MASTER

Development of testing times and sample size for inspection operations of a group of identical items

Unel, S.

Award date:
2012

Eindhoven, May 2012

Development of Testing Times and Sample Size for Inspection Operations of a Group of Identical Items

by

Sevgi Unel

BSc Industrial Engineering – Middle East Technical University, 2010
BSc Chemical Engineering – Middle East Technical University, 2009
Student identity number 0757837

in partial fulfillment of the requirements for the degree of

Master of Science
in Operation Management and Logistics

Supervisors:
Dr. T. Tan, TU/e, OPAC
Dr. H. Peng, TU/e, OPAC
Ir. B. Huisman, NedTrain

TUE. School of Industrial Engineering.
Series Master Theses Operations Management and Logistics

Subject headings: condition based maintenance, condition monitoring, inspection planning, sampling

Abstract

This master thesis describes a research project conducted within the Maintenance Development department at NedTrain. A decision support model is developed for determination of inspection sequence and number of parts that are going to be inspected at each inspection. For that purpose, first the definition of problem is given, and the related literature is presented. Then, the model is explained, and solved for the business case. Sensitivity analysis is conducted to verify the model. The thesis concludes with a discussion on main findings, limitations and possible future extensions.

Preface

This Master Thesis document presents the result of my graduation project for the Master of Science program in Operations Management and Logistics at Eindhoven University of Technology. This study was carried out from September 2011 to May 2012 at Maintenance Development Department of NedTrain in Utrecht, Netherlands. I would like to devote this page to thank all people who have contributed to this thesis.

First of all, I would like to thank my first supervisor from TU/e, Dr. Tarkan Tan for his support, helpful comments and advices. His suggestions were very valuable for me and assisted me throughout the project. Furthermore I would like to thank Dr. Hao Peng, my second supervisor from TU/e, for her feedback and opinions. Within NedTrain, I would like to thank Bob Huisman for giving me the opportunity to work on this project in the company. His feedback, extensive knowledge and experience enabled me to include many aspects in the project. Also, I would like to thank my other colleagues in NedTrain for their interest and discussion on all kinds of topics during the coffee-breaks.

I also want to thank to people who are not directly involved in this project. I would really like to thank Serhat Colak, not just for his incredible programming skill but also for being such an interesting person. I would like to thank Mustafa Oztemel for his help in the programming part of the thesis and patience against my silly questions and mistakes. He was very nice and helpful all the time, I cannot imagine any other person who would respond so positively to every call I made asking for help. I would like to thank my roommate Omer Turan, for being the nicest roommate and patient about my whining about everything. I would like to thank Tuba Uluc for always being so kind and helpful to me. While she was living in spaceboxes, my life was a lot easier, since I had the change to knock the door without any notice, drink my perfectly prepared and served Turkish coffee and had a nice conversation about anything. Then she moved out and my life turned to a nightmare. I traveled a lot in the last two years, but I had the best trip with her. That five days we had in Italy and my wonderful statement ‘one way ticket to Portofino’ is memorable. I would like to thank Emre Ersu and Onur Can Saka for being with me in Eindhoven for a year. Maybe I should thank to the people who arranged the double degree program between TU/e and METU, since they are the reason of their short term existence in Eindhoven. If they

had not joined this program, I would not have been able to enjoy my life in our small town at all. The dinners we mostly had at De Lismortel 112, and our fantastic Europe trips are unforgettable for me. I am so happy for knowing both of them and spending a year together. Additionally, I would thank them for their ideas and comments about my thesis.

Now, I would like to mention some other people whom I am in a ‘long distance relationship’. I would like to start with Ece Ozceri, with who I shared the smallest dormitory room for three and a half year. When I met her, she was just a ‘fledgling’, and then I witnessed her ‘evolution’, and hopefully affected it somehow. She was always like a little sister that I have never had. Thanks to her, I was introduced to a new world, which is named as “Beseri Kantini”. During the times when I went to see her to Beseri, I could escape my “boring, analytical, and ideal” engineering world and had a lot of fun. I met two wonderful women in this new world; Ilke Akvarup and Irem Karamanci. Knowing these two ladies expanded my horizon and always inspired me. There was another fantastic woman who joined us with one year delay, Burcin Akin. Our introduction had a very bad timing; I was in my junior year and struggling with ten courses. Probably, in the first days I scared her a bit, but I was so lucky that this did not affect our friendship. Later, she became a little sister for me, too. She even came to Netherlands after me. Besides the times we had in ODTU, our Paris trip, Tilburg, Queen’s Day, Izmir, Selcuk, Kusadasi, Sirince and hopefully our future adventures will be memorable times for me. I will always be grateful to Ece and Burcin; because when I got bored during my long lasting working hours in CS, and took a ‘tea break’, which was usually longer than my working period but whatever, I went back to room 614 and they were always there for me, so we could talk and laugh about anything for hours. The days we spent in 614 will always be unforgettable.

I would like to thank Sevinc Kahya, whom I knew since the first days in ODTU. While I was writing my thesis, simultaneously she was writing hers. As a result of this coincidence, we talked a lot about our studies; she always supported me and made me believe that I could finish it. Besides, I want to thank her for the earlier times we spent in Ankara, especially at Eskiyeni, those days were wonderful. I would like to thank Betul Cavdar for her long distance support. She was always a good friend for me, and she did not deprive her interest and support even I was thousand kilometers away from her. Although she is a very nice person, she did a terrible mistake and chose a life in Ankara. Thus, I have to endure that ugly city for her sake at the end

of my life. I would like to thank Duygu Sengun, who is known as “Sevgi’s sister” for a long time. I am still not sure about whether she was really looking like me, or she started to look like me after I announced her as my little sister. If the latter is true, this means that I have changed her life forever because she resembles me day by day. I really see her as sister and want her to be very happy. I am very grateful that she was always with me during my difficult times while conducting this research. I would like to mention Emine Kurnaz, who was my teammate for Ismail Tosun’s awesome thermodynamics course, member of exam studying committee for separation process course, brain of group for chemical engineering design course, and close friend. I would like to thank her for the free bed and breakfast service she provided every time I came to Ankara, being so nice even at the times when I was unbearable, and supporting me all the time during this study. I hope we stay as friends forever. I would like mention another wonderful woman in my life, Ozge Mercan. We met in prep class in ODTU, but I did not have the chance to know her closely at that time. I am very regretful about those days now. She is one of the most interesting persons I have ever known. It was always be my pleasure to share/discuss/talk about anything/everything with her. All the conversations we had until now have been both amusing and at the same time very inspiring. I think it is a-hard-to-achieve combination. I would like to thank her for the long distance support, even she was quite busy with her MBA courses, she did not reject my calls. Maybe we should not lose our hope; the day that we will solve/understand everything and be happy will come. I would like to mention a man’s name here, Hakan Demir, but I will not thank him, because he did not thank me in his preface.

There is a man in my life for almost eleven years. My worlds are deficient to describe my happiness about knowing him. I do not want to imagine how my life would be without him. We met when we were kids, at IAFL, and as we tolerated each other until know, I believe that we will be friends forever. When I moved to Ankara and he stayed in Istanbul, we had a ‘pause’ in our friendship, but when I came to Netherlands last year, the problem has solved. Thanks to him, I had great times in Europe. Our trips, especially the ones to Sicily and Paris were unforgettable. I am very hopeful and inpatient about our future adventures. I do not want to give details about our long lasting conversations, but I can say that they have always been and will be very inspiring. I would like to thank one of the most important people in my life, Ugur Ipek, for being such a wonderful person, being my ‘grandfather’ forever and everything. There is another man in

my life for eleven years, I feel guilty if I do not mention his name here, Erdem Kombak. When I moved to Ankara after high school to be a ‘successful engineer’, he has chosen to stay in Istanbul to be a ‘successful doctor’. I am not sure whether I have achieved my goal or not, but I am sure that he has achieved his. Due to the distance between us, we experienced a ‘pause’ in our friendship, like the one I experienced with Ugur. However, this did not change anything, because Erdem was and always will be one of the most important people in my life. I would like to thank him for being my friend.

The last, but not the least I would like to express my deepest appreciation to my mom, dad and brother. Without their continuous support and endless love, I would have never been the person that I am now and this work would have never been existed. Finally, I dedicate this master thesis to my grandmother, who could not live long enough to see my graduation, but always loved me more than anything.

Executive Summary

This project has been conducted in Maintenance Development department at NedTrain. NedTrain is a 100% subsidiary of the NS group and has specialized in rolling stock maintenance, servicing, cleaning and overhaul; they maintain railroad passenger cars and locomotives 24/7. NedTrain is one of the first class rolling stock maintenance companies in Europe.

According to Dekker (1996), next to energy cost, maintenance spending is the largest part of the operational budget. Hence, the maintenance concept for capital goods has gained more importance and among the maintenance concepts, Condition Based Maintenance, which supports right-on-time maintenance based on tangible reasons, has become prominent. Condition Based Maintenance is based on the following idea; if the deterioration of the system can be directly measured and if the system is subject to failure only if it deteriorates beyond a given threshold level, then the deterioration level can be monitored and maintenance decisions can be planned according to the actual deterioration of the system. As it can be concluded from this idea, condition monitoring is the main issue in condition based maintenance. While monitoring the condition of an item, it could be possible to detect an incoming failure.

There are two types of condition monitoring techniques; continuous time condition monitoring and discrete time condition monitoring. While in continuous time condition monitoring, the condition of the system is recorded continuously, in discrete time condition monitoring inspections are performed at previously determined discrete inspection times. Hence, these inspection times gain importance for discrete time condition monitoring, since if the intervals between inspection times are very small, the system is observed very frequently, on the other hand, if the intervals between inspection times are very big, then the system is observed very rarely. Hence, finding the proper times for condition monitoring is very important in order to make use of the advantages of condition monitoring and condition based maintenance.

The other issue about condition monitoring is that; it can be performed to one part as well as a group of identical parts. If the latter one is the case, then monitoring operation may be performed to every part, or alternatively a sample may be chosen to represent the population and monitoring operation may be executed to this sample only.

NedTrain is executing condition monitoring operations to oil in gearboxes. They are taking oil samples from each gearbox (mostly) every 3 months. It is believed that sampling every 3 months is too frequent and for the gearboxes that belong to the same type of train a sample which represents the whole population can be found in order to gain some idea about the system.

Hence, the objective of this thesis is *understanding and modeling the degradation process, finding the new optimum inspection times and number of parts to inspect.*

The main research question for the given problem will be;

How does the degradation process behave?

What are the optimal inspection times and sample size that gives the minimum maintenance cost for condition monitoring process?

In order to answer the first research question, data analysis is performed. In the given business case, the selected control parameters to monitor are the amount of polluting materials in oil. In order to monitor the condition of the oil, at each inspection time oil samples are taken from each gearbox and the amount of iron, copper, lead, tin, nickel, chrome, aluminum, and silicon are measured. These materials accumulate in the oil as time passes, and this accumulation process is analyzed. Linear relation, which means that the amount of contaminating material in oil is increased linearly with time, are found to be the better one to represent the process.

Then, in order to answer the second research question, a literature survey is conducted, and the idea developed by Chelbi et al. (1999) is found to be useful for the given case. The optimum inspection sequence, the one which minimizes the expected total cost (sum of inspection cost, failure replacement cost and preventive replacement cost), is found by using the idea suggested by Chelbi et al. (1999) and some changes are made to represent the given situation. If different cost parameters are used, then different inspections sequences are obtained.

$c_I = 10000$, $c_g = 70000$, $c_p = 15000$

The optimum inspection sequences are given below;

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
760	805	834	855	871	885	896	906

The result of this analysis is that, if the failure replacement cost is high, then inspections are done at earlier times.

Then, in order to determine the required sample size for each inspection, a model is developed, and by minimizing the total monitoring cost the optimum numbers of parts that should be inspected is calculated. This total monitoring cost has two components; inspection cost (which equals to unit inspection cost times number of parts inspected), and uncertainty cost. While doing estimation for a population by using information gained from a sample, this estimation will have a confidence interval. The length of this interval shows the power of the estimation and represents how certain one is about the population. If one wants to be surer about the population, he should have a narrow interval, and then he should assign a high cost for the length of the interval. However, if he tolerates some uncertainty, he may have wide interval, and assign a low cost. If the desire for certainty increases, then more parts have to be inspection, in order to provide more information about the population. The results are given below;

	Inspections							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	118	135	230	276	298	332	352	384
Interval Cost	1000	1500	2000	2500	3000	3500	4000	4500

The result support the theoretical idea, in order to be obtain more information about the population, more parts have be inspected.

A user interface is developed in MATLAB to support NedTrain to use the output of the proposed prediction model. Sensitivity analysis and possible future extensions are represented.

Table of Contents

List of Figures

List of Tables

Abbreviations

CBM	Condition Based Maintenance
CM	Condition Monitoring
ppm	Parts per million
E(AC)	Average cost per unit time
E(C)	Average total cost during one cycle
E(T)	Expected cycle duration
E(I)	Expected number of inspections during the cycle
P_g	Probability that replacement is due to failure
P_p	Probability that replacement is preventive
c_i	Inspection cost
c_p	Preventive replacement cost
c_g	Failure replacement cost

1. Introduction

1.0 Introduction

The availability and reliability have been important issues for manufacturing companies and service organizations. As a result the maintenance concepts for capital goods have gained more importance. Condition based maintenance (CBM) has become famous since it is a proactive maintenance strategy and eliminates over maintenance cost. The main issue related to CBM is the condition monitoring. The purpose of condition monitoring is to collect data to make it possible to detect incoming failure of equipment. Hence, the time point when the data is collected gains importance in view of the fact that while too late monitoring can be useless since the equipment can fail before the monitoring operation, too early monitoring will result in vague prediction of incoming failure. In addition to safety issues, condition monitoring operations have a monetary value. If the system is monitored very frequently, it may provide good information about the system but this will cause a high monitoring cost, on the other hand doing less frequent monitoring operations can decrease the monitoring cost but one will have limited information about the system. Moreover, if the condition monitoring operations are executed on a population of identical parts not on a single part, then the sample size that represent the population will be important since more parts will provide better information related to the population and cause a high monitoring cost, on the other hand monitoring the condition of less parts will decreases the cost but result in uncertainty associated with the condition of population. Hence, this master thesis aims to develop a model which determines optimum condition monitoring times and sample size for this operation.

This chapter gives an introduction to this Master thesis. It starts with a description of NedTrain in Section 1.1; describing its business structure, the maintenance operations conducted, services provided and different kinds of service locations. Then, in Section 1.2, the problem definition is given; general idea about condition monitoring, condition monitoring problem, objective, research questions, research scope and research methodology are explained. In Section 1.3, the business case is explained. Finally, the outline of this Master thesis report is presented in Section 1.4.

1.1 Company Information

1.1.1 Company Description

NS is the biggest Dutch railway operator, welcoming over a million people into its trains every year and NedTrain is a subsidiary of NS Group. NS is a state-owned holding company that consists of three branches; the passenger transportation branch, the stations development and exploitation branch and the infrastructure branch. Figure 1 shows an overview of the organization.

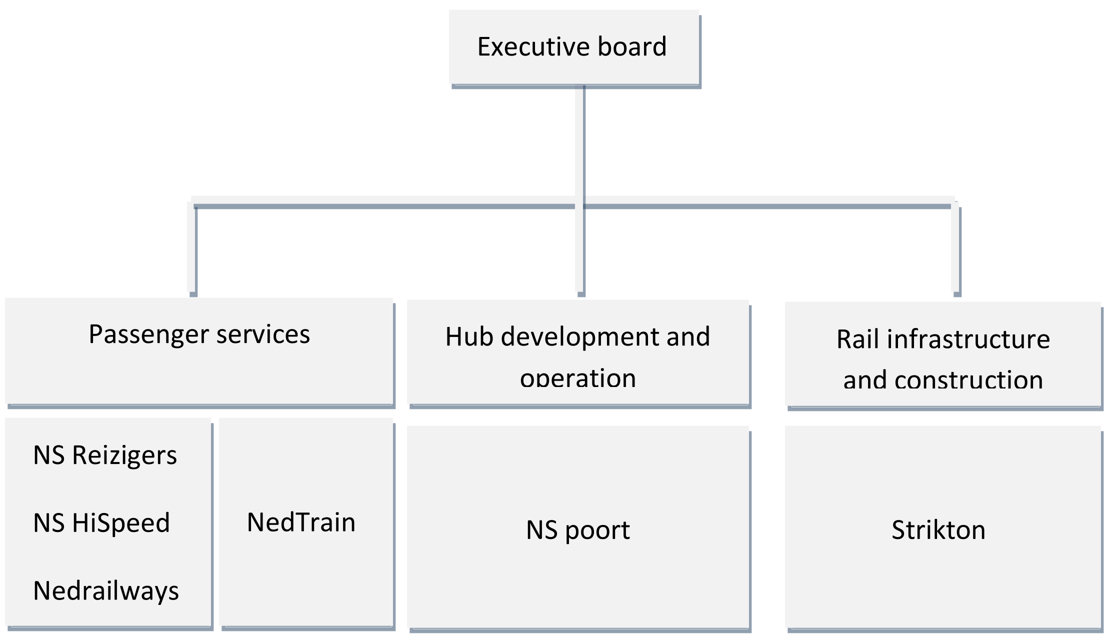


Figure 1: Organizational chart of NS Group

Strukton is a big railway contractor, operating not only in The Netherlands but also in many other European countries. NS Poort runs the exploitation of the train stations and NS Reizigers and NS HiSpeed run the exploitation of the trains.

NedTrain "has specialized in rolling stock maintenance, servicing, cleaning and overhaul, they maintain railroad passenger cars and locomotives 24/7" (www.nedtrain.nl). Although NedTrain is founded officially in the early 90s, its origin extends to 19th century when the first railroad in the Netherlands was founded. In 1938, NS was founded by the merger of the two largest Dutch railway companies; in the early 90s, NS has been privatized and NedTrain has been separated within NS as the company responsible for the maintenance operations of the rolling stock in the

Netherlands. Currently, NedTrain is one of the first class rolling stock maintenance companies in Europe. Some key figures about NedTrain are:

- Number of full time employees: 3400
- Turnover: €475 million/year
- Kilometric performance: 15500 million/year
- Installed base: >2850 coaches (>250000 seats)

The organization of NedTrain is given in Figure 2.

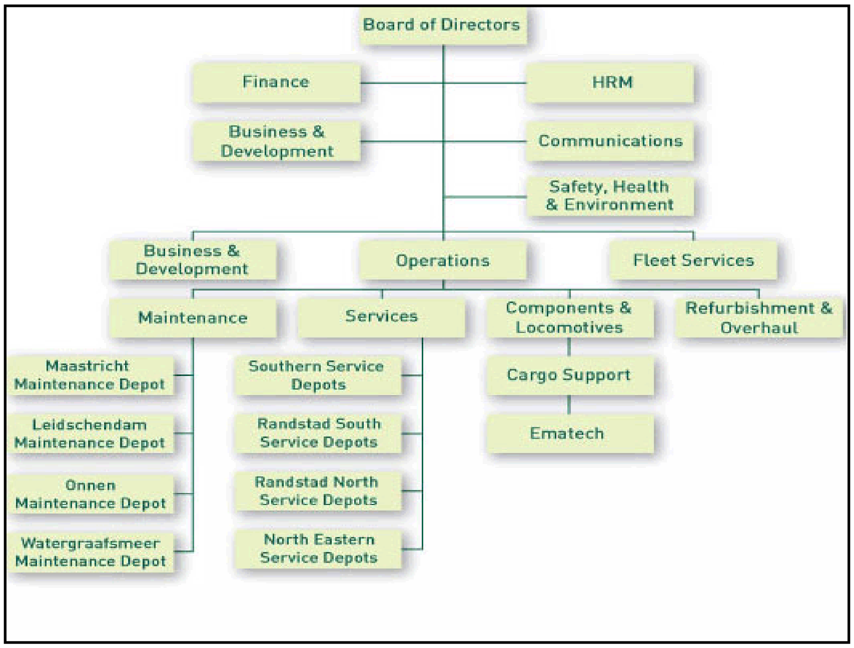


Figure 2: Organization of NedTrain

1.1.2 Current Maintenance Operations

There are three different types of maintenance with respect to parts in NedTrain;

- Used based maintenance: In this type of maintenance activity, a part is replaced after it is used for a specific time period or distance. This type of maintenance is planned preventive maintenance.

- Condition based maintenance: This type of maintenance activity is preventive maintenance with planned inspection and condition replacement, in which a part is replaced after its performance falls below a specific level.

- Failure based maintenance: These are unplanned corrective maintenance activities, in which part is replaced or repaired when it breaks down.

1.1.3 Services and Different Kind of Service Locations

NedTrain is responsible for maintaining, cleaning and revision of rolling stocks. These activities are performed in three different service locations; service depot, maintenance depot and refurbishment and overhaul depot. There are five types of services provided by Nedtrain such as;

- First line service: It is performed on daily basis; rolling stocks are inspected, repaired and cleaned. This is done in one of the service depots. All these locations where all rolling stock stays overnight if it is not in maintenance or used in the night shift. During the night, the rolling stock is cleaned and some small tests are done to check for failure in the system. Small repairs are done in the service depot during the same night. If during the checks problems are detected that are too complicated for the service depot to handle, the rolling stock will be scheduled to visit the maintenance depot.
- Short cyclical periodic maintenance: It is carried out after certain mileage or period of time. During this action, the focus is on checking and cleaning the components. These inspections ensure the safety, quality and reliability of the rolling stock in the line. This service is performed in maintenance depots. There are four different MD located across the Netherlands and different types of rolling stock are repaired in different locations of the MDs. All types of rolling stock are only sent to one, mostly two different locations of the MDs.
- Long term maintenance: During this service, interior upgrades, major alterations to train sets and overhauling or replacement of worn parts are executed. This is service is performed in the refurbishment and overhaul workshop. There are two ROWs in the Netherlands, one in Haarlem and one in Tilburg.
- Failures: If there is failure during the usage of the rolling stock, before acting to it, the criticality of the part is considered. If it is a critical part, then rolling stock is sent to a repair facility and replaced, however if it is a noncritical part, the replacement is done during the next scheduled short cyclical periodic maintenance.
- Damage repair: When a rolling stock is damaged after an accident, NedTrain inspects it and estimates repair costs and then decides whether to repair or discard it.

A schematic representation of the maintenance structure is given in Figure 3.

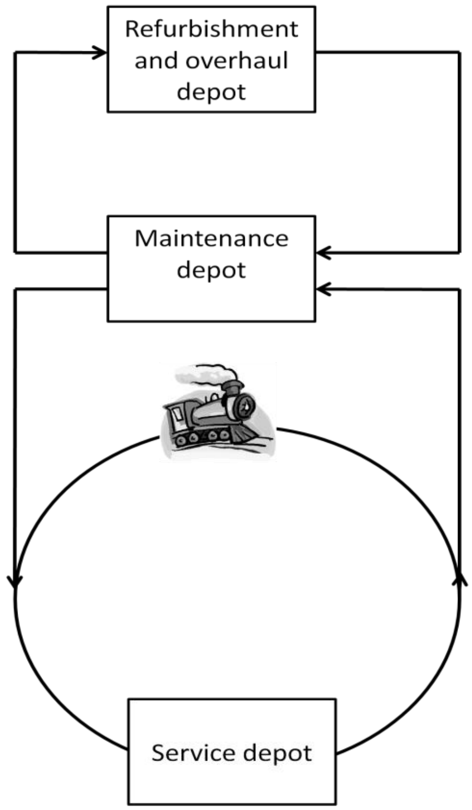


Figure 3: Schematic representation of maintenance process

1.2 Problem Definition

1.2.1 General Idea About Condition Monitoring

Even though, in early 1900s, the maintenance activities were thought to be just some necessary problems to be overcome, these days they are integral part of the business process and they create additional values. According to Dekker (1996), next to energy cost, maintenance spending can be the largest part of the operational budget. Hence these activities have to be planned properly considering both economical and safety issues. For this reason the interest is switched from corrective maintenance to preventive maintenance. Condition based maintenance, which is a preventive maintenance method, is highly used in industry. The main idea behind the condition based maintenance is that; if the deterioration of the system or a control parameter that is strongly correlated with the state of the system can be directly measured and if the system is subject to failure only if it deteriorates beyond a given threshold level, then the deterioration level or the control parameter can be monitored and maintenance decisions can be planned according to the actual deterioration of the system. Thus, condition monitoring is the main issue in condition based maintenance. The purpose of monitoring the condition of an item is to collect data to make it possible to detect incoming failure. There are two types of condition monitoring; continuous time condition monitoring and discrete time condition monitoring. In continuous time condition monitoring, the condition of the system is recorded continuously with the use of sensors, on the other hand in discrete time condition monitoring inspection times should be determined in advance and condition of the system is measured at these specific inspection times. Continuous time condition monitoring gives a better understanding of the system, but it may not be applicable to every system and it is highly costly. On the contrary, discrete time condition monitoring is cheaper compared to continuous time condition monitoring, but it collects information only at discrete times, hence it is limited representing the system.

In order to capture the condition of the system accurately using discrete time condition monitoring, the inspection times have to be established very carefully since they are the specific time points that the condition of the system can only be observed. If the intervals between inspection times are very small, the system is observed very frequently, in fact the discrete time condition monitoring action in this case may approximate to continuous time condition monitoring. However, as mentioned before, this may bring very high monitoring cost and also it may not be very suitable to disturb the process very frequently to monitor the condition. On the other hand, if the intervals between inspection times are very big, the system is observed very rarely. Thus, little information is obtained related to the condition of the system, and the failure will be unexpected in this case. This situation is contradicting with the general idea of preventive maintenance, and can cause problems related to safety issues. Hence, finding the proper times for condition monitoring is very important in order to make use of the advantages of condition monitoring and condition based maintenance.

The condition monitoring operation can be executed for a population of identical parts as well as it is executed for a single part. For instance, a company may have hundreds of identical machines and it may want to monitor the condition of all machines. In this case, monitoring operation may be performed for every part, or alternatively a sample may be chosen to represent the population and monitoring operation may be executed for this sample only, then the conclusion will be applicable to the whole population. There are some issues which have to be considered in this case; monitoring every specific machine will give perfect information about the whole population, but it may cause very high monitoring cost since the same monitoring operation is executed hundreds of times, alternatively observing only a sample can decrease the monitoring cost but it will bring some uncertainty about the condition of the whole population. Hence, if the problem is examining the condition of whole population, correct number of parts that is going to be monitored has to be calculated.

1.2.2 Condition Monitoring Problem

The issues related to condition monitoring problem are given in Figure 4 and explained afterwards.

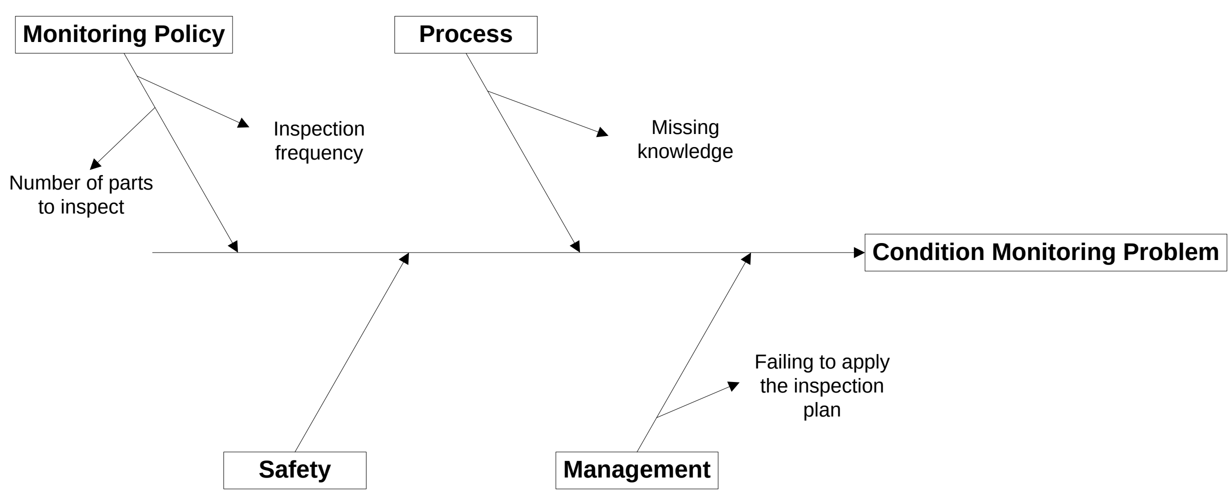


Figure 4: The issues related to condition monitoring

The issues related to condition monitoring problem are categorized under four groups; process, management, monitoring policy and safety. The 'process' is related to missing understanding/knowledge of the excising system. If there is not enough information or a good understanding of the degradation process of the system, it will be difficult to plan the monitoring operations properly. The 'monitoring policy' is related to the frequency of inspection operations and if the problem is related to monitoring the condition of a population the number of parts to inspection. As mentioned before, more frequent inspections will help to monitor the system in a better way but cause a high monitoring cost, and similarly monitoring many parts will increase the understanding of the system herewith a high cost. The ‘safety’ issue should always be considered in condition monitoring since the failure of a system may result in very serious outcomes. Finally, ‘management’ is related to the application of the monitoring plan. The inspection plan should be appropriate for the given system and it has to be applied correctly to gain the best result.

Hence, in order to solve the condition monitoring problem and find the optimum monitoring policy, one has to consider the four issues stated above. First, the system should be analyzed and understood properly. Then the optimum inspection policy should be found which represents the optimum inspection times and number of parts to inspect. Finally, these operations have to be done considering both the safety and management issues.

1.2.3 Objective

The objective of this thesis is that; *understanding and modeling the degradation process, finding the new optimum inspection times and number of parts to inspect.*

1.2.4 Research Questions

The main research question for the given problem will be;

How does the degradation process behave?

What are the optimal inspection times and sample size that gives the minimum maintenance cost (which is sum of inspection cost, preventive maintenance cost and failure cost) for condition monitoring process?

The decision variables will be sample size and inspection times, the parameters (i.e. related cost values, threshold values related to component’s condition level) will be given.

1.2.5 Research Scope

The main research question for the project is finding the correct inspection times and corresponding sample sizes so that the information gained by inspection can be used to understand the condition of the part and then maintenance decision related to entire population can be made. The model generated to solve the given problem will be a general model, it may be applicable to any system if the deterioration of the system or a control parameter that is strongly correlated with the state of the system can be directly measured and if the system is subject to failure only if it deteriorates beyond a given threshold level. Application of the model will be done for gearbox oil sampling for NedTrain and results will be given in Chapter 5.

1.2.6 Research Methodology

This Master of Science project on maintenance management is a Business Problem Solving (BPS) project that aims finding the optimal time and sample size for each inspection operation while minimizing the average total cost. Van Aken et al. (2007) indicate that "Problem solving projects aim at the design of a sound solution and at the realization of performance improvement through planned change." Moreover, they claim that a business problem solving project has to satisfy the following criteria;

Performance focused: the main of the project should be to improve the actual performance

Design oriented: the project steps are controlled by a project plan

Theory-based: theories from the existing literature should be contextualized for the problem, analysis and design activities should be realized

Client centered: the requirements of the company should be identified and taken into account

Justified: the solution should be executed to just justify its convenience

A framework is provided in Figure 5 below, in order to conduct a BPS.

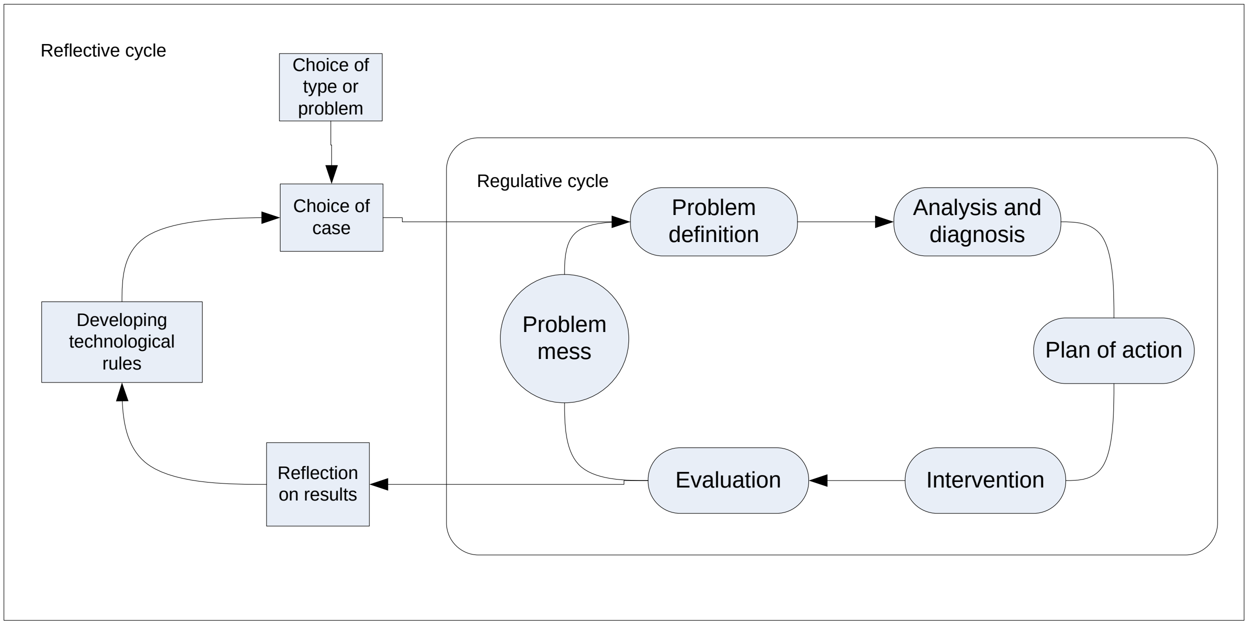


Figure 5: Research design

As can be seen in the Figure 5, there are mainly two cycles: a reflective cycle and a regulative cycle. The regulative cycle is the inner circle element of the reflective cycle. It is the classic problem-solving cycle as indicated by Van Strien (1997).

In this study, the first step is defining the problem, which is, ‘How to minimize the inspection and maintenance costs while keeping the system in a specified condition?’. The second step, analysis and diagnosis, constitutes the analytical part of the project. At this step, the detailed analysis of the maintenance actions and strategies of NedTrain is conducted. During the third step, the solution for the problem is designed. A systematic review of the literature should result in a range of solution concepts to solve the business problem in the ideal scenario. Among the alternative solutions found in the literature, an appropriate one can be chosen and a variant of it can be adapted to the specific problem. The intervention step refers to a change part, in which the redesign is realized through changes in organizational roles and routines, plus the possible implementation of new tools or information systems. In the last step of regulative cycle, the client organization learns to operate within the new system and with the new instruments, and learns to realize the intended performance improvement. After intervention, the next step is to plan a formal evaluation in order to assess whether the intended improvement is achieved or not.

The project will focus on the design part of the regulative cycle, which consists of the first three steps of the cycle.

1.3 Business Case

Railway wheels sit on the rails and the shape and location of wheels and rails on straight track can be seen in Figure 6 below.

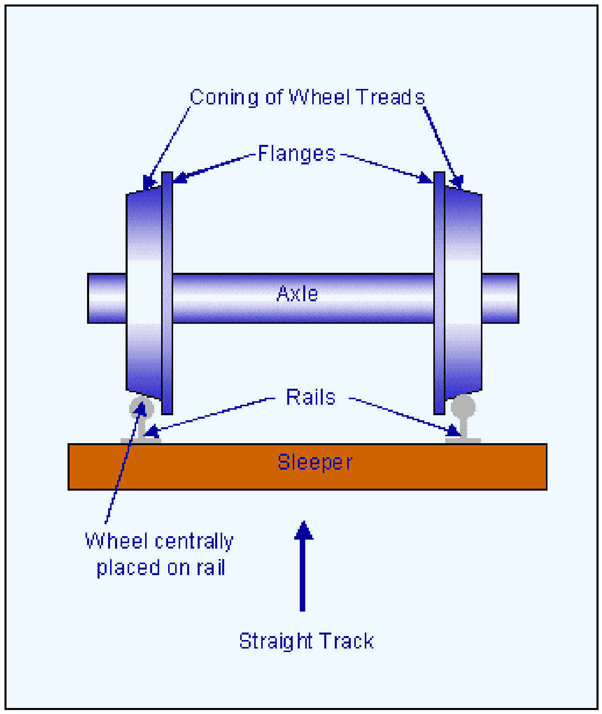


Figure 6: The shape and location of wheels

A pair of train wheels is rigidly fixed to an axle to form a wheelset. Normally, two wheelsets are mounted in a bogie. Most bogies have rigid frames as shown below.

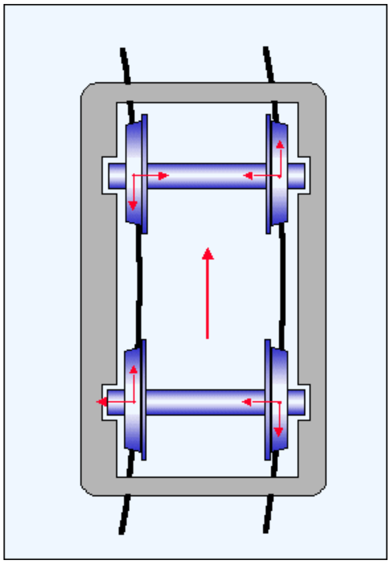


Figure 7: A standard rigid bogie on curved track

The bogie comes in many shapes and sizes. An example is shown in Figure 8, which is a modern design as the motor bogie of an electric locomotive. Here it has to carry the motors, brakes and suspension systems all within a tight envelope.

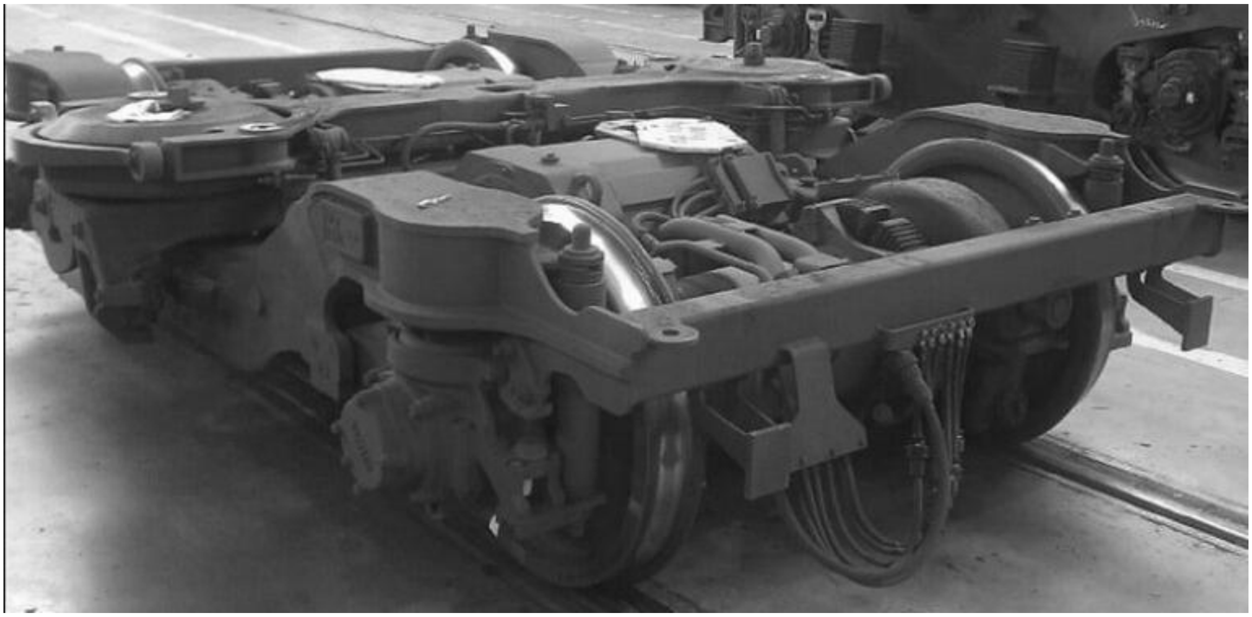


Figure 8: A photograph of a bogie

With all other parts, bogie contains the gearbox which is a device that varies the gear ratio between the engine and the road wheels so that the appropriate level of power can be applied to the wheels. It contains the pinion and gearwheel which connects the drive from the armature to the axle. Figure 9 represents a gearbox

Figure 9: Shape of a gearbox

The efficiency of a gearbox is affected by friction, which can increase the wearing process and cause overheating and premature failure of the gearbox. In order to reduce this friction, usually a fluid is used which can keep two solid parts in the gearbox from touching to each other. In addition to reducing the friction, this lubricant can be used as an indicator of the condition of the gearbox. In order to apply preventive maintenance to a gearbox, the condition of the lubricant can be measured and depending on the results a maintenance strategy can be developed. Moreover, gearbox life is often much better when it is kept clean and well-lubricated.

The current situation in NedTrain for gearbox oil condition monitoring is as follows;

- The company, NedTrain, is taking oil samples from each gearbox (mostly) every 3 months.
- These samples are sent to another company, DeltaRail, for the analysis.
- DeltaRail is doing the experiments and analysis, and then if the oil is contaminated with metal particles or include too much water, they advice NedTrain to change the oil in the gearbox.
- Then, when NedTrain has got the advice from DeltaRail, they do not change the oil immediately, they wait for the next oil sampling/inspection time (which is usually 3

months) and then when the train comes to the maintenance depot for the next time, they change the oil.

It is believed that sampling every 3 months is too frequent and for the gearboxes that belong to the same type of train there is no need to analyze all gearboxes, a sample which represents the whole population can be found in order to gain some idea about the system.

The problem can be stated as follows;

The reliability of the rolling stocks and safety of the passengers are the first concern of NedTrain. By doing inspections, information related to the health of the system can be obtained. The more the inspection is done, the much the information is gained. However, this brings the high inspection cost herewith. On the other hand, decreasing the number of inspections will definitely decrease the inspection cost but the inspection may be done too late and immediate replacement can be required for the system and this will disturb the production plan in the maintenance depot terribly or even in the worst case an accident may happen. In addition to the frequency of inspection, the number of components that are inspected is also important for NedTrain. With the assumption that all components in the system behave in the same way and their conditions are always equal to each other, the information gained from sample inspection is used for deciding for the whole population. Hence, bigger sample size will represent the population in a better way, but it will increase the inspection cost. Then, there should be a trade-off between the inspection cost and rolling stock reliability and whole system smoothness. The optimal inspection times and sample size for each inspection have to be determined.

1.4 Report Outline

The rest of the report is organized as follows; in Chapter 2, literature about maintenance operations, condition based maintenance, condition monitoring and inspection planning is clarified. In Chapter 3, description of the model used to solve the problem is presented. In Chapter 4, the way how the historical data analyzed and the result of this analysis is represented. In Chapter 5, the results of the business case are given which is conducted by using the model given in Chapter 3. Finally, in Chapter 6, conclusion and recommendation are presented.

2. Literature Review

This Chapter gives the literature related to the given problem. It starts with the general information about maintenance in Section 2.1. Then in Section 2.2 condition based maintenance is explained. In Section 2.3 information about condition monitoring is represented. Finally, in Section 2.4 the literature about inspection planning is given.

2.1 General Information About Maintenance

Maintenance, repair and operations is defined by European Federation of Natural Maintenance Societies as follows; "All actions which have the objective of retaining or restoring an item in or to a state in which it can perform its required function. The actions include the combination of all technical and corresponding administrative, managerial, and supervision actions." According to Parida et al. (2006), in early 1900s, maintenance activities were necessary evil, there was no way to gain value from them. Between 1950 and 2000, maintenance was important support function; it can be planned and controlled. Then, in these days it is integral part of the business process, it creates additional value. In other words, as it is stated by Wang et al. (2007), for the first generation of maintenance activities the idea was 'fix it when it broke'. Then, with the invention of computers the idea shifted to 'scheduling and planning maintenance activities' for the second generation. Finally, in the third generation the ideas like 'condition monitoring, design for reliability and maintainability, failure modes and effect analyses are generated.

According to Dekker (1996), next to energy cost, maintenance spending can be the largest part of the operational budget. Zhao et al. (2010) state that the annual cost of maintenance goes up to 15% for manufacturing companies, 20%–30% for chemical industries, and 40% for iron and steel industries. Moreover, Bengtsson (2007) states that as much as one third of the total maintenance cost is spent unnecessarily because of circumstances such as bad planning, overtime costs, poor usage of work order systems and limited or misuse of preventive maintenance. Therefore, importance of maintenance increases significantly and there is a continuous search for a better maintenance policy which provides economic efficiency with higher system reliability, availability and safety.

Maintenance models may be divided into two distinct categories; corrective maintenance and preventive maintenance.

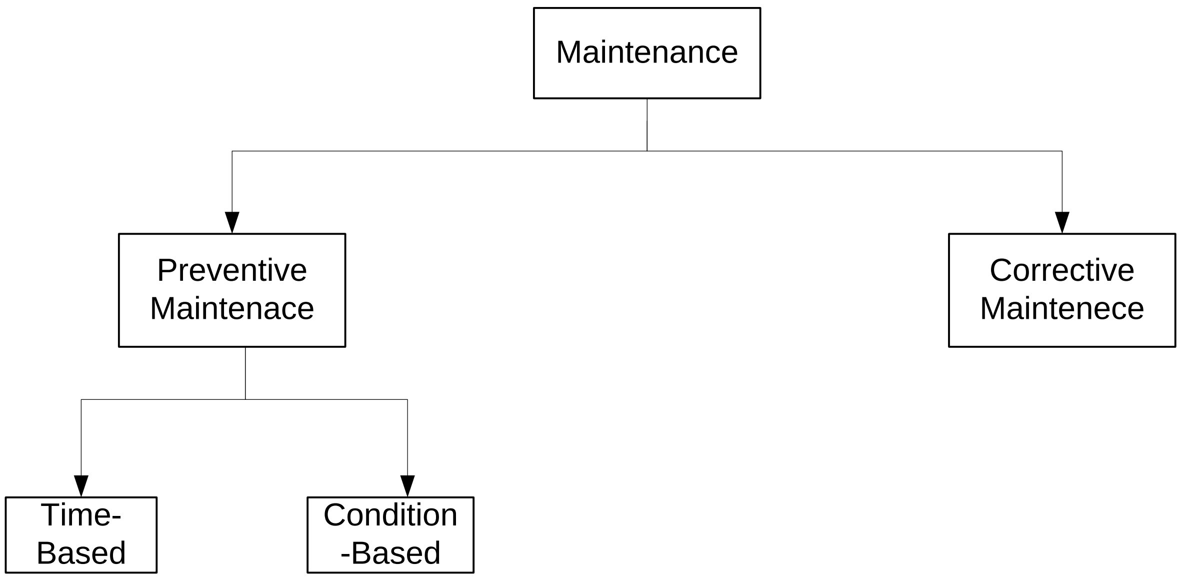


Figure 10: Maintenance Techniques

Corrective maintenance, which is similar to repair work, is undertaken after a breakdown or when obvious failure has been occurred. In Bengtsson (2007), corrective maintenance is defined as: "maintenance carried out after fault recognition and intended to put an item into a state in which it can perform a required function". On the other hand, preventive maintenance is defined as: "Maintenance carried out at predetermined intervals or according to prescribed criteria and intended to reduce the probability of failure or the degradation of the functioning of an item." Moreover, preventive maintenance is divided into two types, time based maintenance (predetermined maintenance) and condition based maintenance. Time based maintenance is scheduled and planned without the occurrence of any monitoring activities and this scheduling can be based on the number of hours in use, the number of times an item has been used, the number of kilometers the items has been used, and so on. It is best suited to an item that has a visible age or wear-out characteristic and where maintenance tasks can be made at a time that for sure will prevent a failure from occurring. Unlike time based maintenance, condition based maintenance, does not utilize predetermined intervals and schedules. Instead, it monitors the condition of items in order to decide on a dynamic preventive schedule.

According to Wang et al. (2007), corrective maintenance is considered as a feasible strategy in the cases where profit margins are large. For the first category of preventive maintenance, which is time based maintenance, it is often difficult to define the most effective maintenance intervals because of lacking sufficient historical data. In many cases when time based maintenance strategies are used, most machines are maintained with a significant amount of useful life remaining. This often leads to unnecessary maintenance, even deterioration of machines if incorrect maintenance is implemented. In the second category of preventive maintenance, which is condition based maintenance, limitations and deficiency in data coverage and quality reduce the effectiveness and accuracy of the strategy.

A major challenge nowadays for maintenance department is not only to know the new techniques can do, but also to choose the proper one for their organizations. According to Wang et al. (2007) the selection of maintenance strategies is a typical multiple criteria decision-making (MCDM) problem. Several MCDM methods have been developed, such as the weighted-sum model (WSM), the weighted-product model (WPM), the TOPSIS method and the AHP. The AHP is one of the most popular MCDM methods.

2.2 Condition Based Maintenance

Condition based maintenance (CBM) is defined by Bengtsson (2007) as “preventive maintenance based on performance and/or parameter monitoring and the subsequent actions.” It is a maintenance type that utilizes on-condition tasks in order to monitor the condition of the system over time and usage. This is done in order to give input to decide maintenance actions dynamically. According to Chen et al. (2001), for CBM the action taken after each inspection is dependent on the state of the system. It could be no action, or minimal maintenance to recover the system to the previous stage of degradation or major maintenance to bring the system to as good as new state. According to Bengtsson (2007), condition based maintenance is performed to serve the following two purposes;

1) to determine if a problem exists in the monitored item, how serious it is, and how long the item can be run before failure

2) to detect and identify specific components in the items that are degrading and diagnose the problem.

Bengtsson (2007) states that the need for condition based maintenance was revealed as early as in the 1960s through a study performed during the development of the preventive maintenance program for the Boeing 747. The study's purpose was to determine the failure characteristics of aircraft components. It was found that a relatively small part of all components (11%) had clear ageing characteristics, which enables a schedule overhaul (that is predetermined maintenance). The rest of the components (89%) did not show such ageing characteristics (that is, they were more or less random failures) and consequently not applicable to schedule overhauls. Page (2002) presents similar conditional-probability curves within the manufacturing industry. He states that only 30% of all components have clear ageing characteristics, and that this percentage decreases as complexity and technology increases. Evidently, the ageing feature of a component is not the best approach, and in some applications not even possible, when planning appropriate maintenance schedules. This fact introduces condition based maintenance and condition monitoring as one solution to the issue.

Then, according to Grall et al. (2002), if the deterioration of the system or a control parameter strongly correlated with the state of the system can be directly measured (through vibration analysis, cumulative wear monitoring, corrosion/erosion level etc.) and if the system is subject to failure only if it deteriorates beyond a given threshold level, it is more appropriate to base the maintenance decision on the actual deterioration state of the system rather than on its age. This leads to the choice of a CBM policy. It is also proved that CBM policy minimizes the cost of maintenance, improve operational safety and reduce the quantity and severity of in-service system failure.

According to Neves et al. (2011) in order to apply condition based maintenance, first of all the parameters which are going to be monitored have to be selected. Then the inspection dates has to be decided, these dates can be either fixed or they can be state dependent. Finally, the warning limit (critical threshold) has to be established.

2.3 Condition Monitoring

According to Bengtsson (2007) condition monitoring is the main issue in CBM. The interval for condition monitoring can be either continuous or periodic. The purpose of monitoring the condition of an item is to collect data to make it possible to detect incipient failure. With this information, maintenance tasks can be planned at a proper time. Another purpose of condition monitoring is to increase the knowledge of failure cause and effect and deterioration pattern. Tsang (1995) states that condition monitoring tries to serve for the following purposes; determine if a problem exists in equipment, how serious the problem is, and how long the equipment can run before failure. Condition monitoring techniques can be classified according to the type of symptoms they are designed to detect: a) dynamic effects such as vibrations and noise levels, particle released into the environment, chemicals released into the environment, b) physical effects such as cracks, fractures, wear and deformation, temperature rise in the equipment, electrical effects such as resistance, conductivity, dielectric strength etc. Moreover, it is stated in Christer et al. (1995) that it is convenient to classify information gained from condition monitoring into two classes, namely direct information and indirect information. In direct information, by condition monitoring the variable which directly determines condition/failure of component is observed and direct information is gained; for instance the thickness of a brake pad or the wear in a bearing. Indirect information provides information which is influenced by the component condition, but is not a direct measure of the failure process; for example an oil analysis or a vibration frequency analysis. Even if the method is different in both cases, the point of concern is to predict the conditional failure time distribution as input to modeling maintenance practice.

As it is mentioned before maintenance issues are receiving greater attention from researchers and engineers due to the fact that equipments are becoming increasingly sophisticated and play a growing role in modern industrial life. According to Grall et al. (1997), in order to limit the failures, inspections and renewals of critical components should be done regularly. However, since both manpower and components are becoming expensive, a trade-off must be found to minimize the total cost incurred by inspections, the renewals and the failures of the system. Wang (2003) mentions that among the decisions made in relation to condition monitoring, one decision is to choose a condition monitoring interval. More frequent monitoring checks cost money, on the other hand a longer monitoring interval may save the monitoring cost but may also increase the risk of a failure between the monitoring checks. There is clearly an optimization problem to balance the trade-off between more and fewer monitoring checks.

2.4 Inspection Planning

Inspection times must be chosen so that undetected failure costs and inspection costs are balanced optimally. A common used inspection policy with constant intervals between inspections is not optimal relative to a certain cost model; optimum inspection policies have decreasing inspection intervals for ageing systems (Leung, 2001).

Interest in optimal inspection schedules for the maintenance of stochastically failing or deteriorating systems originated with the basic model presented by Barlow et al. (1963). They developed a simple model capturing the tradeoff involved in the choice of optimum inspection schedule, frequent check increases the cost of inspections but decreases the cost of late detection of failure. However, it may be difficult to find such optimum policy. To avoid this difficulty, Munford and Shahani (1972) suggested a nearly optimal checking interval. They presented the nearly optimal inspection policy by assuming that the conditional probability, which is the probability of the failure occurrence between successive inspections is constant. Then, this method was applied to a case with Weibull distribution in Munford and Shahani (1973). Tadikamall (1979) discussed a case with Gamma distribution using the same method. Keller (1974) made the problem tractable by supposing that checking is so frequent that it can be described by a continuous density n(t) of checks per unit time. Kaio and Osaki (1984) developed Keller's method using the smooth density and obtained the more exact inspection policy. Leung (2011) developed four optimum inspection policies using inspection density idea. Nakagawa et al. (1980) proposed an approximate calculation of optimal checking procedures which computes successive check times backwards supposing that an appropriate check time is previously given after a large number of checks.

Dieulle et al. (2003) and Dieulle et al. (2001) suggested determining the inter-inspective durations by the help of an inspection scheduling function. Wang (2000) developed a model to determine the optimal critical level and monitoring interval for the criterion of interest. Random coefficient growth model, which is established by Lu and Meeker (1993), is used to describe the deterioration process of the monitored item. A simple model developed by Christer et al. (1995) allows the future wear pattern to be dependent upon the history of wear and minimizes the expected cost per unit time over the time interval between the current inspection and the next inspection. Bérenguer et al. (2003) proposed a multithreshold policy to choose sequentially the best maintenance actions and to schedule the future inspections using the online monitoring information. Bahoe (2002) developed a periodic inspection strategy for three modes system, normal, abnormal and failure. Grall et al. (1997) modeled the problem by a semi-Markov decision process which minimizes the long run average cost incurred by inspections, preventive maintenance and unexpected failures. Ohnishi et al. (1986) obtained an optimal inspection and replacement policy minimizing the expected total discounted cost over infinite horizon for a discrete time Markovian deterioration process. Chelbi et al. (1998) presented two inspection strategies; the first one was a simple inspection strategy which is an extension of the one developed by Munford and Shahani (1972), the second one was a predictive oriented one based in the systematic control of the equipment. Chelbi et al. (1999), proposed a mathematical model and a numerical algorithm to generate an optimal inspection sequence, which is defined as the inspection sequence minimizing the expected total cost per unit time over an infinite span. This model was found to represent the situation in NedTrain properly, and the numerical algorithm was applicable for a practical case. Detailed information about this model is given in Chapter 3.

3. Model Description

This chapter has two sections. In the first section, the description of the model that can be used to determine the inspection times is presented. Then, in the second section, the idea than can be used to determine the sample size is given.

3.1 Inspection Time Determination Model

Case 1: When failure is not self-announcing

As mention in Chapter 2, Chelbi et al. (1999) defined the optimal inspection sequence as the one minimizing the expected total cost and proposed a mathematical model and a numerical algorithm to generate this optimal inspection sequence which is applicable for a practical case.

The condition of the system is monitored through inspections, at each inspection indirect information is obtained by a measure of selected control parameters which are strongly correlated with the state of the system which is subject to failure only if it deteriorates beyond a given threshold level. Two different threshold levels are defined. The first one is the 'alarm threshold' which represents that the condition of the equipment is not good enough to work and a replacement has to be scheduled, this situation can be seen as a 'soft failure' which is not a real failure but it is indicates that the equipment is not at a good condition anymore. The second threshold level is the 'failure threshold' which represents that it is not possible to use the equipment anymore; it can be seen as a 'hard failure'.

It is assumed that inspections do not affect the physical process of deterioration. If the i^{th} inspection performed at instant x_i reveals that the equipment has failed (failure can only captured by inspection), then it is immediately replaced by a new identical one. If it has not failed but the control parameter is found to have exceeded the alarm threshold level, a preventive replacement is scheduled at time $x_i + H$. The time interval H depends on the logistic and any previously scheduled inspections within this period H are cancelled. This is shown in below.

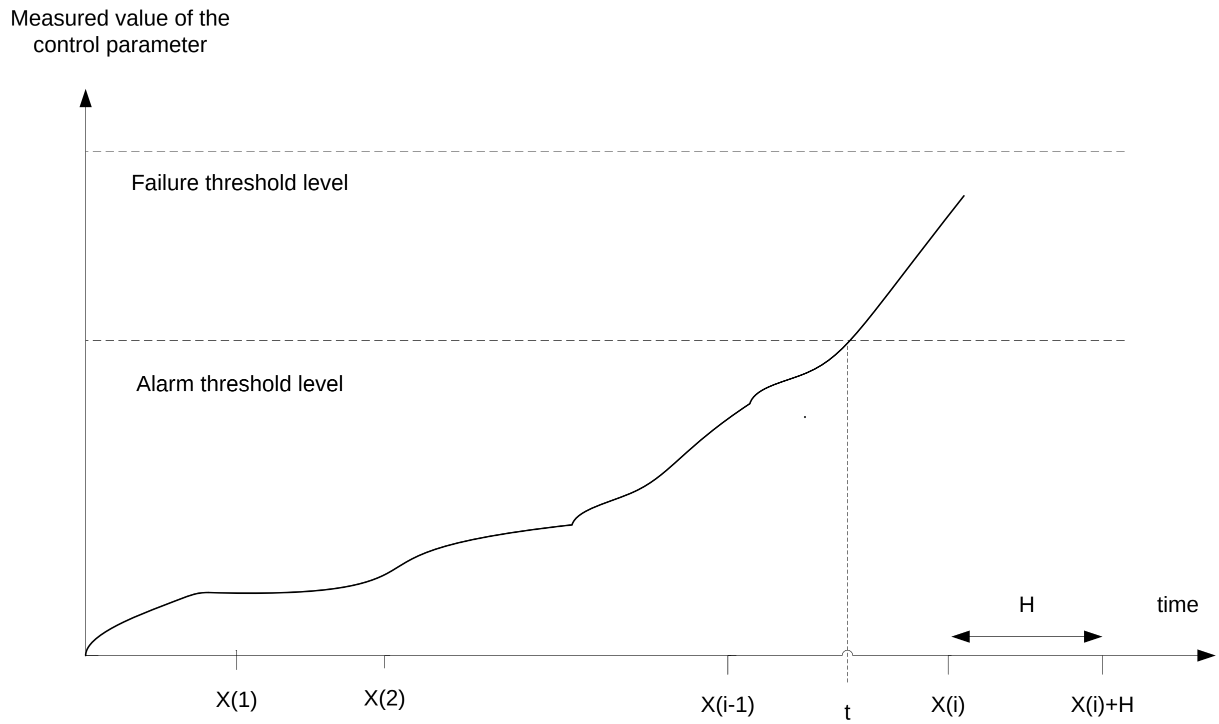


Figure 11: Control parameter evolution with time

In order to take into account this physical behavior of the equipment, two probability density functions Φ(.) and f(.) are considered. $\phi(.)$ is associated with a random variable τ which corresponds to the time at which the alarm threshold is exceeded, whereas f(.) is associated with the random variable t which represents the failure time. Their corresponding cumulative distribution functions are Φ(.) and F(.). They are assumed to be independent.

The best inspection strategy is defined as the inspection sequence which minimizes the average total cost per time unit over an infinite horizon. The cost is represented as;

$$E(AC) = \lim_{T\to\infty} \frac{\{E[C(T)]\}}{T} \rightarrow E(AC) = \frac{E(C)}{E(T)} \tag{3.1}$$

where

$E(C) = average\ total\ cost\ during\ one\ cycle$

$E(T) = expected\ cycle\ duration$

The expected total cost, E(C), is representes as;

$$E(C) = c_i E(I) + c_g P_g + c_p P_p \tag{3.2}$$

where

$c_i = inspection\ cost$

$c_p = preventive\ replacement\ cost$

$c_g = failure\ replacement\ cost$

$E(I) = expected\ number\ of\ inspections\ during\ the\ cycle$

$P_g = probability\ that\ replacement\ is\ due\ to\ failure$

$P_p = probability\ of\ having\ to\ replace\ the\ equipment\ preventively\ at\ the\ end\ of\ cycle$

The expected number of inspections in one cycle E(I);

$$E(I) = \sum_{i=1}^{\infty} ih(i) \tag{3.3}$$

where h(i) is the probabilty that i inspections are performed before replacement

$$h(i) = [1 - \Phi(x_{i-1})] \int_{x_{i-1}}^{x_i} \phi(\tau)\big[1 - \big(F(x_i) - F(\tau)\big)\big]d\tau + [1 - \Phi(x_i)] \int_{x_i}^{x_{i+1}} [f(\tau)]d\tau \tag{3.4}$$

For $h(1)$ the formulation is given below as an example;

$$h(1) = \int_0^{x_1} \phi(\tau)\big[1 - \big(F(x_1) - F(\tau)\big)\big]d\tau + [1 - \Phi(x_1)] \int_{x_1}^{x_2} [f(\tau)]d\tau \tag{3.5}$$

First term is to indicate that alarm threshold is exceeded at time τ which is before the first inspection and failure threshold is not exceeded until first inspection. If this is the case, at the first inspection this situation will be observed, a replacement will be scheduled at time $x_1 + H$, and the cycle will end experiencing only one inspection. Second term is to indicate that alarm threshold is not exceeded until first inspection and failure is occurred until the second inspection. If this is the case, at the first inspection, the equipment will be in a good condition, but it will fail before the second inspection, at the second inspection time that situation will be observed but since the equipment is in the failed state, this observation will not be count as an inspection, and the cycle will end with only one inspection.

The probability P_g that the cycle ends with a failure evet;

$$P_g = \sum_{i=1}^{\infty}[1-\Phi(x_{i-1})]\left[\int_{x_{i-1}}^{x_i} F(\tau)d\tau + \int_{x_{i-1}}^{x_i} \phi(\tau)[1-\big(F(x_i)-F(\tau)\big)][F(x_i+H)-F(x_i)]d\tau\right]$$

First term is to indicate that alarm threshold is not exceed until x_{i-1}, and then second term is to indicate that failure will occure between x_{i-1} and x_i, so at the time x_i cycle will end with a failure, and the last term is to indicate that alarm threshold is exceed at time τ which is between two consequtive inspections but failure is not observed, so a replacement is scheduled after H points of time, then failure is occurred between ith inspection and scheduled replacement time, so the cycle will finish with a failure.

The probability P_p that the cycle ends with preventive replacement;

$$P_p = 1 - P_g \tag{3.7}$$

Finally, the cycle average duration, E(T) can be found considering the two following situations;

1. Failure occurs within the time interval [0, x_i], in which case the cycle average duration would be;

$$\sum_{i=1}^{\infty}\left[x_i \int_{x_{i-1}}^{x_i} F(\tau)d\tau\right] \tag{3.8}$$

2. The equipment does not fail within the time interval [0, x_i], but the alarm threshold is exceeded, replacement is performed at time x_i + H. The cycle average duration would be;

$$\sum_{i=1}^{\infty}\left[(H+x_i)\int_{x_{i-1}}^{x_i} \phi(\tau)[1-(F(x_i)-F(\tau))]d\tau\right] \tag{3.9}$$

Hence, E(T) is given by;

$$E(T) = \sum_{i=1}^{\infty} \left[x_i \int_{x_{i-1}}^{x_i} F(\tau) d\tau + (H + x_i) \int_{x_{i-1}}^{x_i} \phi(\tau)[1 - (F(x_i) - F(\tau))] d\tau \right] \tag{3.10}$$

A conditional probability is defined such as;

p_i
$= conditional\ probability\ that\ the\ alarm\ threshold\ is\ exceed\ within\ the\ time\ interval\ [x_{i-1}, x_i]$

The expression for p_i;

$$p_i = \frac{\Phi(x_i) - \Phi(x_{i-1})}{1 - \Phi(x_{i-1})} \tag{3.11}$$

Given the value of p_i, one can obtain the inspection times x_i, using the following recursive formula;

$$x_i = \Phi^{-1}\{p_i + (1 - p_i)\Phi(x_{i-1})\} \tag{3.12}$$

with $x_0 = 0,\ \Phi(0) = 0$ and $p_1 = \Phi(x_1)$; p_1 being the conditional probability that alarm threshold is exceeded before the first inspection. While Munford et al. (1972) proposes to use constant p_i values for computational convenience, Chelbi et al. (1999) suggests using $p_{i+1} > p_i$ since if the inspection action does not affect the failure rate, the reliability of the equipment remains a decreasing function. Chelbi et al. (1999) suggest the next expression;

$$p_i = p_1^{\frac{1}{i}} \tag{3.13}$$

The inspection sequence can be expressed as a function of a single variable p_1

$$x_i = \Phi^{-1}\left\{p_1^{\frac{1}{i}} + \left(1 - p_1^{\frac{1}{i}}\right)\Phi(x_{i-1})\right\} \tag{3.14}$$

Then, for a given p_1 value, the inspection times, x_i's, can be found using the formula above, and then the corresponding average cost can be calculated. By changing p_1 value, different inpection sequences can be found and average costs related to these sequences can be calculated. The inspection sequence which gives the minimum average cost will be the optimal inspection times.

Some small changes are done in the formulations (Eqs.3.4-Eqs.3.10) compared to the ones proposed by Chelbi et al. (1999). In the given business case, there are two thresholds, soft and hard failure; if the soft failure threshold is exceeded a preventive replacement is scheduled, and if the hard failure threshold is exceeded a failure replacement is executed. In the case Chelbi et al. (1999) studied, if the threshold is exceeded, this increases the failure probability but the cycle continues until the failure. For instance, the probability h(.) that i inspections are made in the cycle is given in Chelbi et al. (1999) as follows;

$$h(i) = \int_{0}^{x_{i+1}} \phi(\tau)F(x_{i+1} - \tau)d\tau - \int_{0}^{x_i} \phi(\tau)F(x_i - \tau)d\tau$$

In this formulation, alarm threshold can exceed at any time between zero till x_{i+1}, but failure is waited to end the cycle. In the given business case, if the alarm threshold is exceed, and preventive replacement is scheduled and the cycle ends. Then, this is represented in Eq.(3.4).

Case 2: When failure is self-announcing

The case represented above is valid for the situations when the failure is not self- announcing and the equipment stays in a 'failed' state until it is observed by an inspection. While this assumption can be valid for some systems, it may not be applicable to others. Thus, an alternative model is represented in this section which may be applicable for the systems where the failure is self-announcing and the equipment is replaced as soon as it fails.

This alternative model requires some small changes in the formulations which are given in the previous section. The general assumptions made for the previous case are still valid for this one. For instance, the condition of the system is monitored through inspections, and like before two different threshold levels are defined, one for 'soft failure' and one for 'hard failure' and finally inspections do not affect the physical process of deterioration

For Case 1, if the i^{th} inspection performed at instant x_i reveals that the equipment has failed, then it is immediately replaced by a new identical one. On the other hand, for Case 2, since the failure is self-announcing the exact moment of the failure will be known and the equipment will be replaced at this, it does not have to wait until the ith inspection to reveal this failure. If it has not failed but the control parameter is found to have exceeded the alarm threshold level, a preventive replacement is scheduled at time $x_i + H$. For Case 1, even if the part experience a failure between the time period x_i and $x_i + H$ (the time that it is observed that the control parameter is above the threshold and the time a replacement is scheduled, respectively), the replacement operation will take place at time $x_i + H$, since it not possible to observe the failure. However, for Case 2, a failure which occurs between x_i and $x_i + H$ will be known, and a replacement can be done at the failure time which can be earlier than $x_i + H$. Finally, for both cases, if there is not any failure occurred, a preventive replacement will be done at time $x_i + H$. The time interval H depends on the logistic and any previously scheduled inspections within this period H are cancelled.

The probability density functions will represent the similar idea as in Case 1, Φ(.) is associated with a random variable τ which corresponds to the time at which the alarm threshold is exceeded, whereas f(.) is associated with the random variable t which represented the failure time. Their corresponding cumulative distribution functions are Φ(.) and F(.) and they are independent.

The definition of best inspection strategy is the same as Case 1, so the Eq.(3.1) and Eq.(3.2) will be valid for Case 2. Moreover, Eq.(3.3) and Eq.(3.4) are also valid which represents expected number of inspections in one cycle and probability of number of inspections are performed before the cycle ends respectively.

The probability P_g that the cycle ends with a failure event will be different then the previous case;

$$P_g = \sum_{i=1}^{\infty}[1 - \Phi(x_{i-1})]\left[\int_0^{x_i - x_{i-1}} F(x_{i-1} + k)dk + \int_{x_{i-1}}^{x_i} \phi(\tau)\big[1 - \big(F(x_i) - F(\tau)\big)\big]\int_0^H [F(x_i + k) - F(\tau)]dk\, d\tau\right] \tag{3.15}$$

First term is to indicate that alarm threshold is not exceed until x_{i-1}, and then second term is to indicate that failure will occure at a time point which is between x_{i-1} and x_i, so that cycle will end with a failure, and the last term is to indicate that alarm threshold is exceed at time τ which is between two consequtive inspections but failure is not observed, so a replacement is scheduled after H points of time, then failure is occurred between ith inspection and scheduled replacement time, so the cycle will finish with a failure. The probability P_p that the cycle ends with preventive replacement is given Eq.(3.7).

Finally, the cycle average duration, E(T) can be found considering the three following situations;

1. Failure occurs within the time interval $[x_{i-1}, x_i]$, in which case the cycle average duration would be;

$$ET(1) = \sum_{i=1}^{\infty} \left[\int_{0}^{x_i - x_{i-1}} (x_{i-1} + k) F(x_{i-1} + k) dk \right] \tag{3.17}$$

2. Failure occurs within time interval $[x_i, x_i + H]$, in which case the cycle ends when failure is observed;

$$ET(2) = \sum_{i=1}^{\infty} \int_{x_{i-1}}^{x_i} \phi(\tau) \int_{0}^{H} (x_i + k)[F(x_i + k) - F(\tau)] dk d\tau \tag{3.18}$$

3. The equipment does not fail within the time interval but the alarm threshold is exceeded, replacement is performed at time $[x_i + H]$. The cycle average duration would be;

$$ET(3) = \sum_{i=1}^{\infty} \left[(H + x_i) \int_{x_{i-1}}^{x_i} \phi(\tau) \left[1 - \left(F(x_i + H) - F(\tau)\right)\right] d\tau \right] \tag{3.19}$$

Hence, E(T) is given by;

$$E(T) = ET(1) + ET(2) + ET(3) \tag{3.20}$$

The definition of conditional probability p_i which represent the conditional probability that the alarm threshold is exceed within time interval $[x_{i-1}, x_i]$ is the same and given in Eq.(3.11). Then, for a given p_1 value, the inspection times, x_i's, can be found using Eq.(3.14), and then the corresponding average cost can be calculated. By changing p_1 value, different inpection sequences can be found and average costs related to these sequences can be calculated. The inspection sequence which gives the minimum average cost will be the optimal inspection times.

3.2 Sample Size Determination

A population can be defined as including all people or items with the characteristic one wishes to understand. Since there is very rarely enough time or money to gather information from everyone or everything in a population, the goal becomes finding representative sample of that population. Sampling is concerned with the selection of a subset of individuals from within a population to estimate characteristics of the whole population. The three main advantages of sampling are that the cost is lower, data collection is faster, and since the data set is smaller it is possible to ensure homogeneity and improve the accuracy and quality of the data.

There are different sampling methods. Simple random sampling implies that any particular sample of a specified sample size has the same chance of being selected as any other sample of the same size. This minimizes bias and simplifies analysis of results. The term sample size simply means the number of elements in the sample (Walpole et al., 2002).

A point estimate of some population parameter θ is a single value $\hat{\theta}$ of a statistic $\hat{\Theta}$. Clearly, it is desired that the sampling distribution $\hat{\Theta}$ to have a mean equal to the parameter estimated. An estimator possessing this property is said to be unbiased. Even the most efficient unbiased estimator is unlikely to estimate the population parameter exactly. It is true that the accuracy increases with large samples, but there is still no reason why a point estimate from a given sample is expected to be exactly equal to the population parameter it is supposed to estimate. There are many situations in which it is preferable to determine an interval within which it is expected to find the value of the parameter. Such interval is called an interval estimate (Walpole et al., 2002).

As it is mention in Chapter 1, the condition monitoring operations can be executed for a population of identical parts. In this case, monitoring operations may be performed to every part, or alternatively, since there is rarely enough time or money for that, a sample may be chosen to represent the population and monitoring operations may be executed for this sample. Monitoring every part will provide the perfect information about the population, but it will cause a high monitoring cost. On the other hand, observing a sample may decrease the monitoring cost but it will cause some uncertainty related to condition of population. Hence, a trade-off between sampling cost and uncertainty related to the population condition has to be done.

If all the parts in a population are not monitored but only a sample is, the uncertainty related to that situation can be represented by the confidence interval of parameters estimate. The more parts are observed, the better the estimation and the smaller the interval is. However, if fewer parts are observed, estimation will be poor and interval will be bigger. In the former case, interval will be smaller but the monitoring cost will be higher compared to the latter case where the interval is bigger but the monitoring cost is lower. Thus, the trade-off can be done between the length of the confidence interval and monitoring cost.

Confidence interval for μ;

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal population (which is a valid assumption due to Central Limit Theorem), a $(1-\alpha)100\%$ confidence interval for μ is;

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}} \tag{3.29}$$

where $t_{\alpha/2}$ is the t value with $\vartheta = n-1$ degrees of freedom (Walpole et al., 2002).

Then the length of the interval will be;

$$2t_{\alpha/2}\frac{s}{\sqrt{n}} \tag{3.30}$$

The monitoring cost will be;

$$cn \tag{3.31}$$

where c is the inspection cost per one item (the set up is ignored here since the main goal is finding sample size, not deciding on whether we should do inspection or not where the setup cost may be relevant, but since we are doing inspection anyway, it is irrelevant in this case).

Then, total monitoring cost will have two components; the first one is to represent the penalty of having a big confidence interval, the second one is to represent the inspection cost.

$$Monitoring\ cost = a \times interval\ length + cn$$

$$Monitoring\ cost = a \times 2t_{\alpha/2}\frac{s}{\sqrt{n}} + cn \tag{3.32}$$

where a is the cost associated with the length of the confidence interval (uncertainty cost).

By increasing the sample size (n), the first term in monitoring cost will decrease since for higher sample size values smaller interval lengths will be obtained, but the second term will increase since more parts will be inspected. Hence, the sample size which minimizes the monitoring cost given above will be optimum sample size.

4. Data Analysis

In this chapter the analysis of historical data is presented. This analysis is done for every train type that NedTrain has, and the main goal is to understand the oil contamination process.

As a starting point, the trains that NedTrain owns were grouped into six categories depending on some technological differences; ICM 1-2, ICM 3, SGMM, MDDM, VIRM and VIRM-4. Then, the data analysis was done one by one for each train type.

The data sheet taken from NedTrain contains last two years oil analysis data for each gearbox. The information represented in the data sheet includes gearbox code, train type, train number, oil sampling date, oil analysis date, amount of iron, copper, lead, tin, chrome, nickel, aluminum and silicon and whether oil sample is rejected or not (which is related to oil changing; if the oil sample is rejected then the corresponding gearbox has to experience an oil change).

After extracting the related information for each train type from the main data sheet, the first analysis done was about understanding the degradation/metal accumulation process. Metal accumulation in the oil is one of the main indicators of oil quality. In order to understand its mechanism, a relation has to be found between time passed since the last oil change and the amount of metal accumulated in it. In the main data sheet, there was not any information about the oil changing time; hence, one has to do some further search to understand the oil accumulation mechanism because it is not possible to relate the time and metal amount without knowing the starting time. Therefore, as a solution for this situation, it was decided that for further analysis the gearboxes which did not have any oil rejection have to be ignored. The reason behind this decision is that; if there is not any oil rejection for a gearbox that means that this gearbox did not experienced any oil change (at least during the time that has records in the main data sheet) and since the last oil changing time was not know, it is not possible to find a relation between time and metal accumulation because the starting time/time zero for oil change is not known. On the other hand, if a gearbox is experienced an oil rejection, then it is known that the oil in this gearbox will be changed and this will be the starting time/time zero for metal accumulation, assuming that there is not any metal in time zero, hence the relation between time and metal accumulation can be found. For instance, if there is a rejection for oil sample in 27

Nov 2010, then it can be assumed that the oil was changed and until the next sampling time, 1 March 2011, 90 days passed and 20 ppm of iron was accumulated. Thus, a relation can be observed between time and metal accumulation. Moreover, in order to find a general relation for a specific metal accumulation in a specific train type (e.g. iron accumulation in ICM 1-2) the information that comes from whole gearboxes for that train type was combined together. For instance, for gearbox 1, it was found that after 93 days and 179 days passed from oil rejection 185 ppm and 200 ppm of iron was accumulated respectively, for gearbox 2, it was found that after 90, 176, and 254 days passed, 160, 230 and 310 ppm of iron was accumulated respectively. Then, while finding the general relation for iron accumulation for these two gearboxes, 5 data points (coming from gearbox 1 and 2) were used together and then the relation would be valid for both gearbox 1 and 2. This is a valid application because the gearboxes for the same train type are assumed to be identical, and while using the whole data points, one can have more information about the system since there is very limited information on the basis of gearbox. As mentioned before, this was done for each train type, the data points for ICM 1-2 can be seen in the Appendix A as an example.

Secondly, after computing the data points for time and metal accumulation from the main data set, a statistical analysis was done to find the best fit for these data points and represents metal accumulation with respect to time. First of all, an outlier analysis was conducted. Extreme events, data entry errors etc. can cause outliers and they can dramatically change the results of analysis. Outliers can be univariate, bivariate and multivariate. There are different identification methods for each of them. For univariate outliers, one has to determine standardized data. Graphical examination can be used to determine bivariate outliers. Finally, for determining multivariate outliers, Mahalonobis distance (D^2) can be used (Hair et al., 2009). For the project, SPSS is used to determine univariate and multivariate outliers. In order to determine univariate outliers, standardize data is checked.

$$z = \frac{x_i - \bar{x}}{sd}$$

If sample size is greater than 80, the data points which have z-scores greater than 3 are considered to be an outlier. If it is less than 80, z-scores greater than 2.5 are considered to be an outlier (Hair et al., 2009).

In order to determine multivariate outliers Mahalonobis distance is calculated. It adjusts for both variability and correlation between variables. If $p < 0.001$, than the corresponding combination of data points is assumed to be an outlier (Hair et al., 2009).

$$D^2 = \frac{1}{1-r^2}\left[\frac{(x_{i1}-\overline{x_1})^2}{s_1^2} + \frac{(x_{i2}-\overline{x_2})^2}{s_2^2} - \frac{2r(x_{i1}-\overline{x_1})(x_{i2}-\overline{x_2})}{s_1 s_2}\right]$$

The univariate and multivariate outliers for ICM 1-2 can be seen in Appendix B as an example. These multivariate outlier points are deleted from the whole data set. In addition to this, if there are extra univariate outliers for a specific metal type, they are also deleted for further analysis.

In order to find the best relation between metal accumulation and time, different relations were tested; linear, logarithmic, inverse, quadratic, cubic, power, growth, and exponential, using SPSS's Curve Estimation option under Regression tool box. The SPSS output for each train type is given in Appendix C, and the results are presented below.

Linear relations were fit best for each metal type in every train group. In some cases, the R-squared values were higher for other relations, which can be seen as a better fit, however these relations did not represent the physical behavior properly. For instance, a quadratic relation with a high R-squared value which is claiming a decrease in metal accumulation with respect to time cannot be a good fit, since it does not represent the physical behavior of the system.

In the equations, x represents the days passed and y represents ppm of corresponding metal accumulated.

Intercept is set to zero, since it is assumed that there is not any metal particles in the oil when it has changed.

Table 1: Relation between time and metal accumulation for ICM 1-2

	Equation	R-Squared
Fe	$y = 0{,}6x$	0,652
Cu	$y = 0{,}081x$	0,646
Pb	$y = 0{,}007x$	0,348
Sn	$y = 0{,}002x$	0,204
Cr	$y = 0{,}004x$	0,556
Al	$y = 0{,}013x$	0,295
Si	$y = 0{,}1x$	0,449

Table 2: Relation between time and metal accumulation for ICM 3

	Equation	R-Squared
Fe	$y = 0{,}526x$	0,643
Cu	$y = 0{,}07x$	0,608
Pb	$y = 0{,}006x$	0,269
Sn	$y = 0{,}002x$	0,115
Cr	$y = 0{,}004x$	0,551
Al	$y = 0{,}012x$	0,280
Si	$y = 0{,}12x$	0,449

Table 3: Relation between time and metal accumulation for SGMM

	Equation	R-Squared
Fe	$y = 0{,}430x$	0,607
Cu	$y = 0{,}09x$	0,765
Pb	$y = 0{,}007x$	0,317
Sn	$y = 0{,}001x$	0,105
Cr	$y = 0{,}004x$	0,496
Al	$y = 0{,}017x$	0,301
Si	$y = 0{,}098x$	0,433

Table 4: Relation between time and metal accumulation for MDDM

	Equation	R-Squared
Fe	$y = 0{,}218x$	0,549
Cu	$y = 0{,}190x$	0,574
Pb	$y = 0{,}011x$	0,3
Sn	$y = 0{,}001x$	0,0071
Cr	$y = 0{,}004x$	0,569
Al	$y = 0{,}004x$	0,179
Si	$y = 0{,}031x$	0,409

Table 5: Relation between time and metal accumulation for VIRM

	Equation	R-Squared
Fe	$y = 0{,}253x$	0,596
Cu	$y = 0{,}042x$	0,674
Pb	$y = 0{,}003x$	0,211
Cr	$y = 0{,}002x$	0,432
Al	$y = 0{,}006x$	0,175
Si	$y = 0{,}067x$	0,277

Table 6: Relation between time and metal accumulation for VIRM-4

	Equation	R-Squared
Fe	$y = 0{,}123x$	0,422
Cu	$y = 0{,}035x$	0,605
Pb	$y = 0{,}002x$	0,206
Cr	$y = 0{,}002x$	0,385
Al	$y = 0{,}001x$	0,114
Si	$y = 0{,}024x$	0,488

After finding the relation between time and metal accumulation for each train category and every metal type, the expert suggested that the amount of accumulated metal in the oil may depend on the age of gearbox in addition to time passed until the last oil change. The reason behind this ideas was that; the metal particles in the oil are coming from some parts of the gearbox (e.g. iron may be coming from either closing or bearing, nickel is coming from bearing etc.). Then, if a gearbox is new, deterioration of its parts will be low and they will produce fewer amounts of metal particles compared to an old gearbox. Hence, even if the same amount of time passed, one may expect to observe more metal particles in the oil for an old gearbox.

In order to include this idea into the analysis, company database was checked to find the relevant information. It was found that there is information for only MDDM type of trains, the exact day of gearbox's first usage is recorded. The gearboxes that were bought in the same year were grouped together, and their ages were found. There have been five groups; the ones bought in year 1996, 1997, 1998, 2000 and 2004. Then, multiple regression analysis is performed considering time and age as independent variables and metal accumulation amount as dependent variable. The results are shown below, details are given Appendix D.

In the equations, x represents the days passed, a represents the age of gearbox and y represents ppm of corresponding metal accumulated.

Table 7: Multiple linear regression for metal accumulation for MDDM

	Equation	R-Squared
Fe	$y = 3{,}762a$	0,756
Cu	$y = 0{,}02x + 2{,}889a$	0,719
Pb	$y = 0{,}05x + 0{,}095a$	0,326
Si	$y = 0{,}578a$	0,626

The R-squared values were improved for this multiple regression analysis. Therefore, it may be suggested to NedTrain that, keeping records for the age of gearboxes for other types of train may be useful for their operations.

5. Model Application

This chapter has three sections. In the first section, the description of the system that represents the business case given in the Chapter 1 is presented. Then, in the second section, the results for this case are given. Finally, in the last section some sensitivity analysis related to model is clarified.

5.1 System Description

The model described in Chapter 3 will be used to solve the business case given in Chapter 1. As it was mentioned in Chapter 2, Chelbi et al. (1999) defined the optimal inspection sequence as the one that minimizes the expected total cost and proposed a mathematical model and a numerical algorithm to generate this optimal inspection sequence which is applicable for a practical case. This idea will be applied to NedTrain's case.

In the given business case, the system which deteriorates is the oil in gearboxes, and the selected control parameters to monitor are the amount of polluting materials in it. In order to monitor the condition of the oil, at each inspection time oil samples are taken from each gearbox and the amount of iron, copper, lead, tin, nickel, chrome, aluminum, and silicon are measured. These materials accumulate in the oil as time passes, and their accumulation processes are analyzed in Chapter 4. Linear relations are found to be the better one to represent the process. According to Gabraeel et al. (2009), linear degradation model is typically used for modeling degradation processes. During this data analysis, it is observed that some of the given control parameters (copper, lead, tin, chrome, aluminum and silicon) do not provide useful information about deterioration process, their accumulation rates are close to zero. Then, iron is selected as the informative one, and used in the further analysis. This idea is commonly used in literature, for instance at Bonjevic et al. (2006), measurements for iron, aluminum and magnesium are taken but iron is used for analysis since it is the most informative and significant one. If the given control parameter accumulates beyond a specified value, then the oil will be useless. Two different threshold levels are defined. The first one is the 'alarm threshold' which represents that the condition of the equipment is not good enough to work and a replacement has to be scheduled, this situation can be seen as a 'soft failure' which is not a real failure but it is indicates that the equipment is not at a good condition anymore. The second threshold level is the 'failure threshold' which represents that it is not possible to use the equipment anymore; it can be seen as a 'hard failure'. For the NedTrain's case, the 'equipment' that is going to be monitored is the 'oil in the gearboxes'. It is not possible to talk about a 'hard failure' for oil, hereby the hard failure concept will have a different meaning; if the oil is contaminated too much, it will start to damage the gearboxes, so this situation will be considered as 'hard failure'. The idea about 'soft failure' is applicable for NedTrain's case, if the contamination level is above the alarm threshold, then this will means that oil is not good enough to be used anymore, and an oil replacement has to be scheduled. It is obvious that the failure events are not self-announcing since even if the oil is contaminated too much it is not possible to know that without an inspection.

It is assumed that inspections do not affect the physical process of deterioration and failure can only captured by inspection. For NedTrain's case, oil sampling times are the inspection times and inspections are nondestructive since taking oil samples will affect neither the gearbox nor the existing oil in it. If the i^{th} inspection performed at instant x_i reveals that oil is not in a good condition and a replacement is necessary, a preventive replacement is scheduled at time $x_i + H$. The time interval H depends on the logistic and any previously scheduled inspections within this period H are cancelled. NedTrain waits for next visit of the train to the maintenance depots, which is usually 3 months later, hence the parameter H in the model can be considered to be 3 months.

To represent the physical behavior of the system, two probability density functions Φ(.) and f(.) are considered. ϕ(.) is associated with a random variable τ which corresponds to the time at which the alarm threshold is exceeded, whereas f(.) is associated with the random variable t which represented the failure time. Their corresponding cumulative distribution functions are Φ(.) and F(.) which are independent. In the given business case, the condition monitoring operation related to the oil in all gearboxes for a given type of train is considered. Then, in this case, the physical behavior that is going to be considered should represent the whole population. The system can be represented as in Figure 12 below; each line is showing the evolution of control parameter with time for a specific gearbox.

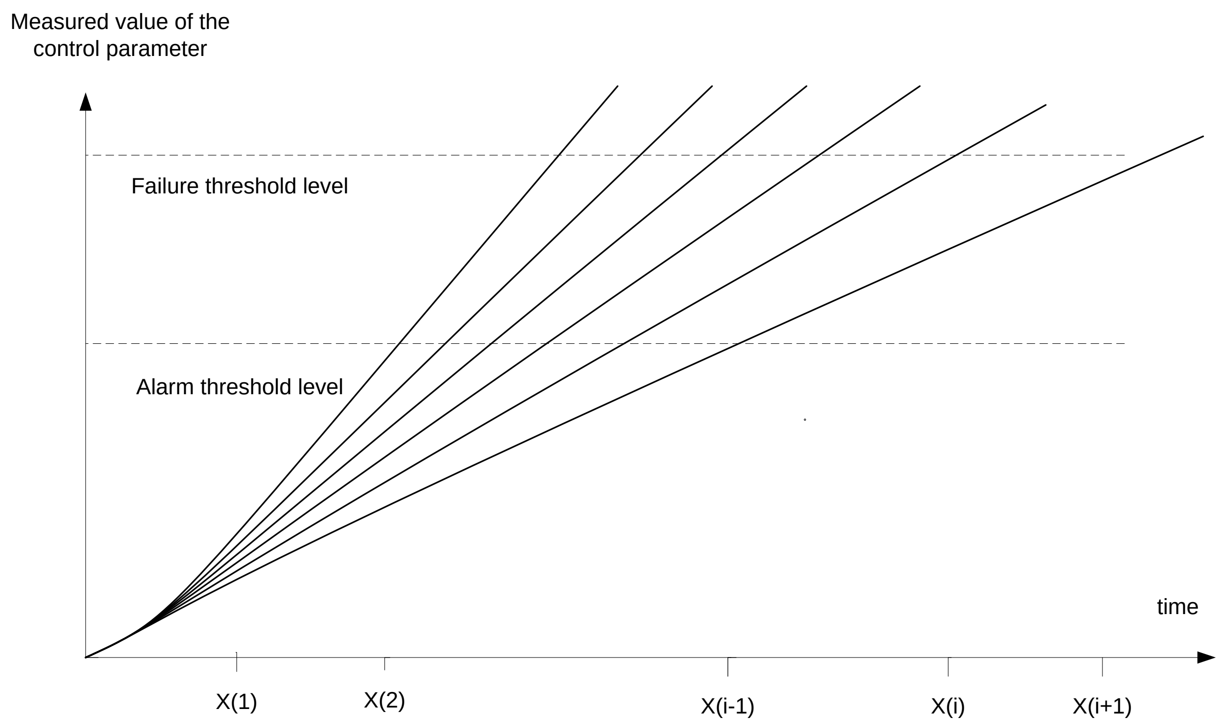


Figure 12: Control parameter evolution with time for population

Then, in order to determine the two probability density functions defined above, maintenance policy defined as follows;

The time when 50% of the population exceeds the alarm threshold will be recorded as the 'soft failure' time of the population. Similarly, the time when 50% of the population exceeds the failure threshold will be recorded as the 'hard failure' time of the population. For instance, if the population is formed from 100 parts, and if control parameter of 50 parts is above the alarm threshold, then it can be concluded that the population has experienced a 'soft failure' and the time recorded as soft failure time. The idea is the same for 'hard failure'. Then, two probability density functions, $\Phi(.)$ and $f(.)$, which are explained before, will represent the whole population.

Then, since the failure is not self-announcing, the ideas and definitions given in Chapter 3 as Case 1 are valid for the business cans and the optimum inspection sequence can be calculated by using Eqs. (3.1-3.14).

5.2 Results for Business Case

5.2.1 Inspection Times

The results are given for train type ICM 1-2. For other train types the results are tabulated in Appendix E.

As a starting point, the rate of metal accumulation for oil in each gearbox is found. As it is mentioned in Chapter 4, linear relation is the best one to represent the amount of metal in the oil which increases with time. Hence, for each gearbox the situation can be represented as follows;

$$y = ax \tag{5.1}$$

where y is the amount of metal in oil, x is the time passed until the last oil change and a is the slope/metal accumulation rate. The intercept set to zero since at time zero (after the oil change) the oil is clean and there is not any metal particles in it. Thus, as a starting point, the a's are found for each gearbox.

For the given train type, there are 64 data points from different gearboxes, and accumulation rates for those 64 gearboxes are found. Then, in order to use in the further analysis, a probability distribution is fit to those rates which is $lnN{\sim}(\mu, \sigma)$ where $\mu = -0.325\ and\ \sigma = 0.472$.

Then, in order to estimate failure times, 100 random numbers are generated using the distribution $lnN{\sim}(-0.325{,}0.472)$ which are assumed to be the slopes of Eq.(5.1) for different gearboxes. The threshold values have to be assumed in order to estimate the 'soft' and 'hard' failure times. These threshold values are assumed as 500ppm and 800ppm respectively, for iron accumulation. When the historical data is analyzed, it is seen that the oil is usually rejected around the given values, so they are assumed to be the threshold values. Then, since the contamination occurs in a linear manner, the slopes are found and thresholds are set to specific value, failure times can be estimated. For each gearbox, if y is set to be the 'soft failure' threshold, which is 500ppm, and a is the slope represented by random number, then x will be the 'soft failure' time, and similarly if y is set to be the 'hard failure' threshold, which is 800ppm, x will be the 'hard failure' time. This is repeated for every gearbox, and then soft and hard failure times for 100 gearboxes will be estimated. Since the soft and hard failure times of the population are defined as "time when the control parameter of 50% of the population exceeds the alarm threshold, and time when the control parameter of 50% of the population exceeds the failure threshold", then the time when the 50 gearboxes exceed the alarm threshold and failure threshold can be found, and these times will be the 'soft failure' and 'hard failure' times of the given population.

The procedure given above is repeated 30 times and 30 different soft failure and hard failure times for given population are found. Then, probability distributions are fit to these failure times to represent the probability of soft failure and hard failure with respect to time, and these distributions are used in the Eqs.(3.1-3.14).

Soft failure $\Phi \sim Weibull(738{,}15)$

Hard failure $F \sim Weibull(1179{,}18)$

The Weibull distribution usually provides the best fit of life data. This is due in part to the broad range of distribution shapes that are included in the Weibull family (Abernethy (2006)).

After estimating the failure time distributions, Eq.(3.14) can be used to find inspection times.

$$x_i = \Phi^{-1}\left\{p_1^{\frac{1}{i}} + \left(1 - p_1^{\frac{1}{i}}\right)\Phi(x_{i-1})\right\} \tag{3.14}$$

with $x_0 = 0$, $\Phi(0) = 0$ and $p_1 = \Phi(x_1)$; p_1 being the conditional probability that alarm threshold is exceeded before the first inspection. For a given p_1 value, the inspection times, x_i's, can be found using Eq.(3.13), and then the corresponding average cost can be calculated. By changing p_1 value, different inpection sequences can be found by using Eq.(3.14) and average costs related to these sequences can be calculated by using Eqs.(3.1-3.10).

In this business case p_1 is changed from 0.01 to 0.99, and 99 different inspection sequences are found. Then, for each sequence the expected number of inspections in one cycle is calculated by using Eqs.(3.3,3.4). The probabilities that the cycle ends with a failure replacement and preventive replacement are found by using Eq.(3.6) and Eq.(3.7) respectively. Now, the expected total cost can be calculated by Eq.(3.2), but the cost of inspection, failure replacement and preventive replacement have to be determined. They are assumed to be as follows, respectively;

$$c_I = 10000$$

$$c_g = 70000$$

$$c_p = 15000$$

Then, the expected total cost calculated by using Eq.(3.2);

$$E(C) = c_i E(I) + c_g P_g + c_p P_p \tag{3.2}$$

Finally, the cycle duration has to be found, which has two components. The first one is calculated by using Eq.(3.8) and the second one is calculated by using Eq.(3.9). Then, the average cycle duration is given by Eq.(3.10);

$$E(T) = \sum_{i=1}^{\infty} \left[x_i \int_{x_{i-1}}^{x_i} F(\tau) d\tau + (H + x_i) \int_{x_{i-1}}^{x_i} \phi(\tau) [1 - (F(x_i) - F(\tau))] d\tau \right] \tag{3.10}$$

Hence, the average total cost per unit time can be estimated by Eq.(3.1) for each inspection sequence;

$$E(AC) = \lim_{T \to \infty} \frac{\{E[C(T)]\}}{T} \to E(AC) = \frac{E(C)}{E(T)} \tag{3.1}$$

Then, the one which gives the smallest average total cost will be the optimum inspection sequence. The results are found by coding in MATLAB, and the optimum inspections times (which are in days) for the given situation above are;

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
760	805	834	855	871	885	896	906

5.2.2 Sample Size

As it is mentioned before, monitoring every part will provide the perfect information about the population, but it will increase the monitoring cost. On the other hand, observing a sample may decrease the monitoring cost but it will cause some uncertainty related to condition of population. Hence, a trade-off between sampling cost and uncertainty related to the population condition has to be done. For the given business case, the condition of oil in the gearboxes for train type ICM1-2 is monitored. Thus, a sample size for these monitoring operations has to be determined. In Chapter 3, it is stated that for finding the optimum sample size a trade-off can be done between the length of the confidence interval of mean and the monitoring cost. The interval length, inspection cost and total monitoring cost are given as Eq.(3.30), Eq.(3.31) and Eq.(3.32) respectively.

The inspection sequence is determined in the previous section. Then, the sample size for each inspection can be estimated by using Eqs.(3.30-3.32). First, the metal amount in the oil is estimated for different numbers of samples and the interval length is calculated by Eq.(3.30). Then, the inspection cost per one item is assumed to be 200 and uncertainty cost is assumed to be 1000. Hence, the required sample is found to be 118, which means for the given cost parameters taking oil samples from 118 gearbox will represent the population.

5.3 Sensitivity Analysis

5.3.1 Effects of Different Cost Parameters to Inspection Sequence

As it is stated before, the inspection sequence which gives the minimum average cost is the optimal inspection times. The average total cost per unit time can be estimated by Eq.(3.1) for each inspection sequence;

$$E(AC) = \lim_{T\to\infty} \frac{\{E[C(T)]\}}{T} \to E(AC) = \frac{E(C)}{E(T)} \tag{3.1}$$

where

$E(C) = average\ total\ cost\ during\ one\ cycle$

$E(T) = expected\ cycle\ duration$

The expected cost has three components; inspection cost, failure replacement cost and preventive replacement cost. It is representes as;

$$E(C) = c_i E(I) + c_g P_g + c_p P_p \tag{3.2}$$

Then, E(T) is given by;

$$E(T) = \sum_{i=1}^{\infty} \left[x_i \int_{x_{i-1}}^{x_i} F(\tau) d\tau + (H + x_i) \int_{x_{i-1}}^{x_i} \phi(\tau)[1 - (F(x_i) - F(\tau))] d\tau \right] \qquad (3.10)$$

For a given p_1 value, the inspection times, x_i's, can be found using Eq.(3.14), and then the corresponding average cost can be calculated. For a given inspection sequence, *E(T)* will be same, on the other hand, changing the cost parameters in Eq.(3.2) can change *E(C)*. Hence, if different inspection cost, failure replacement cost and preventive replacement cost are used, different optimum inspection times can be obtained.

There is a trade-off between these cost components. If the failure replacement cost too high, then in order to decrease the probability of having a failure, the inspection sequence will start at an earlier time and more inspections are executed. On the other hand, if the failure replacement cost is not high compared to others, then the inspection operations will be delayed and sequence will start at a later time.

In order to observe this behavior, three different levels are assigned for each cost parameter; low, medium and high. Then, the program is run for 27 cases and results are presented and compared below. The cost parameters are as follows;

$$c_I = (low, medium, high) = (10000, 50000, 100000)$$

$$c_g = (low, medium, high) = (200000, 3000000, 5000000)$$

$$c_p = (low, medium, high) = (20000, 200000, 400000)$$

For the first three cases c_I and c_p are set to their low values and c_g is changed;

Case 1: (low, low, low)

Case 2: (low, medium, low)

Case 3: (low, high, low)

The optimum inspections are given in Table 8 below;

Table 8: Inspection times for different cost parameters

Inspections						
	1	**2**	**3**	**4**	**5**	**6**
Case 1	760	805	834	855	871	885
Case 2	543	639	689	723	748	768
Case 3	543	639	689	723	748	768

As it can be seen from the examples above, if the failure replacement cost is high, then inspections are done at earlier times, which is something expected.

For the next three cases c_I is fixed at its lowest level, c_p is at medium level and c_g is changed;

Case 4: (low, low, medium)

Case 5: (low, medium, medium)

Case 6: (low, high, medium)

The optimum inspections are given in Table 9 below;

Table 9: Inspection times for different cost parameters

Inspections						
	1	**2**	**3**	**4**	**5**	**6**
Case 4	817	859	885	904	919	932
Case 5	543	639	689	723	748	768
Case 6	543	639	689	723	748	768

Similar to previous cases, as the failure replacement cost increases, inspections are performed at earlier times in order to prevent a possible failure. Additionally, if one compares Case 1 and Case 4, then inspections can be performed at later times because at that times failure replacement probability is higher than preventive replacement probability and since preventive replacement cost is higher than failure replacement cost this is situation is preferable.

For the next three cases c_I is fixed at its lowest level, c_p is at its high level and c_g is changed;

Case 7: (low, low, high)

Case 8: (low medium, high)

Case 9: (low, high, high)

The optimum inspections are given in Table 10 below;

Table 10: Inspection times for different cost parameters

Inspections						
	1	**2**	**3**	**4**	**5**	**6**
Case 7	817	859	885	904	919	932
Case 8	797	840	867	887	902	915
Case 9	543	639	689	723	748	768

Similarly, when the failure replacement cost is high, inspections are performed at earlier times in order to prevent a possible failure. Moreover, since the preventive replacement cost is higher than the previous cases, inspections are performed at later times since as time passes failure replacement probability increases and preventive replacement probability decreases.

For the next three cases c_I is fixed at its medium level, c_p is at low level and c_g is changed;

Case 10: (medium, low, low)

Case 11: (medium, medium, low)

Case 12: (medium, high, low)

The optimum inspections are given in Table 11 below;

Table 11: Inspection times for different cost parameters

Inspections						
	1	**2**	**3**	**4**	**5**	**6**
Case 10	751	797	826	847	864	878
Case 11	747	794	823	844	861	875
Case 12	743	790	819	841	858	872

Similarly, as the failure replacement cost increases, inspections are performed at earlier times. Additionally, inspections are performed some days later compared to Case 1 and Case 2 when the inspection cost is higher.

For the next three cases c_I is fixed at its high level, c_p is at low level and c_g is changed;

Case 13: (high, low, low)

Case 14: (high, medium, low)

Case 15: (high, high, low)

The optimum inspections are given in Table 12 below;

Table 12: Inspection times for different cost parameters

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 13	751	797	826	847	864	878
Case 14	749	796	825	846	863	877
Case 15	747	794	823	844	861	875

When the failure replacement cost increases, inspections are performed at earlier times. Additionally, inspections are performed some days later compared to Case 1, Case 2, Case 11 and Case 12 when the inspection cost is higher.

For the next three cases c_I and c_p are set to their medium levels and c_g is changed;

Case 16: (medium, low, medium)

Case 17: (medium, medium, medium)

Case 18: (medium, high, medium)

The optimum inspections are given in Table 13 below;

Table 13: Inspection times for different cost parameters

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 16	817	859	885	904	919	932
Case 17	768	813	841	861	877	890
Case 18	760	805	834	855	871	885

Inspections are performed at earlier times when the failure replacement cost is high. Moreover, if the inspection cost and preventive replacement cost are higher, then inspections are performed at some days later compared to Case 1, Case 2 and Case3.

For the next three cases c_I is set to its medium level and c_p is set to its high levels and c_g is changed;

Case 19: (medium, low, high)

Case 20: (medium, medium, high)

Case 21: (medium, high, high)

The optimum inspections are given in Table 14 below;

Table 14: Inspection times for different cost parameters

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 19	817	859	885	904	919	932
Case 20	808	851	877	896	911	924
Case 21	802	845	872	891	906	919

When the failure replacement cost is high, inspections are performed at earlier times. Moreover, when the preventive replacement cost is higher compared to previous cases, then inspections are performed at later time because as time passes preventive replacement probability decreases.

For the next three cases c_I is set to its high level and c_p is set to its medium levels and c_g is changed;

Case 22: (medium, low, high)

Case 23: (medium, medium, high)

Case 24: (medium, high, high)

The optimum inspections are given in Table 15 below;

Table 15: Inspection times for different cost parameters

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 22	817	859	885	904	919	932
Case 23	760	805	834	855	871	885
Case 24	755	801	830	851	868	882

As expected when the failure replacement cost is high, inspections are performed at earlier times. Additionally, increase in preventive replacement cost delays the inspection operations.

For the next three cases c_I is set to its high level and c_p is set to its high levels and c_g is changed;

Case 25: (medium, low, high)

Case 26: (medium, medium, high)

Case 27: (medium, high, high)

The optimum inspections are given in Table 16 below;

Table 16: Inspection times for different cost parameters

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 25	817	859	885	904	919	932
Case 26	787	831	858	878	894	907
Case 27	772	816	844	864	880	893

Inspections are performed at earlier times if failure replacement cost is high. Moreover, when the preventive replacement cost is higher compared to previous cases, then inspections are performed at later times.

5.3.2 Inspection Sequence for Single Item

Another issue that affects the inspection times is the soft failure and hard failure probability functions. In the business case given above, the inspection times that are considered are for the population. The soft failure and hard failure for the population are defined as; if 50% of the population is above the alarm threshold, then the soft failure will occur for the population, and similarly if 50% of the population is above the failure threshold, then the hard failure will occur for the population. Then, while finding the soft failure and hard failure probability functions with respect to time, these definitions are considered, and since these population failures usually happen at similar times, the variances of these functions are very low. As a result, the inspection times are grouped around the expected life time. However, if the 'equipment' that is going to be considered is a single part not a group of identical parts, then the variability for the failure times will be higher and inspections will be done at earlier times.

In order to test this idea, the failure time distributions are estimated for a single item. First, with the help of historical data, the soft failure and hard failure distributions of oil from 63 gearboxes are estimated. Then, soft failure and hard failure distributions are fit to these data;

Soft failure $\Phi \sim Weibull(885.4{,}2.03)$

Hard failure $F \sim Weibull(1416.7{,}1.97)$

The trade-off between the cost components is also applicable for this case. If the failure replacement cost too high, then in order to decrease the probability of having a failure, the inspection sequence will start at an earlier time and more inspections are executed. On the other hand, if the failure replacement cost is not high compared to others, then the inspection operations will be delayed and sequence will start at a later time.

If the cost of inspection, failure replacement and preventive replacement are assumed to be as follows; $c_I = 500$, $c_g = 50000$, $c_p = 1000$

The optimum inspections are found by using Eqs.(3.1-3.14) and given below;

	Inspections					
	1	**2**	**3**	**4**	**5**	**6**
Case 1	386	617	876	1178	1534	1955

5.3.3 Effects of Different Cost Parameters to Sample Size

The other point related to the sensitivity analysis is about sample size. It is mentioned before that the sample is found by making a trade of between the interval length of mean estimation and monitoring cost. So, if these cost factors change, the optimum sample size will also change. For instance, if the interval length is wanted to be shorter, then a higher cost for that can be assigned and then the optimum sample size will be bigger since if the more samples are observed, the more confident about the condition of the population and the shorter the confidence interval.

In order to observe this idea, the cost related to the interval length is assumed to be at higher levels and the inspection cost per unit item is assumed to be 200. For each inspection different interval costs are used since for the first inspections it may be tolerable to have a wide interval and a smaller interval cost can be used, but when time passes and later inspections are done it may be desired to have narrow interval and a higher interval cost can be used. In Table 17, the sample size which minimizes the monitoring cost given in Eq.(3.32) for each inspection is given.

Table 17: Sample size corresponding to inspections

	Inspections							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	118	135	230	276	298	332	352	384
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

5.3.4 Effect of Maintenance Policy

In order to represent the population behavior, a maintenance policy is defined in previous section, which is; "the time when 50% of the population exceeds the alarm threshold will be recorded as the 'soft failure' time of the population. Similarly, the time when 50% of the population exceeds the failure threshold will be recorded as the 'hard failure' time of the population". Then, soft failure and hard failure probability functions are found using those failure times. The percentage used in calculations is a decision parameter, and using different percentages will result in different inspection sequences. For instance, if a lower percentage is used, then 'failures' will occur at earlier times and as a result inspections will be executed at earlier time points. In order to test this idea, failure of 10% percent of the parts is used as a failure event of the population. Then, probability distributions are fit to these failure times to represent the probability of soft failure and hard failure with respect to time, and these distributions are used in the Eqs.(3.1-3.14).

Soft failure $\Phi \sim Weibull(390{,}14)$

Hard failure $F \sim Weibull(625{,}14)$

The cost of inspection, failure replacement and preventive replacement are assumed to be as follows, respectively; $c_I = 10000$, $c_g = 70000$, $c_p = 15000$ and the optimum inspections times for the given situation are;

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
397	423	440	452	462	470	477	483

5.3.5 Validation of Results

According to Irobi et al.(2004) face validity is one of the most commonly used validation technique which is used to know if the logic used in the conceptual model is correct and if the input-output relationship is reasonable. This has to do with asking knowledgeable people if the system model behavior is reasonable. In this case Bob Huisman, manager maintenance development by NedTrain, is asked about it. The procedure that is suggested in this research is confirmed by him and said to be useful for the company and will be implemented in the future.

6. Conclusion and Recommendations

In this chapter an overview of the conclusion and recommendations that can be made based on the research presented in the previous chapters is given. In the first section the conclusion of the research is presented. Then, in the second section some future research options are explained.

6.1 Conclusion

This project is held in cooperation with NedTrain Maintenance Development Department and aimed at understanding the oil contamination process in gearboxes and determining optimum inspection times for oil sampling and required sample size for each inspection. Then, the research question is stated as follows;

How does the contamination process behave?

What are the optimal inspection times and sample size that gives the minimum maintenance cost for condition monitoring process?

Maintenance cost includes inspection cost, preventive replacement cost and failure replacement cost.

In order to answer the first research question, data analysis is conducted. The historical data which belong to six different types of trains are analyzed. In order to monitor the condition of the oil, samples are taken from each gearbox and analyzed for the existence and amount of eight different materials (i.e. iron, copper, lead, tin, chrome, nickel, aluminum and silicon) in the oil. At the end of the analysis, it is found the linear relation is the best fit for the accumulation of each material with respect to time in the oil for all train types. The other result of this analysis is that, besides iron, all other materials accumulate in a very slow rate, and iron is the informative one for further analysis.

In order to answer the second research question, firstly the existing literature is searched for inspection planning models. The model suggested by Chelbi et al.(1999) is found to be suitable since a numerical algorithm is used to generate the optimal inspection sequence and it is applicable for a practical case. This model is explained in detail in Chapter 3. The optimum inspection sequence is defined as the one which gives the minimum expected total cost. Then, using the model presented in Chapter 3, optimum inspection sequence is determined. For the other part of the second research question, a model is thought to find the optimum sample size for each inspection operation. This model which minimizes total monitoring cost is explained in detail in Chapter 3, and by assigning relevant cost parameters, the optimum sample size is determined. Thus, the research questions are answered.

After determining the inspection sequence and sample size for each inspection, a sensitivity analysis is conducted. As it is mention before, the optimum inspection sequence is found by minimizing the expected total cost, and the cost is composed of inspection cost, preventive replacement cost and failure replacement cost. As a result of sensitivity analysis, it is seen that if the failure replacement cost is higher compared to other costs, then inspections are executed at earlier times. Hence, if failure of the equipment results in a serious problems and failure replacement cost is high, then in order to estimate failure before it happens, inspections should be done at earlier times. A similar analysis is conducted for sample size determination. If one is more sensitive about the precision of the data collected to represent the population, then they have to analyze more parts, which means a bigger sample size.

6.2 Implementation

In order to answer the research questions given in Chapter 1, a literature survey is conducted and the idea proposed by Chelbi et al. (1999) is used in the further analysis. The optimal inspection sequence is defined as the one which minimizes the expected total cost and a mathematical model and a numerical algorithm to generate this sequence is proposed which is applicable for a practical case.

In addition to the optimum inspection sequence, sample size for each inspection operation is determined. NedTrain has six different train types, and oil samples are taken from every gear for each train types. With the assumption of identical gearbox for each train type, sample sizes are determined to represent the population. Hence, taking oil samples from a part of the population will be enough to gain information about the whole group.

In order to apply the ideas given in this report at NedTrain, an initialization processes is needed. In the current case, the operations are done in the gearbox bases; oil samples are taken from each gearbox for condition monitoring operations and the age of the oil is different for every gearbox. In order to apply the population idea, first the oil in the all gearboxes has to be changed. Then, the condition will be same for each gearbox and the population idea can be used.

The results are obtained with the help of MATLAB. The formulas and parameters given in Chapter 3 are coded and entered to MATLAB and calculations are done for different cases and the optimum result is obtained. Moreover, by using GUI option of MATLAB, a user interface is designed. It is a decision support tool, and depending on the choice of the train type made by the user, it automatically calculates the inspection sequence and sample size. An image of this user interface can be seen in Appendix F. In order use this interface, MATLAB software has to be installed.

Finally, the parameters (e.g. failure replacement cost, inspection cost, preventive replacement cost, alarm threshold etc.) are fixed values in the program. If a change is needed, this has to done in the main code.

6.3 Recommendations

In this section some discussions about the research is given and some recommendations for future work are presented.

To begin with, in data analysis part of the project, there is some limitations about the existence of data, for instance for MDDM type of train, only sixteen data points are used. The analysis can be improved if it is done by using more data. Moreover, as it is mentioned in Chapter 4, when multiple regression analysis is conducted, it is seen that the model which represents the metal accumulation in oil is improved if additional information about the age of gearbox is added to it. Hence, the age of each gearbox can be recorded and used in the analysis to improve the understanding related to the contamination process of oil. Moreover, during the data analysis, it is observed that even if three months passed from the previous inspection, the parts in oil can decrease, which is something not logical. Then, the reason for this is investigated, and expert opinion is taken, the reason for this decreases caused by letting the train wait in the maintenance depot too long, so the contaminating parts in oil sink and when the oil samples are taken they will contain less parts. Hence, the samples should be taken as soon as possible when the train comes to maintenance depot.

The oil analyses are conducted by DeltaRail. The samples are taken at NedTrain's maintenance depot and sent to DeltaRail. After the analyses are done, DeltaRail informs Nedtrain about the conclusion, whether the oil is in a good condition or not. However, in order to apply condition based maintenance policy efficiently, NedTrain has to have more information about the condition of the oil. Knowing whether the condition is good or bad is not enough to do prediction for future.

Condition based maintenance is the new trend in maintenance, and then it should be considered to be applied for other parts. First of all, the degradation process of the related equipment has to be understood clearly; for instance if the failure rate decreases with time, then CBM may not be the proper strategy. The control parameters have to be decided very carefully, it should represent the condition of equipment. Then, the monitoring strategy has to be determined; the condition of the equipment can be monitored continuously or with discrete time inspections, this depends on the importance of the equipment, the cost of monitoring operation and characteristics of the equipment. The indicator of the failure has to be identified; there should be some values for each control parameters that if they are above these values, it will be a sign for bad condition of the equipment. If discrete time condition monitoring is used, then the correct times for monitoring operations have to be determined. These monitoring times can either be periodic or not, for instance if it is periodic then inspections are done in equidistant times, if it is aperiodic then inspections are done at the previously determined times which can be less frequent at the beginning of the lifetime of equipment and more frequent towards the expected lifetime. Finally, a strategy has to be identified for maintenance actions if a problem is observed related to the equipment.

After the data analysis, amount if iron parts in oil is chosen to be the informative one out of eight control parameter to represent the contamination of oil. This way is followed in the research because it is a frequently used in literature (Bonjevic et al. (2006)), and the available literature is dominated by one parameter models. Thus, for the future work, adding other control parameters to the model can provide a better representation of system and for academics the gap in the literature can be closed.

In order to answer the research questions, literature survey is conducted and the model suggested by Chelbi et al. (1999), which is explained in Chapter 3, is used. They developed their model to determine the optimum inspection sequence of a single part. However, the research question is finding optimum inspection sequence for a group of identical parts. Hence, some changes are made to existing model with the extension about optimum sample size for each inspection. Thus, this is a contribution to literature.

While answering the second research question, which is 'what is the optimal inspection times and sample size that gives the minimum maintenance cost for condition monitoring process?', it is divided into two parts. Firstly, the inspection sequence is determined, and then for each inspection the number of sample that is going to be inspected is decided. The optimum inspection sequence is determined by minimizing the expected total cost. This cost has three components and one of these components is the inspection cost. This inspection cost basically has two components; set up cost of the inspection process (which is a fixed cost) and total inspection cost of each item (which is a variable cost and increases in proportion with sample size). In this research, an average cost value is assigned as 'inspection cost' in the model without considering the effect of sample size and optimum inspection sequence is determined. Then, after finding the inspection sequence, the sample size is calculated. As a further research area, this two optimization problem can be combined. An algorithm can be generated as follows;

Step 1: Assume an average value for inspection cost and determining inspection sequence

Step 2: Determine the sample size for the inspection sequence found in Step 1

Step 3: Assign a new value for inspection cost which depends on the sample size found in Step 2 and solve the model again using this cost parameter and determine the new inspection sequence

Step 4: Repeat this until it converges to same number

In this research, the set up cost is ignored in calculations since the inspection times are determined before, and the main goal is finding the sample size for that inspection not deciding on whether to do inspection or not. However, if the two optimization problem is going to be combined, then the set up cost may be relevant and should not be ignored because at that time the decision will be about doing inspection or not at any day.

For the last part of the research question, which is determining the sample size for each inspection, a model is presented in Chapter 3. This model is based on the idea that, if the total monitoring cost is minimized, then optimum sample size can be found. While writing the total monitoring cost, a cost parameter, a, is defined as the cost related to the interval length of the estimated mean of iron accumulation. This represents the wish that one has to have a wide or narrow interval for the estimation, which is directly related to how certain one wants to be about the population. Hence, it is a very tentative cost. In the further research, a detailed analysis for this cost parameter can be conducted.

References

Abertnethy, R.B., (2006), "The New Weibull Handbook", 5th Edition, Dr. Robert. Abernethy.

Barlow, R.E., Hunter, L.C., Proschan, F., (1963), "Optimum checking procedures", *Journal of the Society Industrial and Applied Mathematics* 11, 1078-1095.

Baohe, S., (2002), "An optimal inspection and diagnosis policy for a multimode system", *Reliability Engineering and System Safety* 76, 181-188.

Bengtsson, M., (2007), "On condition based maintenance and its implementation in industrial settings", *Malardalen University Press Dissertation* No.48

Berenguer, C., Chu, C., Grall, A., (1997), "Maintenance inspection planning: an application of semi-Markov decision process" *Journal of Intelligent Manufacturing* 8, 467-476

Bonjevic, D., Jardine, A.K.S., (2006), "Calculation of reliability function and remaining useful life for a Markov failure time process", *IMA Journal of Management Mathematics* 17, 115-130.

Castainer, B., Bérenguer. C., Grall, A., (2003), "A sequential condition based repair/replacement policy with non periodic inspections for a system subject to continuous wear", *Applied Stochastic Models in Business and Industry* 19, 327-347.

Chelbi, A., Ait-Kadi, D., (1998), "Inspection and predictive maintenance strategies", *International Journal of Computer Integrated Manufacturing* Vol.11, 226-231.

Chelbi, A., Ait-Kadi, D., (1999), An optimal inspection strategy for randomly failing equipment, *Reliability Engineering and Systems Safety* 63, 127-131.

Chen, D., Trivedi. K.S., (2001), "Closed-form analytical results for condition-based maintenance", *Reliability Engineering and Systems Safety* 39, 43-51.

Christer, A.H., Wang, W., (1995), "A simple condition monitoring model for a direct monitoring process", *European Journal of Operational Research* 82, 258-269.

Dekker, R., (1996), “Application of maintenance optimization models: a review and analysis”, *Reliability Engineering and Systems Safety* 51, 229-240.

Dieulle, L., Bérenguer, C., Grall, A., Roussignol, M., (2001), “Continuous predictive maintenance scheduling for a deteriorating system”, *Annual Reliability and Maintainability Symposium.*

Dieulle, L., Bérenguer, C., Grall, A., Roussignol, M., (2003), “Sequential condition based maintenance scheduling for a deteriorating system”, *European Journal of Operational Research* 150, 451-461.

Evers R., (2010), “Algorithms for Scheduling Train Maintenance”, Master Thesis, Delft University of Technology.

Gabraeel, N., Elwany, A., (2009), “Residual life predictions in the absence of prior degradation knowledge”, *IEEE Transactions on Reliability* Vol.58, No.1.

Grall, A., Berenguer, C., Dieulle, L., (2002), “A condition-based maintenance policy for stochastically deteriorating systems”, *Reliability Engineering and System Safety* 76, 167-180.

Hair, J.F., Black, B., Babin, B., Anderson R.E., (2009), “Multivariate Data Analysis”, 7th Edition, Prentice Hall.

http://www.railway-technical.com/rstock.shtml, last visited on March 2012.

http://en.wikipedia.org/wiki/Sampling_(statistics), last visited on April 2012.

Irobi, I.S., Andersson, J., Wall, A., (2004), “Correctness criteria for models’ validation-A philosophical perspective”, *In Proceeding of MSV/AMCS’* 2004, pp 475-476.

Kaio, N., Osaki, S., (1984), “Some remarks on optimum inspection policies”, *IEEE Transaction on Reliability* Vol.R-33 N.4, 277-279.

Keller, J.B., (1974), “Optimum checking schedules for systems subject to random failure”, *Management Science* Vol.21 N.3, 256-260.

Kusters, R.A.M., (2011), “The Design of a Logistic Support System”, Master Thesis, Eindhoven University of Technology.

Leung, F.K., (2001), “Inspection schedules when lifetime distribution is unknown”, *European journal of Operational Research* 132, 106-115.

Lu, C.J., Meeker, W.Q., (1993), “Using degradation measures to estimate a time to failure distribution”, *Technometrics* Vol.35, 161-174.

Munford, A.G., Shahani, A.K., (1973), “An inspection policy for the Weibull case”, Operational Research Quarterly, Vol.24, 453-458.

Munford, A.G., Shahani, A.K., (1972), “A nearly optimal inspection policy”, *Operational Research Quarterly*, Vol. 23, 373-379.

Nakagawa, T., Yasui, K., (1980), “Approximate calculation of optimal inspection times”, *Journal of the Operational Research Society* Vol.31 551-553.

Neves, M. L., Santiago, L. P., Maia, C. A., (2011), “A condition-based policy and input parameter estimation for deteriorating systems under periodic inspection”, *Computers & Industrial Engineering* Article in press.

Ohnishi, M., Kawai, H., Mine, H., (1986), “An optimal inspection and replacement policy under incomplete state information”, *European Journal of Operational Research* 27, 117-128.

Ozkan E., (2010), “New Control Mechanisms for the Repairables in NedTrain”, Master Thesis, Eindhoven University of Technology.

Page, R. (2002), “Maintenance Management and Delay Reduction”, *Maintenance and Asset Management*, 17, pp. 5-13.

Parida, A. and Kumar, U., (2006), “Maintenance Performance Management: issues and challenges”, *Journal of Quality in Maintenance Engineering* Vol. 12 No. 3, pp 239-251.

Tadikamalla, P.R., (1979), “An inspection policy for the Gamma failure distribution”, *Journal of the Operational Research Society* 30, 77-80.

Tsang, A.H.C., (1995), “Conditional-based maintenance: tools and decision making”, *Journal of Quality in Maintenance Engineering* Vol.1 No.3, pp 3-17.

Van Aken, J.E., Berends, J.J., & Van der Bij, J.D. (2007), “Problem Solving in Organizations”, Cambridge: Cambridge University Press.

Van Strien, P.J. (1997), “Towards a Methodology of Psychological Practice”, *Theory and Psychology*, Vol. 7 N.5, 683-700.

Vernooij K., (2011), “An Aggregate Planning for Preventive Maintenance of Bogies by NedTrain”, Master Thesis, Eindhoven University of Technology.

Wang, L., Chu, J., Wu, J., (2007), “Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process”, *International Journal of Production Economics* 107, 151-163.

Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., (2002), “Probability & Statistics for Engineers & Scientists”, 7th Edition, Prentice Hall.

Wang, W., (2000), “A model to determine the optimal critical level and the monitoring intervals in condition based maintenance”, *International Journal of Production Research* Vol.38 N.6, 1425-1436.

Wang, W., (2003), “Modeling condition monitoring intervals: a hybrid of simulation and analytical approaches”, *Journal of the Operational Research Society* 54, 273-282

Zhao Z., Wang F., Jia M., Wang S., (2010), “Predictive maintenance policy based on process data”, *Chemometrics and Intelligent Laboratory Systems* Volume 103, Iss 2, pp 137-143.

APPENDICES

Appendix A: Data points computed for ICM 1-2 from main data set

Table A.1: Time (days) passed and corresponding metal accumulation (ppm) for ICM 1-2

time(days)	FeSNEL	XCu	XPb	XSn	XCr	XNi	XAl	XSi
80	110	25	2	0	1	0	3	10
83	80	10	5	3	1	0	5	10
83	60	10	4	2	0	0	0	0
83	95	10	1	0	0	0	2	35
83	40	5	1	0	1	0	0	35
84	85	10	0	0	0	0	0	60
85	65	15	0	0	1	0	0	5
85	40	10	0	0	1	0	0	10
86	65	20	0	0	1	0	0	15
87	65	20	0	0	1	0	5	2
87	105	30	3	0	1	0	0	10
88	90	15	5	1	0	0	1	0
89	95	15	2	0	1	0	0	30
90	160	10	3	0	1	0	10	65
90	105	15	5	0	1	0	10	35
90	160	15	0	1	1	0	4	45
90	70	4	1	1	0	0	3	10
90	265	5	2	0	1	0	2	210
90	165	5	2	0	1	0	3	20
91	85	15	3	2	0	0	0	5
91	105	25	3	1	1	0	1	10
91	125	35	5	0	1	0	1	20
91	85	15	3	0	1	0	5	15
91	215	30	1	0	1	0	0	15
91	90	15	3	0	0	0	0	10
91	70	10	2	2	1	0	3	0
91	115	5	2	4	1	0	10	15
91	25	5	3	1	0	0	5	5
91	65	5	0	0	0	0	0	5
92	85	20	0	2	1	0	0	195
92	115	15	3	0	1	0	0	20
92	140	4	0	0	1	0	1	45

92	60	4	1	0	1	0	0	20
93	185	15	2	0	2	0	5	10
93	195	20	0	1	2	0	4	25
93	85	10	2	0	1	0	15	30
93	195	30	1	2	3	0	3	15
93	50	3	3	0	1	0	0	2
93	215	4	3	0	1	0	10	20
93	120	20	3	4	1	0	0	15
94	20	10	1	0	0	0	3	4
94	70	10	5	0	1	0	0	50
95	45	10	1	4	0	0	0	0
95	95	35	0	0	1	0	0	10
95	65	5	4	1	0	0	0	20
96	85	15	5	1	1	0	5	20
97	110	15	1	0	1	0	3	10
97	80	35	1	0	1	0	5	3
98	60	20	1	0	0	0	1	40
100	25	15	3	0	1	0	0	0
101	75	15	4	0	0	0	0	15
103	90	15	5	1	1	0	10	50
103	90	5	3	2	1	0	3	30
108	105	20	1	1	1	0	4	20
123	100	20	0	0	1	0	0	35
130	55	25	0	0	0	0	5	10
138	125	35	1	0	1	0	0	65
139	135	40	0	0	2	0	0	20
143	100	20	0	0	0	0	10	15
162	45	10	0	0	0	0	0	5
164	115	20	0	0	1	0	0	20
165	100	5	0	0	0	0	0	260
169	75	10	0	0	1	0	0	10
169	110	5	0	0	1	0	1	75
172	80	20	1	0	1	0	0	15
175	105	15	0	0	1	0	4	30
176	230	20	0	0	1	0	4	55
176	75	15	0	0	1	0	1	5
176	255	20	0	0	1	0	10	20
178	190	25	0	0	1	0	15	30
178	130	5	0	4	1	0	0	40
178	40	5	0	0	0	0	0	3
179	200	20	0	0	2	0	0	25

179	120	15	5	0	1	0	0	3
180	150	15	1	0	1	0	10	35
180	105	40	1	0	1	0	10	35
181	60	20	1	0	1	0	4	30
181	365	5	5	2	1	0	0	165
181	55	5	0	0	1	0	0	10
183	50	10	3	3	0	0	0	4
183	165	30	2	0	0	0	3	15
183	245	4	4	0	1	0	0	35
183	270	20	4	3	1	0	5	50
184	100	25	0	0	0	0	0	20
184	125	15	1	1	1	0	0	25
184	50	15	0	0	0	0	5	10
184	50	15	4	2	1	0	1	1
187	80	10	0	0	1	0	0	3
188	110	20	0	0	1	0	10	30
189	80	15	0	0	0	0	0	50
190	130	15	3	0	0	0	0	20
190	70	15	3	1	0	0	0	0
190	75	15	2	1	0	0	0	0
192	40	10	3	0	0	0	0	5
192	50	15	2	3	0	0	0	5
192	145	15	0	0	1	0	0	50
192	50	5	1	2	0	0	0	0
198	120	15	0	0	1	0	0	3
198	135	5	0	0	1	0	0	15
199	75	40	1	0	1	0	0	10
202	135	10	0	2	1	0	0	35
202	155	15	1	4	1	0	0	40
204	235	25	1	1	1	0	3	30
205	50	15	1	0	1	0	4	10
211	45	5	1	0	0	0	0	10
212	20	5	0	0	0	0	2	3
220	85	20	0	0	1	0	5	5
226	70	15	1	0	1	0	2	15
248	120	25	0	0	1	0	65	50
254	310	20	0	0	1	0	0	40
254	75	15	0	0	0	0	0	15
254	315	20	1	1	2	0	4	35
255	165	20	0	0	1	0	0	20
263	90	25	0	0	0	0	0	5

266	50	15	0	1	0	0	0	15
267	50	15	4	1	1	0	4	0
268	415	25	2	1	3	1	2	25
270	55	10	0	0	1	0	5	4
270	435	5	2	0	2	1	5	195
274	545	25	2	2	2	0	3	45
274	190	15	3	1	1	1	4	15
275	255	25	5	1	1	0	10	90
275	245	30	2	1	2	0	10	15
281	360	30	3	0	2	0	5	55
282	90	15	3	0	0	0	4	10
282	140	5	3	3	1	1	5	15
283	145	15	3	0	0	0	4	25
354	115	20	1	1	0	0	0	10
359	85	20	0	0	1	0	0	5
381	115	4	3	1	1	0	10	10
444	145	20	3	0	2	0	5	10
466	80	5	1	1	1	0	1	15

Appendix B: Outliers for ICM 1-2

Table B.1: Multivariate outliers for ICM 1-2

time(days)	FeSNEL	XCu	XPb	XSn	XCr	XNi	XAl	XSi
92	85	20	0	2	1	0	0	195
165	100	5	0	0	0	0	0	260
248	120	25	0	0	1	0	65	50
268	415	25	2	1	3	1	2	25
270	435	5	2	0	2	1	5	195
274	545	25	2	2	2	0	3	45
274	190	15	3	1	1	1	4	15
282	140	5	3	3	1	1	5	15

Table B.2: Univariate outliers for ICM 1-2

time	XSn	XCr	Xsi
178	4		
95	4		
202	4		

91	4		
93	4	3	
181			165
90			210

Appendix C: Relations between time and metal particles

ICM 1-2

Iron

Model Description

Model Name		MOD_1
Dependent Variable	1	fe
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		time
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

a. The model requires all non-missing values to be positive.

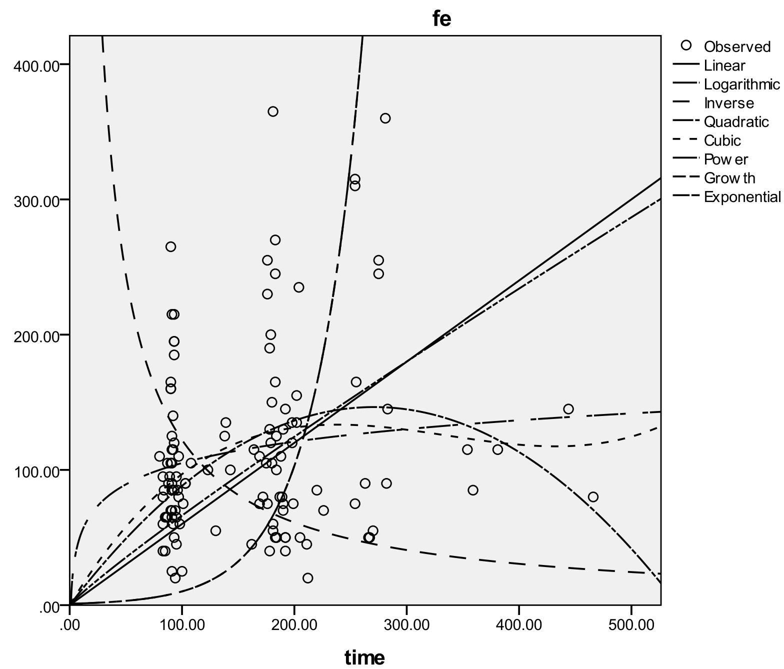
Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.652	230.393	1	123	.000	.600		
Logarithmic	.735	341.472	1	123	.000	22.814		
Inverse	.562	158.093	1	123	.000	12220.773		
Quadratic	.728	163.411	2	122	.000	1.081	-.002	
Cubic	.733	110.890	3	121	.000	1.369	-.004	4.415E-6
Power	.980	5916.902	1	123	.000	.911		
Growth	.813	536.012	1	123	.000	.023		

Exponential	.813	536.012	1	123	.000	.023		

The independent variable is time.



Cupper

Model Description

Model Name		MOD_2
Dependent Variable	1	cu
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timecu
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

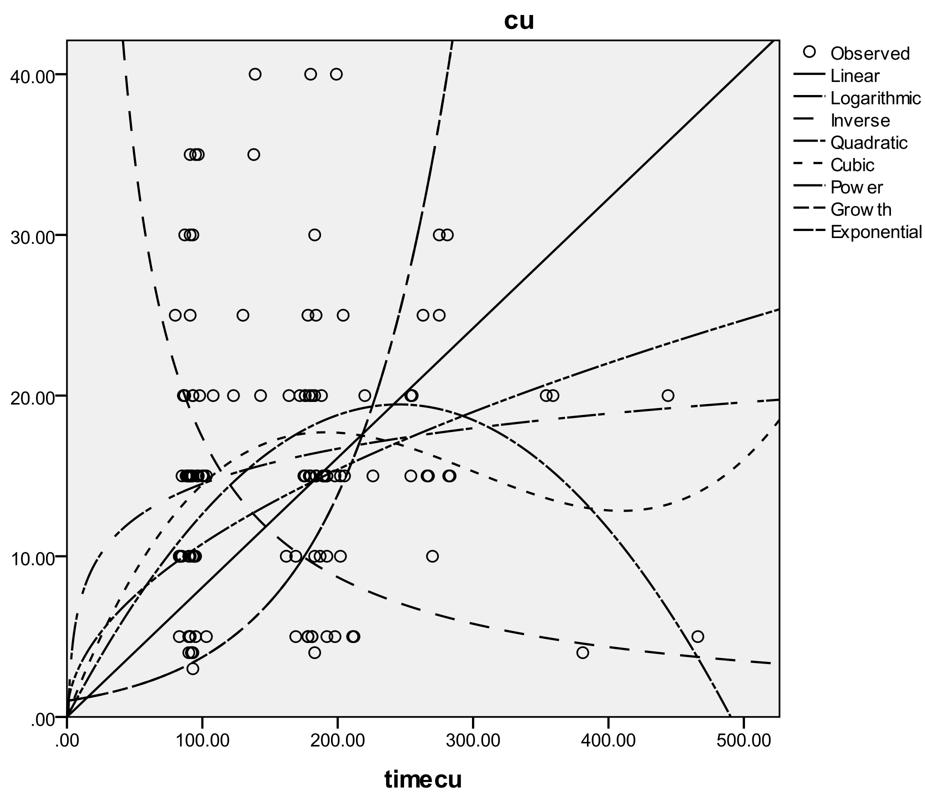
a. The model requires all non-missing values to be positive.

Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.646	224.656	1	123	.000	.081		
Logarithmic	.770	410.887	1	123	.000	3.151		
Inverse	.626	205.507	1	123	.000	1739.789		
Quadratic	.756	189.098	2	122	.000	.159	.000	
Cubic	.769	134.033	3	121	.000	.220	-.001	9.356E-7
Power	.942	1996.229	1	123	.000	.516		
Growth	.782	441.210	1	123	.000	.013		
Exponential	.782	441.210	1	123	.000	.013		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Description

Model Name		MOD_4
Dependent Variable	1	pb
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timepb
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

a. The model requires all non-missing values to be positive.

Model Summary and Parameter Estimates

Dependent Variable:pb

	Model Summary					Parameter Estimates		
Equation	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.348	65.772	1	123	.000	.007		
Logarithmic	.472	110.173	1	123	.000	.312		
Inverse	.508	127.249	1	123	.000	199.723		
Quadratic	.415	43.213	2	122	.000	.015	-3.183E-5	
Cubic	.485	38.000	3	121	.000	.033	.000	2.787E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

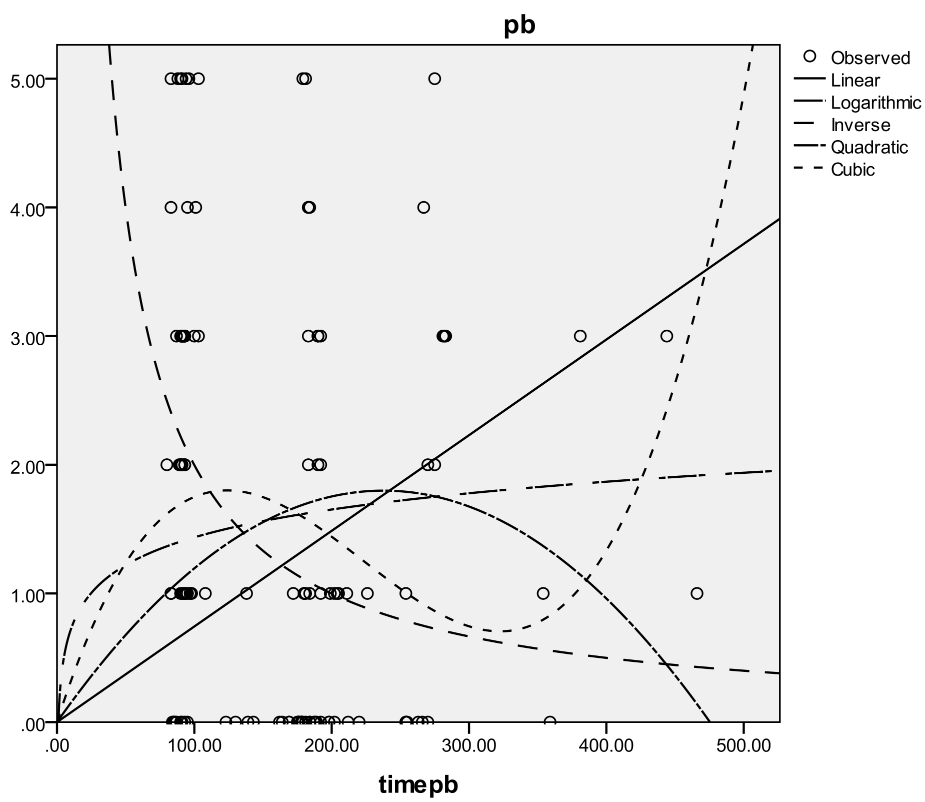
Model Summary and Parameter Estimates

Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.348	65.772	1	123	.000	.007		
Logarithmic	.472	110.173	1	123	.000	.312		
Inverse	.508	127.249	1	123	.000	199.723		
Quadratic	.415	43.213	2	122	.000	.015	-3.183E-5	
Cubic	.485	38.000	3	121	.000	.033	.000	2.787E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Description

Model Name		MOD_5
Dependent Variable	1	sn
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timesn
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

a. The model requires all non-missing values to be positive.

Model Summary and Parameter Estimates

Dependent Variable:sn

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.204	30.197	1	118	.000	.002		
Logarithmic	.241	37.434	1	118	.000	.087		
Inverse	.205	30.427	1	118	.000	49.515		
Quadratic	.228	17.249	2	117	.000	.004	-7.393E-6	
Cubic	.240	12.215	3	116	.000	.007	-3.274E-5	4.551E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

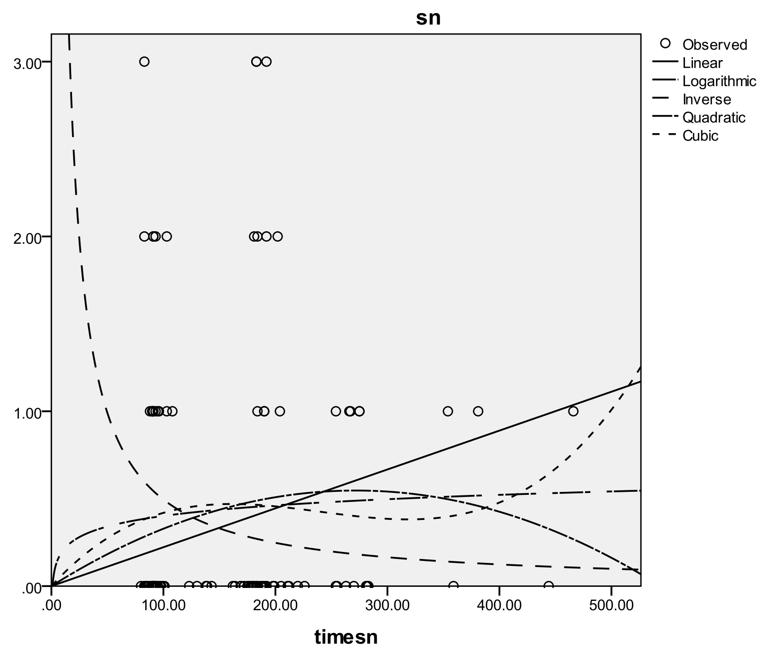
Model Summary and Parameter Estimates

Dependent Variable:sn

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.204	30.197	1	118	.000	.002		
Logarithmic	.241	37.434	1	118	.000	.087		
Inverse	.205	30.427	1	118	.000	49.515		
Quadratic	.228	17.249	2	117	.000	.004	-7.393E-6	
Cubic	.240	12.215	3	116	.000	.007	-3.274E-5	4.551E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Description

Model Name		MOD_6
Dependent Variable	1	cr
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timecr
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

a. The model requires all non-missing values to be positive.

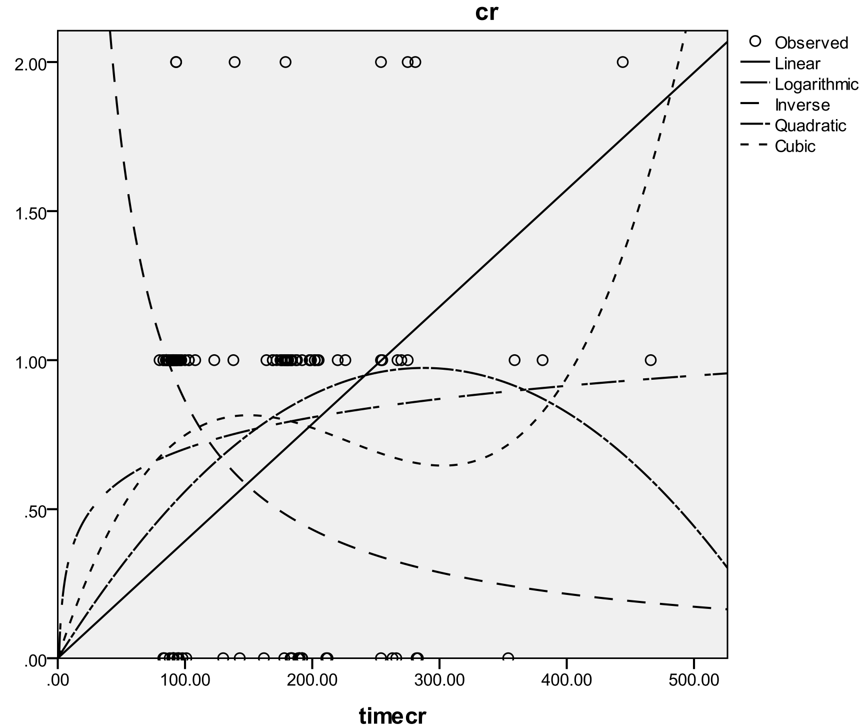
Model Summary and Parameter Estimates

Dependent Variable:cr

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.556	153.037	1	122	.000	.004		
Logarithmic	.651	227.681	1	122	.000	.153		
Inverse	.552	150.604	1	122	.000	86.366		
Quadratic	.609	94.259	2	121	.000	.007	-1.177E-5	
Cubic	.656	76.375	3	120	.000	.013	-6.479E-5	9.539E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Description

Model Name		MOD_7
Dependent Variable	1	al
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timeal
Constant		Not included

Variable Whose Values Label Observations in Plots	Unspecified
Tolerance for Entering Terms in Equations	.0001

a. The model requires all non-missing values to be positive.

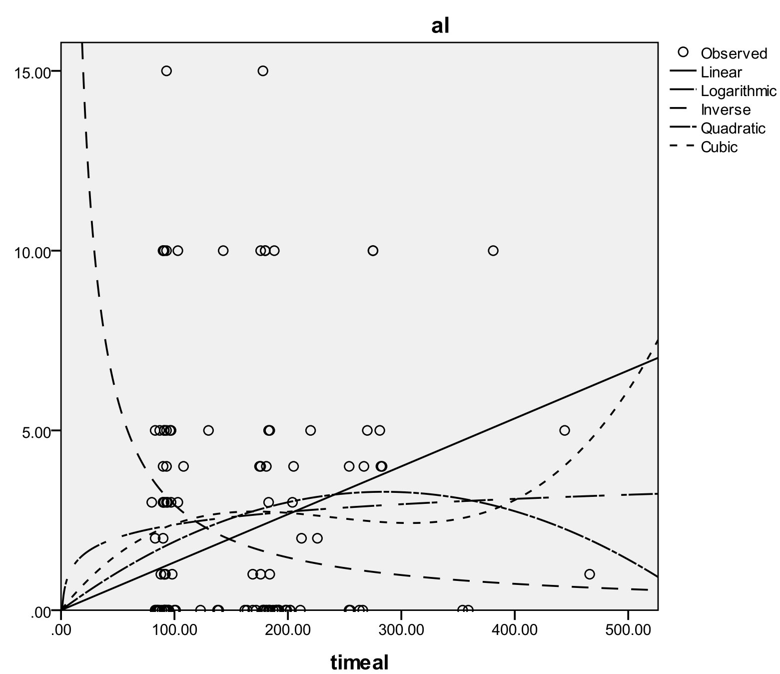
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.295	51.577	1	123	.000	.013		
Logarithmic	.346	65.105	1	123	.000	.517		
Inverse	.297	51.942	1	123	.000	293.113		
Quadratic	.324	29.289	2	122	.000	.023	-4.062E-5	
Cubic	.340	20.732	3	121	.000	.039	.000	2.507E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Description

Model Name		MOD_8
Dependent Variable	1	si
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power[a]
	7	Growth[a]
	8	Exponential[a]
Independent Variable		timesi
Constant		Not included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

a. The model requires all non-missing values to be positive.

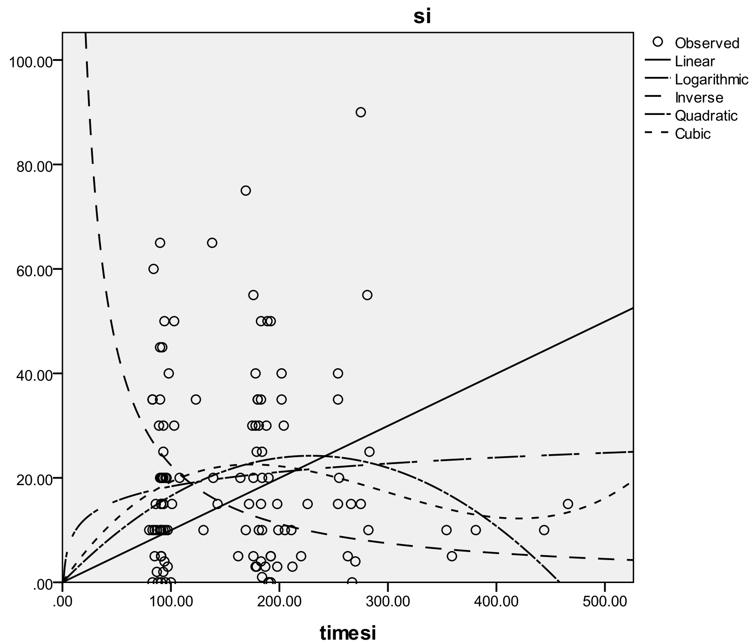
Model Summary and Parameter Estimates

Dependent Variable:si

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.449	98.795	1	121	.000	.100		
Logarithmic	.557	151.887	1	121	.000	3.987		
Inverse	.467	105.991	1	121	.000	2240.644		
Quadratic	.551	73.698	2	120	.000	.211	.000	
Cubic	.564	51.223	3	119	.000	.301	-.001	1.374E-6
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



ICM 3

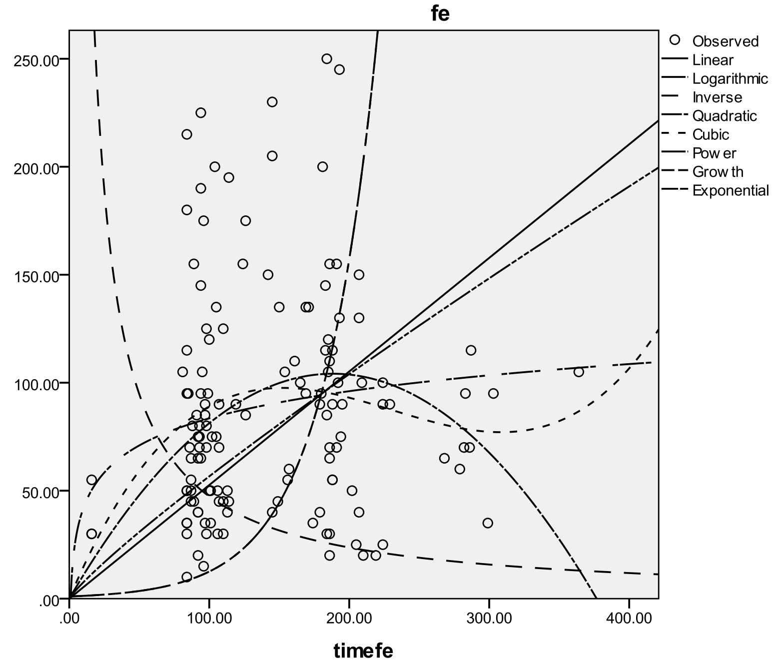
Iron

Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.643	242.622	1	135	.000	.526		
Logarithmic	.742	388.450	1	135	.000	18.131		
Inverse	.328	66.040	1	135	.000	4736.217		
Quadratic	.740	190.859	2	134	.000	1.106	-.003	
Cubic	.750	133.178	3	133	.000	1.531	-.007	1.084E-5
Power	.973	4917.955	1	135	.000	.877		
Growth	.833	675.297	1	135	.000	.025		
Exponential	.833	675.297	1	135	.000	.025		

The independent variable is timefe.



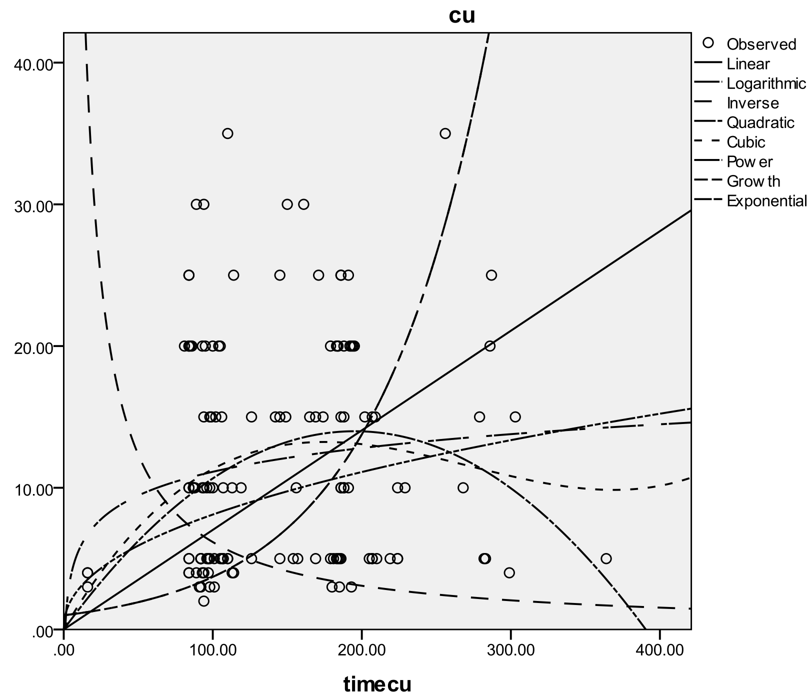
Cupper

Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.608	209.468	1	135	.000	.070		
Logarithmic	.690	300.912	1	135	.000	2.416		
Inverse	.291	55.326	1	135	.000	617.108		
Quadratic	.688	148.068	2	134	.000	.143	.000	
Cubic	.692	99.768	3	133	.000	.180	-.001	9.280E-7
Power	.908	1339.490	1	135	.000	.454		
Growth	.788	500.748	1	135	.000	.013		
Exponential	.788	500.748	1	135	.000	.013		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Summary and Parameter Estimates

Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.269	50.134	1	136	.000	.006		
Logarithmic	.430	102.474	1	136	.000	.256		
Inverse	.359	76.332	1	136	.000	92.213		
Quadratic	.453	55.920	2	135	.000	.021	-7.446E-5	
Cubic	.517	47.856	3	134	.000	.041	.000	5.079E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

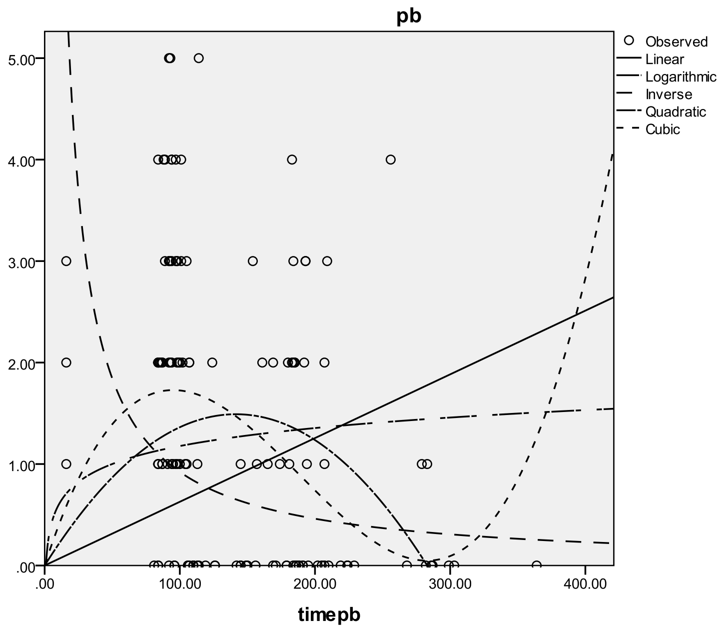
Model Summary and Parameter Estimates

Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.269	50.134	1	136	.000	.006		
Logarithmic	.430	102.474	1	136	.000	.256		
Inverse	.359	76.332	1	136	.000	92.213		
Quadratic	.453	55.920	2	135	.000	.021	-7.446E-5	
Cubic	.517	47.856	3	134	.000	.041	.000	5.079E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

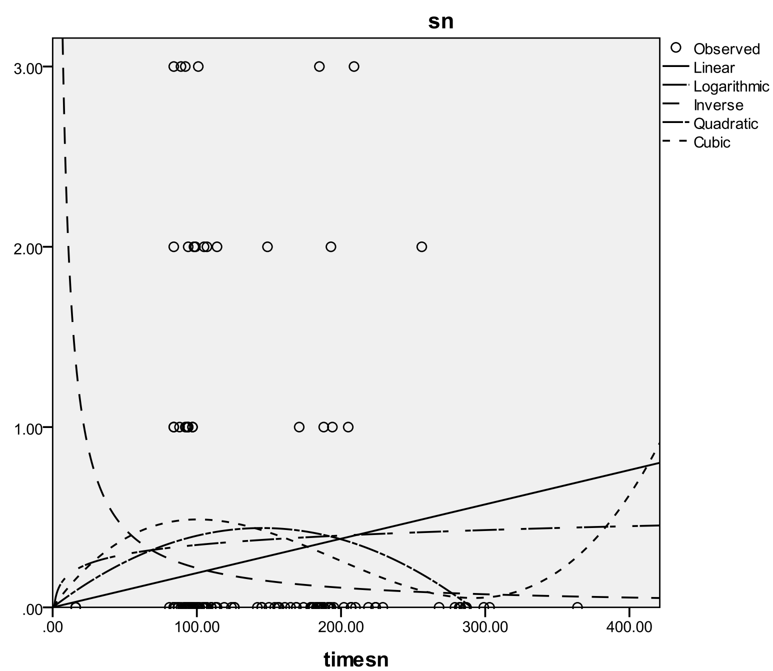
Model Summary and Parameter Estimates

Dependent Variable:sn

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.115	17.403	1	134	.000	.002		
Logarithmic	.171	27.654	1	134	.000	.075		
Inverse	.093	13.687	1	134	.000	21.804		
Quadratic	.181	14.732	2	133	.000	.006	-2.085E-5	
Cubic	.199	10.935	3	132	.000	.011	-7.316E-5	1.242E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

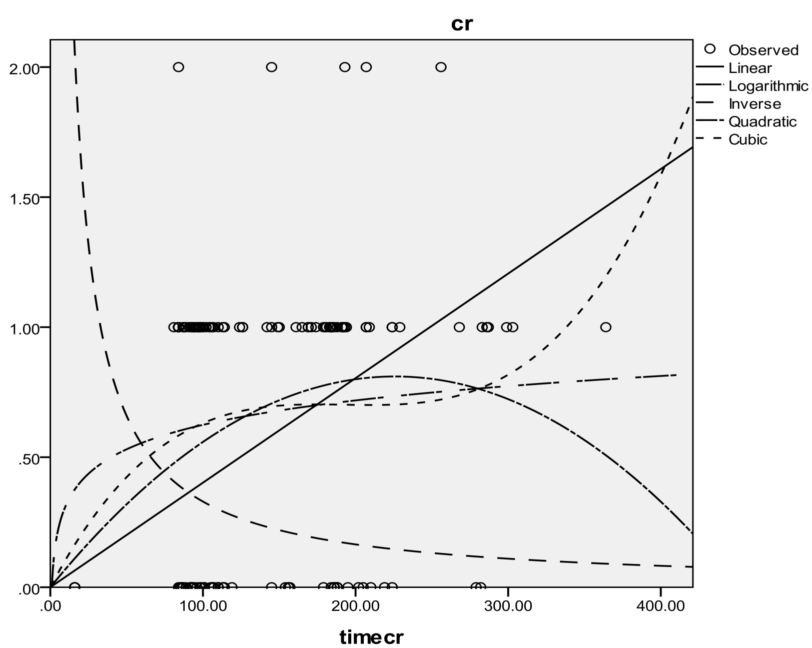
Model Summary and Parameter Estimates

Dependent Variable:cr

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.551	166.771	1	136	.000	.004		
Logarithmic	.603	206.535	1	136	.000	.136		
Inverse	.229	40.388	1	136	.000	32.945		
Quadratic	.593	98.154	2	135	.000	.007	-1.588E-5	
Cubic	.605	68.464	3	134	.000	.011	-5.831E-5	1.010E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

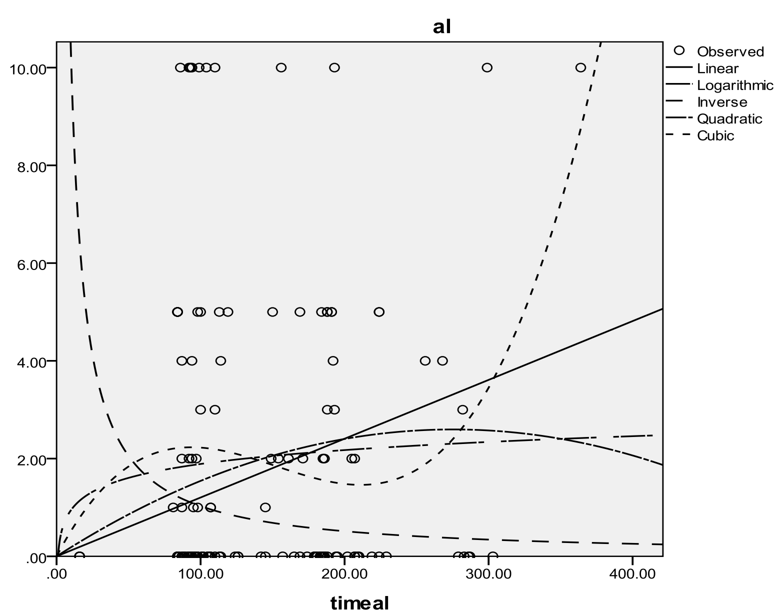
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.280	52.157	1	134	.000	.012		
Logarithmic	.312	60.736	1	134	.000	.410		
Inverse	.127	19.426	1	134	.000	102.743		
Quadratic	.291	27.318	2	133	.000	.019	-3.425E-5	
Cubic	.355	24.170	3	132	.000	.056	.000	9.465E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

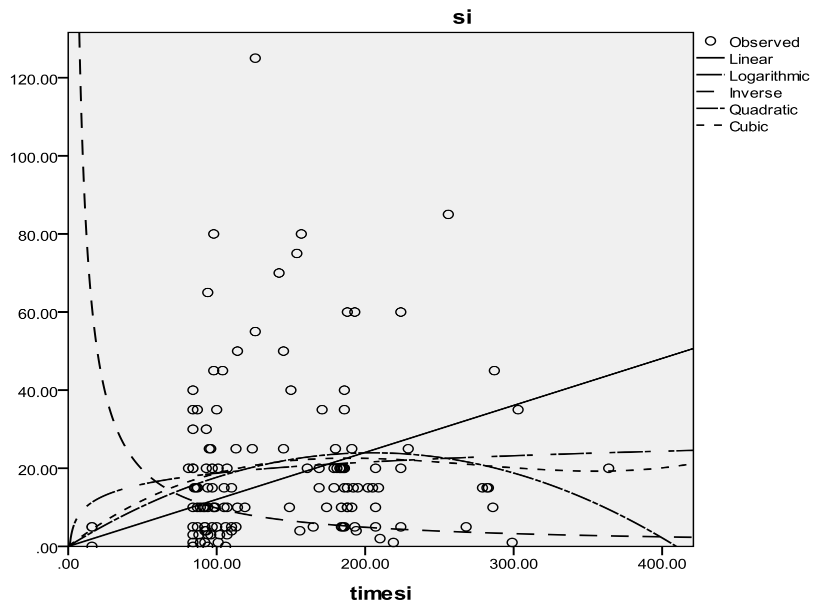
Model Summary and Parameter Estimates

Dependent Variable:si

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.449	110.008	1	135	.000	.120		
Logarithmic	.493	131.491	1	135	.000	4.069		
Inverse	.185	30.699	1	135	.000	980.595		
Quadratic	.498	66.536	2	134	.000	.234	-.001	
Cubic	.501	44.444	3	133	.000	.290	-.001	1.441E-6
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



MDDM

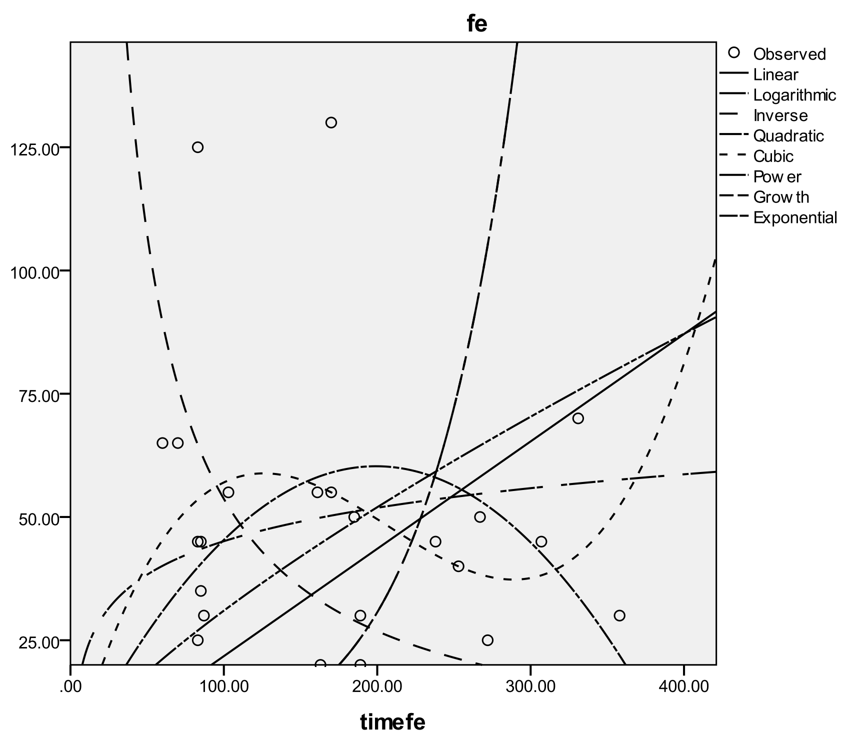
Iron

Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.549	26.786	1	22	.000	.218		
Logarithmic	.742	63.226	1	22	.000	9.791		
Inverse	.663	43.221	1	22	.000	5367.500		
Quadratic	.715	26.379	2	21	.000	.606	-.002	
Cubic	.766	21.818	3	20	.000	1.092	-.006	9.994E-6
Power	.968	660.728	1	22	.000	.746		
Growth	.763	70.773	1	22	.000	.017		
Exponential	.763	70.773	1	22	.000	.017		

The independent variable is timefe.



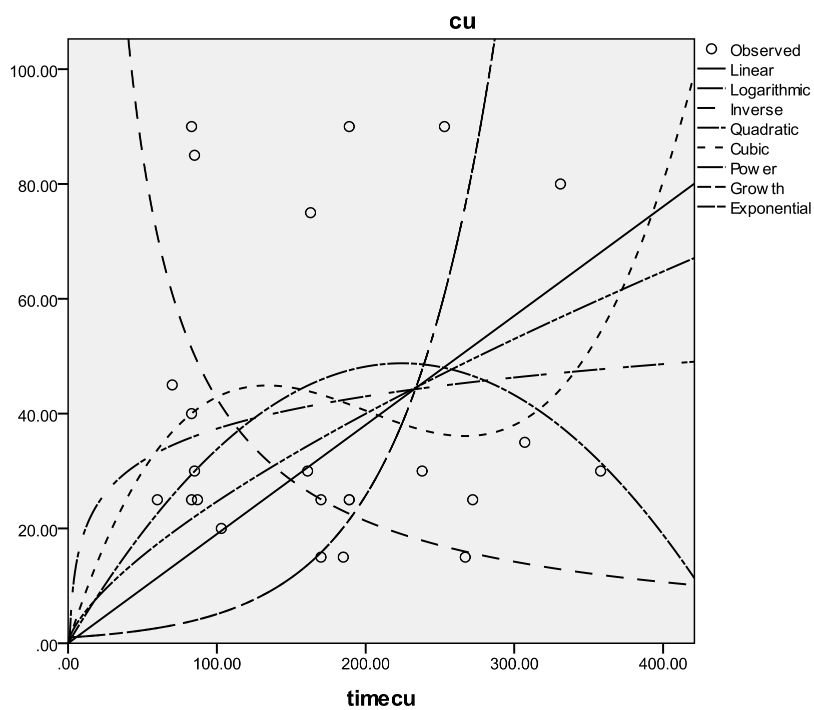
Cupper

Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.574	30.936	1	23	.000	.190		
Logarithmic	.703	54.315	1	23	.000	8.116		
Inverse	.559	29.159	1	23	.000	4262.715		
Quadratic	.665	21.814	2	22	.000	.435	-.001	
Cubic	.703	16.578	3	21	.000	.805	-.005	7.512E-6
Power	.961	571.955	1	23	.000	.696		
Growth	.778	80.452	1	23	.000	.016		
Exponential	.778	80.452	1	23	.000	.016		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

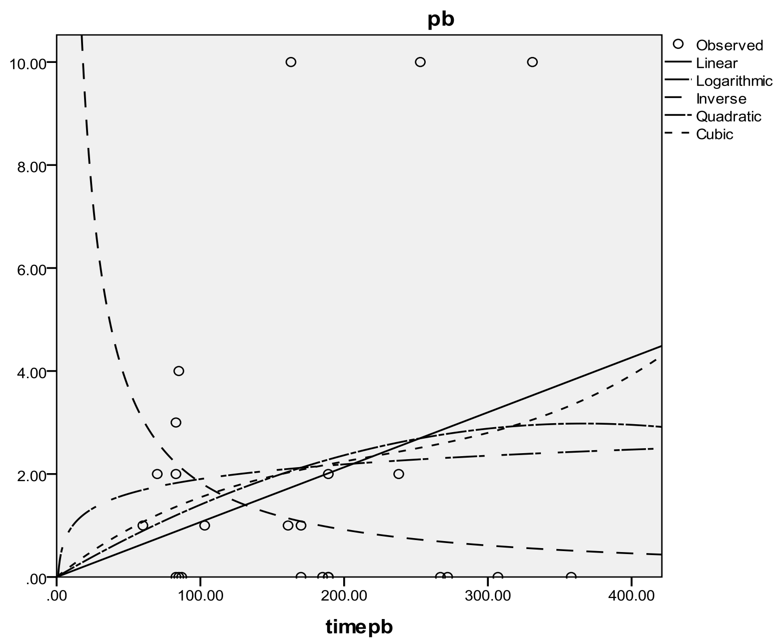
Model Summary and Parameter Estimates

Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.300	9.859	1	23	.005	.011		
Logarithmic	.303	9.992	1	23	.004	.413		
Inverse	.173	4.798	1	23	.039	183.542		
Quadratic	.308	4.896	2	22	.017	.016	-2.216E-5	
Cubic	.310	3.139	3	21	.047	.022	-7.793E-5	1.181E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

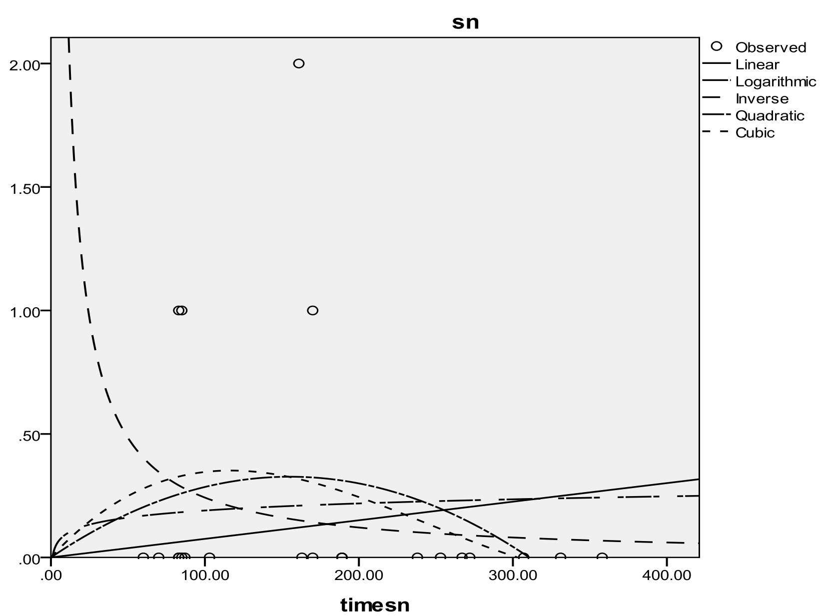
Model Summary and Parameter Estimates

Dependent Variable:sn

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.071	1.680	1	22	.208	.001		
Logarithmic	.142	3.654	1	22	.069	.041		
Inverse	.146	3.755	1	22	.066	24.229		
Quadratic	.213	2.845	2	21	.081	.004	-1.355E-5	
Cubic	.228	1.969	3	20	.151	.007	-3.795E-5	5.187E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

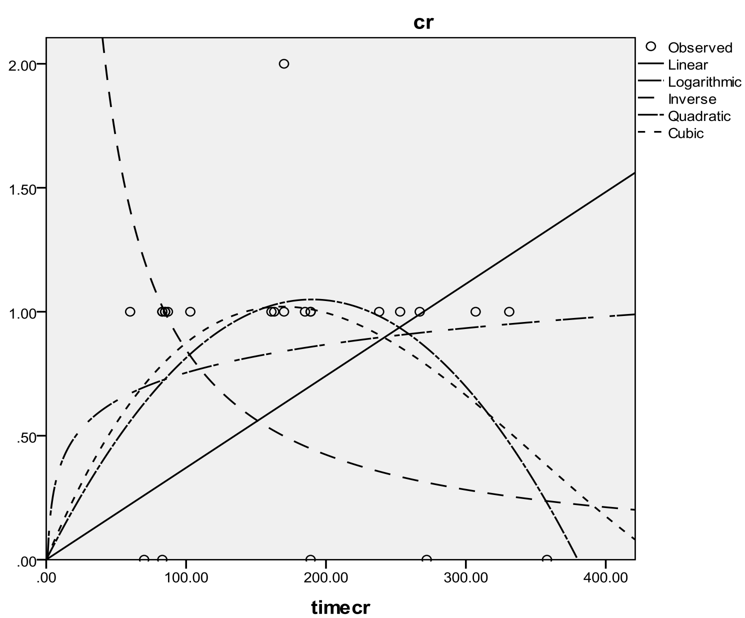
Model Summary and Parameter Estimates

Dependent Variable:cr

	Model Summary					Parameter Estimates		
Equation	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.569	30.420	1	23	.000	.004		
Logarithmic	.747	67.792	1	23	.000	.164		
Inverse	.576	31.240	1	23	.000	84.669		
Quadratic	.784	39.949	2	22	.000	.011	-2.907E-5	
Cubic	.788	26.096	3	21	.000	.014	-5.257E-5	4.978E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

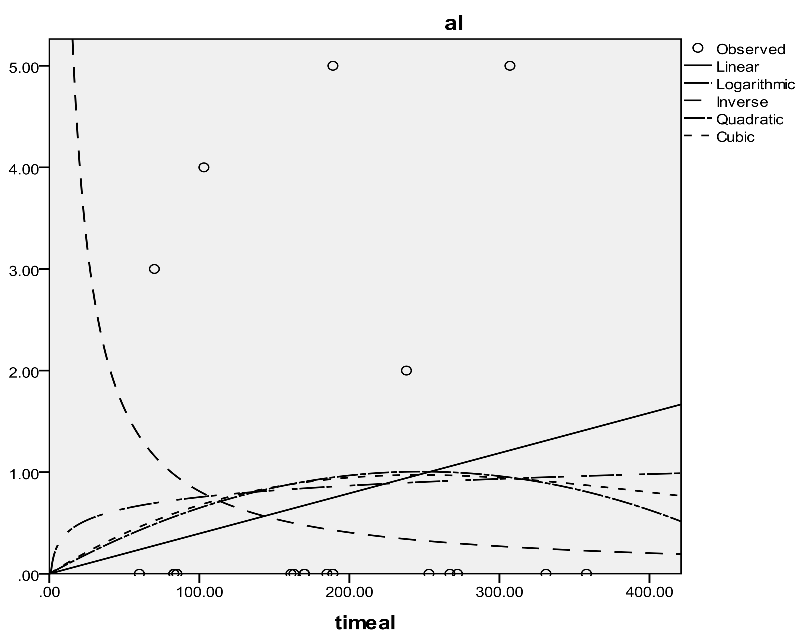
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.179	4.806	1	22	.039	.004		
Logarithmic	.201	5.542	1	22	.028	.164		
Inverse	.137	3.480	1	22	.076	81.226		
Quadratic	.198	2.584	2	21	.099	.008	-1.636E-5	
Cubic	.198	1.643	3	20	.211	.009	-2.674E-5	2.181E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

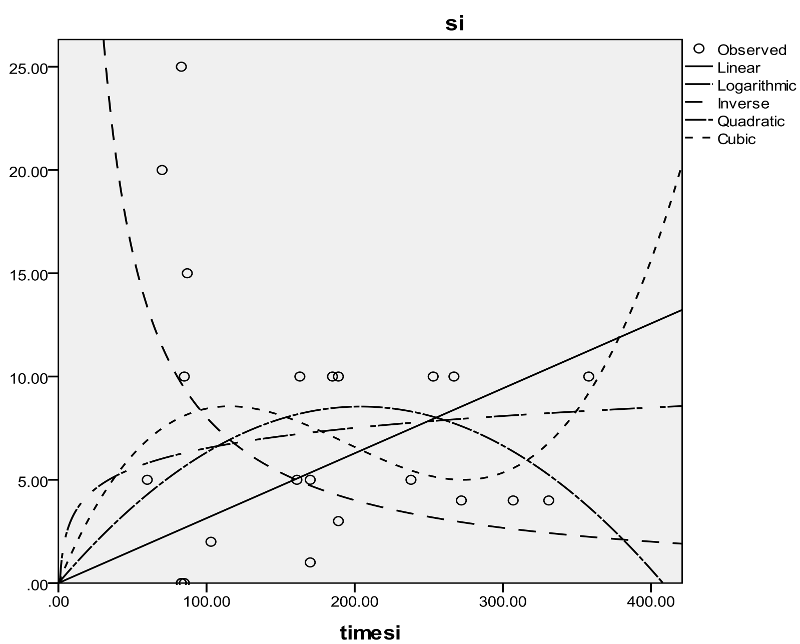
Model Summary and Parameter Estimates

Dependent Variable:si

	Model Summary					Parameter Estimates		
Equation	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.409	15.235	1	22	.001	.031		
Logarithmic	.557	27.700	1	22	.000	1.418		
Inverse	.530	24.854	1	22	.000	802.525		
Quadratic	.518	11.268	2	21	.000	.084	.000	
Cubic	.578	9.123	3	20	.001	.172	-.001	1.819E-6
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



SGMM

Iron

Warnings

The dependent variable (fe) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

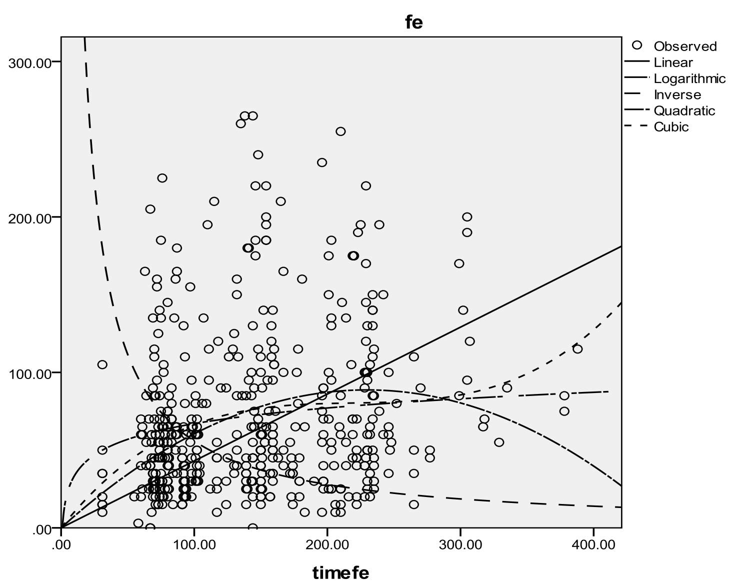
Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.607	745.506	1	482	.000	.430		
Logarithmic	.661	938.647	1	482	.000	14.555		
Inverse	.419	348.152	1	482	.000	5581.213		
Quadratic	.654	454.414	2	481	.000	.773	-.002	
Cubic	.663	314.716	3	480	.000	1.074	-.005	7.409E-6
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timefe.

a. The dependent variable (fe) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



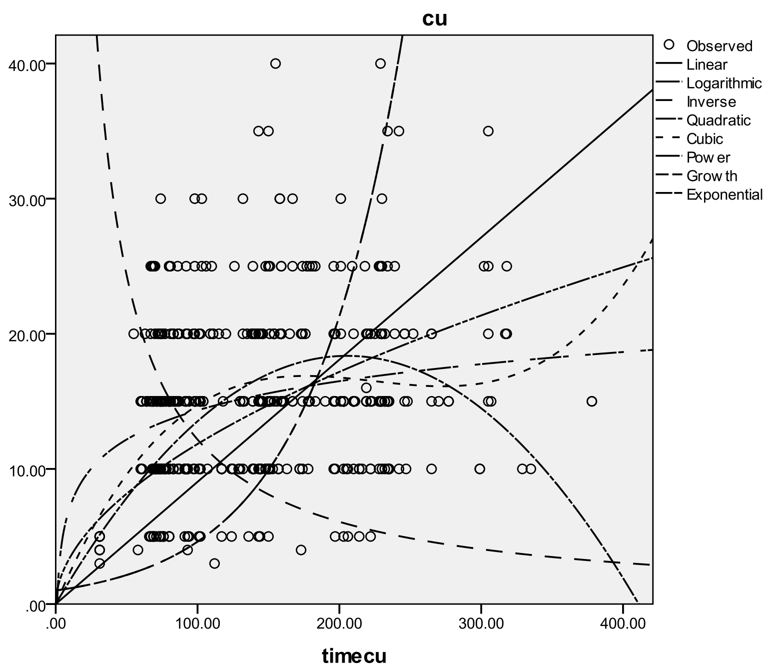
Cupper

Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.765	1578.093	1	484	.000	.090		
Logarithmic	.860	2977.067	1	484	.000	3.113		
Inverse	.565	629.649	1	484	.000	1217.416		
Quadratic	.849	1358.152	2	483	.000	.179	.000	
Cubic	.860	989.134	3	482	.000	.244	-.001	1.695E-6
Power	.971	15995.661	1	484	.000	.537		
Growth	.831	2378.496	1	484	.000	.015		
Exponential	.831	2378.496	1	484	.000	.015		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

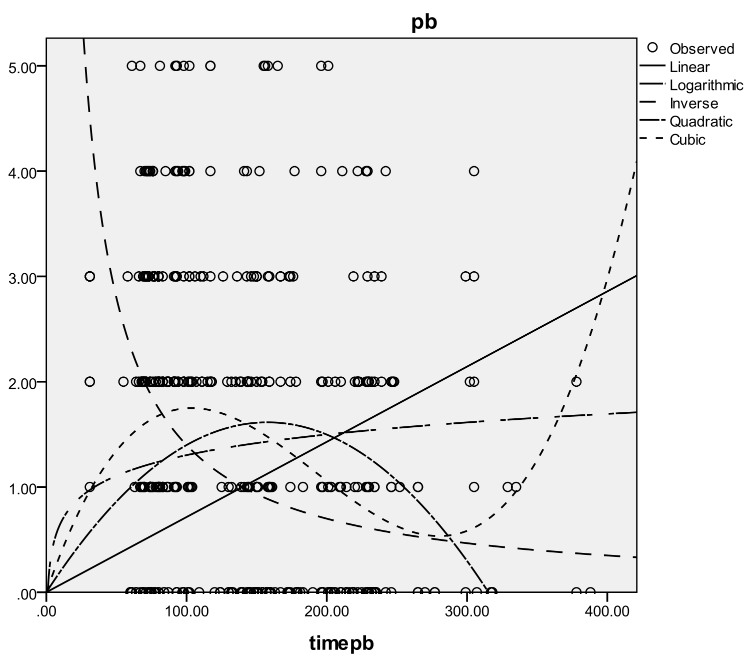
Model Summary and Parameter Estimates

Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.317	225.481	1	486	.000	.007		
Logarithmic	.467	426.315	1	486	.000	.283		
Inverse	.490	466.630	1	486	.000	139.769		
Quadratic	.450	198.687	2	485	.000	.020	-6.495E-5	
Cubic	.511	168.393	3	484	.000	.039	.000	4.421E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

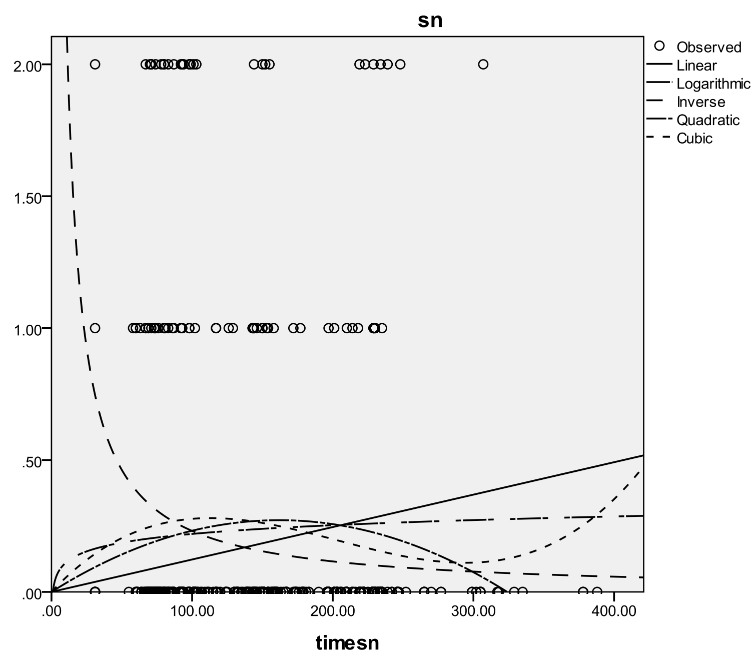
Model Summary and Parameter Estimates

Dependent Variable:sn

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.105	56.865	1	484	.000	.001		
Logarithmic	.149	84.878	1	484	.000	.048		
Inverse	.151	85.849	1	484	.000	23.129		
Quadratic	.143	40.386	2	483	.000	.003	-1.036E-5	
Cubic	.155	29.364	3	482	.000	.006	-3.479E-5	5.696E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

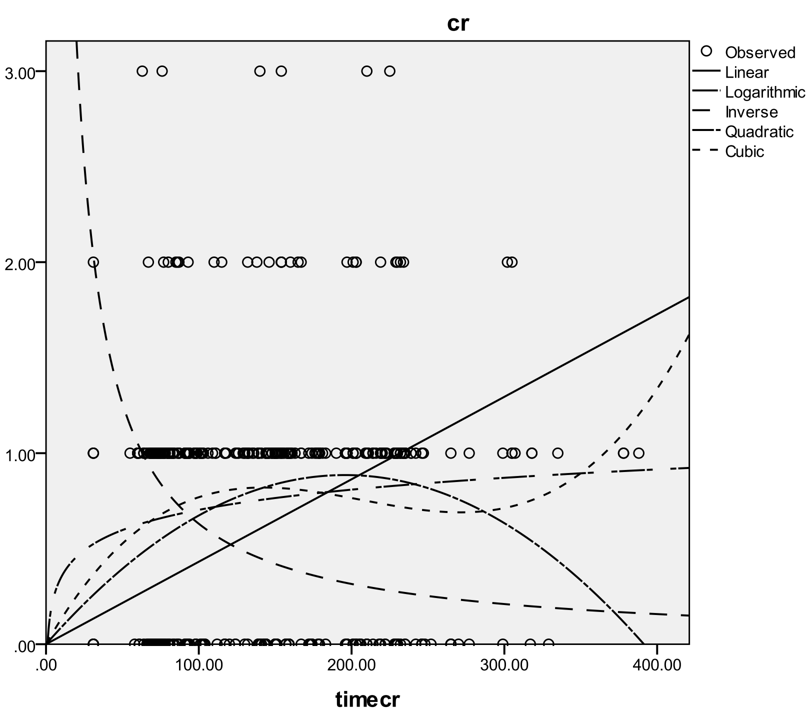
Model Summary and Parameter Estimates

Dependent Variable:cr

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.496	473.881	1	481	.000	.004		
Logarithmic	.587	682.795	1	481	.000	.153		
Inverse	.431	364.473	1	481	.000	63.047		
Quadratic	.568	316.147	2	480	.000	.009	-2.313E-5	
Cubic	.588	227.862	3	479	.000	.014	-7.495E-5	1.209E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

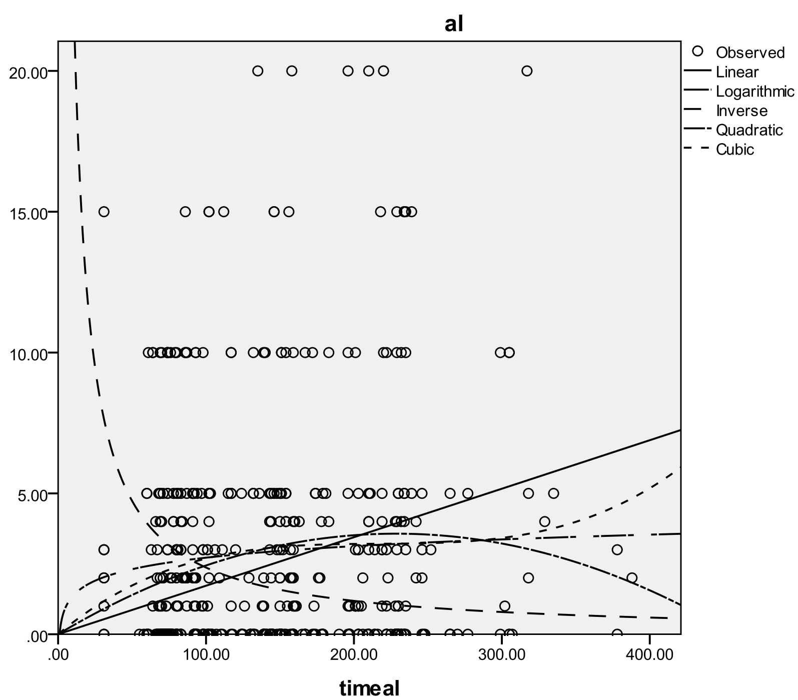
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.301	208.419	1	485	.000	.017		
Logarithmic	.333	241.908	1	485	.000	.590		
Inverse	.230	145.177	1	485	.000	236.772		
Quadratic	.325	116.407	2	484	.000	.031	-6.837E-5	
Cubic	.330	79.171	3	483	.000	.044	.000	3.099E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

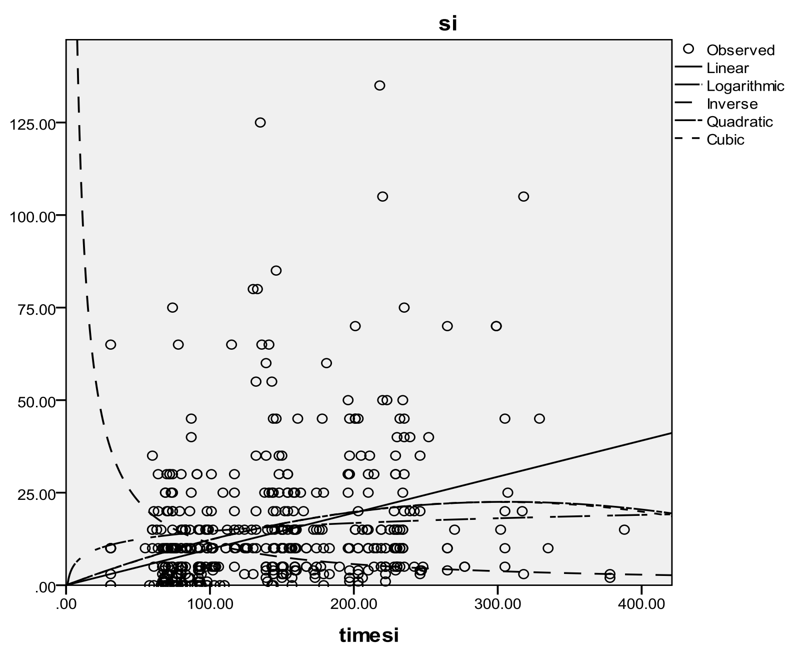
Model Summary and Parameter Estimates

Dependent Variable:si

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.433	368.913	1	483	.000	.098		
Logarithmic	.430	364.878	1	483	.000	3.171		
Inverse	.240	152.235	1	483	.000	1139.243		
Quadratic	.446	194.290	2	482	.000	.147	.000	
Cubic	.446	129.261	3	481	.000	.145	.000	-4.818E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



VIRM

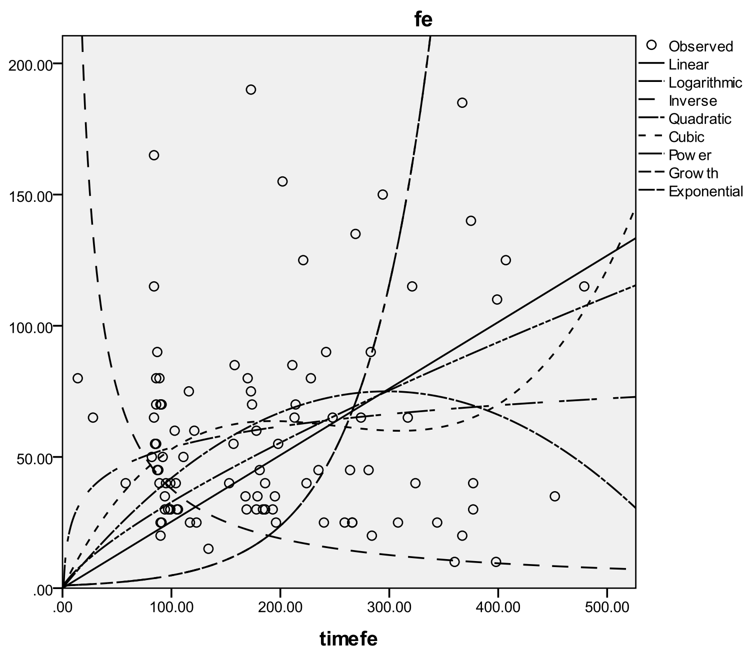
Iron

Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.596	144.424	1	98	.000	.253		
Logarithmic	.707	236.221	1	98	.000	11.632		
Inverse	.354	53.777	1	98	.000	3785.096		
Quadratic	.668	97.573	2	97	.000	.504	-.001	
Cubic	.696	73.104	3	96	.000	.833	-.004	4.649E-6
Power	.961	2417.487	1	98	.000	.758		
Growth	.745	286.487	1	98	.000	.016		
Exponential	.745	286.487	1	98	.000	.016		

The independent variable is timefe.



Cupper

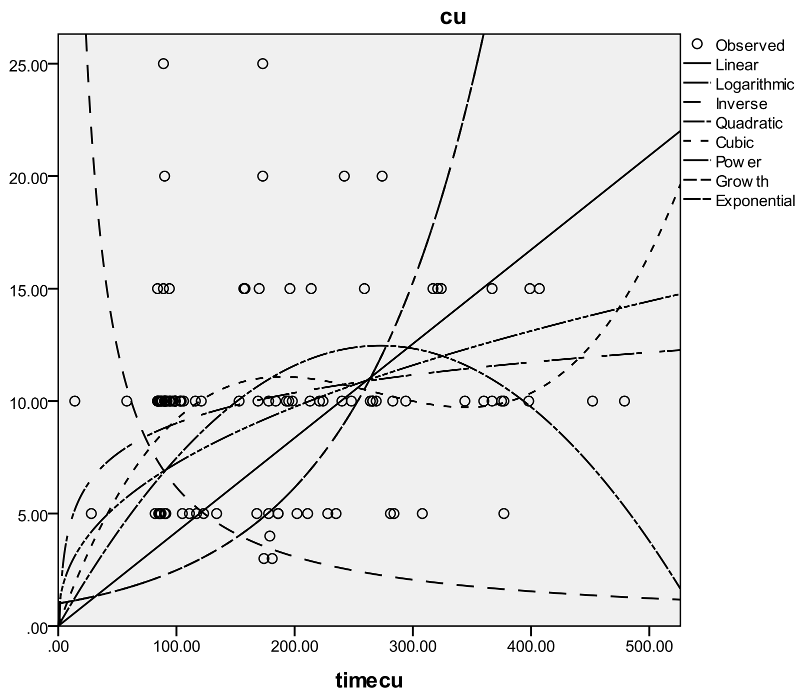
Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3

Linear	.674	202.222	1	98	.000	.042		
Logarithmic	.831	482.658	1	98	.000	1.957		
Inverse	.391	62.905	1	98	.000	616.941		
Quadratic	.791	183.519	2	97	.000	.091	.000	
Cubic	.819	144.808	3	96	.000	.143	-.001	7.277E-7
Power	.950	1871.928	1	98	.000	.430		
Growth	.755	301.684	1	98	.000	.009		
Exponential	.755	301.684	1	98	.000	.009		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Summary and Parameter Estimates

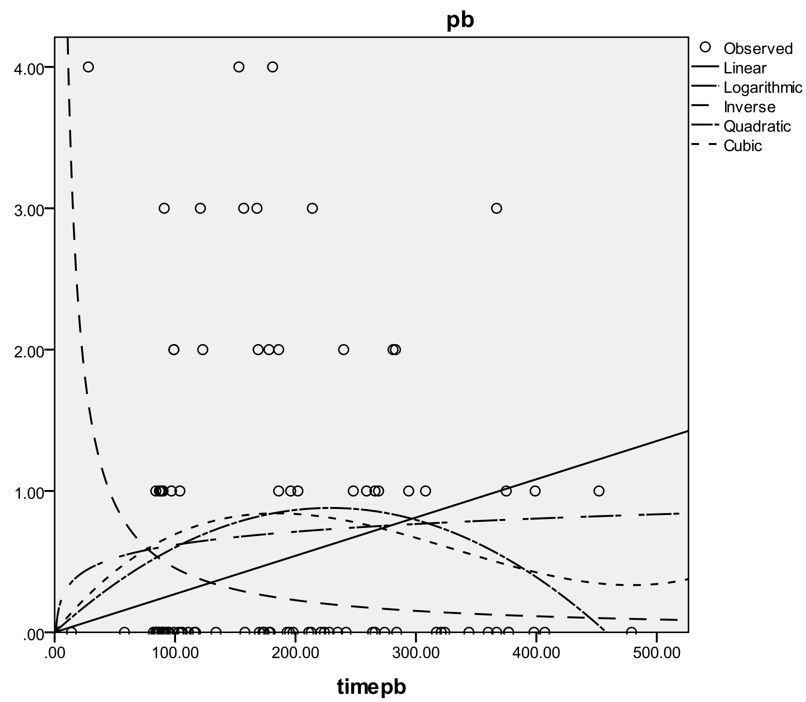
Dependent Variable:pb

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.211	25.907	1	97	.000	.003		

Logarithmic	.291	39.885	1	97	.000	.134			
Inverse	.160	18.411	1	97	.000	45.496			
Quadratic	.298	20.386	2	96	.000	.008	-1.676E-5		
Cubic	.304	13.853	3	95	.000	.010	-3.935E-5	3.954E-8	
Power[a]	.	.	.	.	.	.			
Growth[a]	.	.	.	.	.	.			
Exponential[a]	.	.	.	.	.	.			

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Summary and Parameter Estimates

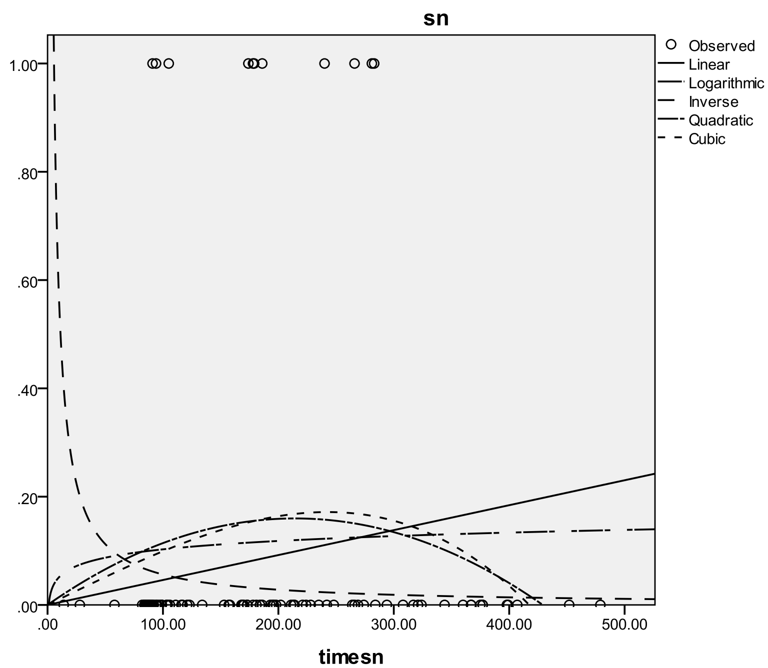
Dependent Variable:sn

Equation	Model Summary					Parameter Estimates			

	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.087	9.238	1	97	.003	.000		
Logarithmic	.115	12.625	1	97	.001	.022		
Inverse	.035	3.517	1	97	.064	5.618		
Quadratic	.141	7.883	2	96	.001	.001	-3.470E-6	
Cubic	.143	5.302	3	95	.002	.001	2.185E-7	-6.460E-9
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

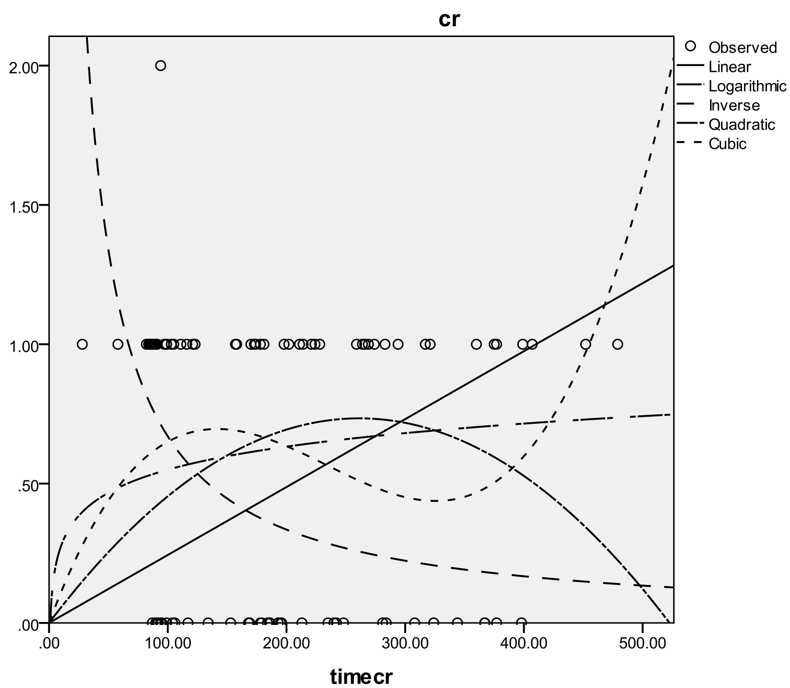
Model Summary and Parameter Estimates

Dependent Variable:cr

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.432	73.703	1	97	.000	.002		
Logarithmic	.582	134.999	1	97	.000	.119		
Inverse	.506	99.262	1	97	.000	66.969		
Quadratic	.523	52.536	2	96	.000	.006	-1.075E-5	
Cubic	.591	45.806	3	95	.000	.011	-5.816E-5	8.297E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

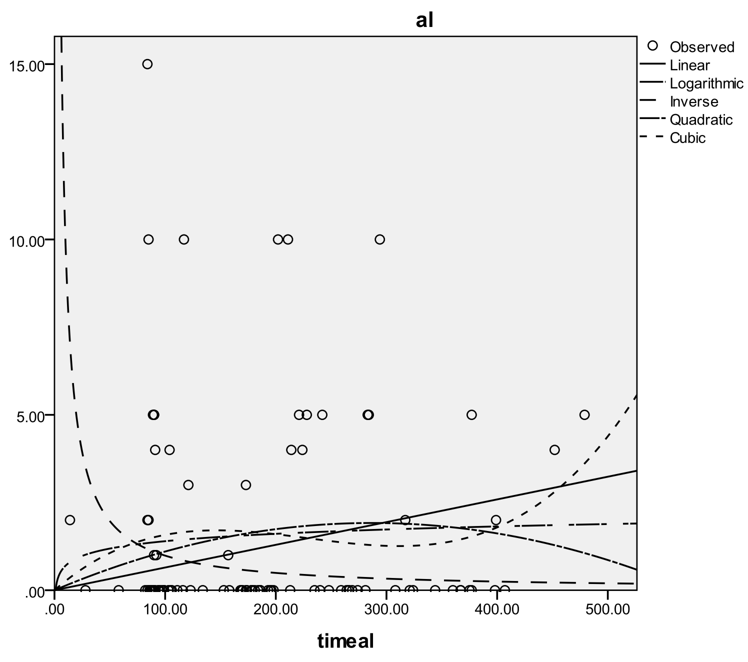
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.175	20.778	1	98	.000	.006		
Logarithmic	.217	27.115	1	98	.000	.304		
Inverse	.108	11.907	1	98	.001	98.628		
Quadratic	.199	12.080	2	97	.000	.013	-2.327E-5	
Cubic	.223	9.170	3	96	.000	.028	.000	2.016E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

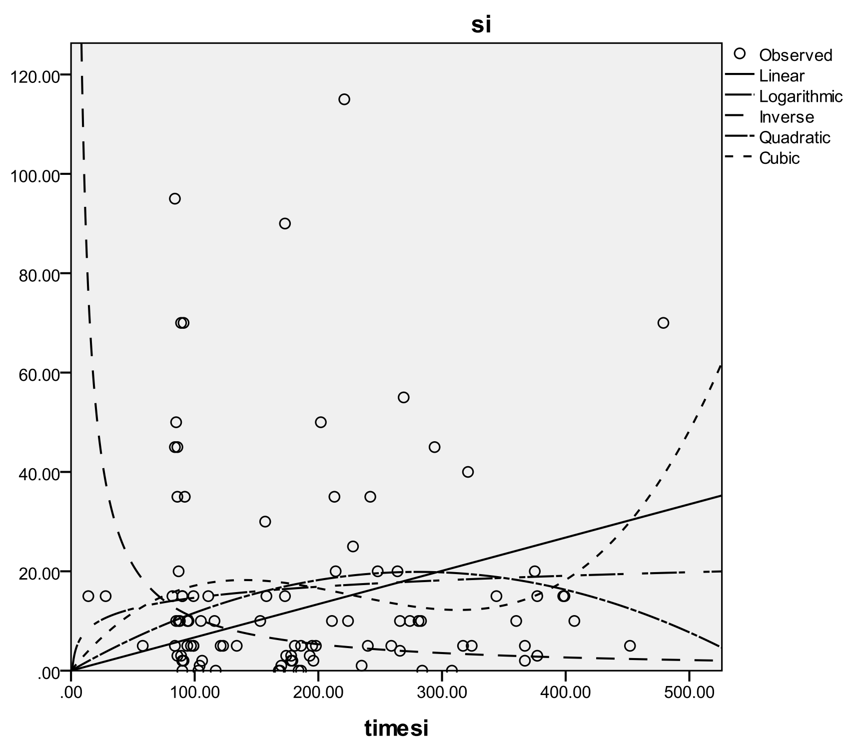
Model Summary and Parameter Estimates

Dependent Variable:si

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.277	37.455	1	98	.000	.067		
Logarithmic	.353	53.433	1	98	.000	3.188		
Inverse	.186	22.372	1	98	.000	1063.483		
Quadratic	.319	22.723	2	97	.000	.142	.000	
Cubic	.365	18.370	3	96	.000	.306	-.002	2.321E-6
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



VIRM-4

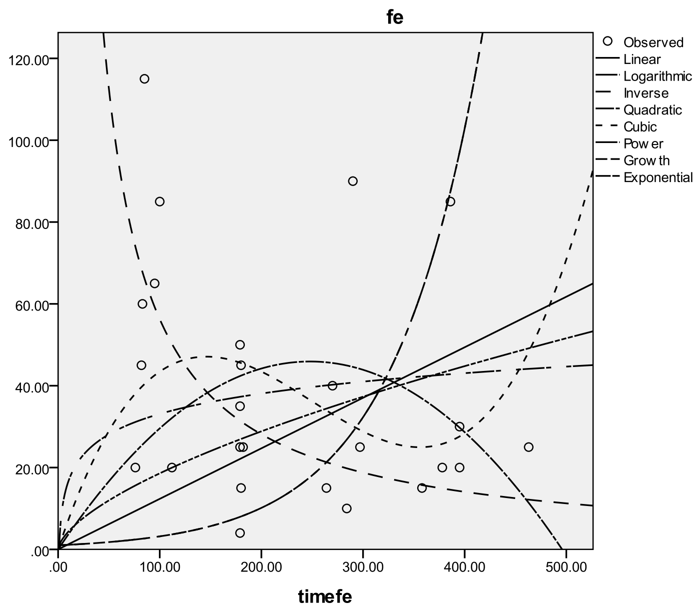
Iron

Model Summary and Parameter Estimates

Dependent Variable:fe

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.422	17.523	1	24	.000	.123		
Logarithmic	.619	38.943	1	24	.000	7.188		
Inverse	.655	45.513	1	24	.000	5626.534		
Quadratic	.568	15.099	2	23	.000	.370	-.001	
Cubic	.635	12.768	3	22	.000	.748	-.004	4.800E-6
Power	.929	312.336	1	24	.000	.635		
Growth	.715	60.284	1	24	.000	.012		
Exponential	.715	60.284	1	24	.000	.012		

The independent variable is timefe.



Cupper

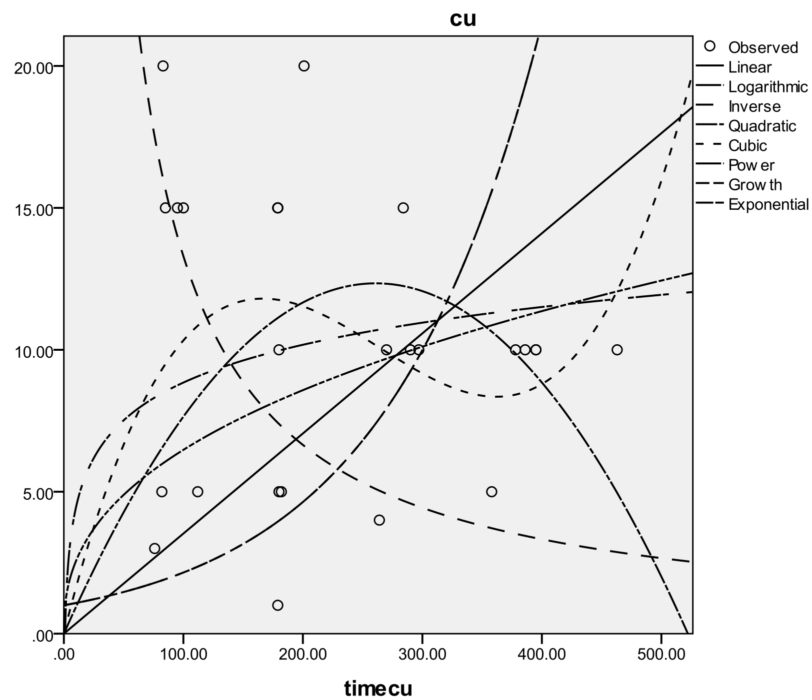
Model Summary and Parameter Estimates

Dependent Variable:cu

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.605	38.265	1	25	.000	.035		

Logarithmic	.788	92.922	1	25	.000	1.921		
Inverse	.640	44.510	1	25	.000	1330.038		
Quadratic	.758	37.581	2	24	.000	.095	.000	
Cubic	.801	30.880	3	23	.000	.167	-.001	9.262E-7
Power	.905	237.450	1	25	.000	.406		
Growth	.737	70.032	1	25	.000	.008		
Exponential	.737	70.032	1	25	.000	.008		

The independent variable is timecu.



Lead

Warnings

The dependent variable (pb) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Summary and Parameter Estimates

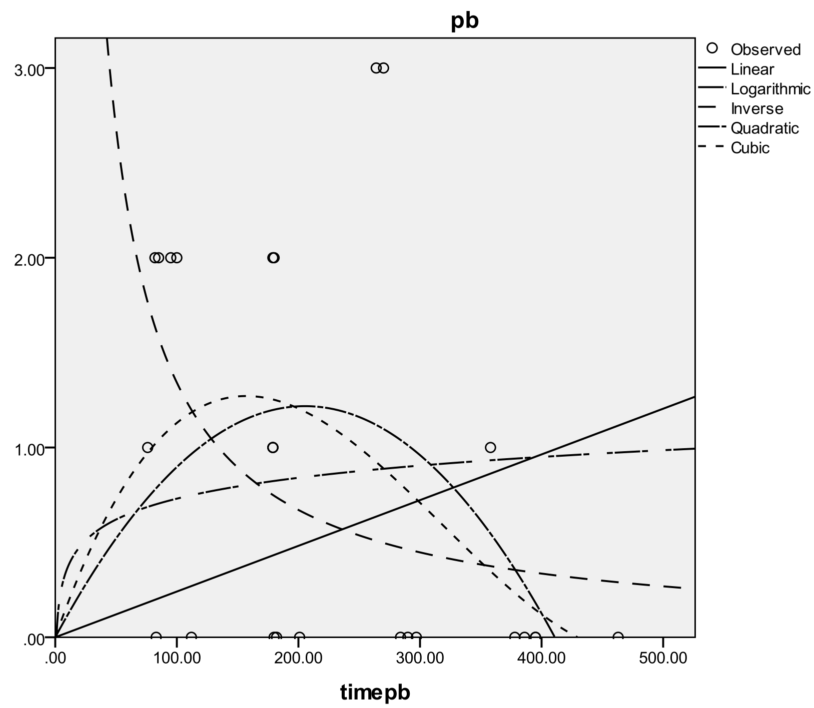
Dependent Variable:pb

	Model Summary					Parameter Estimates		
Equation	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.206	6.478	1	25	.017	.002		
Logarithmic	.392	16.101	1	25	.000	.159		

Inverse	.475	22.645	1	25	.000	134.213			
Quadratic	.489	11.500	2	24	.000	.012	-2.886E-5		
Cubic	.513	8.073	3	23	.001	.018	-7.673E-5	8.016E-8	
Power[a]	.	.	.	.	.	.			
Growth[a]	.	.	.	.	.	.			
Exponential[a]	.	.	.	.	.	.			

The independent variable is timepb.

a. The dependent variable (pb) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Tin

Warnings

The dependent variable (sn) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

Model Summary and Parameter Estimates

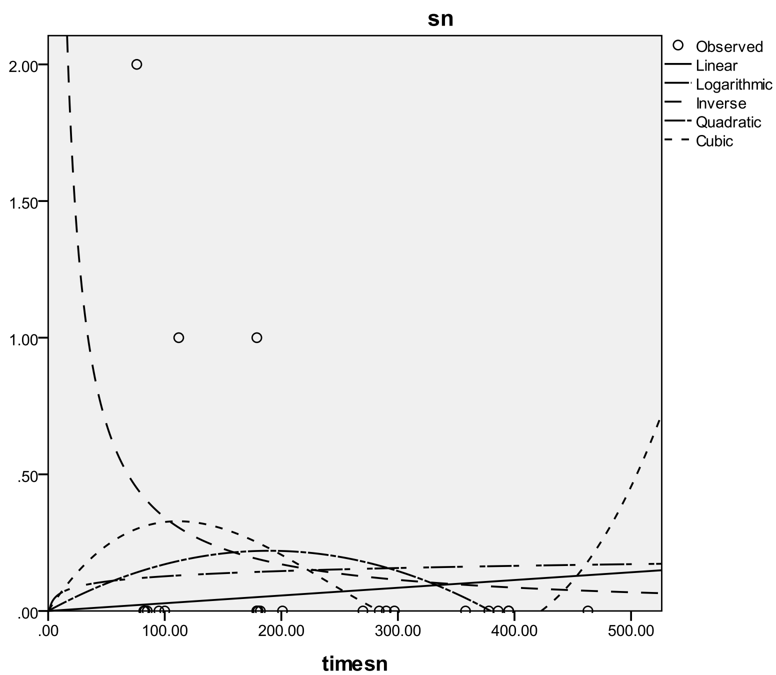
Dependent Variable:sn

Equation	Model Summary	Parameter Estimates

	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.021	.491	1	23	.491	.000		
Logarithmic	.085	2.144	1	23	.157	.028		
Inverse	.233	6.971	1	23	.015	34.180		
Quadratic	.114	1.417	2	22	.264	.002	-6.114E-6	
Cubic	.197	1.715	3	21	.195	.007	-3.870E-5	5.472E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesn.

a. The dependent variable (sn) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Chrome

Warnings

The dependent variable (cr) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

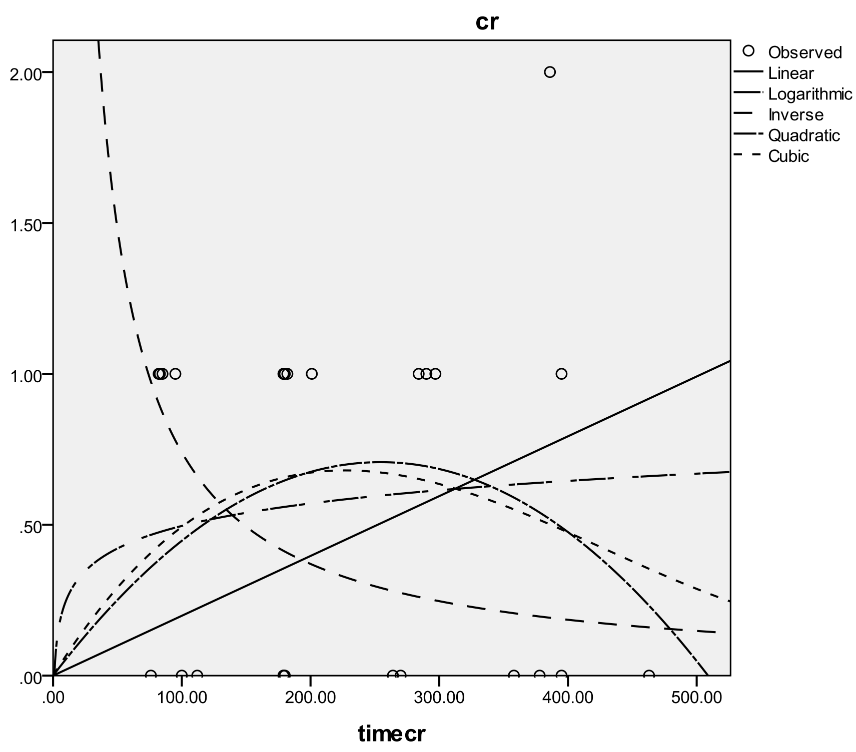
Model Summary and Parameter Estimates

Dependent Variable:cr

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.385	15.669	1	25	.001	.002		
Logarithmic	.499	24.915	1	25	.000	.108		
Inverse	.400	16.637	1	25	.000	74.010		
Quadratic	.498	11.890	2	24	.000	.006	-1.093E-5	
Cubic	.500	7.675	3	23	.001	.007	-2.041E-5	1.588E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timecr.

a. The dependent variable (cr) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Aluminum

Warnings

The dependent variable (al) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

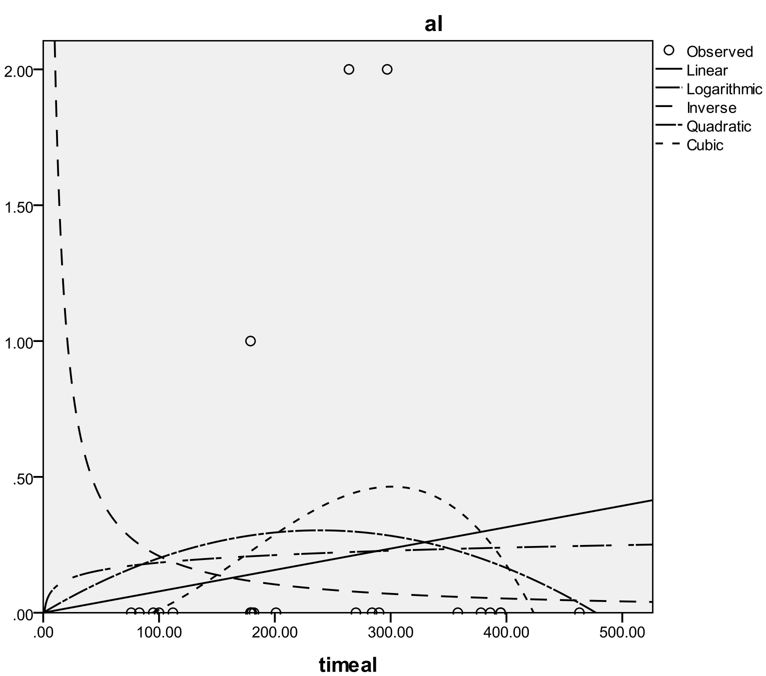
Model Summary and Parameter Estimates

Dependent Variable:al

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.114	2.953	1	23	.099	.001		
Logarithmic	.123	3.236	1	23	.085	.040		
Inverse	.046	1.113	1	23	.302	20.877		
Quadratic	.162	2.120	2	22	.144	.003	-5.326E-6	
Cubic	.225	2.030	3	21	.140	-.002	3.155E-5	-6.092E-8
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timeal.

a. The dependent variable (al) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Silicon

Warnings

The dependent variable (si) contains non-positive values. The minimum value is .000. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.

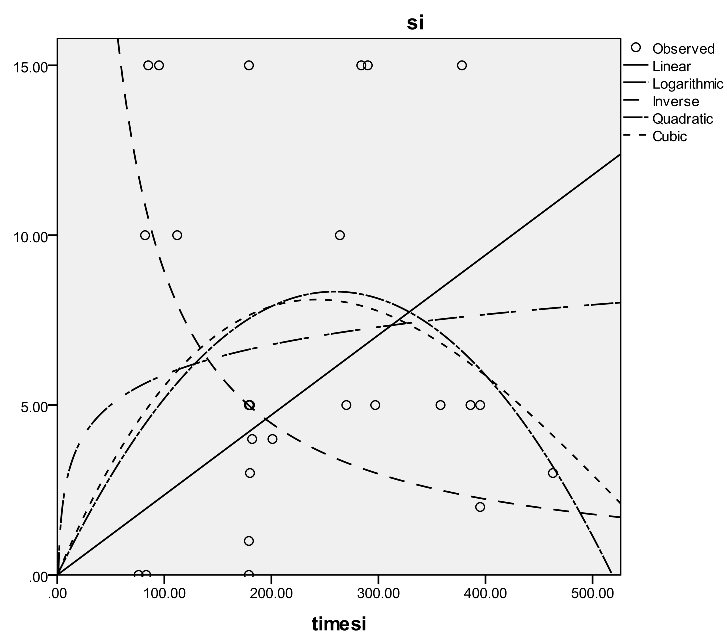
Model Summary and Parameter Estimates

Dependent Variable:si

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.488	22.907	1	24	.000	.024		
Logarithmic	.619	38.956	1	24	.000	1.280		
Inverse	.483	22.450	1	24	.000	892.717		
Quadratic	.615	18.385	2	23	.000	.064	.000	
Cubic	.616	11.788	3	22	.000	.074	.000	1.218E-7
Power[a]	.	.	.	.	.	.		
Growth[a]	.	.	.	.	.	.		
Exponential[a]	.	.	.	.	.	.		

The independent variable is timesi.

a. The dependent variable (si) contains non-positive values. The minimum value is .00. Log transform cannot be applied. The Compound, Power, S, Growth, Exponential, and Logistic models cannot be calculated for this variable.



Appendix D: Multiple Regression Analysis for MDDM

Iron

Model Summary

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
1	.856[a]	.733	.709	33.99056

a. Predictors: agemddm, timemddm

b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

ANOVA[c,d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	69807.117	2	34903.559	30.210	.000[a]
	Residual	25417.883	22	1155.358		
	Total	95225.000[b]	24			

a. Predictors: agemddm, timemddm

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Dependent Variable: femddm

d. Linear Regression through the Origin

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	timemddm	-.012	.073	-.038	-.169	.868
	agemddm	4.254	1.076	.889	3.955	.001

a. Dependent Variable: femddm

b. Linear Regression through the Origin

Cupper

Model Summary

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
1	.848[a]	.719	.693	27.10497

a. Predictors: agemddm, timemddm

b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

ANOVA[c,d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	41287.048	2	20643.524	28.099	.000[a]
	Residual	16162.952	22	734.680		
	Total	57450.000[b]	24			

a. Predictors: agemddm, timemddm

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Dependent Variable: cumddm

d. Linear Regression through the Origin

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	timemddm	.020	.058	.080	.345	.733
	agemddm	2.889	.858	.777	3.368	.003

a. Dependent Variable: cumddm

b. Linear Regression through the Origin

Lead

Model Summary

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
1	.571[a]	.326	.265	3.25046

a. Predictors: agemddm, timemddm

b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

ANOVA[c,d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.560	2	56.280	5.327	.013[a]
	Residual	232.440	22	10.565		
	Total	345.000[b]	24			

a. Predictors: agemddm, timemddm

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Dependent Variable: pbmddm

d. Linear Regression through the Origin

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	timemddm	.005	.007	.260	.727	.475
	agemddm	.095	.103	.330	.925	.365

a. Dependent Variable: pbmddm

b. Linear Regression through the Origin

Silicon

Model Summary

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
1	.749[a]	.561	.521	8.16397

a. Predictors: agemddm, timemddm

b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

ANOVA[c,d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1870.692	2	935.346	14.034	.000[a]
	Residual	1466.308	22	66.650		
	Total	3337.000[b]	24			

a. Predictors: agemddm, timemddm

b. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

c. Dependent Variable: simddm

d. Linear Regression through the Origin

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	timemddm	-.008	.017	-.131	-.455	.654
	agemddm	.771	.258	.860	2.983	.007

a. Dependent Variable: simddm

b. Linear Regression through the Origin

Appendix E: Inspection Sequences for Train Types

ICM 3

Soft failure $\Phi \sim Weibull(860,13)$

Hard failure $F \sim Weibull(1380,15)$

Cost parameters; $c_I = 10000$, $c_g = 70000$, $c_p = 15000$

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
890	951	990	1018	1040	1059	1075	1089

	Inspections							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	166	241	241	252	252	461	461	461
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

MDDM

Soft failure $\Phi \sim Weibull(1175,13)$

Hard failure $F \sim Weibull(1880,13)$

Cost parameters; $c_I = 10000$, $c_g = 100000$, $c_p = 15000$

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1220	1303	1355	1393	1423	1448	1470	1489

	Inspections							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	226	255	255	414	414	414	414	599
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

SGGM

Soft failure $\Phi \sim Weibull(1051,15)$

Hard failure $F \sim Weibull(1683,15)$

Cost parameters; $c_I = 10000$, $c_g = 70000$, $c_p = 15000$

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1081	1145	1186	1216	1239	1258	1275	1289

Inspections								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	216	267	316	316	338	348	478	478
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

VIRM

Soft failure $\Phi \sim Weibull(1278,13)$

Hard failure $F \sim Weibull(2045,13)$

Cost parameters; $c_I = 10000$, $c_g = 150000$, $c_p = 15000$

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1327	1415	1470	1511	1543	1570	1593	1613

Inspections								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	244	312	317	394	460	467	483	483
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

VIRM-4

Soft failure $\Phi \sim Weibull(2132,9)$

Hard failure $F \sim Weibull(3411,12)$

Cost parameters; $c_I = 10000$, $c_g = 200000$, $c_p = 15000$

Inspection Times							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1307	1694	1911	2064	2182	2278	2359	2429

	Inspections							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample Size	106	156	156	252	252	252	986	986
Uncertainty Cost	1000	1500	2000	2500	3000	3500	4000	4500

Appendix F: User Interface