ANALYZING PRODUCTION EFFICIENCY IN SEMICONDUCTOR MANUFACTURING: A SIMULATION APPROACH USING DOWNTIME DATA

CHAI YU QIAN

UNIVERSITI TUNKU ABDUL RAHMAN

ANALYZING PRODUCTION EFFICIENCY IN SEMICONDUCTOR MANUFACTURING: A SIMULATION APPROACH USING DOWNTIME DATA

CHAI YU QIAN

**A project report submitted in partial fulfilment
of the requirements for the award of Bachelor of Engineering
(Honours) Industrial Engineering**

**Faculty of Engineering and Green Technology
Universiti Tunku Abdul Rahman**

May 2026

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature : _________________________

Name : Chai Yu Qian

ID No. : 21AGB03076

Date : 15 May 2026

APPROVAL FOR SUBMISSION

I certify that this project report entitled "**ANALYZING PRODUCTION EFFICIENCY IN SEMICONDUCTOR MANUFACTURING: A SIMULATION APPROACH USING DOWNTIME DATA"** was prepared by **CHAI YU QIAN** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Engineering (Honours) Industrial Engineering at Universiti Tunku Abdul Rahman.

Approved by,

Signature : ______________________________

Supervisor : Ts. Dr. Cheah Chew Keat

Date : 15 May 2026

Signature : ______________________________

Co-Supervisor: Ir. Ts. Dr. Joshua A/L Jaya Prakash

Date : 15 May 2026

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to everyone who has contributed to the successful completion of this project.

First and foremost, I would like to thank my supervisor, Ts. Dr. Cheah, for his consistent guidance, invaluable supervision and continuous encouragement throughout the entire research process. His expert knowledge and constructive feedback were instrumental in shaping the direction and quality of this study. I am deeply grateful for his patience, dedication, and constant support in guiding me through every challenge encountered during the completion of this project.

I would also like to extend my heartfelt appreciation to my co-supervisor, Ir. Ts. Dr. Joshua, for his support and technical guidance, particularly in helping me understand and apply the WITNESS simulation software. His assistance was fundamental to the successful development of the simulation model in this project.

I would also like to express my gratitude to the engineers and staff of Company X for their cooperation in providing the necessary production data and technical knowledge that made this research possible. Their willingness to share their expertise greatly contributed to the quality and relevance of this study.

Finally, I am deeply grateful to my parents and friends for their unwavering encouragement and moral support throughout this journey. Their belief in me and constant motivation have been a tremendous source of strength in the completion of this project.

ANALYZING PRODUCTION EFFICIENCY IN SEMICONDUCTOR MANUFACTURING: A SIMULATION APPROACH USING DOWNTIME DATA

ABSTRACT

This study investigates production performance at an automated die testing line, known as the Multi-die Pick and Place Inspection (MPI) line, at Company X in Penang, Malaysia. Real downtime data from the MPI tracking system was incorporated into a Discrete Event Simulation (DES) model developed using WITNESS software, adopting a data-driven approach to overcome limitations of studies relying on assumed downtime values. The downtime data was analyzed and categorized into major error types to identify key factors affecting production efficiency. Overall Equipment Effectiveness (OEE), comprising Availability (A), Performance (P), and Quality (Q), were used as the primary performance metric. Results showed that Performance and Quality remained stable, while Availability was reduced due to frequent downtime, confirming it as the main limiting factor of system performance. A WITNESS-based simulation model representing 10 MPI machines was developed and validated against actual production data, achieving an average deviation of 0.02%, indicating high reliability. The model was then used to analyze operational scenarios, including throughput and revenue forecasting, product mix optimization, and performance improvement strategies. Results demonstrated a strong relationship between OEE, throughput, and revenue, highlighting availability improvement as the most effective strategy. Additionally, AI-based tools were developed to visualize results and support data-driven decisions. This study shows that integrating DES, OEE, and AI tools enhances production efficiency in semiconductor manufacturing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

A	Availability
AI	Artificial Intelligence
CT	Cycle Time
DES	Discrete Event Simulation
KPI	Key Performance Indicator
MPI	Multi-die Pick and place Inspection
MTBF	Mean Time Between Failures
MTTR	Mean Time to Repair
ML	Machine Learning
OEE	Overall Equipment Effectiveness
P	Performance
Q	Quality
SMED	Single-Minute Exchange of Dies
t	Time
T	Total production time
USD	United States Dollar
WIP	Work-In-Process
WW	Work Week

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter provides the introduction for the study on analyzing and estimating production efficiency in semiconductor manufacturing using downtime data through a simulation approach. It begins with Section 1.2, which presents the background of the study, followed by Section 1.3 that outlines the problem statement. Section 1.4 highlights the research objectives, and Section 1.5 defines the scope of the study. Finally, Section 1.6 presents the overall structure of the thesis.

1.2 Background

The semiconductor industry plays a critical role in modern technological advancement. Technology has further facilitated products like computers, mobiles, automobiles, medical equipment, communication etc. Semiconductor manufacturing requires high production cost. It goes through multiple processes including wafer preparation, photolithography, etching, doping, testing and packaging processes. In order to maintain quality and yield, all these processes should be carried out consistently and accurately.

In a competitive environment, manufacturers need to improve the efficiency of various processes which enhance costs, product quality and reliability. According to Semiconductor Industry Association (SIA), equipment downtime is one of the major factors affecting production efficiency. When the equipment is not functioning due to failure, scheduled maintenance, calibration, set-up or process adjustment, we will refer it to downtime. For semiconductor manufacturers, since many manufacturing operations are interdependent, the downtime of one manufacturing operation will affect the product flow of the entire facility. This will lead to extended cycle times, increased work-in-progress (WIP) inventory, and challenges in delivery times (Jeon & Lee, 2017; Negahban & Smith, 2014).

The sophisticated photolithography, etching, and other equipment used inside semiconductor plants are very expensive. A little disruption can lead to huge production loss as well as a high operating cost (Beaverstock et al., 2017). One common technique for measuring equipment performance is Overall Equipment Effectiveness (OEE). According to Díaz-Reza et al. (2019), OEE is calculated based on three components: Availability (A), Performance Efficiency (P) and Quality (Q), as shown in Equation 1.1. Six big losses are linked to the three components.

$$OEE = A \times P \times Q \qquad (1.1)$$

where
OEE = Overall Equipment Effectiveness
A = Availability
P = Performance Efficiency
Q = Quality Rate

WITNESS is one of the simulation tools commonly used in manufacturing research and process improvement, as it can effectively represent detailed process flows, machine interactions and resource allocation (Michlowicz & Smolinska, 2018). Key steps in simulation modeling, such as problem description, data collection, model development and verification are important as they help ensure reliable and practical outcomes (Skoogh et al., 2017). As a result, those responsible for the company production decisions can obtain quantitative information about potential yield and revenue improvements without modifying the actual production areas.

Shou, Wang, and Wu (2021) emphasize that simulation plays a vital role in lean manufacturing by enabling engineers to test scenarios and optimize system performance without disrupting real operations. This underlines the main advantage of simulation as a useful tool for analyzing and improving complex production systems efficiently and cost-effectively. Companies can use WITNESS simulation models to evaluate improvement methods under both predictive and realistic operating conditions. These may include modifying process flows, adjust downtime and resource allocation or optimizing performance and maintenance schedules. In a virtual environment, companies can assess the effectiveness of these strategies before implementation of any changes. WITNESS also enables the incorporation of different variables into simulation models, ensuring that the simulated system reflects actual operations as accurately as possible. As a result, strategies such as optimizing preventive maintenance plans, reallocating workforce, balancing WIP levels, or redesigning processes, waste reduction, and performance improvement can be simulated without additional time and cost for the changes.

1.3 Problem Statement

The semiconductor manufacturing industry requires a significant amount of capital and time. Any equipment downtime due to planned or unplanned maintenance will result in losses in production, yield, and operational efficiency. In a fully automated environment, machine downtime reduces OEE and production output. In a connected production environment, if multiple machines are down, completing all operations may take longer, resulting in more WIP and ultimately causing delays in customer shipments.

The machine downtime in the semiconductor industry is associated with increased operational costs and reduced efficiency (Phongmoo et al., 2025). Machine downtime can have significant financial consequences, with studies indicating that in some cases, the combined planned and unplanned downtime can exceed 10% of total operating costs (Dunscombe et al., 2000).

One of the most popular methods for simulating semiconductor manufacturing systems is through discrete event simulation (DES). Most previous studies, in order to simplify machine reliability, did not use actual machine downtime history but instead made some assumptions. This led to inaccurate DES outputs that could not truly reflect the current state of the system, making it difficult to validate with production data. Therefore, they are still ineffective in providing information about the system production and design decisions.

Therefore, it is necessary to create a data-driven simulation method that utilizes real-time downtime data to provide a more accurate representation of actual manufacturing conditions. Integrating empirical production data into DES models will allow for a more accurate assessment of the impact of machine downtime on system performance. This allows for a more reliable evaluation of improvement strategies, such as reducing downtime, optimizing maintenance, and production scheduling, without disrupting actual production operations.

1.4 Objectives

The main goal of this project is to develop a reliable model to simulate equipment downtime and analyze the results and impacts so that improvements can be carried out to address the issues. Downtime data from a semiconductor company, Company X in Penang, was provided and used for data analysis. The analyzed data is then used to develop a WITNESS-based simulation model. This model will support detailed analysis of downtime errors and provide a virtual environment for simulating and evaluating various scenarios without disrupting actual production.

In line with this goal, the study will focus on the following objectives:

1. To analyze and categorize equipment downtime data in a semiconductor manufacturing line.
2. To develop a WITNESS-based simulation model to evaluate downtime reduction scenarios.
3. To validate the simulation model by comparing its output with actual equipment performance using Overall Equipment Effectiveness (OEE) as the benchmark.
4. To use the simulation model to evaluate improvements, particularly in terms of throughput and revenue, under different product testing environments to provide a basis for the development of useful AI tools and decision-making.

1.5 Project Scope

The project scope is limited to a specified production line involving MPI machines and is subject to the availability of the data. It does not extend to operations with other types of machines or to data not included in the data set provided by Company X. An MPI machine is a fully automatic LED die prober equipped with a vision system, a pick-and-place module and a pin-ejection module. Its functions include testing LED wafer dice, recording test data, and sorting the dice based on the test results. The project scope includes data analysis, model development and model validation where simulated outcomes are compared with actual conditions or improvements observed in the actual manufacturing lines. This ensures that the model generates reliable results for analysis and decision making. Finally, the validated model is used to simulate different production and performance scenarios in the MPI line, and the results are used to evaluate potential throughput and revenue gains.

1.6 Outline of the Thesis

This thesis is divided into five chapters. Chapter 1 introduces the study by presenting the background, problem statement, objectives, and scope of the project. Chapter 2 reviews related literature, where past research and journal articles are studied to understand existing approaches in semiconductor manufacturing and simulation modelling. Chapter 3 describes the research methodology, including the analysis of data collected from the factory and the process of developing the simulation model using Witness. Chapter 4 presents the results and discussion, which include the findings from the simulation and the evaluation of the model performance. Finally, Chapter 5 concludes the thesis by summarizing the main outcomes of the study and providing recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents the literature review related to this study. The review mainly focuses on equipment downtime in semiconductor manufacturing and the application of simulation modelling to analyze and improve production performance. The review is organized into four main subsections. Subsection 2.2 examines production efficiency in semiconductor manufacturing, with emphasis on performance indicators such as cycle time, throughput, and OEE. Subsection 2.3 discusses downtime in semiconductor fabs, highlighting its types, causes, measurement metrics, and implications for operational and financial performance. Subsection 2.4 explores simulation-based approaches for analyzing and improving efficiency under downtime conditions. This includes applications of DES, modeling approaches for downtime, and the role of specific tools such as WITNESS and other simulation software, and the integration of AI tools to enhance simulation analysis and decision-making. Lastly, Subsection 2.5 presents the discussion and insights drawn from the literature review, highlighting research gaps that this study aims to address.

2.2 Production Efficiency in Semiconductor Manufacturing

In semiconductor manufacturing, production efficiency has a significant impact on cost, competitiveness, and productivity. Semiconductor manufacturing is very complicated and expensive because of steps like photolithography, etching and deposition which need special equipment that must work in very clean rooms. Due to the complexity of semiconductor manufacturing, many manufacturers rely on using key performance indicators (KPIs), such as cycle time, throughput, and OEE, to measure the performance of their equipment and identify opportunities for improvement.

Cycle time is defined as the time required for a wafer to complete all production processes. Cycle time can be affected by process variations as well as equipment failures. Due to small delays in the operation of bottleneck equipment such as photolithography machines, cycle time may be extended, leading to a decrease in yield and an increase in defect rates. Cycle time and OEE jointly affect the throughput of semiconductor manufacturing systems. Throughput is defined as the number of wafers produced within a specific unit of time and is an indicator of overall production efficiency. Previous studies have shown that improvements in scheduling, predictive maintenance, and real-time monitoring can increase yield by 10% to 15% (Chong and Ng, 2016). However, semiconductor manufacturing systems continue to face challenges such as equipment failures, process variability, labor shortages, and supply chain disruptions, all of which contribute to inefficiencies and yield losses (Musa et al., 2025).

OEE is an essential metric used by many manufacturers to achieve one of the important lean manufacturing strategies, which is production with zero waste and breakdown. The OEE measure was originated from the Japanese framework of Total Productive Maintenance (TPM). It is one of the five pillars within TPM and is used to measure TPM activities. OEE is one of the important indicators of the success of TPM programme. Ahuja and Khamba (2008) suggest that OEE is a measure that supports the strategic outcome of a TPM implementation through the various metrics of manufacturing to reduce waste, which is also the main objective of lean manufacturing.

OEE was first introduced by Seiichi Nakajima in 1988 with the main objectives of measuring the productivity of individual production equipment and eliminating the 'six big losses'. OEE exposes the deficiency of equipment and acts as an improvement driver to generate continuous improvements. As noted by Dal *et al.* (2000), OEE is being used to track and trace improvements or decline in equipment effectiveness. It is also a managerial tool for unleashing hidden capacity, for reducing production lost times as well as for extending a major capital investment (Muchiri and Pintelon, 2008). The key goals are minimizing production losses, improving productivity and maximizing profits. Muthiah *et al.* (2008) reports that the OEE metric is a powerful tool that can be used to measure performance and perform diagnostics at the equipment level. OEE is also employed as an indicator of the process improvement activities within a manufacturing environment (Dal *et al.*, 2000). Fleischer *et al.* (2006) state that the competitiveness of manufacturing companies depends on the availability and productivity of their production facilities. The intensifying competitions and cost pressures have led to an increasing number of manufacturers using OEE, especially in large-scale automation productions. Huang *et al.* (2003) states that OEE has become increasingly popular and has been used as a quantitative tool essential for measurement of productivity in semiconductor manufacturing operations, due to an extreme capacity constrained manufacturing environment and an increasing concern about return on capital facility investment.

In considering OEE, Nakajima (1988) defines 6 big losses in production that a company must strive to eliminate, they are (1) Breakdown losses, which are the time losses due to equipment failure or breakdown. (2) Setup and adjustment losses, which are the time losses due to setup or changeover. The losses occur when production of one item ends and setup and adjustment are made to meet the requirements of the next item. (3) Idling and minor stoppage losses, which are the time losses due to equipment is idling or interruption of production caused by minor equipment malfunction. (4) Reduced speed losses, which are the losses due to the difference between equipment design speed and actual operating speed. (5) Defects and rework, which are the losses in product quality caused by malfunctioning equipment. (6) Reduced yield during start-up, which are the yield losses that occur from equipment start-up to stabilization.

Breakdown losses and setup and adjustment losses are categorized as the down time losses. Idling and minor stoppage losses and reduced speed losses are categorized as speed losses. Lastly defects and reduced yield during start-ups are categorized as quality losses. By identifying and reducing the six big losses, the OEE will improve. OEE is measured in these six big losses that can be further expressed as a function of availability, performance efficiency and quality rate as illustrated in Figure 2.1.

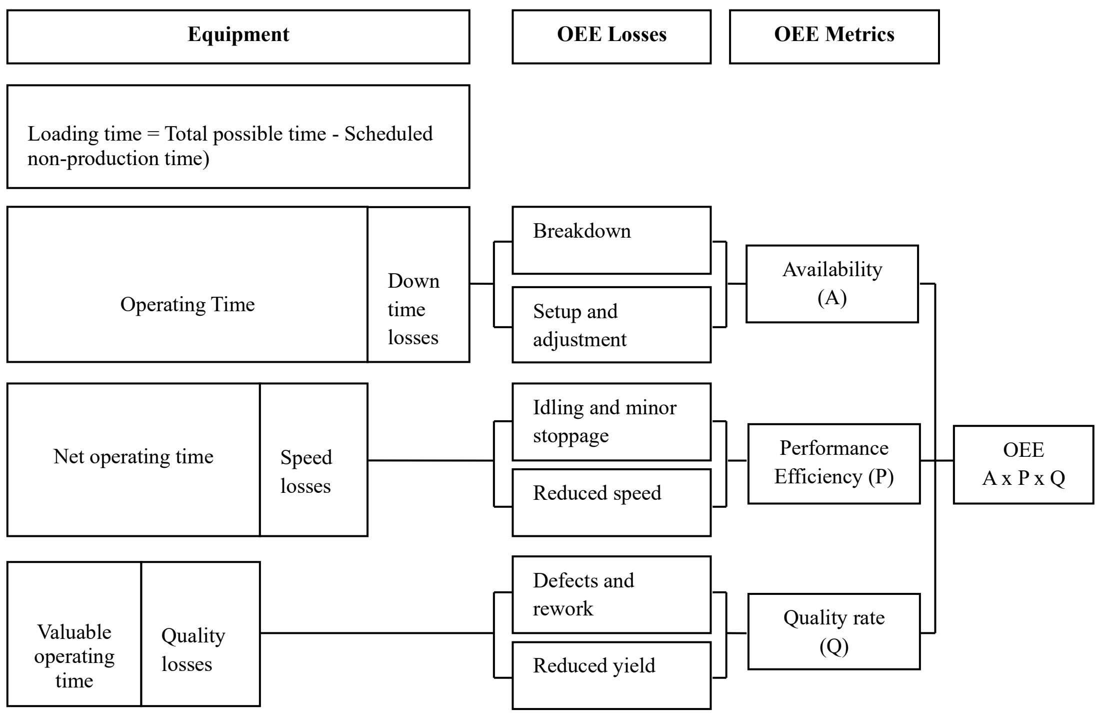


Figure 2.1: OEE and the six big losses (Nakajima, 1988)

OEE is composed of three key components: Availability (A), Performance (P), and Quality (Q) (Nakajima, 1988). OEE is calculated by multiplying A, P, and Q. Availability is the actual production time of the equipment based on the planned production schedule, minus the time lost due to breakdowns and setups. Performance efficiency rate is the comparison between the actual operating speed of the equipment and the maximum speed of the equipment based on the planned production, minus the downtime caused by speed reductions and minor stoppages. Quality rate is the ratio of the number of qualified products that meet production specifications to the total number of products produced. Therefore, the quality rate provides an indication of the effectiveness of the production process in producing defect-free products.

Losses are considered as activities that absorb resources but create no added value (Bamber *et al.* 2003). Production losses are caused by manufacturing disturbances, which in turn cause low equipment utilization. Sporadic disturbances occur quickly and are more obvious since they deviate significantly from the normal state. Chronic disturbances are small, hidden and are the result of several concurrent causes. They are difficult to detect as they happen in the normal state. Due to the obvious irregular effects, sporadic disturbances are easier to detect than the chronic disturbances that can only be identified through comparison of actual and theoretical equipment capacity. These losses not only reduce productivity but also incur additional costs to production. In this regard, OEE can be used to identify and eliminate these losses.

Furthermore, OEE is a performance measurement metric, but it also provides an effective support tool for manufacturers' strategic continuous improvement efforts. Manufacturers can use OEE to monitor equipment performance, identify areas causing losses, and reveal potential capacity in their production processes. (Muchiri & Pintelon 2008). Therefore, OEE is beneficial in the decision-making process as it supports efforts to reduce equipment downtime, improve operational efficiency, and increase output. Research data shows that the OEE levels in semiconductor manufacturing operations average between 60% and 80%, which leaves very limited opportunities for further improving current performance (Braglia, Frosolini, and Zammori, 2008).

In summary, achieving high production efficiency in semiconductor manufacturing requires a comprehensive understanding of equipment performance, production losses, and process variability. Key performance indicators, such as OEE, cycle time, and yield, can provide an in-depth analysis of system performance, while the use of predictive maintenance, process improvement, and automation is crucial for minimizing equipment downtime and maximizing overall production capacity.

2.3 Downtime in Semiconductor Manufacturing

Any period during which production equipment is unable to perform normal manufacturing operations is often referred to as a downtime. In semiconductor manufacturing, downtime is primarily categorized into planned and unplanned downtime. It is important to note that in OEE calculation, unplanned downtime is not included. Planned downtime encompassed scheduled tasks such as engineering runs, equipment calibration, and preventive maintenance. Unplanned downtime usually stems from unexpected events such as equipment failures, contamination incidents, wafer handling errors, software or control issues and shortages of consumables (Pecchia et al., 2023). In semiconductor factories, unplanned downtime was particularly disruptive, although both types reduced available capacity. Given the high capital costs and repetitive nature of semiconductor manufacturing, even brief stoppages on production can cause delays that may propagate across multiple process stages. These disruptions often destabilized production operations and resulted in costly scraps or rework (Mourtzis, Panopoulos, and Fotia, 2021).

In manufacturing, metrics are often used to measure downtime. Among the most common ones are MTTR, which measures the average duration required to repair equipment after a failure. MTBF measures the average operating time between consecutive failures. Failure frequency is defined as the number of downtime events per unit operating time, and availability loss in OEE is expressed as the percentage of downtime relative to planned equipment availability. Both are also important metrics. Empirical research indicated that availability loss in front-end semiconductor fab led to a decline in OEE exceeding 30% (Mourtzis, Panopoulos, and Fotia, 2021). To address this issue, many factories employed Pareto analysis to identify and tackle the few critical failure modes responsible for most of the downtime, enabling more efficient utilization of maintenance resources (Wang, Zhang, and Li, 2023).

For semiconductor factories, downtime has significant operational and financial implications. The opportunity cost of idle time is substantial, given the high specialization and cost of equipment. Research indicates that each hour of unplanned downtime for high-throughput equipment can result in revenue losses reaching six figures (Ndokwu, 2025). Case studies further validate these financial losses. For instance, 10% of unplanned downtime at Intel's wafer fab due to equipment failures resulted in substantial scrap and revenue losses. The fault detection system reduced unplanned downtime by 66%, saying an estimated value of up to $16 million per facility (Colby et al., 2023).

The operational consequences of downtime usually go beyond delayed production. Unplanned downtimes affect cycle times, increase work-in-process (WIP) and may cause missed delivery schedules. In some cases, these may contribute to yield loss. Simulation-based research showed that ignoring downtime in fab models could lead to an underestimation of cycle time by more than 20%, emphasizing the importance of incorporating downtime into production analysis (Dequeant et al., 2017). To reduce losses, most manufacturing companies implement root-cause analysis methods such as fault tree analysis, the 5-Whys and Pareto chart to facilitate waste prioritization and reduction. Research indicated that this hybrid strategy significantly improved equipment reliability and reduced the frequency and duration of unplanned downtime (Amato et al., 2023). This highlights that effective downtime reduction is essential for sustaining efficiency and competitiveness in semiconductor manufacturing.

2.4 Simulation-Based Analysis of Downtime and Efficiency in Semiconductor Manufacturing

Semiconductor manufacturing is a highly complex and capital-intensive process. As such, equipment downtime has a severe impact on its productivity and revenues. Even a small interruption in a process can spread through the system, creating bottlenecks and heavy financial losses (Kalir, 2018). DES has become an important tool for analyzing such complex manufacturing systems. DES enables scenario testing to maximize efficiency without disrupting real-world production operations. DES can simulate processes, machine interactions and random events such as failures. This provides insights into the performance of manufacturing systems. In addition, this can save time and costs while facilitating decision-making and improvement actions.

2.4.1 Applications in Semiconductor Manufacturing

Simulation models can be used to effectively address main issues in semiconductor manufacturing, such as supply chain disruptions, bottlenecks, poor scheduling and energy consumption. This is usually done by simulating production processes and evaluating optimization strategies using the models. Subsequent case studies below will show examples of specific manufacturing issues, related simulation-based solutions and the outcomes.

The 200mm wafer manufacturing facility of Infineon Technologies faced severe bottlenecks, resulting in constrained output and extended cycle times by approximately 20% due to variations in batch sizes and limited tool availability. Domaschke, Grewal, and Rose (1998) developed a DES model to evaluate factory scheduling rules to optimize equipment utilization and reduce queueing times. By prioritizing high-yield batches and testing multiple scenarios, they successfully used the simulation to remove bottlenecks and increased throughput. After implementing the optimal scheduling rules derived from simulation, Infineon increased weekly throughput from 1,200 wafers to 1,320 wafers, a 10% gain while reducing average batch cycle time from 2,500 hours to 2,125 hours, i.e. a 15% decrease. The solution avoided additional capital expenditures, and it validated the cost advantages of simulation technology.

GlobalFoundries that had several large-scale 300mm wafer factories, struggled with high-energy usage, a significant cost driver compounded by idle tools during downtime. Hu, Chou, and Chen (2022) applied a DES simulation to improve energy-saving tool scheduling, with the objective of minimizing energy losses from tools in standby mode. The simulation tested multiple scheduling schemes on varying tool power states according to production requirements and downtime patterns. The findings indicated that monthly energy usage was reduced by 25%, from 1.2 GWh to 0.9 GWh, while throughput was maintained at 30,000 wafers a month. The study showed that simulation is a powerful means of addressing sustainability and efficiency concerns. However, the study did not consider downtime variations in real time that could potentially further improve energy savings.

The Taiwan Semiconductor Manufacturing Company (TSMC) had experienced losses in efficiency due to supply chain disruptions arising from global semiconductor shortages. This reduced its on-time delivery reliability to 75% and increased cycle time variability. Zeidler et al. (2022) used a simulation of Discrete Event System Specification (DEVS) with model predictive control to model fab supply chain dynamics. The simulation evaluated different scenarios of material shortages to optimize inventory management and production scheduling. TSMC improved delivery reliability by 20%, achieved 90% on-time deliveries and reduced cycle time variability by 12% by applying simulation-suggested strategies. It is observed that although the study focused on external sources of disruption like material shortages, it did not consider internal sources of disruption like equipment downtime. As such, the project fell short of fully applying its solutions to broader areas.

Samsung Electronics faced inefficiencies in its wafer test facility, where scheduling variability and setup time had reduced throughput by about 15%. Yoon and Chae (2019) employed a DES model to conduct simulations on scheduling schemes, while considering the debugging time and equipment failures. As a result, throughput increased from 10,000 wafers per week to 11,500 wafers. This represented 15% growth and 10% cycle time reduction. This was achieved by deriving an optimal scheduling strategy that minimized downtime and prioritized high-value batches. By modeling real-time disturbance factors, the study showed that optimization through dynamic data integration using simulation was feasible.

2.4.2 Modeling Downtime in Simulations

Downtime, including planned maintenance and unplanned equipment failure, is a major efficiency barrier to semiconductor manufacturing that often introduces 20 to 50% losses of OEE (Leachman et al., 2007). The subsequent case studies describe how simulations tackle semiconductor fab downtime, with specific challenge, solution, and result details.

At a Samsung memory chip manufacturing facility, production schedules were impacted and OEE dropped from 85% to 65% at due to significant unplanned downtime caused by equipment failures. Ramirez-Hernandez et al. (2008) applied a DES simulation to improve preventive maintenance (PM) scheduling by modeling failure patterns from historic downtime records. The simulation varied several PM intervals to find a balance between maintenance frequency and production uptime. An optimal schedule that reduced interruptions was created. By adopting simulation recommendations, Samsung improved OEE by 12%, from 65% to 73%. It achieved monthly downtime saving of 15%, down from 200 hours to 170 hours. The simulation enhanced production stability with little disruptions to the production. However, the model was developed based on static historical data, hence it could not adapt to real-time failure modes such as those triggered by equipment aging or operational errors.

Intel Corporation suffered from downtime issues in their photolithography process because of frequent reticle setups with 10% added cycle times and slower wafer processing. Ehm et al. (2015) applied a DES model to simulate reticle management policy scenarios to reduce setup frequency and improve reticle availability. A simulation of historic setup data modeled downtime activities to reduce reticle changes. They managed to increase throughput by 8%, from 25,000 to 27,000 wafers per month and reduced downtime by 15%, from 50 hours to 42.5 hours. It worth pointing out that the static data set used in the study did not include dynamic factors that could further enhance efficiency through integration.

Micron Technology's DRAM manufacturing plant experienced unpredictable equipment failures, resulting in a 30% loss in overall equipment effectiveness and reduced output. To create a digital twin simulation model which leverages time-series data from equipment sensors to predict downtime. Heger and Voss (2022) integrated digital twin simulation (DES) with machine learning technologies. The simulation validated predictive maintenance techniques, determining failure patterns to avoid disruptions ahead of time. Using the simulation's optimized schedule for maintenance, Micron minimized unplanned downtime by 18%, from 150 to 123 hours a month, and yield was increased near 10%, from 90% to 99% defect-free wafers. It was noted that it was challenging to validate the ML model with real-time data, citing a need for further robust integration of data to validate accuracy.

In a typical generic semiconductor wafer fab, high WIP levels induced by tool down-time widened cycle time variation by 20%. Wang et al. (2023) applied DES and machine learning to simulate WIP control, modeling down-time caused by tool failure and maintenance. The simulation subjected maintenance plans and WIP limits to simulation and was able to identify ways to stabilize production flow. By adopting the recommendations of the simulation, the fab brought down cycle time variation by 18%, from 20% to 2% deviation, and increased throughput by 10%, from 20,000 to 22,000 wafers per month. The aggregated-data basis of the model restricted its ability to reflect real-time variation, implying a role for IoT-based integration of data.

2.4.3 WITNESS in Semiconductor Manufacturing

WITNESS is a DES program which has the reputation of having a nice user interface and strong stochastic modeling capabilities and is suitable for modeling process downtime and machine use with the semiconductor industry. Its flexibility for representing complex systems and evaluating maintenance strategies is consistent with optimizing efficiencies by analysis of downtime data, like this thesis is attempting to do. The following are case studies to illustrate the use of WITNESS in semiconductors and related manufacturing environments.

Within the semiconductor fabrication process of electrodeposition, frequent equipment failures have caused severe interruptions to operation, thus increasing production costs and delaying schedules by approximately 20%. Lee and Kim (2020) utilized WITNESS to build a simulation model of the electrodeposition process that was fed sensor-based downtime figures to identify failure patterns. The simulation tested predictive maintenance methods, analyzing varied schedules of maintenance that sought to minimize disruptions. Through implementation of the optimized schedule of maintenance identified through WITNESS, the plant realized a monthly reduction of 25% of downtime, reducing from 80 to 60 hours, while at the same time extending component lifespan by 30%, from 1,000 to 1,300 cycles. WITNESS's user-friendly visualization functions enabled detailed failure scenarios to be analyzed; however, the focus of the research on a singular sub-process reflects potential for broader extensions of applications to fabrication plants.

In a surface-mount technology (SMT) production line that was relevant to semiconductor manufacturing, frequent machine failures decreased overall equipment effectiveness (OEE) by 20%, hence affecting throughput. Tan et al. (2019) used WITNESS to model the corrective maintenance strategies by simulating repair-related failure events and repair activities. The simulation tested several repair schedules to arrive at the best option that was to react faster to critical failures. It increased machine availability from 80% to 92%, an additional 15% over the previous condition and reduced the downtime from 120 months to 96 hours or a reduction of about 20% overall. WITNESS's features, such as modeling stochastic events and visualization of the results, supported more sighted decision-making. However, its use was limited to SMT rather than semiconductor fab application.

In a SMT production line, the consequence of low system availability caused by planned and unplanned maintenance events resulted in only a 32% availability rate, ultimately disrupting the production schedules. In the study by Lee, Park, and Shin (2023), WITNESS has been utilized for simulation for maintained events even for preventive and corrective maintenance; therefore, modeled downtime events and assessed different maintenance periods. The simulation outcomes led to the identify an optimized preventive maintenance approach that considered both uptime and maintenance costs which led to an increase in availability from 32% to 61% and decrease our monthly downtime from 150 to 117 hours or 22%. The advantage of using WITNESS is the option to schedule and simulate stochastic, or unexpected events for downtime. In addition, the limited area of WITNESS in application for full semiconductor fabs provides an update valuable opportunity for evaluation in this thesis.

2.4.4 Other Simulation Software in Semiconductor Manufacturing

Other discrete-event simulation software such as Arena, Simio, and FlexSim, are utilized within the semiconductor manufacturing domain because of their optimization algorithms, generation of scenarios, ability to scale with the fab, and generally built with production data. These simulation tools generally not only complement WITNESS, but each provides a strong solution and performance modeling support specific to big fab production and scheduling but are separated from each other by the system complexity and ultimately diverse applications. A few case studies below illustrate the usage of Arena, Simio, and FlexSim in the semiconductor world.

At NXP Semiconductors, their mixed-signal chip fabrication facility was experiencing production scheduling problems that were caused by unwanted tool downtimes. The unwanted downtimes were causing weekly throughputs to drop by 12%. To improve the production scheduling, Mönch et al. (2018) turned to Arena to produce a DES model for dynamic scheduling aimed to optimize the allocation of tools to minimize the setup times. The simulation functioned from a few scheduling policies that prioritized high throughput lots to create easier production flow streams. In the end, NXP was able to implement their optimized schedule and increase its weekly throughput from 10,000 wafers to 11,500 wafers, achieving an increase of 15%, while reducing the monthly downtime from 100 hours to 90 hours, representing a decrease of 10%. Arena's ability to integrate the optimization algorithms contributed to an improved quality of scenario analysis, but it is important to note that Arena required complicated preprocessing of data before being able to convert simulations to real-time models.

Next, in a typical 300mm semiconductor fab, delays in cycle time from production bottlenecks added 18% of cycle time due to tool downtime and impacted wafer processing time. Fowler and Rose (2004) used Simio to simulate their lot release policies to demonstrate how a particular simulation can model the flow of production. Their simulation experimentation optimized a lot, sequencing production issues such as downtime which contributed to the bottlenecks. Based upon the recommendations from the simulation model, the 300mm fab was able to save 14% in cycle time from 3,000 hours per lot to 2,580 hours per lot, and improve throughput per month by 12%, from 28,000 wafers to 31,360 wafers. Simio's ability to model large models was a benefit, but the ability to perform in a timely manner with substantial computer processing and performance emphasized the need to handling simulation data that is processed and quickly returned.

A semiconductor manufacturer, STMicroelectronics, experienced a 10% throughput reduction in its wafer fabrication planning system due to inefficiencies resulting from unproductive usage of the tools in the system. Dabbas and Chen (2001) used FlexSim to simulate the production scheduling process through the modeling of tool use and setup times. The simulation studied the scheduling options to optimize tool loading and use, minimize idle time and isolated tests, and found favorable simulation results identifying a scheduling policy including including the highest demand products first. Based on the simulation and scheduling policy, STMicroelectronics was able to increase its throughput 13% from 15,000 to 16,950 wafers per month and reduce monthly scheduled-related downtime due to launching new versions from 80 to 73.6 hours, which was an 8% reduction in downtimes.

2.4.5 Use of AI Tools for Simulation Analysis and Decision Making

Simulation technology is often used in production to simulate various options without disrupting actual production, but simulation is often time-consuming and requires specialized expertise. Recently, many new AI software applications have made it easier than ever to access a wide range of affordable AI-based manufacturing solutions. Although this trend is occurring across all types of businesses, from small and medium-sized enterprises (SMEs) to large manufacturers, every business can benefit from using AI-based manufacturing solutions.

Most AI platforms are free, while some advanced features can be accessed through paid subscriptions. This enables organizations to conduct more advanced data analysis and interpretation without the need for additional investment in more complex applications (Tishtykbayeva, 2023). In a manufacturing environment, AI tools can support the analysis of production-related data, help interpret simulation results, and provide insights, while also being suitable for those without extensive analytical or programming skills.

Commonly used AI tools, such as ChatGPT, Gemini, Claude, and similar platforms can assist in analysing data and generating insights quickly. This makes them practical for various applications (Soldatos, 2024). In manufacturing simulations, the AI tools can reduce dependence on expensive software with additional simulation features and allow internal staff to perform analysis and make decisions more efficiently. As a result, simulation results can benefit from AI-driven insights.

In this work, AI tools have improved the interpretation and visualization of simulation results, particularly in evaluating the relationships between OEE, throughput, and revenue across different scenarios. This indicates that AI can support discrete event simulation (DES) by converting raw simulation results into useful, actionable insights.

In summary, when combined with simulation, AI provides an effective decision support tool. Simulation provides a method for detailed modelling of systems and evaluation of various operational scenarios, AI enhances result interpretation, visualization, and decision-making efficiency. Combining discrete event simulation (DES) with artificial intelligence (AI) enhances the practicality of simulation-based analysis and enables more informed, data-driven decisions in the operation of manufacturing systems (Cardoso et al., 2023).

2.5 Discussion of Review

The reviewed literature was categorized into three main scopes: downtime analysis, efficiency optimization, and simulation modeling for systematic comparison of the case studies. Each scope was further divided into more detailed methodological approaches. Table 2.1 provides a structured overview of the 29 reviewed papers. It includes the scope and methods each study addressed, highlights the different approaches used and hence provides a useful basis for this project.

Table 2.1 Summary of Reviewed Literature on Downtime and Efficiency in Semiconductor Manufacturing

Scope	Method/Approach	Paper No.																												
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Production Efficiency	OEE mearsurement			/	/	/	/	/																						
	Cycle time analysis	/	/						/																/	/				
	Throughput measurement	/	/						/																/	/	/			
	Lean/ TPM framework			/	/	/	/	/																						
Downtime Analysis	Downtime classification									/	/																			
	MTTR/MTBF metric									/	/	/																		
	Root cause analysis											/	/																	
	Financial/ operational impact										/	/	/																	
Simulation Modelling	Discrete event simulation (DES)	/												/	/	/	/	/	/	/	/			/	/					/
	Downtime modelling in DES																/	/	/	/										
	Schedulling optimization	/												/	/	/	/					/		/	/					
	Preventive maintenance modelling																/	/	/				/	/						
	WITNESS software used																					/	/	/						
	Other DES software (Arena/Simio/FlexSim)																								/	/	/			
AI tools Analysis	AI tools for result interpretation																								/			/	/	/
	ML/ presictive maintenance integration																		/	/										/

<u>References</u>

1–Chong & Ng (2016) | 2–Musa et al. (2025) | 3–Ahuja & Khamba (2008) | 4–Dal et al. (2000) | 5–Muchiri & Pintelon (2008) | 6–Muthiah et al. (2008) | 7–Huang et al. (2003) | 8–Braglia et al. (2008) | 9–Pecchia et al. (2023) | 10–Mourtzis et al. (2021) | 11–Ndokwu (2025) | 12–Colby et al. (2023) | 13–Domaschke et al. (1998) | 14–Hu et al. (2022) | 15–Zeidler et al. (2022) | 16–Yoon & Chae (2019) | 17–Ramirez-Hernandez et al. (2008) | 18–Heger & Voss (2022) | 19–Wang et al. (2023) | 20–Ehm et al. (2015) | 21–Lee & Kim (2020) | 22–Tan et al. (2019) | 23–Lee et al. (2023) | 24–Mönch et al. (2018) | 25–Fowler & Rose (2004) | 26–Dabbas & Chen (2001) | 27–Tishtykbayeva (2023) | 28–Soldatos (2024) | 29–Cardoso et al. (2023)

The literature presented in Sections 2.2 to 2.4 illustrates the important nature of downtime in the context of semiconductor manufacturing. It shows how simulation modelling could be used to obtain and improve manufacturing performance. The literature review reveals that downtime arises from multiple factors, including operational issues that directly affect throughput, cycle time, and cost efficiency. In the following section below, the key findings relevant to this study are discussed.

There are several common methods for measuring and analyzing downtime. Measures such as Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR) and Overall Equipment Effectiveness (OEE) are the most used metrics. These measures help to establish average levels of reliability and monitor the trends in downtime in manufacturing systems. It is important to note that these measures have limitations as they usually assume stable operating conditions. This makes it difficult to quickly respond to complex and dynamic processes in semiconductor manufacturing. This may affect management seeks to predict the impact of changes quickly. In addition, other useful analytical tools such as root cause analysis and Pareto charts are commonly used for categorizing downtime, but their effectiveness depends heavily on data quality. Poor downtime tracking could lead to incomplete or inaccurate diagnoses of downtime issues.

Discrete Event Simulation (DES) is the most prominent simulation method as it allows users to emulate the nature of semiconductor processes. The DES model enables users to examine competing scheduling policies, buffer sizes and machine arrangements in a simulated environment without disrupting real operations. This facilitates the process of simulating downtime to become a decision-support system, meaning the production is no longer reactive to situations but more proactive to address the issues. However, the establishment of DES models and the calibration of typical real operations required time and effort. It is common that data with machine specific failures and repair times is not adequate, making validation of complex models difficult.

Most studies have looked at integrating simulation with downtime reduction methods. In this regard, it is found from the review of the papers that there is a lack of a model integrating OEE as the key metric for simulation, followed by using AI tools to improve decision making. These findings demonstrate there are potential benefits by combining both the developed simulation model and new AI tools to visualize and analysis the effects of changes in production, especially in terms of overall OEE score, throughput, and revenue.

Within this context, there is a rationale for using WITNESS simulation software for this study. Compared to other simulation software, WITNESS has been demonstrated to be appropriate for manufacturing studies as it has user-friendly interface, is flexible to model re-entrant production flow, and has good visualization capabilities. In addition, WITNESS possesses built-in functions for downtime modelling and performance evaluation, which reduces the need for complex coding and makes it more practical for academic research that values both accuracy and usability.

In summary, the literature reviewed showed that downtime in semiconductor manufacturing was a complex and multidimensional problem that requires effective measurement, analysis and reduction. Although simulation, particularly Discrete Event Simulation (DES), offers advanced capabilities to emulate the complex nature of a production line, there remains issues with availability of data and model development. This affects the ability to use simulation results for decision making that initiates impactful improvements. Addressing these gaps through the combination of simulation with real-time data, throughput and financial considerations provides the case for the present study as well as for future research.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter discusses the research methods employed in this project. This methodology describes a clear and orderly approach to ensure the achievement of project objectives. It is divided into the following sections: Section 3.2 outlines the general methods for conducting the study, including research design and the flow of the project. Section 3.3 describes the preparation of the research framework and outlines the five stages that guide the research from start to finish. Section 3.4 provides a detailed introduction to each part of the stage. It includes collecting and analyzing data, building and validating the simulation models, testing scenarios, and concluding simulation studies.

3.2 Methodology to Conduct the Study

To achieve the aim and objectives of this study, the research methodology was carefully planned. The objectives are to analyze and categorize equipment downtime data in a semiconductor manufacturing line, to develop a WITNESS-based simulation model for evaluating downtime reduction scenarios, and to validate the model by comparing the simulation outputs with real production data. The study followed a structured methodology to ensure that the research remained on the correct track and achieved its stated objectives. The methodology used to conduct the study is shown in Figure 3.1.

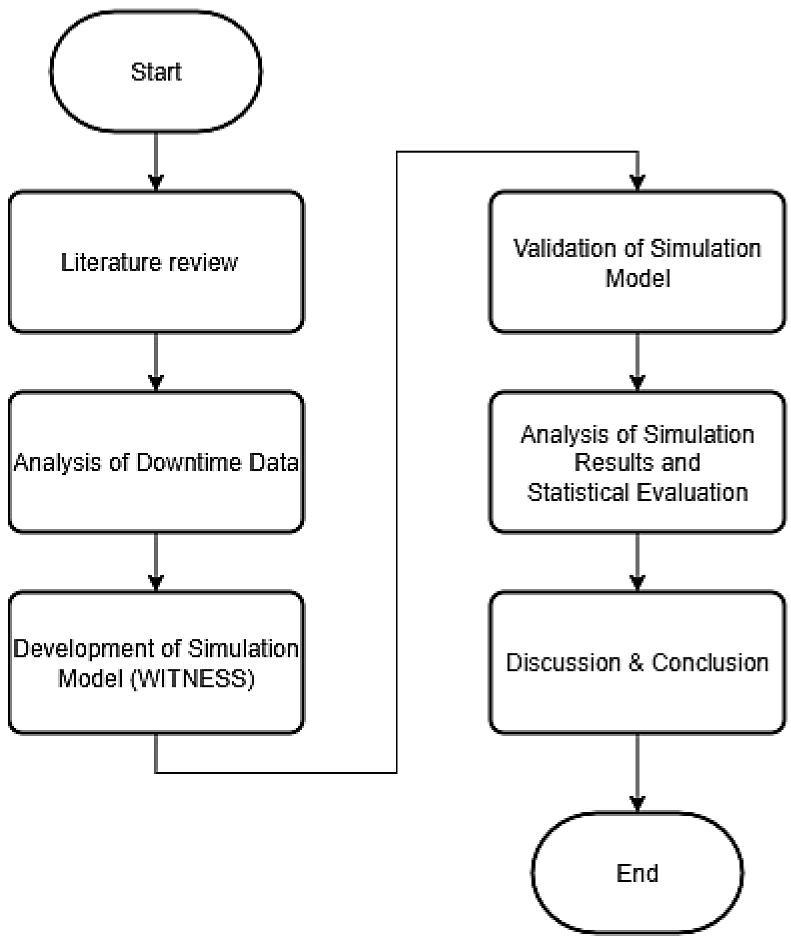


Figure 3.1: Methodology to Conduct the Study

The first phase is literature review on downtime analysis, simulation modeling and production efficiency in semiconductor manufacturing. The review aimed to identify established downtime metrics, including Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR) and OEE. At the same time, it also covers some methods such as root cause analysis, Pareto analysis, and predictive maintenance techniques. The literature review also addressed simulation methodologies, particularly DES in order to demonstrate the usefulness of simulation tools in the industry.

A detailed analysis and categorization of the data provided by Company X was carried out to analyze machine downtimes and their overall impact on throughput and revenue. This involved the categorization of downtime events into categories such as equipment failure, setup time and scheduled maintenance. This enables the production team to focus on types of major issues that impacted the MPI line the most. In addition, OEE was applied as the key metric of the simulation model. A discrete-event simulation model was then developed using WITNESS software based on data analysis. WITNESS was chosen because of its user-friendly interface, its ability to emulate processes in semiconductor industries and its built-in downtime modelling functions.

Validation was then conducted to ensure the accuracy of the simulation model. Validation involved comparing simulation outputs to actual production data. Throughput rate, machine utilization, and distribution of downtime were the main inputs used to develop the model. If substantial differences were identified, then the parameters of the model were modified until the outputs of the simulation resembled the actual outputs as closely as possible. This step was important for establishing the reliability of the simulation results for decision-making.

Once the model was validated, it was used to evaluate different downtime reduction, product mix and performance and quality scenarios. The simulation results were analyzed to assess the impacts of different scenarios. Baseline performance was compared with improved scenarios to determine potential returns, especially in throughput and revenue.

The final step involved visualizing and interpreting the simulation results with the aid of AI tools. The implications of the downtime reduction, the strengths and limitations of using simulation as a decision-support too, and the advantages of incorporating AI tools into the simulation process were examined. Conclusions were then made to summarize the key contributions of the project and assess whether the objectives of the project have been achieved. Finally, recommendations were provided for future research.

3.3 Development of Framework

In Figure 3.2, a framework was developed to systematically analyze and improve production efficiency in semiconductor manufacturing using downtime data. It consists of five stages: Initiation, Data Analysis, Modelling, Experimentation & Analysis, and Termination. Each stage has clear objectives and defined steps, providing a structured approach to identify, simulate, and implement strategies for reducing downtime and enhancing overall line performance.

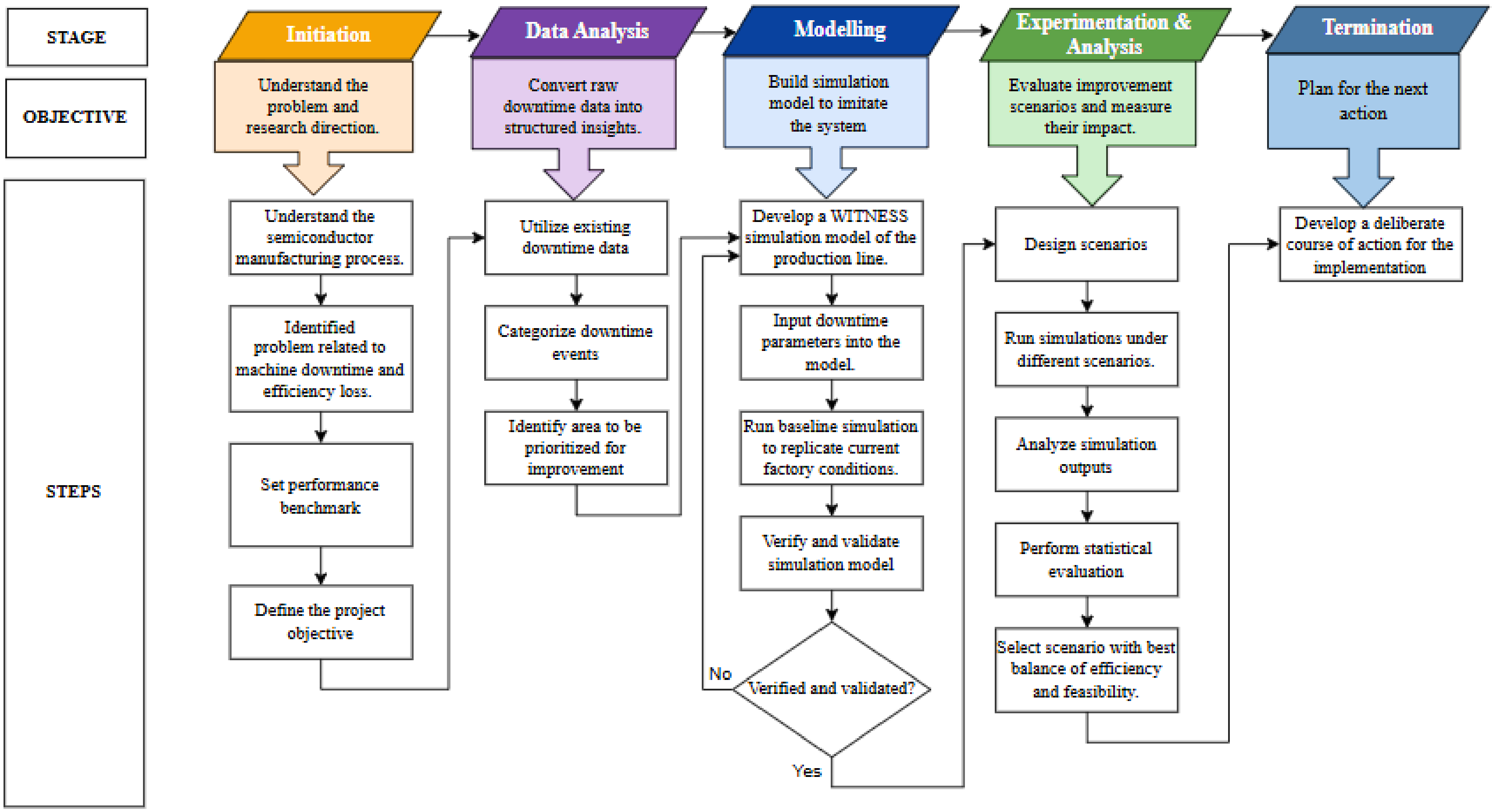


Figure 3.2: Framework of the study

3.4 Explanation of the Framework

Based on the framework outlined in Figure 3.2, five phases are involved in the developed framework. Each phase of the framework is explained step by step in the following sub-subsection.

3.4.1 Stage 1: Initiation

The first step is understanding the semiconductor manufacturing process. The background of the manufacturing system, such as production system, manufacturing processes, process flow and existing resources must be distinguished within the MPI testing line. A comprehensive understanding of these elements will help ensure that the simulation model can accurately replicate the actual manufacturing system. It is important to ensure that the knowledge possessed about the manufacturing system is sufficient to carry out further investigation. Semiconductor factories involve random factors, including variations in processing times and equipment reliability, making this knowledge important for accurate follow-on modeling.

Next, problems causing machine downtime and low efficiency are identified. This starts with looking for obvious issues, like machines breaking down often, which can slow production or increase WIP inventory. Downtime is categorized into basic things such as mechanical failures or operational delays based on production logs or on-site observations. In semiconductor manufacturing for example, unplanned downtime will affect the entire production line due to repetitive steps. After recognizing unplanned downtime overtime, simple tools, such as fishbone diagrams can help identify main causes, and decide which issues to investigate more closely.

The third step is to set a performance benchmark, which is a measurement that shows how well the areas identified in the previous step are performing. The benchmark provides a reference point to check if the improvement project is working by comparing the current performance with future results. This benchmark also helps in defining the project's goals in the next step. The team can choose any performance measure as the benchmark, if it directly relates to the issues they want to improve.

The fourth step consists of identifying the process improvement goal based on the problems defined in the semiconductor production line, as well as the performance benchmark set up in a previous step. A situationally specific goal provides the process improvement project with a direction, as well as inspiration and focus. A process improvement project goal should be clear, unambiguous, measurable, and realistic; usually convenient frameworks, such as SMART (Specific, Measurable, Achievable, Relevant, Time-bound), are used to help make the project goal more specific and actionable. The project goal should articulate the targeted level of improvement, clearly specifying deviations in the key performance measures such as planned downtime using machinery or line efficiency. For instance, if the identified issue involves unplanned downtime at a specific process station, the improvement target should be defined in measurable terms, such as reducing unplanned downtime by a specified percentage within a defined timeframe. It is also important that the improvement target is realistic and aligned with the existing system performance as well as management expectations, in order to ensure that the defined goal is both achievable and meaningful within the context of the study.

3.4.2 Stage 2: Data Analysis

During the data analysis phase, unstructured downtime data is organized and arranged so the relevant areas for potential improvement can be easily identified. This phase is critical in semiconductor manufacturing because production lines usually create large amounts of log data every day. This data can be analyzed to find useful patterns and trends. The main task is to turn the raw data into meaningful OEE metrics that can enable the design of simulation models. The three-week data was provided to capture different patterns to establish a more consistent downtime baseline for the MPI machines.

Subsequently, downtime events are categorized into major and minor errors. A Pareto analysis approach is used to identify causes that are most prevalent or have the most impact on the MPI line. This ensures that efforts can be focused on the most significant issues. Clustering or grouping techniques are used to simplify categories of downtime so that analysis of data becomes less time consuming.

Lastly, the critical factors for improvement are determined. Metrics such as MTBF, MTTR, frequency of downtime and total downtime in terms of their impact are used to determine the needed actions. Priorities are generally set according to impact or cost. In semiconductor operations, simply focusing on critical downtimes can often significantly improve throughput and revenue.

3.4.3 Stage 3: Modelling

The third stage involves modeling the MPI line using WITNESS simulation software. The objective of this phase is to ensure that the constructed simulation model accurately replicates the actual operation of the MPI line. This is crucial to ensure the simulation results are reliable. The modeling stage creates a virtual version of the MPI line using downtime data, and this allows the production team to simulate multiple scenarios without affecting real production. The main purpose for this state is to create a simulation model that emulates the actual production system as accurately as possible in order to obtain reliable insight into different scenarios for better decision making and actions.

At this stage, a WITNESS simulation model of the production line is established. It includes entities like wafers, machines, and buffers, along with process routes, times, and constraints. The software helps visualize flows and identify problems, and it is well-suited for complex semiconductor lines with unpredictable events.

Next, downtime data can then be entered into the model as statistical distributions (this could mean exponential distributions for machine failure) so the WITNESS simulation can simulate interruptions that would statistically occur in a real factory condition. A few preliminary tests should be conducted to confirm that the model is stable before running the model in the baseline simulation.

A baseline simulation is run in order to reproduce the current factory condition. In multiple runs, the key performance indicators (KPI) including throughput and machine utilization are measured in order to summarize outputs and recognize some degree of randomness. The simulated factory outputs from WITNESS simulation model can then be compared to real factory output.

Moreover, step four is to verify and validate the simulation model. Verification of the model of the simulation contains two parts. Verification ensures that the simulation program is functioning correctly and operating under the appropriate inputs and assumptions. According to Carson (2002), verification can be performed through several approaches, including structured walkthroughs, logical checks, and manual calculations to confirm that the simulation produces the expected output under predefined conditions. This can include manual calculations or otherwise checking against timelines to confirm that the simulation produces the expected results. Validation of the model involves establishing if the simulation is a valid replication of the real manufacturing system (Sargent, 2013). When attempting to validate the model, key performance measures such as flow time and throughput are compared between the simulation and the actual system. If the simulation deviates by less than 15% for these measures of system performance compared to the real system, then we consider the model to be accurate (Law & Kelton, 2007).

If the model verification or validation is unsuccessful, this means the simulation model is incorrect. Practitioners would need to go back to the modeling stage for the related steps which will require the practitioners to verify if there is sufficient and accurate data collected, ensure the data is entered into the simulation correctly, and ascertain if there are no mistakes in the simulation itself.

3.4.4 Stage 4: Experimentation and Analysis

This stage examines possible enhancements using simulation to determine the optimal solutions. The purpose is to test scenarios against strategies where performance is located, as a basis for taking decisions on the data made available.

The first step is creating multiple scenarios based on Availability (A), Performance (P) and Quality (Q). It involves changes in downtime, product mix, cycle time and yields. The next step is connecting to the simulation results, normally through charts to illustrate improvements, especially in terms of throughput and revenue. AI tools are used to help compare and evaluate scenarios with baseline performance. In short, a scenario with the best outcome is selected where the decision made would be based on quantitative simulation results.

3.4.5 Stage 5: Termination

The final stage of the framework involves concluding the simulation study and translating the findings into practical insights for the manufacturing system. In this study, termination refers to the completion of simulation runs after sufficient data have been collected to represent system performance accurately.

The simulation model is executed over a predefined production period based on actual operational data. Upon completion, the simulation results will be evaluated against a series of system performance metrics such as availability, OEE, throughput, and revenue generated under different scenarios.

The simulation results help the company make decisions. The model can identify possible improvements based on the scenario evaluation, such as reductions in machine downtime, better performance of machines and quality. These findings also provide the company with an understanding of how various operational changes will affect the performance of the entire system. The model can be used as a reference to guide improvement efforts and optimize overall production efficiency.

While the simulation model does not implement changes in the physical system, companies may use the simulation results to assist with decision making and planning for future improvements. This allows engineers and management the ability to evaluate the degree of feasibility, expected benefits of improvement and potential risk prior to making the improvement.

CHAPTER 4

RESULTS AND DISCUSSION

4.0 Introduction

This chapter presents and discusses the results obtained from the downtime data analysis, the WITNESS simulation model, and the comparison with real production improvement data. The first section presents the background of the semiconductor production system and the MPI testing line. The second section focuses on the analysis of downtime data, including categorization and identification of major downtime contributors. The third section mainly describes the development of the simulation model using WITNESS software. The fourth section involves using the simulation to evaluate different scenarios with the aid of AI tools. Finally, the fifth section discusses the findings of the study and provides discussion based on simulation results.

4.1 Background of the Company

The case study was conducted in Company X, a global semiconductor manufacturing company located in Penang, Malaysia. The company operates globally and has customers in more than 30 countries worldwide. It serves many different types of industries such as automotive lighting and consumer electronics, as well as providing general illumination. As a key player in the semiconductor industry, the company operates in a highly competitive and demand-driven environment, where production requirements are closely affected by market demand.

Customer demand can fluctuate significantly due to the nature of the semiconductor supply chain. As such, the semiconductor companies need to maintain a high level of responsiveness and production flexibility. The manufacturing system in semiconductor industry follows a high-volume production environment with continuous operations to meet global demand. As a result, the factory usually operates 24 hours a day, 365 days a year, to ensure consistent production output and timely delivery.

Production processes largely depend on multiple interlinked operations within the production floor. All operations including material handling, inspection, testing and assembly must be carefully coordinated to maintain product quality and production efficiency. In such a complex manufacturing environment, equipment performance and reliability play an important role in ensuring smooth operations.

Consequently, to maintain overall production efficiency, it is essential to minimize production interruptions and ensure high equipment availability. This highlights the importance of analyzing equipment downtime and identifying areas for improvement, which forms the basis of this survey.

4.2 Stage 1: Initiation

This study focuses on the MPI testing line in a semiconductor manufacturing system. An MPI machine is a fully automated LED die prober consisting of a vision system, a pick-and-place module, and a pin-ejection module. The process involves die pick-up, placement onto the testing platform, electrical testing, and transfer of tested dies to output.

Figure 4.1: MPI Machine (Source: Company X)

Figures 4.2 and 4.4 illustrate the die probe manufacturing process and the main components of the MPI equipment.

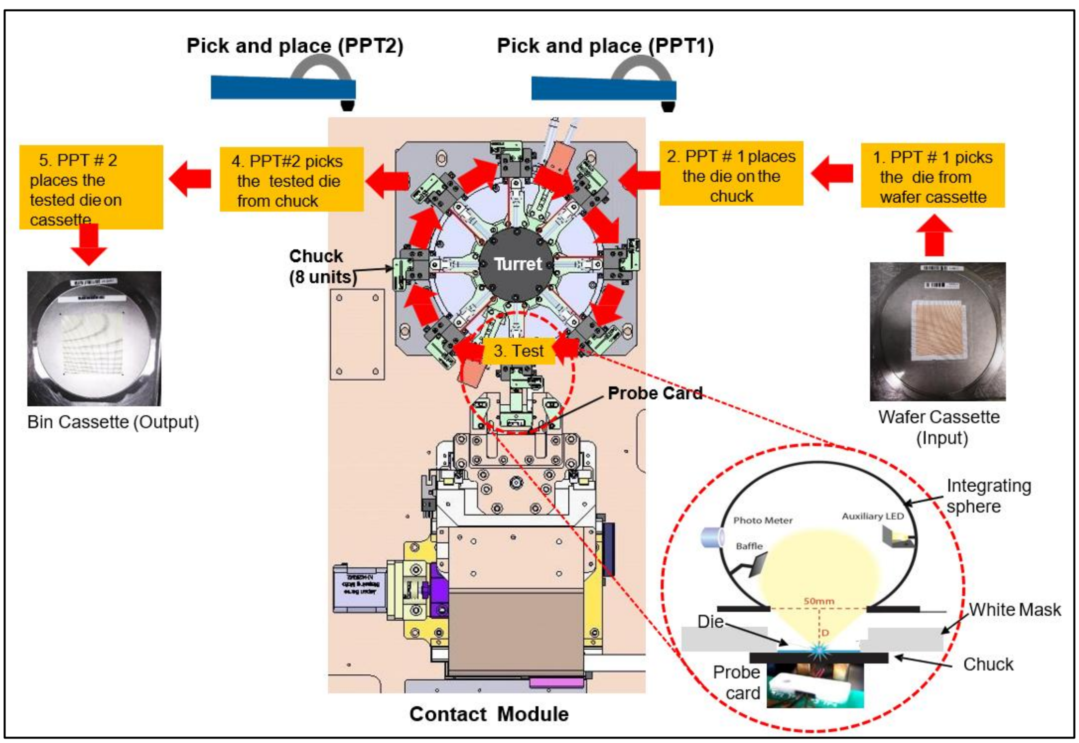


Figure 4.2: The die probing process flows in MPI (Source: Company X)

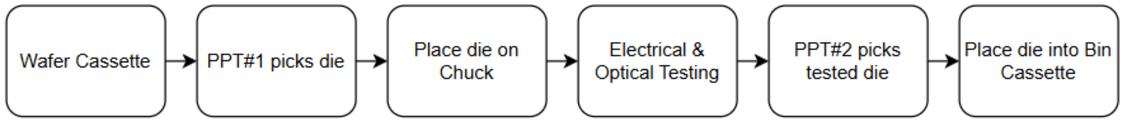


Figure 4.3: Process Flow of MPI Testing Operation

As shown in Figure 4.3, the MPI testing process consists of sequential steps from die pick-up to final binning.

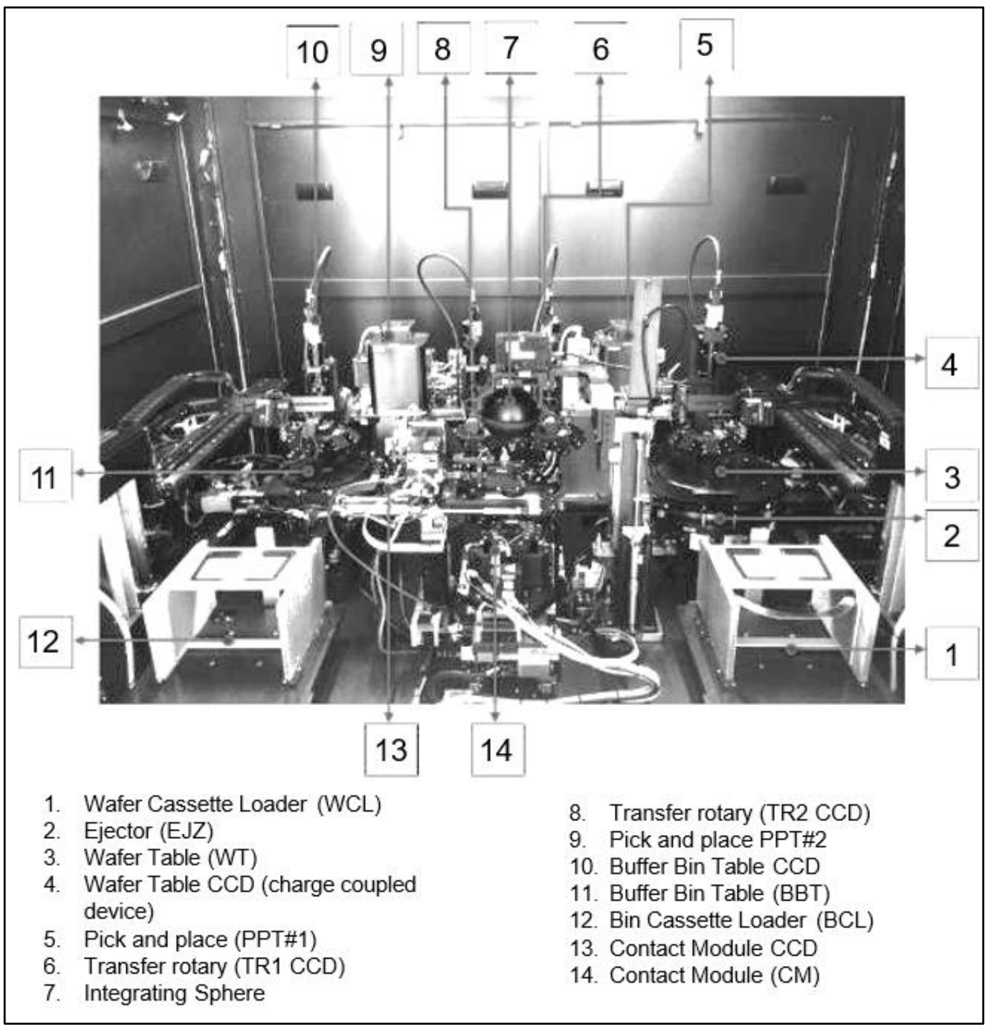


Figure 4.4: Main Components in MPI Machine (Source: Company X)

The MPI machine operates as an integrated system for testing and sorting LED dies based on their electrical and optical performance. The process begins with wafer loading through the Wafer Cassette Loader (WCL), followed by wafer alignment and pre-scan at the Wafer Table (WT) using a vision system. An ejector mechanism then pushes individual dies upward for pick-up. The pick-and-place module (PPT#1) uses a vacuum collet to pick individual dies from the wafer and place them onto the chuck located in the contact module. The chuck holds the die in position and provides thermal control during testing. Electrical and optical testing are performed using a probe card and integrating sphere, where the die performance is measured and classified. After testing, the second pick-and-place module (PPT#2) transfers the tested die from the chuck to the Bin Buffer Table (BBT), where it is sorted and placed into the appropriate bin cassette. Vision systems are utilized throughout the process to ensure alignment accuracy and positioning during die transfer. The entire operation runs continuously, where dies are sequentially picked, tested, and sorted until all dies are processed. Because of the high volume and highly automated nature of this operation, the system is very sensitive to machine stoppages and process disturbances.

In the actual production system, a total of 43 MPI machines operate in parallel, distributed across different sections to process products for various customers. For this study, 10 MPI machines are considered, as they are dedicated to processing products for Customer Y. As all MPI machines share identical configurations and operating characteristics, the production data are aggregated and appropriately scaled. This approach reduces model complexity while preserving the overall behaviour and representativeness of the actual production system. Each machine produces an average output of approximately 5,500 dies per hour. Due to this high-volume operation, any machine interruption results in significant output loss. It has been determined that the MPI test line is facing productivity challenges, primarily due to excessive machine downtime. Therefore, machine downtime has been identified as the main issue limiting production efficiency in the system.

To evaluate system performance, Overall Equipment Effectiveness (OEE) is selected as the primary performance indicator. OEE is a comprehensive statistical metric that evaluates equipment efficiency based on three components: availability, performance efficiency, and quality rate. Among these components, availability is one of the components directly affected by machine downtime, making OEE a suitable metric for assessing the impact of downtime on production performance. In the MPI testing line, high downtime reduces machine availability, which subsequently lowers overall OEE and limits production throughput. Therefore, improving OEE through downtime reduction is essential for enhancing system efficiency and equipment utilization. OEE is selected as the benchmark in this study because it provides a comprehensive evaluation of machine performance by capturing the combined effects of different types of production losses. Compared to individual metrics such as throughput or utilization, OEE offers a more reliable and holistic measure of overall system performance.

Based on the observed performance gaps, this study aims to improve the OEE of the MPI machines by reducing downtime-related losses. Although the targeted improvement may appear small, even a marginal increase in OEE can result in significant output gains due to the high production rate of approximately 5,500 dies per hour.

In the next section, the data from the tracking system was analyzed and categorized to identify the major contributors affecting machine performance. Based on the identified downtime issues, a simulation model is developed using WITNESS to represent the operation of the MPI system. This enables the evaluation of different improvement strategies aimed at reducing downtime without disrupting actual production operations.

4.3 Stage 2: Data Analysis

The downtime data were collected from the MPI testing line of Company X over a three-week period, from workweek 43 (WW43) to workweek 45 (WW45). The data used in this study was sourced from a real-time automated performance tracking system, which continuously records equipment status and production performance, including measurements of throughput and operational status. However, to conduct a comprehensive performance evaluation, additional data extraction and analysis are required. The collected dataset was analyzed based on the three key components of Overall Equipment Effectiveness (OEE), namely availability (A), performance efficiency (P), and quality rate (Q), as shown in Figure 4.5.

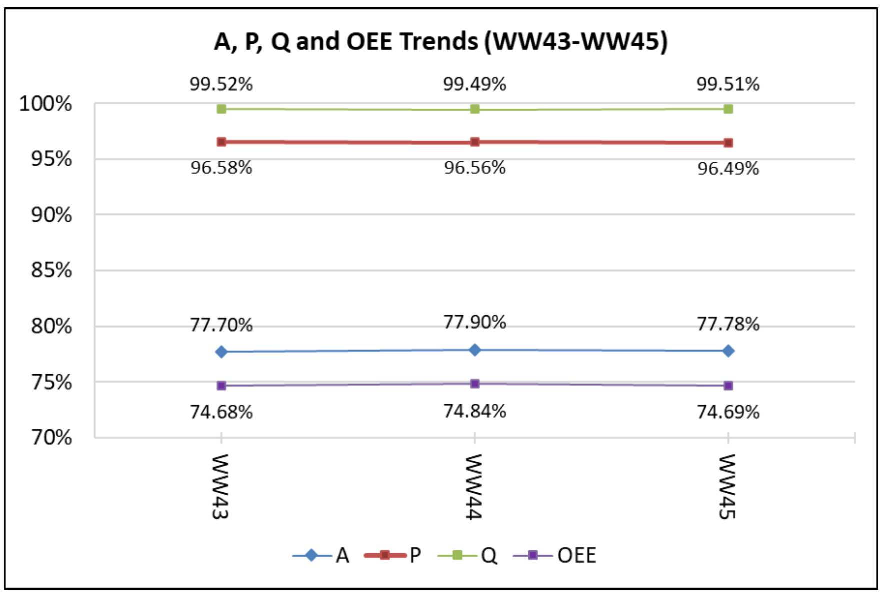


Figure 4.5: A, P, Q and OEE Trends of MPI Line from WW43 to WW45

As shown in Figure 4.5, availability (A) remains around 77.80%, which is below the benchmark value of 90% suggested by Nakajima (1988). In contrast, performance (P) and quality (Q) are consistently high, averaging 96.54% and 99.51%, respectively. Since P and Q remain stable, variations in OEE are primarily driven by changes in availability. Therefore, improving availability is identified as the key focus for enhancing overall system performance. Therefore, given that A is the only metric with the greatest potential for improvement, minimizing losses in the A aspect will significantly enhance the overall OEE performance of the MPI production line.

Availability losses in the MPI testing line are primarily attributed to unplanned downtime, which includes unexpected equipment failures, process interruptions, and operational disturbances. In the context of the MPI test process, these interruptions directly impact equipment availability, as they cause the test process to pause and reduce effective production time. The availability tracking data collected in the MPI tracking system comprises a few codes that define the types of equipment interruptions or issues associated with equipment downtime. To support systematic analysis, all available equipment downtime can be categorized into major groups, including contact resistance, contact yield, conversion, pick-and-place faults (PPT), printer faults, and others, as shown in Table 4.1. For modelling purposes, equipment downtime items with a minor impact are grouped under the ‘Other’ category to simplify model complexity whilst maintaining the representativeness of the entire MPI test system. This would not affect the overall results of the simulation, as the downtime categorized under 'Others' was not omitted and was still included in the simulation. These classifications facilitate a clearer identification of the primary factors contributing to the loss of A and enable a more in-depth analysis of downtime behavior. Figure 4.6 illustrates the breakdown of these major downtime categories during the period from WW43 to WW45.

Table 4.1: Types of Major Downtime Errors

No.	Downtime Error
1	Contact resistance
2	Contact yield
3	Conversion
4	Others
5	PPT1 problem
6	PPT2 problem
7	Printer problem

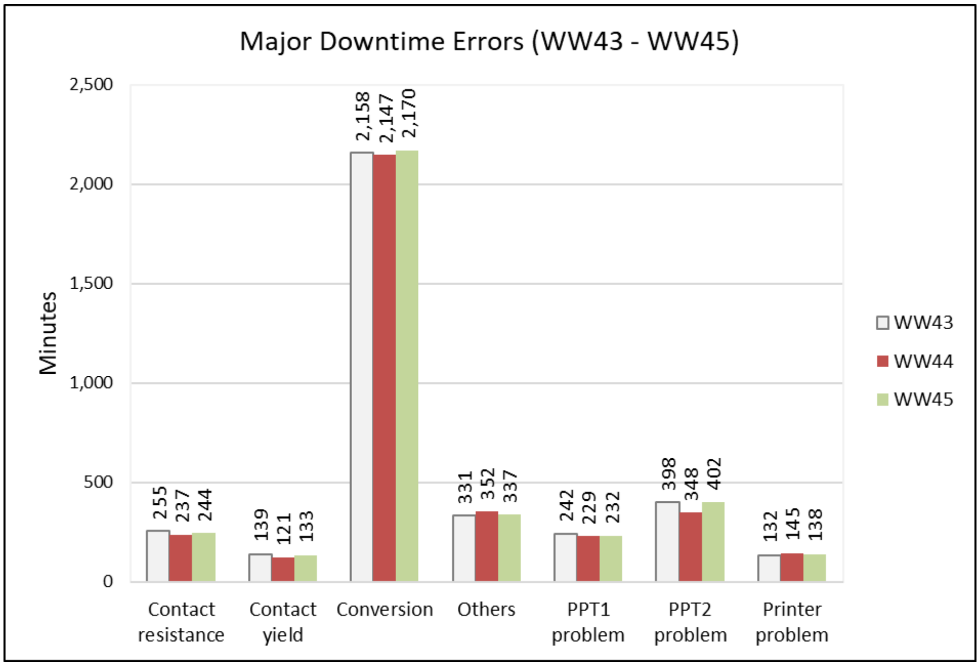


Figure 4.6: Major Downtime Errors from WW43 to WW45.

In the MPI testing line of Company X, a thermal chuck is integrated into the MPI machine to maintain the device under test at a controlled temperature throughout the testing process. If the device under test exceeds the maximum temperature, the machine stops temporarily to allow cooling. While the thermal chuck provides a stable testing environment for temperature-sensitive devices, it also introduces interruptions that negatively affect machine availability. On MPI test lines, downtime caused by temperature-related interruptions is referred to as cold shutdowns and is factored into overall availability calculations. Based on the average data from WW43 to WW45, cold shutdown accounts for approximately 18% of total available production time and contributes about 84% of total downtime. This indicates that thermal instability is a primary bottleneck in the MPI testing process. Consequently, cold shutdowns have a direct impact on equipment availability and overall OEE performance. Figure 4.7 illustrates the definition of availability loss on the MPI production line, whilst Figure 4.8 presents the downtime associated with cold shutdowns during the period from WW43 to WW45.

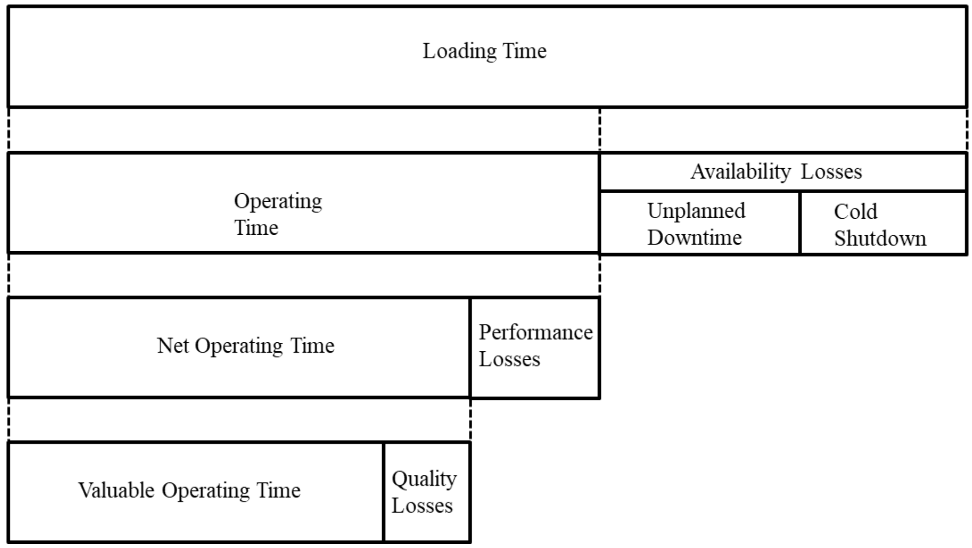


Figure 4.7: Definition of A losses in the MPI Line

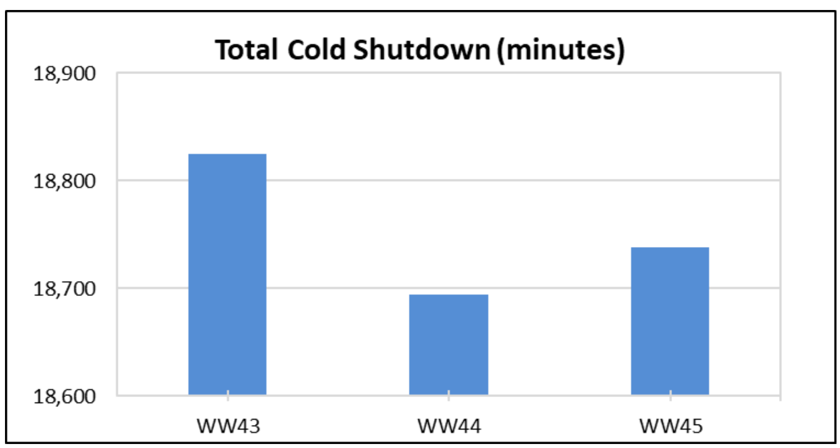


Figure 4.8: Downtime due to Cold Shutdown (WW43 - WW45)

Furthermore, unplanned downtime for each MPI machine in WW46 was analyzed using Pareto charts, as shown in Figure 4.9. The Pareto charts indicate that downtime distribution varies across machines. In WW46, conversion is dominant in MPI #1, #2, and #4, while other machines are mainly affected by PPT-related issues, contact resistance, and other operational factors. This suggests that downtime behaviour is machine dependent rather than uniform across the system. Although all machines process products for Customer Y, variations in workload, parameter adjustments, and machine condition may lead to differences in dominant downtime types.

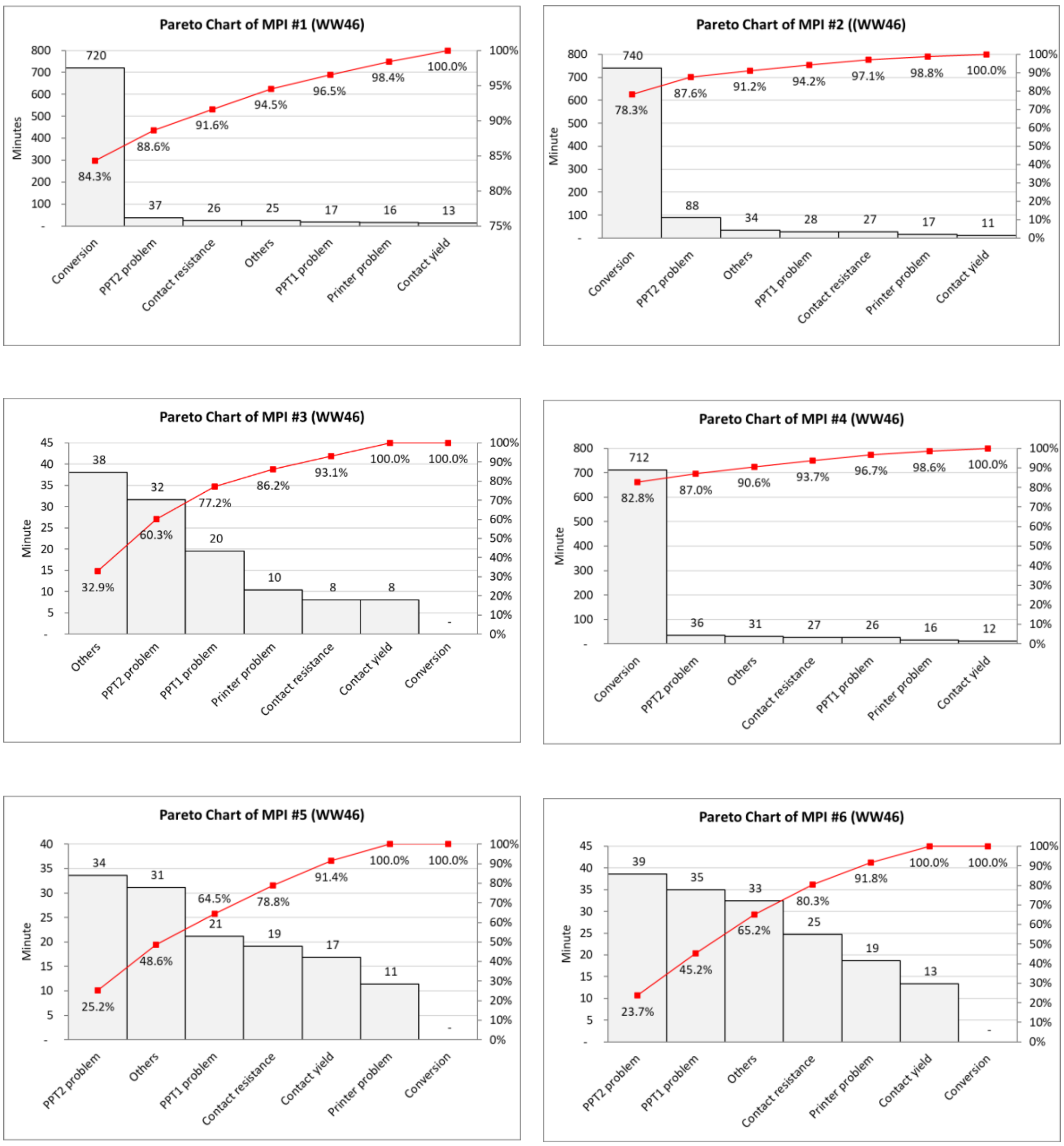


Figure 4.9: Unplanned Downtime by Error Type for Each MPI

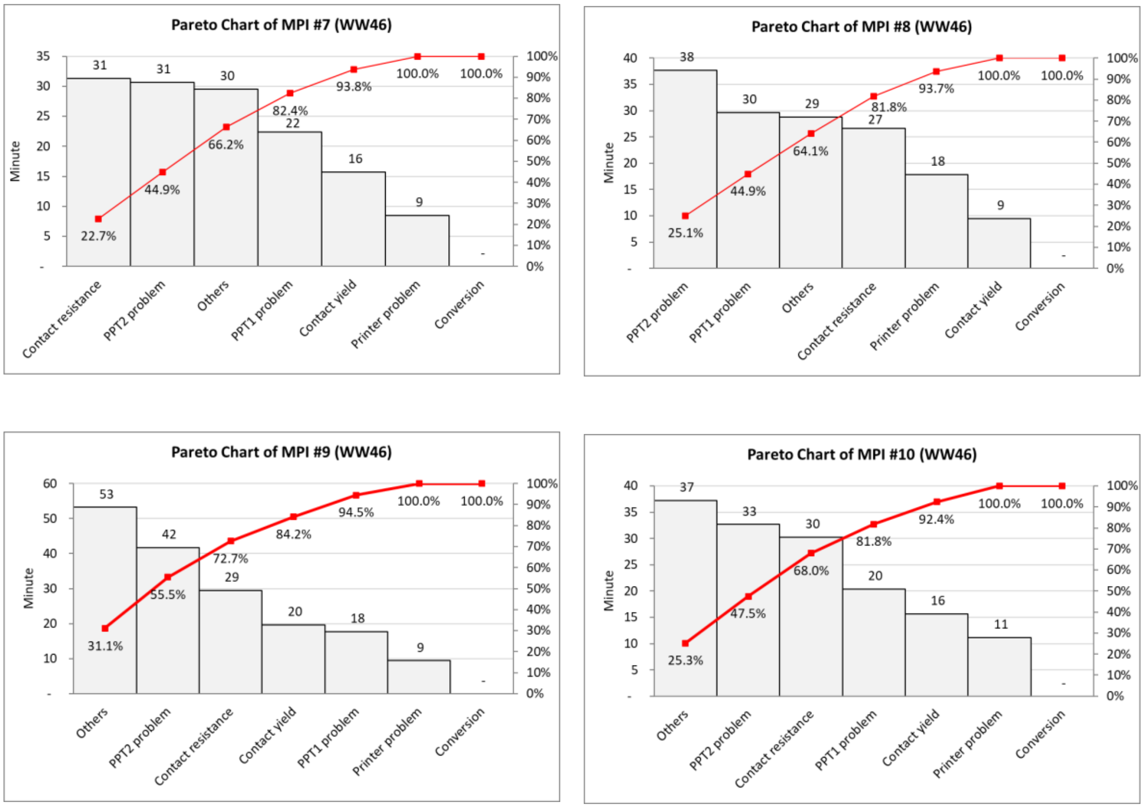


Figure 4.9 (continued): Unplanned Downtime by Error Type for Each MPI

While the cold shutdown time for each MPI in the same workweek was as shown in Figure 4.10. Cold shutdown occurs consistently across all machines, with relatively small variation, confirming that it is a system-wide issue rather than a machine-specific problem. Consequently, cold shutdowns are the primary cause of availability loss across the MPI testing line.

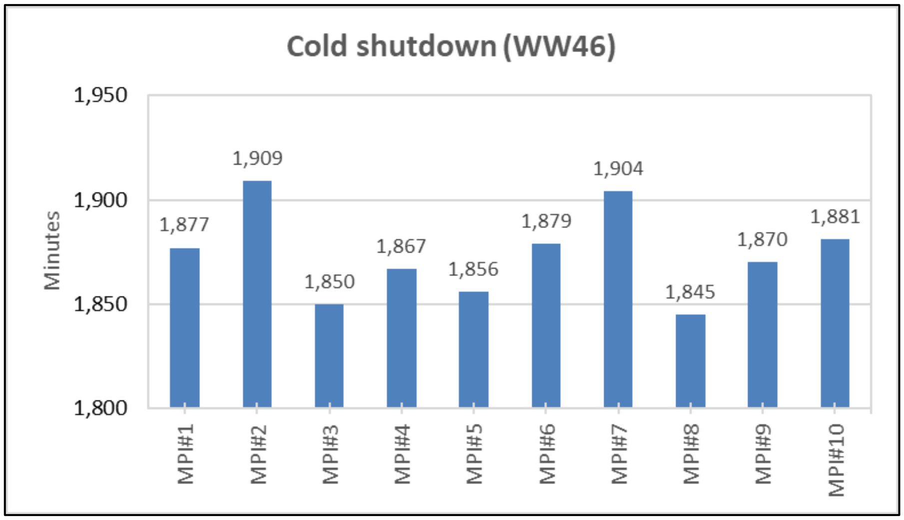


Figure 4.10: Cold Shutdown Time for Each MPI

The average overall OEE value obtained from the analysis for the 10 MPI machines was 74.70%. The A, P, Q, and OEE trends are shown in Figure 4.11. Availability shows a similar trend to OEE, while performance and quality remain stable. This confirms that OEE performance is primarily influenced by availability.

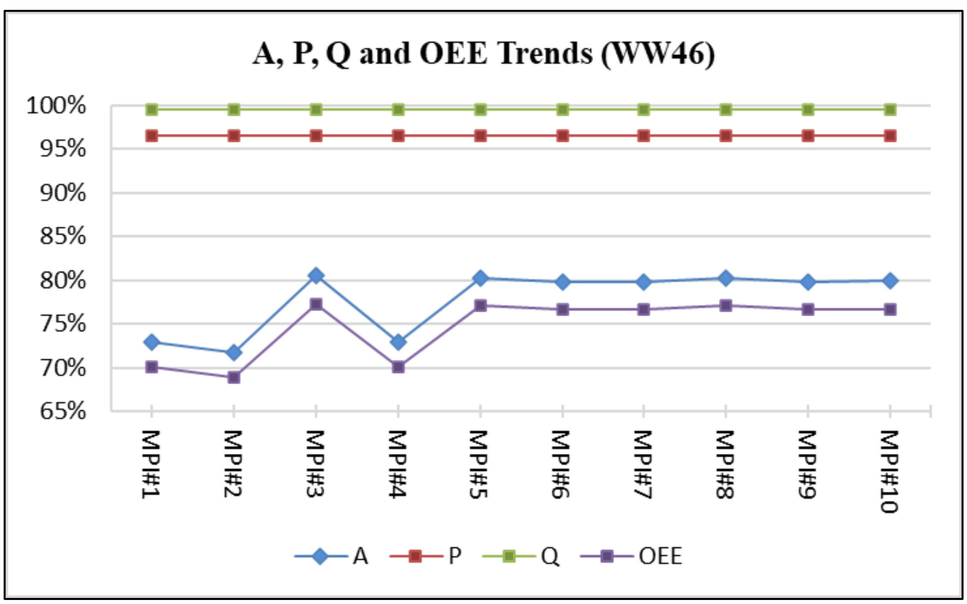


Figure 4.11: A, P, Q and OEE Trends for Each MPI (WW46)

Overall, the analysis indicates that availability losses are the primary factor limiting the OEE of the MPI test line. Cold shutdowns appear to be a significant contributor to overall losses across the entire test line system, while conversion and other operational factors contribute for fluctuations attributable solely to individual equipment. This analysis provides a solid foundation for the development of the simulation model described in stage 3.

4.4 Stage 3: Modelling

In this stage, a discrete-event simulation (DES) model was developed using WITNESS to represent the MPI testing process of Company X. As shown in Figure 4.12, the simulation model consists of 10 MPI machines operating in parallel, representing the actual production system of Company X. This approach allows us to capture the behaviour of a real-world production system while avoiding excessive model complexity.

Each MPI machine in the model is represented individually (M1–M10). Each machine has its own set of key performance indicators, such as availability, uptime and Overall Equipment Effectiveness (OEE). The model can also calculate the total performance (total output) of the entire system, as well as the system's overall OEE. It is therefore capable of measuring the performance of both individual machines and the system. There are no buffers in the simulation model, as each MPI machine operates as standalone in-test-out system, neither dependent on upstream processes nor affected by downstream processes. The testing process is independent of other processes, and there are no intermediate storage or queues between machines that could cause interruptions in the material flow. The model assumes a continuous inflow of wafers, thereby ensuring that machines are always fed in accordance with standard production practices. In this way, the model can focus on machine-related performance characteristics without being affected by physical stock shortages or material logistics constraints. Based on the stability observed in historical data, the performance efficiency (P) and quality rate (Q) have been set at 96.54% and 99.51% respectively. Consequently, the model focuses primarily on the loss of availability caused by equipment downtime.

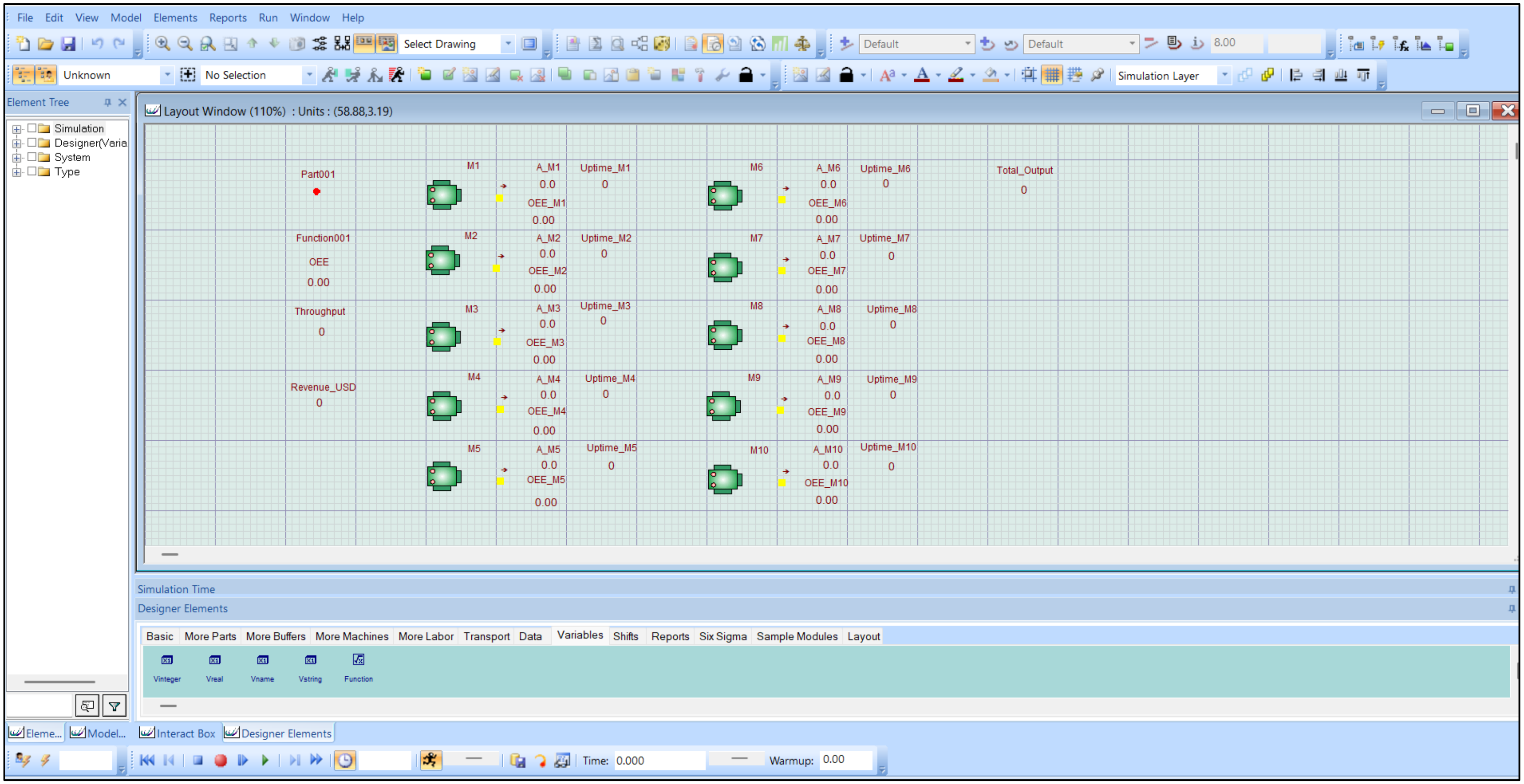


Figure 4.12: WITNESS Simulation Model of the MPI Testing Line

Based on the analysis data from Stage 2, all input parameters for the simulation model were determined. There are 7 types of downtime errors and cold shutdown. Each type of downtime has been incorporated into the model based on its observed frequency and duration, to ensure that the simulation results accurately reflect the actual operating conditions of the MPI testing line. In this study, the MTBF for each downtime category was determined through analysis of the failure frequency of downtime events. Within the simulation model, the frequency and duration of downtime events were iteratively adjusted until the simulated downtime closely matched the observed average downtime from the three-week dataset. This ensures that the model accurately represents real system behaviour at the system level. The estimated MTBF values effectively reproduce observed downtime patterns and provide a reliable basis for evaluating the impact of downtime on OEE and throughput.

The average throughput of each MPI was 5,500 units. The average cycle time was calculated as shown below:

$$\text{Average cycle time} = \frac{60}{5500}$$

$$= 0.01091 \text{ minutes/unit}$$

Using this cycle time and assuming the initial A, P, and Q were 100%, the ideal simulation results for one week are as shown in Figure 4.13. The run time was derived as shown below.

$$\text{Total production time} = 7\text{days} \times 24 \text{ hours} \times 60 \text{ minutes}$$

$$= 10080 \text{ minutes}$$

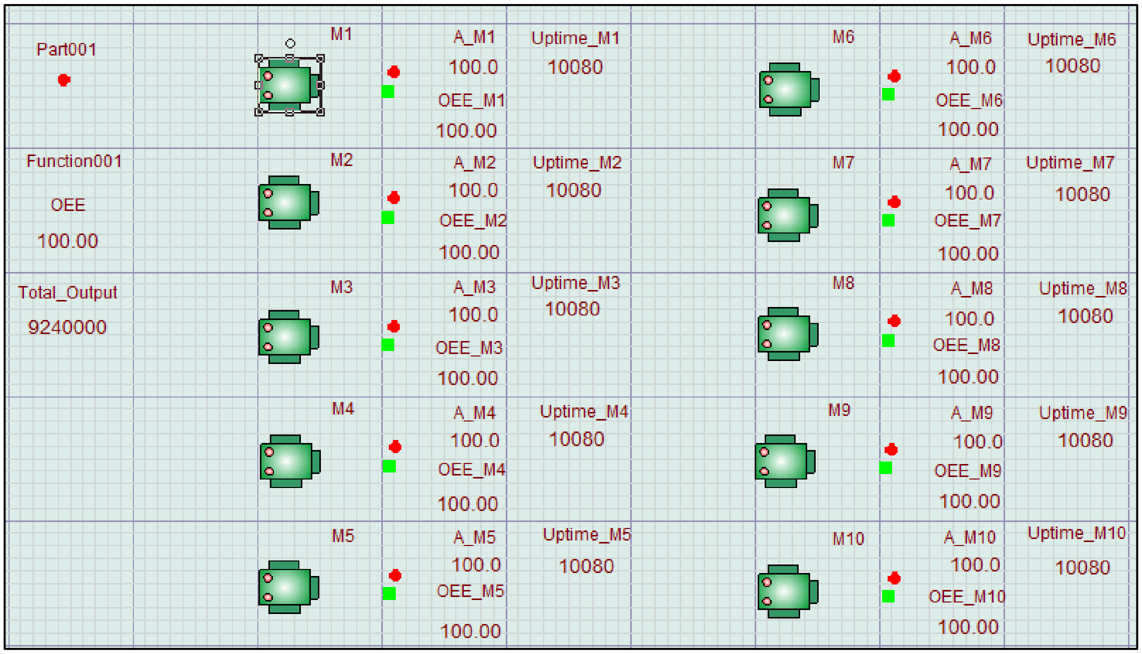


Figure 4.13: Ideal MPI Line Model in Witness

The final step of this stage involves the verification and validation of the simulation model to ensure its accuracy in representing the behaviour of the actual MPI testing line. It can be observed that the uptime for each MPI machine is 10,080 minutes, resulting in an availability (A_MX, where X denotes the MPI number) of 100%. Under these conditions, both the overall OEE and availability reach 100%, while the total output (Total_Output) is 9,240,000 units, calculated based on an average production rate of 5,500 units per hour over 7 days across 10 MPI machines. These results indicate that the model represents an ideal operating environment with no downtime or performance losses.

In addition, the downtime data for major errors and cold shutdowns in WW46 were inserted into WITNESS and the simulation results are shown in Figure 4.14. Following verification, the model was carried out by comparing the model results with actual production data from WW46. The simulation included performance metrics for each MPI machine, such as availability, uptime and OEE, as well as overall system performance metrics. The simulation results were compared with the calculated values obtained from the actual production data. The comparison shows that the simulated overall OEE of 74.71%

is highly consistent with the calculated OEE of 74.70% as shown in Stage 2, with only minimal deviation observed. The percentage difference in Availability (A) ranges from –0.44% to 0.12%, further confirming the accuracy and reliability of the simulation model.

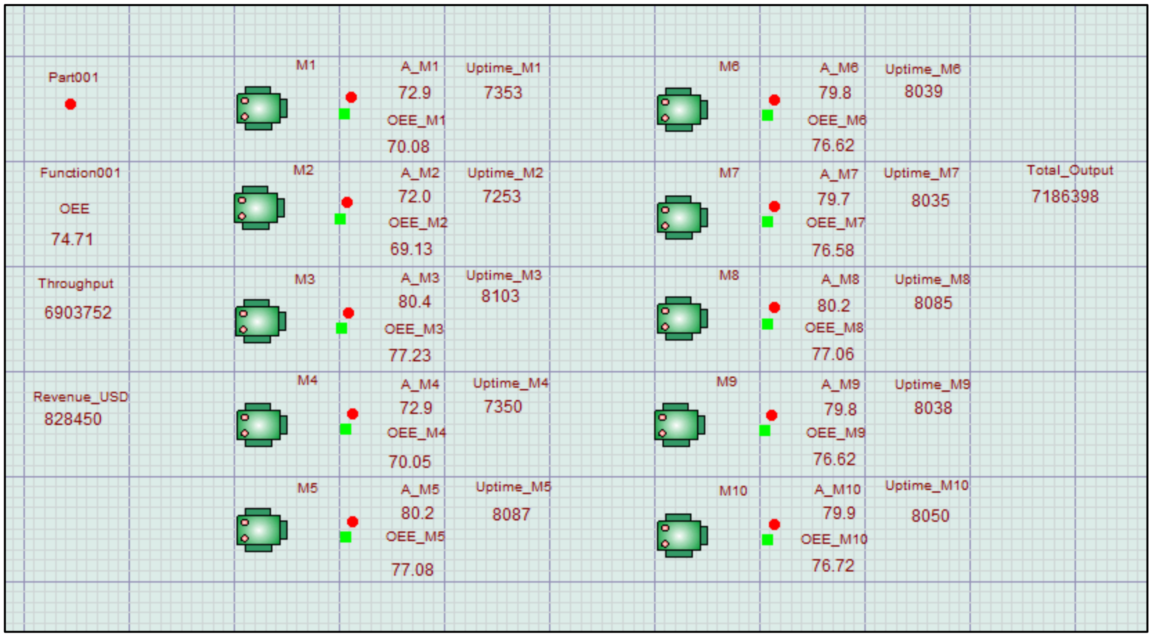


Figure 4.14: Simulation Model After 10080 Minutes (WW46)

Table 4.2 illustrates the difference between the simulated and actual availability of MPI on each machine. The results indicate that the discrepancy between simulated and actual values is very small, with an average deviation of just 0.02%. Therefore, the model is considered valid and can be used as a reliable virtual representation of the MPI testing line for further analysis.

Table 4.2: Comparison of Actual and Simulated Availability for Each MPI (WW46)

MPI Machine	Actual Availability	Simulated Availability	% Difference
M1	72.91%	72.90%	-0.01%
M2	71.69%	72.00%	0.44%
M3	80.50%	80.40%	-0.12%
M4	72.95%	72.90%	-0.07%
M5	80.27%	80.20%	-0.08%
M6	79.74%	79.80%	0.07%
M7	79.74%	79.70%	-0.05%
M8	80.21%	80.20%	-0.01%
M9	79.75%	79.80%	0.06%
M10	79.88%	79.90%	0.03%
Average	77.76%	77.78%	0.02%

4.5 Stage 4: Experimentation & Analysis

Based on the validated simulation model developed in the previous phase, stage 4 focuses on experimentation and analysis to evaluate potential improvement strategies for the MPI test line. In this phase, different scenarios will be studied to analyze their impact on system performance, particularly on availability and overall OEE.

4.5.1 Conversion Time Improvement and Evaluation

In this stage, the validated simulation model was used to evaluate potential improvement strategies aimed at enhancing the performance of the MPI testing line. The model will be reevaluated after an improvement project is conducted by Company X. Based on the findings from Stage 2, the primary factor affecting the overall OEE of the MPI test line is loss of availability, which is mainly attributable to equipment downtime caused by conversion and cold shutdowns. Among the identified factors, the conversion time is easier to optimize than cold shutdowns which require external technical resources and are subject to thermal control constraints, the focus will be on improving conversion time. According to time study of Company X, the traditional conversion process takes an average of approximately 720 minutes as shown in Table 4.3. The analysis indicates that alignment of coaching versus sniper pattern tasks accounts for most of the total conversion time.

Table 4.3: Conventional Conversion Time Study

No	Chuck Replacement Steps	Time (minutes)
1	Draw sniper pattern on current chuck position through TR1 CCD	2
2	Remove the cover and the chuck	48
3	Clean the chuck base to remove the older heat compound	24
4	Apply the new heat compound evenly	40
5	Install the new chuck and the cover	64
6	Levelling check and alignment	22
7	Alignment of coaching vs sniper pattern/chuck mechanical adjustment	320
8	Chuck height teaching and adjustment	80
9	Perform temperature calibration	90
10	Run dummy dies and releases the machine to production	30
	Total Time Taken	**720**

To investigate the issue, a root cause analysis was carried out. The causes were categorized into three groups: Man, Machine, and Method, while Material and Measurement were not considered major factors. Based on these findings, a fishbone diagram was developed as shown in Figure 4.15.

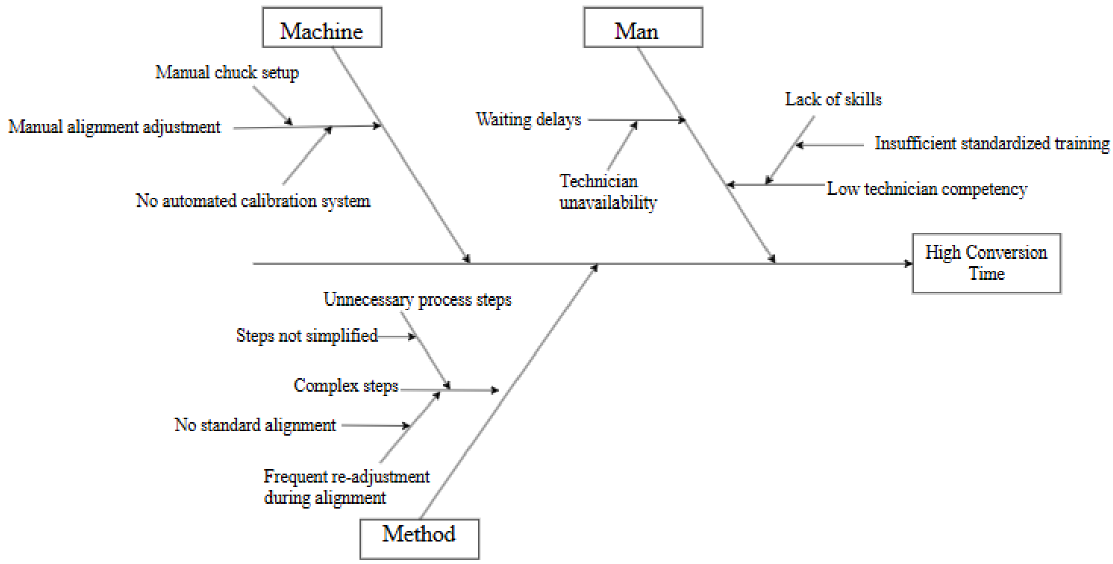


Figure 4.15: Fishbone Diagram

Factors related to the machine indicated that due to the need for manual adjustments and the absence of an automatic calibration system, the alignment step took a long time to complete, requiring multiple fine-tunings, thereby increasing processing time.

The analysis under man indicated that reliance on technicians was a factor contributing to the extended time required for the alignment and adjustment steps. The operators' level of experience, insufficient training, and absenteeism could all lead to delays and variations in the time required to complete these steps.

The method indicated that the design efficiency of the conversion process was low, with complex and unoptimized steps, a lack of standardized alignment processes, and frequent readjustments, all of which directly led to an extended time required for alignment and calibration, as shown by the conversion time study. A brainstorming session was held to generate potential solutions to the problem in Table 4.4.

Table 4.4: Brainstorming of Potential Solutions

No	Brainstorm potential solutions	Advantages	Disadvantages	Potential best solution
1	Increase the efficiency of the technicians.	Shortening the redundant conversion time caused by conversion skill deficiency.	Insignificant impact to overall conversion time (expected <5%).	No
2	Reduce waiting time due to unavailability of technicians.	Shorten response time.	Insignificant impact to overall conversion time (expected <5%).	No
3	Reduce unnecessary conversion steps and manual adjustment/alignment.	Expected significant conversion time reduction (≥30%) by simplifying steps 3-8.	Chuck needles might be out due to software change. Need evaluation.	Yes

With reference to Table 4.4, the company has determined that Solutions 1 and 2 have had little effect in terms of improving technician efficiency and reducing waiting times as these options are expected to improve the overall conversion time by less than 5%, they are not considered to be significant factors contributing to the increase in conversion time. In contrast, Solution 3 aims to eliminate unnecessary conversion steps and reduce the need for manual alignment and adjustment of parts, thereby speeding up the conversion process. This option was selected by Company X as the best solution. It is anticipated that this solution will reduce conversion time by at least 30%. The team decided to apply the Single-Minute Exchange of Dies (SMED) approach and automation to address the problem.

The validated simulation model was then used to evaluate the impact of conversion time reduction on system performance. Moreover, we created several different scenarios in which we reduced conversion time by 15%, 30%, and 45% respectively, while keeping all other established parameters unchanged. As shown in Figures 4.16, 4.17, and 4.18, the simulation results indicate that when the conversion time is reduced, the overall OEE will increase over time. Specifically, the OEE increased from the baseline value of 74.71% to

75.20%, 75.32%, and 75.63% for the 15%, 30%, and 45% reduction scenarios, respectively. These are based on an average of three conversions per week.

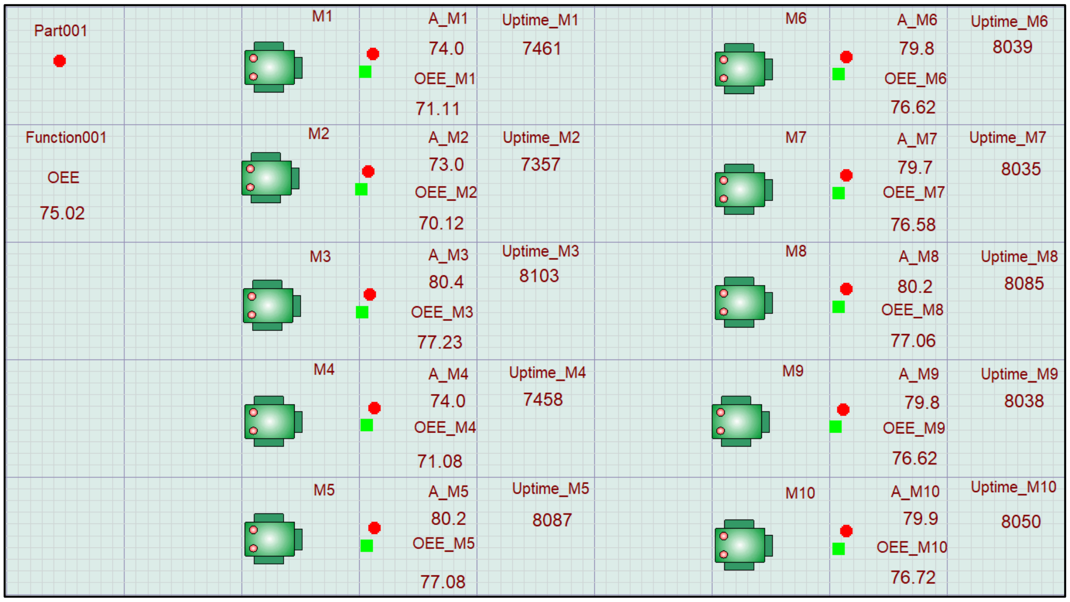


Figure 4.16: Simulation of a 15% Reduction in Conversion Time

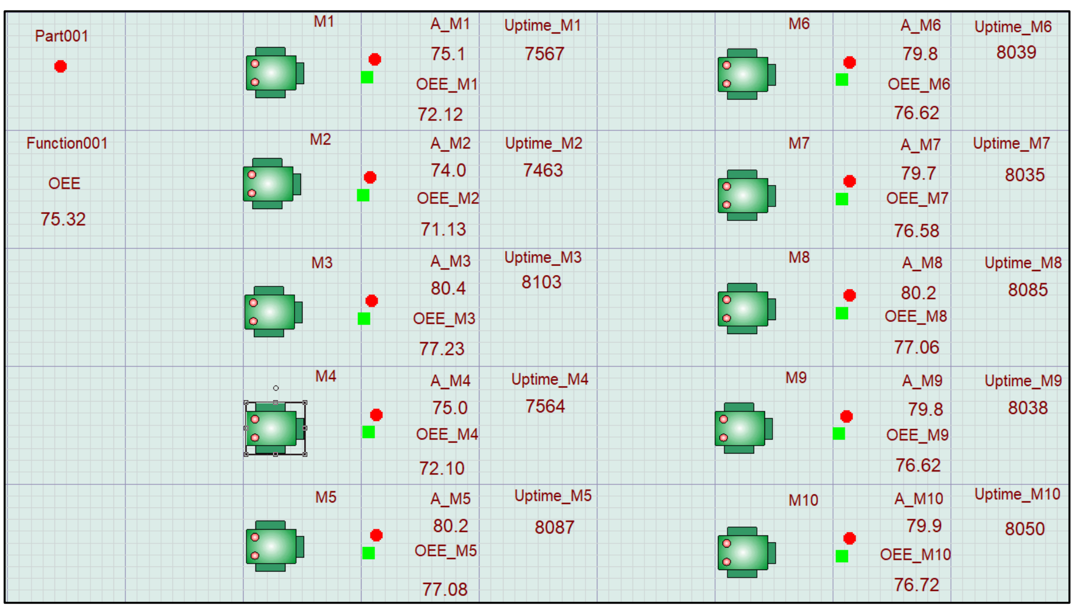


Figure 4.17: Simulation of a 30% Reduction in Conversion Time

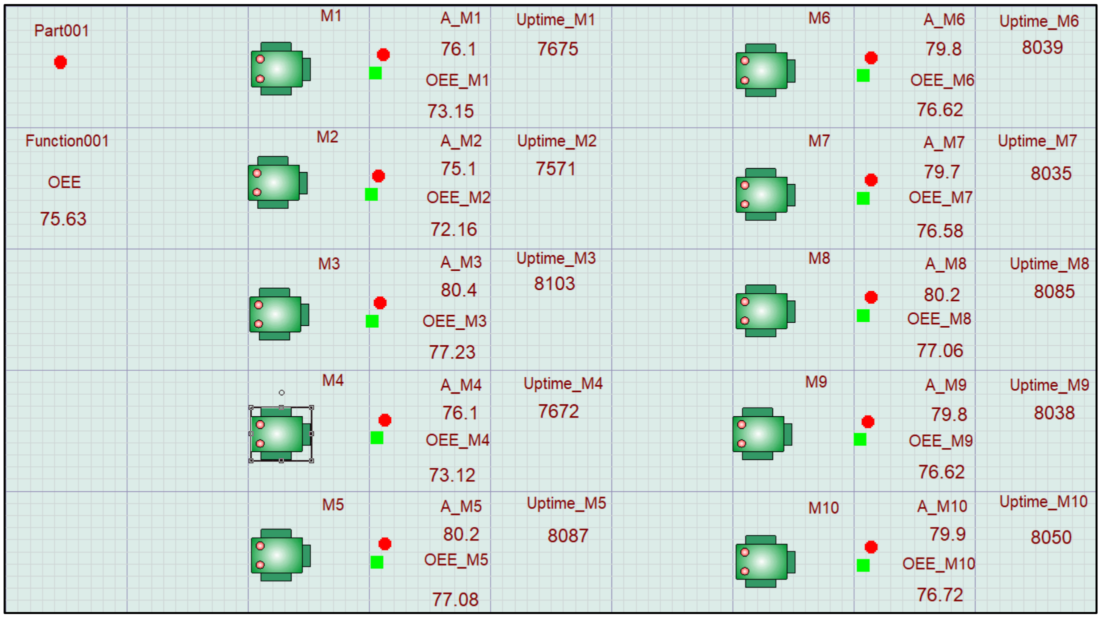


Figure 4.18: Simulation of a 45% Reduction in Conversion Time

The simulation model was used to estimate the potential gains under different improvement scenarios, particularly with respect to the impact of different setup time reductions on overall OEE improvement. The estimated improvements were crucial in prioritizing the most effective solution following the root cause analysis. The simulation results show that once the conversion time is reduced, the OEE performs better over time, allowing team members to see potential performance enhancements and set realistic improvement goals. More importantly, the model supported the estimation of the project payback period based on the potential OEE gain as the project involved external service engineers and experts. The team used the simulation results to estimate the potential revenue gain, prepared a solid technical justification and successfully convince management to greenlight the project.

The chosen improvement plan is solution 3, which includes modifications to the software of the chuck. The testing of the improved software was conducted using two MPI machines. The results of this experiment are satisfactory, with the conversion time successfully reduced from 720 minutes to 390 minutes, representing a 46% reduction.

Additional verification was conducted under actual production conditions, and the updated study results are presented in Table 4.5.

Table 4.5: Conventional Conversion Time Study (After Improvement)

No	Chuck replacement Steps	Average Time (minutes)
1	Draw sniper pattern on current chuck position through TR1 CCD.	2
2	Remove the cover and the chuck.	48
3	Install the new chuck and the cover.	64
4	Levelling check and alignment.	22
5	Alignment coaching vs sniper pattern (EZ chuck).	14
6	Alignment by EZ chuck software.	40
7	Chuck height teaching and adjustment.	80
8	Perform temperature calibration.	90
9	Run dummy and release the machine to production.	30
	Total Average Time	**390**

In order to assess the effectiveness of the proposed changes, the simulation model was re-evaluated after the improvement in conversion time. Data on unplanned downtime for each MPI for WW02 was categorized according to the types of downtime errors to facilitate further analysis of availability losses as shown in Figure 4.19, while the cold shutdown time for each MPI in the same workweek was as shown in Figure 4.20.

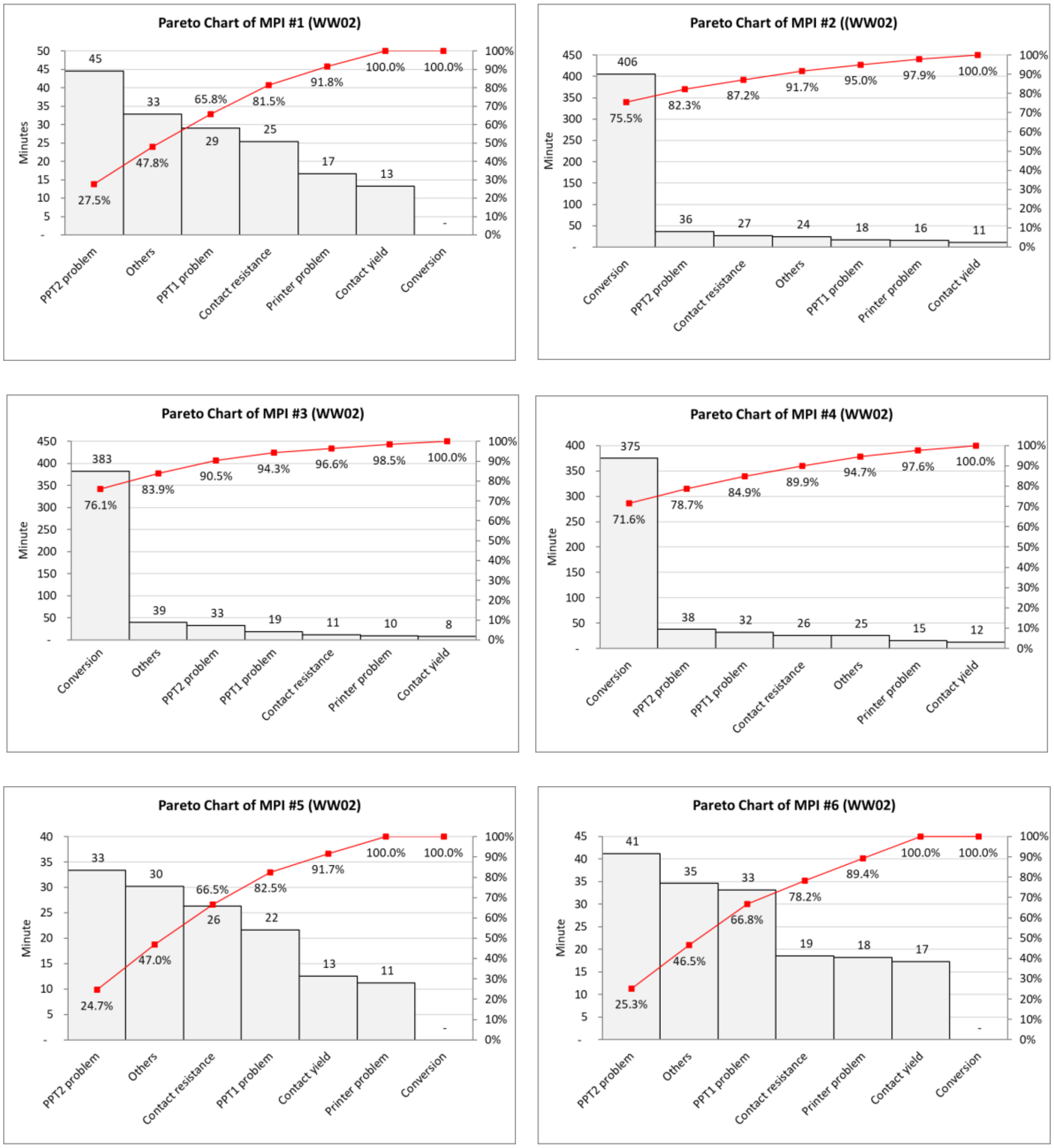


Figure 4.19: Unplanned Downtime by Error Type for Each MPI (WW02)

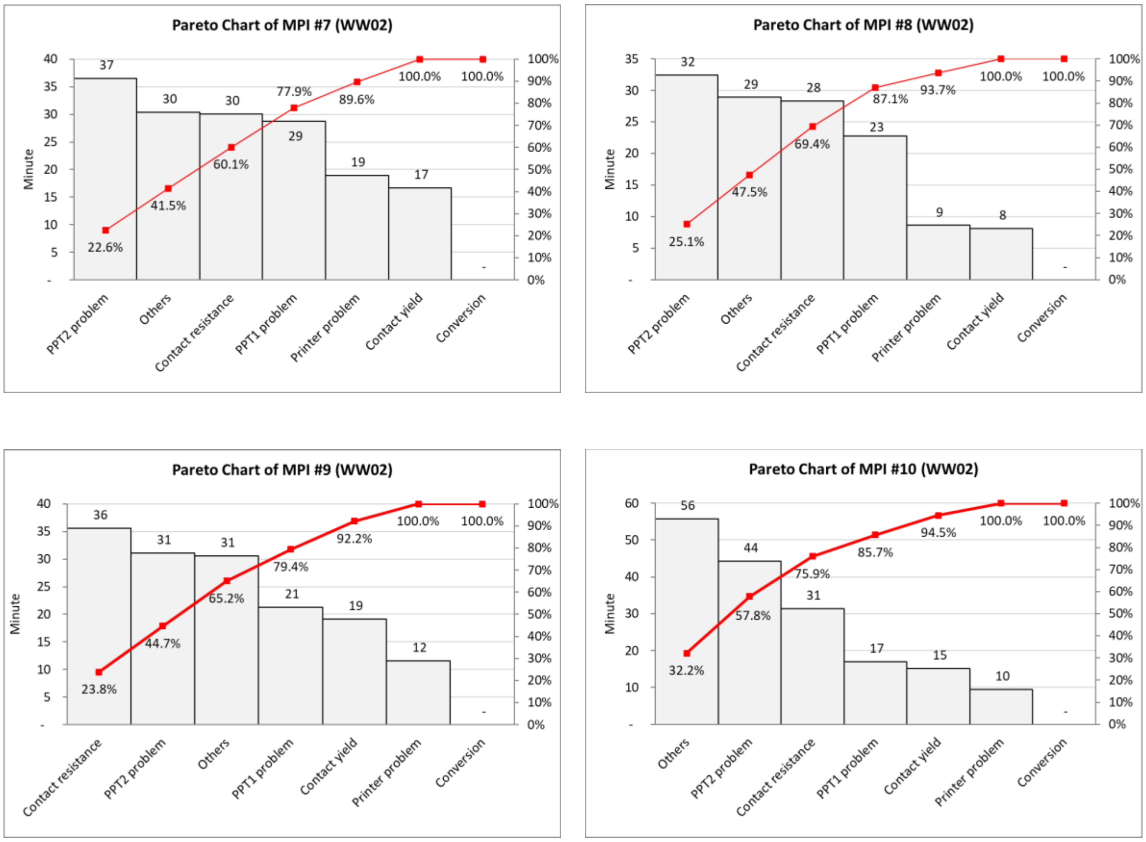


Figure 4.19 (continued): Unplanned Downtime by Error Type for Each MPI (WW02)

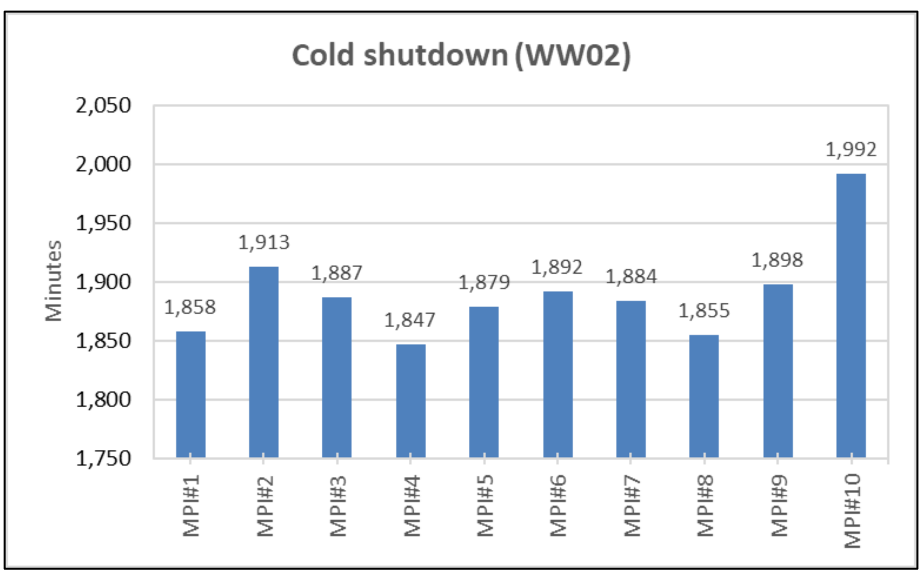


Figure 4.20: Cold Shutdown Time for Each MPI (WW02)

Figure 4.21 shows the performance indicators of WW02 including availability, performance, quality, and OEE.

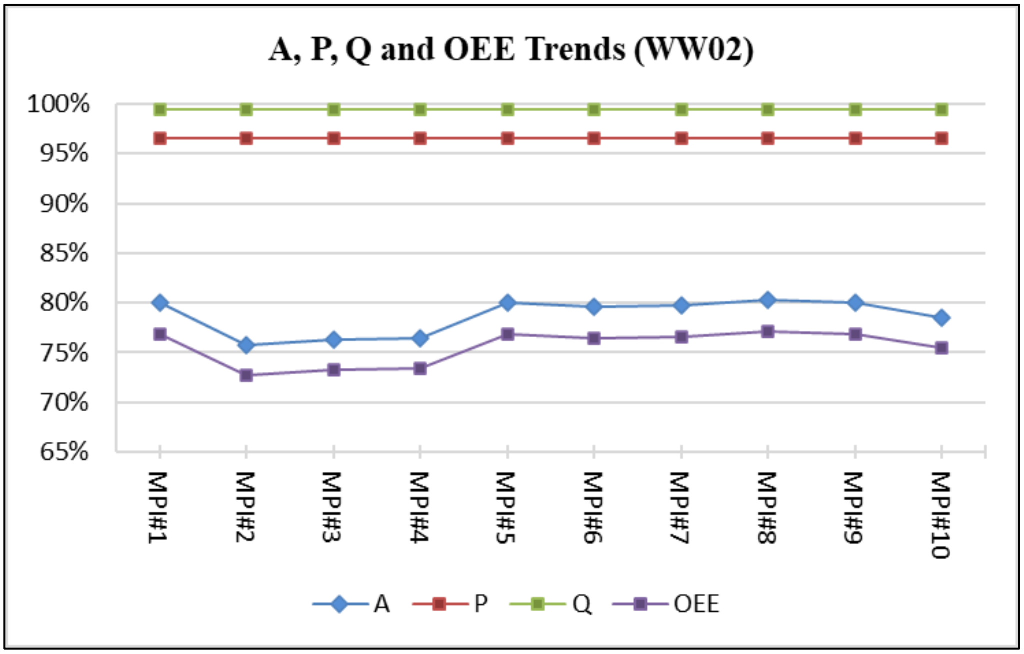


Figure 4.21: A, P, Q and OEE Trends for Each MPI (WW02)

Compared to the baseline results from WW46, it highlights the significant improvement in availability and the overall OEE after the reduction in conversion time. The downtime data for major errors and cold shutdowns in WW02 were inserted into WITNESS and the simulation results are shown in Figure 4.22.

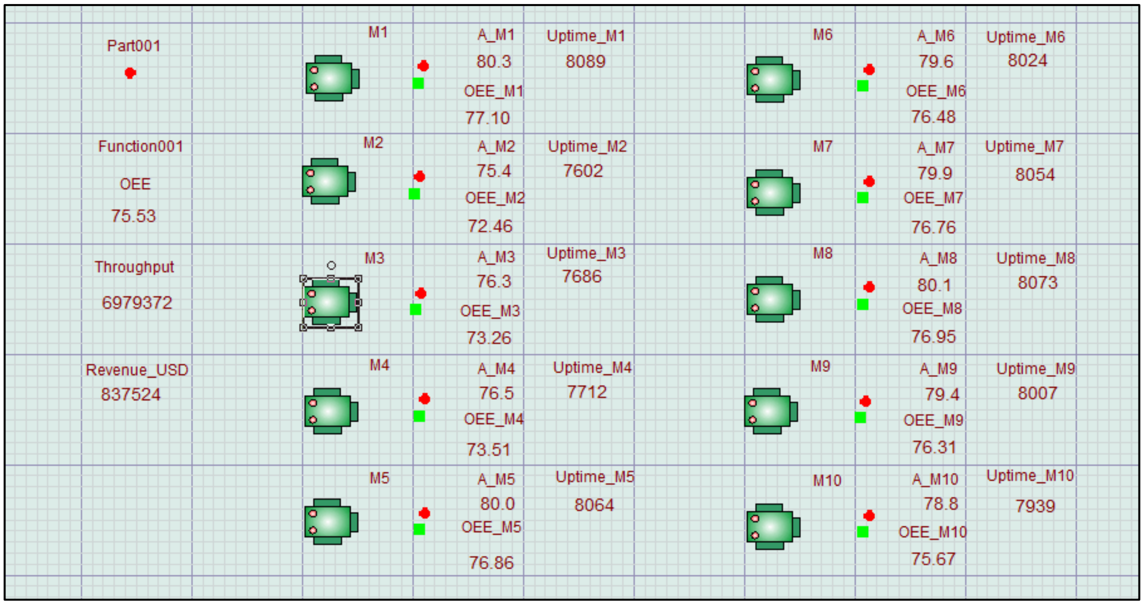


Figure 4.22: Simulation Results (WW02)

The simulated results for availability (A) were compared with the actual production data, as summarized in Table 4.6. The comparison shows that the deviation between simulated and actual values is minimal, ranging from –0.80% to 0.42%, with an average difference of only –0.04%. In addition, the calculated OEE was 75.57% and the simulated OEE was 75.53%, indicating only a negligible difference.

Table 4.6: Comparison of Actual and Simulated Availability for Each MPI (WW02)

MPI Machine	**Actual Availability**	**Simulated Availability**	**% Difference**
M1	79.96%	80.30%	0.42%
M2	75.69%	75.40%	-0.39%
M3	76.29%	76.30%	0.01%
M4	76.48%	76.50%	0.02%
M5	80.02%	80.00%	-0.02%
M6	79.61%	79.60%	-0.02%
M7	79.71%	79.90%	0.24%
M8	80.32%	80.10%	-0.27%
M9	80.04%	79.40%	-0.80%
M10	78.52%	78.80%	0.36%
Average	78.66%	78.63%	-0.04%

In summary, the reevaluation of the model after the conversion improvement in WW02 shows that there are only slight differences between the calculated and simulated availability (A) and overall OEE values. The findings are consistent with the similar comparison to WW46. Hence, it was validated that the developed model was reliable for the simulation of the MPI line.

4.5.2 Scenario Analysis

Following validating and evaluating the simulation model, it was used to simulate a series of operational scenarios to assess the system's performance under different conditions. These scenarios include the evaluation of throughput, revenue, and machine utilization, as well as the impact of variations in product type, cycle time, and performance (P) and quality (Q) rates on the overall performance of the MPI testing line. The simulation results presented here indicate that the model can analyze operational conditions and support decision-making by predicting the impact of different types of improvement strategies and system changes.

Scenario 1-Throughput and Revenue Forecast

Simulations were conducted to evaluate the relationship between OEE, throughput, and revenue for the MPI testing line serving Customer Y. The cost per tested die was assumed to be USD 0.12. Figure 4.23 presents the system performance under ideal conditions, while Figures 4.24 shows the corresponding throughput and revenue at OEE levels of 74.71% (before improvement) and 75.53% (after improvement). The results indicate that both throughput and revenue increase with higher OEE values. This can be attributed to the improvement in equipment efficiency, which reduces downtime, increases productive running time, and allows for more equipment within the same time. Therefore, even a relatively small improvement in OEE can significantly increase overall output and final revenue, thereby highlighting the sensitivity of high-output production systems to changes in availability.

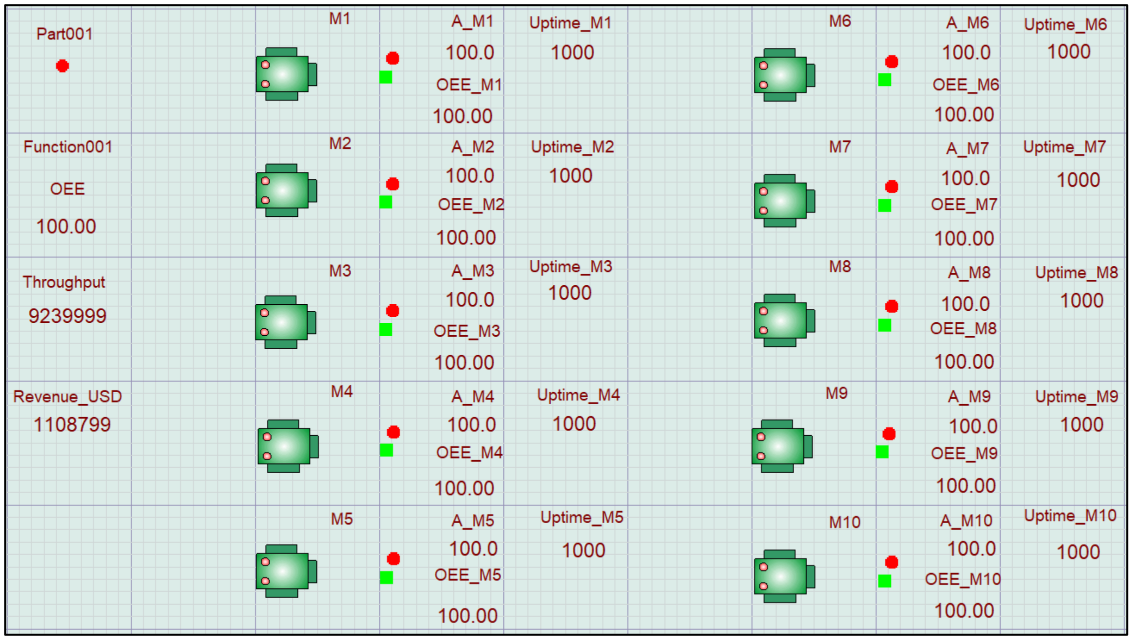


Figure 4.23: Throughput and Revenue of the MPI line under Ideal Conditions

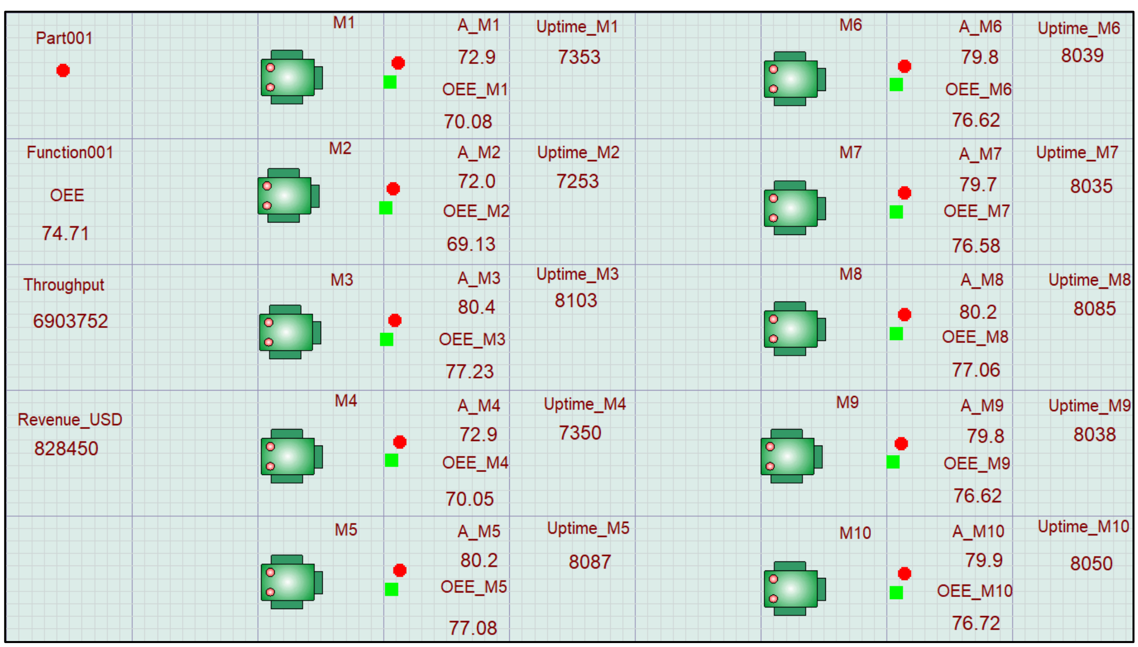


Figure 4.24: Throughput and revenue of the MPI line at OEE 74.71% and 75.53%

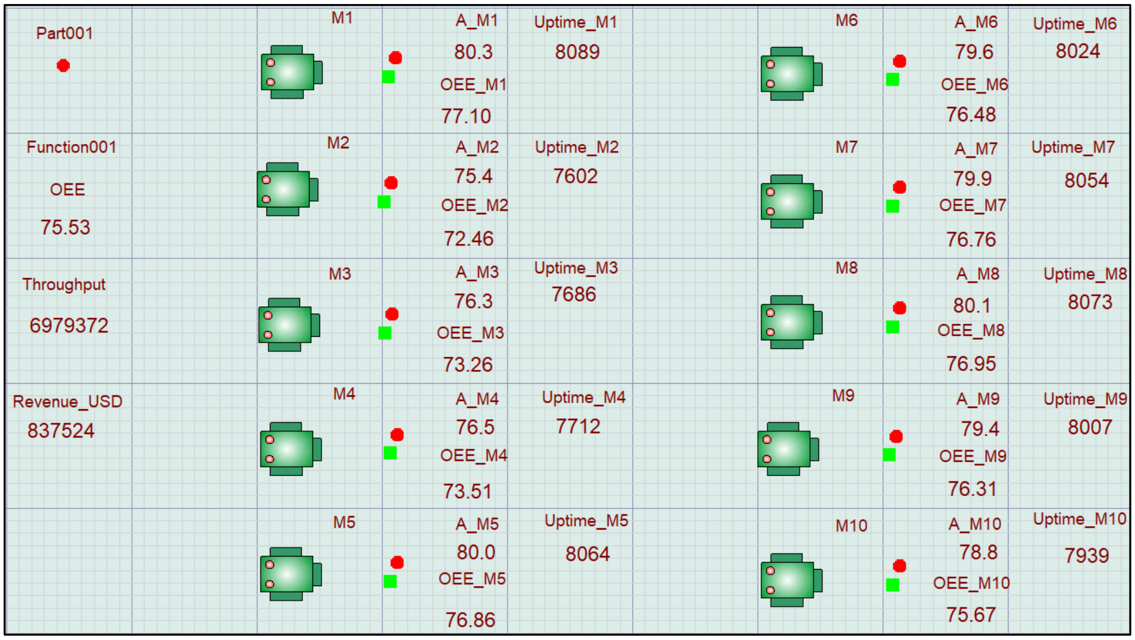


Figure 4.24 (continued): Throughput and revenue of the MPI line at OEE 74.71% and 75.53%

Figures 4.25 and Figure 4.26 illustrate both throughput and revenue increase proportionally with OEE, indicating a strong positive linear relationship. This is because a higher OEE means better equipment performance, especially in reducing downtime, and will increase the available production time. As a result, more units will be produced during this period, resulting in greater throughput and more revenue. This explains the consistent linear trend observed in throughput and revenue. Since revenue directly depends on the number of units produced, these two variables exhibit similar patterns. This indicates that OEE can be a reliable means of predicting the production quantity and output value of a single MPI machine.

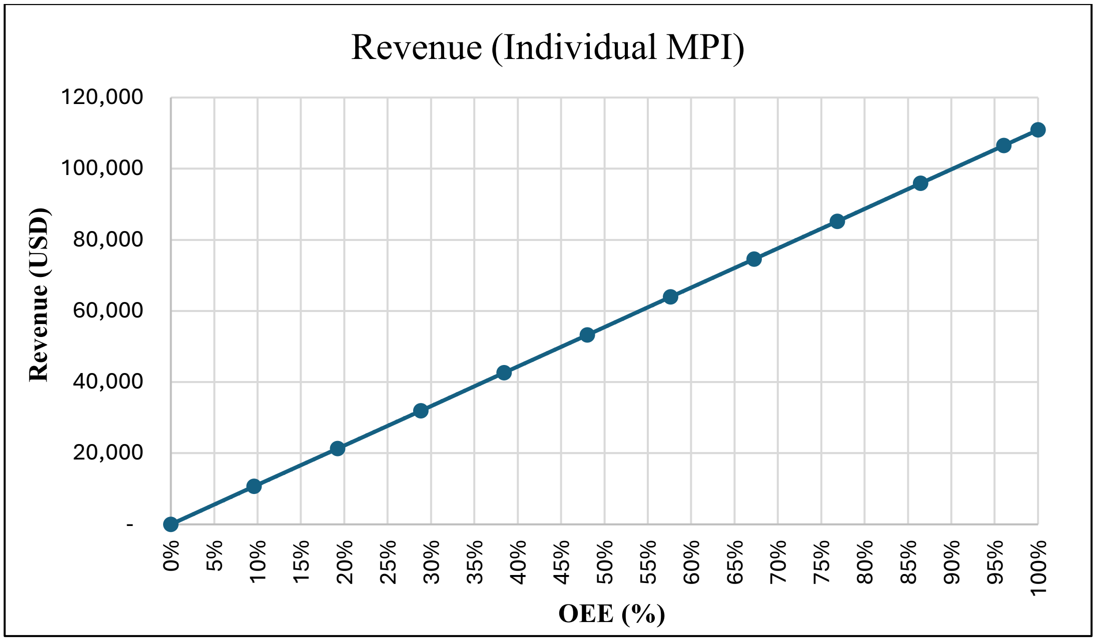


Figure 4.25: Relationship between Revenue and OEE for an Individual MPI Machine

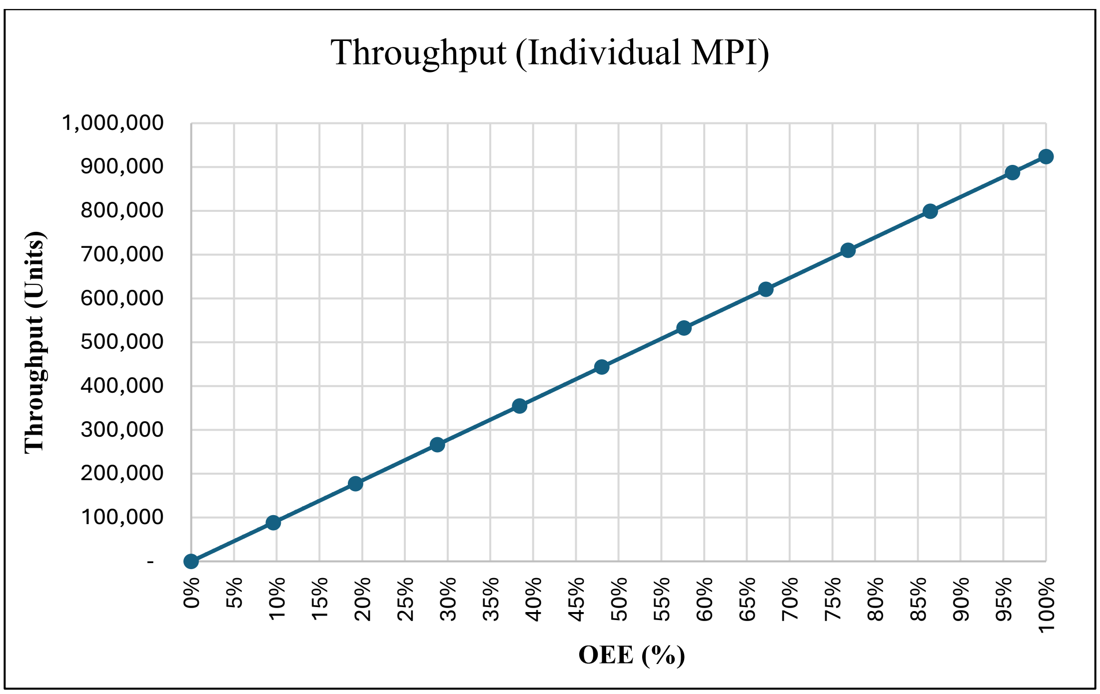


Figure 4.26: Relationship between Throughput and OEE for an Individual MPI Machine

Furthermore, this relationship is also consistent at the system level. As shown in Figure 4.27, the overall output and revenue of the MPI production line increase proportionally with OEE. This indicates that the improvement in equipment efficiency not only enhances the performance of individual machines but also significantly boosts the production efficiency and financial performance of the entire production line. This relationship is valuable for forecasting the potential benefits of OEE improvement initiatives.

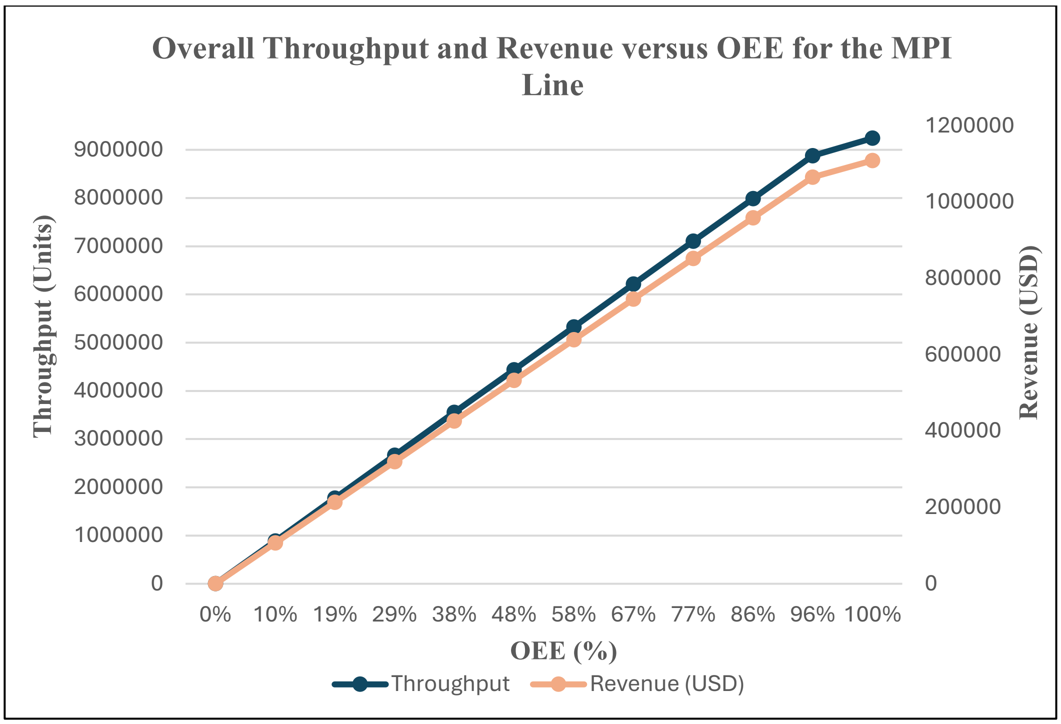
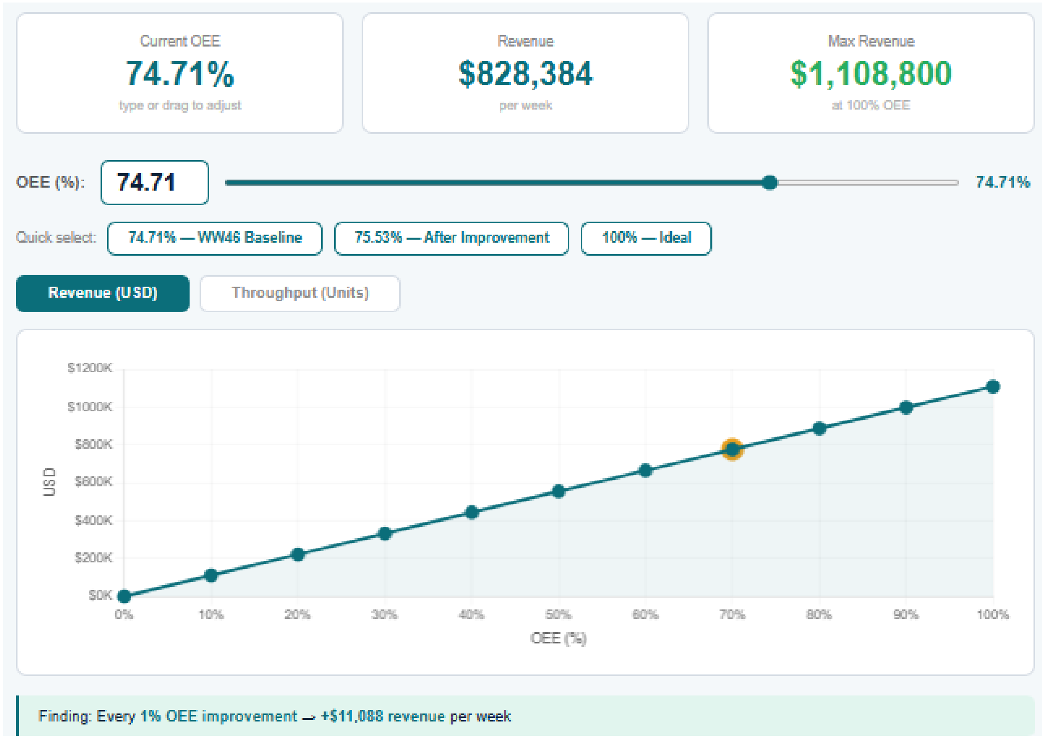


Figure 4.27: Overall Throughput and Revenue versus OEE for the MPI Line

Based on these charts, engineers and management can clearly visualize the relationship between OEE, throughput, and revenue. At the operational level, OEE is a very useful metric for measuring equipment in terms of efficiency and financial performance. The improvement of OEE is the result of three key performance indicators: availability (A), quality (Q), and performance (P). P and Q will remain relatively stable within the system, therefore, the main driver of OEE changes will be improvements in the availability metric. In this case, any improvement in availability should significantly increase output and revenue. This will enable the organization to identify and prioritize improvement opportunities and reduce downtime. To enhance this analysis, an interactive tool was created to demonstrate the potential outcomes of the relationship between OEE, throughput, and revenue. As shown in Figure 4.28, this tool allows users to adjust the OEE value to reflect its impact on throughput and revenue.



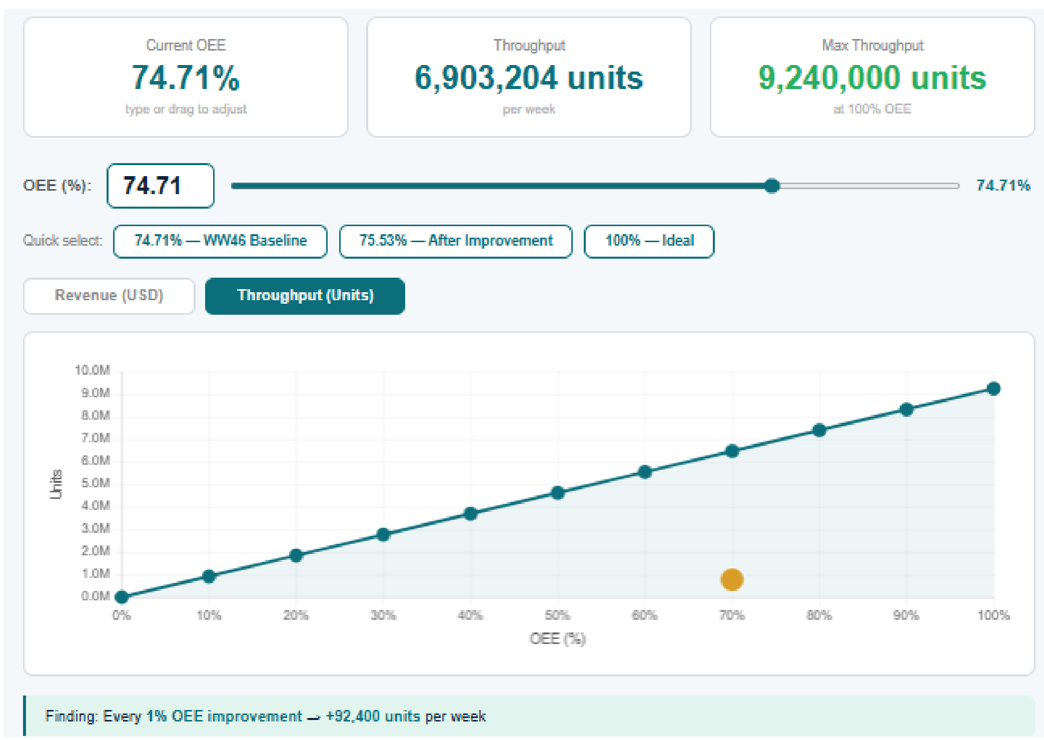


Figure 4.28: Interactive Visualization of OEE, Throughput and Revenue for the MPI Line

This visualization supports the decision-making process by allowing improvement teams to evaluate how different OEE levels affect the overall production of the organization. In addition, it demonstrates how an increase in equipment effectiveness can lead to increased revenue through the ability to produce more products. Based on the results from the simulation, the team will have a practical reference to help them in determining which improvement actions should be given the highest priority and how to best optimize total production capabilities of the organization.

Scenario 2- Product Mix and Allocation Forecast

The developed simulation model was used to study different machine allocation scenarios for testing multiple products for Customer Y using a total of 10 MPI machines. This scenario focuses on assessing how variations in cycle time, downtime characteristics, and testing cost influence overall system performance in terms of throughput, OEE, and revenue. When multiple products are processed on the MPI line, machine allocation becomes a critical decision variable that directly affects both production efficiency and financial performance. In this study, two products with distinct characteristics were considered, as summarized in Table 4.7. Product A has a shorter cycle time, enabling faster processing, whereas Product B has a longer cycle time but generates higher revenue per unit due to its higher testing cost.

Table 4.7: Product Mix Scenario

Customer Y	**Cycle Time (min)**	**Testing Cost (USD)**
Product A	0.0109	0.12
Product B	0.0120	0.15

Different distribution scenarios were simulated to assess the effect of machine allocation. Figures 4.29 show the baseline case, where five MPI machines are allocated to produce product A and five MPI machines are allocated to produce product B. This provided a balance point to compare other alternatives for distribution and their impact.

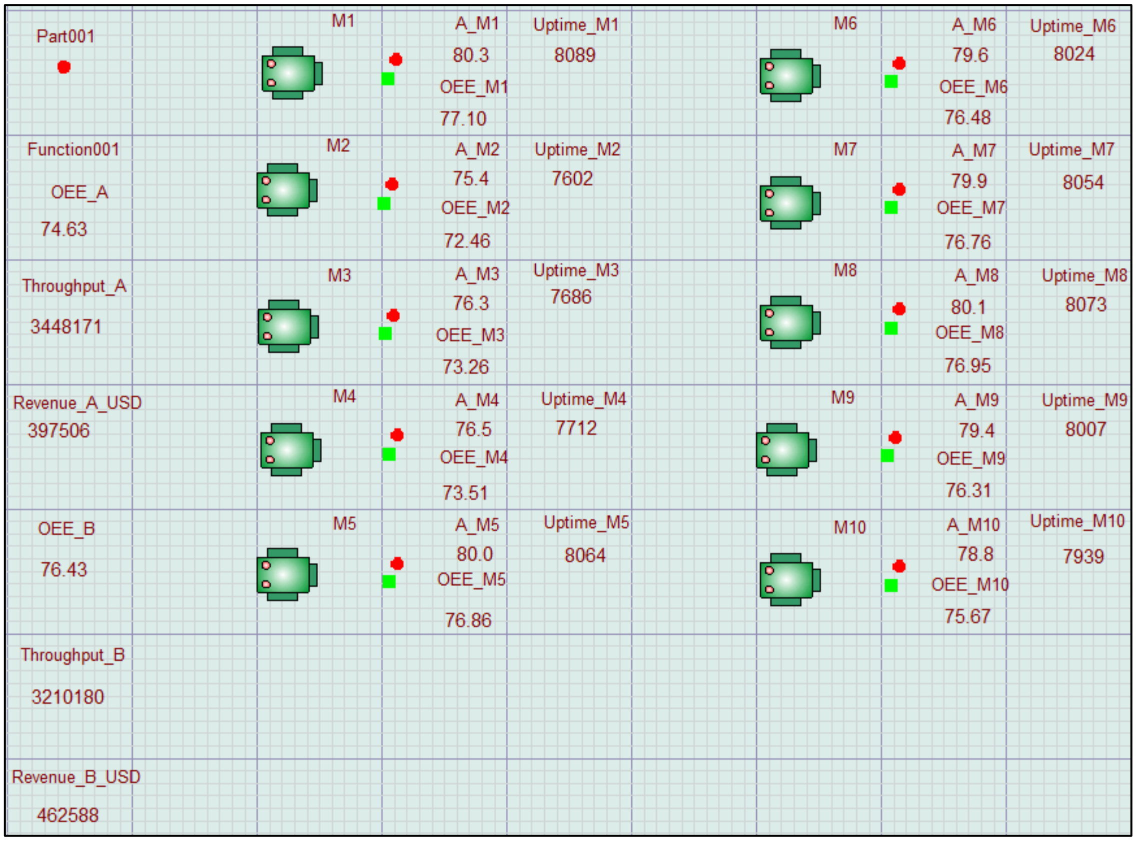


Figure 4.29: MPI Product Allocation 5A:5B

To further investigate allocation sensitivity, additional scenarios were analyzed. Figure 4.30 provides an example of a distribution scenario where an organization is allocated more machines for Product A rather than Product B (7A:3B). Conversely, Figure 4.31 provides an analysis of the opposite scenario, where product B is allocated 7 machines, while product A only receives 3 machines (3A:7B).

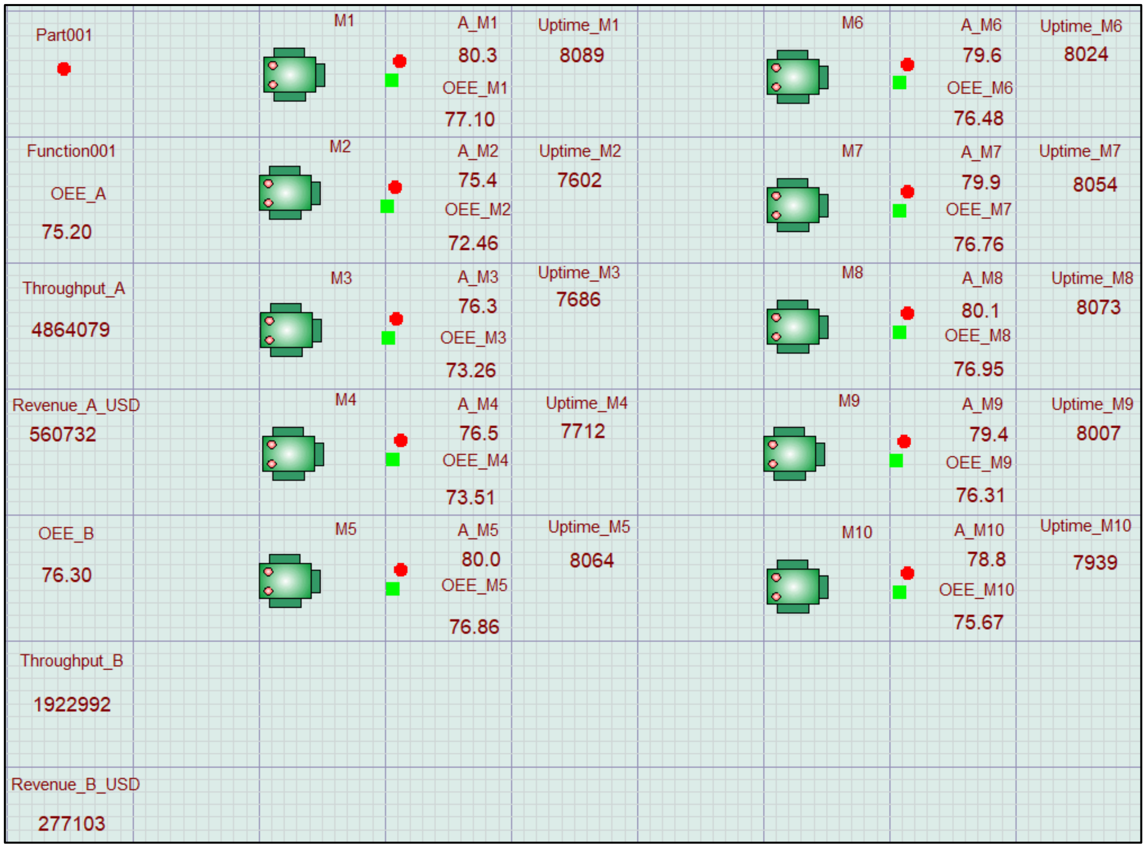


Figure 4.30: MPI Product Allocation 7A:3B

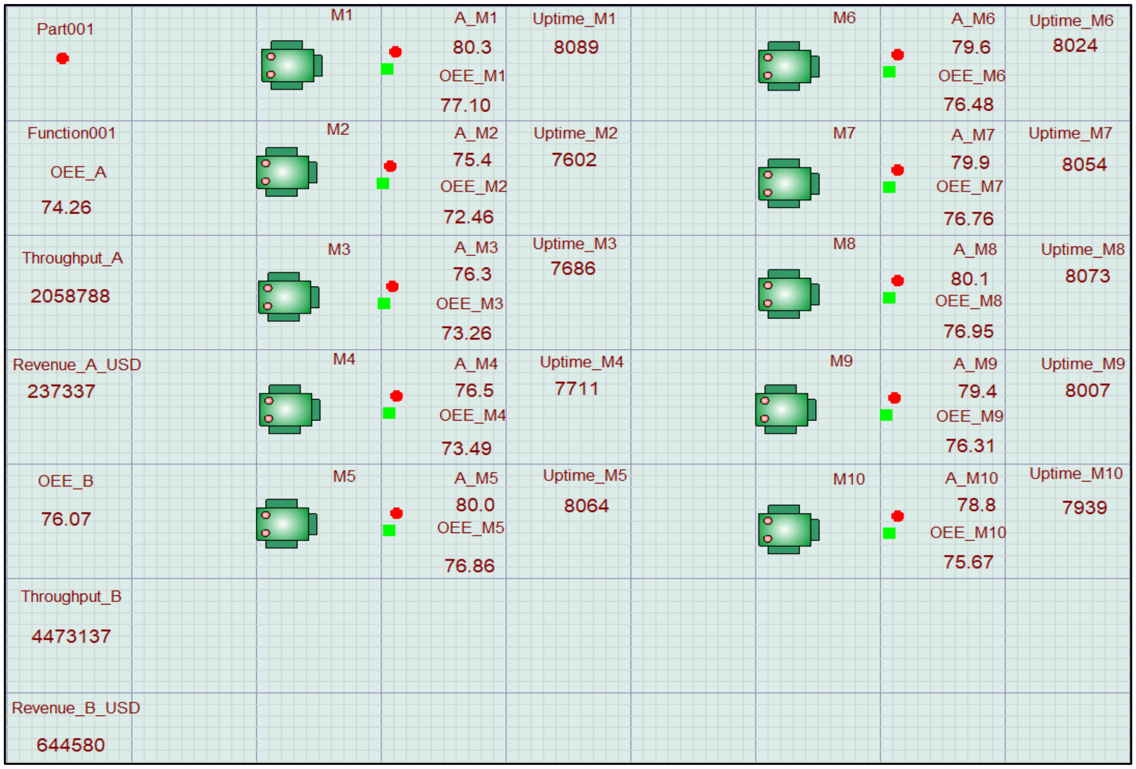


Figure 4.31: MPI Product Allocation 3A:7B

The overall results of all scenarios have been presented in Table 4.8. The results of these example scenarios clearly indicate that, given the shorter cycle time for producing product A, increasing the machine allocation for product A will achieve higher output than increasing the machine allocation for product B. However, the financial return from the unit sales output of product B is higher than that of product A. This highlights a clear trade-off between maximizing production output and maximizing revenue. While Product A improves production efficiency, Product B contributes more significantly to financial performance. Slight variations in OEE were observed across the different allocation scenarios due to differences in downtime characteristics between the two products. This indicates that machine allocation decisions should consider not only cycle time and testing cost, but also the impact of downtime on overall system performance.

Table 4.8: Simulation Results under Different MPI Machine Allocation Scenarios

Scenario	Machine Allocation (A:B)	Throughput A (Units)	Throughput B (Units)	Total Throughput (Units)	Revenue A (USD)	Revenue B (USD)	Total Revenue (USD)
1	5A : 5B	3,448,171	3,210,180	6,658,351	397,506	462,588	860,094
2	7A : 3B	4,864,079	1,922,992	**6,787,071**	560,732	277,103	837,835
3	3A : 7B	2,058,788	4,473,137	6,531,925	237,337	644,580	**881,917**

In addition, based on the analysis of the simulation results, an interactive visualization tool was created, allowing users to dynamically change machine allocation and OEE to see how these changes affect overall yield and revenue, as shown in Figure 4.32, without repeatedly executing the simulation model.

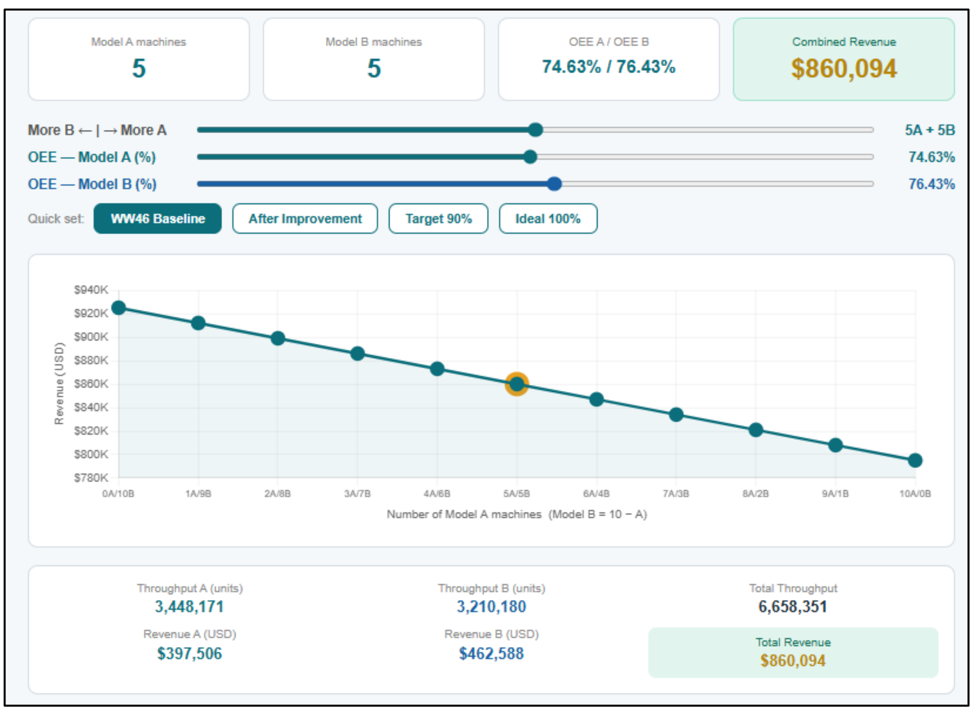


Figure 4.32: Interactive Visualization of Machine Allocation, OEE, and Combined Revenue

This visualization shows how machine allocation affects revenue performance. Based on the differences in the number of machines allocated to each product, there are significant variations in revenue performance, highlighting the sensitivity of financial performance to the method of machine allocation. Therefore, it enables production planners to easily identify the machine allocation configuration that maximizes revenue based on the conditions of machine usage.

In summary, this visualization also demonstrates the effectiveness of simulation models and interactive visualizations as decision-making tools for product portfolio planning. It enables production planners to analyze trade-offs and determine the best machine allocation strategy based on operational efficiency and financial goals.

Scenario 3- Cycle Time, Performance, and Quality Improvement Forecast

The developed simulation model was further utilized to evaluate the combined effects of performance (P) and quality (Q) improvements on system performance. This scenario aims to quantify how incremental changes in operational efficiency influence OEE, throughput, and revenue under a fixed product mix condition. Using the product allocation scenario of 5A:5B from Scenario 2 as the baseline, a 1% increase in performance for both Product A and Product B was first simulated, as illustrated in Figure 4.33.

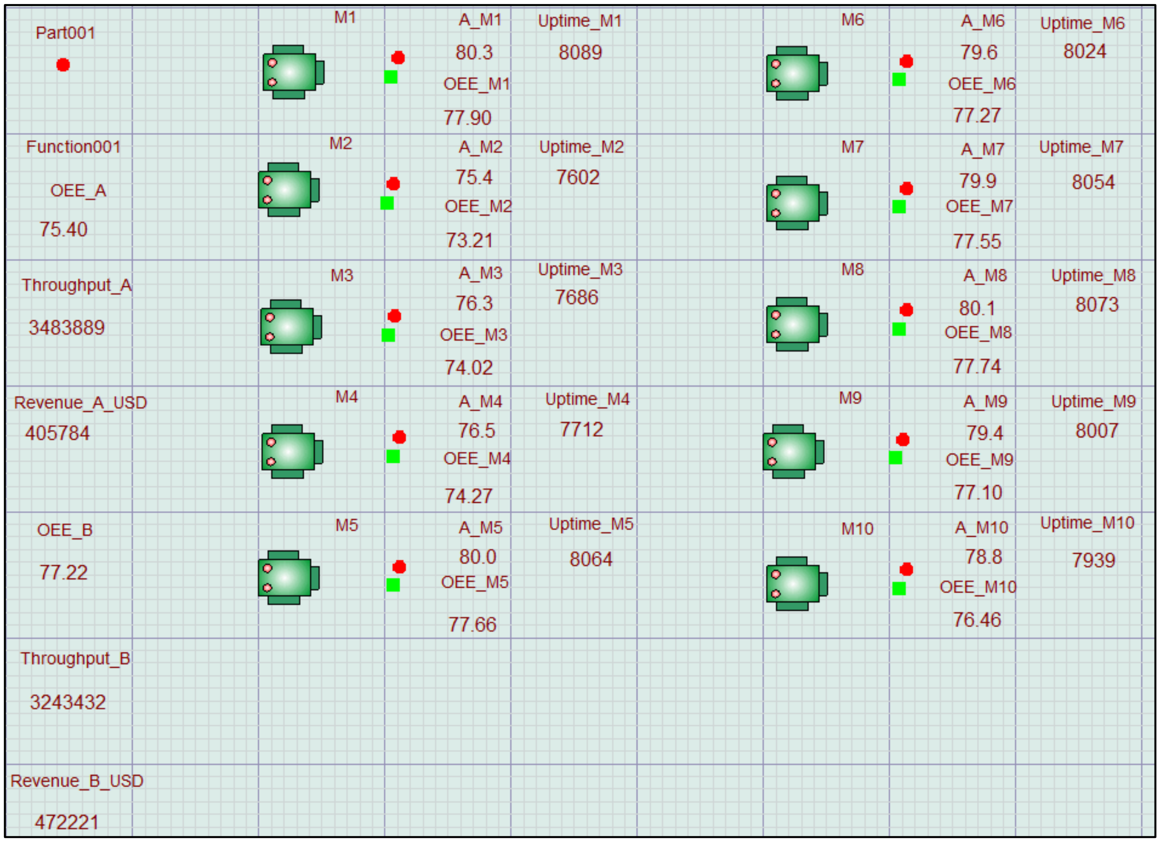


Figure 4.33: Simulation of 1% Performance Increase Using MPI 5A:5B Model

The results indicate that the proportional improvement in system performance is a direct result of the enhancement in system performance. Compared to the baseline scenario, OEE of product A improved from 74.63% to 75.40% with throughput improving from 3,448,171 to 3,483,889 units and overall revenue increasing from \$397,506 to \$405,784. Product B also showed similar results in terms of OEE, increasing from 76.43% to 77.22%, with production rising from 3,210,180 units to 3,243,432 units, and revenue increasing from \$462,588 to \$472,221. These data indicate that even minor improvements in performance can lead to significant increases in output and revenue.

To further evaluate system sensitivity, a mixed scenario combining performance variations with quality variations was also evaluated to understand the sensitivity of the system. In terms of performance, Product A was reduced by 1% while Product B was improved by 1%. In terms of quality, Product A had an increase of 0.1%, and Product B decreased by 0.1%. The simulation results for this scenario are presented in Figure 4.34.

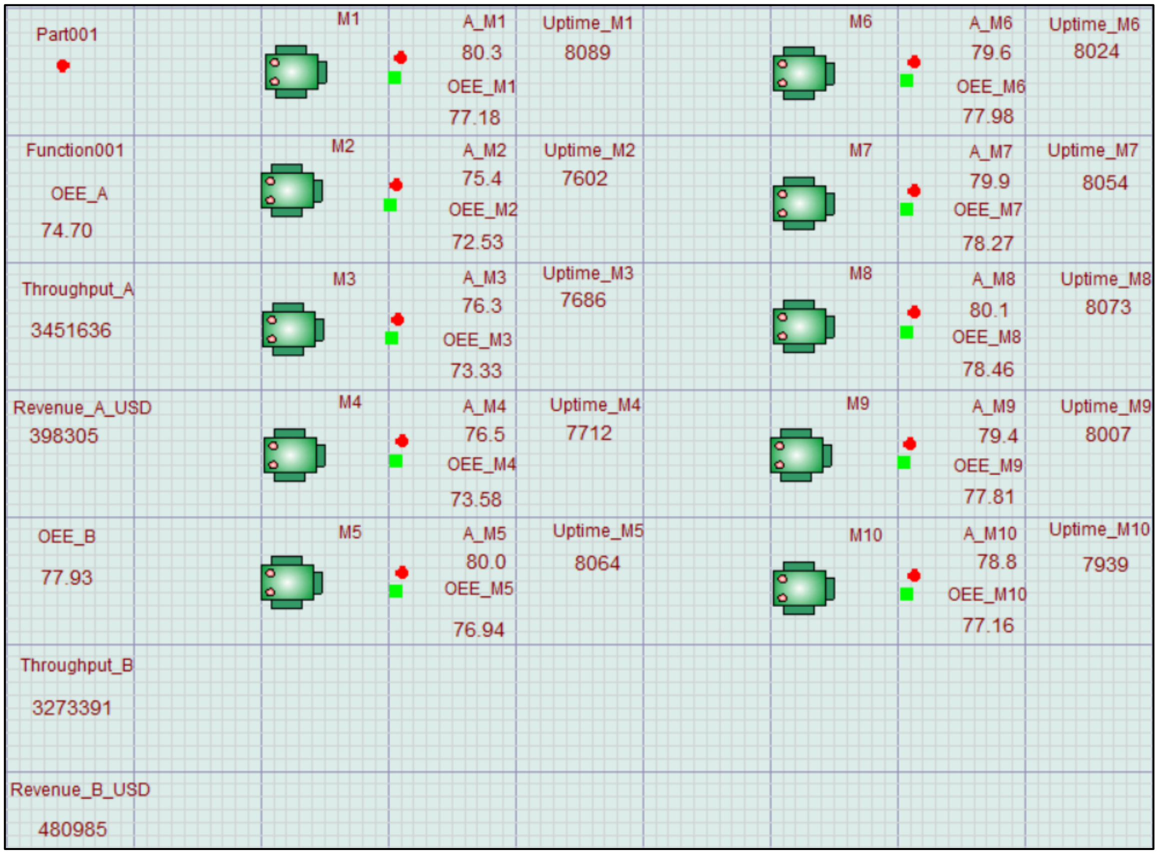


Figure 4.34: Simulation of Performance and Quality Changes for Products A and B

The simulations confirmed the role of performance and quality in determining the overall behaviour of the system. In the mixed scenario, by making changes to both performance and quality, it was possible to see results as evident from the measures of OEE, throughput, and revenue for Product A and Product B. For Product A, the OEE has increased slightly to 74.70%, with throughput of 3,451,636 units, and revenue of USD 398,305. Product B, on the other hand, saw a significant improvement with an OEE of 77.93%, throughput of 3,273,391 units, and revenue of USD 480,985. The results further suggest that small changes in performance and quality (±1% and ±0.1%) will produce measurable changes in overall system output. The improvements observed in Product B offset the slight differences in Product A and indicate that the system overall has an output affected by the combination of the multiple product streams compared to the performance of just one individual product. Based on these findings, it appears that performance improvements will tend to have a larger impact on the system output compared to minor variations in quality under the given conditions. This is because P directly affects the speed of operation of the machine, while the quality will provide a small reduction in the effective output of the machine. Consequently, when optimizing production processes, P and Q should be assessed together, as they have mutual effects on OEE, throughput, and ultimately the revenue generated from those products.

Furthermore, interactive visualization tools were developed based on simulation results to support operational analysis and decision-making. Figure 4.35 illustrates the relationship between performance, cycle time and throughput. As performance increases, cycle time decreases, resulting in a corresponding increase in throughput. This inverse relationship indicates that higher machine efficiency enables more units to be processed within the same production period. Figure 4.36 presents the integrated relationship between Performance, Quality, OEE, throughput and revenue, allowing users to evaluate how adjustments in P and Q simultaneously affect overall system performance. These visualization tools provide an effective approach for conducting sensitivity analysis, enabling decision makers to assess operational trade-offs and performance improvement strategies without the need to re-run the simulation model repeatedly

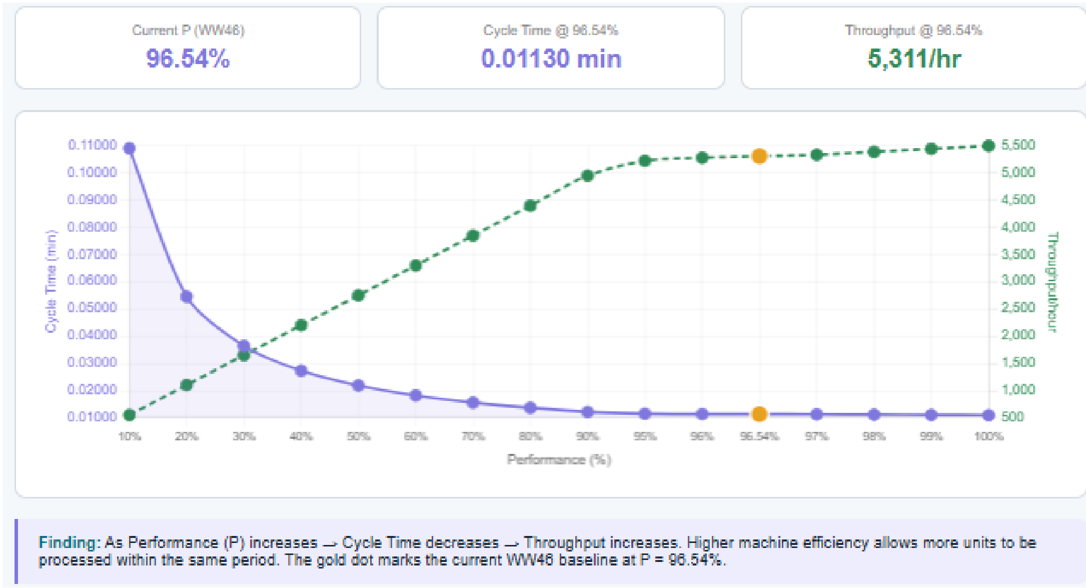


Figure 4.35: Cycle Time and Throughput Interactive Chart

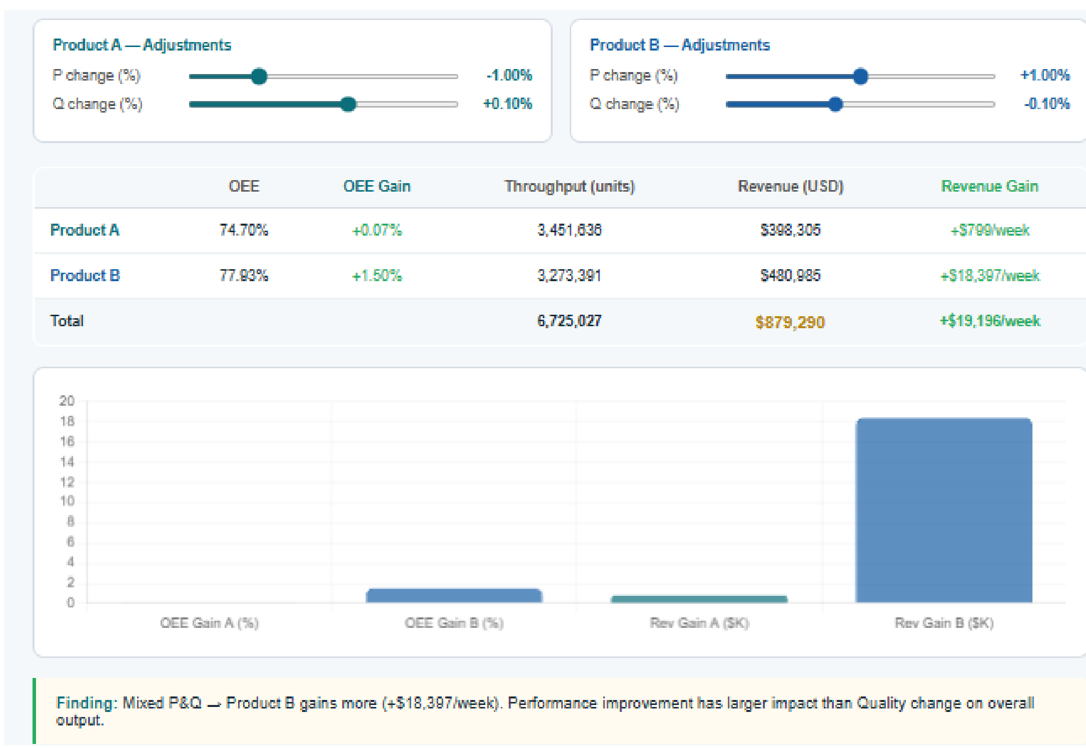


Figure 4.36: Interactive Simulator for Performance and Quality Adjustment with OEE, Throughput and Revenue Impact

Scenario 4- Model Replication and Scalability Forecast

The simulation model was adapted to expand its capabilities by utilizing the MPI configuration in a larger production environment. As shown in Figure 4.37, the new simulated manufacturing environment is more realistic and complex than the original simulation content and includes multiple MPI machines. The replication of machines allows the model to analyze the performance of the entire system such as throughput and revenue when the number of machines is increased. By scaling from a single-machine production and multi-machine configurations in the model, the scalability of production capacity and the system performance under increased production loads can be evaluated.

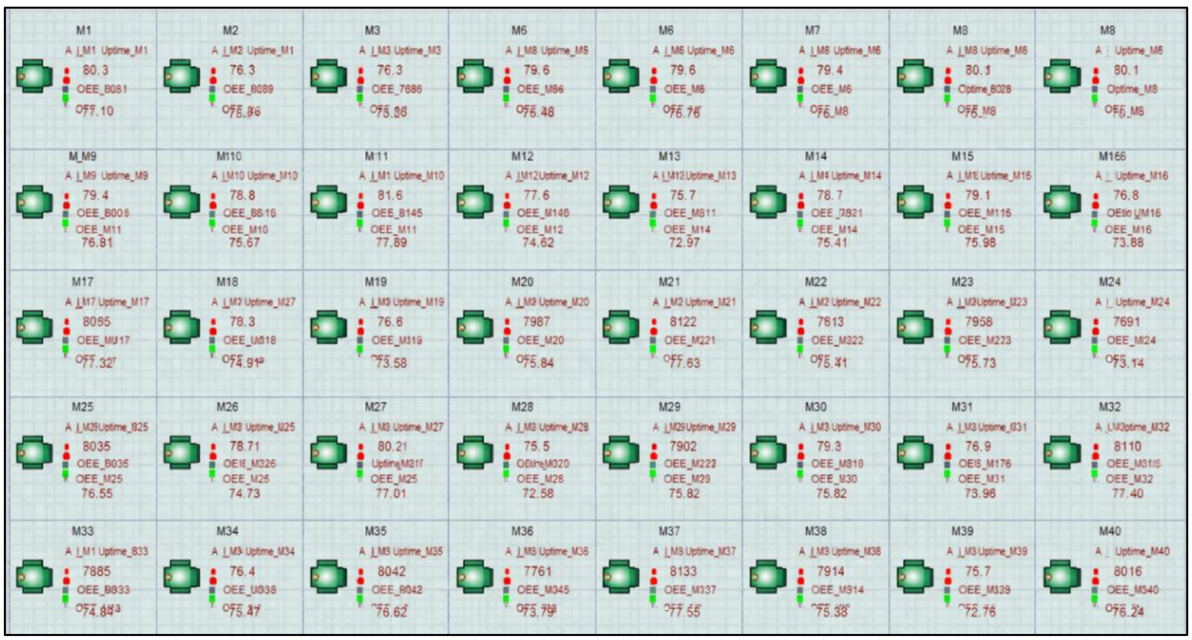


Figure 4.37: Scalability of the Simulation Model

In addition, the model can be extended beyond two products to simulate multiple products with different cycle times, testing costs and demand levels. This enables the production team to evaluate machine allocation strategies across different product types and prioritize products based on their revenue and throughput contributions. The model also serves as a performance evaluation tool that assesses the impact of changes in quality, performance and cycle time at various levels including individual machines, selected product groups or the entire production line. This provides a more comprehensive perspective on how improvement strategies affect overall system performance under different operating conditions. To further support scalability analysis, an AI-based multi-machine simulator was developed, as shown in Figure 4.38, enabling simulation of up to 43 MPI machines with individually configurable parameters for each machine.

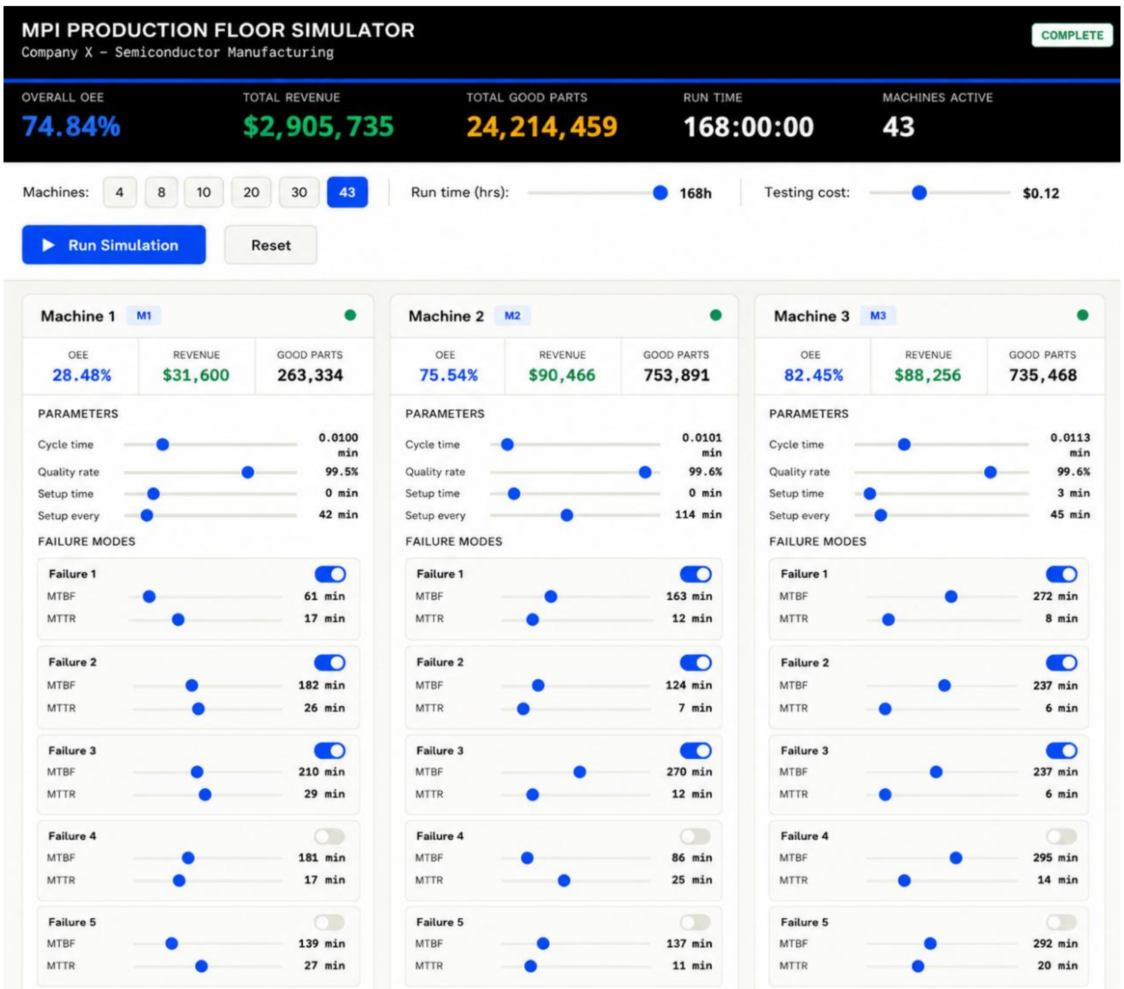


Figure 4.38: AI Multi-Machine Simulator

The AI multi-machine simulator introduced to the production team was much simpler and more user-friendly compared to some simulation software, where they could easily apply it by adjusting sliders such as cycle time, quality rate, setup time, MTBF and

MTTR according to their data. The team could not only use it to simulate the line quickly but could also add more functions as it progressed to further improve the tool or create new tools on their own.

In summary, the scalability of this model enhances its application in current operations and future planning. This means it can support decision-making in the following scenarios such as adding new machines to operations, introducing new products to the product line, or expanding production capacity. Therefore, this model will enable management to leverage the impact of machine allocation, process improvement activities, and investment planning, while estimating the effects of these activities on throughput and revenue.

4.5 Stage 5: Termination

The simulation model was run over a fixed production time that matched the actual manufacturing data used in this study, with each simulation run representing 7 days of production time. Therefore, ensuring consistency between the simulated and real production environments. The simulation was terminated upon completion of the defined production period, allowing the collected results to represent the steady-state behaviour of the system.

To provide the simulation with greater reliability, four measurement variables were tracked through multiple simulation runs. The variables include availability, OEE, throughput, and revenue. Multiple simulation runs were conducted when required to validate the reliability and consistency of results and reduce the influence of variability in the simulation model.

The created simulation model is a decision-support tool for evaluating various strategies prior to application. Certain improvements such as reduction in conversion time, allocation of machines by product, modifications to performance and quality can be evaluated within the model for estimating the effect on system performance and profitability.

However, the implementation of these improvements within the actual manufacturing environment may present challenges such as operator adaptation to changes in processes and inter-department coordination. Thus, it is suggested that the company follows a phased approach for the implementation of these improvements. The initial testing of pilot programs at a select group of MPI machines will provide validation of expected improvements prior to full-scale implementation, and thereby minimize the risks associated with operational improvement.

The termination criteria adopted in this study ensure that the simulation outputs are statistically stable and suitable for comparison with actual manufacturing data. The validated results provide a quantitative basis for evaluating system performance and assessing the impact of different operational scenarios, particularly in terms of downtime reduction, performance improvement, and production planning optimization.

The simulation model can also be modified and repurposed for future analysis based on new production data to facilitate ongoing evaluation of new alternative scenarios for the organization and thereby provide support for long-term future decisions.

4.6 Findings and Discussion

The findings of this study show that combining discrete event simulation with OEE provides a practical and reliable way to analyze and improve production performance in the MPI testing line. By using actual downtime data from the MPI test line, the developed model helps represent the actual behaviour of the manufacturing system and allows for accurate evaluation of improvement measures.

The validation results indicate that the simulated performance metrics closely match the actual performance. The difference in availability between the model and actual data is negligible, with an average deviation of less than 0.05%, while the simulated OEE closely matches the actual OEE values. This shows that the model is reliable and suitable for further analysis. A well-validated model is important because it ensures that the results from different scenarios can be trusted for decision-making.

The primary finding from this study is that availability is the most significant factor affecting OEE performance on the MPI testing line. The results of analysis in Stage 2 indicate that both performance and quality are relatively steady, whereas availability is considerably lower than the benchmarked value described in the literature, which is approximately 90% for world-class performance (Nakajima, 1988). This indicates that downtime is the primary source of inefficiency. Cold shutdown and conversion processes contribute significantly to availability losses. These results are consistent with existing research which argues that equipment downtime is a critical factor limiting productivity in the semiconductor manufacturing environment (Huang et al., 2003). The research demonstrates that making improvement in availability will have the most significant positive impact on overall system performance.

The results from Scenario 1 establish a clear and strong positive relationship between OEE, throughput, and revenue. As OEE increases, both throughput and revenue increase proportionally due to the reduction in downtime and the increase in effective production time. This finding extends the conventional use of OEE beyond a performance measurement tool by demonstrating its capability as a predictor of production output and financial performance. In a high-volume manufacturing environment such as the MPI testing line, even a small improvement in OEE can result in substantial gains in production output and revenue. This highlights the economic significance of operational efficiency improvements.

Scenario 2 demonstrates the trade-offs between production efficiency and financial performance when selecting product mixes and allocating machines. For example, if more machines are allocated to Product A, which has a shorter cycle time, the overall output of the factory will increase, however, if more machines are allocated to Product B, which has higher testing costs, the overall revenue of the factory will also increase. This demonstrates that optimal decisions on the manufacturing line it is necessary to find a balance between processing efficiency and economic returns. The findings highlight that production planning should not rely solely on throughput maximization but must also consider profitability and system-level performance.

Scenario 3 shows that performance has a greater impact on system output compared to small variations in quality. This is because performance directly affects processing speed and cycle time, while quality mainly influences yield. The results also demonstrate that the interaction between P and Q leads to combined effects on OEE, throughput, and revenue. Therefore, these parameters should be evaluated collectively rather than independently when optimizing system performance. This provides a deeper understanding of the internal dynamics of OEE components and their relative influence on production outcomes.

The scalability analysis in Scenario 4 demonstrates that the developed simulation model is flexible and can be extended to represent more complex production environments. By replicating machines and incorporating additional products, the model can be used for capacity planning, resource allocation, and long-term production strategy. Beyond the practical applications of the model demonstrated in this study, it suggests that the model should be considered for use as a strategic decision-making tool in the semiconductor industry.

Moreover, the integration of interactive visualization tools with simulation outputs represents another valuable contribution to this study. These tools enable users to explore the dynamic relationships between key variables such as OEE, throughput, and revenue without the need to run multiple simulations, thereby improving analytical efficiency and providing faster and more intuitive decision-making. This will be particularly beneficial in industrial environments, as decisions are time-sensitive and require speed.

Despite these findings, the analysis is based on specific production data and simplified assumptions, such as constant performance and quality rates, which may not fully capture real manufacturing variability. These limitations are further discussed in Chapter 5.

CHAPTER 5

CONCLUSION

5.0 Introduction

This chapter discusses the summary of this research, as well as limitations and recommendations for future research. In Section 5.1 summarize the primary findings of the analysis and simulation results. Section 5.2 will describe the limitations of the study and Section 5.3 will suggest areas of future research that would potentially improve the adoption and usability of the simulation model.

5.1 Summary of Research

This study analyzed production performance in an automated die testing line, specifically the MPI line, at Company X by integrating downtime data with a DES model developed using WITNESS This study aimed to analyze and categorize equipment downtime, develop and validate a WITNESS-based simulation model using OEE as a benchmark, and evaluate different operational scenarios with the support of AI tools to enhance production performance and decision-making. Downtime data were extracted from the MPI tracking system, processed in Excel, and categorized into major and minor error types to simplify model development while maintaining data integrity. OEE was used to develop the model, and three key performance indicators (KPIs), namely availability (A), performance (P), and quality (Q) were used as part of the model development process.

The analysis of the production data collected over a three-week period showed that both performance and quality were stable, with average values of 96.54% performance and average values of 99.5% quality. These values were verified by engineers from the MPI line and were consistent with historical records. However, availability was significantly lower than the benchmark value of approximately 90%, indicating that downtime is the main factor limiting system performance. In particular, cold shutdown and conversion processes were identified as the major contributors to availability loss.

A simulation model representing 10 MPI machines dedicated to Customer Y was successfully developed, with each machine characterized by its own downtime behaviour. The simulated results were then compared to the calculated results, and the validation results showed that the simulated OEE of 74.71% closely matched the calculated OEE of 74.70%. The model was validated by comparing simulated results with actual production data, showing only slight differences, with an average deviation of approximately 0.02% in availability.

A second simulation was conducted after a conversion improvement project was implemented in the MPI line. The conversion time was reduced by approximately 46%, resulting in an improvement in both Availability and OEE. The model was revalidated using the new actual production data from the post-improvement period. The simulated and actual results remained closely aligned, with percentage differences ranging from –0.80% to 0.42%, further confirming the robustness and accuracy of the model. In addition, the simulated OEE of 75.53% was compared against the calculated OEE of 75.57%, indicating only a negligible difference between the two values. This further verified that the model remained an accurate representation of the MPI line in terms of OEE and availability after the improvement was implemented.

The validated model was then applied to simulate several operational scenarios:

(1) Throughput and Revenue Forecast

By tracking the weekly throughput trend, revenue projections can be made and the relationship between revenue and downtime fluctuations can be determined. This enabled the engineers and managers to establish a current performance baseline against improved throughput and potential revenue gain, linking operational performance to financial outcomes.

(2) Product Mix and Allocation Forecast

The model was used to simulate throughput differences between two products with different cycle times. This provided insights into throughput per product, revenue per product, and the effective utilization of the MPI line. The results from this simulation method support decision-making in determining the optimal product mix under fixed capacity constraints.

(3) Cycle Time, Performance, and Quality Improvement Forecast

The model was used to investigate the impact of cycle time reduction and quality improvements on throughput, and to estimate the potential OEE improvement apart from availability improvements.

(4) Model Replication and Scalability Forecast

The model can be replicated and expanded to include more MPI machines to simulate a more complex testing environment. This helps estimate and predict the impact of different product types across multiple customers on overall line or performance. It also helps determine the capacity utilization limits and the future expansion of the MPI line.

In addition, the developed model was provided to the production team at the MPI line in Company X. The model helped the production team to estimate throughput and revenue based on changes in availability losses, enabling them to predict the outcomes and impacts of their improvement projects. This also supported better planning and prioritization of these projects.

The AI tools developed from simulation results enabled the team to quickly analyze and visualize the impact of equipment downtime, performance rate, quality rate, and product mix. This helped save time by reducing the need to run simulation model multiple times. The ability of the developed WITNESS simulation model and AI tools to quickly evaluate different scenarios helped improve the team in planning and prioritizing improvement activities based on the best estimated outcomes in terms of throughput and revenue. The team responded positively to the model and AI tools as they found them useful and practical in supporting data analysis and decision-making.

From an academic perspective, this study demonstrates that conventional OEE remains a practical and effective performance metric without requiring modification. OEE has undergone significant evolution and expansion beyond its origins as a measure of individual equipment performance. This evolution has been largely driven by the shortcomings and limitations that emerged during its application across various industries over the years. Today, OEE varies from one industry to another based on different applications and approaches. With the basis of measuring effectiveness derived from the original concept, OEE has been customized to fit specific industrial requirements (Ng Corrales et al., 2020).

The extension of conventional OEE has resulted in the modification of its formulation, leading to the adoption of different terms in both literature and industrial practice. With the inclusion of more performance elements, components and loss classifications, the application scope of modified OEE has been broadened. However, the introduction of new performance elements moves the metric away from the standard OEE definition. Several modified versions have been proposed in the literature, including the

Overall Equipment Effectiveness of a Manufacturing Line by Braglia et al. (2008), the incorporation of yield as an additional OEE loss factor by Raja and Kannan (2007), the introduction of monetary losses by Wudhikarn (2012) and the inclusion of material losses by Braglia et al. (2018). These modifications have altered the original OEE definition to varying degrees. As a result, the new metrics deviate from the conventional OEE formulation and pose challenges to its standardized acceptance in the industry. Furthermore, the change of time base in these modified metrics deviates from the standard OEE definition (Gibbons and Burgess, 2010).

The simulation models developed in this study show that OEE can be linked to throughput and revenue, beyond its primary function of measuring Availability, Performance and Quality. Moreover, OEE can be extended to evaluate overall line performance under different product mixes and across different types of equipment. The results also demonstrate that one of the main advantages of conventional OEE, specifically its function as a performance benchmark, remains applicable and practical in real-world manufacturing environments. For instance, for production lines with similar equipment configurations, the models can adopt the best-performing OEE value as an improvement target for other comparable lines. These findings collectively indicate that conventional OEE, without modification or the addition of new elements, remains one of the most reliable and practical indicators for both simulation-based analysis and real-world manufacturing decision-making.

Revisiting the objectives of this study, the project has successfully achieved all of them.

1. To analyze and categorize equipment downtime data in a semiconductor manufacturing line.

2. To develop a WITNESS-based simulation model to evaluate downtime reduction scenarios.

3. To validate the model by aligning simulation outcomes with actual equipment performance using OEE as the benchmark.

4. To use the simulation model to evaluate improvements, particularly in terms of throughput and revenue, under different product testing environments to provide a basis for the development of useful AI tools and decision-making.

In summary, the real implications of the project are as follows:

(1) The developed model was used to evaluate the payback period of the conversion time reduction project, enabling the team to develop a financial justification for its implementation, with a conversion time reduction of 46% and cost savings of approximately USD 10,000 per week based on three conversions per week in the MPI line.

(2) The simulation results enabled the development of useful AI tools for the production team without the need for simulation software that would incur additional costs. These tools offer a cost-effective alternative for the team to facilitate analysis and decision-making. For businesses and companies to stay competitive and survive today, AI is a must, not an option, and Company X will benefit from using these AI tools as a start.

(3) Academically, the models prove that OEE remains a powerful equipment performance indicator along with its benchmarking capability, even without the introduction of new elements or modifications.

This study demonstrates that integrating DES with OEE provides a systematic and data-driven approach for evaluating manufacturing performance. It enables the identification of key areas for improvement and supports more effective decision-making in semiconductor production systems. The project introduced MPI engineers to applying and building their own AI tools, initiating AI adoption in line with global trends.

5.2 Limitations of the Study

The study acknowledges that there are limitations to the tracking system. First, company X provided much of the dataset used for developing the simulation, and some of this data was filtered for confidentiality purposes. As a result, it was limited access to detailed raw data, therefore, the extent of validation and the ability to perform more comprehensive analysis.

Second, the simulation model represents a standalone automated MPI testing process and does not incorporate any buffers, interaction with labor, and the upstream and downstream processes that occur in a manufacturing facility. Therefore, the model does not necessarily represent all the complexities associated with a typical integrated production system.

There were also several assumptions made when developing the model, such as constant performance (P) and quality (Q) rates, which were provided and verified by Company X based on historical production data. These assumptions simplify the modeling of the overall testing process. These assumptions make the modelling process easier but do not really show what happens in the real world when things are being made.

In addition, the MTBF value used in this model is derived from calibration estimates based on a limited dataset and does not directly calculate from detailed failure data. Due to the lack of data in the available database, although it can only approximate the actual downtime characteristics but cannot accurately reflect the true variability of machine failures.

Furthermore, external factors like changes in demand, worker availability and supply chain constraints were not considered in this study. These factors can affect how well a system works in manufacturing.

Lastly, using WITNESS 20 to create simulation models created limiting constraining factors at their maximum capacity for modelling, leading to the limitations due to constraints imposed by the software. Therefore, the results of this study should be interpreted within the scope of these assumptions and modelling limitations.

5.3 Recommendations for Future Research

Future research can extend this study by developing more comprehensive simulation models that incorporate additional manufacturing processes beyond the standalone MPI testing operation. For example, future models may include upstream and downstream processes, assembly operations, and semi-automated systems involving labour interaction and buffer management. This will better represent actual manufacturing systems and allow for the analysis of more complex production scenarios.

In addition, future research may improve the accuracy of downtime modelling by adding detailed failure data such as actual real-time failure distribution and machine-specific behaviour. This will improve the reliability of the simulation results and bring them closer to actual operating conditions.

Furthermore, future research could consider incorporating more dynamic operational variables, such as demand fluctuations, workforce performance, and supply chain constraints. Including these types of variables will make the simulation model more realistic and provide better methods to evaluate the performance of the production system under changing conditions.

The use of different discrete event simulation (DES) software, such as FlexSim or ProModel, can be explored to enhance model capability, visualization, and flexibility. In addition to improvements in these areas, by integrating artificial intelligence (AI) and real-time data analysis, simulation models can also enhance their ability to perform predictive analysis and make faster decisions.

REFERENCES

Ahuja, I.P.S. and Khamba, J.S. (2008) 'Total productive maintenance: literature review and directions', International Journal of Quality & Reliability Management, 25(7), pp. 709–756.

Amato, F., Peterson, B., Phillips, C. and Thipsawad, T. (2023) Predictive maintenance and process optimization in semiconductor manufacturing. IEEE Transactions on Semiconductor Manufacturing, 36(1), pp. 45–56.

Bamber, C.J., Castka, P., Sharp, J.M. and Motara, Y. (2003) 'Cross-functional team working for overall equipment effectiveness (OEE)', Journal of Quality in Maintenance Engineering, 9(3), pp. 223–238.

Beaverstock, M., Greenwood, A., Lavery, E. and Nordgren, W. (2017) Applied Simulation: Modeling and Analysis with FlexSim. Orem, UT: FlexSim Software Products.

Braglia, M., Castellano, D., Frosolini, M., & Gallo, M. (2018) 'Overall material usage effectiveness (OME): a structured indicator to measure the effective material usage within manufacturing processes', Production Planning & Control. Taylor & Francis, 29(2), pp. 143–157.

Braglia, M., Frosolini, M. and Zammori, F. (2008) 'Overall equipment effectiveness of a manufacturing line (OEEML) An integrated approach to assess systems performance', Journal of Manufacturing Technology Management. Emerald Group Publishing Limited, 20(1), pp. 8–29.

Cardoso, M.G., Ares, E., Ferreira, L.P. and Peláez, G., 2023. The use of simulation and artificial intelligence as a decision support tool for sustainable production lines. *Advances in Science and Technology*, *132*, pp.405-412.

Carson, J.S. (2002) 'Model verification and validation', Proceedings of the 2002 Winter Simulation Conference, pp. 52–58.

Chong, C.S. and Ng, K.H. (2016) Simulation-based scheduling optimization for semiconductor manufacturing. International Journal of Production Research, 54(12), pp. 3632–3645.

Colby, J., Peterson, B., Phillips, C. and Thipsawad, T. (2023) Fault detection systems for predictive maintenance in semiconductor wafer fabs. IEEE Transactions on Semiconductor Manufacturing, 36(1), pp. 45–56.

Dabbas, R.M. and Chen, H.-N. (2001) Mining semiconductor manufacturing data for productivity improvement—an integrated relational database approach. Computers in Industry, 45(1), pp. 29–44.

Dal, B., Tugwell, P. and Greatbanks, R. (2000) 'Overall equipment effectiveness as a measure of operational improvement – a practical analysis', International Journal of Operations & Production Management, 20(12), pp. 1488–1502.

Dequeant, K., Vialletelle, P., Lemaire, P. and Espinouse, M.-L. (2017) A review of simulation models applied to downtime management in semiconductor manufacturing. IFAC-PapersOnLine, 50(1), pp. 13536–13541.

Díaz-Reza, J.R., García-Alcaraz, J.L. and Avelar-Sosa, L. (2019) Overall Equipment Effectiveness: Impact on the Competitiveness and Profitability of Companies. Cham: Springer.

Domaschke, J., Grewal, S. and Rose, O. (1998) Simulation-based performance evaluation of dispatching rules in a semiconductor fab. Proceedings of the 1998 Winter Simulation Conference, pp. 1003–1008.

Dunscombe, P., Roberts, G. and Valiquette, L., 2000. Preventative maintenance and unscheduled downtime from an economic perspective. *Journal of Applied Clinical Medical Physics*, *1*(2), pp.68-75.

Ehm, H., Wenke, L., Mönch, L., Ponsignon, T. and Forstner, L. (2015) Towards a supply chain simulation reference model for the semiconductor industry. Proceedings of the 2015 Winter Simulation Conference, pp. 2322–3333.

Fleischer, J., Weismann, U. and Niggeschmüdt, S. (2006) 'Failure cost calculation and avoidance for proactive maintenance', CIRP Annals, 55(1), pp. 25–28.

Fowler, J.W. and Rose, O. (2004) Grand challenges in modeling and simulation of complex manufacturing systems. Simulation, 80(9), pp. 469–476.

Gibbons, P. M. and Burgess, S. C. (2010) 'Introducing OEE as a measure of lean Si Sigma capability', International Journal of Lean Six Sigma. Emerald Group Publishing Limited, 1(2), pp. 134–156.

Heger, J. and Voss, T. (2022) Optimal scheduling in semiconductor manufacturing using a digital twin based simulation approach. IFAC-PapersOnLine, 55(10), pp. 2233–2238.

Hu, H.-H., Chou, Y.-C. and Chen, T. (2022) Energy-efficient scheduling in semiconductor manufacturing via discrete event simulation. Sustainable Computing: Informatics and Systems, 35, p. 100737.

Huang, S.H., Dismukes, J.P., Shi, J., Su, Q., Razzak, M.A., Bodhale, R. and Robinson, D.E. (2003) 'Manufacturing productivity improvement using effectiveness metrics and simulation analysis', International Journal of Production Research, 41(3), pp. 513–527.

Jeon, B. and Lee, J. (2017) 'Machine learning-based tool condition monitoring and prognosis in semiconductor manufacturing', International Journal of Precision Engineering and Manufacturing, 18(3), pp. 373–382.

Kalir, A. (2018) Industry 4.0 and its impact on semiconductor manufacturing. Journal of Semiconductor Technology and Science, 18(5), pp. 543–550.

Law, A.M. and Kelton, W.D. (2007) *Simulation Modeling and Analysis*. 4th edn. New York: McGraw-Hill.

Leachman, R.C., Ding, S. and Chien, C.-F. (2007) Economic efficiency analysis of wafer fabrication. IEEE Transactions on Automation Science and Engineering, 4(4), pp. 501–512.

Lee, J. and Kim, Y. (2020) Simulation-based analysis of downtime in semiconductor electrodeposition using WITNESS. International Journal of Advanced Manufacturing Technology, 108(5–6), pp. 1503–1512.

Lee, S., Park, J. and Shin, H. (2023) Optimizing maintenance strategies in manufacturing using WITNESS simulation. Journal of Manufacturing Systems, 67, pp. 123–134.

Michlowicz, E. and Smolinska, K. (2018) 'WITNESS simulation software in logistics systems analysis', LogForum, 14(2), pp. 161–170.

Mönch, L., Fowler, J.W., Dauzère-Pérès, S., Mason, S.J. and Rose, O. (2018) A survey of scheduling problems in semiconductor manufacturing. European Journal of Operational Research, 271(2), pp. 391–414.

Mourtzis, D., Panopoulos, N. and Fotia, S. (2021) A holistic approach for the assessment of downtime impact in semiconductor manufacturing. Procedia CIRP, 104, pp. 1543–1548.

Muchiri, P. and Pintelon, L. (2008) 'Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion', International Journal of Production Research, 46(13), pp. 3517–3535.

Musa, A., Gunasekaran, A., Yusuf, Y. and Azhikhanov, A. (2025) Addressing supply chain disruptions and labor shortages in semiconductor manufacturing: A simulation approach. International Journal of Production Economics, 270, p. 109191.

Muthiah, K.M.N., Huang, S.H. and Mahadevan, S. (2008) 'Automating factory performance diagnostics using overall throughput effectiveness (OTE) metric', The International Journal of Advanced Manufacturing Technology, 36(7–8), pp. 811–824.

Nakajima, S. (1988) Introduction to TPM: Total Productive Maintenance. Cambridge, MA: Productivity Press.

Ndokwu, C. (2025) Financial implications of downtime in semiconductor manufacturing. Journal of Manufacturing Systems, 74, pp. 12–20.

Negahban, A. and Smith, J.S. (2014) Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems, 33(1), pp. 241–261.

Ng Corrales, L.D.C., Lambán, M.P., Hernandez Korner, M.E. and Royo, J. (2020) 'Overall equipment effectiveness: Systematic literature review and overview of different approaches', Applied Sciences. Multidisciplinary Digital Publishing Institute, 10(18), p. 6469.

Pecchia, A., Cotroneo, D., Cataldo, A. and Russo, S. (2023) 'Automated identification of equipment anomalies using operational data in semiconductor manufacturing', IEEE Transactions on Semiconductor Manufacturing, 36(1), pp. 12–21.

Phongmoo, S., Suedumrong, C., Kuensaen, C., Sinthavalai, R. and Leksakul, K., 2025. Predictive maintenance in semiconductor manufacturing: Comparative analysis of machine learning models for downtime reduction. *Computers & Industrial Engineering*, *205*, p.111211.

Raja, P. N. and Kannan, S. M. (2007) 'Evolutionary programming to improve yield and overall equipment effectiveness of casting industry', Journal of engineering and applied sciences, 2(12), pp. 1735–1742.

Ramirez-Hernandez, J.A., Crabtree, J., Yao, M., Fernandez, E. and O'Connor, M. (2008) Optimal preventive maintenance scheduling in semiconductor manufacturing. Proceedings of the 2008 Winter Simulation Conference, pp. 2157–2165.

Sargent, R.G. (2013) 'Verification and validation of simulation models', Journal of Simulation, 7(1), pp. 12–24.

Shou, W., Wang, J. and Wu, P. (2021) 'The application of simulation in lean production research: a critical review and future directions', *Engineering, Construction and Architectural Management*, 28(8), pp. 2119–2154.

Skoogh, A., Perera, T. and Johansson, B. (2017) ‘Input data management in simulation – industrial practices and future trends’, Simulation Modelling Practice and Theory, 73, pp. 18–26.

Soldatos, J., 2024. *Artificial intelligence in manufacturing: Enabling intelligent, flexible and cost-effective production through AI* (p. 505). Springer Nature.

Tan, Y., Li, X. and Zhang, Q. (2019) WITNESS-based simulation for surface mount technology production line optimization. Procedia Manufacturing, 39, pp. 876–884.

Tishtykbayeva, A.Z., Gelashvili, N.N. and Turusbekov, A.E., 2023. Artificial intelligence implementation in small businesses. *BUKETOV BUSINESS REVIEW*, *111*(3), pp.125-132.

Wang, Z., Zhang, Y. and Li, J. (2023) Pareto-based maintenance optimization in semiconductor manufacturing. IEEE Transactions on Semiconductor Manufacturing, 36(2), pp. 189–198.

Wudhikarn, R. (2012) ‘Improving overall equipment cost loss adding cost of quality’, International Journal of Production Research. Taylor & Francis, 50(12), pp. 3434–3449.

Yoon, H. and Chae, J. (2019) Simulation-based scheduling optimization in semiconductor wafer testing facilities. International Journal of Production Research, 57(14), pp. 4567–4582.

Zeidler, C., Kopp, D., Ponsignon, T., Ehm, H. and Frazzon, E.M. (2022) Model predictive control and hybrid systems towards autonomous semiconductor supply chains. IFAC-PapersOnLine, 55(10), pp. 2451–2456.