CHALMERS
UNIVERSITY OF TECHNOLOGY

Production Data Analytics – To identify productivity potentials

MUKUND SUBRAMANIYAN

Department of Product and Production Development
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden - 2015

Master Thesis - 2015

Production Data Analytics – To identify productivity potentials

Mukund Subramaniyan

CHALMERS

Department of Product and Production Development

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden, 2014

Production Data Analytics – To identify productivity potentials
MUKUND SUBRAMANIYAN

Department of Product and Production Development
Chalmers University of Technology
SE-41296 Gothenburg
Sweden
Telephone - +46 (0) 31-772 1000

Gothenburg, Sweden 2015

ABSTRACT

Manufacturing industries are always under constant pressure to improve the productivity. Many manufacturing companies started to capture the shop floor data to the time scale of seconds. Consequently, the challenge is to harness the value from the data and identify the ways in which the value extracted could improve the productivity.

In this thesis, two sets of Manufacturing Execution Data (MES) data consisting of shop floor data were used to identify the productivity potentials. The first set of data had the MES information was derived from a common data source of 23 industries consisting of 884 machines. The second data set was more specific to one manufacturing line. The methodology to analyse both data sets includes data cleaning, data preparation and data modelling.

The outcome of the analysis of the first data set was the impact of operator influenced loss times on Overall Equipment Efficiency (OEE). The outcomes of the analysis of the second data set were identification of static and momentary bottlenecks in the production line from the real time data and to develop algorithms for those. Also, the Key Performance Indicators (KPI) were modelled to determine the pattern and to predict their behaviour.

Identifying the productivity potentials (operator influenced loss times, bottlenecks detection and predicting the behaviour of the KPI) from the real time data is very useful to make fact based decisions which reduces the value at risk of making these decisions which in turn helps to improve productivity.

Keywords: Operator Influenced Loss Times, Bottleneck Detection, Data Driven Analytics, Big Data, OEE

Acknowledgement

First and foremost, I am deeply indebted to my thesis examiner Anders Skoogh; his contributions of time, ideas which have been invaluable to me and which made my thesis experience productive and stimulating. He created an opportunity for me to work in a research project StreaMod. He has provided guidance at the key moments and allowed me to work independently on the thesis. I would also like to thank Peter Almstrom for giving me an opportunity to work in a research project and for his guidance and extensive support. A special thanks to both Peter Almstrom and Anders Skoogh for allowing me to combine the two separate research projects and formed a very interesting thesis. It has been an honour to be as a student under both of them.

I also thank my supervisors, Maheshwaran Gopalakrishnan and Richard Hedman, for their constant encouragement and support. The challenge in completing this thesis resulted in countless enjoyable discussions with them. They significantly influenced the way I think about the research projects. Both patiently corrected my mistakes and cared about my work at every step to successfully complete this thesis. Should you find this thesis an interesting study; it is very much because of their valuable insights.

In addition to my supervisors I would also like to thank for the extensive help received from Hanna Atieh, Partner Company. She helped in understanding the data files and to validate the results.

I would also like to thank the other research members of the StreaMod project for their inputs by asking critical questions which made me to improve the quality of my analysis.

Finally, special thanks for all the help received from the very helpful staffs at Product and Production Development Department at Chalmers University.

Lastly, I would like to thank my parents. They did provide me more than my genetic material. Without their continuous encouragement and support, my masters would have been just a dream. Also, I would like to thank all my friends for their support.

Mukund Subramaniyan
Gothenburg, 2015

Contents

List of Figures

List of Tables

1. Introduction

In this chapter, the background of the problem is presented followed by the thesis aim and finally the research questions.

1.1. Background

The key factor which drives the improvement of the manufacturing industries today is the need for competitiveness and the ability to handle the threats and opportunities in a flexible manner. The manufacturers are always under constant pressure to make more efficient use of resources. Today they find it difficult to improve the productivity in industries which already have an efficient process(James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, 2011). Though lean and six sigma techniques are deployed by the manufacturers to reduce the waste and variability of the production process, there are sometimes extreme swings in the variability. Moreover, the production systems could be complex in nature, multi-stepped with multiple loops with strong interrelationships and also consists of many activities that influence the productivity. In this scenario, the manufacturers need more granular approach in diagnosing and correcting the process flow and to reduce the variability. One of the methods that promise to deliver significant productivity gains and explains why the variability exists is the application of data analytics(Lechevalier, Narayanan, & Rachuri, 2014). Nowadays, with the advancements in the technology and systems integration, the capturing of the real time information of the manufacturing operations has reached a new dimension(Krumeich et al., 2014). Manufacturing is one of the largest data generators today(Data, Group, & Greenplum, n.d.). Real time data and information recorded by the Manufacturing Execution Systems (MES) are the primary assets of the manufacturing industries. For the companies to run smarter, more agile, the collected data should be transformed into meaningful information with the data driven discoveries e.g. analytics (Davenport, 2006).

Big data is a growing torrent among the manufacturing industries. From this big data which are collected, in-depth insights into the production process can be drawn within and compare them with other similar production systems to identify the key improvement areas. Apart from finding improvement potentials, the big data, provides us with the insights that help the companies in day to day decision making process. According to McAfee and Brynjolfsson (2012),

"..*the more companies, characterised themselves as data-driven, the better they performed on objective measures of financial and operational results. In particular, the companies in the top third of their industry in the use of data driven decision making were, on average 5% more productive and 6% more profitable than the competitors."* (p.64).

The manufacturers collect the data and use them either for tracking purposes or to use them to improve and optimise the production processes. However, recent trend show that the use of big data in manufacturing is still in the primary phase when compared to the other types of industries like finance, service etc.(Yang & Nurtam, 2013). Advanced analytics over the big data can give the companies an explicit picture of the impact of different variables on the overall productivity of their operations. First, it delivers insights on the most efficient way to control the production systems given the system constraints (e.g. better management of bottlenecks in the production). Secondly, it highlights the largest opportunities to improve the performance by highlighting the key losses of the production system. Moreover, analytics explores the correlations to identify patterns and predicts the relationships among variables and also quantifies the applicability of the identified insights statistically. With the use of increasing availability of the data and the tools to analyse them, tremendous value could be captured from the new treasure trove and the facts from the data could fundamentally change the way that the management takes decisions.

1.2. Overall Purpose

The purpose of this thesis is to use the production data to achieve increased productivity of the production system. This is done by identifying the potentials through data analytics from the real time data.

Two sets of big data are analysed in this thesis. Those are described below:

- Data set I: It is derived from a common data source which records the MES information of machines of different industries.
- Data set II: It is derived from a supervisory system that monitors, collects and stores the information of the automated assembly line at a factory.

1.3. Background and Objectives of Data Set I

Overall Equipment Efficiency (OEE) is a measure that quantifies how the well a manufacturing unit performs by comparing what the equipment produced to what the equipment could have potentially produced. Improving the OEE is an important factor in the different types of manufacturing industries today where shrinking margins, consolidation and fierce competition have driven the need to reduce cost and improve the efficiency. OEE is a valuable tool to unleash the hidden capacity (Muchiri & Pintelon, 2008). OEE accounts for the losses due to availability, performance and quality. This metric in-turn measures different types of production losses and indicate areas of potential improvement. The losses of the production system are calculated from one of the sources: real time manufacturing data collected by Manufacturing Execution System (MES) or similar resources which tracks and stores the real time production data, from the manual files that records the production details maintained by the production team. By eliminating the losses from the production system, the unplanned downtime of the machines are reduced, thus increasing the OEE. Though OEE is a popular measure of productivity, it can only be used to measure the performance of the individual machines(Muchiri & Pintelon, 2008).

OEE was first introduced by Nakajima in 1988.Many researches and practitioners have argued about the use OEE since then in many different ways over the years. With the evolution of OEE, different modifications were done to OEE to fit a broader perspective as supposed important for the companies(Muchiri & Pintelon, 2008). Those are Overall Factory Effectiveness (OFE), Overall Plant Effectiveness (OPE), Overall Throughput Effectiveness (OTE), Production Equipment Effectiveness (PEE), Overall Asset Effectiveness (OAE) and Total Equipment Effectiveness Performance (TEEP) (Muchiri & Pintelon, 2008). While OEE measures the equipment's efficiency against the scheduled time, TEEP measures the OEE against the calendar time. PEE gives a weight to the factors of OEE i.e. PEE doesn't give equal importance to all the three factors of OEE. OFE, OPE and OAE raise the OEE measure to the factory level. However they are raised by synthesizing the subsystem level metrics and capturing their interconnectivity information(Huang & Keskar, 2007). To summarise, TEEP and PEE are measures of individual equipment's performance whereas OAE, OFE and OPE are extended to the factory level.

The competitiveness of the manufacturing facilities not only depends on the utilization of the equipment but also depends on operator productivity(Chien, Zheng, & Lin, 2013). The human factor (operators) is the most important and a critical factor which influences the machine productivity(Dvořák, Malkovský, & MacKů, 2008). Research has showed that the interference problems (i.e. allowing one operator to operate or repair several machines) results in machine idleness reduces the production system performance .This interference time is the potential time lost in the total planned production time.

The performance measurement metrics: OEE, TEEP, OFE, OPE, PEE, OAE measure only the equipment's efficiency. But on the other hand, the utilisation of the operators affects the utilisation of the equipment and vice versa (Hedman, Sundkvist and Almstrom, 2014). But these equipment performance metrics doesn't provide any information on the amount of manual work done by the operator or the machine interference time. As human decisions and actions also affect the overall performance of the production system(Bailey & Barley, 2005), it is necessary to analyse the amount of manual work or the operator influenced loss times (like machine interference time) in order to improve the overall productivity.

To remain competitive, apart from increasing productivity, companies should compare themselves with the peers. This method, which is otherwise called as benchmarking, allows comparing the practices and process with peer companies which helps them to identify areas of competitive advantage and disadvantages (Boxwell & Robert, 1994). Also, this benchmarking helps analysing the underlying reasons behind the variations in performance. Moreover, it will strengthen the credibility and demonstrate the performance to the company's stakeholders. On the other hand, the benchmarking exercise help researchers in order to assess and focus on the reasons for the performance gaps within or across the different industry groups which could underpin new wave for productivity potentials.

This study is initiated by the researchers at Chalmers University after the findings from the master thesis at Aros electronics on "Increasing the productivity of a surface mounting line"(Bergstrom& Palmkvist, 2014) that the operator also influences the OEE. Good Solutions AB is a partner in this study and supported this study by providing the empirical data set. Good Solutions AB is a Swedish Software and Service company that has developed its own product called RS Production. They implemented this tool in many Swedish Industries. This tool is a used for practical improvements and real time visualisation of production status on the shop floor.

1.3.1. Significance of the Study

This study is intended to make to the advancement of knowledge to improve the productivity in Swedish manufacturing industries. The overall OEE of the 23 Swedish Industries could be used as base performance metric for the future research conducted in the area of productivity improvements. Also, the effect of the operator influence loss times is not explicitly shown by the OEE and other derived performance metrics of OEE. The results from this study are intended to show the proportion of the operator influenced loss times from the MES data and to show how those could be seen as improvement potential in order to increase the OEE.

1.3.2. Purpose

The purpose of the study is to identify the ways by which higher OEE of the machines can be achieved in a production system.

1.3.3. Objective

The objective of the study is to determine the current OEE and to assess the impact of the operator influenced loss times on OEE.A set of research questions are formulated to specify the objective of the study and to maintain its scope. The research questions are formulated with the aim to answer them with the results got from this study. The following three research questions are stipulated:

RQ: How can the big data be used to determine the OEE and the loss levers of OEE?

This research question is framed in order to find out the levers of the OEE by analysing the different losses from big data in order to find out the OEE. To be further specific, this research question is split into three research questions RQ1, RQ2 and RQ3.

RQ 1 : What is the difference in OEE among the different industry groups?

This research question is framed in order to get the overview of the performance of the companies in different industry groups. Also, the aim is to analyse which of the losses contribute to the OEE. This forms the basis for RQ 2.

RQ 2 : What is the average overall OEE of the industries from the given MES big data?

This research question is framed to evaluate the average overall performance of industries by combining the performance of the individual industry groups and to identify the losses contributing to the OEE.

RQ 3 : How large is the operator disturbance portion of OEE?

The third research question is stated in order to assess how much of the total loss time is operator influenced. This research question also involves the investigation on how the operator tasks could affect the OEE.

1.3.4. Delimitations

In order to limit the analysis of the study and ensure the adherence to the objectives, the following delimitations are made:

- The MES data from Good Solutions AB is assumed to be reliable data and no further explicit validation will be done to validate the data
- The basis of the selections of the machines in each company was not described. Due to this fact it was assumed that the data is given for all the machines under each company
- Though the data file had the cycle time factor which was measured, this was not taken into the analysis as the cycle time was not measured for all machines. This was done to maintain the uniformity in the analysis.
- The level of performance at individual company level is not the focus of the study but rather the focus is on the general industry groups
- Only the focus is on evaluating how the operator influences the OEE and not on other OEE factors
- No recommendations based on the OEE of the industrial groups are made as the OEE represents the aggregated performance of the various machines. This is due to the fact that the production flows or bottleneck machines are not identifiable from the empirical data set

1.4. Background and Objectives of Data Set II

To remain competitive most of the manufacturing companies should take steps to increase the productivity and also establish greater operational stability. The manufacturing companies should manage the production networks efficiently and this includes the key challenge of increasing the efficiency of the shop floor operations. There are many variables which affects the performance of the system. The throughput of the production system is affected by the capacity of the machines in the system(Sundkvist & Sundkvist, 2014). Depending on the nature of the production system, some machines disrupt the flow of products across the production system and affect the overall throughput. The limitations of the production system can be traced to limitations of one or two machines which are called as “bottlenecks”(Christoph Roser, 2001). To maximize the throughput from the production line, the throughput of the bottlenecks needs to be improved (Goldrat, 1992). Efforts and resources should therefore be focused on bottleneck machines to get an improved throughput from the line. The bottlenecks could be due to frequent failures of the machines, long setup times etc. Moreover, the other variables like incoming raw material fluctuations; product mix etc. can also shift the bottlenecks from one day to other day. Identification and management of the bottlenecks in the production system is a challenging task. Another important component of the production system which has a direct impact on the improvement of the overall production performance is the maintenance. According to Mishra (2012) if the right kind of maintenance is not chosen then it may lead to over maintenance or under maintenance which might increase the cost and reduce the productivity. Therefore, the cost effective and right maintenance at right time will boost the productivity in a production system by reducing the total breakdown time and by reducing the frequency of breakdowns.

The companies have grown remarkably in sophistication bristling with MES systems which monitor the machine activity almost every instant of the time. This results to the accumulation of the machine data. Figure 1, shows the average data rows collected per year from a machine in the shop floor of an automotive manufacturing company in Sweden. If 500 000 data rows are collected per machine, and if the production line has ten such machines, then the amount of data collected for the production system is 5 million data rows per year which is really big.

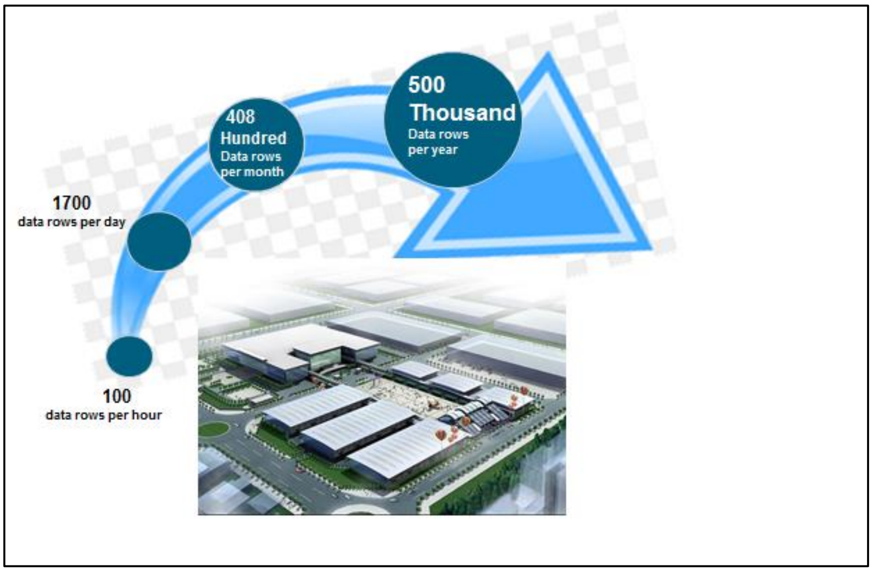


Figure 1: Average data rows of data collected per machine

Tapping the valuable information from this vast troves of data which otherwise would go unused could help manufacturer to gain valuable insights and this leads to profitable operation. The analytics on the big data collected can help to identify the bottlenecks and the selection of the maintenance strategies. The analytics could be conducted on myriad of ways over the big data which could identify the correlations and hidden patterns, thus supporting the fact based decision making process in the production environment. Therefore, manufacturers must take the advantage of this real time data collected by MES to attack their biggest challenges and the most important objectives.

This thesis is done as a part of the research project "StreaMod" within the Production Engineering Department at Chalmers University of Technology with the involvement of several industrial and academic partners.

1.4.1. Significance of the study

This study is intended to make to an advancement of knowledge to improve the productivity in Swedish manufacturing industries using real time production data collected from the shop floor. The algorithms for the bottleneck detection from the real time data could be used an alternative to simulation of production systems which demands much more skills and time. The traditional simulation studies takes from four weeks and more("Application within : FFI , Hållbar produktionsteknik Streamlined Modeling and Decision Support for Fact-based Production Development," 2013) and by using the data driven analysis will take much lesser time. Also, other useful insights on the key performance indicators of the production system derived from the real time analytics will serve as a base for decisions made to improve the production performance and thus making a data driven decision making process in manufacturing industries.

1.4.2. Purpose

The purpose of the study is to identify the ways in which productivity of the production line by controlling and managing the system constraints. The system constraints may be due to bottlenecks or frequent breakdowns of the machine etc. Before controlling and managing the system constraints, identifying the system constraints is important. By addressing and managing the system constraints the overall availability of the production line is increased. To support this analysis the real time data of an assembly line of an automotive manufacturing company in Sweden is derived from the company's supervisory system that monitors, collects and stores the machine information and analyzed.

1.4.3. Objectives

The objective is to detect the bottlenecks and to predict the breakdown pattern of the machines from real time data. A set of research questions was formulated to specify the objective of the study and to maintain its scope. The following three research questions were formulated,

- RQ 1 : How can the real time data be used to visualize bottlenecks and downtime parameters?

  The first research question is framed to try out the different bottleneck techniques and with the sample data set and visualize the results from the analysis to identify the bottlenecks. Also, an attempt is made to visualize the downtime parameters e.g. frequency of breakdowns, total down time etc. But identification of bottleneck machines itself is not enough to improve the overall performance of the system. Predicting the nature of the behavior of the bottleneck machine and other machines of the production system are important. This leads to RQ 2.

RQ 2 : How can predictive analytics deliver value for tactical decision making process?

The second research question is stated in order to prove that the predictive analytics could also be done with the help of the real time data and the results from the predictive analytics could be used for fact based decision making process.

1.4.4. Delimitations

In order to limit the analysis of the study and ensure the adherence to the objectives, the following delimitations are made:

- The basis of the selection of production line for analysis is not within the scope of the thesis. The selection was done by the industrial partners of the research project
- The process of the data collection by the supervisory system on the machines during the production run is not explained
- The extraction process of the data from the supervisory system which monitors and stores the machine data is not described
- The thesis does not evaluate the different bottleneck analysis methods. Rather the focus is on how to use the different bottleneck analysis over the real time data

1.3. Structure of the Report

The structure of the report is as follows:

- Data Set 1
  - Chapter I 1 describes the frame of reference with a special focus on OEE and operator influenced tasks in a production system
  - Chapter I 2 explains the systematic methodology adopted to carry out the data analysis
  - Chapter I 3 presents the results from the data analysis and quantifies the operator disturbance portion of the OEE and identifies the major contributors to OEE
  - Chapter I 4 discusses the results with respect to the frame of reference emphasizing on the effects of operator influenced tasks on OEE
  - Chapter I 5 provides the conclusions drawn from the study
  - Chapter I 6 presents the scope of future work

- Data Set 2
  - Chapter II 1 describes the framework of reference with a focus on bottleneck detection, maintenance data modelling and predictive analytics
  - Chapter II 2 explains the systematic methodology adopted to carry out the data analysis
  - Chapter II 3 explains the experimental plan explaining the different bottleneck detection techniques, maintenance frequency and total down time and predictive analytics which are carried out on the real time data
  - Chapter II 4 presents the results from the data analysis with simple visualisation technique to identify the bottlenecks, data modelling of maintenance indicators and predictive analytics

- Chapter II 5 explains the creation of generalised algorithms for bottleneck detections discusses and the results with respect to frame of reference
- Chapter II 6 provides the conclusion drawn from the study and also presents the scope of the future work
- Chapter II 7 presents the future work

- Chapter 2 presents the reflections on two types of data analysis and overall conclusion emphasizing the importance of production data analytics to under pin new wave of productivity potentials

The report also includes five appendices which supplements the information contained in the report

- Appendix A: The losses description and the classification of the losses into the three levels of operator influence
- Appendix B: Layout of AAA and BBB Line

Data Set I

.

I.1. Frame of Reference

This chapter will present the frame of reference. Firstly, the fundamentals of Overall Equipment Effectiveness (OEE) are presented including the different methods to calculate the OEE. Thereafter, the theories of the operator influenced loss times are described successively. The existing benchmarking standards of OEE are finally reviewed.

In order to gain the understanding of OEE and the operator influenced loss times on OEE, a literature review was conducted. The literature were collected from the following scientific databases which are accessed through Chalmers Library

- Science Direct (sciencedirect.com)
- Google scholar (scholar.google.com)
- Scopus(scopus.com)
- ProQuest (proquest.com)
- Books 24/7 (books24x7.com)

To search for the relevant literature regarding the OEE and operator influence on OEE, the following key words are used,

- Overall Equipment Effectiveness
- Disturbances
- Machine interference
- Lean tools
- Operator disturbance
- Productivity
- Capacity
- Resource utilisation
- Operator utilisation
- Performance measurement

Additional literature were given from the supervisors and based on the most cited references in literature. These literatures are used to gain deeper insights on OEE and operator influence on OEE which are used to discuss and validate the results.

I.1.1. Overview of Overall Equipment Effectiveness (OEE)

Today some manufacturing plants started to closely monitor the production performance through Manufacturing Execution Systems (MES).The measurement of the manufacturing resources utilization indicates the equipment performance(Costa & Lima, 2002). OEE is an useful indicator and is seen to be the fundamental way of measuring performance efficiency in a comprehensive way(Puvanasvaran, Kim, & Siang, 2012). OEE is basically the ratio of the actual time the machine is producing by achieving the quality and specifications criteria to the time the machine is scheduled for production(Costa & Lima, 2002). OEE is also a measure of equipment availability, performance and the efficiency losses as a result from rework and yield losses (Nakajima, 1988) as shown in Equation 1.

OEE = Availability x Performance X Quality **Equation 1**

Hence, OEE could be viewed as a combination of operation, maintenance and management of manufacturing resources(Dal, Tugwell, & Greatbanks, 2000). On the other hand, there are some losses which reduce the performance of the equipment and it becomes important to study those losses. These losses are due to production disturbances. These production disturbances can be classified into two categories: Chronic and Sporadic disturbances as defined by Jonsson & Lesshammer, 2005. The former category is very difficult to identify as they are hidden in the production system. They are very small disturbances in the production system and they are seen in the normal state. The latter category is those which occur quickly and as a result there are large deviations from the current state. They occur irregularly and are much easier to detect than chronic disturbances. Comparing the chronic and sporadic disturbances, it is the chronic disturbance that leads to the lower utilisation of the machine.

There are six major losses to be addressed in the production system in order to achieve a higher OEE as defined by Nakajima, 1988. Those are,

- **Equipment Failure** : The losses are when the productivity is reduced and there is a volume loss due to repairing or replacement of machine parts to function
- **Setup and Adjustment** : These losses occurs from end of the production of one product and a changeover and setup of new tools takes place and the machine is adjusted to meet the requirements of the new product.
- **Idling and Minor Stoppage**: They occur when there is a temporary malfunction of the machine i.e. losses which could occur by the removal of abnormal work pieces etc. or when a machine is idling
- **Reduced Speed** : It is the difference between the machine design speed and the actual operating speed) i.e. the time loss when the standard cycle time of the machine is at 50 seconds and the actual operation takes 60 seconds, then the speed loss is 10 seconds.
- **Reduced Yield:** It is the time loss from the machine start-up to stabilisation. For example, time losses after a repair in the machine, time losses after lunch breaks etc.
- **Defect in the Process:** This causes loss of time and losses in quality of the product caused by malfunctioning of the production equipment

The term “losses” arise due to the chronic and sporadic disturbances which absorb these resources and hence contributing to the above six losses. The first two losses, the equipment failure and setup and adjustment are collectively known as downtime losses and are used to calculate the true availability of the machine. The losses, idling and minor stoppage and reduced speed, are used to calculate the performance efficiency of the machine. The last two losses, the reduced yield and the defect in process, are used to calculate the quality efficiency. The higher the number of defects the lower is the quality efficiency. Figure 2 is a representation of the detailed definition and calculation of OEE with all key losses which affects the final result.

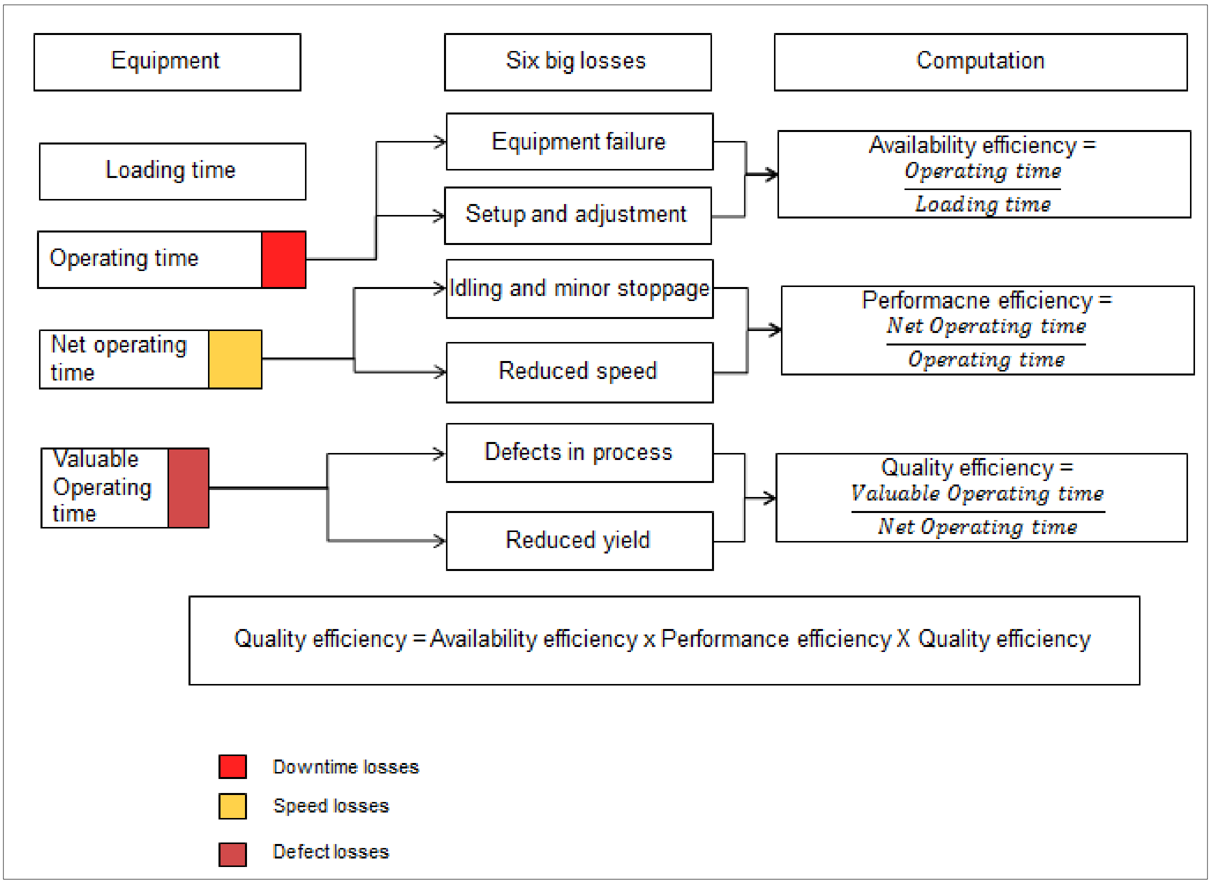


Figure 2: OEE Computation and Procedure (Adapted from Nakajima, 1988)

Though the definition of OEE remains the same, the calculation methodology of the factors: availability, performance and quality differs based on applications(Jonsson & Lesshammar, 2005). Table 1 shows the different calculations of OEE by two different authors Nakajima, 1998 and De Groote, 1995.

Table 1: Comparison of different ways to calculate OEE

Factors	Nakajima, 1988	De Groote, 1995
Availability (A)	$\frac{Operating\ time}{Loadng\ time}$	$\frac{Planned\ Production\ Time - Unplanned\ Down\ Time}{Planned\ Production\ time}$
Performance (P)	$\frac{Net\ perating\ time}{Loading\ time}$	$\frac{Actual\ amount\ of\ production}{Planned\ Production}$
Quality (Q)	$\frac{Valuable\ operating\ time}{Net\ operating\ time}$	$\frac{Actual\ amount\ of\ production - non - accepted\ amount}{actual\ amount}$

OEE is a good tool to calculate the utilisation of a machine for a given manufacturing process. However, the successful calculation of OEE is highly dependent on the input data(Jeong, Jeong, Phillips, & Phillips, 2001). The manufacturing process are unique and the nature of the machines in the different production systems are different and hence the method used to collect the data needs to be validated and justified in order to compute the OEE. The six big losses as shown in the Figure 2 corresponds to the equipment state(Jeong et al., 2001). Furthermore, every company may define their own equipment states of the machine which is based on the company's data collection ability and the level of accuracy needed. For example, capital intensive industries the losses can be classified into eleven categories(Jeong et al., 2001) as shown in Table 2.

Table 2: Ten categories of OEE losses (Jeong et al., 2001)

Serial number	Categories	Definition
1	Non Scheduled time	Time duration of which the equipment is not scheduled to operate
2	Scheduled maintenance time	Time spent for preventive maintenance
3	Unscheduled maintenance time	Time spent for breakdown
4	R&D time	Time spent for the purpose of research and development
5	Engineering	Time spent for the purpose of improvement activity
6	Setup and adjustment time	Time spent for setup and adjustment for operation
7	Engineering usage time	Time spent for engineering check-up
8	Work In Progress (WIP) starvation time	Time for which equipment is operating when there is no WIP to process
9	Idle time without operator	Time for which WIP is ready, however there is no operator available
10	Speed loss	Time loss due to equipment that is operating under standard speed
11	Quality loss	Time for which equipment is operating for unqualified products

Jenong et al. 2001 combines the ten categories of losses and used Nakajima, 1988 approach and proposed a new method to calculate the OEE as shown in Figure 3. The OEE has the factors time efficiency, speed efficiency and quality efficiency. The time efficiency includes non-scheduled maintenance, un- scheduled maintenance, R&D usage, Engineering time, setup and adjustment, WIP starvation, idle time without the operator and other loss times. The speed efficiency includes the speed loss and the quality efficiency includes the quality loss.

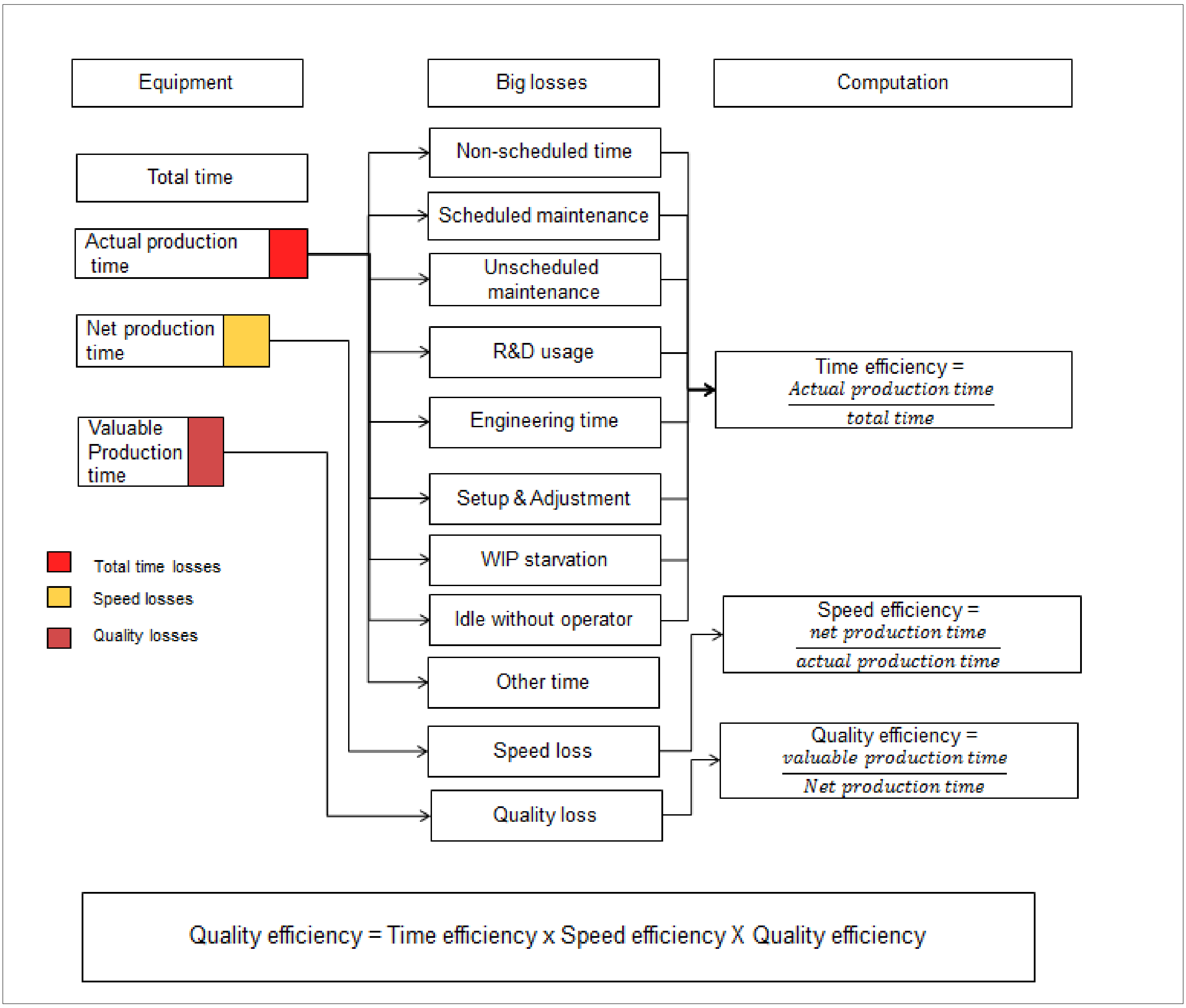


Figure 3: Calculation of OEE (Adapted from Jeong et al., 2001)

The planned production time(De Groote, 1995) which is a part of the total time could be calculated in different ways by grouping the losses differently(Andersson & Bellgran, 2015). Figure 4 explains different ways of calculating the planned production time.

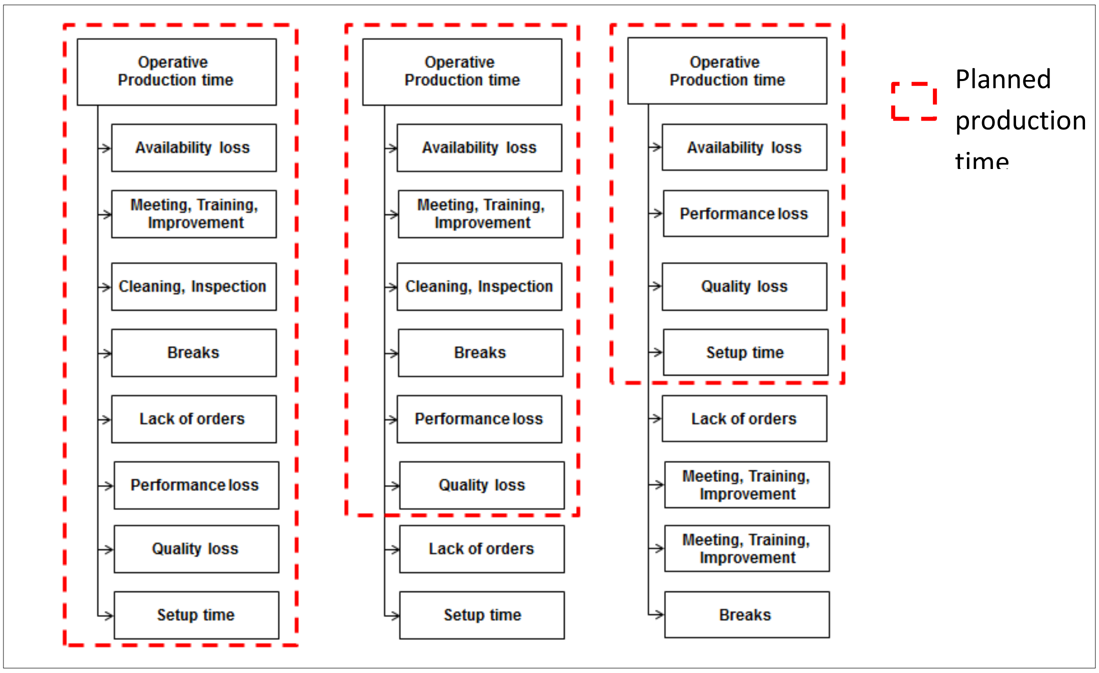


Figure 4: Different scenarios for classifying the losses to calculate planned production time (Adapted from Andersson & Bellgran, 2015)

From Figure 4, it could be inferred that the planned production time may include all the losses or it might exclude the losses which are due to the company policy, organisation losses etc. For example, breaks, cleaning and inspection, meetings, trainings and improvements etc. are planned production losses under the total time for production and these could be excluded for the calculation of planned production time. The other planned production losses may include waiting due to completion of current orders, Total Productive Maintenance (TPM), planned maintenance activities(Costa & Lima, 2002).

There are many benefits in increasing the OEE by eliminating the losses and one important benefit is the increase in capacity without major investments (Konopka & Trybula, 1996). There is a strong link between the OEE and the financial ratios of the company (Hansen, 2001). An increase in OEE from 60% to 66% of a company gave a 21% increase on Return of Assests (ROA), increased the capacity by 10% and increased the operating income by 21 % (Hansen, 2001).

I.1.2. Operator Influence on OEE

Operators play a significant role in the production shop floor. However, the operators are assigned to multiple machines to increase the utilisation of them(Chien et al., 2013) e.g. one operator is responsible for operating several machines or one operator may be responsible for carrying out maintenance related activities in several machines. When a machine stops or breakdowns etc., it will not start to produce until the problem is fixed by the operator(Stecke & Aronson, 1985.). As one machine can be serviced at a time by an operator and if other machines in the same production line is down at the same time, then the remaining machines need to wait for the operator until the first machine is repaired to resume its production. This will offset the utilisation of machines. This waiting time of machines is termed as machine interference time (Stecke, 1982). If this interference time is large, it reduces the available capacity of the machines (Desruelle & Steudel, 1996) by affecting the machine efficiencies and production rates (Stecke & Aronson, 1985).

The machine interference problem in the production lines could be solved by assigning the correct number of machines to an operator(Stecke & Aronson, 1985.). However, when assigning too many machines to an operator to improve the operator utilisation, then the operator is overloaded with large number of machines. According to Stecke (1982), this will lead to operator fatigue which might slow the pace of working by the operator. If the operator tries to speed up the pace of working, it might lead to quality problems or defective output as a result of which the production could go down. On the other hand, assigning too few machines to the operator will lead to minimal efforts needed to keep the machines running which results in idleness of the operator and work standards also drops (Stecke, 1982).

The machine interference time is highly influenced by the work tasks of the operator carried out during the interference time(Almström, Hansson, & Samuelsson, 2014). Apart from the direct causes of the OEE which are affected by the factors like breakdowns, setups etc. one important factor that affects the OEE is the operator work tasks(De Ron & Rooda, 2006). The operator work tasks during the disturbance however may not completely be influenced only by the nature of the disturbance like breakdown, setup etc. but also includes the activities outside the actual workflow like searching of tools, documents etc. (Hedman, Sundkvist and Almstrom, 2014). These activities interrupt the workflow during a activity. By reducing the time spent on activities like searching etc. the machine interference time could be reduced. The time periods for those specific work tasks of the operator can be seen as an improvement potential from work place design and standards point of view (Hedman, Sundkvist and Almstrom, 2014). Also, if the nature of work tasks which are completely influenced by the disturbance, for example setup time activities, are not standardised, then the time taken to perform those tasks will increase the interference time.

In one example case study described by Hedman, Sundvikst and Almstrom (2014) in an automatic surface mount assembly, it is noted that the preparations of the setup was done when the machines were waiting for changeover. This could be seen as an improvement potential as some of the internal activities of setup could be converted into external activities. One established lean technique to reduce the setup time is using Single Minute Exchange of Dies (SMED) system. This technique is widely used by the industries in order to increase the OEE by reducing the changeover time i.e. reduce the time of the machine when the machine is down due to setup activities. Shingo, 1985 divided the setup operations into two types: internal operations and external operations. Internal operations can be performed when the machine is stopped and the external operations are performed when the machine is operating. But converting most of the internal setup operations to external, leads to less machine interference time as machine waiting time for the setup is reduced. This will also increase the utilisation of operators. On the other hand, the standardisation of the work activities during the actual setup time will result in lower downtime of the machine and hence less interference time. One more key finding by Bergstrom and Palmvkvist (2014) is that the majority of the losses can be related to the manual work and hence the development of work and time standards is more important to reduce the time spent on the disturbance handling. Apart from SMED, which is used to specific for the setup time, other lean tools like Andons, Kanbans (Liker, 2014) etc. could be used in order to decrease the failure reporting time by the operators, material replacement time etc., which will reduce the machine downtime.

The machine interference time is the valuable time lost under the planned production time and hence leads to low OEE. Studying the operator influenced work activities becomes more important in order to minimise the machine interference time and maximise the OEE.

I.1.3. OEE Benchmarks

The benchmarking is defined as a systematic approach through which the organisations can measure the performances against the best in class organisations (Attiany, 2009). There are many different types of benchmarks according to Andersen (1999). Those are internal benchmarking, competitive benchmarking, functional benchmarking, generic benchmarking, and performance benchmarking, process benchmarking.

OEE can be used as a performance benchmark between the manufacturing industries for measuring the performance of manufacturing(Dal et al., 2000). There are different opinions regarding the acceptable OEE performance(Dal et al., 2000). An OEE measure greater than 50% is considered to be more realistic and an acceptable target for manufacturing industries (Kotze, 1993). The acceptable OEE performance can vary between 30% - 80% (Ericsson, 1997). Since there are varying norms across industries in capturing and accounting of the losses, it would be difficult to establish an optimal reference for OEE(Dal et al., 2000) and difficult when using OEE for external benchmarking(Jonsson & Lesshammar, 2005). Also, OEE aggregates to a larger extent and could be inappropriate for benchmarking (Liker, 2014).

According to Nakajima (1988), under ideal conditions the organisations should have availability greater than 90%, performance greater than 95% and the quality rate greater than 99%. Thus the resulting OEE should be greater than 84% as shown in Equation 2 and is a good benchmark for a typical manufacturing capability.

$$90\%(\text{Equipment Availability})\text{x } 95\%(\text{Performance efficiency})\text{x } 99\%\ (\text{Rate of quality}) = 84.6\ \%\ \text{OEE} \qquad \textbf{Equation 2}$$

Also, under ideal conditions, a batch type production unit have a world class OEE greater than 85%, discrete process is greater than 90% and for continuous process have OEE greater than 95%(Hansen, 2001).

I.2. Methodology

The aim of the analysis is to find the current OEE of the 23 companies and the operator influence on OEE and this analysis was carried out on a data set representing the production losses given by the company Good Solutions AB. The production data of the machines was collected by Good Solutions AB in agreement with various Swedish Companies.

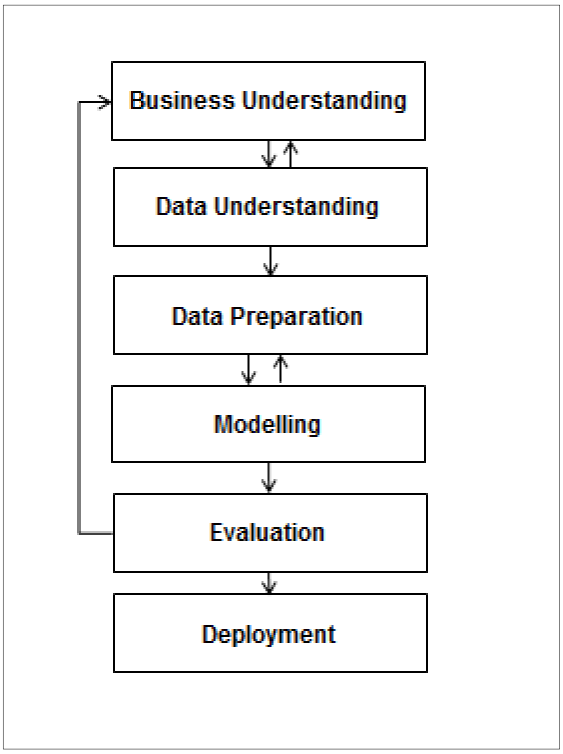


Figure 5 : Phases of CRISP – DM Model (Adapted from Shearer, 2000)

The CRISP-DM (Shearer, 2000) model as shown in Figure 5 was used as a reference model in this thesis. The six phases of the model which are business understanding, data understanding, data preparation, modelling, evaluation and deployment are explained below.

a. Business Understanding

   The main objective of this data analytics is to determine the current OEE and to build a model to assess the impact of operator influenced loss times on OEE. The success of this can be measured by validating the models by comparing it with the previous research on the same topic and by getting the feedback from the participants of the research. The data quality was a constraint in this thesis.

b. Data Understanding

   The single MS Excel data file which had the data rows for the machines in each company was provided. This file had the losses description of each machine and also had the time lost for each loss. The time lost was in the form of elapsed time and not the event times. The data was provided for six months starting from October 2013 to March 2014. The data file had 1339 data rows in total which had the information of 23 companies and 884 machines. The file had 29 columns of description recorded for each machine.

c. Data Preparation – Data Analysis

As a first step, each data column was studied to understand their meaning. As a step towards cleaning the data, the data column named "not ok" in the data file indicated that there were few rows in the file which had junk data. These data rows were revisited and found that there were no data recorded in these data rows. As a result these data rows were deleted in order to ensure an error free qualified data is available for modelling. As a second step, the months for which there is no data available was checked for each machine and found that there were no missing data for the month. Thirdly, the duplication of data rows were checked and found that there were no rows of data repeated. Finally, after performing these steps, a clean and qualified data was available for modelling.

d. Modelling

OEE representation model and the operator influenced tasks model was built from the clean and qualified data. The detailed steps in building the model are defined in the Results section.

e. Evaluation – Validation

OEE representation model and the operator influenced tasks model was built from the clean and qualified data. The clean data means that there are no junk values in the data set. Qualified data means there is no repetition of the data rows and no data rows have any missing value against each factor recorded. The detailed steps in building the model are defined in the Results section.

f. Deployment

From the outcomes of the model, recommendations could be made on how the OEE could be improved by better utilising the manufacturing resources. Additionally, the quantification of link between the Operator influenced loss times and the OEE is established emphasizing on the standardisation of the operator work tasks.

I.2. Results and Analysis

In this chapter, the empirical findings from the data analysis are presented. In addition to the empirical findings from the data, OEE levers framework was developed from the theoretical study and the data file.

I.2.1. Calculation of OEE

The data collected from 884 machines of 23 companies are categorised into four industrial groups based on the nature of production system and the products they manufacture.

- Food and Beverage
- Mechanical Workshop
- Other Automated Discrete Production
- Polymeric (Rubber and Plastics)

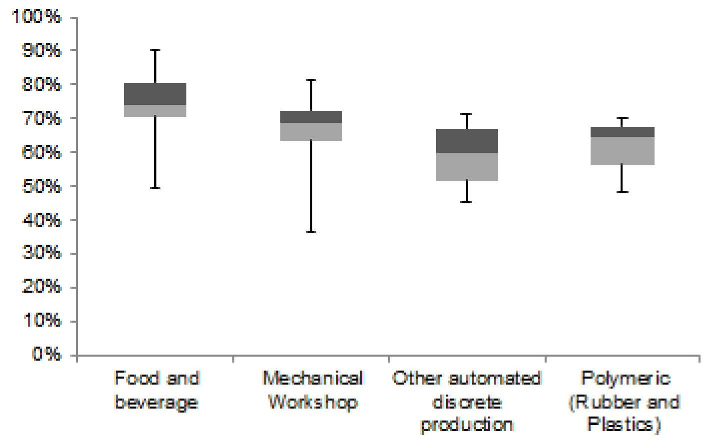
The number of companies and the number of machines under each industry group is presented in Table 4.

Table 3: Number of companies and machines in each Industry Category

Industry	Food and Beverage	Mechanical Workshop	Other Automated Discrete Production	Polymeric (Rubber and Plastics)
Count Companies	7	9	4	3
Count Machines	244	364	119	157

From the data file, it can be observed that the performance rate recorded for 702 machines out 884 machines has the value 100% and 796 out of 884 machines has a recorded value of 100% for quality rate.

The distribution of performance in OEE across the various industrial groups as shown in Figure 6 contains a deep set of rich data. It is to be noted that the OEE measures have been calculated in relation to the scheduled production time. Also, the average and the median OEE are calculated for all industry groups. The mean is simple the average whereas the median is the middle number of the series arranged in a rank order. From Figure 6, it could be seen that across industries, the OEE distribution is not symmetrical and using average will not reflect the true average as it is significantly influenced by the outliers. On the other hand, the median is also a form of average which gives a better idea of the central tendency of the data.

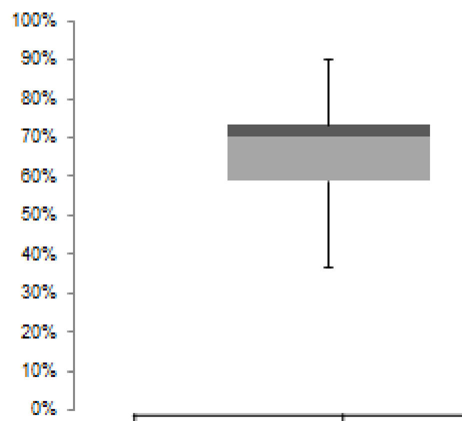


Industry	Food and Beverage	Mechanical Workshop	Other Automated Discrete Production	Polymeric (Rubber and Plastics)
Average OEE (%)	74	65	59	61
Median OEE (%)	74	68	59	65
Top Performance (%)	90	82	71	70

Figure 6: OEE comparison of different industry groups

It could be inferred from Figure 6 that Food and Beverage Industry type has the highest median OEE of 74% with most of the companies performing above the median OEE. The lower quartile is tightly grouped. Mechanical Workshop Industry type has a median OEE of 68% which is the second highest among the four industry types. The lower quartile group is spread out indicating there is a room for potential improvement. Other Automated Discrete Production Industry type has the lowest median performance of 59% .Adding on; it has a broad distribution in performance indicating a higher variability and a less consistent performance. Polymeric (Rubber and plastic) Industry type has 65% as median OEE and top performance as 70%. Adding on, it has a tighter distribution compared to other industry types. Also, it has most of its performance well below the median indicating skewness in the performance and there are improvement potentials to increase the average performance up to the median level at 95% confidence level.

Aggregating the OEE of all industry groups, the overall OEE of the population is determined. This is shown in the below Figure 7.



Industry	
Count companies	23
Count Machines	884
Average OEE %	65
Median OEE %	70
Top performance %	90

Figure 7: Overall OEE of 23 companies

The overall median OEE of 23 companies is 70% whereas the average OEE is 65%, indicating a positively skewed performance with more spread in the lower quartile region. Also, it could be said that there are improvement potentials to raise the overall performance by arresting the various losses.

I.2.2. Classification of Losses

From the literature review, it is inferred that, the total production time could be divided into two categories: planned down time and planned production time as shown in Figure 8. The planned down time is the time loss due to the management policies or organisation policies. These includes non-scheduled time (e.g. weekends, days not scheduled due to lack of orders), scheduled maintenance (e.g. Planned maintenance activities, TPM, machine cleaning and operator maintenance), R&D usage(e.g new equipment installation and trials), engineering time(e.g. process improvement activities), breaks, meetings and operator trainings .

The planned production time includes the time spent for producing good products and the time lost in unplanned downtimes. These unplanned downtimes include setup time, measurement and adjustment, equipment failure, idling and minor stoppage, scrap and rework. The setup and adjustment loss could be considered as unplanned downtime loss as these losses could also occur from poor production planning practices.

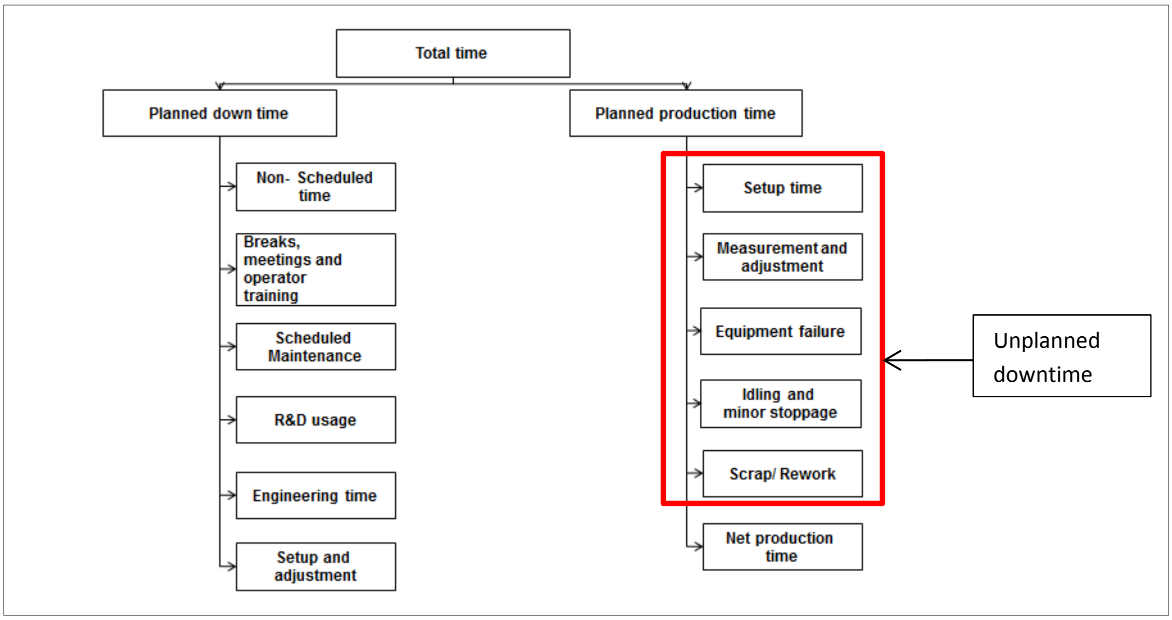


Figure 8: Classification of losses to calculate the OEE

Furthermore, from the data file, the main categories of unplanned downtime time losses as shown in Figure 8 is further divided in sub-categories in order to get deeper insights on the specifics of the loss categories. This is shown in Figure 9. Also, the categories *other down time losses* and *unclassified losses* are included to capture the losses from the data file which cannot be classified under the loss categories as shown in Figure 8.

Planned downtime	Setup time	Measurement And Adjustment	Equipment Failure	Idling and Minor stoppage	Other down time losses	Scrap / Rework	Unclassified
Breaks	Changeover time	Measurement And Adjustment	Breakdowns And Maintanence	Blocked / Starved	Lack of Materials (Internal)	Quality Problems	Unclassified
Education and Training			Other equipment error	Idle time with operator	Lack of Materials (External)	Rework	
Operator Meetings				Minor loss / other loss	Lack of operators		
Planned Shutdown					Planning loss		
Preventive maintenance					Start-up losses		
R&D and Engineering Time							
Shift Changeover							
Shutdown due to no demand							

Losses derived from literature

Grouped Losses from data file

Figure 9: Classification of production losses

In the Figure 9, the planned downtime, setup, measurement and adjustment, equipment failure, idling and minor stoppage, other down time losses, scrap/rework, unclassified are the losses derived from literature. The other sub categories of the main losses are the losses grouped from the data file.

Figure 10 displays the components of OEE for the overall industrial groups showing key improvement opportunities. The planned downtime losses are excluded in the calculation as those are planned production losses and are influenced by the company management policies. The other seven categories are collectively called as unplanned production downtime. As stated, the OEE is calculated on the basis of planned production time.

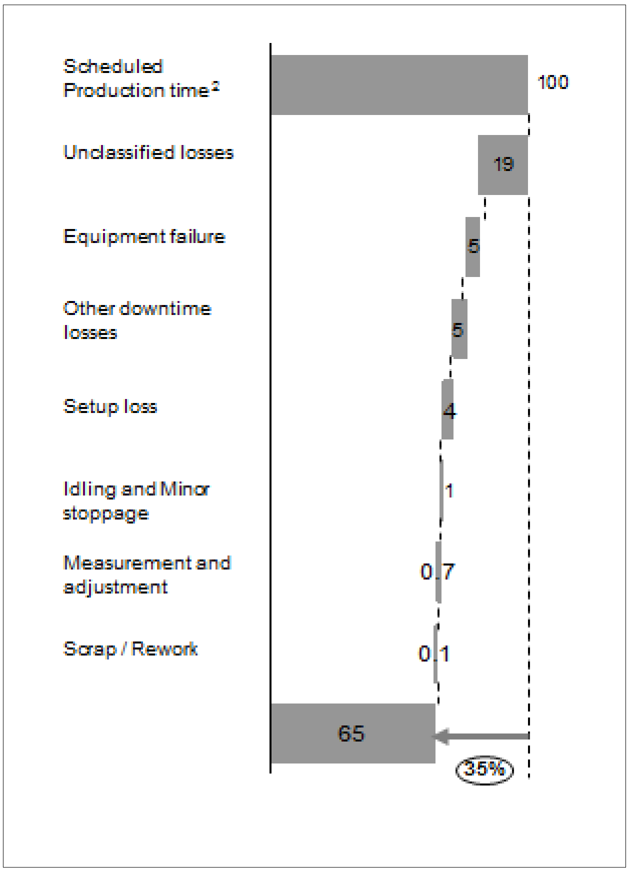


Figure 10: Distribution of OEE levers as a percentage of scheduled production time in percentage

About 19% of the loss times are not classified or are poorly described in words in the data file which makes the classification difficult. For example, the losses with description like “reason missing”, “other”, “uncategorized”, “not decoded stop”, “false stops”, “portal” etc. The second major loss is the equipment failure which is 5% and other downtime losses which is also 5%. Moreover, combining the setup losses and measurement and adjustment losses which could be due to poor quality of setup activities constitutes 4.7% which is the third major loss.

Combining the factors of OEE (Equipment failure (5%), other downtime loss (5%), set up loss (4%) and measurement and adjustment loss (0.7%)), the availability is calculated. The availability is 85.3%. The performance efficiency is calculated from the idling and minor stoppage (1%) and the unclassified losses (19%). The unclassified losses could be included in the performance efficiency calculation under the assumption that those losses occurs very frequently and the variety of the losses is more for them to be classified under the standardised losses. These performance losses in turn affect the machine performance. Then, the performance efficiency is 80%.

The quality efficiency is calculated from scrap and reworks (0.1%) and is 99.9%.

I.2.3. Effects of Operator Influenced Loss Times on OEE

The work tasks carried by the operator during the machine interference time have a direct effect on OEE. The work tasks of the operator not only depends on the nature of the disturbance like setup, breakdown etc. but also includes the other operator tasks like searching for tools during the disturbance etc. It is important to note that the operator themselves don't cause the losses but when the losses occur, the operators is a crucial factor in influencing the loss time. The effects of operator influence loss times on OEE were determined by classifying the 499 production losses times which are described in the data file into three levels: operator influenced loss time, may be operator influenced loss time and not operator influenced loss times. The complete list of losses is found in Appendix A.

- ***Operator influenced loss time*** are those tasks consisting of manual activities and operators has the power of causing an effect on that task in a direct or indirect way. For example, the loss 'mechanical failure' time is assumed to be captured by the company from the start instant of the failure till the time the machine is up again for production. This elapsed time between the start of the failure till the up time of machine is completely dependent on the tasks performed by the operator i.e. the time when the machine is waiting for operator to be repaired and the repair times are the aspects influenced by the operator as shown in Figure 11. Hence this activity is completely operator dependent. Some of the other operator influenced loss times are machine error, pneumatic error, search, order replacement with setup, product change, checking and adjusting.

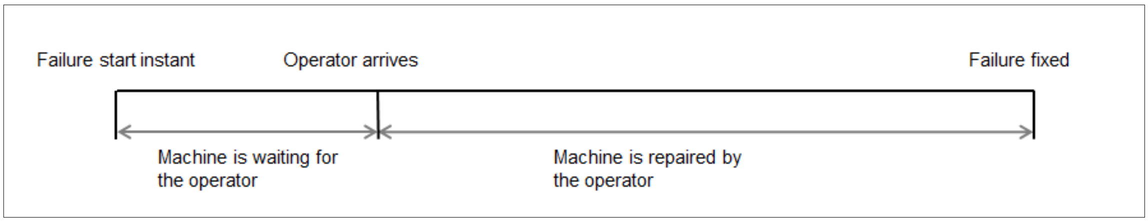


Figure 11: Representation of an equipment failure fixation in time instant

- ***May be operator influenced loss time*** are those in which the operator may have the power of causing an effect on the tasks. For example, the loss window shortage could be due to two aspects i.e. the operator didn't fill the pallet with windows for the production or no windows is available in the inventory to be refilled. Hence, with this type of loss where the loss time could be as a result of operator tasks is classified as May be Operator Influenced loss time. The other examples of may be operator influenced loss time description are stop previous shift, lacking input, fully in conveyor buffet, materials wait, waiting jobs, packaging machine – outlet, waiting for P4, internal materials, micro stop, return materials

- ***Not operator influenced loss times*** are those losses description, from which it can be inferred for sure that the losses are completely external to the scope of the operator, then those losses are classified as *Not Operator Influenced* loss time. For example, the loss external material missing, is the loss time which is not within the scope of operator and is due to the supplier issue. The other examples include lack of materials from supplier, purchased material, heating loss, start-up losses due to machine, external delivery.

Classifying the tasks based on the operator influence and grouping them under each OEE component category is shown in Figure 12.

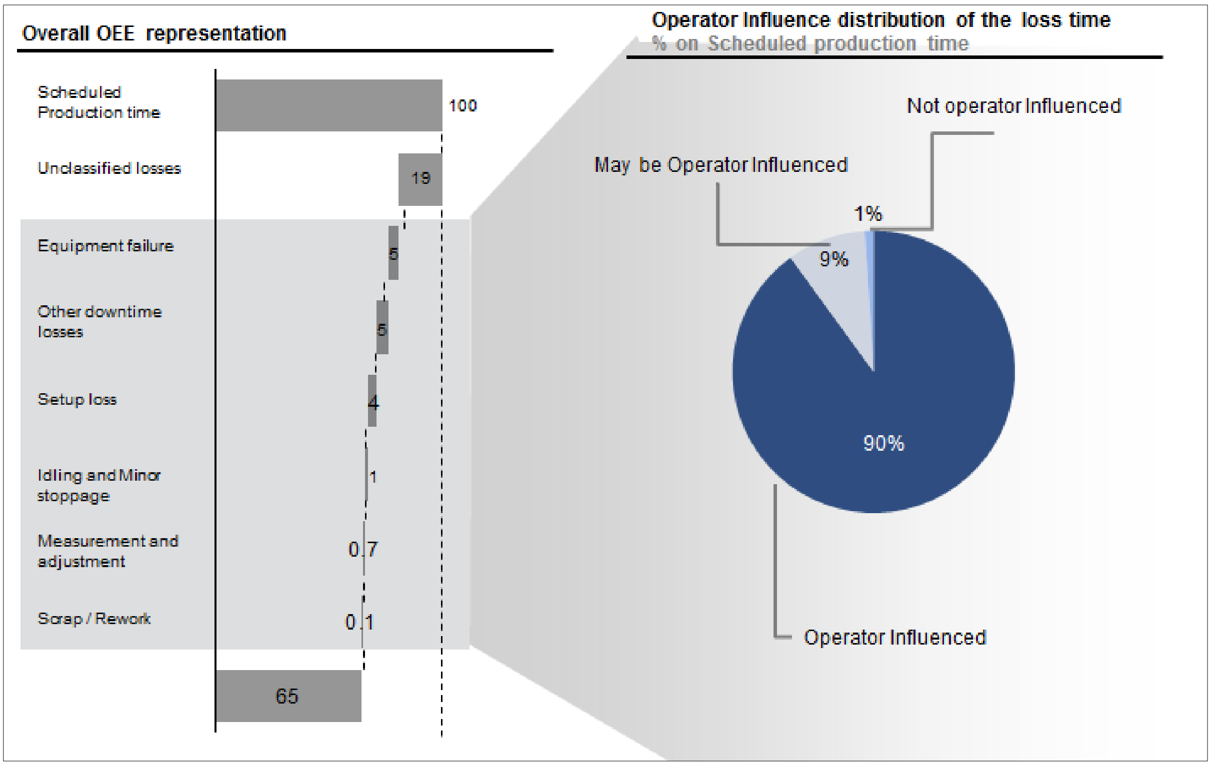


Figure 12: Classification of loss time into level of operator influence

Figure 12 show the operator influenced work tasks under the unplanned downtime losses: equipment failure, other downtime losses, setup loss, idling and minor stoppage, measurement and adjustment and scrap/rework. The unclassified losses are excluded as it lacks clear description of the loss in order to understand the cause of the loss. From Figure 12 it could be inferred that the operator influenced loss time are 90% of the scheduled production time. Adding on, the loss time which may be influenced by the operator corresponds to 9% of the scheduled production time. Only 1% of the loss time is not influenced by the operator. Furthermore, the *operator influenced* and *may be operator influenced* is considered as one group in the further analysis because the latter category is very small compared to the former category.

Visualising the overall operator influenced tasks time under each factor of OEE is shown in Figure 13. It could be inferred from Figure 13, that the operator influenced tasks times are more in equipment failure (35%), other downtime losses (27%) and setup loss (26%). The other three factors: Idling and minor stoppage (4%), measurement and adjustment (8%) and scrap/rework (0.7%) contains less number of operator influenced tasks compared to other factors. However, it can be noted that the measurement and adjustment loss times could be due to poor setup.

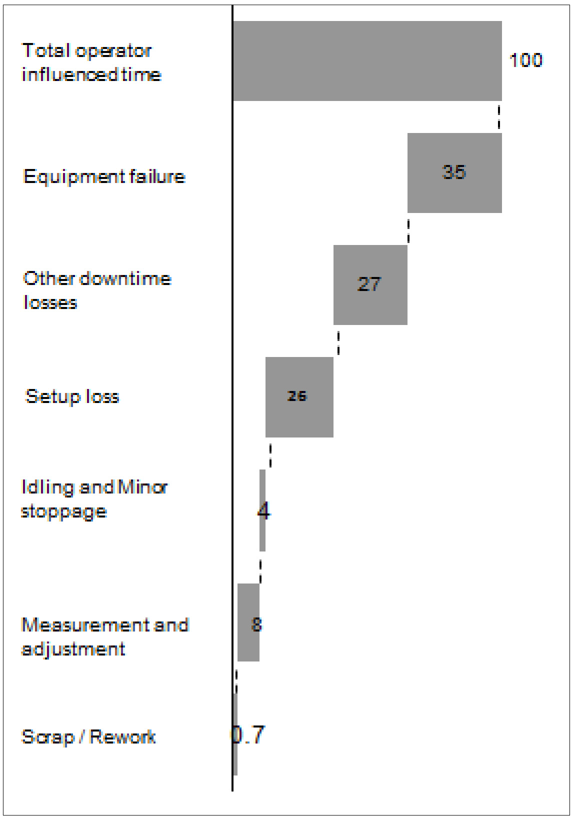


Figure 13: Representation of operator influenced time of each OEE lever in percentage

Taking the operator influenced loss times under each component of OEE and visualising it across the different industry types as shown in the Figure 14, shows the major operator influenced loss time under each industry type.

The following are the inferences from Figure 14:

Food and Beverage industry type has 46% of operator influenced production loss time as equipment failure. The second highest is the setup time which is 41%. The number of companies for which the data has been collected for this industry type is seven and all the companies has the changeover time data recorded. Five out of seven companies only monitor the breakdowns and maintenance, and other equipment error under the equipment failure category. Taking this aspect into account, it can be said that the OEE component of the operator influenced setup time is more significant loss time in this industry group.

Mechanical Workshop industry type has 47% of the operator influenced production loss time as set up time and 38 % under the equipment failure. Eight out nine companies monitor changeover time and hence changeover is the major component in this Industry type. On the other hand, under the equipment failure component of OEE, six out of nine companies have breakdowns and maintenance time recorded and eight out of nine companies have the loss other equipment error loss. So taking this aspect into consideration, along with change over time, other equipment error aspect under equipment failure are the most significant loss times in this industry group.

Other Automated Discrete Production industry type has 38% of the operator influenced production loss time as set up component of OEE. The second highest is the equipment failure which constitutes 28%. The data has been collected from four companies under this Industry type. All four companies have the data for the changeover times. But under the equipment failure component only two companies monitor the breakdowns and maintenance. Also, it could be noted that, the other downtime losses accounts for 24% and this is due to the lack of operator time which is recorded as a loss time by all the companies. So the significant loss time in this industry group is the operator influenced set up time, breakdowns and maintenance and lack of operators.

Polymer (Rubber and Plastic) industry type, 65 % of the operator influenced production loss time as equipment failure. The second major, OEE component of operator influenced time is other down time. This is different compared to other industry groups. Out the total of three companies, all the three companies record the time lost due to breakdowns and maintenance. On the other hand, all the three companies have the lost time due to lack of operators and lack of internal materials. Hence, breakdowns and maintenance and planning losses which includes manpower and materials are the significant operator influenced loss categories across this industry type.

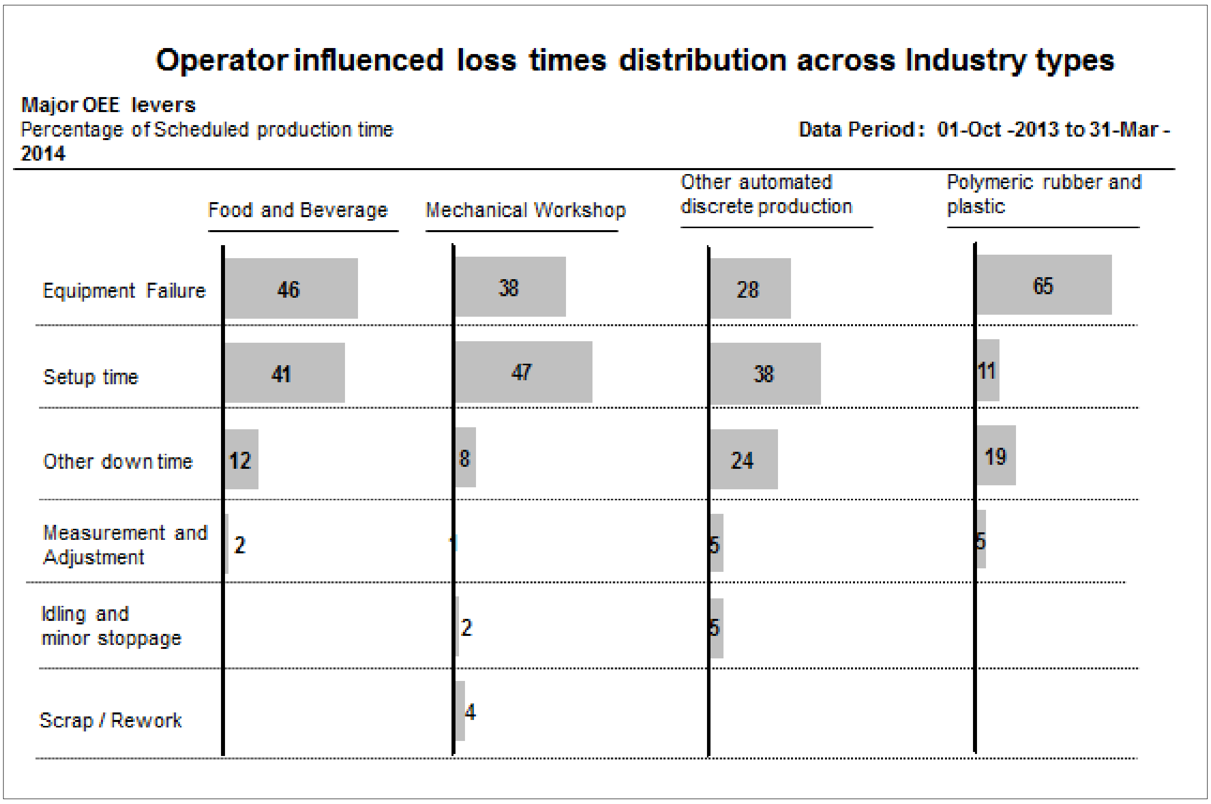


Figure 14: Classification of operator influenced loss times grouped into OEE components across industry groups

I.4. Discussion

In this section the results from the empirical data analysis of OEE are discussed in relation to literature. This is done in order to explain the findings and validate them.

I.4.1. OEE calculation

According to (Muchiri & Pintelon, 2008), "*the OEE is a measure of total equipment performance, that is the degree to which the equipment is doing what is supposed to do*". OEE is a result of equipment availability efficiency, performance efficiency and quality efficiency.

The availability is a measure of the machine uptime against the planned production time. The machine uptime is the difference between the planned production time and the unplanned downtime (De Groote, 1995). From the empirical data, the loss times: breaks, education and training, operator meetings, planned shutdowns, preventive maintenance, R&D and engineering time, shift changeover and shutdown due to no demand are excluded in the calculation of the planned production time. These loss times are excluded in the analysis under the assumption that all these losses are due to organisational or company policy influenced and these loss intervals are usually paid for the operators. However, the explicit information about the paid and unpaid time intervals cannot be interpreted from the empirical data set. On the other hand, capturing some of the above mentioned loss times under planned production time will expose the opportunities for improvement. For example, it could be possible to schedule the production line using relief operators during the breaks, operator meetings or the machine could be loaded with enough material through the breaks, operator meetings, and shift changeover so that the machine keeps on working for that short time interval. These improvements will increase the OEE. Furthermore, the setup loss is regarded as loss but not as productive time. This is because the total setup time is a factor of number of setups and the average time per setup. The number of setups depends on the batch size; the average setup time depends on nature and the skill level of operator's work tasks which is a potential improvement factor in order to reduce the average setup time.

The performance efficiency depends upon the standard cycle time of an operation and the actual cycle time of the operation. It could also be defined as the ratio of actual amount of production by planned production (De Groote, 1995). The empirical data set didn't have any information on the actual production from the machines or the actual cycle time data. Also, there was no information on standard cycle time. Hence it wasn't possible to calculate the performance losses. Moreover, the empirical dataset had the performance rate defined for each machine. The values for this particular factor were recorded as 100% for 702 machines out of 884 machines which is 80%. This means either the machines perform at its theoretical maximum speed or the companies do not measure the cycle time or the production speed and the default value recorded was 100% or the speed was measured with the most skilled operator working on the machine. If the production speed is 100%, then also the performance factor could be improved by increasing the speed of the process as stated by Nakajima (1988) that *"if speed loss is 100%, then bring actual operation speed up to design speed; then make improvement to surpass design speed"*. However, if the process could be made faster, then the standard cycle time should be proportionally reduced. Lunjberg (1998) reported that many companies are not aware and does not focus on performance loss which is identical from the inference made from the data set in this study.

The Quality efficiency is defined as the accepted amount of production to the actual amount of production (De Groote, 1995). The empirical data set didn't have any information on the scrap or the total amount of production recorded for each machine. Thus the quality efficiency calculation as defined by De Groote cannot be applied. Moreover, the empirical data set had quality efficiency values defined for each machine. 796 out of 884 machines had values 100% which means either the all the machines produced acceptable products and hence there is no scrap or the quality factor was not measured for the machines and the default value recorded was 100%.

If the OEE is calculated as per Nakajima (1988) with the performance and quality rates as 100% then, the resulting OEE is only a true measure of availability. On the other hand, the empirical data file had the losses description and those losses could be classified in to the main loss categories as described by (Jeong et al., 2001) and (Nakajima, 1988) to calculate the availability, performance efficiency and the quality efficiency. The availability is calculated from the losses: equipment failure, set up loss, measurement and adjustment and unscheduled maintenance time(Jeong et al., 2001). The performance efficiency is calculated from the losses idling and minor stoppage, reduced speed(Nakajima, 1988) and the quality efficiency is calculated from the time spent on producing the scrap or rework(Jeong et al., 2001). Hence it could be inferred that for the given empirical data set, the standalone methods: Nakajima (1988), De Groote (199)5 and Jeong et al. (2001) method of calculating OEE could not be directly applied but with a combination of these methods the OEE could be calculated.

The average OEE of 884 machines of 23 companies is 65%. The top loss contributors to the OEE are the unclassified losses which are 19%. The unclassified losses are those loss descriptions which are poorly described and are difficult to classify. This indicates that the companies need to pay more attention to capture the losses in a standardised manner. The calculation of the losses and the collection of them seem to be easy as described by the theory. Though MES is a very good tool to collect the data of a machine, if the descriptions of the losses are not clear, then it becomes hard to interpret the underlying reasons for the losses. Though the MES tracks the length of the stoppage time, the reason for the stop as interpreted from the loss descriptions in the empirical data set reveals that the operator records the reasons for the losses in the MES. Ljunberg (1998) described that the computerised systems capture both the frequency and the length of the different stoppage times and the reason for the stoppage is described by the operator. It could be inferred that this trend remains the same over the past decade. In order to capture the reasons for the losses precisely, standardised losses descriptions should be incorporated into MES. This facilitates not only easy interpretation of losses in a big data but also enables to benchmark the production system within and across similar production systems of different companies. Also, this unclassified loss distorts the exact impact of other OEE factors. Hence, companies should start capturing the losses in standardised manner to precisely measure the OEE.

Yet another important point to note is that the empirical data collected from different companies was only for a period of six months. However, this has an impact on the final OEE measure. As found in the results, the OEE of the 23 companies is 65 % and this measure is got only by collecting the data for six months. These six months could be the best months where the demand was very high for the company. So, in order to cover the demand pattern, at least one year data needs to be collected to absorb the seasonal variations in the demand. Moreover, the OEE calculated from the empirical data set is a result of aggregation of various machines in one company. This according to

Liker (2014) is the not the appropriate method of calculating OEE. But as neither the production flow nor the bottleneck machines are indicated in the empirical data set, this method of aggregating the individual machines data to represent the OEE of the company was made.

I.4.2. Effects of Operator Influenced Loss Times

The losses from the empirical data were classified into *operator influenced loss times, may be operator influenced loss times and not operator influenced loss time*s based on their meaning. It can be seen that 90% of the total loss time is operator influenced, may be operator influenced tasks are 9% and not operator influenced tasks are 1% in relation to the planned production time, and not including the unclassified losses. From this it could be inferred that, the operators could influence the loss times by working in a non-systematic way which leads to increase in the total downtime. However, the operator influenced loss times, which contains the operator response time to the disturbance and the actual operator time addressing the disturbance, are not distinguished explicitly from the empirical data. One more interesting inference which could be drawn when comparing among the operated influenced loss times among different industry groups is that, the operator influenced loss time is more in setup across industries. Higher setup time, could be argued for higher number of setups. On the other hand, it could be explained that the average setup time per changeover could be high and this could be due to the operator influenced tasks. The other OEE lever which is high across industry groups is equipment failure. Though the nature of the breakdowns determines the time required to repair the machine, the failure reporting process and nature of the operator tasks during the repair time could be influenced by the operator. Overall, the operator influenced loss times could be completely related to the operator work tasks. Also, one key note is that the OEE is an aggregated measure and does not explain the operator influenced loss times separately.

These findings raises the question regarding to what extent the current operator tasks are standardised. Also, these results show an indication that there are chunks of unproductive time which is wasted within the planned production time. Thus there is a significant improvement potential in decreasing the machine interference time by developing the work standards for the tasks like setup, equipment failure etc. The importance of developing the work standards is that the variations in the time taken by the operators during setups, equipment failure etc. are reduced. (Hedman, Sundkvist and Almstom, 2014). In case of setup times, though the number of changeovers is not under the production department control, the average time per setup could be decreased by applying SMED technique which also involves creating standardised tasks for setup activities (Shingo, 1985). For example, many machines in Food and Beverage industries may not designed for easy changeover (as it is a process industry and may require a lot of cleaning between the products changeover) and may involve complex setup techniques which could be operator influenced. In this case SMED could be applied to simplify the setup process and to reduce the setup time. In case of equipment failure, the failure reporting process which is an operator influenced task could be made faster for example by pulling the Andon cord (Liker, 2004).

Certain loss times like search time, replenishment time are present in all industry groups. These also could be viewed as unproductive times under the planned production time. These times are completely dependent on operator. Implementation of 5S and Kanban system to reduce the search time and the material replenishment time respectively (Liker, 2004) could eliminate these operator influenced loss times.

One more key observation from the results is that, the equipment downtime time and the operator influenced loss times are proportional. These results are in line with the previous research conducted by Hedman, Sundkvist and Almstom (2014). Though, the results of the research are based on work sampling method of a particular machine in one company, the results from this study of empirical data of 884 machines also points towards the same pattern of the findings between the operator dependent tasks times and the machine down times. As the results from the empirical data also signify the same, it cross validates the findings.

Addressing the above mentioned aspects would improve the operator influenced tasks and thereby decreasing the machine interference time and thus improving the OEE. By improving the OEE, the real capacity could be increased (Hansen, 2001).

I.4.3 OEE Benchmarks

The OEE of the 23 Swedish companies was around 65% against the world class benchmark under ideal conditions of 85% (Smith & Hawkins, 2004). Though there are many different types of benchmarking as explained by Andersen, 1999, the performance benchmarking was the only type that could be done with the empirical data set.

As the empirical data doesn't contain the details of bottleneck machines or the production flow, the benchmarking calculation is done by aggregating the measures of the individual machines in a company. This benchmarking sets the foundation for the performance improvement when compared to the other companies of the same industry type and hence enhances the competiveness. Also, these benchmarking questions the survival of the business if the company's performance is low when compared to the other companies in the same industry type and also exposes the gap between the top performer and the current state.

While these benchmarking is helpful on a high level, on the other side they might also obscure deeper insights that would rise if a company compares the performance of individual steps in the production process. For example, one company in Food and Beverage industry type may have the packaging process as more efficient and also have other inefficient machines which will offset the aggregated OEE values. Also, as OEE benchmarks are only a mere comparison of the performance, they don't reveal anything about the circumstances based on which the companies performed. Though there are down sides of this way of benchmarking, the performance benchmarking of OEE is still useful to visualise the performance gap among the different companies of each industry group and to see the overall improvement potentials (Boxwell & Robert, 1994).

However, data collection for more number of companies under each Industry group is required to improve the quality of the results and to comment on the overall performance of each industry type. Any new company's performance under a particular industry group could be compared with the results got in thesis against the median as well as top quartile performance of the OEE.

One more key understanding from the data analysis is that, OEE is quite difficult to benchmark as different companies' measure the losses differently and also more the true insights could be drawn only if the bottleneck machines are analysed and compared among the industries. As pointed out by Hansen (2001), the OEE benchmarking is useful when the OEE of the bottleneck machines are considered.

I.4.4. Summarising the results discussion

In order to calculate the availability factor of OEE, the losses need to be captured in a standardised manner in the MES. And if the losses are captured in the standardised manner across industries, then the quality of the benchmarking will improve. Also, a standardised form should also be used when calculating the availability especially when it comes to including the company policy influenced loss times. The insights from the empirical data raises question to what extent the MES captures the performance and quality efficiencies as it is observed that majority of the companies monitor these two factors at 100%. The companies need to start monitoring the actual cycle time and the deviation with respect to the standard cycle time in MES in order to calculate the performance factor of OEE. Adding on, the quality efficiency should also be precisely monitored by the MES. Also, Nakajima (1988) and De Groote (1995) standalone methods of calculating the OEE cannot be directly applied to the empirical data set. This is due to fact that factors captured in the data set don't correspond to the factors described in these methods.

The results in the analysis have shown that majority of the loss times are influenced by the operator. Standardisation of the operator tasks and implementation of lean tools to reduce the operator tasks outside the normal work flow is highly effective in order to decrease the machine interference time and to improve the OEE. As a supporting fact, this study has shown improvement potential across industry types in reducing the operator influenced loss times. Also, it could be inferred that the OEE measure is a highly aggregated measure and doesn't explicitly gives information on operator influenced loss times.

Finally, the benchmarking of the aggregated OEE analysis reveals the overall performance gaps across companies. Though this benchmarking cannot reveal more insights within a company, it is a good indicator to compare the performance across the companies. Also, from the overall big data analysis, it is found that the OEE is much difficult to benchmark when the bottleneck machine or the production flow is not known and also due to the fact that the companies monitor losses in different ways.

I.5. Conclusion

In this chapter, the conclusions drawn from the study are presented. This is done by answering the stipulated research questions.

RQ: How can the big data be used to determine the OEE and the loss levers of OEE?

The losses which are described in the empirical data set are grouped and a framework was made to calculate the OEE. Then, the OEE calculations are made with respect to the loss framework.

RQ 1 : What is the difference in OEE among the different industry groups?

It has been found that the among the four industry groups: Food and Beverage, Mechanical workshop, other automated discrete production and Polymeric, the Food and Beverage Industry has the highest median OEE of 74%. The second performer being the mechanical workshop followed by the polymeric and lastly is the other automated discrete production.

There are however, clearly contextual factors that influence the OEE calculations, especially when it comes to the accounting the planned losses in availability (in this study the planned production losses are excluded from the calculations). Adding on, the performance and quality efficiencies are calculated from the loss descriptions due to the lack of availability of the necessary data. Also, this OEE value is calculated by aggregating the values of the individual machines and assuming that all the companies under one industry group have same type of production system.

RQ 2 : What is the average overall OEE of the industries from the given MES big data?

The average OEE of the 23 Swedish companies is 65% and the median OEE is 70% indicating that, there is a significant potential to improve the OEE to the median level.

Also, the major loss to the OEE is the unclassified loss which is 19%. This indicates that the loss definitions needs to be standardised and should be classified in a standardised manner.

RQ 3 : How large is the operator disturbance portion of OEE?

This study shows that, around 90% of the loss time of OEE could be categorised as operator influenced. Also, it is proved that, the impact of the operators on the loss times is not explicitly captured by the equipment performance metrics like OEE.

This shows an enormous amount of potential regarding the standardisation of the operator tasks and elimination of the manual work by the operators outside the normal workflow.

In addition to the findings, benchmarking was made by comparing the OEE factors: Availability, performance and quality which are calculated for the 23 companies with the world class OEE values. This shows that there is significant potential improvement in the performance factor of the OEE. The starting point for the companies to improve on performance efficiency is to compute the actual cycle time of the process and the standard cycle time of the process.

As a final note, the hope is that the outcomes of this study can stress the importance of the capturing the losses in the standardised way and to uncover the fact that, apart from improving the equipment downtime, improving the operator influenced loss times could also improve the OEE.

I.7. Suggestions for Future Research

As there is huge variation in the way the companies are monitoring the losses, a standardised methodology is to be created in order to classify the losses. Further research could be conducted on the standardisation of the losses and the procedure to implement in the industries. This will not make the data analyst easy to interpret but also would help to benchmark the performance with their competitors to the level of the exact loss.

Despite the fact that the operator influenced loss times effects the OEE, a mathematical expression for the OEE could be derived which includes this factor along with the availability, performance and quality. This will be helpful for the companies to see quantitatively how much was the exact lost time under the planned production time due to operator.

.

Data set II

II.1. Frames of Reference

This chapter presents the theoretical study of different theories used in order to identify the bottlenecks in the production system. This chapter starts with the general description of the production system with more focus on the bottlenecks. Thereafter, more detailed explanation of bottleneck detection approaches are described in detail. Furthermore, predictive modelling importance and approaches are explained in detail.

In order to gain the understanding of the different bottleneck techniques and the downtime parameters and to gain knowledge on the statistical analytics a literature review was conducted. The literature were collected from the following scientific databases which are accessed through Chalmers Library

- Science Direct (sciencedirect.com)
- Google scholar (scholar.google.com)
- Scopus(scopus.com)
- ProQuest (proquest.com)
- Books 24/7 (books24x7.com)

To search for the relevant literature regarding the OEE and operator influence on OEE, the following key words are used,

- Bottlenecks in production
- Bottleneck methods
- Disturbances
- Real time data analytics
- Maintenance parameters
- Productivity
- Capacity
- Predictive analytics
- Reliability analytics
- Monte carlo methods
- Algorithms
- Flowcharts

Additional literature were given from the supervisors and based on the most cited references in literature. This literature is used to gain deeper insights on bottleneck detection techniques and downtime parameters which are used to discuss and validate the results.

II.1.1. Bottlenecks in Production System

The raw material is converted into a finished product by undergoing value added activities in production line (Boysen, Fliedner, & Scholl, 2007). There are many challenges faced by the production industry in today’s competitive environment. One of the challenges is to manage the product flow in the production line which is often disrupted .These disruptions in any machine will result in the blocking or starving of the upstream or downstream process in the production system. The disruptions in the production system are due to breakdowns, setup, operator failure, lack of material supply etc. If these disruptions takes place frequently in a machine of the production system, then that machine is the bottleneck for the entire production line as it disturbs other machines and hence the production system performance. To improve the performance of the production system, the key is to mitigate the bottlenecks. The system production rate is only 60-70% of the system capacity and this is due to the bottlenecks (S.-Y. Chiang, Kuo, & Meerkov, 1999). In order to improve the system performance rate, the throughput rate in the bottlenecks of the production system should be improved(Goldrat & Cox, 2004). Therefore, the bottleneck identification is the most important critical first step in order to mitigate and manage the bottlenecks.

There are a variety of definitions found in the literature regarding the bottlenecks. The bottlenecks could be classified into **simple bottlenecks, multiple bottlenecks** and **shifting bottlenecks**. In a fixed interval of time, if there is only one machine which is the bottleneck in the entire production system, then it is called simple bottleneck (Grosfeld-Nir, 1995). If two or more machines are the bottlenecks in a fixed interval of time, then it is called multiple bottlenecks (Aneja & Punnen, 1999). The bottlenecks in the production line will be shifting at different run times from one machine to another machine. This is termed as shifting bottlenecks (C. Roser, Nakano, & Tanaka, 2003). The bottlenecks could also be classified based on the durations of the bottleneck machines. Those are **long term bottlenecks** and **short term bottlenecks**. The machines which affects the performance of the system for a short time interval is termed as short term bottlenecks while on the other hand, the machine which affects the performance of the system for a longer interval of time is called as long term bottlenecks. Also, Roser, Nakano and Tanka (2001) distinguish the bottlenecks between **primary, secondary** and **non-bottlenecks**. Primary bottlenecks are the machines which have the largest effect on the system and the secondary bottlenecks are those which limit the performance but to a small extent, while the non-bottlenecks do not have any influence on the production system performance. Also, defined are the **static** and **dynamic** bottlenecks(Chwif, 2008). Static bottlenecks influence the system all the time whereas the dynamic bottlenecks influence the production system over a specific time frame. Yet another type of bottleneck is **momentary** and **average bottleneck**(Christoph Roser, 2002)**.** Bottleneck at a specific time are called momentary bottlenecks whereas the average bottleneck is related to all momentary bottlenecks and considers the most significant machine as the primary factor to improve the system performance.

There are many method proposed in the literature to identify the bottlenecks. Those methods are summarised below:

- Queue length (Lawrence and Buss, 1994): The machine with the largest queue length before the machine is the bottleneck machine. The queue length is expressed as numbers.
- Lowest production rate (Kuo, Lim and Meerkov, 1996): The machine whose production rate is low when compared to other machines in the production system is the bottleneck machine as it controls the line throughput. The production rate is expressed as numbers per unit time.
- Sum of blockage and starvation times (S. Y. Chiang, Kuo, & Meerkov, 1998): The machine which has the lowest sum of blockage and starvation time in a production line is the bottleneck machine. The blockage and starvation time are measured in time units.
- Utilization (Law and Kelton, 2000): The machine which has the highest utilization in a production line is the bottleneck machine. The utilization is measured in percentages.

- Waiting time before the machine (C. Roser et al., 2003): The machine which has the longest waiting time of the product in queue before the machine is the bottleneck machine. It is measured in time units.
- Active period method (C. Roser et al., 2003): The machine which has the longest active period in a production line is the bottleneck machine. The active period is expressed here as percentage of time or in time units. This method is also suitable for dynamic bottleneck detections.
- Shifting bottlenecks (C. Roser et al., 2003): The machine which has the highest sum of duration of sole active state (without any interruptions during that time interval) is the bottleneck. It is expressed in percentage of time. This method is also suitable for dynamic bottleneck detection methods.

All the previous methods were either built on mathematical models or by using simulation as support tool as shown in Figure 15. None of the previous research had been done in using these techniques over the real time data measured on the factory floor. The data driven bottleneck detection of manufacturing systems was introduced by Li et al. (2009) using the turning point method i.e. the machine with the smallest sum of blockage and starvation times is the bottleneck machine. However, the author doesn't describe the background of the data collection process. As a result, the author assumes that all the companies collect the blockage and starvation times explicitly for every machine. Though the generalised algorithm is created for this turning point method, the same algorithm cannot be used in all scenarios. For example, when the production system has parallel machines, then this method application will be questionable.

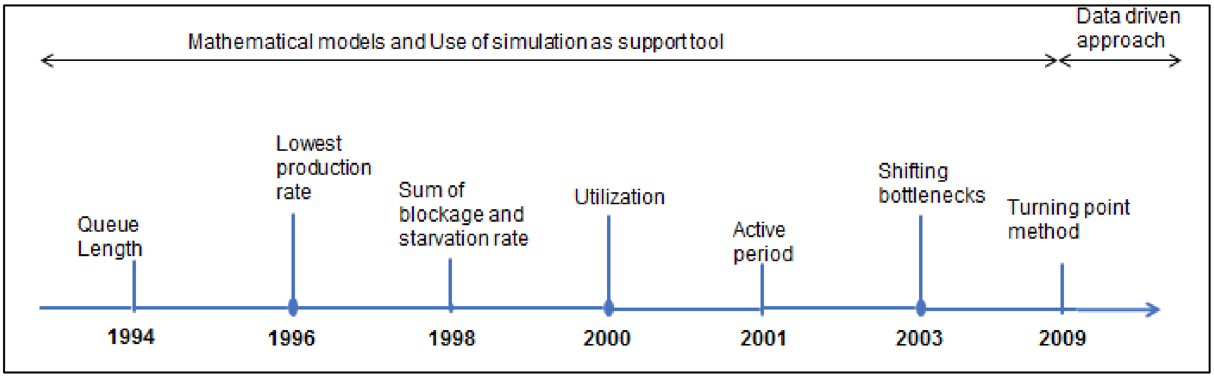


Figure 15: Timeline of bottleneck approaches

Though there are several ways to detect the bottlenecks, the results got from these applications of methods differ with respect to the types of the production system. The bottlenecks are not that easily identified by using conventional approaches (Chwif, 2008). Since the production systems are highly dynamic in nature, the constraints could also be highly time dependent (C. Roser et al., 2003).With the various detection techniques available, the situations where the various techniques could be used is shown in Table 5.

Table 4: Situation dependent bottlenecks detection recommendation (Chwif, 2008)

Method Situation	Queue length	Utilization	Waiting time	Active time period	Shifting bottleneck method
Low mix; Lower number of station; Low fluctuation	Recommended but queue size should be infinite	Recommended	Recommended but queue size should be infinite	Recommended	Recommended if other methods are not applicable
High mix; Lower number of station; Low fluctuation	Recommended but queue size should be infinite	Recommended	Recommended but queue size should be infinite	Recommended	Recommended if other methods are not applicable
Low mix; Lower number of station; High fluctuation	Low recommended (especially if queues are not infinite)	Low recommended	Low recommended (especially if queues are not infinite)	Low recommended	Recommended
High mix; High number of station; High fluctuation	Non recommended	Low recommended	Non recommended	Low recommended	Recommended

Table 5 shows that shifting bottlenecks could be applied in almost all scenarios. On the other hand, the queue length method and the waiting time method becomes inefficient when there is no buffer in between the machines. Furthermore, the utilisation and active period method could also be used in all scenarios even though it is not highly recommended for high mix and high demand fluctuation scenarios. Moreover, the active period method encompasses the utilisation method as in utilisation method only the working state is considered for the analysis, whereas in the active period method, in addition to producing state, the other conditions of the machines are considered like breakdowns etc as shown in Table 6.

The two mostly recommended methods: active time period, shifting bottlenecks are described in detail.

II.1.1.1. Active time period Methods

Each machine in the production line has only two states, active and not active. Table 6 shows the active state categories and the inactive states categories.

Table 5 : Active – Inactive states of different machines (Christoph Roser, 2001)

Machine	Active	Inactive
Processing machine	Working, in repair, changing tools, serviced	Waiting for part, waiting for service, blocked
AGV	Moving to a pick up location, moving to a drop off location, recharging, being repaired	Waiting, moving to a waiting area
Human worker	Working, recovering	Waiting
Supply	Obtaining new part	Blocked
Output	Removing a part from the system	Waiting

Active Period Percentage Method

The sum of the active durations of the machines over the time interval yields percentage active time of the machine. The machine with the highest active percentage duration is the bottleneck machine. This method can only identify the long term bottlenecks and primary bottlenecks but it cannot identify the secondary and the non-bottlenecks. Also, this method cannot detect short term bottlenecks. An example active period utilisation graph for eight machines arranged in a sequential manner with a buffer size three in between the machines is shown in Figure 16.

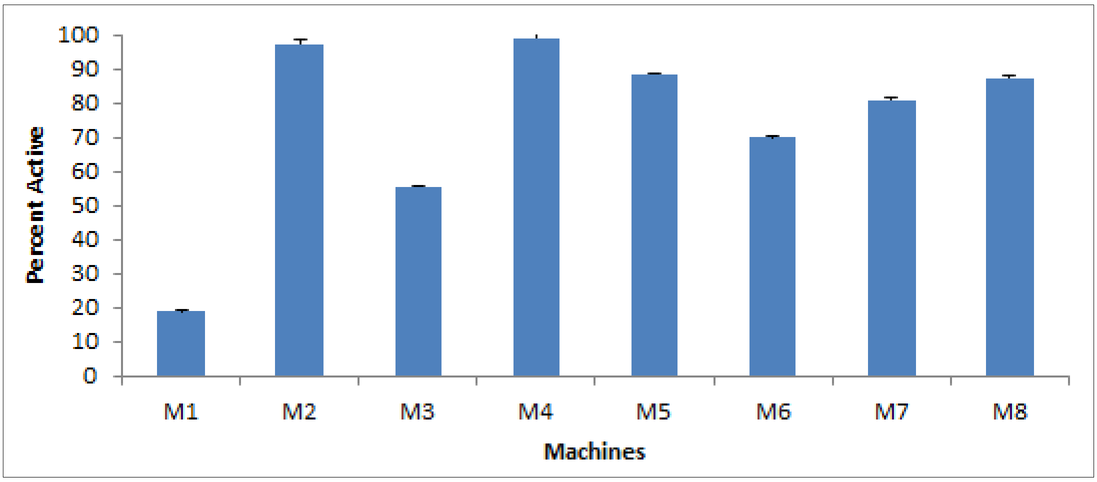


Figure 16: Active period percentages of machines (Adapted from Christoph Roser, 2001)

From Figure 16, it can be inferred that M4 machine has a workload of 99% which is the highest among all the machines. M2 machine has the second highest workload percentage of 97%. On the other hand, their confidence limits overlap with each other. Therefore, the bottleneck identification between these two machines cannot be determined with certainty(Christoph Roser, 2001).

Mean Active Period Method

An improved version of this method is the mean active period method (Christoph Roser, 2001). The active durations of the machines are collected in a specific interval of time and then the average of those active durations is computed. Since the active durations of the machines could be widely distributed, the standard deviation of those values is calculated. In addition to this, a confidence interval of 95% is calculated. Due to this, the results of the bottleneck detection have high confidence level and the bottleneck is detected with higher accuracy compared to the percentage active duration method. Moreover, this method could be applied to the historical data (Christoph Roser, 2001). Figure 17 show the graphs of mean active period and confidence intervals of the same production setup as described in active period percentage method section.

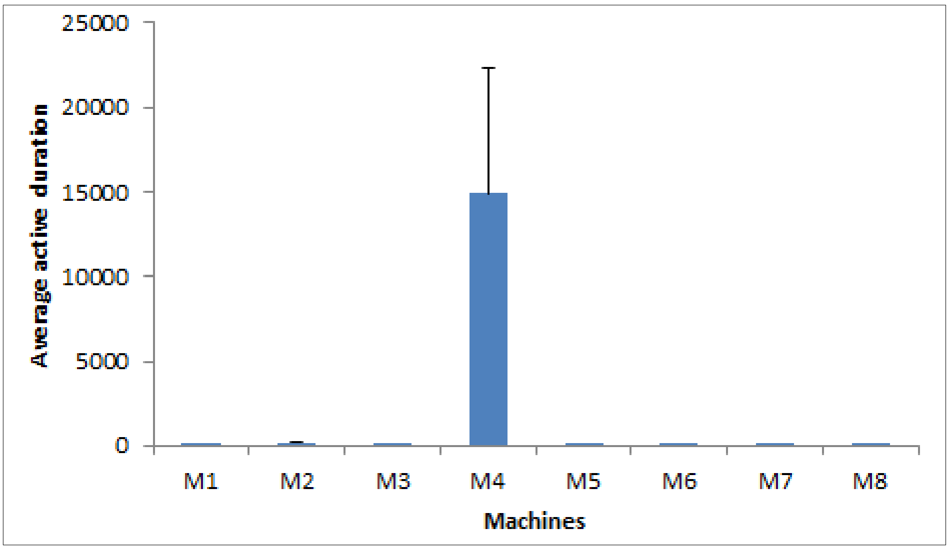


Figure 17: Mean duration of bottlenecks (Adapted from Christoph Roser, 2001)

From Figure 17 it can be inferred that M4 has the longest mean active period of 14885.2 seconds among all the machines. This indicates that M4 is the most probable bottleneck machines when compared to other machines.

Sole and Shifting bottlenecks

Sole Bottlenecks

The machine with the longest un interrupted active period is termed as the sole bottleneck and the system is constrained by this machine (C. Roser et al., 2003). Sole bottleneck detection are useful as additional resources could be allocated to the sole bottleneck in order to improve the overall performance of the system (Christoph Roser, 2000). For example, the active period of two machines over time is shown in Figure 18.

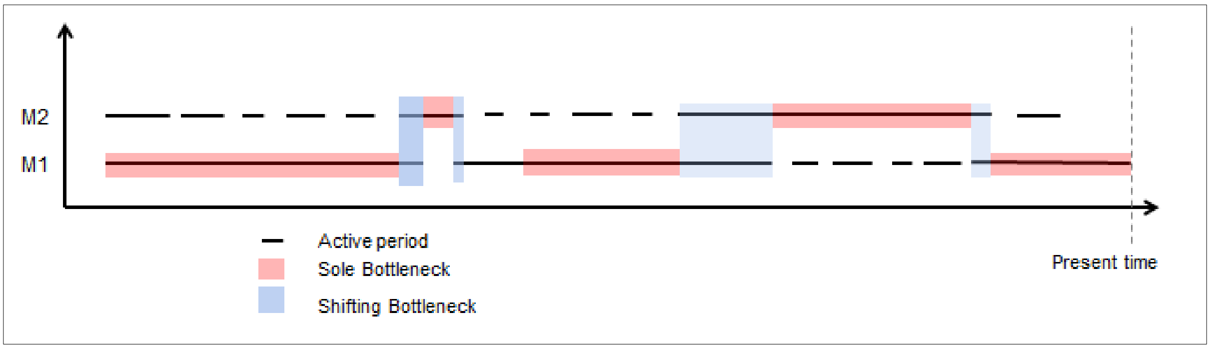


Figure 18: Sole and Shifting Bottlenecks Illustration (Adapted fromChristoph Roser, 2000)

The present time is the time at which the bottlenecks are determined. It could be from Figure 18 that at the present time the machine M2 is the current bottleneck as it has the longest uninterrupted active period when compared to machine 1.

Shifting Bottlenecks

On the other hand, when the active period of one bottleneck machine overlaps with the active period of the next bottleneck machine is termed as shifting bottlenecks (C. Roser et al., 2003). The difference between the sole and shifting bottleneck is that, the sole bottlenecks are those which do not overlap with the previous bottlenecks (Christoph Roser, 2000). From Figure 19, it can be inferred that M1 was the shifting bottleneck during the third active period of M2 and once the M1 becomes inactive, then M2 becomes the sole bottleneck.

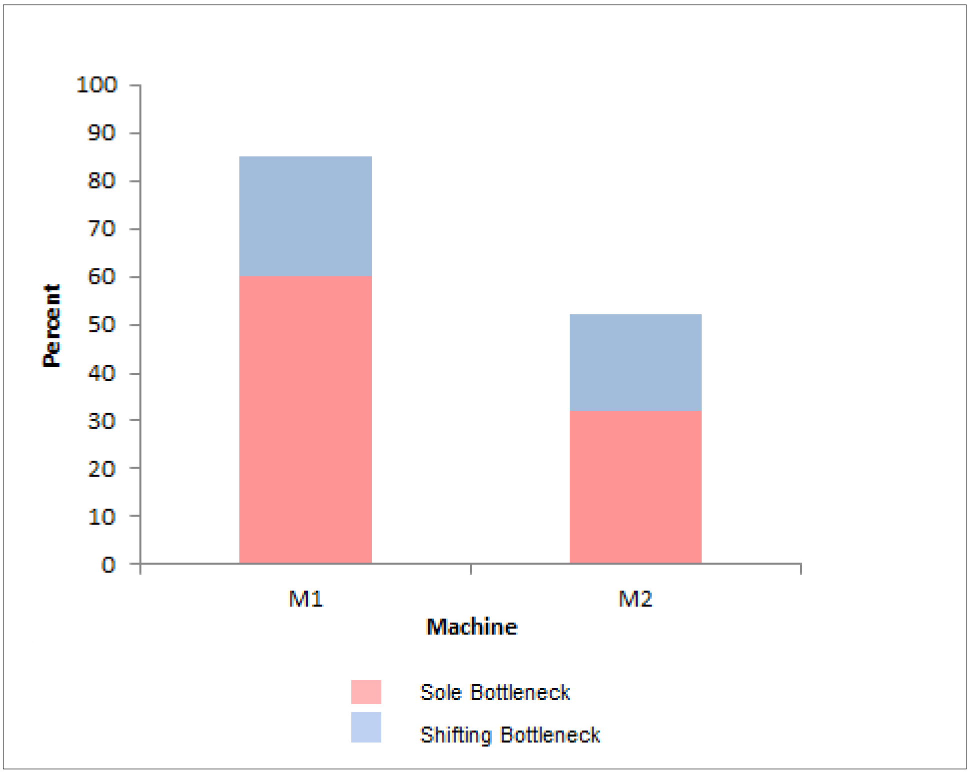


Figure 19: Sole and Shifting percentages of machine (Adapted from Christoph Roser, 2000)

The machine which has the largest sum of the sole and shifting bottleneck percentages is the primary bottleneck. Generalising it, the larger the percentages, the larger is the effect of respective machine slowing down or stopping the system(Christoph Roser, 2000).

Roser et al. (2003) compared the methods: active period percentage method and shifting bottleneck approach as described in Table 7. From Table 7, it could be understood that the shifting bottleneck is more superior to the active period percentage method.

Table 6: Bottleneck detection methods comparison (Roser et al. (2003))

Method	Active period percentage	Shifting bottlenecks
Accuracy	Medium	Excellent
Understand ability	Medium	Excellent
Required Data size	Large	Small
Long term Bottleneck	Yes	Yes
Medium term Bottleneck	No	Yes
Short term Bottleneck	No	Yes
Primary Bottleneck	Yes	Yes
Secondary Bottleneck	No	Yes
Non – Bottleneck	No	Yes
Implementation	Very Easy	Medium
System Limitations	Moderate	Few

Bottleneck Detection using Cycle time

Weindahl and Hagenscheidt (2002) the cycle time of the every individual station do have an effect on the entire system utilisation. In other words, the cycle time is also a useful parameter in order to detect the bottlenecks in the production line. The machine with the highest cycle time is the bottleneck.

II.1.2. Failure data modelling

Maintenance is a support group function which is important to support the production related processes. The breakdowns in the machines of the production line affect the throughput. The usage of maintenance differs among the companies. Maintenance are of three categories : preventive, corrective and predictive maintenance (Dhillon & Liu, 2006). The definitions of the three categories according to Dhilon & Liu (2006) are:

- The maintenance activities which are well planned and is carried out on a periodic basis in order to maintain the working condition of the machine.
- Corrective maintenance is done after the breakdown has happened in the machine. It is a reactive maintenance strategy.
- Predictive maintenance is to continuously monitor and diagnose the machine condition during the operation.

Yet another definition of predictive maintenance is the process of collecting the past information on the various breakdowns of the machine and also about the state of the machine (Niebel, 1994). This collected information is used to predict the breakdown pattern and accordingly plan and schedule the maintenance activities.

One of the important strategies to carry out an effective and efficient maintenance is forecasting the breakdown patterns of the machines. The two important aspects of the maintenance capacity are the number of people involved in the maintenance and the skill set of those people to carry out the maintenance activities (M.Ben –Daya, et al.2009). The breakdowns are highly uncertain in nature and hence the corrective maintenance load and higher repair time and due to this uncertainty the forecasting of the breakdown of the machines are important (M.Ben –Daya, et al.2009).. There are two approaches on which the forecasting is based on. One is quantitative data and the other one is qualitative data. Qualitative data is collected in the form of interviews and expert personnel opinion and this is done when the historical data on breakdowns cannot be collected for the machine. On the other hand, the quantitative data approach is used when this historical data on the machines are available. The forecasting model could be built by using this quantitative data. The following steps are suggested when developing a quantitative forecasting model (M.Ben –Daya, et al.2009):

- Define the variables and identify the causality
- Collect and validate the data
- Search for major trends and seasonality
- Propose different forecasting models
- Validate the models and select the best one
- Improve the performance of it

Preventive maintenance could also be called as periodic maintenance. The basic idea of both the terms is that certain maintenance activities are done at periodic intervals before the machine breakdown. In this preventive maintenance strategy the failure characteristics of the machine plays an important role. Mean Time Between failure (MTBF) is one of the common failure characteristic used in forecasting modelling. MTBF is described as the sum of Mean Time to Repair (MTTR) and Mean Time to Failure (MTTF) for repairable systems. MTTR is the average time taken to bring up the machine after the failure and MTTF is the average time the machine is working after the machine has been brought up from the failure till the next failure occurs. Therefore, MTBF is the time elapsed between two consecutive equipment failures i.e. the time between the start of the failure to the start of the next failure. This failure characteristic is calculated and analysed from the failure times recorded and stored in the maintenance database (Ahmad & Kamaruddin, 2012). In certain companies the maintenance database is linked to the Manufacturing Execution Systems (MES) which captures all the production losses. Another definition of MTBF is that MTBF is the reciprocal value of failure rate (Troyer, 2009) and is shown in Equation 3 .

$$MTBF = \frac{1}{Failure} = \frac{total\ running\ time\ during\ the\ period\ of\ investigation}{total\ number\ of\ failures\ with\ in\ the\ population}$$ **Equation 3**

There are two methods to determine the value of MTBF (Rahman & Kadirgama, 2009):

1. **Estimate MTBF:** The value of MTBF could be found from the historical data.
2. **Predict MTBF:** This method is used when the historical data is not available. The value of MTBF is found out based on the reliability design of the system

The statistical distributions like Weibull distribution, normal distribution, lognormal distribution, exponential distributions etc. are used to analyse and predict the failure characteristics of the machines. After the MTBF data is modelled and the machine breakdown trends are found, the maintenance strategy could be formulated.

The accuracy of the model depends upon the sample size and the amount of the failure data considered (Antony, 2008). For instance, when there is no enough data are available there will be a high uncertainty associated with the output from the model.

II.1.3. Monte Carlo Simulation Technique

The analog – simulation tools with statistical capabilities is a very useful tool for engineers(Johnson, 2011). One such method is Monte Carlo method which provides a large amount of useful and crucial information on how the system will operate. This method works by performing multiple simulations(Johnson, 2011). In the problems of combinatorial analysis and the theory of probabilities, an analogous situation exists(Metropolis & Ulam, 1949). In the game of solitaire, the probability of a successful outcome is an intractable task(Metropolis & Ulam, 1949). In this case, the probability of success could be determined by producing a large number of examples and from this, relative portion of success could be determined (Metropolis & Ulam, 1949). In other words, this method could be described as a method to solve statistical problems in combination with virtual representation of the problem using simulation. Monte Carlo method can be used to identify the outcomes of different scenarios by performing multiple simulations. During this process, the variable's parameters are varied with respect to their statistical distributions. The outcomes are a simple representation of the way the system will operate over a number of design builds(Johnson, 2011). From these outcomes of the simulations and interpreting it various decisions could be made.

The outcome of the Monte carlo analysis are probabilistic results which explains how likely each outcome would be and this is determined without any approximation. Yet another advantage is that, the outcomes are based on all the possibilities that have already happened i.e. the monte carlo method have the ability to factor in a range of values for various inputs. This method is widely used in manufacturing and production environments in order to determine the tolerances for operations and also used to make strategic decisions.

II.2. Methodology

In this chapter the methods used in this thesis are discussed. First the overview of the approach in this thesis is presented followed by the reliability and the validity of the findings is discussed.

II.2.1. Overall Methodology

The CRISP-DM (Cross Industry Standard Process for Data Mining) approach was followed in this thesis, which is the industry standard methodology for data mining and predictive analytics. The advantages of this methodology is that it makes the data mining much faster, reliable, manageable and cheaper (Shearer, 2000).

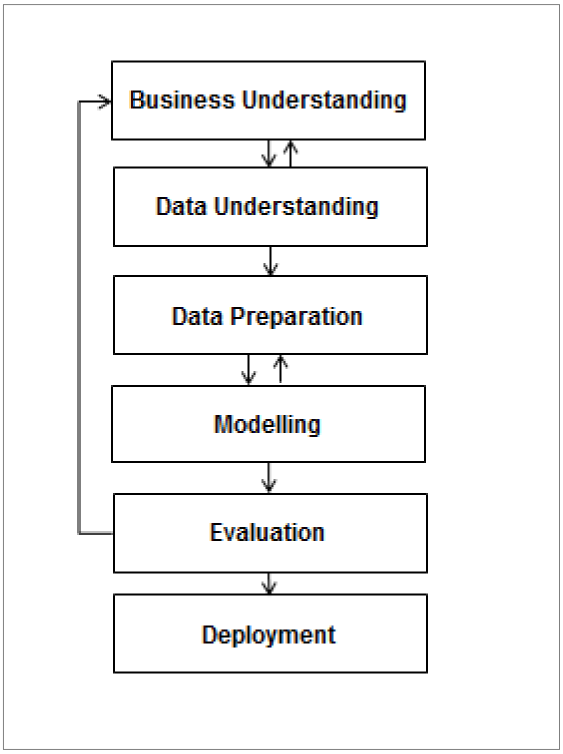


Figure 20: Phases of CRISP – DM model (Adapted from Shearer, 2000)

There are six phases of the CRISP – DM model as shown in the Figure 20. Those are business understanding, data understanding, data preparation, modelling, evaluation, deployment. These phases help to guide a data analytic project. The detailed explanation of how these phases are used in this thesis is given below (Shearer, 2000):

a. Business Understanding
   The main objective of this thesis is to detect the bottlenecks and to predict the breakdown pattern of the machines. The success of this thesis can be measured by validating the bottleneck detection and failure prediction model which increases the machine availability. Also, in order to develop the model, large number of qualified data is collected from a leading automotive manufacturing company's Manufacturing Execution System (MES) database. The data quality was a significant constraint for this thesis.

b. Data Understanding

The event log files of all the machines from in the production line were provided. This event log file consists of the condition data which includes the states of the machines and the event data which records the time events of these states. The data was structured and the structure was uniform across the machines.

Two months data were provided for 18 machines in the AAA line and 14 machines in the BBB line, for the months September 2014 and October 2014, respectively. Two months data were given in a single MS Excel file for each machine. Each file had 17 columns of description. The total number of data rows in the Excel files is 1.38 million. These Excel files have the condition data as well as the event data. In addition to these data files, the layout of the production facility was also given by the company as shown in Appendix B.

c. Data Preparation – Data Anlaysis

Error free and high quality data is required to generate a strong decision support model. Though the data taken from MES is precise which monitors the machines condition and records the events for all time, errors and outliers will still occur. The aim of this step is to get an error free qualified data for the modelling.

Firstly the data for all machines for the two months period is checked to identify the missing days in these two months for which the data is not captured. For the AAA line, 15 days data in the two month period were missing and for BBB line it is 13 days. The missing pattern of data was uniform across the machines in AAA lines and BBB lines. Secondly, the duplication of data rows in the data file of every machine is removed and checked whether all data rows have the events and the conditions recorded which are crucial for the analysis. But there were no missing data. Thereafter the average scheduled production hours were checked for all days in the two month period for all machines and the most frequent scheduled time of the lines across the days was calculated. After the exact time interval for production is calculated, the data rows outside this time interval were excluded and were not taken into the analysis.

After these steps, the cleaned qualified data is made available for modelling and further analysis.

d. Modelling

Various bottleneck detection and maintenance forecasting methods have been used to develop models .The cleaned and qualified data as a result of data preparation is used to develop bottleneck modelling and maintenance forecasting which are further explained in the Results Sections.

e. Evaluation - Validation

The results got from the model are validated using the face validation technique(Sargent, 2010). The expert from the Industry who has the full knowledge about the production system was asked to verify outcomes of the bottleneck detection, maintenance and throughput predicting models. Specifically to the throughput predicting model, internal validity technique was used by running the model several times to determine the variability(Sargent, 2010).

f. Deployment

In this step, the algorithms and flowcharts for bottleneck detection methods and the failure data models are created. The algorithms are designed as per the specifications explained by Bruno & Steiglitz (1972). The algorithms contain the inputs and the outputs and describe the step by step procedure to solve the problem.
The algorithms, after necessary coding, could be integrated with the MES as an add on option. Furthermore, the decision support model which is the integration of insights and knowledge gained from bottleneck prediction and failure prediction models are created.

II.2.2. Bottleneck Detections Model

The active period percentage, mean active period, median cycle time and sole and shifting bottlenecks are carried out using MS Excel 2010. It is to be noted that no additional software was used to model these methods.

II.2.3. Failure Data Modelling

The step to build the failure data model is shown in the Figure 21. The failure time data set is extracted from the real time data using MS Excel 2010. The Statistical Fit of the failure time data set was found out using Minitab 17 statistical software.

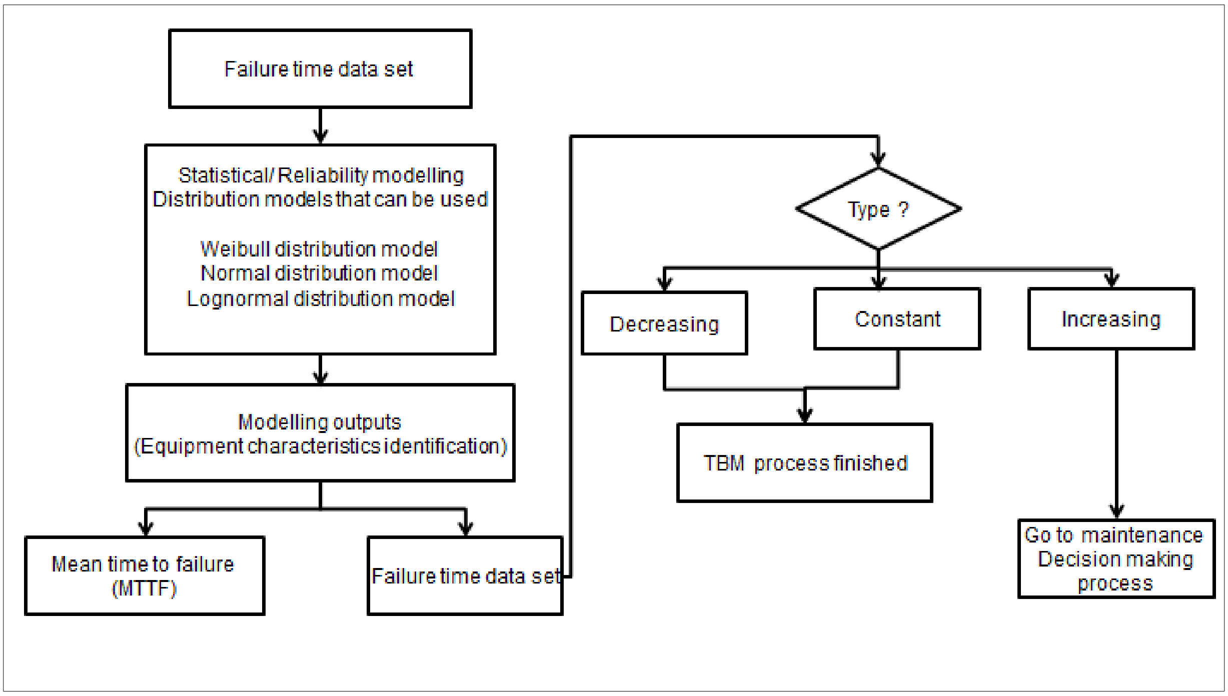


Figure 21: Steps in Failure data modelling(Adapted from Ahmad & Kamaruddin, 2012)

II.2.4. Monte Carlo Method

A step by step procedure of Monte Carlo method is described in the Figure 22. The statistical distribution of the data set is found by using Minitab 17 software. The estimation model was built in MS Excel 2010 and the simulation was also carried out using MS Excel 2010.

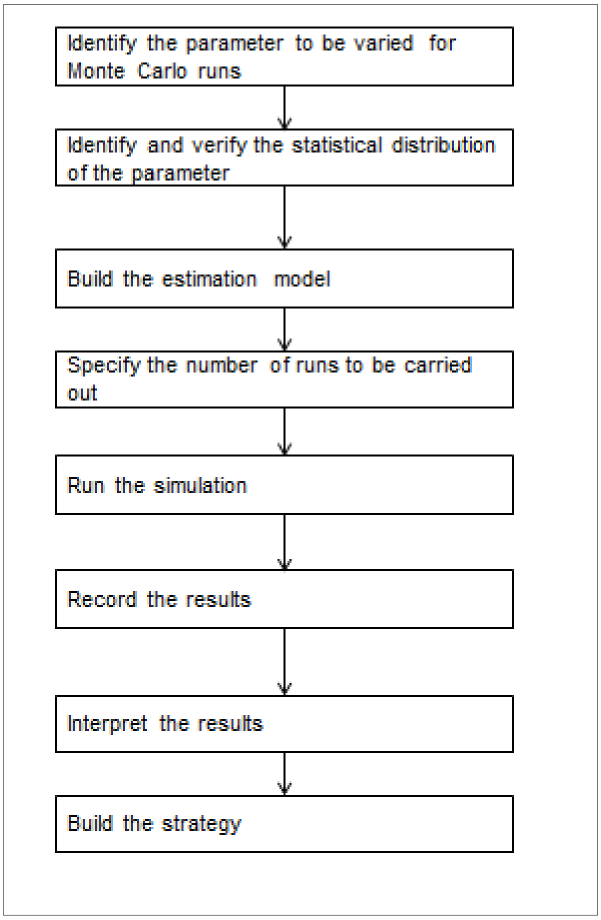


Figure 22: Monte Carlo Simulation Steps(Adapted from Johnson, 2011)

II.3. Experimental Plan

Table 8 shows the bottleneck detection methods, frequency and total downtime of the breakdowns and the predictive analytics which are carried out on the real time data. The states of the machine are divided into two categories: active and not active. The active states of the machine from the event log files of the machine includes *producing, part changing, comlink down , comlink up, error.*The not active states of the machine include *waiting and not active*.

Table 7: Experimental plan of the bottleneck detection methods and predictive analytics

Theme	Method	Description
Bottleneck detection	Active period percentage	The active period percentages are compared between the machines.
	Mean active period	The mean of the all the active periods of the machine are compared with other machines
	Median cycle time method	The median of the different *processing* times of the machine are compared with other machines
	Shifting bottlenecks	The sole and the shifting bottleneck, called collectively as momentary bottlenecks are computed at a particular instant by plotting the active period of different machine
Frequency and total down time	Frequency of breakdowns	The number of breakdown occurrences of a machine during a particular interval of time is calculated and compared with other machines
	Total down time	The total down time due to the breakdown of the machine during a particular interval of time is calculated and compared with different machines
Predictive modelling	MTBF data modelling	The MTBF is calculated from the real time data and the statistical distribution of the data is found out. Thereafter, the trend of the MTBF data is plotted for the shift using probability distribution function
	Breakdown as a percentage of scheduled hours	The breakdowns as a percentage of scheduled hours are plotted on a daily basis in order to identify the trend
	Confidence level of the throughput estimator	The confidence level of the expected throughput against the demand is calculated to better plan the production scheduling

The order of the Results chapter is in the same order as the experimental plan.

II.4. Results

In this chapter the results from the data are presented. The chapter starts with the results from the static bottleneck detections from the real time data, followed by dynamic bottleneck detection. In addition to those, the predictive modelling of the maintenance and the production KPI is also presented.

II.4.1. Bottleneck Detection from Real Time Data

The static and dynamic bottlenecks methods which were presented in literature study were tried out on the real time data. The bottlenecks are determined using the machines states as described in the data file. All the machines in the sample production line have eight states. Those are *producing, part changing, and error, comlink down, comlink up, waiting, and not active, empty run*. The *producing* state of the machine is the state where the machine is engaged in producing the product. The *part changing* state of the machine is the state where there is a setup time due different product in that machine or the changeover between the tools in that particular machine. The *error* state, *comlink down and comlink up* is the assigned when the machine is down due to a maintenance action.. The *waiting* state is when the machine is waiting for the product to be produced or the machine is blocked from producing due to problems with the downstream process. *Not active* state of the machine is when the machine stops apart from the other above mentioned reasons. *Empty run* state is assigned to a machine when the new products or trial products are processed in that machine.

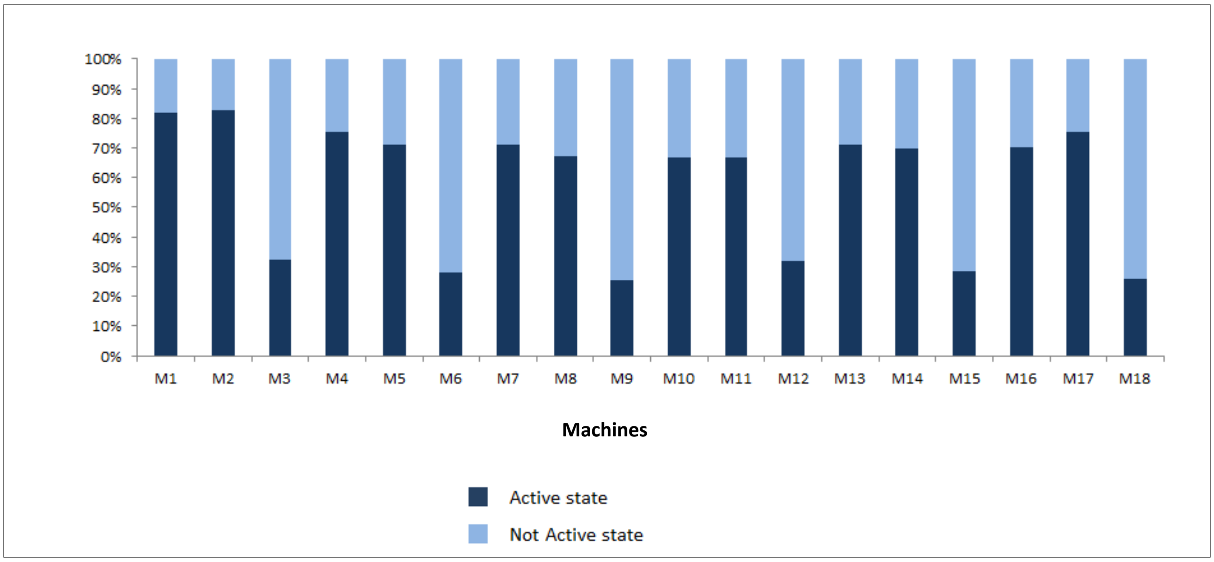
The production line of 32 machines is decoupled into two lines (AAA line and BBB line) as there is a huge difference between the cycle times of the machines between the two lines and as a result there is a huge buffer between the two lines in the stored in a conveyor. The AAA line is scheduled for 44 days in a two month period and the BBB line is scheduled for 47 days in a two month period. Also, the scheduled start time of the two lines on each day is at 06:30:00 and the ends at 23:30:00.This corresponds to total scheduled time of 748 hours for AAA line and 799 hours for BBB line in a two month period.

II.4.1.1. Static Bottlenecks

Three different static bottleneck approaches were tried out over the real time data. Those are Active Period Percentage Method, Mean Active Period Method, and Cycle Time method. Each one of the methods and the results are described further.

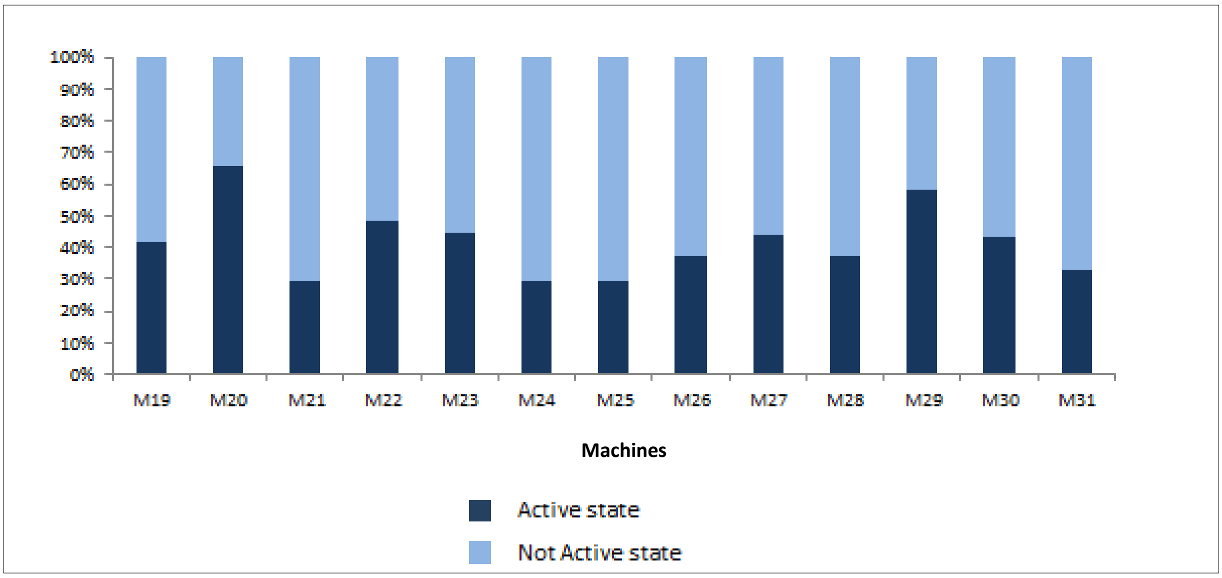
II.4.1.1.1. Active period method Percentage Method

Applying the active period percentage method for the machines in the AAA line and BBB line, the percentage active period of each machine is got as shown in the Figure 23 and Figure 24 respectively.



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	44
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

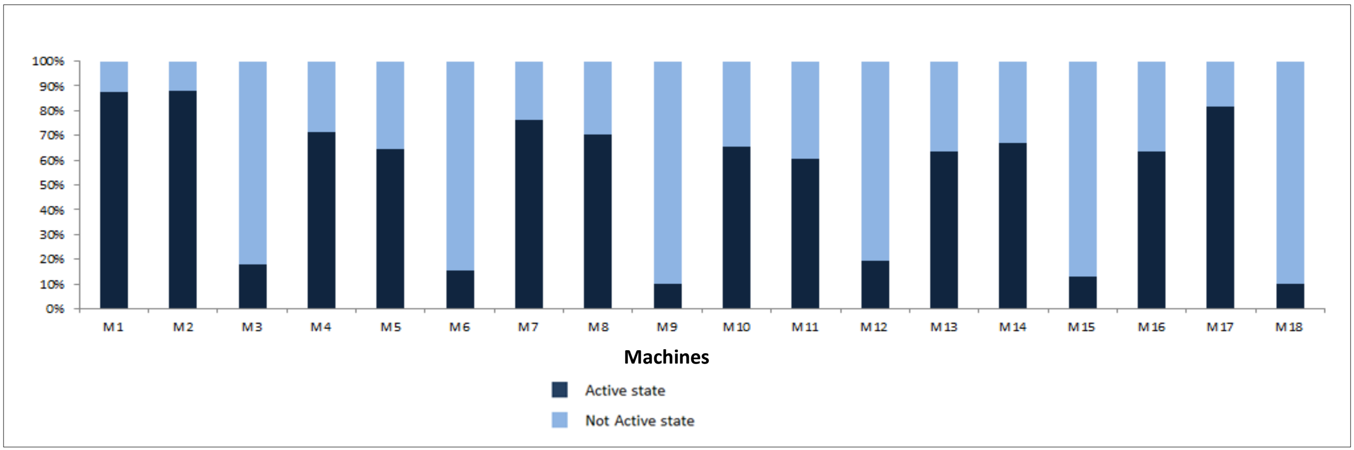
Figure 23: Percentage Active period of AAA line machines



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	47
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	13

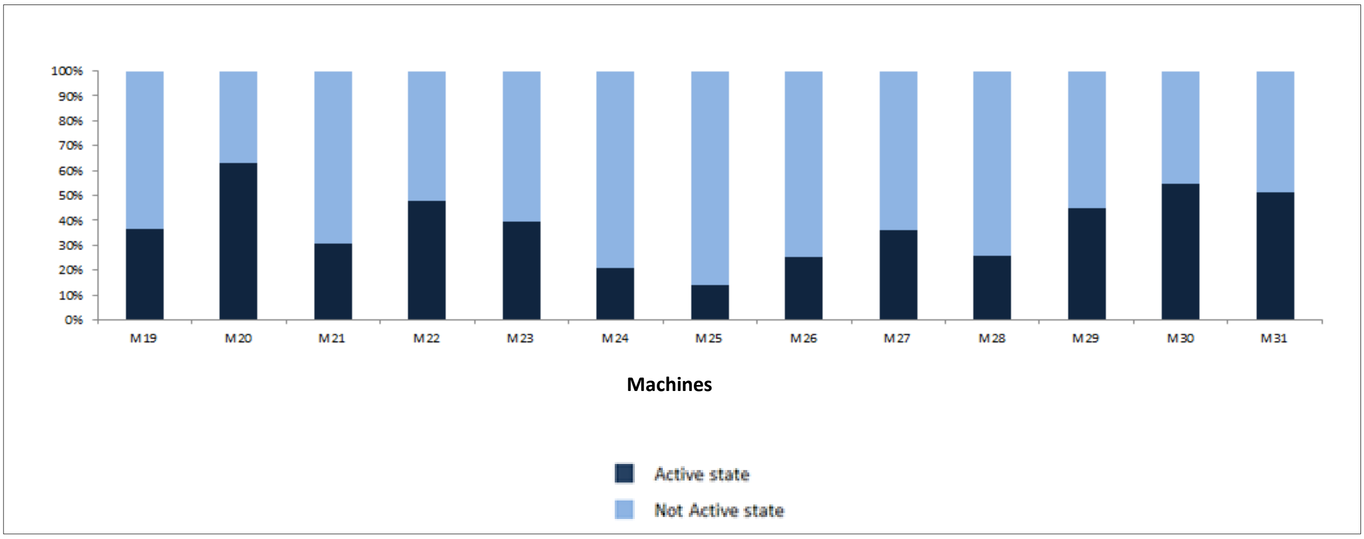
Figure 24: Percentage Active period of BBB Line machines

From Figure 23, it can be inferred that machines M1 and M2 are the bottlenecks as they are active 82% of the time. But on the other hand, it cannot be said with certainty out of M1 and M2 which of the two machines is the bottleneck. Similarly, from Figure 24, it can be inferred that M20 is the bottleneck as it active 66% of the total scheduled production hours. This method is also scaled down to determine bottlenecks on a day level. Figure 25 and Figure 26 shows the bottlenecks of the AAA and BBB line respectively.



Interval Period	01-Sep-2014
Number of scheduled days	1
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 25 : Percentage Active period of AAA line machines for one particular day



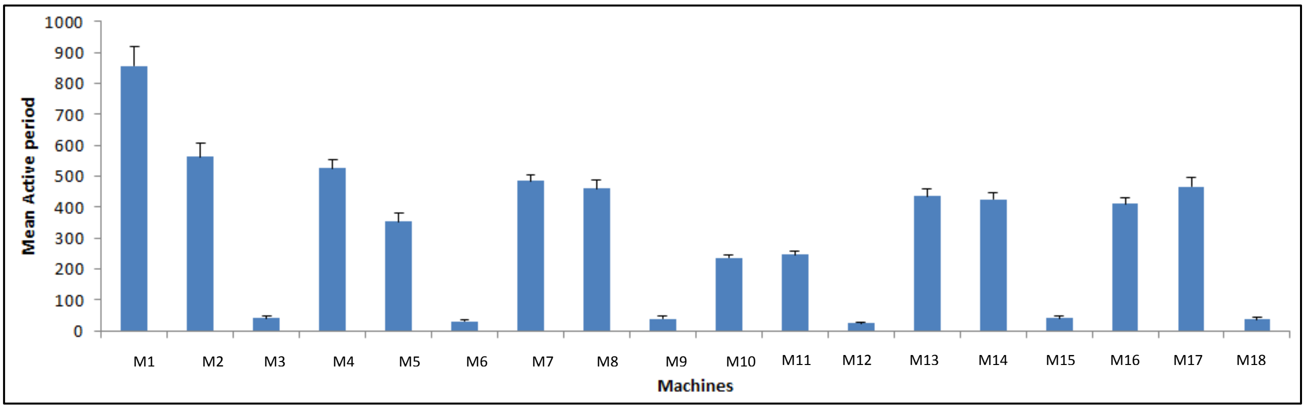
Interval Period	01-Sep-2014
Number of scheduled days	1
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	13

Figure 26 : Percentage Active period of BBB Line machines for one particular day

From Figure 25, it can be inferred that machines M1 and M2 are active 88% of the time and they are the bottlenecks in AAA line. But again, it cannot be said with certainty that out of M1 and M2 which of the two machines is the bottleneck. From Figure 26, machine M20 is active for 63% of the time and hence it is the bottleneck for BBB line.

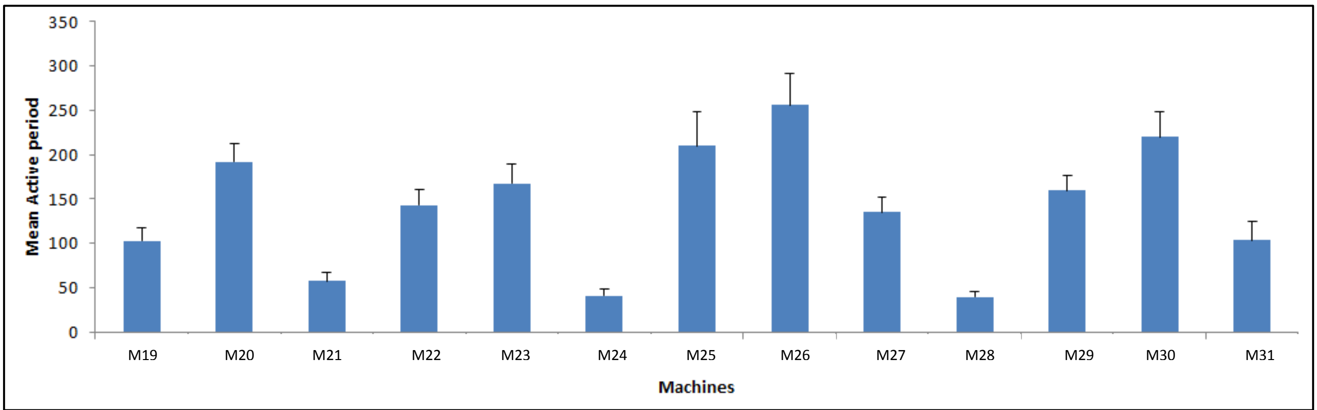
II.4.1.1.2. Mean Active Period Method

The time of the active states are collected and the average time is then calculated. In addition to that, to estimate the accuracy of the bottleneck detection a 95% confidence interval is calculated. Applying the mean active period method for the machines in AAA and BBB line, the mean active period of each machine is got as shown in Figure 27 and Figure 28 respectively.



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	44
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 27: Mean active period of AAA line machines with 95% confidence interval

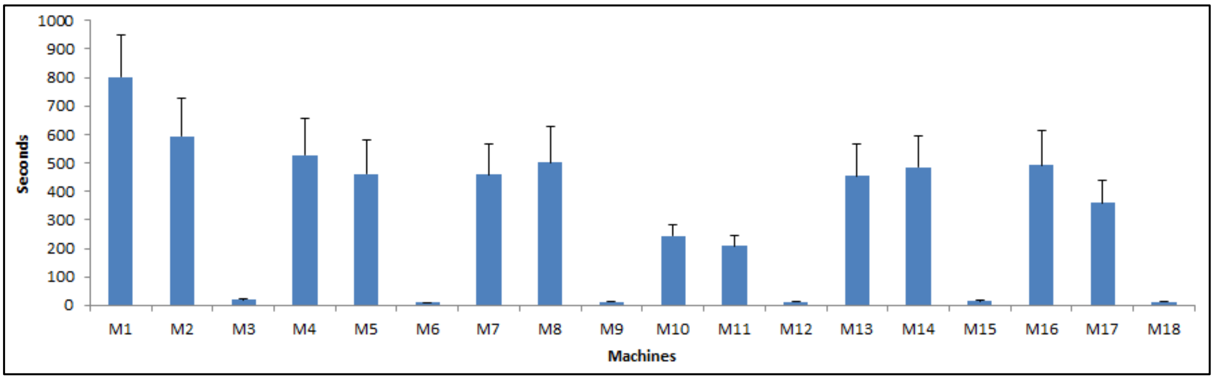


Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	47
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	13

Figure 28 : Mean active period of BBB line machines with 95% confidence interval

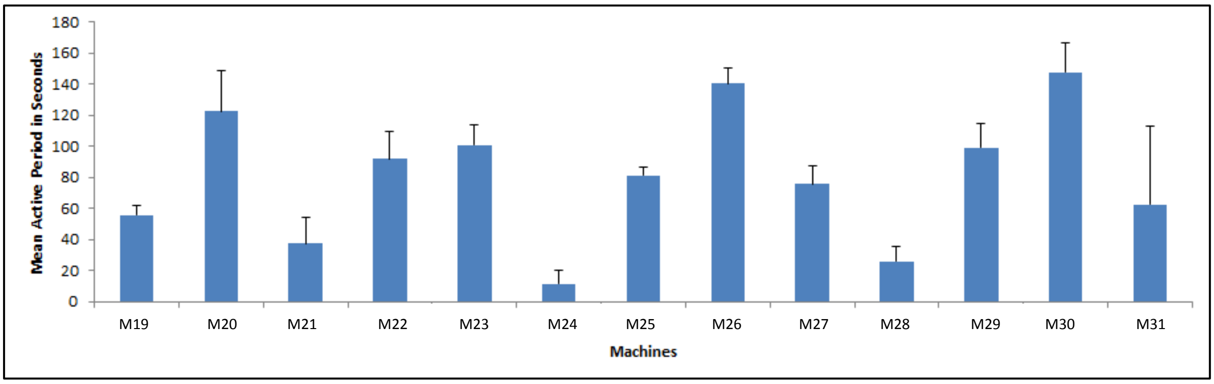
Figure 27 shows that the machine M1 is the bottleneck in AAA line as on average it is the machine with longest active duration compared to all the machines in the line. Moreover, the confidence intervals of this machine doesn't overlap with other machine which indicated that machine M1 is the primary bottleneck and this machine should be improve to improve the AAA line performance. In other words, it could be described as M1 was working for an 856 seconds before being interrupted by the waiting time. Also, the confidence interval of the average active period of machine M1 shows that the average duration fluctuates. Therefore, it can be concluded that machine M1 has the longest active period over 99% certainty and it is the primary bottleneck of the system. Similarly Figure 28 shows that machine M26 is the bottleneck in BBB line. But again it could be inferred that M25 upper confidence interval is the same as machine M26 lower confidence interval. This could be an ambiguous situation in determining the most certain bottleneck. As a next level, machines M20 and 240 have confidence interval overlapping with machine M25. Overall, the machine M26 has the highest sum of the mean active period with the confidence interval, it could be said that machine M26 is a bottleneck with a lower accuracy.

This method is also scaled down to a day basis to determine the bottlenecks on a day level. Figure 29 and Figure 30 shows the bottlenecks of AAA and BBB line individually.



Interval Period	01-Sep-2014
Number of scheduled days	1
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 29 : Mean active period of AAA line machines with 95% confidence interval for one day



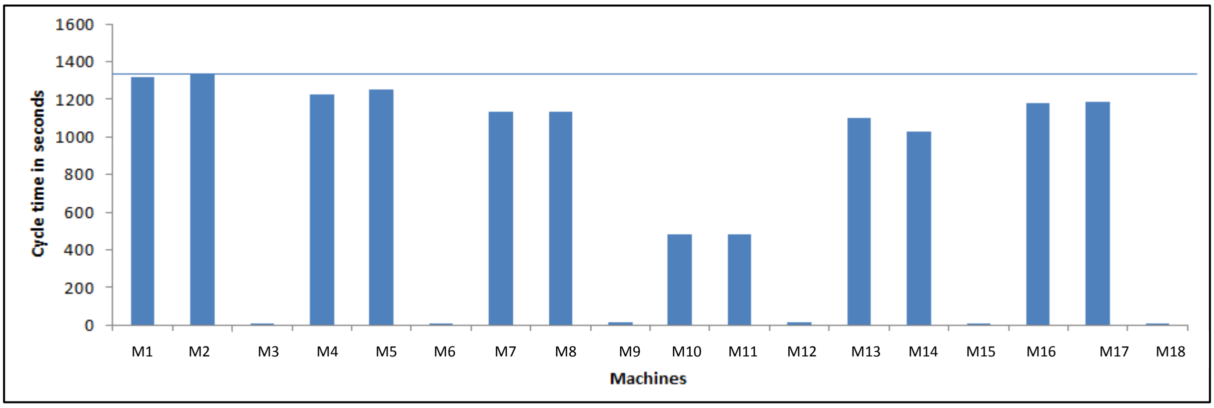
Interval Period	01-Sep-2014
Number of scheduled days	1
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	13

Figure 30: Mean active period of BBB line machines with 95 % confidence interval for one day

From Figure 29, it can be inferred that M1 is the primary bottleneck of AAA line. From Figure 30, machine it could be said that, M30 is the primary bottleneck of the BBB line. Also, it could be said that machine M20 could be the bottleneck as the confidence interval of machine M20 overlaps with the confidence interval of machine M30, but then the degree of overlapping is very small (~ 1%). Also, machine M26 confidence interval also overlaps with machine M30, but again the degree of overlapping is small (~2%). From these values of small values of overlapping, it could be concluded that machine M30 is the primary bottleneck on that particular day.

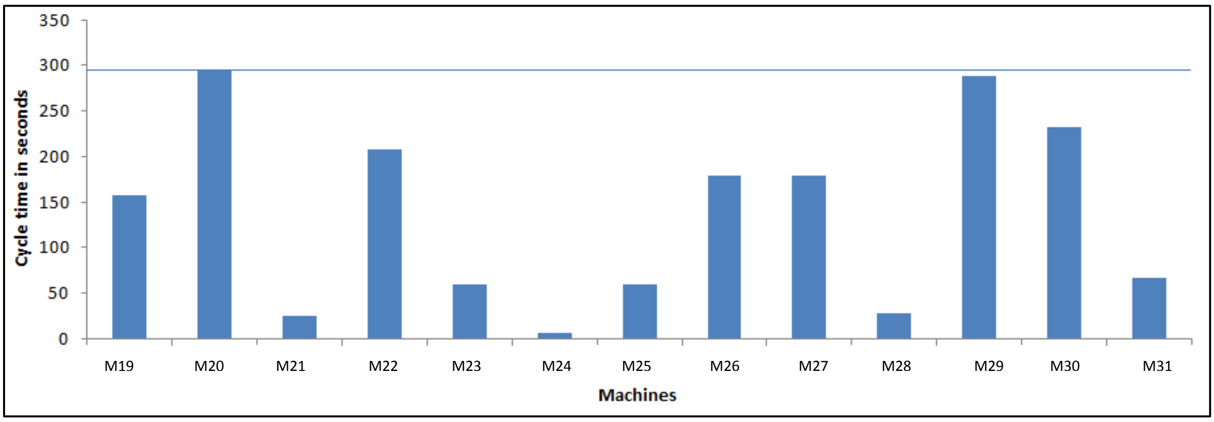
II.4.1.1.3. Cycle Time Method

The cycle time is a good indicator to identify the bottleneck. The cycle time for each machine is found out from the data file. The elapsed time of the *producing* state of each machine is determined for each shift and for each day. To avoid the skewness effect i.e. to avoid very large or very small values can distort the average, the median indicator is more appropriate to compare the cycle times of different machines. From the MES data for AAA and BBB line, the median cycle time for each machine is found out and compared with other machines to find the bottleneck as shown in Figure 31 and Figure 32.



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	44
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 31 : Median Cycle time of AAA line Machine



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	47
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	13

Figure 32 : Median cycle time of BBB line machines

From Figure 31 and Figure 32, it can be inferred that machine M2 in the AAA line and M20 in the BBB line are the bottleneck machines as they have the highest cycle time of 1336 seconds and 296 seconds respectively.

Table 8 and Table 9 summarises the outputs from different bottlenecks detection methods applied over the AAA line and the BBB line.

Table 8: Comparison of bottleneck detection machines of AAA Line

Bottleneck rank	Active period percentage method	Mean active period method	Median Cycle time method
1	M1, M2	M1	M2
2	M4	M2	M1
3	M5, M7, M13	M4	M4

Table 9: Comparison of bottleneck detection results of BBB Line

Bottleneck rank	Active period percentage method	Mean active period method	Median Cycle time method
1	M20	M26	M20
2	M29	M30	M29
3	M22	M25	M26, M27

It is clear from Table 9, that the active period percentage method and the mean active period method points the machine M1 and hence it is concluded that machine M1 is the primary bottleneck. Also, it is clear that the active period percentage method does not detect the bottleneck with certainity (Roser et al., 2001) as there are two machines which equal active period percentages. But with mean active period it could be concluded that the machine M1 is the primary bottleneck as it has the longest mean active duration with 95% confidence interval without overlapping with other intervals. But, according to the cycle time method, M2 is the primary bottleneck and this method does not consider the other active states expect the producing states.

From Table 10, it can be inferred that the active period percentage method and the cycle time method points to the machine M20 as the bottleneck whereas the mean active period method shows M26 is the bottleneck. But according to the mean active period method, machine M26 is not a very certain bottleneck as there is an overlapping with the confidence interval of machine M25. Moreover, this arises as the mean active period is used which might have a skewed effect on the result. So considering these factors, it could be said that machine M20 is the primary bottleneck of the BBB line as the other two methods point to this machine.

All the results drawn from the static bottlenecks analysis are validated with the production system expert in the industry and the expert agreed to the outcomes.

II.4.1.2. Momentary Bottlenecks

The bottlenecks at any given instant of time could also be determined from the real time data. The sole bottlenecks and shifting bottlenecks are the two approaches to determine the momentary bottlenecks. For the momentary bottleneck approach, only the AAA line is taken for calculations due to the time constraint. The same calculations can be repeated for BBB line.

II.4.1.2.1 Sole Bottlenecks

The machine with the longest uninterrupted active period at any given instant is the sole bottleneck. To determine this bottleneck from the real time data, the active time for each machine is plotted on a uniform time scale, starting from start of the shift .This is done to map the state of the machine at every instant of the time and to visualize the states of all machines at the same time. Figure 33 shows an example of the plot showing the active times of all the machines of AAA line from the start of the shift to the end of the shift on a particular day.

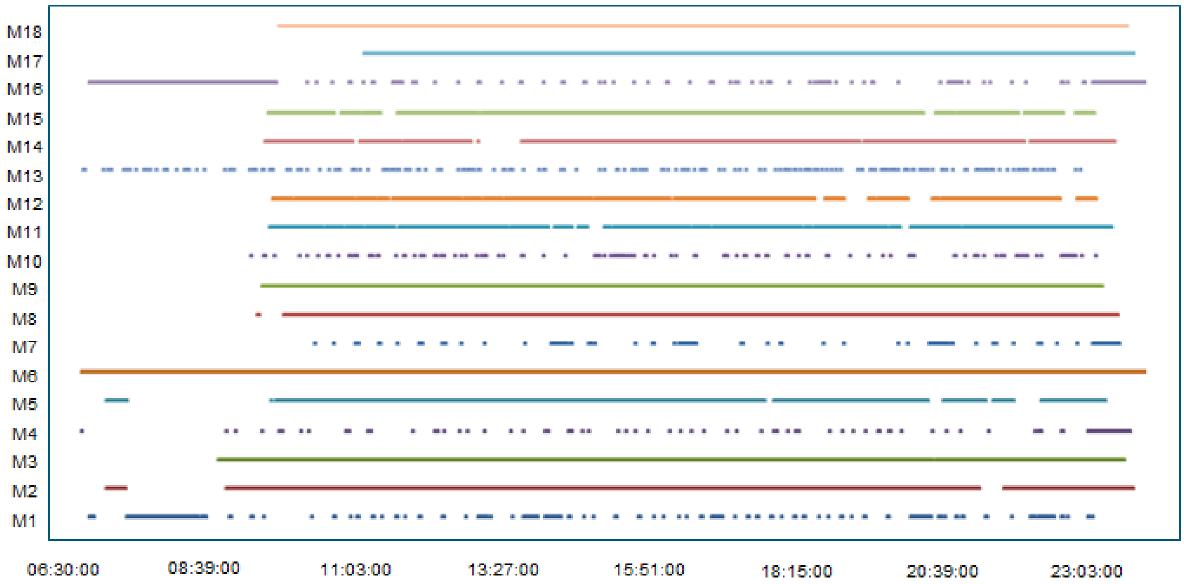


Figure 33 : Plot of the active period of the machines of AAA line

The sole bottleneck is found by determining the longest active period of the machine at the any given instant. Figure 35 shows the active period plot of machines on a particular day from the start of the shift, 06:30:00 until 06:53:48. At 06:53:48, by comparing the length of the uninterrupted active period starting at that instant dating to the past, the sole bottleneck machine is found out. From Figure 34, it can be seen that machine M5 has the longest uninterrupted active period at 06:53:48 and hence the bottleneck at this instant is the machine M5.

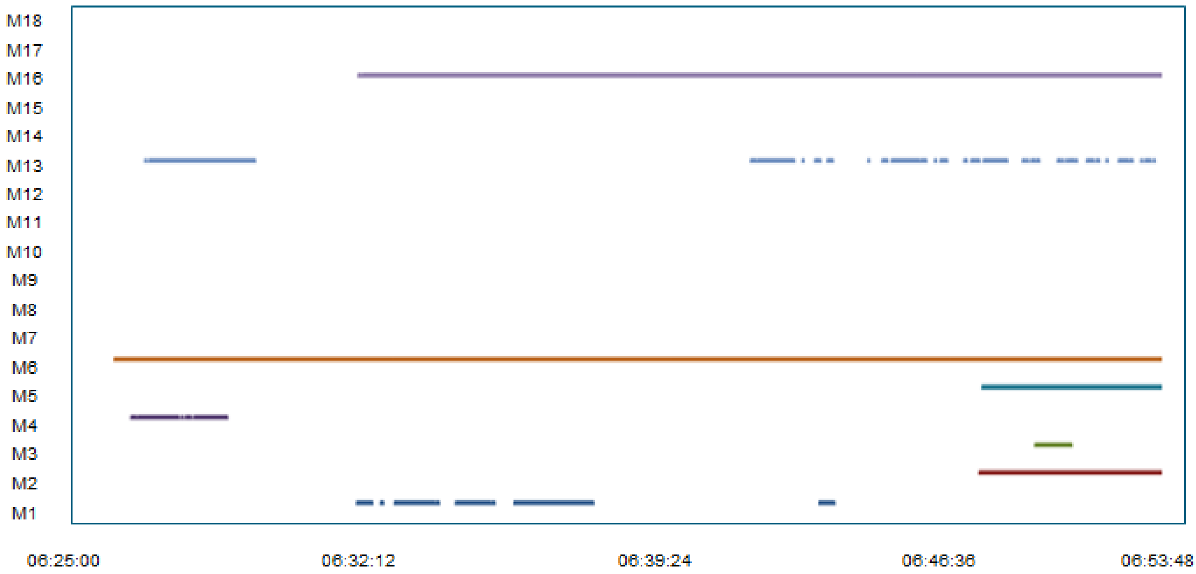


Figure 34 : Plot of active period of machines of AAA line within a time frame

Furthermore, Figure 35 shows the active period plot with an extended time frame of Figure 34 i.e. starting from the start of the shift, 06:30:00 till 09:46:36.

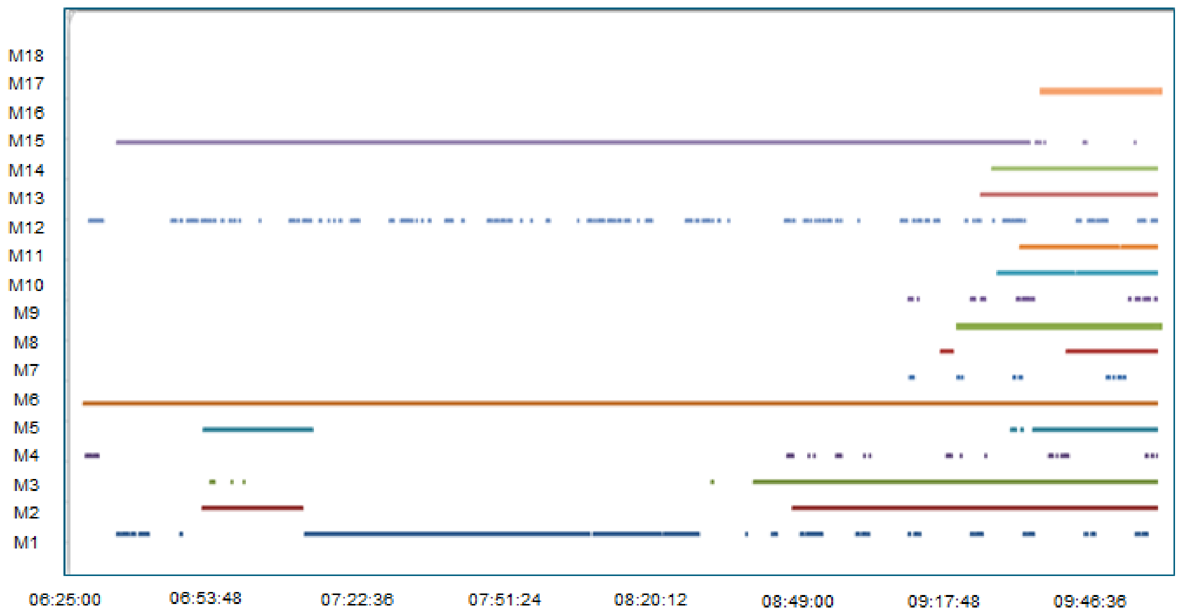


Figure 35 : Extended time frame plot of active period of machines of AAA and BBB line

From Figure 35, it could be inferred that at 09:46:36, machine M5 has the longest uninterrupted active period which makes that machine as the sole bottleneck.

II.4.1.2.1. Shifting Bottlenecks

The bottlenecks in the real time change frequently. To determine the shifting bottlenecks from the real time data, the machine active period is plotted on a uniform time scale starting from the start of the shift. The first step to find out the shifting bottlenecks is to find out the sole bottlenecks. Then the other machine is compared with respect the sole bottlenecks in order to find the shifting bottlenecks. If there are no sole bottlenecks at a given time instant, then the shifting bottlenecks for each machine is calculated with respect to other machines. It could be seen from Figure 35 that the shifting bottlenecks is not observed in the real case as some machine are active for the entire shift. This was checked for three random days and the observation was same.

To demonstrate this method, six machines from the AAA line where the shifting pattern could be observed among the machines are taken for calculation. The same procedure to determine the shifting bottlenecks could be repeated in other days where the shifting bottlenecks exists in order to get the shifting percentage for all machines. Figure 36 shows the visual representation of the shifting bottlenecks over the time period starting from 06:30:00 to 09:24:00.

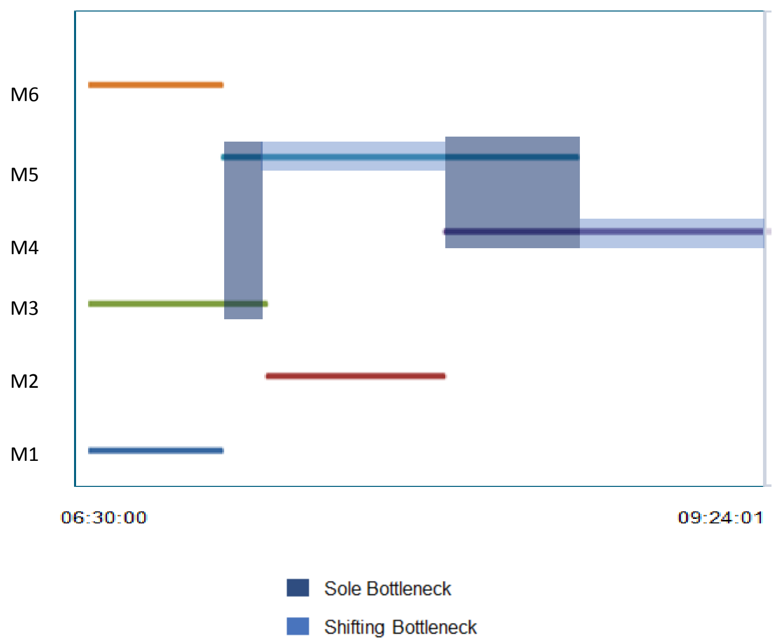


Figure 36: Shifting bottlenecks illustration

From Figure 36, the percentage time the machines were in shifting and sole period is also calculated as shown in Figure 37.

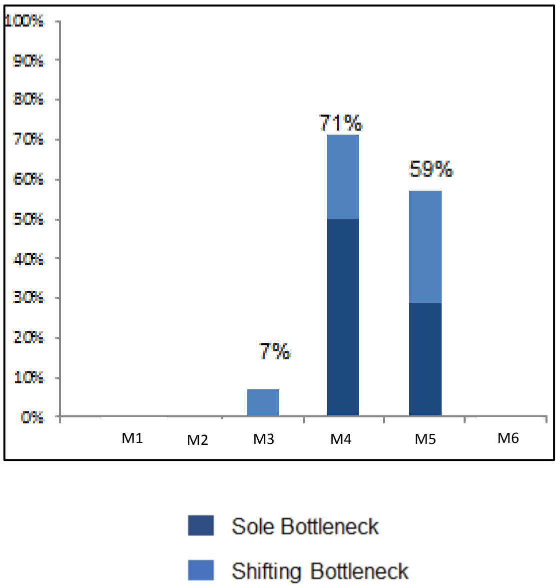
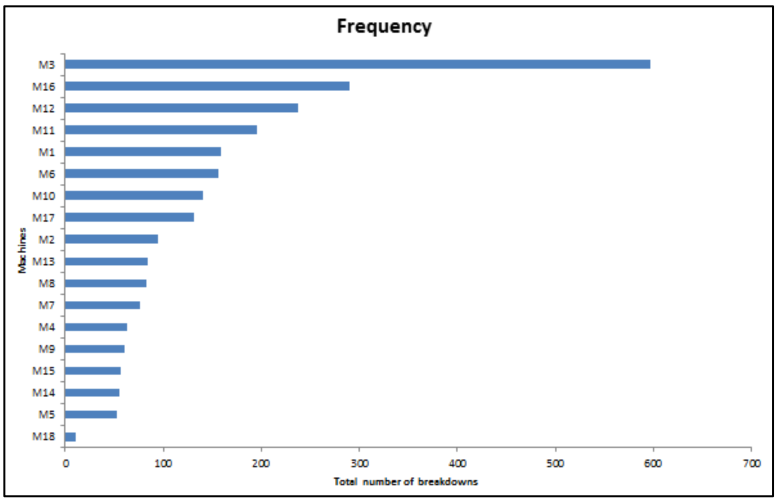


Figure 37 : Sole and Shifting bottlenecks as a percentage of the time frame

From Figure 37, it can be seen that machine M4 is the primary bottleneck over the specified period of time as it is a bottleneck for 71% of the time.

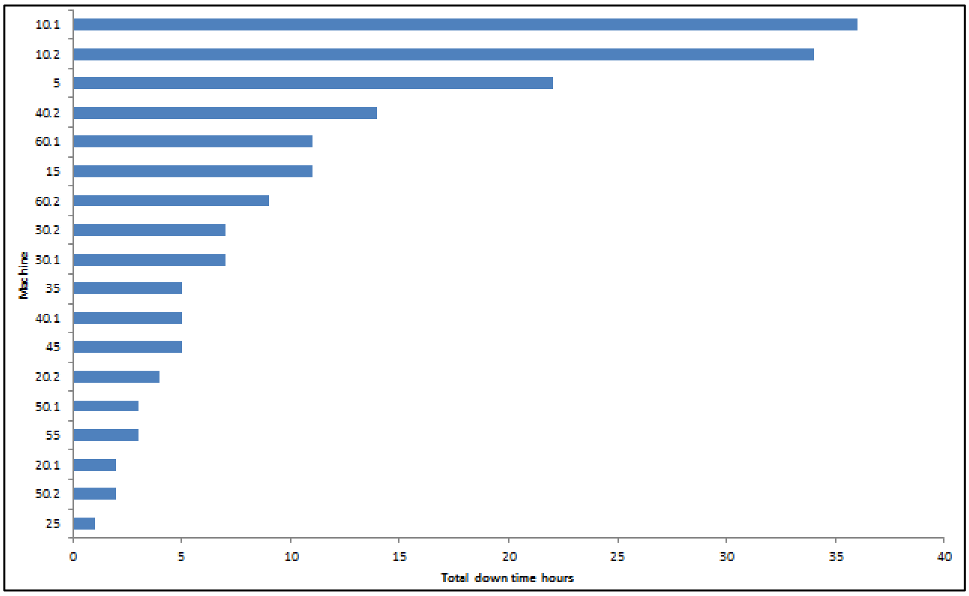
II.4.2. Frequency of Breakdowns and Total down Time of the Machines

The frequency of breakdowns and the total down time of the machines is also determined from the real time data. The frequency of the machine and the total down time of the machines in AAA line is shown in the Figure 38 and Figure 39.



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	44
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 38: Total number of breakdown occurrences in AAA line



Interval Period	01-Sep-2014 to 30-Oct-2014
Number of scheduled days	44
Start time of the shift	06:30:00
End time of the Shift	23:30:00
Number of machines	18

Figure 39 : Total down time of Grovel line machines

From Figure 38 and from Figure 39, it could be seen that the machine with highest number of breakdowns is M3 and its corresponding total down time is 22 hours which is not the highest. This indicates that the MTBF is less when compared to other machines. On the other hand, the machine M1 has 36 hours total down time but only 159 occurrences. This indicates that the machine M1 has larger MTTR when compared to other machines. The same analysis can also be replicated for the machines in the BBB line.

II.4.3. Predictive Modelling of Production Indicators

The predictive modelling is a collection of mathematical techniques to in a mathematical relationship and to predict the future values(Dickey, 2012). The predictive modelling of the production line is done from two aspects: Maintanence and line throughput. The aim of the maintenance predictive modelling is to identify the futuristic breakdown pattern of the different machines and the likely hood of the breakdown trend of different machines. On the other hand, with the breakdown pattern taken into consideration, the confidence level of production line meeting the set target production is also found out. These models will be helpful for the production and maintenance teams to design their action strategies.

II.4.3.1. MTBF data modelling

The futuristic breakdown pattern of the different machines is found by modelling the MTBF data. From the historical data, the MTBF is calculated for all the machines i.e. the time from the start of the failure till the start of the next failure is calculated from the data. The statistical distribution of the MTBF data for each machine is found through EasyFit 5.6 Professional software. The probability density function is also found for each fit. The Kolmogorov-Smirnov (KS) tests are used to find out the statistical distributions. The MTBF data was calculated for the machines in the AAA line and the probability density plot for each machine is drawn for the length of the shift as shown in Figure 40.

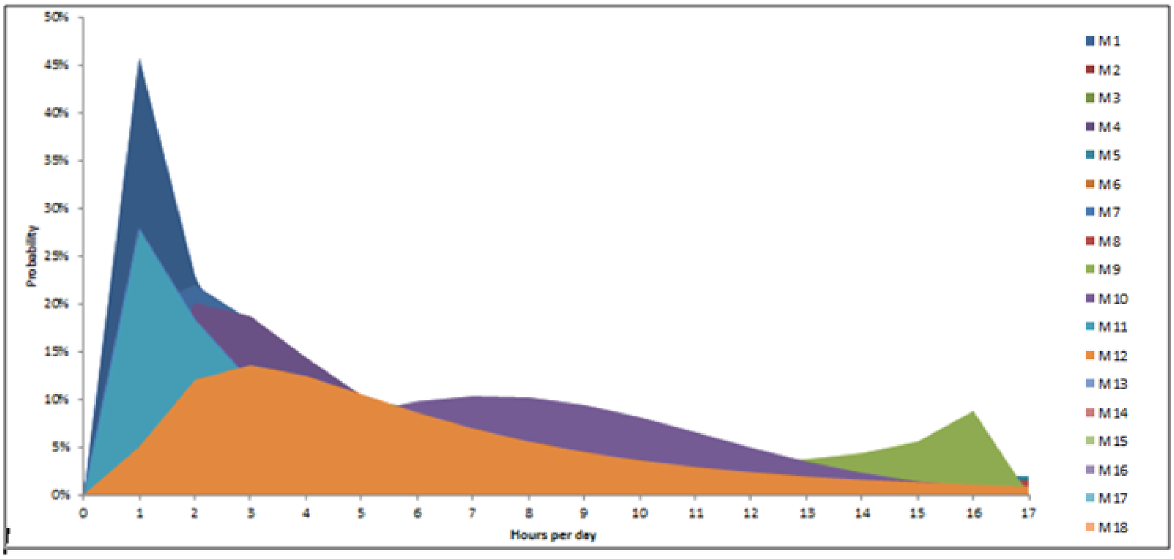


Figure 40 : Probability density plot of AAA Line machines

From Figure 40, it can be seen that that during the first hour of production from the start time of the shift. This statistical distribution fit will be helpful for maintenance team in order to design their tactical maintenance strategies.

II.4.3.2. Breakdown as a percentage of scheduled hours

The breakdown as a percentage of scheduled hours is an indicator of the time lost due to breakdowns. With the plot of this indicator over a time period, the trend of the breakdown is visualised. Moreover, having control limits on this breakdown time, it is possible to predict the general causes of the breakdown. Figure 41 shows the breakdown pattern of the machine M1, which the bottleneck machine was as pointed out by mean active period method, over the two month time period plotted in Minitab Statistical Software. It could be inferred from the Figure 41 that majority of the downtimes are due to common causes as they are within the three sigma level. The moving range chart shows the variation as calculated from the ranges of two successive breakdown percentages while on the other hand, the individual chart represents individual breakdown percentage. As all the point in the moving range chart is within the control limits, one can be sure that the breakdown pattern doesn't have an unstable variation. Following this, one point is outside the control limit in the individual plot, shows that on that particular day, a major breakdown or huge number of small breakdowns have occurred in the machine and it has shifted more than three sigma levels from the average of the breakdowns. One more observation from the individual chart is that, a roller coaster trend is seen towards the end which needs more attention in detailing out the reasons for the losses as some problems seems to be causing a uniform drift to both sides of the centre line.

IMR chart – Machine M1

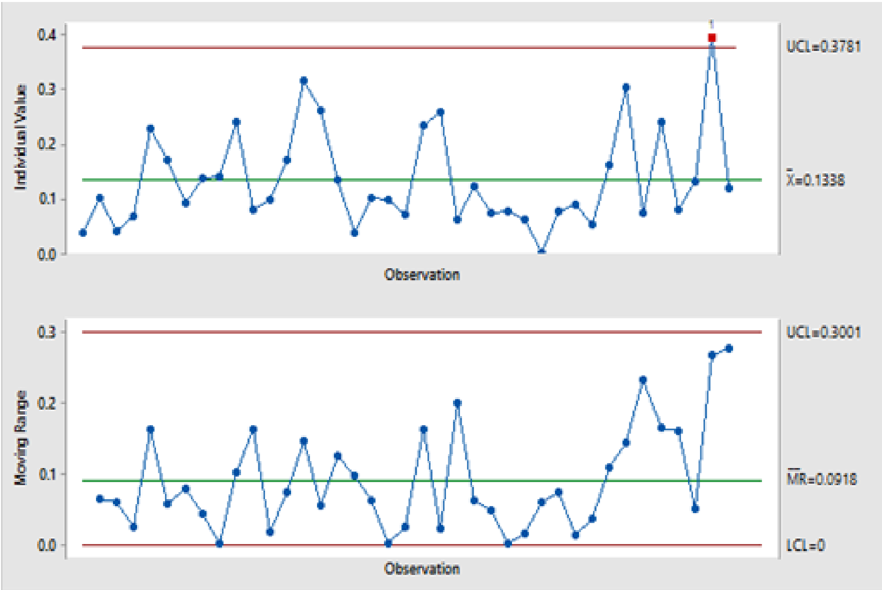


Figure 41: IMR trend chart for breakdown as a percentage of time

II.4.3.3. Throughput estimator

The throughput is one of the key parameter that is important for production engineers. From the past real time data, the expected throughput or the run rate can be calculated. Also, by using the probability approaches, the confidence level of the throughput is also calculated. The last machine in the AAA line is M18. Though the throughput numbers from this machine is not available, the *producing* state details of the machine are used to calculate the throughput of the AAA line. The count of *producing* states is the throughput. The historical throughput data was then modelled and the statistical distribution of the throughput data is obtained. Although, there are various statistical distributions could be fit for the data, the Weibull distribution is used as it is the most flexible distribution and is a representative of all other distributions by changing the shape and the scale parameter. The Weibull distribution of the historical throughput as shown in the Figure 42 is then - used to predict the throughput on a daily basis as shown the Figure 43.

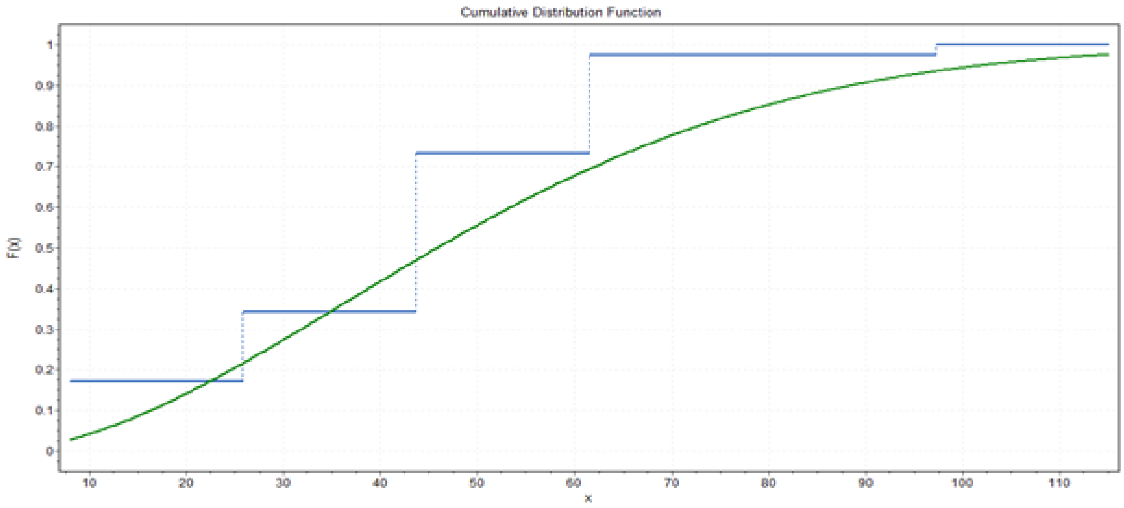


Figure 42 : Cumulative Weibull distribution function of the historical throughput data

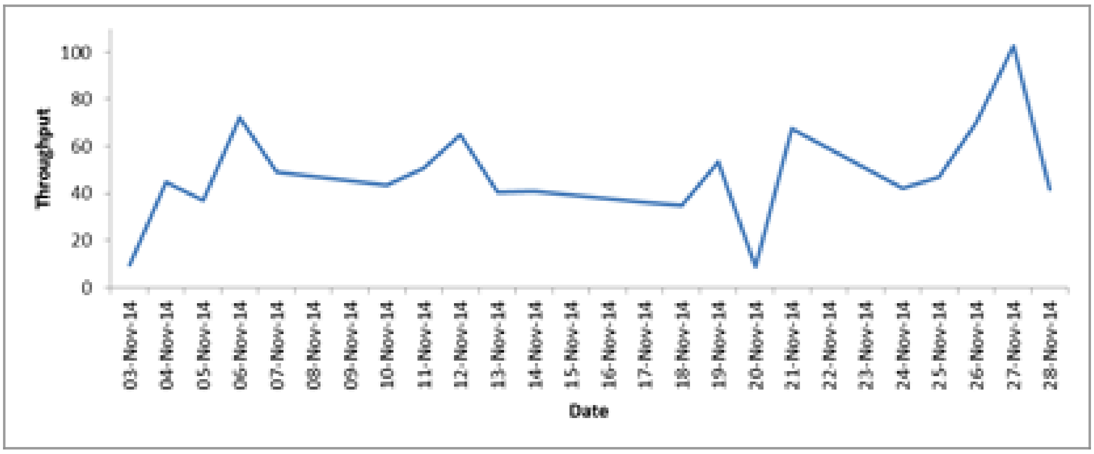


Figure 43: Daily throughput prediction using the Weibull distribution from one sample iteration

100 iterations were run and the average daily productions are added to give the monthly production. Figure 44 show that, the AAA line could handle a monthly volume of 870 at 90% probability level. With the volume greater than 870, the probability of achieving the production drops indicating more strategic plans need to be made by the production team in order to achieve the monthly production.

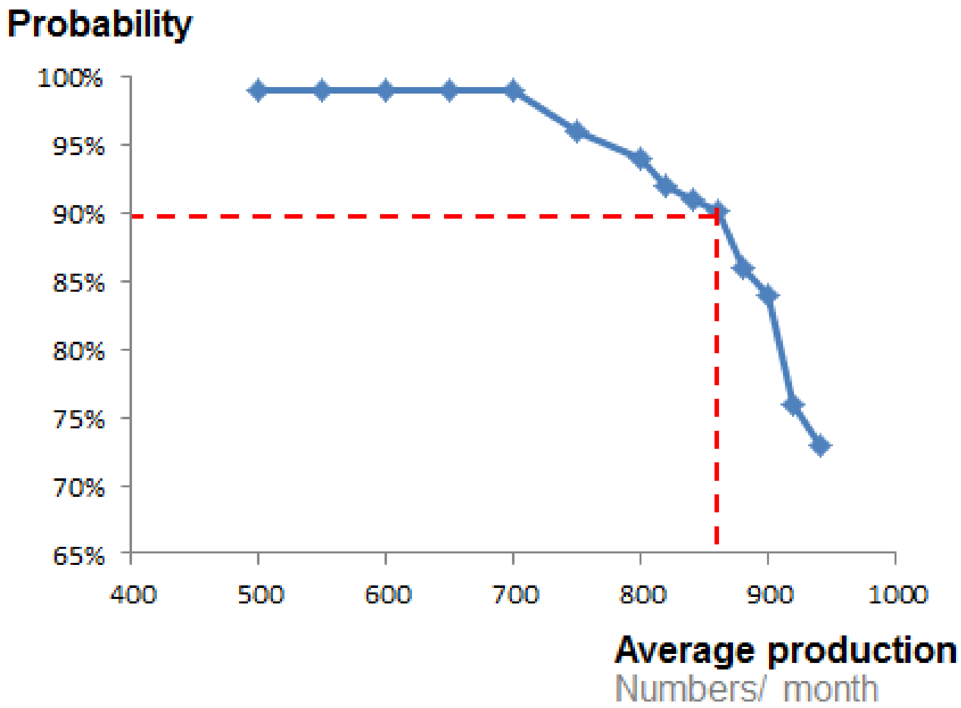


Figure 44: Throughput estimation with probability

II.5. Algorithms and Discussion

This chapter explains the creation of generalised algorithms of the bottleneck detection methods and the maont

II.5.1. Static bottleneck detection methods

The three static bottleneck detection methods were possible to carry out from the real time data: Active period, mean active period and the cycle time method. Algorithms are defined for each one of these method. An assumption is made here that, the MES records the data for each individual machine and the output from the MES data files are individual for each machine as described in the Methodology Section.

II.5.1.1. Active period percentage method

This method takes only the active states of the machine. Below is the generalised algorithm to calculate the active period percentage of a machine.

Algorithm

Step 1 : Start
Step 2 : Specify the time interval
Step 3 : Calculate the sum of active periods of the machine
Step 4 : Calculate the active period percentage according to the Equation 4

$$A_i = \left(\sum_{1}^{k} a_{i,k} \div (T-t)\right) \times 100$$

Equation 4

Where A_i= Active period percentage

$a_{i,k}$= Individual active periods

T = Start of the time interval

t = End of the time interval

Step 5 : Stop

Similarly, the same algorithm is used for all the machines in the production line to compute the active period's percentages of the entire machines in the production line.

The inputs are the time instants of the machine states and the tie intervals. The model output is the active period percentages of the machine. The algorithm is validated over the random data generated and compared the results with the manual calculations.

The advantages of this algorithm are that the algorithm is defined for each machine and thus this algorithm can be implemented independently of the structure of the production system. Adding on it uses only the data recorded in the event log files and therefore it was easy to calculate from the event log files. Though this algorithm is trained on the given data set, a sample random test data representing the problem was created and this algorithm was tested. To validate the results, the manual calculations were also done on the test data set and the results were compared.

Also as pointed out by Wang et al (2011) that this method assumes that the throughput from each machine is the same when the machines process different products during different active periods. This aspect is widely affected when there are different processing times, different product based breakdown patterns, setup times etc. Moreover, with this active period method percentage method, the small and the large active durations cannot be interpreted. For example, with this method it cannot be said for sure that the machine is active for the time interval continuously or the active period is discontinuous but occurs more frequently

II.5.1.2. Mean active period method

Like active period percentage method, this method also takes only the active states into consideration in determining the bottlenecks. Furthermore, the generalised algorithm is same as that of the active period percentage method until the active periods of machine *i* is stored in 1 x n matrix.

Algorithm

Step 1 **:** Start
Step 2 **:** Specify the time interval
Step 3 **:** Calculate the mean of the active period of the machine according to Equation 5
Step 4 **:** Calculate the standard deviation according to Equation 6
Step 5 **:** Calculate the confidence intervals according to Equation 7
Step 6 **:** Stop

The mean active period of the machine *i* is calculated using Equation 2. This equation is adapted from Roser et al. (2001).

$$\overline{A}_i = \frac{\sum_{k=1}^{k=n} a_{i,k}}{n}$$

Equation 5

where $\overline{A}_I$= Mean active period

$a_{i,k}$= Individual active periods

n = number of active periods

The standard deviation is calculated according to Equation 3. This equation is adapted from Roser et al. (2001).

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^{k=n} |(a_{i,k} - \overline{Ai})|}{n-1}}$$

Equation 6

The confidence intervals are then calculated at 95% confidence level by Equation 4. This equation is adapted from Roser et al. (2001).

$$\mathrm{CI}_i = Z_{\alpha/2} \frac{\sigma_i}{\sqrt{n}}$$

Equation 7

Similarly, the same algorithm is used for all the machines in the production line to compute the mean active period of all the machines.

The inputs are the time instants of the machine states and the tie intervals. The model output is the mean active period of the machine. The algorithm is validated over the random data generated and compared the results with the manual calculations.

Unlike active period percentage method, the mean active period could detect the bottleneck accurately at 95% confidence level. The similarity being with the active period percentage is that it follows the same algorithm until the active period matrix is calculated. Adding on, it's relatively simple to calculate from the data set, only the information from the log files were used as pointed out by Roser (2001).

Like active period percentage method this method also assumes the throughput from each machine is same and when each machine produces different products during different active periods(Wang, Chen, Wang, Zhang, & Sun, 2011). Though the mean active period includes a confidence interval, the primary bottleneck could be detected only when the confidence interval of the machines doesn't overlap. Otherwise it cannot be said for sure which machine is the bottleneck. Moreover the effect of the average concept will have skewed results when the active period set has a larger distribution of the data. Both the active period percentage and mean active period assumes that operator was not the bottleneck. For example, the machine could be waiting for the maintenance operator to be addressed when the machine is down. So the decision has to be taken whether the operator/ mechanic should be considered as a separate entity in the bottleneck detections using this method.

II.5.1.3. Cycle time method

This method is also very useful method in order to detect the bottlenecks more quickly from utilisation point of view. The generalised algorithm for this method is described below.

Algorithm

Step 1 : Start
Step 2 : Specify the interval of time
Step 3 : Calculate the median time of the producing modes of the machine
Step 4 : Stop

Similarly, the same algorithm is used for all the machines in the production line to compute the median cycle time of all the machines.

The inputs are the time instants of the machine states and the time intervals. The model output is the median cycle time of the machine. The algorithm is applied over the real data set and the sample random data. The algorithm is validated over the random data generated and compared the results with the manual calculations.

The same algorithm is run for all the machines in the production line and the median cycle time is compared to find out the bottlenecks especially when there is a large varieties of the products in the production line and the machines MES doesn't monitor the type of product rather it only monitors the machine states. The cycle time method is a very easy method to detect bottlenecks from the real time MES data. But on the other hand, it cannot detect the secondary bottlenecks and short term bottlenecks.

The median cycle times when compared across the machines takes the different product variants processing times of the machines into consideration. This is one of the strength when compared to the former methods on bottleneck detection. On the other hand, it takes only the machine cycle time (including loading time to machine, unloading time and the machine processing time) into consideration, and not the other active states of the machine like breakdowns, setups etc.. So to conclude, this method only points the machine which improves the throughput only from the machine cycle time perspective. Also, this method is a good method when applied to machines which has technical availability of 100%. Yet another situation is that, it can use in situations where the work station activities are completely manual as the manual cycle times could be computed and the median could be taken.

II.5.2. Momentary Bottlenecks

The bottlenecks at any given instant of time can be determined from the real time data by detecting the sole and shifting bottlenecks. Unlike the static bottlenecks algorithms design for each individual machine and then comparing the results across the machines in the production line, the momentary bottlenecks requires different approach. The first step in determining the momentary bottlenecks is to aggregate the data of all machines in a production line into one data file for the required time interval. An algorithm is defined over this aggregated data in order to identify the sole and the shifting bottlenecks.

II.5.2.1. Sole Bottlenecks

The algorithm to determine the sole bottleneck machine in a production line is explained below.

Algorithm

Step 1 : Start
Step 2 : State the current instant *t* and the number of the machines
Step 3 : Check whether the which machines are active at *t*
Step 4 : Determine the machine having the longest uninterrupted active period at *t*
Step 5 : The machine with the longest uninterrupted active period is the sole bottleneck
Step 6 : Stop

The inputs are the current time instants, number of machines and the machines states and their time instants. The output is the bottleneck machines at that time instant. This algorithm is useful in finding out the sole bottlenecks at the defined time instants. The algorithm is applied over the real data set and the sample random data. The algorithm is validated over the random data generated and compared the results with the manual calculations.

II.5.2.2. Shifting bottlenecks

Shifting bottlenecks algorithms could be considered as an extended part of the sole bottlenecks.

Algorithm

Step 1 : Start
Step 2 : The sole bottlenecks are determined at time instant *t.* Let the sole bottleneck machine be

M1

Step 3 : The starting time instant of the sole bottleneck M1 is determined. Let that be *t1*

Step 4 : At time instant *t1,* determine the sole bottleneck machine and name it as M2

Step 5 : The starting time instant of the sole bottleneck M2 is determined. Let that be *t3*

Step 6 : Determine time instant when the active period of *M2* ends is the time interval of *t1* to *t.* Let this time instant be *t2.*

Step 7 : Similarly repeat the steps 4 to step 6 until the start of the shift

Step 8 : Calculate the shifting and the sole bottlenecks of the machine according to Equation 8 and Equation 9 respectively,

$$\text{Sole percentage machine} = \frac{(non - overlapping\ time\ intervals)}{shift\ interval} \text{x} 100$$

Equation 8

$$\text{Shifting percentage machine} = \frac{(overlapping\ time\ intervals)}{shift\ interval} \text{x} 100$$

Equation 9

Step 9 : Stop

The inputs are the machine states with time instants and the number of machines. The output is the sole and the shifting percentages of the machines. The algorithm is validated over the random data generated and compared the results with the manual calculations.

As explained by Roser et al. (2003), the momentary bottlenecks show the primary, short term, long term and the non-bottlenecks of the production system. Also, it can be observed from the pattern of the real time data of AAA line that, some machines seems to be active for the entire period of the shift and the shifting pattern of the bottlenecks are hardly visually observed. In this case, the machines with the highest uninterrupted active period become the sole bottleneck machines. Also, one more observation is that, when this method is used in a production system which has more number of machines, then the complexity increases to identify the shifting pattern between the machines. One big advantage of this method is that, the scaling down and scaling up approach. This method could be applied to even fifteen minute interval to detect the bottleneck and could be scaled up to hours or even up to one day and it detects the bottlenecks at the defined instant of time with accuracy. Another advantage is that the shifting bottleneck approach can also be used to detect the secondary bottlenecks due to the overlapping functionality. Also according to Roser et al. (2002), this approach of bottleneck detection showed the primary stations which restrain the capacity of the production line. Though the shifting bottleneck approach detects the primary bottleneck machine of the production line, this method cannot reveal what type of action is necessary in order to debottleneck the machine.

II.5.3. Summary of different bottleneck detections methods

The four different methods to detect the bottlenecks are summarised in the Table 11. From Table 11 it could be inferred that the shifting bottlenecks is the most superior bottlenecks among the four types of the bottleneck detection methods and this is because of the accuracy in determining the bottlenecks and the ability to detect the bottlenecks at different time intervals (Roser et al, 2002). The accuracy is high because this type of bottleneck detection techniques is based on the nature of the production systems while the other methods are standalone methods and a comparison is drawn to detect the bottlenecks. The median cycle time is also a good method to detect the bottleneck especially where the cycle times are widely spread due to the product variant. On the other hand, this method doesn't take the availability of the station into consideration. But it could be argued that the availability of the station is highly dynamic in nature and one can't predict the exact availability of the station in spite of efficient and effective operational practices. In this situation, the cycle time method is very useful for bottleneck detection as the machine with the highest cycle time governs the line throughput. Also, it indicates the amount by which the cycle time of the bottleneck machine should be reduced according to Wiendahl and Hegenscheidt (2003).

Table 10: Comparison of the four different bottleneck machines

Method	Active period percentage method (%)	Mean active period method (sec)	Median cycle time method (sec)	Shifting bottlenecks (%)
Factors	All active durations	Mean of the active duration	Median product processing times	Overlapping active duration and longest un interrupted active duration
Algorithm	Independent of production system	Independent of Production system	Independent of production system	Much dependent on production system
Reason for why the machine is a bottleneck	Machine with largest active percentage duration will cause the downstream machine to starve and upstream machine to be blocked	Machine with the largest mean active period is least likely to be interrupted by other machines and has a large effect on the overall system output	Machine with the largest cycle time will govern the overall system output even if the technical availability of the station is 100%	Machine with the largest sum of shifting and sole bottlenecks over a period of time reduces the overall output
Indicator of bottleneck (How one can tell that the machine is the bottleneck)	Largest active period percentage	Largest mean active duration with confidence interval	Highest machine processing time	Sum of percentage of sole and percentage of shifting
Dynamic analysis	No	No	No	Yes
Accuracy of the detection	High – when the machine has the highest active percentage among all other machines Low – if the active period percentages are the same	High - when the confidence interval doesn't overlap with other machines Low –if confidence interval overlaps	High – when the stations have a availability of 100% Low – in other cases as only the processing time is used as an indicator	Very high – when there is a sole bottleneck at the time instant of interest Medium – When no machine is active at that time instant of interest and the sole and shifting are done based on the past data Low – When two or more machines have the same uninterrupted active period at the time instant of interest
Representation of product variants	No	No	Yes to some extent as median indicator is used	No
Representation of shifting bottlenecks	No	No	No	Yes
Ease of algorithm implementation	High	Medium	High	Low

II.5.4. Frequency of Breakdowns and Total time of Breakdowns

The frequency of the breakdowns and the total time of the breakdowns can also be calculated from the real time data.

Algorithm

Step 1 : Start
Step 2 : Specify the time interval
Step 3 : Calculate the number of error/down states to determine the frequency of the breakdowns
Step 4 : Calculate the time elapsed of the error/down states
Step 5 : Calculate the total time of the error/down states according to the Equation 10

$$DO_i = \left(\sum_{k=1}^{k=n} do_i\right)$$

Equation 10

Step 6 : End

The inputs are the time interval, machine states and their time instants. The output is the frequency of the error/down state and the total downtime of the machine. The algorithm is applied over the real data set and the sample random data. The algorithm is validated over the random data generated and compared the results with the manual calculations.

The frequency and the total down time calculations from the real time data is a very useful metric for the maintenance department to develop long term strategies. Also, this analysis will indicate which machine requires condition monitoring. Adding on, this analysis also throws light on the maintenance key performance indicators (KPI's) like MTTR and MTBF. For example, the machine with the highest frequency of breakdowns and has lowest total down time indicates that MTBF for that machine is lower when compared to other machines in the line. On the other hand, if the machine with lower frequency of breakdowns and higher total down time indicates that the MTTR is high for that machines when compared to other machines in the production line.

II.5.5. Predictive Modelling of Production indicators

The performance indicators of production system provide valuable information of the production system. The performance indicators like MTBF, percentage breakdown over scheduled hours, throughput is determined directly from the real time data.

II.5.5.1 MTBF Data Modelling

From the real time data, it was straightforward to calculate the MTBF and to find the statistical distribution. Though the statistics delivers more insights, the disadvantages of this statistical fit are that more data needs to be collected in order to minimise the error during the statistical fit. The MTBF is calculated from the real time data file as shown in the Figure 45.

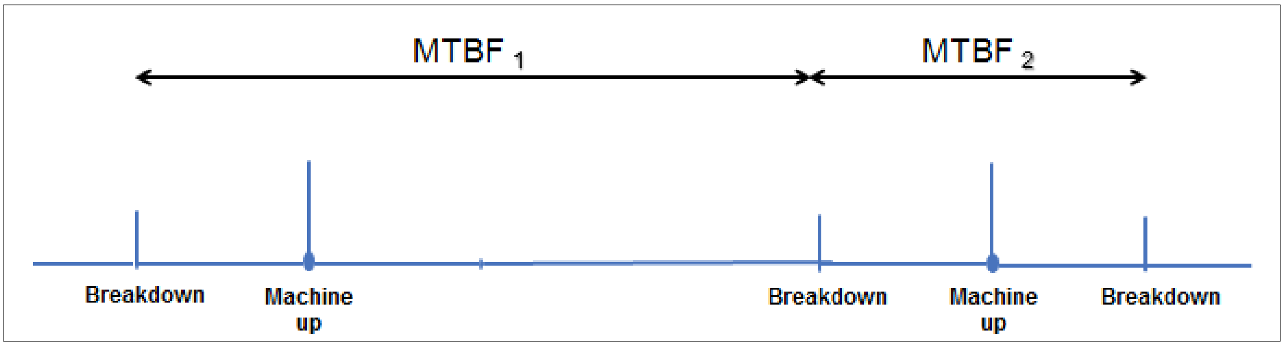


Figure 45: MTBF calculation description from the real time data

The outcome of the statistical distribution of the MTBF data can be used directly as an input to the simulation. Adding on, this method is more superior to the method proposed by Troyer (2009) which gives the average estimation of MTBF. This average estimation is affected by the extreme values if present in the data and does not indicate the correct picture.

As per Ahamed and Kamaruddin (2012), the MTBF statistical probability distributions of different machines could be compared to find out the pattern of the machine failure rate and this was also possible to do from the real time data. Also, the statistical distributions of the MTBF data of the machines, can tell the pattern of the breakdowns during any particular day. One can pin point at the exact time intervals when the machines have the highest probability to the failure. This type of information could be used by maintenance team to develop tactical proactive strategies such as frequent watch or checks on the machine by the maintenance team during the hour in which it has the highest probability to fail.

II.5.5.2. Breakdown as a percentage of scheduled hours

Monitoring breakdown as a percentage of scheduled hours could be a very important KPI for the maintenance team. An overall trend and the localised pattern analysis of this indicator will provide useful information to the maintenance team as without identifying and correcting the factors contributing to the special cause variation, the breakdowns cannot achieve a statistical state of control. Also, the increasing and decreasing trend of the breakdown pattern of a particular machine could help maintenance team to design the long term and short term strategies: for example, implementing condition monitoring of a component if the breakdown is associated with some critical component in the machine or to strengthen the preventive maintenance activities.

II.5.5.3. Throughput estimator

The prediction of the throughput from the historical data is a good indicator for the overall production team in order to assess whether the production system could meet the demand. As explained by Johnson (2011), the monte carlo method can be used in throughput estimation. The strength of the throughput prediction is that the probability density function of the throughput is built over the historical data and this throughput data includes all losses in the production system. In other words, the losses in the production system will have a direct effect on throughput. This means that the historical throughput data is a good indicator to reflect on all the losses. With these losses in the production system, the monte carlo model predicts the confidence level of the production system meeting the required demand. In general, if the confidence level is less, then the production team could plan for an overtime or allot new resources etc. for a shorter span of time in order to meet the required demand.

Traditionally, the production team calculates the success rate of achieving the demand from their experience in the production system. This method of calculating from the experience could be wrong sometimes as some scenarios could be excluded. On the other hand, this monte carlo method of confidence level estimator includes all extreme scenarios. Hence, this method is much more reliable in predicting the success rate to meet the demand.

II.5.6. Summary of the discussion

The argument made by Roser et al. (2001) that the machine with the largest mean active period with confidence intervals contains reasonable faults when compared to the results drawn from this thesis. The most important is that Roser et al. (2001) doesn't argue for instances when the confidence intervals of different machines overlap. So it can be concluded that the machine with the largest mean active period need not be the primary bottleneck if it has overlapping confidence interval. Also, Roser et al. (2001) didn't acknowledge the fact that average parameter is not a good indicator when the machine's active period intervals are not equal and are widely spread. This was also seen in the results drawn from this thesis.

Similarly, Roser et al (2002) argument on the machine with the highest sum of shifting and sole bottleneck percentages is the bottleneck machine is concluded with overreaching assumptions. The Roser et al (2003) also draws the same conclusion in analysing yet another case in which the pattern of shifting bottlenecks is almost the same as the first case. The author doesn't include the case where two machines have equal uninterrupted active period. The argument might have been strengthened if the analysis of the latter case also leads to same conclusion. The results from this thesis point to the fact that in a complex production system some machines are always active from the start to the end of the shift, in which case the shifting pattern is not observed.

Also it could be found that the turning point method for bottleneck detection from the real time data which was proposed by Li et al. (2009) cannot be applied to the empirical data set as the data set doesn't have the blockage and the starvation times of the machine explicitly. Though the author claims to have developed the generalised algorithm for this method, the algorithm could not be used in the empirical data set in this thesis and this is due to the fact of the lack of the necessary data collection of the process.

The bottleneck approaches should be selected based on the nature production system and the type of the goal that is aimed for. For example, if the goal is to get the maximum throughput from the production line of machines which has very high reliability, then cycle time method is a good indicator of the bottlenecks.

MTBF is very useful metric and could be derived from the real time data. The method of calculation presented in this thesis is based on computing the exact time intervals. This is in contraction to Toyer (2009) method of MTBF calculation which was based on again the averages parameter. Also, the percentages of breakdowns trend and the monte carlo simulation of the throughput parameter was proved to be good pointers based on which the tactical decisions could be made.

II.6. Conclusion

In this chapter the conclusion drawn from the study are presented. This is done by answering the research questions.

RQ 1 : How can the real time data be used to visualize bottlenecks and downtime parameters?

The effective bottleneck detection techniques: Active period, mean active period, cycle time method and shifting bottlenecks are tried over the real time. It was possible to carry out these methods over the real time. Moreover, the numerical results from these methods were converted into simple graphs in order to detect the bottlenecks visually. Also, the frequency of breakdowns in and the total down time in a given interval of time are the useful indicators to develop maintenance strategies.

Though different bottleneck methods were tried out, the application of the bottleneck detection methods should be based on the nature of production system.

RQ 2 : How can the predictive analytics deliver value for tactical decision making process?

The predictive analytics in terms of confidence level predictor of the expected service level of the production system to the forecast demand, maintenance trend analytics delivers valuable information to the production and the maintenance team in order to better plan their tactical activities.

Though the predictive analytics is very useful, one has to understand that it is the forecast and this could change from the reality as the production disturbances like sudden breakdowns etc. cannot be foreseen with accuracy.

From the results of this thesis, it is understood that the variability elitists in the production system and the real time data analytics could help in understand the causes of the variability and help the production and the maintenance team to develop their tactical and long term strategies. This could be seen from the below Figure 46 which shows that the daily output of the given production line variation across the days.

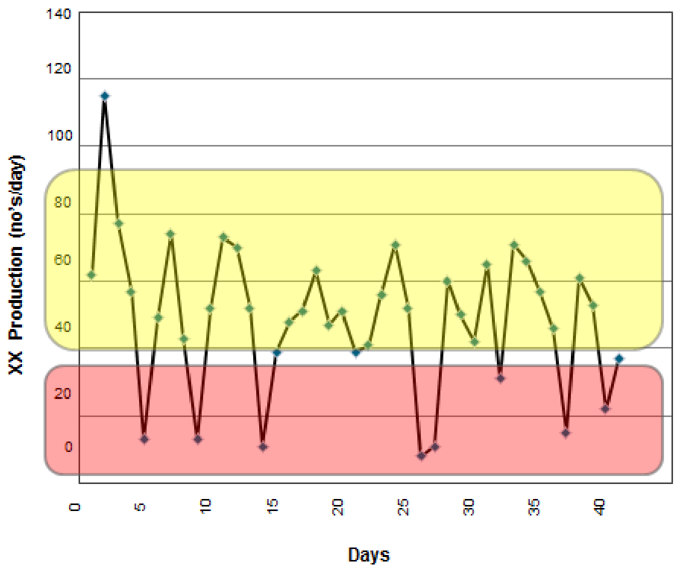


Figure 46: Daily output of BBB line

From Figure 46 , the first highlighted yellow area shows that the production per day varies mostly between 40-50. This variation could be due to the short term bottlenecks or disturbances. The second highlighted red area shows that there are some critical losses affect the system and pushes down the throughput to the level of 10. Analytics is a powerful operation on the real time data, to identify the reasons behind this variation which is proved in this thesis.

II.6. Future Work

The algorithm drawn from this thesis could be tested on different production system real time data. Adding on, the automation of algorithm over the real time data is an important step for having continuous decision support. Also, more predictive analytics could be extended further to determine the future active states of the machines during the consecutive production run day and to predict the bottlenecks from this forecast data of active states. This will further make the production team to plan better for the future.

Predictive analytics tell what will happen in the future. Perspective analytics means what must be done which is an extension of predictive analytics. This aspect could be explored for autonomous decision making by machines with the help of the real time data.

2. Overall Discussion

In the two parts of this thesis, it was proved that analytics over the different real time data sets provided valuable insights. Firstly the difference between the data sets is that the first data set was unstructured and a framework was developed to make it structured and the second data set was more structured and was uniform across the machines. In both these cases the data analytics yielded valuable insights. The analysis on the first data set reveals how important is to measure the operator influenced loss times in order to improve the productivity and the importance to capture the losses in a standardised format. The analysis on the second data set reveals how the data could identify the bottlenecks, show the trends of maintenance KPI's etc.

The interesting part of the analytics is that the decisions could be made based on the facts that are derived from analytics and not by rough judging or guessing. However, one of the key important finding in working on two data sets is that, the data analytics is best applied to questions which has uncertain answers and answering that question will have add a huge value to the company. On the other hand, it could be argued that, for the questions which have low level of uncertain answers and answering those doesn't add much value, then judgemental way of making decisions can be used. For example, it could be seen that finding the bottlenecks was a critical and question with an uncertain answer due to the complex production system and the variety of activities. Identifying the bottlenecks with certainty is possible using analytics which will help to improve the throughput from the production line to a large extent. In this case data analytics will show the trends and identify the bottleneck machines in the production line. Simulation could also be argued as an alias to data analytics but it often takes more time to build the model and interpret the results.

The two different strategies of the analytics in two different data sets yield important insights. Reflecting on those, it could be generalised that three different aspects are important when working with such type of big production data. Those are the importance on avoiding the sampling error, too much focus on numerical precisions and stability. The sampling error is that picking up of a small sample from the population and analysing instead of analysing the entire population data in order to identify the trends that affect the production system performance. The second one is that the evaluation of trends is more important in the production analytics rather than focusing on numerical precisions of the results. The third one is that use of the data insights to assess the impact of each event which are likely to occur and to develop strategy using those inputs rather than only planning for exclusive scenarios and ignoring others. Also, it is important to note that analytics itself will not provide the right answers. It is only about getting the trends and it needs to be used as an enabler to initiate the scenario based thinking. Yet another important insight is that, the KPI's needs to be identified even before looking at the data so that the data analysis could be made to look only that particular KPI trend.

From analysing the two sets of data and assessing the value it can yield, it is imperative that manufacturers could use the real time data to discover the new waves of productivity potentials. The common theme of analytics and the use of it in the two different data sets actually reveal the type of behaviour of the machines and the type of behaviour of the losses in the production system. The analysis on the past data and the predictive analytics on the data, which is in the form of huge number of log files of the machines, could throw new trends that is hidden in the production system and which is not always thought by the companies. Without the analytics on the real time data, looking for the areas of potential improvement is like looking for a needle in a haystack.

3. Conclusion

The findings from the Data Set 1 are the operator influenced loss times and their impact on OEE. The findings from the Data Set 2 are the bottleneck detections and the predictive analytics. These results are the productivity potentials which were identified by data analytics.

The operator influenced loss times, the bottleneck detection techniques and the KPI trends could be directly derived from the real time data using analytics and this method is much faster and hence the fact based decisions could be made and quickly. By improving the operator influenced loss times, the OEE could be increased. Similarly, by detecting the bottlenecks in the production line, it could be managed efficiently in order to achieve higher throughput. Adding on, the predictive analytics could help to foresee the failure pattern of the machine and helps to design the right maintenance strategy, thus increasing the overall availability of the system.

The operator influenced loss times, detecting the bottlenecks and identifying the failure patterns are some of the ways to manage and improve higher productivity. This thesis demonstrates that the data analytics is a valuable tool to identify the productivity potentials from the massive influx of big data, as well as to develop analytical platforms that can run over the real time data to deliver key insights.

References

Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers and Industrial Engineering*, *63*(1), 135–149. doi:M1016/j.cie.2012.02.002

Almström, P., Hansson, E., & Samuelsson, J. (2014). How to imporve productivity by 160%, *1*(M).

Andersen. (1999). Industrial benchmarking for competitive advantage. *Human Systems Management Journal, Vol. 18.* doi:3/4

Andersson, C., & Bellgran, M. (2015). On the complexity of using performance measures: Enhancing sustained production improvement capability by combining OEE and productivity. *Journal of Manufacturing Systems*, *35*, 144–154. doi:M1016/j.jmsy.2014.12.003

Aneja, Y. P., & Punnen, a. P. (1999). Multiple bottleneck assignment problem. *European Journal of Operational Research*, *112*, 167–173. doi:M1016/S0377-2217(97)00392-5

Antony, G.G. (2008). How to determine the MTBF of grearboxex. Available online: http://www.powertransmission.com/issues/0408.antony.pdf. Accessed: 15 May 2015

Application within : FFI , Hållbar produktionsteknik Streamlined Modeling and Decision Support for Fact-based Production Development. (2013), 1–15.

Attiany,M. (2009). The role of benchmarking in improving institutional performance of the jordanian pharmaceutical fims. PHD,dissertation, not published, *Arab acdemy for banking and financial services,* Amman, Jordan

Bailey, D. E., & Barley, S. R. (2005). Return to work: Toward post-industrial engineering. *IIE Transactions*, *37*(8), 737–752. doi:M1080/07408170590918308

Ben-Daya, M., Duffuaa, S.O., Raouf, A., Knezevic.J., & Ait-Kadi, D. (2009). Handbook of Maintenance Management and Engineering. Vol . XXVII, 2-426

Bergström, F., Plamkvist, N. (2014). An analysis to increase the productivity of a surface mounting line. *Master thesis report in Production Engineering*. Chalmers University of Technology

Boxwell, Robert, J., Jr. (1994). Benchmarking for competitive advantage./ *New York McGraw-Hill*. ISBN 0-07-006899-2

Boysen, N., Fliedner, M., & Scholl, A. (2007). A classification of assembly line balancing problems. *European Journal of Operational Research*, *183*, 674–693. doi:M1016/j.ejor.2006.10.010

Bruno, J., & Steiglitz, K. (1972). The Expression of Algorithms by Charts. *Journal of the ACM*, *19*(3), 517–525. doi:M1145/321707.321720

Chiang, S. Y., Kuo, C. T., & Meerkov, S. M. (1998). Bottlenecks in Markovian production lines: A systems approach. *IEEE Transactions on Robotics and Automation*, *14*(2), 352–359. doi:MM209/70.681256

Chiang, S.-Y., Kuo, C.-T., & Meerkov, S. M. (1999). C-Bottlenecks in Serial Production Lines: Identification and Application. *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No.99CH36304)*, *1*(December), 456–461. doi:MM209/CDC.1999.832820

Chien, C. F., Zheng, J. N., & Lin, Y. J. (2013). Determining the operator-machine assignment for machine interference problem and an empirical study in semiconductor test facility. *Journal of Intelligent Manufacturing*, 1–13. doi:M1007/s10845-013-0777-3

Chwif, L. (2008). Proceedings of the 2008 Winter Simulation Conference S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds., (time 1), 1746–1751.

Costa, S. E. G. Da, & Lima, E. P. De. (2002). Uses and misuses of the “overall equipment effectiveness” for production management. *IEEE International Engineering Management Conference*, *2*, 816–820. doi:MM209/IEMC.2002.1038543

Dal, B., Tugwell, P., & Greatbanks, R. (2000). Overall equipment effectiveness as a measure of operational improvement – A practical analysis. *International Journal of Operations & Production Management*, *20*(12), 1488–1502. doi:MM208/01443570010355750

Data, B., Group, I., & Greenplum, E. M. C. (n.d.). An independent report conducted by Big Data Insight Group, commissioned by EMC Greenplum and Bull.

Davenport, T.J. (2006). Competing on analytics. *Harvard Business Review.*

De Groote, P. (1995). Maintenance performance analysis: a practical approach. *Journal of Quality in Maintenance Engineering,* Vol. 1 No. 2, 4-24

De Ron, a. J., & Rooda, J. E. (2006). and Equipment Effectiveness: an Evaluation. *International Journal of Production Research*, *44*(23), 4987–5003. doi:M1080/00207540600573402

Desruelle, P., & Steudel, H. J. (1996). A Queuing Network Model of a Single-Operator Manufacturing Workcell with Machine/Operator Interference. *Management Science*, *42*(4), 576–590. doi:M1287/mnsc.42.4.576

Dhillon, B. S., & Liu, Y. (2006). Human error in maintenance: a review. *Journal of Quality in Maintenance Engineering*, *12*, 21–36. doi:MM208/13552510610654510

Dickey, D. a. (2012). Introduction to Predictive Modeling with Examples. *SAS Global Forum*, (1974), 1–14.

Dvořák, J., Malkovský, Z., & MacKů, J. (2008). Influence of human factor on the time of work stages of harvesters and crane-equipped forwarders. *Journal of Forest Science*, *54*(1), 24–30.

Ecricsson, J. (1997). Disruption analysis: An important tool in lean production. *Department of Production and Materials Engineering.* Lund Univeristy

Grosfeld-Nir, A. (1995). Single bottleneck systems with proportional expected yields and rigid demand. *European Journal of Operational Research*, *80*(94), 297–307. doi:M1016/0377-2217(94)00172-9

Hansen, R. C., (2001). Overall equipment effectivness : A powerful production / maintenance tool for increased profits. *Newyork Industrial Press, Inc.,* ISBN 0-8311-3138-1

Hedman,R.,Sundkvist,R., & Almström, P.(2014). Identification of relationships between operator utlisation and real process capacity in automated manufacturing

Huang, S. H., & Keskar, H. (2007). Comprehensive and configurable metrics for supplier selection. *International Journal of Production Economics*, *105*(2), 510–523. doi:M1016/j.ijpe.2006.04.020

James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, A. H. B. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*, (June), 156. doi:M1080/01443610903114527

Jeong, K.-Y., Jeong, K.-Y., Phillips, D. T., & Phillips, D. T. (2001). Operational efficiency and effectiveness measurement. *International Journal of Operations & Production Management*, *21*(11), 1404–1416. doi:MM208/EUM0000000006223

Jonsson, P., & Lesshammar, M. (2005). of manufacturing performance measurement systems – the role of OEE

Johnson, P. E. (2011). Monte Carlo Analysis in Academic. ?, (or MCMC), 1–67. doi:M1093/oxfordhb/9780199934874.013.0022

Jonsson, P., & Lesshammar, M. (2005). of manufacturing performance measurement systems – the role of OEE.

Konopka, J., & Trybula, W. (1996). Overall equipment effectiveness (OEE) and cost measurement [semiconductor manufacturing]. *Nineteenth IEEE/CPMT International Electronics Manufacturing Technology Symposium*, (512), 137–140. doi:MM209/IEMT.1996.559707

Kotze, D. (1993). Consistency, accuracy lead to maximum OEE benefits. *TPM Newsletter,* 4(2), AITPM, Productivity, Inc., Norwalk, November

Krumeich, J., Werth, D., Loos, P., Krumeich, J., Werth, D., Schimmelpfennig, J., … Jacobi, S. (2014). Advanced Planning and Control of Manufacturing Processes in Steel Industry through Big Data Analytics Case Study and Architecture Proposal, 16–24.

Lechevalier, D., Narayanan, A., & Rachuri, S. (2014). Towards a Domain-Specific Framework for Predictive Analytics in Manufacturing, 987–995.

Liker, J. K. (2004). The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer. New York, McGraw-Hill.

Liker,J. (2014). Is OEE a Useful Key Performance Indicator? . *Industry Week*.

Lungberg, Ö. (1998). Measurement of overall equipment effectiveness as a basis for TPM activities. *Journal of Quality in Maintenance Engineering,* vol. 18, no. 5, 495-507

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The management Revolution. *Hardvard Business Review*

Metropolis, N., & Ulam, S. (1949). Monte carlo method. *Journal of American Statitical Association*. No 247. Vol. 44

Muchiri, P., & Pintelon, L. (2008). Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International Journal of Production Research*, *46*(13), 3517–3535. doi:M1080/00207540601142645

Nakajima, S. (1988). Introduction to TPM. *Productivity Press,* Cambridge, MA

Niebel, & Benjamin, W. (1994). Engineering maintenance management. *New York: Marcel Dekker Inc.*

Puvanasvaran, P., Kim, C. Y., & Siang, T. Y. (2012). Overall Equipment Efficiency (Oee) Improvement Through Integrating Quality Tool : Case Study, (October), 15–16.

Rahman, M. M., & Kadirgama, K. (2009). National Conference on Postgraduate Research (NCON-PGR) 2009 1st October 2009, UMP Conference Hall, Malaysia © Centre for Graduate Studies, Universiti Malaysia Pahang Editors: M.M. Noor; M.M. Rahman and K. Kadirgama. *National Conference on Postgraduate Research (NCON-PGR)*, (October), 165–175.

Roser, C., Nakano, M., & Tanaka, M. (2000). Tracking shifting bottlenecks. *Japan - USA Symposium on Flexible Automation*.

Roser, C., Nakano, M., & Tanaka, M. (2001). A practical bottleneck detection method. *Proceedings of the 2001 Winter Simulation Conference, 2001.*, 1079-1086

Roser, C., Nakano, M., & Tanaka, M. (2002). Shifitng bottleneck detection. *Proceedings of the 2002 Winter Simulation Conference, 2002.*, 1079-1086

Roser, C., Nakano, M., & Tanaka, M. (2003). Comparison of bottleneck detection methods for AGV systems. *Proceedings of the 2003 Winter Simulation Conference, 2003.*, *2*, 1192–1198. doi:MM209/WSC.2003.1261549

Sargent, R. G. (2010). Proceedings of the 2010 Winter Simulation Conference B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, eds. *Simulation*, (2001), 135–150. doi:MM209/WSC.2010.5679148

Shearer, C. (2000). The CRISP-DM model: the new blue print for data mining. *Journal of Data Warehousing, 5*(4), 13-19

Shingo, S. (1985). A revolution in manufacturing: the SMED system. *Productivity Press, Cambridge, MA*

Smith, R., & Hawkins, B. (2004). Lean maintenance. *Butterworth Heinemann.* ISBN: 978-0-7506-7779-0

Stecke, K. E. (1982). Machine interference: the assignment of machines to operators in the handbook of industrial engineering, edited by Gavriel Salvendy. *New York: John Wiely & Son*

Stecke, K. E., & Aronson, J. E. (1985).Review of operator/ machine interference models. *International Journal of Production Research, 23*(1), 129-151

Sundkvist, R., & Sundkvist, R. (2014). Financial benefits of shop floor productivity improvements.

Troyer, D. (2009). Reliability engineering principles for the plant engineer. Available online: http://www.reliableplant.com/Read/18693/reliabilityengineering-plant. Accessed: 14 May 2015

Wang, J., Chen, J., Wang, S., Zhang, Y., & Sun, S. (2011). Shifting economic bottleneck identification. *IEEE International Conference on Industrial Engineering and Engineering Management*, 1760–1764. doi:MM209/IEEM.2011.6118218

Yang, C., & Nurtam, M. R. (2013). Proceedings of the Institute of Industrial Engineers Asian Conference 2013, 827–835. doi:M1007/978-981-4451-98-7

APPENDIX A: Losses Description and Levels of Operator Influence

Loss Description	Level of operator influecne
Fullt i konveyorbuffert	May be operator related
Materialväntan	May be operator related
Materialväntan, B2	May be operator related
Materialväntan, L5	May be operator related
Materialväntan, R3	May be operator related
Materialväntan, R4	May be operator related
Vänta jobb	May be operator related
Breakdown	Operator related
ställtid	Operator related
Reparation (EM)	Operator related
Systemfel	Operator related
Systemfel RS 3	Operator related
Tappmaskin/Rinser K4 - Tekniskt fel	Operator related
Verktygsbrist	Operator related
Verktyg - Haveri	Operator related
Artikelbyte med enkel städning	Operator related
Artikelbyte utan städning	Operator related
Artikelbyte	Operator related
Artikelbyte med helstädning	Operator related
Påspackare (PP)	Operator related
Skivbyte	Operator related
F42-Ecobloc-Fyllare-Fyllning	Operator related
Fyllning	Operator related
Påfyllning zink	Operator related
slipskivebyte	Operator related
Maskinjustering	Operator related
Material (materialbyte)	Operator related
Modellbyte	Operator related
Orderbyte	Operator related
Orderbyte med ställ	Operator related
Sortbyte/ställ	Operator related
Sortbyte med disk	Operator related
Sortbyte utan disk	Operator related
Product change	Operator related
Byte elektroder LGR	Operator related
Byte elektroder Q2RS	Operator related
Byte Film	Operator related
Byte Tråd	Operator related
Byte av coil	Operator related
Rullbyte (RB)	Operator related
Slipskivebyte	Operator related

Inställning/Justering, D1	Operator related
Inställning/Justering, D3	Operator related
Inställning/Justering, L5	Operator related
Inställning/Justering, R3	Operator related
Inställning/Justering, R4	Operator related
Inställning/Justering, S2	Operator related
Ställtid	Operator related
Byte klippdel längdtråd	Operator related
Verktygsbyte	Operator related
Operatör ej stämplat	Operator related
Sök	Operator related
Stopp för packen	Operator related
Artikel saknas	May be operator related
Kapsylbrist	May be operator related
Vagnar/Emballage saknas	May be operator related
Kartong Saknas (KS)	May be operator related
Enternt material saknas	May be operator related
Externt material saknas	Not operator related
Externt Material	Not operator related
Grafitbrist	May be operator related
Gallerbrist	May be operator related
Habiaproducerat material	May be operator related
Inkö tom	May be operator related
Internt material saknas	May be operator related
Internt material	May be operator related
Saknar insatsmaterial	May be operator related
Saknar meterial (burkar, etiketter etc) från leverantör	Not operator related
M-Brist- Gummi	Not operator related
Material brist	May be operator related
Material brist - Tork	May be operator related
Material brist - Externleverans	Not operator related
Saknar Råvara (SR)	May be operator related
M-Brist-Gummi	May be operator related
Ingen aktiv artikel	May be operator related
Ingen produkt (IP)	May be operator related
Ingen Produt (IP)	May be operator related
Order/Material brist	May be operator related
Otorkat material	May be operator related
Väntar på burkar att packa	May be operator related
Fönsterbrist	May be operator related
Fönster brist	May be operator related
Absence - Operator	Operator related
Personal saknas	Operator related
Brist på personal	Operator related

Personal - Ej Fullbemannat	Operator related
Personal brist	Operator related
Personal brist - Annnan Stn	Operator related
Personal brist - Lågprio Stn	Operator related
Personal brist - Ej Fullbemannat	Operator related
Personal brist - Annan Stn	Operator related
Personal brist (PB)	Operator related
Personal brist -Ej Fullbemmanat	Operator related
Justering efter ställ	Operator related
Justering/Inställning	Operator related
Justering klippdel längdtråd	Operator related
Justering, No Hit!, B2 Färgprov	Operator related
Justering, No Hit!, D1 1:a Körning/Colorering	Operator related
Justering, No Hit!, D1 Färgprov	Operator related
Justering, No Hit!, L5 1:a Körning/Colorering	Operator related
Justering, No Hit!, L5 Färgprov	Operator related
Justering, No Hit! L5 1:a Körning/Colorering	Operator related
Verktygsjustering	Operator related
Kontroll och Justering	Operator related
Etikettmaskin K1 - Justering	Operator related
Mätning	Operator related
Meypack K1 - Justering	Operator related
Ingen justering, Hit! Färgprov	Operator related
Prasmatic K4 - Justering	Operator related
SMI K3 - Justering	Operator related
Tappmaskin/Rinser K4 - Justering	Operator related
Tappmaskin K1 - Justering	Operator related
Mikrostopp	May be operator related
Other	Not possible to classify
Övrigt / Orsakskod saknas	Not possible to classify
Annat mekrelaterat UH	Not possible to classify
Annan orsak	Not possible to classify
Processproblem	Not possible to classify
Annan maskin prioriterad	May be operator related
Producerat antal uppnått	May be operator related
Order färdig	May be operator related
Bristande kvalitet	Not possible to classify
M-Brist - Externleverans	Not operator related
Omarbete	Operator related
Ingen produkt från linjen	May be operator related
Linjeuppstart	Operator related
Line Startup	Operator related
Line startup	Operator related
Normal uppstart/montering	Operator related

Produktions Start (PS)	Not possible to classify
produktion Start(PS)	Not possible to classify
Uppstart/Kallstart, R3	Not operator related
Uppstart/Kallstart, R4	Not operator related
Uppstart/Kallstart, S2	Not operator related
Uppstart/uppvärmning	Not operator related
Uppstart/Avslut	Not operator related
1015-M200	Not possible to classify
M200-1015	Not possible to classify
Tryckverk 1-6, D1	Operator related
Referensfel Vision	Operator related
AM16 Andra fel	May be operator related
AMS Brist på majs	May be operator related
Other mechanical maintenance	Operator related
Påspackare	Operator related
Påspackare (PP)	Operator related
Banderoll, S2	May be operator related
pga CNC Berger	Operator related
Bestrykningstork, D3	Operator related
Kartongmaskin	Operator related
Uppstart, avslut	Not possible to classify
Uppstart	Not possible to classify
BOXES	May be operator related
Brättbyte, change of something	Operator related
Byte IPS storlek	Operator related
C&C maskin	Operator related
Kabel + Verktyg	Operator related
Åkvagn stannar	Operator related
Vagnar/Emballage saknas	May be operator related
Kartongresare, D3	Operator related
Orsaks kod finns ej	Not possible to classify
CD	Not possible to classify
C-motor	Operator related
Centrumhylsa Hålsko	Operator related
Byte varumärke	Operator related
Avslut	Not possible to classify
Avslut med problem	Not possible to classify
Avslut med problem (skriv)	Not possible to classify
Klimat Tunnel - Låg Fuktighet	Operator related
Coilbyte	Operator related
Kontaktrör	Operator related
Kontroll provutrustning	Operator related
Omställning	Operator related
Transportör	May be operator related

Transportbanor	Operator related
Lucka	Operator related
Skärstans	Operator related
Klippverktyg	Operator related
Datumprinter	Operator related
DatumBandfel (DB)	Operator related
Avdrag	Not possible to classify
Detaljbrist	May be operator related
Riktning tråd 6mm	Operator related
Hämta/lämna material	May be operator related
Borr 2	Operator related
Körning efter rast	May be operator related
pga Hålslip	Operator related
EJ planerad service	Operator related
Elektriskt fel	Operator related
Elfel	Operator related
Gavelsvets	Operator related
Maskinhaveri	Operator related
Planerat UH stopp	Operator related
Etikettfel (EF)	Operator related
F42-Ecobloc-Blåsmaskin-Blåsfel	Operator related
F43 Packmaskin - Utlopp	May be operator related
F43 trågmärkare	Operator related
F43B Etikettmaskin Anker - Etikett	Operator related
F43B Etikettmaskin Anker - Utloop	May be operator related
Falskt stopp	Not possible to classify
Fel på konveyor	Operator related
Matarverk	Operator related
Film	Not possible to classify
Fixturbyte	Operator related
Planhetskontroll	Operator related
Spolning Foam	Operator related
Fluss övriga fel	Operator related
FM10 Fyllmaskin-Pappersbrott	Operator related
FM10 Pallastare	Operator related
FM2 Fyllmaskin - LS/TS	Operator related
FM2 Trågmäkare	Operator related
FM2 Trågpackare	Operator related
FM3 Pallastare	Operator related
FM3 Trågpackare	Operator related
FM4 Fyllmaskin - Backsystem	Operator related
FM4 Trågpackare	Operator related
FM5 Etikettmaskin - Utlopp	May be operator related
FM5 Pack - Bricka	Operator related

FM7 Fyllmaskin-Fyllstystem	Operator related
FM7 Pallastare	Operator related
FM8 Fyllmaskin-Backsystem	Operator related
FM8 Fyllmaskin-LS/TS	Operator related
FM8 Fyllmaskin-Pappersbrott	Operator related
FM8 Trågpackare	Operator related
FM9 Fyllmaskin-Fyllsystem	Operator related
FM9 Fyllmaskin-Pappersbrott	Operator related
FM9 Pallastare	Operator related
FM9 Trågpackare	Operator related
Folie Bromsen	Operator related
Formsäkring	Operator related
Formläggare	Operator related
FS - Temp. Aggregat	Operator related
FS -Temp.aggregat	Operator related
FUH	Not possible to classify
Fullt i konveyorbuffert	May be operator related
Fastnar i utkast	Operator related
Fastnar i modell	Operator related
Grosning	Not possible to classify
Haveri	Operator related
Höglager	May be operator related
Hydraulikbortfall	Operator related
ILA-pack	Operator related
Formspruta övrigt	Operator related
Inlastningsfel	Operator related
Inpastare, R3	Operator related
Internt material	May be operator related
Trassel	Operator related
Jomet/Metod	Not possible to classify
Justerning, No Hit!, D1 Färgprov	Operator related
Justerning, No Hit!, D3 Färgprov	Operator related
Klipportal	Operator related
Köpdetaljer	May be operator related
Kutsbrist	May be operator related
Etikettutrustning, R3	Operator related
Etikettutrustning, R4	Operator related
Lådbrist	May be operator related
Lådrobot	Operator related
Större ställ	Operator related
Laserstråle	Operator related
Sidolyft mellan décor och 2-fot cell	Operator related
Mindre ställ	Operator related
Vågen	Operator related

Linjeavslut	Operator related
Line training	Operator related
Linjeuppstar	Operator related
Fylla på material	May be operator related
Lång tid Vibrasvets	May be operator related
Tappar vacuum i låda	Operator related
Tappar vacuum i form	Operator related
LS/TS	Not possible to classify
M - Brist - Externleverans	Not operator related
Maskin	Operator related
Machine error	Operator related
Maskinfel Plastmaskin, X1	Operator related
Huvudverktyg	Operator related
Underhåll MEK/EL	Operator related
Manuell hantering	Operator related
Manuell laddning Blohm	Operator related
Material	May be operator related
Material slut	May be operator related
Material, R4	May be operator related
Mekaniskt fel	Operator related
Mekfel	Operator related
Mekfel robot	Operator related
Motorskydd skrotbana	Operator related
Multipond/Metod	Not possible to classify
Ny personal	Operator related
Ingen justering, Hit! 1:a Körning/Colorering	Operator related
Ingen produkt från linjen	May be operator related
Ej planerad service	Operator related
Normalt avslut	Not possible to classify
Ej avkodat stopp	Not possible to classify
Märkmaskin	Operator related
Okagoriserate	Not possible to classify
Okatergoriserat	Not possible to classify
Operatörs UH	Operator related
Opertörs UH	Operator related
Övrigt / Orsakskod saknas	Not possible to classify
Andra Maskinfel (AM)	Operator related
Övrigt mekfel	Operator related
Övrigt, X1	Operator related
Omställning (O)	Operator related
Utmatning	May be operator related
Ugnar 1-5, D3	Operator related
P - brist - Ej Fullbemmanat	Operator related
P-brist - Ej Fullbemmanat	Operator related

P-brist - Övrigt	Operator related
P -Brist - Annan Stn	Operator related
Palletare K3 - Justering	Operator related
Pallastare/Metod	Operator related
Palletering (PA)	Operator related
Pappersbrott, B2	Operator related
Pappersbrott, D1	Operator related
Pappersbrott, S2	Operator related
P-Brist - Ej Fullbemmanat	Operator related
Otillåtet stopp	Operator related
Personal utlånad	Operator related
Plastare	Operator related
Pneumatisk fel	Operator related
Portal	Not possible to classify
Förberedelser sen start	May be operator related
Tryckverk, 1-5, D3	Operator related
Problem med RS3	Operator related
Process/Filtrering	Operator related
Process/Mixern	Operator related
Processproblem	Operator related
Producerat antal uppnått	Operator related
Produktion Avstägning (PA)	Not possible to classify
Programering	Operator related
Punkstvets	Operator related
Köpt material	Not operator related
Kylare	Operator related
Rb - Plockfel	Operator related
Rb - Positionsfel	Operator related
Orsakskod finns ej	Not possible to classify
Orsakskod saknas	Not possible to classify
Orsakskod saknas (andra)	Not possible to classify
Orsakskod saknas Lagret K4	Not possible to classify
Upprullning, D1	Operator related
Reparation mall	Operator related
Ståll	Not possible to classify
RIBBON	Operator related
Robot	Operator related
Robot (Vanilj)	Operator related
Robot 2	Operator related
Robot 3	Operator related
Robot fell	Operator related
Robot Övrigt	Operator related
RS	Not possible to classify
Gummiblandning	Operator related

Rullbyet	Operator related
Prov	Operator related
Provtagning	Operator related
Slipskivebyte	Operator related
Scorotron, X1	Operator related
Repor	May be operator related
Screenverk 1-3 S2	Operator related
Screenverk, D1	Operator related
Givarfel O -ringsränna	Operator related
Givarfel EL	Operator related
Givare	Operator related
Service	Operator related
Ställ	Operator related
Inställning/Justering, R3	Operator related
Plåthantering	Operator related
Hylla fastnar i magasin (Maskinfel)	Operator related
Hylla fastnar i magasin (Skevhet/Mått/Svetsfel)	Not possible to classify
Omställning/sort byte	Operator related
Liten omställning	Operator related
SMI K3 - Justering	Operator related
Mjukvarufel	Operator related
Stacker	Operator related
Staplingsutrustning	Operator related
Personal	Operator related
Personal - Lågprio Stn	Operator related
Trappa	Operator related
Prägling, D3	Operator related
Start och stopp	Not possible to classify
Uppstart/Kallstart, B2	Not possible to classify
Uppstart/Kallstart, D3	Not possible to classify
Uppstart/Kallstart, R3	Not possible to classify
Uppstart/Kallstart, S2	Not possible to classify
Uppstart/uppvärmning	Not operator related
Uppstart, X1	Not possible to classify
Station 1	May be operator related
Stegtransportör	Operator related
Stopp vid nerplock	Operator related
Stopp vid påhängning	May be operator related
Stopp pga annan msk	Operator related
Stoppkod saknas	Not operator related
Stopp i Décor 2 -fot cell	Operator related
Stopp i Décor cell	Operator related
Stopp måleri	Operator related
Stopp föregående skift	May be operator related

Stopp för packen	Operator related
Syncropack	Operator related
Systemfel	Operator related
Systemfel RS 3	Operator related
Tappmaskin/Rinser K4 - Justering	Operator related
Tappmaskin/Rinser K4 - Tekniskt fel	Operator related
Tappmaskin K1 - Tekniskt fel	Operator related
Tekniskt/Maskin	Operator related
Teknik	Operator related
TEFLON - BOTTOM BELT	Operator related
Tempereringsagregat - FS	Operator related
Limmet	Not possible to classify
Fram etikett	Not possible to classify
Normal omställning	Operator related
Order färdig	Operator related
Sen start före pack	Not possible to classify
Trepack (TP)	Not possible to classify
Tid utanför ordinare uppgift	Not possible to classify
Verktygsfel	Operator related
Verktyg	Operator related
Verktyg - Avformningsfel	Operator related
Verktyg - Balansering Dysa	Operator related
Verktyg - Haveri	Operator related
Verktyg - FS	Operator related
Verktyg - Pipbrott	Operator related
Verktygstvätt	Operator related
Transfer	Operator related
Transfervagn	Operator related
Transport Band (T)	Operator related
Trågpackare	Operator related
Vagn/pallbyte	Operator related
Felsökning	Operator related
Truck körning	Not possible to classify
Sortbyte	Operator related
Ugnsproblem	Operator related
Okategoriserat	Not possible to classify
Unerhåll MEK/EL	Operator related
Avrullning, R3	Operator related
Uppstaplare 2/4	Operator related
Verktyg + kabel + Färg	Operator related
Vaccum	Operator related
Vacuum - Rb	Operator related
Vacuumpump	Operator related
Vacuumsläpp	Operator related

Vacuumfel Vibrasvets	Operator related
Verkgsbyte	Operator related
Verktysbyte	Operator related
Verkygsbyte	Operator related
Vinda	Operator related
Vision	Operator related
Väntar på P4	May be operator related
WA-maskin 10-pack	Operator related
Svets Q2RS	Operator related
Hjul fastnar i transportör	Operator related
Rullmaskin, L5	Operator related
Rullmaskin, R3	Operator related
Rullmaskin, R4	Operator related
Rullmaskin, S2	Operator related
Wolfen (W)	Operator related
Inplastare kaka	Operator related
Inplastare, D3	Operator related
Inplastare, R3	Operator related
Inplastare, R4	Operator related
Inplastare, S2	Operator related
Fel ihopträdning	Operator related
Uncategorised	Not possible to classify

APPENDIX B: Layout of AAA and BBB Line