessing the production and cost reliabilities of the operation. The relevant concepts of structural reliability are briefly discussed in the following.

In today's modern mining setup, HEMM equipments are increasingly deployed because of their large hauling capacity. This yields time-saving during loading compared to using smaller equipment, and carries reduced operator cost. However, the risk of deploying HEMM equipment is larger because of the high consequence of failure of one HEMM. It is necessary that downtimes and idle-time (time in non-operation) of such equipments are minimal because loss of production results in high economic (opportunity) costs. Importantly, the overall cost of an earth moving mining operation is significantly affected by the reliability of the HEMM equipments used in the operation. Lower reliability of the equipment alters productivity and drives the maintenance and operating costs; thus affecting profit margins. This motivates the focus on productivity and cost-based limit-states in the subsequent reliability analysis.

3.1. Reliability Formulation

Reliability is defined as the probability that a component will perform its intended function until a specified point in time under encountered use conditions. In mining engineering, reliability is interpreted as the probability of non-failure of the component.

A typical system includes a number of components arranged in series, parallel, or a combination thereof, which are visualized by means of reliability block diagrams. In structural reliability, the preliminary steps in evaluating reliability of the system include: finding the failure modes of individual components (limit-states); determining the set of components whose joint failure constitutes failure of the system (Melchers, 1999). Structural reliability analysis involves the utilization of the concept of limit-state function. A limit-state function is a mathematical representation of an event of interest, such as the component failure, which results from load (demand) on the component exceeding its resistance (capacity). The limit-state function is designated by $g(\mathbf{x})$, where \mathbf{x} is a vector of random variables. Negative and positive values of g signify the failure state and the safe, respectively.

The limit-state-based reliability formulation developed in this thesis computes the reliability of a production system comprising a set of equipment. Specifically, the system comprises production from the ore-hauling operation using a fleet of four HEMM equipments. The production is assessed in the form of equipment production time, i.e., the total time spent by equipment in the operating state. The limit-state function-based reliability formulation for production time of an equipment is

$$g(\mathbf{x}) = T_o(\mathbf{x}) - T_\theta \quad (15)$$

where $T_O(\mathbf{x})$ is the total time spent by equipment in the operating state (production time), and T_θ is a user-selected time threshold. Effectively, a reliability analysis with this limit-state function yields the probability that the time in operation, T_O , is less than the selected threshold, T_θ .

The data analysis in the first part of the paper provides input to the probabilistic characterization of T_O . Because the Bayesian updating assessed the rate of events in *non*-operation, the limit-state function in Eq. (7) is re-formulated to read

$$g(\mathbf{x}) = [T - T_{NO}(\mathbf{x})] - T_\theta \quad (16)$$

where T is the total observation period and $T_{NO}(\mathbf{x})$ is the total time spent by the equipment in any of the non-operating states (downtimes and idle times). The limit-state-based reliability formulation in Eq. (16) represents the production contribution of equipment measured in production time. Notably, the formulation considers not only the equipment downtime due to failure instances, it also considers the idle-time of the equipment caused due to other non-operating instances. The limit-state function of Eq. (16) is evaluated for each equipment separately and then combined to obtain the productivity contributions from all the equipments. This has the novel implication that a problem that is traditionally addressed as a “system reliability problem” (Pukite and Pukite 1998) is here analyzed by means of *one* cumulative production-oriented limit-state function.

The probability of production time being less than a threshold value is given by integrating the PDF of the vector \mathbf{x} over the failure domain (Ang and Tang 1970)

$$p = P[g(\mathbf{x}) \leq 0] = \int_{g(\mathbf{x}) \leq 0} \dots \int f(\mathbf{x}) d\mathbf{x} \quad (17)$$

where $f(\mathbf{x})$ is the joint PDF. Eq. (17) is a component problem (single limit-state function), which can be solved using various numerical techniques such as First-order Second-moment Method (FOSM), First-order Reliability Method (FORM), Second-order Reliability Method (SORM), Importance sampling, and Monte-Carlo sampling (Melchers 1999).

In accordance with the state modeling approach described above, $T_{NO}(\mathbf{x})$ is a function of the number of occurrences and duration of various non-operating sub-states. Figure 12 describes the methodology of calculating $T_{NO}(\mathbf{x})$ for an equipment. The figure illustrates n different non-operating sub-states within the observation period, wherein occurrence and duration characteristics of sub-state $x=1$ (e.g. Lunch break) is deterministic and that of sub-states $x=2, 3 \dots (n-1)$, n are random. Accordingly, the time in sub-state 1 is obtained by multiplying the total number of occurrences with the duration of each occurrence. The total number of occurrences of random sub-states N_x is given by the Poisson probability mass function, $p(N_x)$, which has a mean occurrence rate, λ_x , which in turn has the Gamma distribution from Bayesian updating; $f'(\lambda_x)$. The i^{th} occurrence of the sub-state, where $i = 1, 2, 3 \dots N_x$, lasts for a random duration $d_{x(i)}$, with distribution $f(d_x)$. The time in a random non-operating sub-state is calculated using Eq. (1). The total time in both deterministic and random non-operating sub-states is given by

$$T_{NO}(\mathbf{x}) = D_o + \sum_{x=2}^n D_x \quad (18)$$

where D_o is the total time spent by the equipment in deterministic non-operating sub-states.

Substituting Eq. (18) in Eq. (16) it is observed that the limit-state function comprises a linear combination of discrete (N_x) and continuous random variables $f(d_x)$. This results in a non-smooth limit-state function, making the analysis using FORM practically challenging

(Haukaas and Der Kiureghian 2006). Sampling is adopted as an alternative in the present study, in which the primary objective is to put forward the overall methodology. It is also noted that with the limit-state function definition in Eq. (16) there is a zero likelihood of outcomes in which the total time in non-operation exceeds the observation period, i.e., $T_{NO}(\mathbf{x}) > T$. This is a potential shortcoming of the model, however, the probability of total time in non-operation exceeding the observation period is very small.

As described in Eqns. (1) and (18), $T_{NO}(\mathbf{x})$ is the sum of D_o and D_x , which in turn is obtained by summing random durations $d_{x(i)}$. Furthermore, random durations $d_{x(i)}$ for each sub-state x have different distributions $f(d_x)$, as listed in Table 6. The computation of $T_{NO}(\mathbf{x})$ using sampling should be such that the sum of random durations $d_{x(i)}$ for all sub-states x don't exceed T . This is possible if the following constraint is met

$$\left[D_o + \sum_{x=2}^n D_x \right] + T_o(\mathbf{x}) = T \quad (19)$$

i.e., the sum of the total time spent by the equipment in operating and non-operating sub-states always equals the observation period, T . However, because the operating state dominates any time period beyond, say, one month, this is only an academic issue that has not caused problems in the present study. One approach for computing the failure probability using sampling is to keep the threshold value constant and divide the number of negative outcomes of limit-state function with total number of samples. Typically, the sample outcomes of limit-state function are generated until the coefficient of variation of the failure probability is below, say, 2%. Then, in order to obtain a CDF curve, sampling is performed at different threshold values.

A more efficient approach adopted in this study is to directly generate outcomes of the equipment production time, i.e., without considering the threshold in the limit-state function. These outcomes of time in non-operation are post-processed to obtain distribution plots for production time. A limitation of this sampling approach, however, is the weak inference at the tails. This is because relatively less numbers of random samples get picked from the extremes of the distributions in this approach. In the previous approach, sampling can be performed by selecting the thresholds near the tails to obtain more informed distribution.

Figure 13 shows a flowchart for obtaining sampling outcomes of the total time spent by equipment in non-operating sub-states T_{NO} . Initial values of T_{NO} and sub-state x (now interpreted as the number of the sub-state) is set as 0.0 and 1 respectively. The final value of T_{NO} is computed by adding to it the total durations corresponding to each sub-state D_x , in accordance with Eq. (18). For each sub-state $x=1,2,3...n$, the Bayesian updated parameters k_x'' and v_x'' are inputted to obtain the updated Gamma distribution $f''(\lambda_x)$. Then, a random outcome for mean occurrence rate λ_x is generated from the updated Gamma distribution. Now for the randomly generated mean occurrence rate, Poisson random outcome for number of occurrences of sub-state N_x in period T are obtained. In order to find the total time duration spent by equipment in sub-state x , the initial value of D_x is set as '0'. Now, N_x random outcomes of duration in sub-state d_x are computed. Here, each random duration outcome corresponds to a random occurrence. All the duration outcomes are finally added to obtain the total time duration D_x , which in turn is added to T_{NO} and the sub-state counter, x is incremented to next sub-state. Above steps are repeated for all sub-states of the equipment until the final value of T_{NO} is obtained.

A fleet of four equipment, numbered 6161, 6162, 6163, and 6164 for identification purpose, has been considered for reliability analysis. Complete distribution parameters for the fleet of equipment were derived from the analysis of available database. Specifically the Tables 6 and 11 lists parameters for Bayesian updated Gamma distribution and for theoretical distributions of duration $f(d_x)$ for each sub-state. Based on the data analysis, the reliability formulation described in Eq. (18) is analyzed for all equipments using sampling. A significant number of samples (100,000) for T_{NO} are generated in accordance with the described methodology.

Another novel aspect of the reliability formulation includes an algorithmic formulation of a cost-based limit-state function. Different costs are associated with different states of equipment functioning.

Table 12 estimates the most significant state-specific costs including the operating cost C_{OP} , scheduled repair costs C_{SR} , unscheduled repair costs C_{UR} , and economic cost C_E (Cost of opportunity lost due to non-production). The impact of economic cost is the most significant of all the costs. It is noted that the economic cost for loss of production due to non-operation of equipment only has been considered here. State-specific cost considerations are incorporated in the reliability formulations to study the impacts of equipment downtime and other non-operating instances on the costs of operation. This is done by multiplying costs per unit time with the total time spent in respective sub-states. For example, in the set-up described in Figure 12, the economic and unscheduled repair costs $C(x)$ due to non-production and equipment downtime is calculated as

$$C(x) = \left[\sum_{x=2}^n D_x \cdot C_E \right] + D_{UR} \cdot f(C_{UR}) \quad (20)$$

where C_E is the economic cost per unit time, D_{UR} is the time spent by equipment in unscheduled repair, and $f(C_{UR})$ is a uniform distribution for repair costs per unit time. The corresponding cost-based limit-state function is then

$$g_c(x) = [C(x) - C_\theta] \quad (21)$$

where C_θ is a user-selected cost threshold value. In a similar manner, different types of state-specific cost measures can be incorporated into the basic reliability formulation. The cost measures may also be incorporated in the form of probability distributions derived from historic cost data.

Table 13 describes the empirical estimates of economic cost impact due to the non-operation of one HEMM equipment. The estimates are obtained from the daily production data for CAT-793 trucks at the Highland Valley Copper Mine. The first few rows of the Table 13 list the daily productivity data of one truck. Accordingly, the average daily ore production at the mine is 133000 tonnes, with 76,000 tonnes (57% of daily production) ore getting hauled by the trucks. Twelve trucks are operational at the mine, making the available haulage capacity as 6333 tonnes per truck per day. The yield from the ore is 0.43%, resulting in 27 tonnes of enriched copper being hauled per truck per day. The last three rows of Table 13 provide the estimates of economic cost impact. Accordingly, the median price of Copper is considered to be \$4,445/tonne. This results in an hourly revenue impact of non-production of \$5,044/hour, which is also the economic cost. It is observed that the operating and miscellaneous expenses of the truck have not been considered in the presented reliability formulation. These cost assumptions are inputted in Eq. 21 to obtain results for the fleet of equipment.

Another significant cost related to the operation of HEMM equipments is the cost of repairs. The cost of repair depends upon the nature of repairs and the duration of repair. For example, an UR costs more than a scheduled repair and repair of certain equipment components cost more than others. Considering this level of variation, the costs owing to unscheduled repairs are assumed to follow a uniform distribution in the discussed numerical example. The parameters of the distribution are assumed to range between \$1,000 and \$100,000, for each repair event. These cost assumptions are inputted in Eq. (20) to obtain results for the equipments.

The results obtained from the analysis provide the decision maker with an estimate of the production attainable from the fleet of equipment, based on the historic data. The presented reliability methodology is simple to implement. Further, the discrete sub-state based modeling approach offers considerable flexibility in reliability formulation for a wide number of applications. For example, probabilistic measures of total time spent by equipment in specific non-operating sub-states, such as UR, can be obtained. This provides the decision-makers with relevant information for allocation of future equipment maintenance budget and for choosing appropriate system redundancy levels. Similarly, probabilistic measures of total time spent by equipment in other non-operating states provide useful inputs for improving work schedules, process flows, redundancy, etc.

3.2. Results and Applications

In this section, the results obtained from the analysis of reliability formulations are discussed. The results present probabilistic curves for the productivity contributions and cost impacts from a fleet comprising four numbers of CAT-793® trucks involved in the ore-hauling mining operation. The productivity is assessed in the form of production time, which is the

total time spent by the fleet in operation. The production time is measured in units of truck-days. For example, the fleet operating constantly for one day (24 hours) has a production time of four truck-days. Cost is also explained in a separate section. Sampling is performed for estimating the probabilities of production time for the fleet in half-year (182.5 days), one-year (365 days), and two-year (730 days) observation periods. The production capacity of the fleet in half-year is four times 182.5, i.e., 730 truck-days. Similarly, the production capacity of the fleet for one and two year observation period is 1460 and 2920 truck-days respectively. It is assumed that the fleet is deployed for operation at the mine site during all times within the considered observation periods.

3.2.1. Production Time

A significant number of samples (100,000) for production time of the fleet are generated using a sampling algorithm. The algorithm is based upon the method described in Figure 13 and it takes into account the sampling constraints described in Eq. (19). The duration and occurrence data presented in Tables 6 and 11 serve as input for the algorithm. The sampling algorithm is programmed and executed using a student-licensed version of MATLAB® software. The detailed sampling algorithm is presented in electronic form (CD-ROM) attached with the thesis (Appendix B).

The results obtained from sampling are presented in the form of PDF and CDF plots. Figure 14 shows a PDF plot of the sampled production time of the fleet for half-year period. According to the plot, the total production time for the fleet is estimated to be between 376.15 and 531.74 truck-days, with a standard deviation of 10.1 truck-days and coefficient of variation of 2%. It is observed that the left tail of the distribution is discontinuous and spread out widely between 376 and 450 truck-days. The reason for this is that very few points get

sampled from the tail of the $f(d_x)$ distributions. This is a weakness of the sampling approach when compared with the other approaches for solving the integral described in Eq. (17). The expected value of production time for the fleet in half-year period is 499.72 truck-days. Overall, the results indicate that 68.5% of the installed productive capacity of the fleet (considering 499.72 truck-days of production time from the available production time of truck-730 days for the fleet) is estimated to get utilized within half-year. The associated cumulative frequency plot is illustrated in Figure 15. According to the plot, the probability of total production time of the fleet being at least 500 truck-days within half-year is 0.47. The probability of total production time of the fleet remaining at least 490 and 510 truck-days respectively within the half-year period is 0.15 and 0.87.

Similar results are obtained for assessing total production time of the fleet over one and two-year observation periods. Figure 16 shows a PDF plot of the sampled production time of the fleet for the one-year period. According to the plot, the total production time for the fleet is estimated to be between 805.3 and 1054.3 truck-days, with a standard deviation of 14.8 truck-days and coefficient of variation of 1.48%. The expected value of the production time for the fleet in one-year period is 999.47 truck-days. Considering the expected value of the production time of the fleet in one-year, 68.45% of the installed productive capacity of the fleet is estimated to get utilized, i.e., 995 truck-days of production from the available production time of 1460 truck-days for the fleet. The associated cumulative frequency plot is illustrated in Figure 17. According to the plot, the probability of the total production time of fleet being at least 1000 truck-days within one-year is 0.50. The probability of total production time of the fleet remaining at least 980 and 1020 truck-days respectively within one-year is 0.10 and 0.94.

Figure 18 shows a PDF plot of the sampled production time of the fleet for a two-year observation period. According to the plot, the total production time for the fleet is between 1768 and 2074 truck-days, with a standard deviation of 22.14 truck-days and a coefficient of variation of 1.11%. The expected value of production time for the fleet in two-year period is 1998.9 truck-days. Considering the expected value of production time, 68.45% of the installed productive capacity of the fleet is likely to get utilized. Figure 19 shows the associated cumulative frequency plot. According to the plot, the probability of total production time of the fleet being at least 2000 truck-days within two-years is 0.50. The probability of total production time of the fleet remaining at least 1960 and 2040 truck-days respectively within two-years is 0.05 and 0.98.

According to the CDFs of production time for the half-year, one-year, and two-years periods, the probability of total production time of the fleet being at least 500, 1000 and 2000 truck-days, respectively, is 0.5. However, this does not imply that the probability results for a different time period can be obtained simply by scaling of the results from *one* analysis. To explain this, compare the PDFs of production time for the half-year, one-year, and two-year periods. It is noted that only the expected values; that is, where the CDF value is 0.5, that the production time vary linearly, as noted above. The values in the tails of the distribution do not show a linear correspondence between time period and threshold. In fact, for the three observation periods, the probability of total production time of the fleet remaining at least 510, 1020, and 2040 truck-days respectively is 0.87, 0.94, and 0.98. Furthermore, the standard deviations for the production time over the three observation periods do not vary linearly. A simple calculation reveals that the standard deviation varies approximately by the square root of the observation period. This is expected, because of the nature of the limit-

state function. Specifically, the random quantity in the limit-state function is “time in non-operation.” As described previously, realizations of this random quantity is obtained as the sum of number of occurrences multiplied with the duration of each occurrence. The Poisson distribution governs the number of occurrences in a time period. From this distribution it is readily derived that the expected number of occurrences is proportional to time, while the standard deviation of the number of occurrences is proportional to the square root of the time period under consideration. Hence, this explains the results. The reason that the standard deviation is not exactly proportional to the square root of the considered time periods is the added uncertainty in the *duration* of the occurrences. In other words, the limit-state-function described in Eq. (16) basically comprises of the occurrence rate parameter λ and the distribution for duration, $f(d_x)$. The expected value of the total number of Poisson occurrences, $(\lambda \cdot T)$, varies linearly with the observation period, T . Therefore, the non-linear scaling of production time at the tails of the distributions for the three observation periods is likely to be due to $f(d_x)$.

3.2.2. Economic and Unscheduled Repair Costs

The cost-based reliability formulation described in Eq. (21) is evaluated using sampling. The duration, occurrence and cost data presented in Tables 6, 11, and 13 respectively serve as the input for the sampling algorithm. The results from sampling provide probabilistic measures for the impact of equipment downtimes and non-operating instances on the economic and UR costs for different observation periods.

Figure 20 shows a PDF plot for the sampled economic and UR costs of the fleet in a half-year observation period. According to the plot, the total economic and UR costs for the fleet for the half-year period is between \$1.28 and \$7.4 million, with a standard deviation of

\$ 0.89 million and a coefficient of variation of 29.1%. The expected value of the total economic and UR cost in half-year is \$3.06 million. The dispersion of the distribution is relatively large. One of the reasons for this large dispersion is the assumption of the cost of each UR event to be uniformly distributed with parameters \$1,000 and \$100,000. The associated CDF plot of the economic and UR costs of the fleet for the half-year period is shown in Figure 21. According to the plot, the probability of economic and UR costs for the fleet being at least \$3 millions in the half-year period is 0.48. Further, the probability of economic and UR costs for the fleet being at least \$2 and \$4 millions respectively within the half-year period is 0.13 and 0.83.

Figure 22 shows the PDF plot of sampled economic and repair costs of the equipment fleet in a one-year observation period. According to the plot, the total economic and UR costs for the fleet for the one-year period is between \$2.94 and \$12.9 million, with a standard deviation of \$3.1 million and a coefficient of variation of 39.1%. The expected value of the total economic and UR cost in one-year is \$7.92 million. Figure 23 is the associated CDF plot for the same sampling outcomes. According to this plot, the probability of economic and UR costs for the fleet being at least \$6 million in one-year period is 0.475. Further, the probability of economic and UR costs for the fleet being at least \$4 and \$8 millions respectively within the half-year period is 0.135 and 0.835.

Similarly, Figure 24 and Figure 25 show the PDF and CDF plots of the sampled economic and repair costs of the equipment fleet within a two-year observation period. According to the plot, the total economic and UR costs for the fleet for the one-year period is between \$6.06 and \$24.48 million, with a standard deviation of \$ 5.73 million and a coefficient of variation of 37.5%. The expected value of the total economic and UR cost in

two-years is \$15.2 million. The CDF plot indicates that the probability of economic and UR costs for the fleet being at least \$12 million in two-years period is 0.475. Further, the probability of economic and UR costs for the fleet being at least \$8 and \$16 millions respectively within the half-year period is 0.125 and 0.83. The slope of CDF plots indicates the dispersion of the distribution. A flatter slope means more dispersion.

The above-discussed results find application in the analysis of importance and sensitivity measures of various costs and processes of the mining operation. Clearly, the cost and production time will be more sensitive to some sub-states than others. Sensitivity studies may be performed to investigate the effect of cost or time related parameters on the production probabilities. The probabilities can then be utilized in production optimization applications.

Furthermore, the data analysis approach of the method uses Bayesian statistics in order to obtain the occurrence rates of different states of equipment functioning. This facilitates the incorporation of various uncertainties of mining set up, making the approach more informative as compared to assuming point estimates of the occurrence rates.

The reliability formulations are flexible and can therefore be utilized for different optimization applications. For example, budgetary allocations for operating costs of equipment are generally based on production-planning schedule and historic operating cost patterns of the equipment. This approach, however, does not consider the different costs associated with the different states of equipment functioning. In this context, an important application of the discussed method is to incorporate sub-state specific costs in the reliability formulation to obtain more informative budget estimates. Furthermore, the costs owing to sudden equipment breakdowns result in considerable unplanned cost escalations through

repairs, replacements and the ensuing loss of production, etc. The cost-based reliability formulation, which considers the updated occurrence rate, incorporates the cost escalations owing to sudden equipment breakdown.

Typically, the maintenance schedule of equipment is decided on the basis of equipment manufacturer's recommendations and the usage levels of the equipment. However, despite the suggested routine maintenance, sudden equipment breakdowns happen all too often, owing to the influence of inherent variations in local environment. The Bayesian updated occurrence rate distribution for unscheduled repair events can serve as an important consideration during the planning of maintenance schedule.

Table 12: Costs associated with different states of functioning of equipment.

Cost Measures	Non-Operating states					Operating
	Delay	Operating Delay	Standby	Unscheduled Repair	Scheduled Repair	Operating
Operating cost	-	C_{OP}	-	-	-	C_{OP}
Repair cost	-	-	-	C_{UR}	C_{SR}	-
Opportunity/ Economic cost	C_E	C_E	C_E	C_E	C_E	-

Table 13: Empirical Productivity and Economic Cost Data

Daily Productivity Data (Source: Highland Copper Mine)	
Daily ore production	133000 tonnes/day
Estimated ore-haulage using CAT-793 trucks ~ 57%	76000 tonnes/day
Total number of CAT - 796 trucks operating at the mine site	12
Ore haulage capacity of each truck (Capacity: 240 tonne)	6333 tonnes/truck/day
Percentage of Copper in the ore (Yield)	0.43 %
Enriched Copper obtainable/truck/day	27 tonnes
Estimated Revenue Impact and Economic Cost of Non-operation	
Median Cu price	CND \$ 4445/tonne
Revenue Impact	CNDS \$ 121052/truck/day
Economic cost of non-operation (Opportunity Cost)	CNDS \$ 5044/truck/hour

Figure 12: Methodology of calculating the total time in non-operation, $T_{NO}(\mathbf{x})$ for an equipment

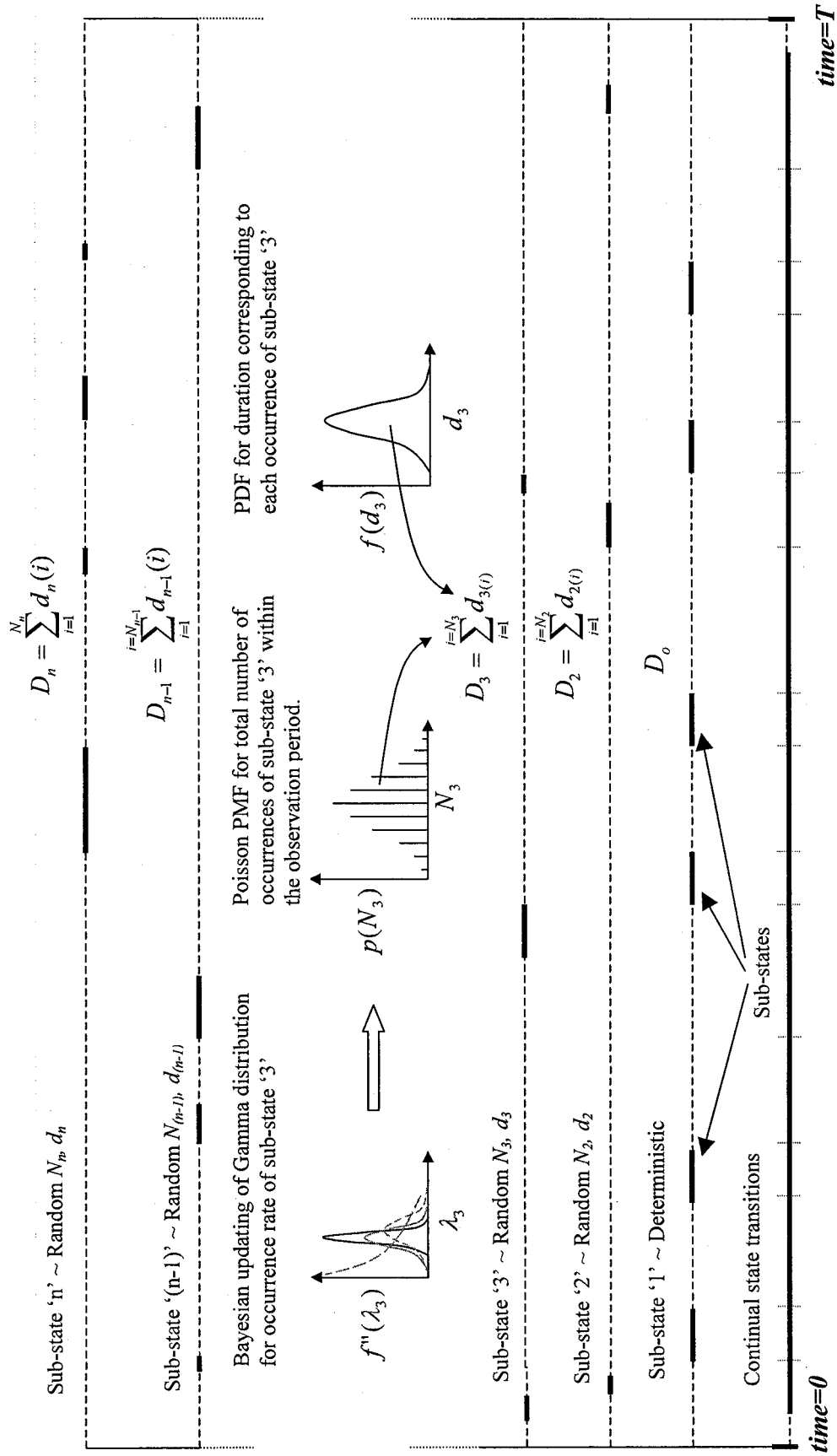


Figure 13: Flowchart illustrating an algorithm for sampling time spent in non-operating state

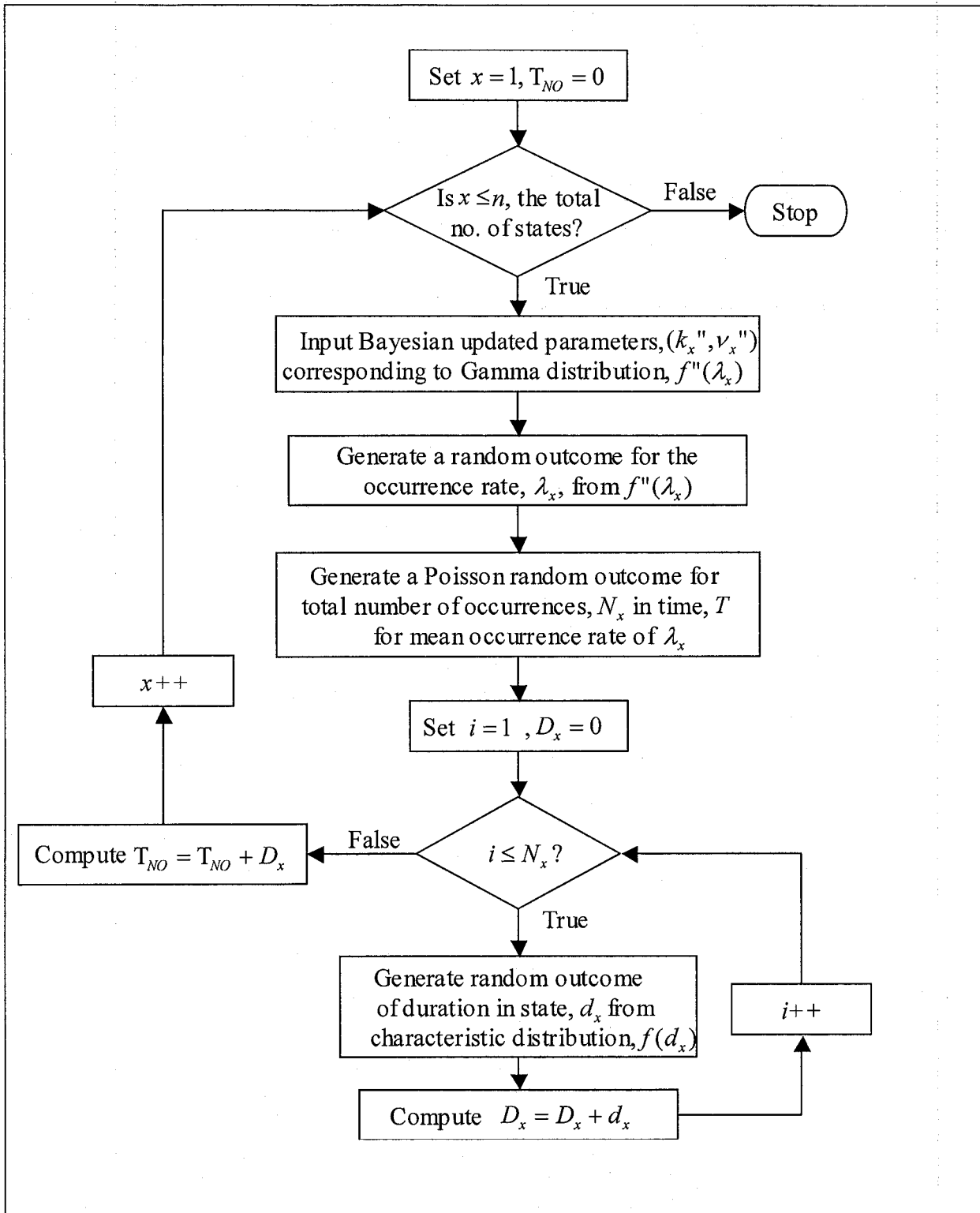


Figure 14: Sampling results - PDF plot for the total production time (units: truck-days) of the four HEMM equipments in half-year.

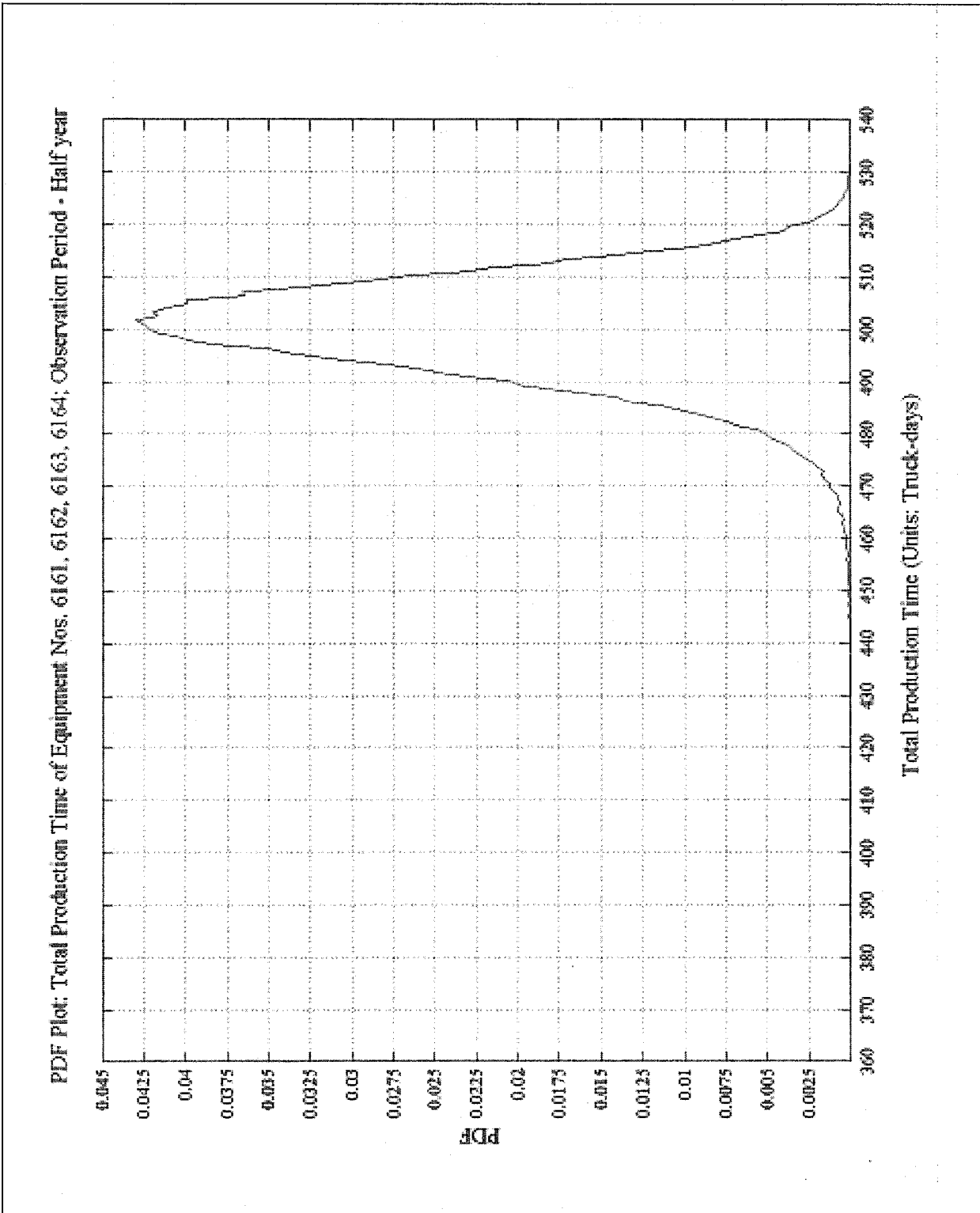


Figure 15: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in half-year.

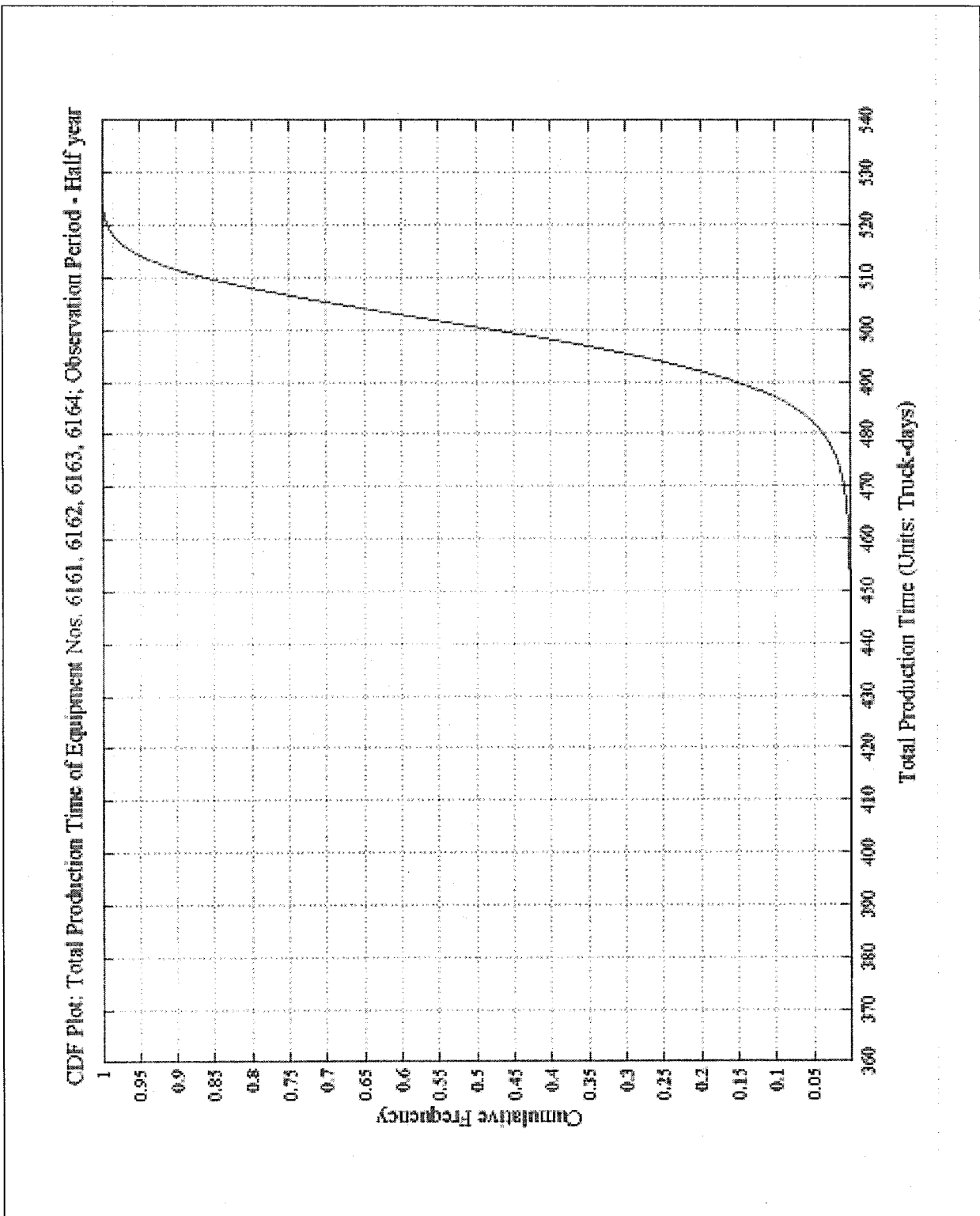


Figure 16: PDF plot for the total production time (units: truck-days) of the four HEMM equipments in one-year.

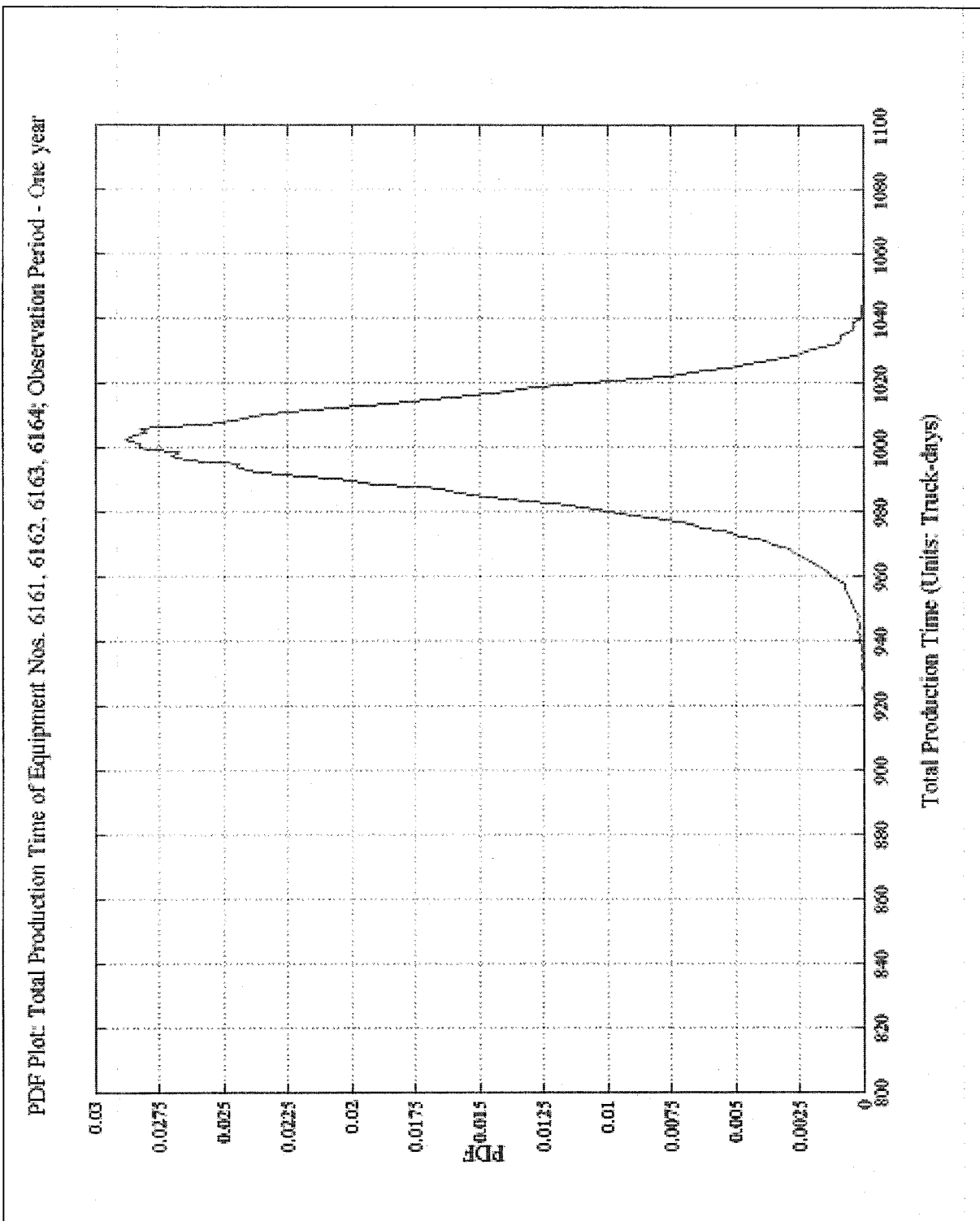


Figure 17: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in one-year.

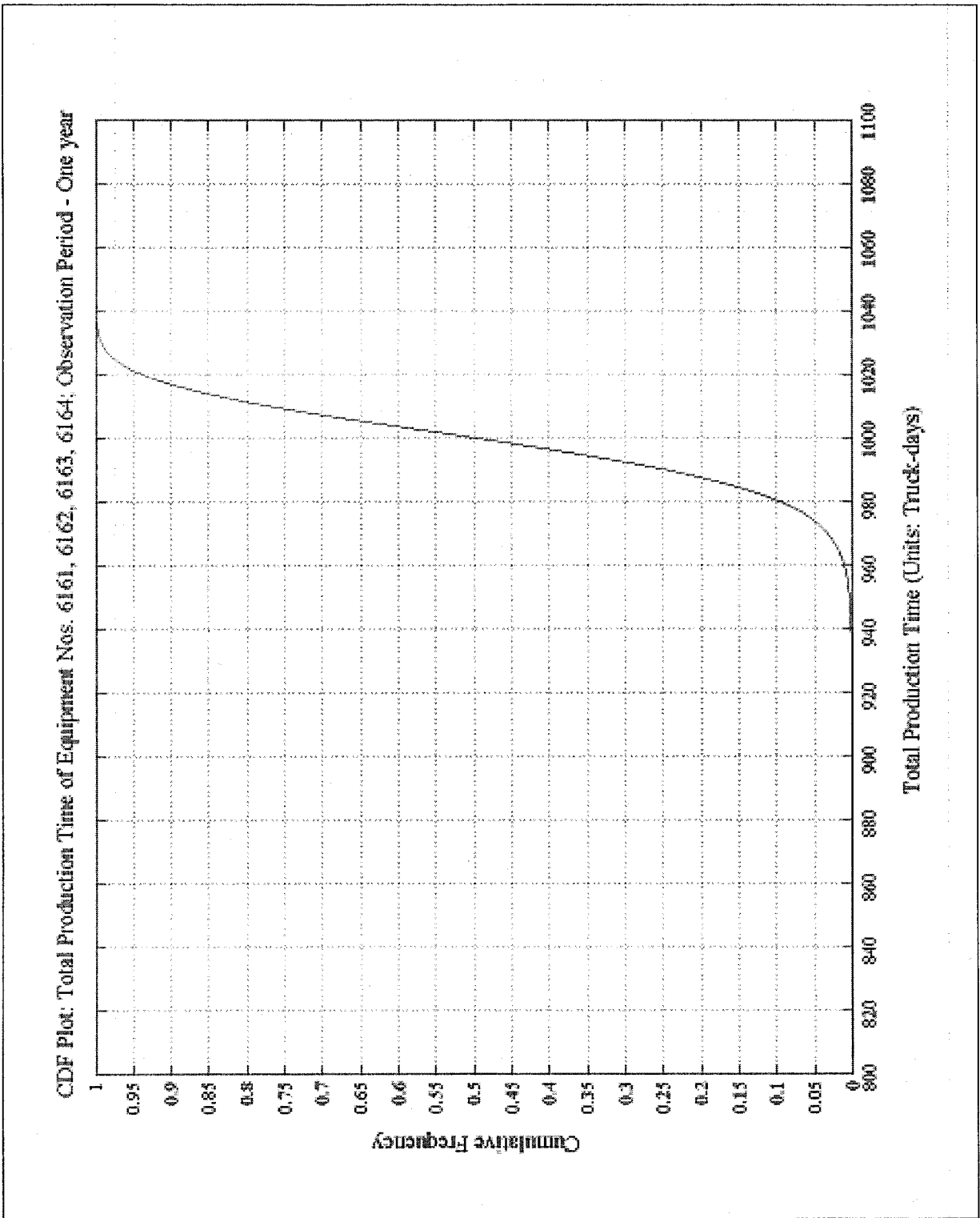


Figure 18: PDF plot for the total production time (units: truck-days) of the four HEMM equipments in two-years.

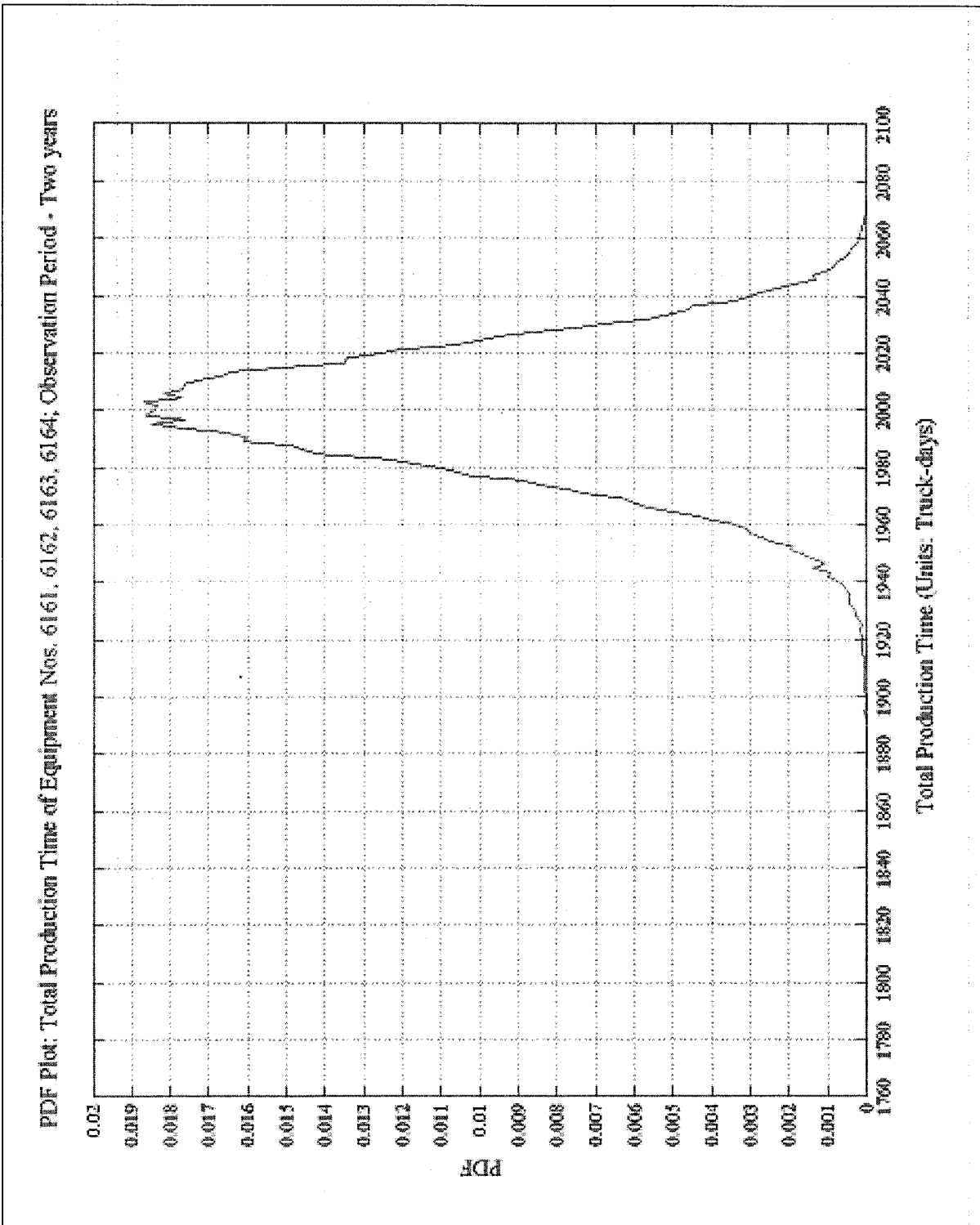


Figure 19: Sampling results - CDF plot for total production time (units: truck-days) of the four HEMM equipments in two-years.

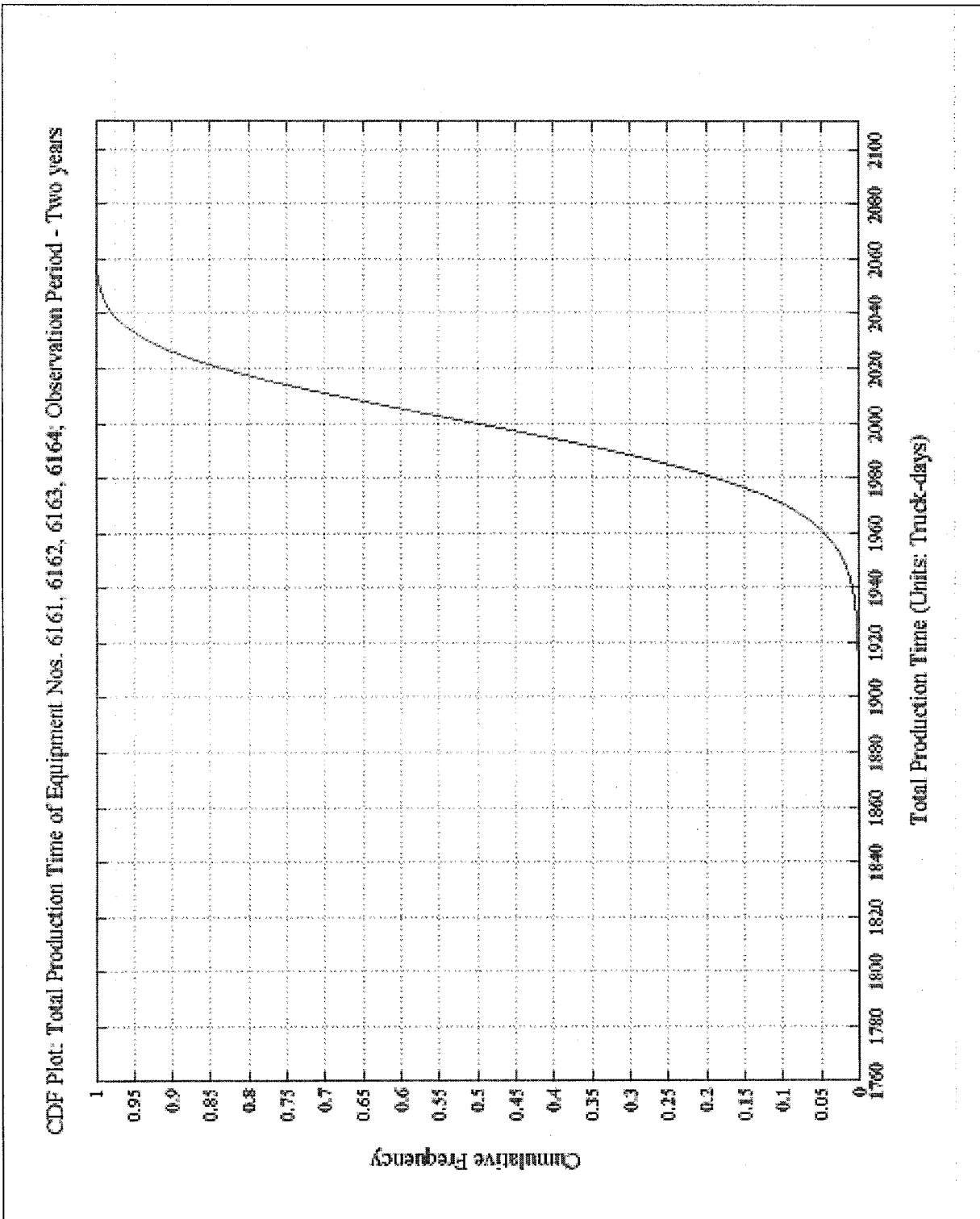


Figure 20: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in half-year.

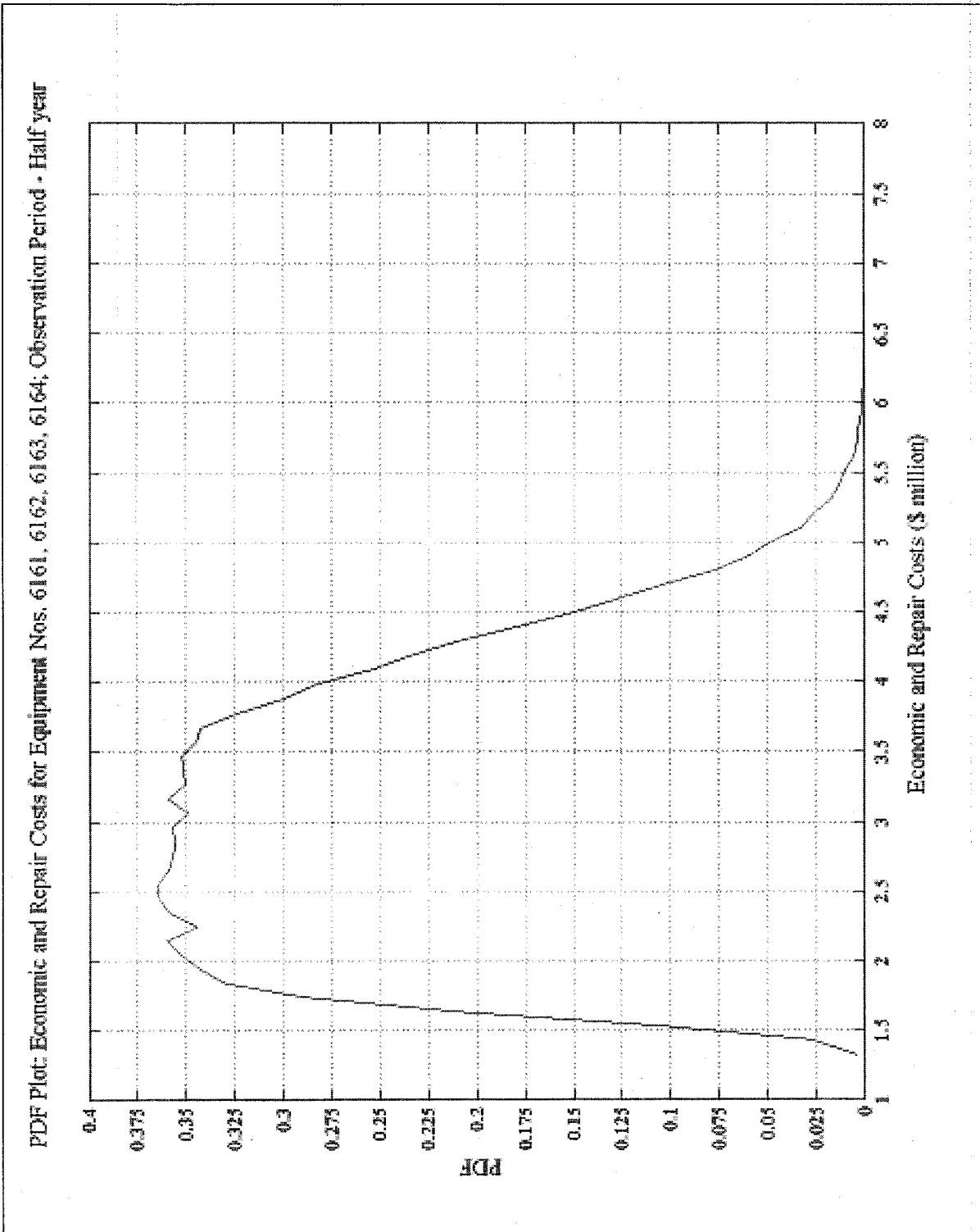


Figure 21: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in half-year.

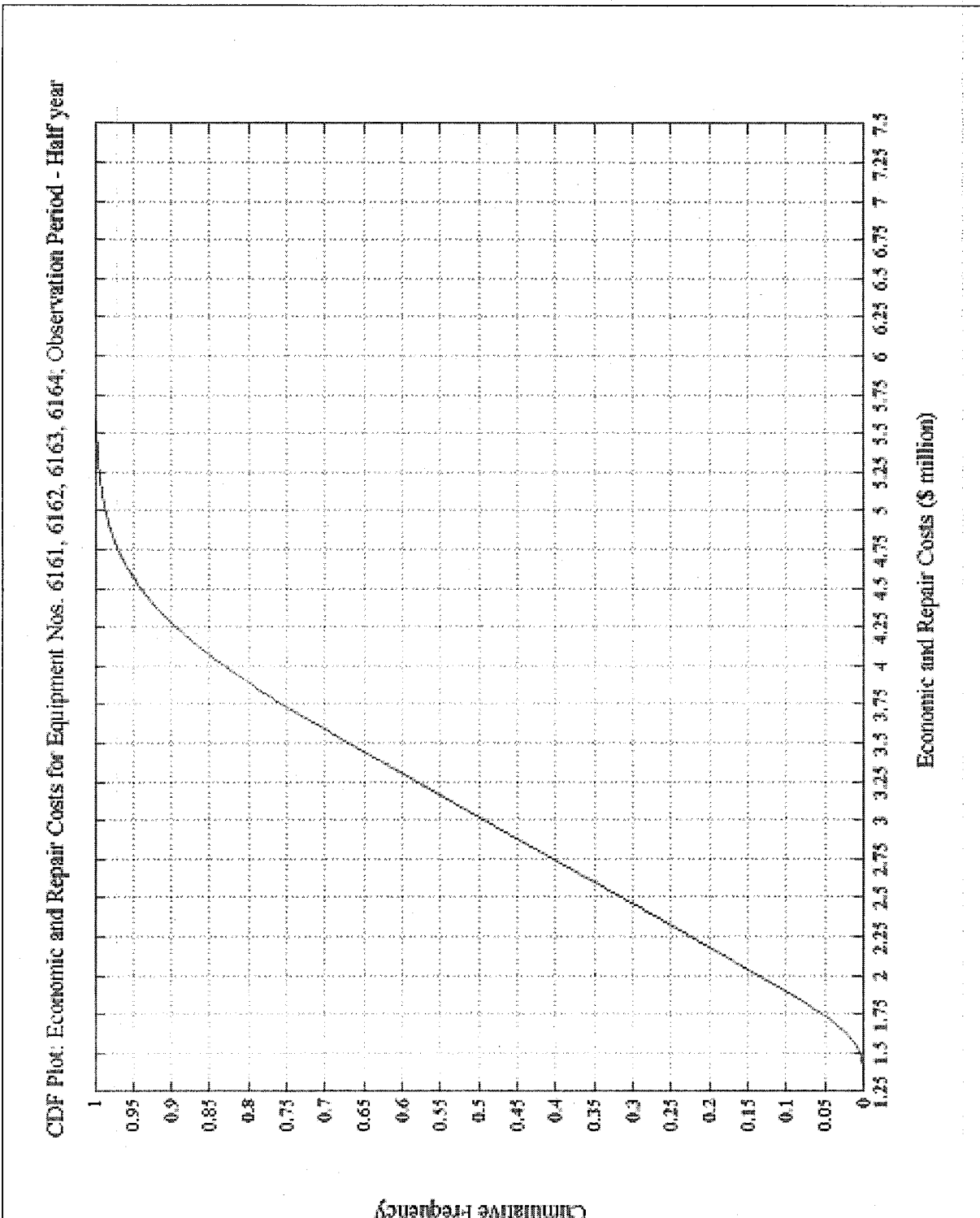


Figure 22: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in one-year.

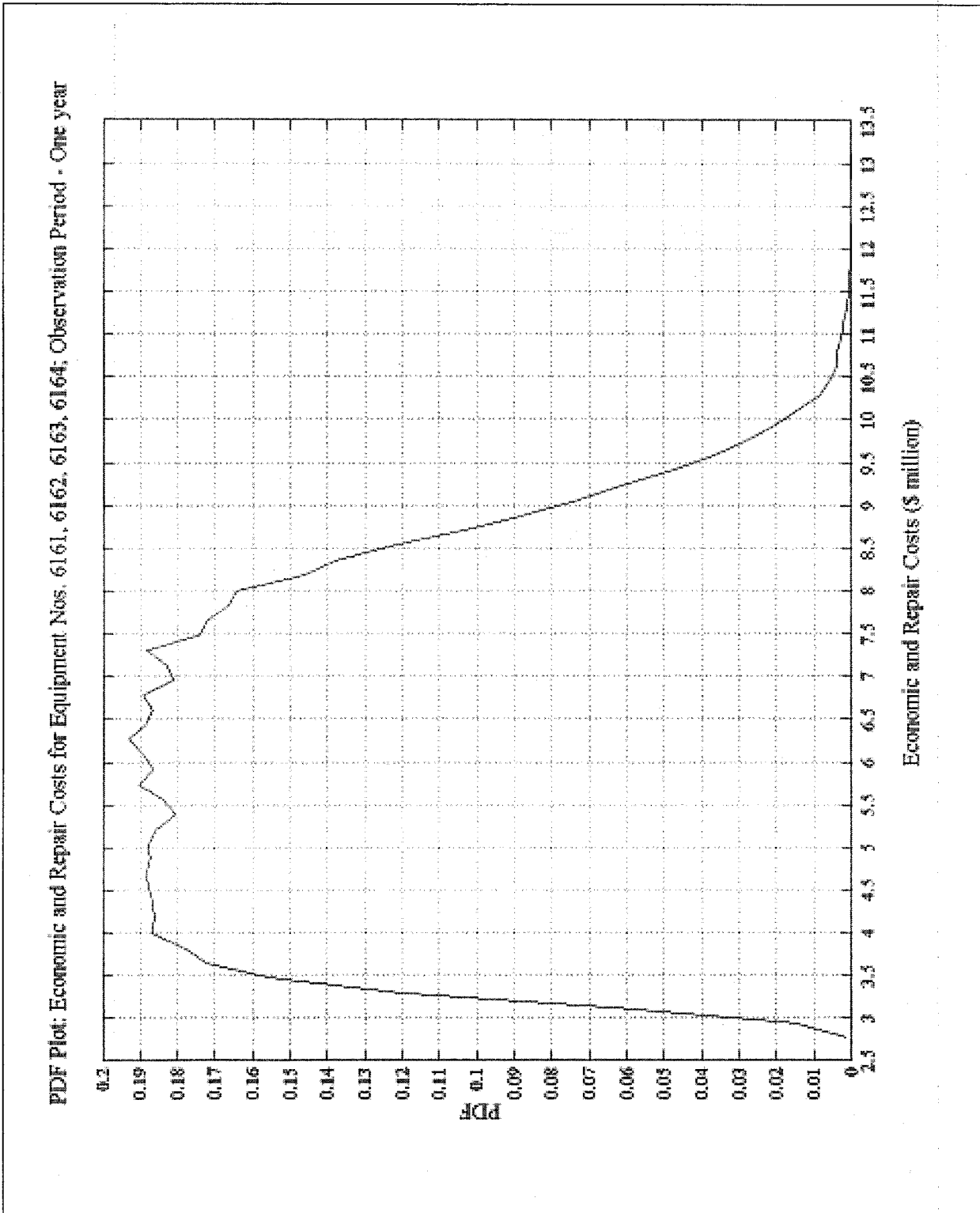


Figure 23: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in one-year.

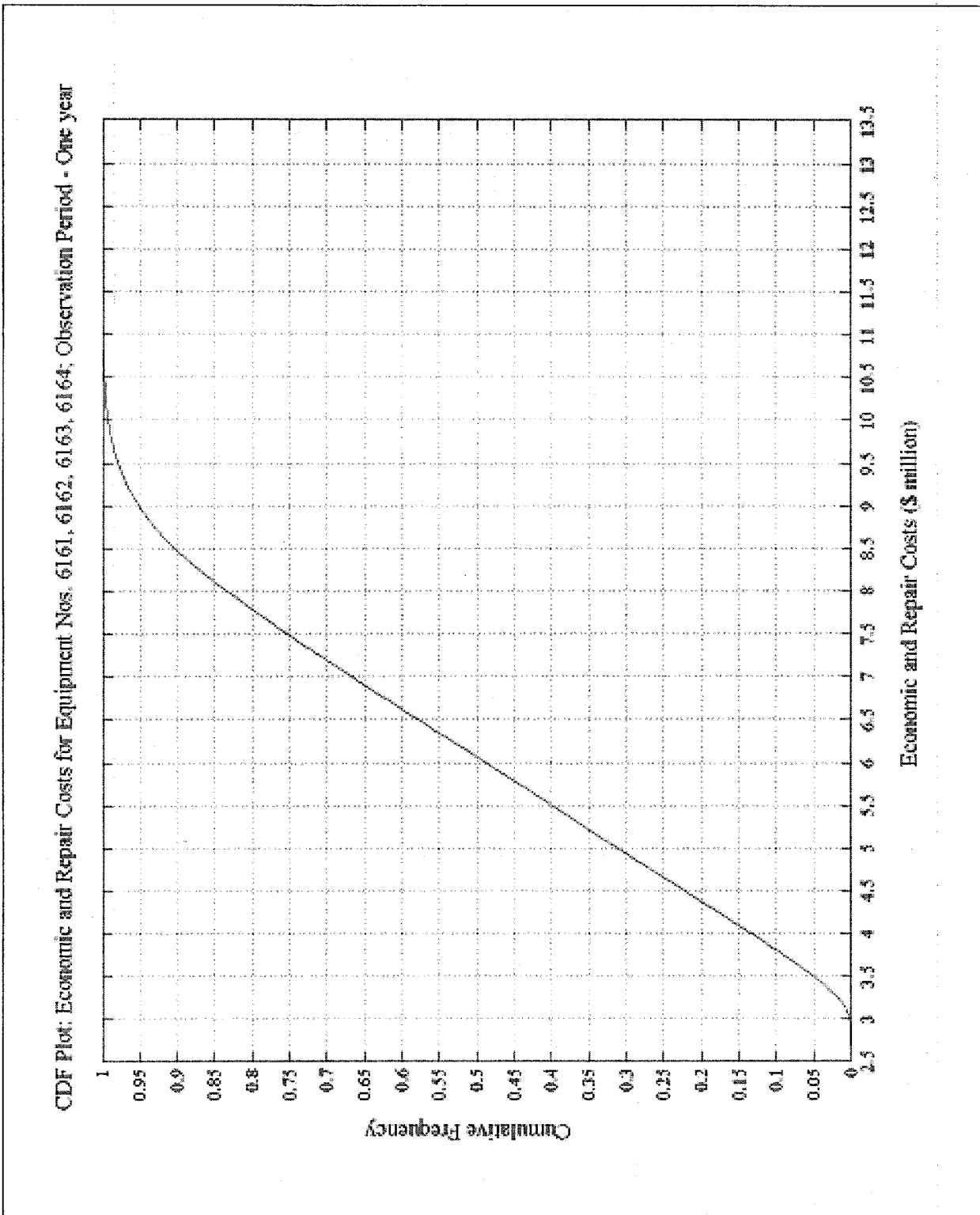


Figure 24: Sampling results - PDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in two years.

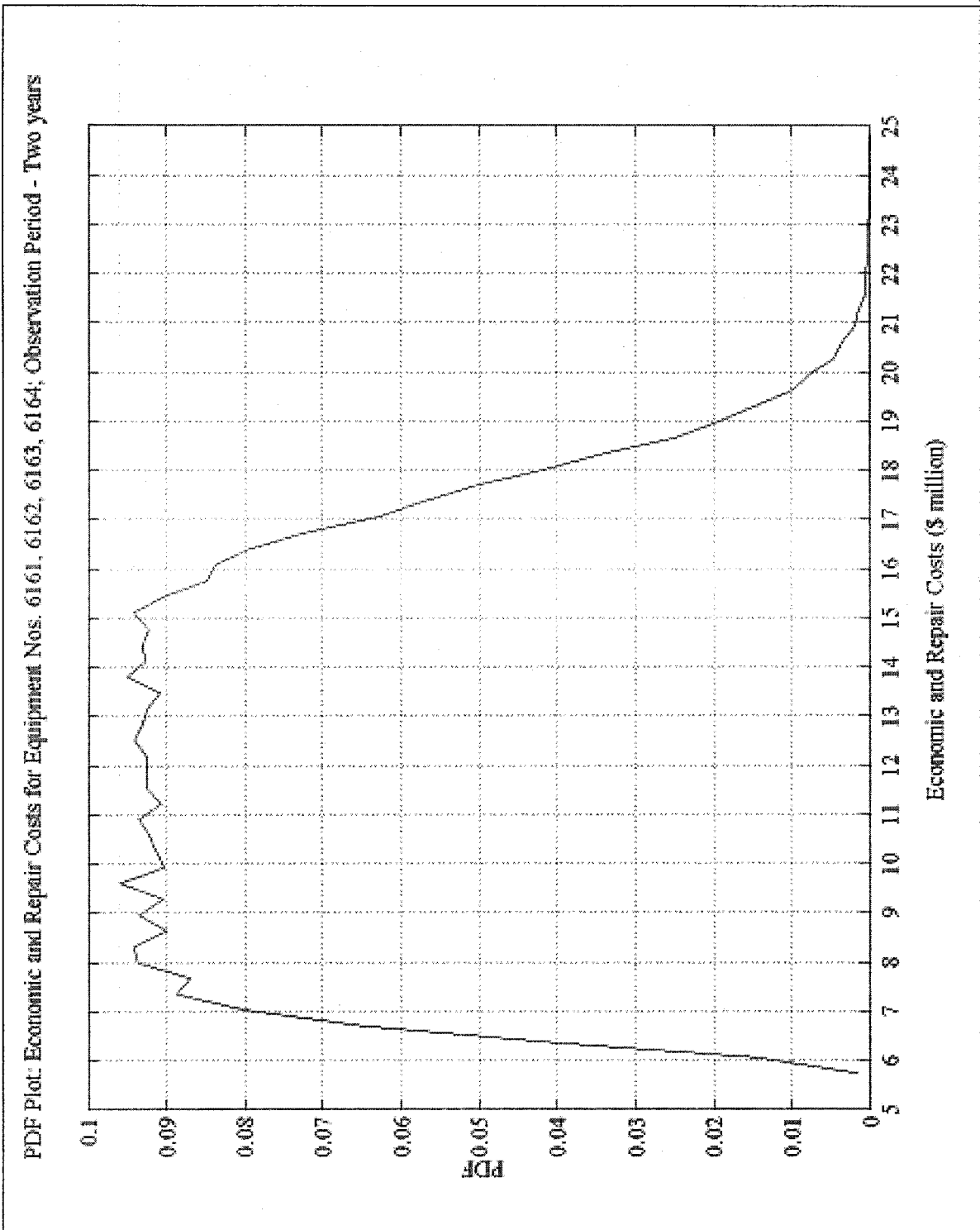
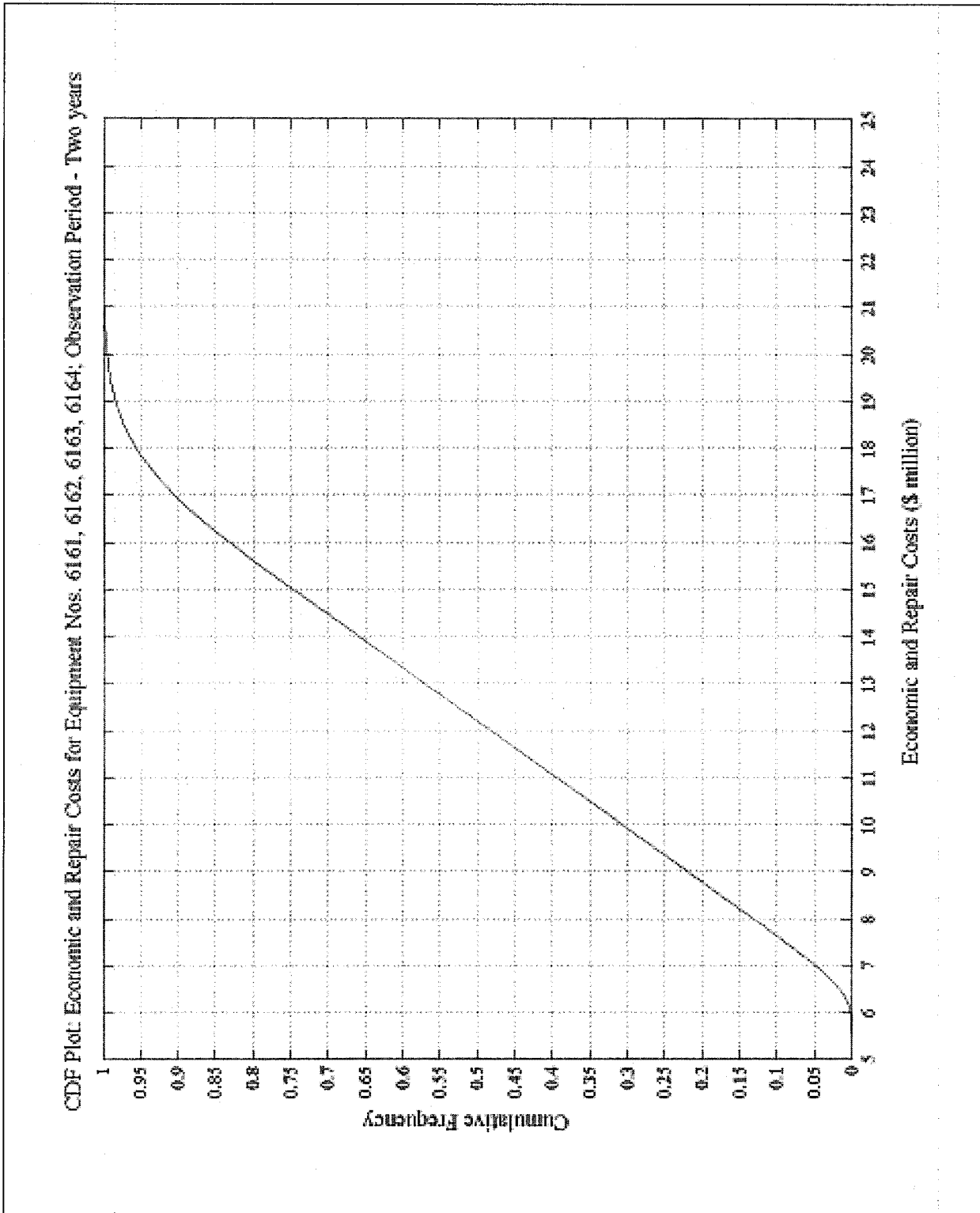


Figure 25: Sampling results - CDF plot for total economic and repair costs (units: \$ million) of the four HEMM equipments in two years.



4 Discussion and Conclusions

This thesis demonstrates a novel application of structural reliability concepts for assessing the production probabilities of ore-hauling mining operation. In particular, the thesis contributes to the state-of-the-art in the field of mining operation productivity assessments as follows

- Data analysis
 - A novel approach for the modeling of equipment states of functioning is presented.
 - Uncertainty in the occurrence rate parameter is addressed by assuming the Poisson random occurrence model and Bayesian Inference. The utilization of Bayesian updated distribution for estimating occurrence rate is an improvement over the traditional technique, wherein, a point-estimate of occurrence rate is considered. In the Bayesian updating process, the distribution of occurrence rate becomes narrower as the number of observations (of occurrences) increase.
- Reliability analysis
 - Utilization of structural reliability concepts - Limit-state-based reliability formulations have been developed and applied for a mining application.
 - The cost and production-time related reliability formulations described in the thesis are simple to evaluate.

References

- Ang, A.H., Tang, W.H. 2002, "Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering", 2nd Edition, Wiley.
- Apeland, S., Scarf, P.A. 2003, "A Fully Subjective Approach to Modeling Inspection Maintenance", *European Journal of Operational Research*, v 148 n 2, p.410-25.
- Benjamin, J.R., Cornell, C.A. 1970, "Probability, Statistics, and Decision for Civil Engineers", McGraw-Hill, New York.
- Christer, A.H., Waller, W.M. 1984, "Reducing Production Downtime Using Delay-time Analysis.", *The Journal of the Operational Research Society*, v 35, p. 499–512.
- Chung, T., Mohamed, Y., and AbouRizk, S. 2004, "Simulation Input Updating Using Bayesian Techniques", *Proceedings of the 2004 Winter Simulation Conference*.
- Department of Defense Guide for Achieving Reliability, Availability, and Maintainability, 2005, Website address: http://www.acq.osd.mil/sse/docs/RAM_Guide_080305.pdf. (Accessed - April 14, 2008)
- Gupta, S., Bhattacharya, J. 1999, "Discrete Markov Chains - An Analytical Tool For Productivity Analysis of Surface Mining Systems", *International Journal of Surface Mining, Reclamation and Environment*, v 13, n 3, p. 111-116.
- Hall, R. A., Daneshmend L. K. 2003, "Reliability and Maintainability Models for Mobile Underground Haulage Equipment", *CIM Bulletin*, v 96, p. 159-165.
- Hall, R. A., Daneshmend L. K. 2003, "Reliability Modelling of Surface Mining Equipment: Data Gathering and Analysis Methodologies", *International Journal of Surface Mining, Reclamation and Environment*, v 17, n 3, p. 139-155.

- Hall, R. A., Daneshmend, L., K. 1998, "Reliability Analysis of Mobile Underground Mining Equipment – A Case Study", Proceedings of the 10th CIM Maintenance/Engineering Conference, Saskatoon.
- Haukaas, T., Kiureghian, A. D. 2006, "Strategies for Finding the Design Point in Non-linear Finite Element Reliability Analysis", Probabilistic Engineering Mechanics, v 21, n 2, p. 133-147.
- Meeker Q. W., Escobar A.L. 2004, "Reliability: The Other Dimension of Quality", Quality Technology & Quality Management, p 1-25 (1).
- Meeker Q. W., Escobar A.L. 2001, "SPLIDA (Splus Life Data Analysis): Software for Reliability Data Analysis and Test Planning", p. 29 (18-21).
- Melchers, R. E. 1999, "Structural Reliability Analysis and Prediction (Civil Engineering)", 2nd Edition, Wiley.
- Priyadarshi, S., Hall, R.A., Streit, L., and Torjusson, G. 2006, "Maintenance and Reliability Analysis of Haul Trucks", Conference Proceeding.
- Procaccia, H., Cordier R., and Muller S. 1997, "Application of Bayesian Statistical Decision Theory for a Maintenance Optimization Problem", Reliability Engineering and System Safety, v 55, n 2, p. 143-149.
- Prokopenko, V. I. 1991, Article translated from "Fiziko-Tekhnicheskie Problemy Razrabotki Poleznykh Iskopaemykh", v 4, p. 85-91.
- Pukite, J., Pukite, P. 1998, "Modeling for Reliability Analysis: Markov Modeling for Reliability, Maintainability, Safety, and Supportability Analyses of Complex Computer Systems", IEEE Press series on engineering of complex computer problems, New York.

- Roxana I., 2005,” Statistical Model for Estimating the Failure Probability in the Field”,
IEEE - RAMS 2005.
- Sturgul, J. R. 2001, “Modeling and Simulation in Mining - Its Time Has Finally Arrived”,
Simulation 2001, v 76, p. 286.
- Vorster, M., Garza, J., 1991, “Consequential Equipment Costs Associated with Lack of
Availability and Downtime”, Journal of Construction Engineering and Management,
v 116, n 4.
- Yuriy, G., Vayenas, N. 2007, “Discrete-event Simulation of Mine Equipment Systems
Combined with a Reliability Assessment Model based on Genetic Algorithms”,
International Journal of Mining, Reclamation and Environment, p. 114.

APPENDIX A

A.1 Random Variable

Random variable represents an event in an analytic form. Unlike a definitive variable, the random variable is defined within a range of possible values. For example, the random variable, $X = x_i$, where $x_i \in (a, b)$ represents an event when X takes on a discrete value, say x_i or a continuum of values within a real sample set, (a, b) . Random variable is denoted with capital case letter, and the values that it can take is denoted with lower case letter.

Random variable having a countable number of possible values is a discrete random variable. For example, if the random variable X represents the sum of two fair dices then x_i can take values ranging from 2 to 12, and the probability of X taking a discrete value 5 is as follows

$$P\{X = 5\} = P\{(1,4), (4,1), (2,3), (3,2)\} = \frac{4}{36} = \frac{1}{9}$$

Random variables can also take a continuum of values. For example, if the time between failures of equipment is considered as a random variable, then the random variable can be defined for a continuum of values say, 1.04 days, 2.2 days, 2.2 days, 45.2 days, and so on, within a predefined sample space $X = (0.1 \text{ days}, 60 \text{ days})$ for instance. Such a random variable is known as continuous random variable. For example, in our case, the time between failures, t , and the mean rate of failure μ_λ have been considered as continuous random variables.

The uncertainty aspect of the random variable X is modeled by a probability distribution, $f(x, \theta)$. The probability distribution is a mathematical function of the values x , that the random variable can take in its sample space and the parameter, θ , that describes the

distribution. For example, the time between failures of equipment t follows an exponential distribution:

$$f(t, \lambda) = \lambda e^{-\lambda t} \quad (\text{A1})$$

The exponential distribution is a function of t the rate parameter λ . The rate parameter can either be a point estimate, or it can be estimated using techniques like Bayesian Inference, as will be described later. If the rate parameter is a point estimate then Eq. (1) becomes a mathematical function of x only. However, in the Bayesian inference, the rate parameter is also treated as a random variable.

The random variables can be assigned probability measures according to the probability law, as discussed below.

A.2 Cumulative Distribution Function

Cumulative Distribution Function (CDF) of random variable X can be defined for any real number x , $-\infty < x < \infty$, as follows:

$$F(x) = P(X \leq x) \forall x \quad (\text{A2})$$

The CDF denotes the probability that the random variable X realizes a value less than or equal to x . Similarly, the probability that the random variable X takes on values which lie in an interval (a, b) .

The random variable X in Eq. (2) can be either discrete or continuous depending upon whether the probability measures are defined only for discrete values of x or for all values of x respectively.

Properties of CDF:

(a). CDF is always non-decreasing, i.e., $\frac{d}{dx} F(x) \geq 0$

(b). $\lim_{x \rightarrow \infty} F(x) = 1$

(c). $\lim_{x \rightarrow -\infty} F(x) = 0$

Figure A1 below illustrates a typical CDF.

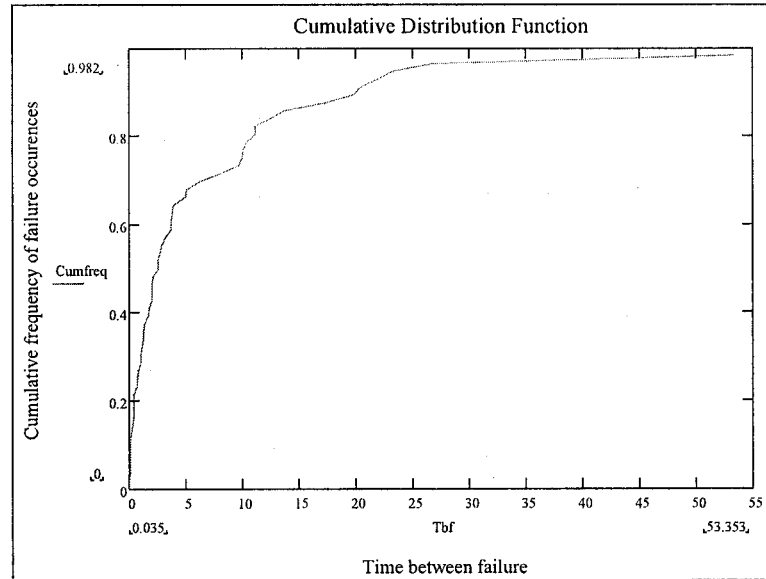


Fig. A1

A.3 Probability Mass Function

The probability distribution for a discrete random variable is described by its Probability Mass Function (PMF). PMF expresses the probability of the random variable taking on each of the discrete values

$$P(X = x_i) = p(x_i) \forall x_i$$

where $i = 1, 2, 3 \dots n$

$$\sum_{i=1}^n p(x_i) = 1 \quad (A4)$$

The CDF for discrete random variables is as follows

$$F(x_j) = \sum_{i=1}^j P(X = x_i) = \sum_{i=1}^j p(x_i) \quad (A5)$$

A.4 Probability Density Function

For continuous random variables, the probabilities are associated with intervals of x . Further, for any specific value of x ; only the probability density is known, i.e., $P(X=x)=0$

We can define the PDF as the gradient of CDF, i.e.,

$$f_x = \frac{dF_x(x)}{dx} \quad (A6)$$

Figure A2 below illustrates a typical PDF.

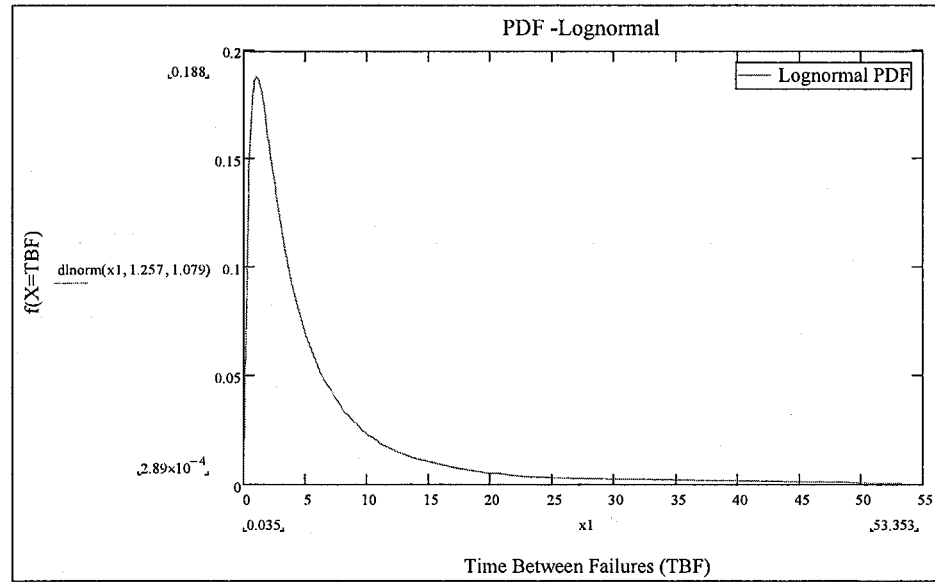


Fig. A2

A.5 Types of Uncertainties

The uncertainties associated with data analysis are of two types: the epistemic (subjective) uncertainty and the aleatory (inherent randomness) uncertainty. These are explained below through examples.

1. Epistemic Uncertainty: Consider that the cost of a new engineering project at its planning stage is to be evaluated. In order to do so, a cost function, which considers the complex interaction of various independent variables, is developed. Some of these variables are listed in the following.

- a). Time for completion – It is influenced by the quality of schedule, level of schedule compliance, management capabilities, etc.
- b). Productivity – is influenced by the timely availability of manpower and equipments; amount of reworks, etc.
- c). External factors such as geo-political and economic environment (interest, foreign exchange and inflation rate regimens, revenue flow profile), third party stakeholders, etc.

There exists uncertainty about the exhaustiveness of the above list. There may still be variables which are not incorporated in the list and which influence the cost function. In such situations, the model may deviate from reality and thus the inferences drawn from the analysis will not be realistic enough. This type of uncertainty is referred to as a model uncertainty. Model uncertainty is an example of epistemic uncertainty, and is reflected by our degree of belief (in the strength or, correctness of the model in this case). Epistemic uncertainty is caused by the lack of knowledge about the factors that characterize a physical system. Epistemic uncertainty can sometimes be reduced by further study, measurements, or

taking an expert opinion. The reduction of epistemic uncertainty helps in coming up with more realistic models.

2. Aleatory Uncertainty: Consider a system comprising a set of equipments, which are operational in an open-pit mining environment. Sudden breakdown of a part or whole of the equipment affects the productivity of the mining operation. The breakdown may be influenced by the occurrence of a number of chanced factors such as the amount of usage of the equipment, quality and schedule of maintenance, operator handling, design quality, weather conditions. The number of observed failures of the equipment within a time interval varies randomly. Similarly, the recurrence time of failure of equipment varies with an inherent randomness. Therefore, the uncertainty associated with the measurement of number of failures in a time interval, or the recurrence time of failure is influenced by chance and is a function of the system. Such uncertainty in the measurement of a variable, which is caused by its inherent randomness, is known as an aleatory uncertainty. This uncertainty cannot be reduced by further measurements, but can be reduced by changing the system.

A.6 Basic Stochastic Processes

The basic stochastic processes are described in the following.

A.6.1 The Bernoulli Sequence:

An engineering process can be modeled by random occurrence or non-occurrence of a particular event. For example, upon inspection, a randomly chosen equipment may or may not be defective, a new construction project may or may not get finished within its scheduled completion time, a fair coin upon flipping may or may not yield a head, etc.

Consider a sequence of trials, at which the event in consideration may or may not occur. The sequence is termed as a Bernoulli sequence, if the following conditions hold:

Only two possibilities in each trial – success or failure;

(i). Trials are statistically independent, i.e., occurrence of each trial is independent of previous trials; and

(ii). The probability of occurrence of the event in each trial is constant, i.e., the occurrence probability is uniform/constant over all trials.

Application of Bernoulli sequence to events occurring in time

The Bernoulli's sequence can be applied to the random events occurring in time, by discretizing time into various intervals (trials) and considering only two possibilities in each interval, i.e., occurrence or non-occurrence of event. Each interval constitutes a trial and the series of finite number of intervals constitute the Bernoulli sequence.

Therefore, if n denotes the number of intervals (each of similar duration, e.g. days/month/year) in consideration, the random variable X represents the number of intervals in which the event occurs, and p denotes the probability of occurrence of the event in each respective interval. The number of occurrences then has the binomial distribution

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad (\text{A7})$$

Mean number of occurrences = $n.p$

Standard deviation of number of occurrences = $n.p.(1-p)$

The corresponding CDF can be written as

$$F_X(x) = \sum_{i=0}^x \binom{n}{i} p^i (1-p)^{n-i} \quad (\text{A8})$$

The number of intervals (trials) until an event occurs for the first time is defined by the geometric distribution. Therefore, if 'N' is the random variable representing the number of intervals before the occurrence of the specific event occurs for the first time, then

$$P(N = n) = p(1 - p)^{n-1} \quad (\text{A9})$$

where $n=1, 2, 3, \dots$

Recurrence time and return period

One important assumption in the Bernoulli sequence is the statistical independence of the trials (time intervals in this case). If it is so then, the time interval till the first occurrence of an event must be same as that of time between any two consecutive events. Referring to Figure A2, the above statement implies that the time t_1 equals time t_2 . Therefore, the probability distribution of the recurrence time is equal to that of the first occurrence time and thus the recurrence time in a Bernoulli sequence is also governed by the geometric sequence.

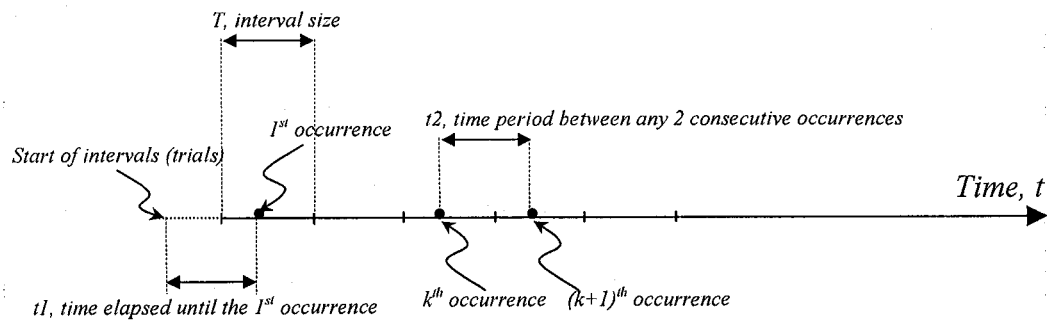


Figure A2 (Illustration of the mean recurrence time in a Bernoulli sequence)

The mean recurrence time (or the return period) between two consecutive occurrences of an event is equal to the reciprocal of the probability of the event within one time interval.

$$\bar{T} = E(T) = \sum_{t=1}^{\infty} t \cdot p \cdot (1-p)^{t-1} = \frac{1}{p} \quad (\text{A10})$$

A.6.2 The Poisson Process - Random occurrence model

In the Poisson process, there is a continuum of opportunity for an event to occur, i.e., theoretically, we can have any number of events (zero to infinity) in any specified interval/period of time, no matter howsoever small. Therefore, when compared to the Bernoulli process, in the Poisson Process, the time intervals have collapse to zero and the number of intervals is increased to infinity.

It can also be demonstrated mathematically that the Bernoulli process approaches the Poisson process as the time interval is decreased.

The underlying assumptions of the Poisson's process are:

- Uniformity: An event can occur randomly at any instant of time, i.e., it is equally likely that the occurrence happens at any instant of time;
- Independence: Any occurrence is independent of what happened before;
- No overlap: Only one occurrence can happen at a particular time.

The probability of observing x occurrences in a time interval T is, according to the Poisson process given by the Poisson probability mass function

$$p(x) = \frac{(\lambda T)^x e^{-\lambda T}}{x!} \quad (\text{A11})$$

T = time period under consideration

x = number of occurrences

λ = mean rate of occurrence (# of occurrences per unit time)

Mean number of occurrences: λT

Variance of the number of occurrences: λT

APPENDIX B

(In CD-ROM)