DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

PROCESS MINING AND VALUE STREAM MAPPING WITH LEAN MANUFACTURING TECHNIQUES: A CASE STUDY AT A FAUCET FACTORY (CASTING AND MACHINING)

by

Göksel SAKA

September, 2024

İZMİR

PROCESS MINING AND VALUE STREAM MAPPING WITH LEAN MANUFACTURING TECHNIQUES: A CASE STUDY AT A FAUCET FACTORY (CASTING AND MACHINING)

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
in Partial Fulfillment of the Requirements for the Degree of Master of Science
in Department of Industrial Engineering, Industrial Engineering Program

by

Göksel SAKA

September, 2024

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"PROCESS MINING AND VALUE STREAM MAPPING WITH LEAN MANUFACTURING TECHNIQUES: A CASE STUDY AT A FAUCET FACTORY (CASTING AND MACHINING) "** completed by **GÖKSEL SAKA** under supervision of **PROF.DR. BİLGE BİLGEN** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof.Dr. Bilge BİLGEN

Supervisor

Prof.Dr. Derya BİRANT

Jury Member

Doç.Dr. Mehmet Ali ILGIN

Jury Member

Prof. Dr. Okan FISTIKOĞLU

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENT

I’m grateful to my thesis advisor, Prof. Dr. Bilge BILGEN, for her help and support during the whole process in the master program. I am very grateful to her for taking the time to carefully review my studies, despite her limited time.

Also, thanks to Nazan GÜNEY for always being very supportive and kind, which made me feel peaceful every time I visited her, even in difficult times. I could not imagine this department without her.

The words cannot express my gratitude to my wife, Seher SAKA who supported, motivated, and encouraged me during my thesis and never lost her faith in me.

Lastly, I would be remiss in not mentioning my family, my daughter Cemre Zeynep SAKA, my son Ahmet Alp SAKA, my mother Hatice SAKA, my father Necmettin SAKA. Their belief in me has kept my self-confidence and motivation high during this process. I also apologize to my entire family for stealing their life during this thesis.

Göksel SAKA

PROCESS MINING AND VALUE STREAM MAPPING WITH LEAN MANUFACTURING TECHNIQUES: A CASE STUDY AT A FAUCET FACTORY (CASTING AND MACHINING)

ABSTRACT

In recent years, manufacturing systems have become increasingly complex due to the reduction in manufacturing lots alongside a rise in product variety. The rapid development of information systems has endowed companies with an abundance of valuable data, crucial for sustaining their competitive edge. Traditional methods for evaluating this critical data, such as Value Stream Mapping, are predominantly manual and lack the dynamic capability to model actual process flows over time. Consequently, they fail to capture the evolving nature of process flows. This thesis proposes a novel approach that integrates Vale Stream Mapping with Process Mining, thereby transforming the static nature of Value Stream Mapping and enabling an examination of the system over extended time intervals for performance analysis. The primary objective of this thesis is to reduce the throughput time and enhance the efficiency of a faucet factory, which in turn will lower work-in-process stock and working capital requirements. This is achieved by analyzing the flow using a combination of Value Stream Mapping and Process Mining. We revise the current state map based on the performance analysis results derived from Process Mining. Subsequently, several process improvement suggestions are formulated according to the revised current state map. As a result, the throughput time has been improved by %27, reducing from 48.8 days to 35.6 days in the Future State Map.

Keywords: Value stream mapping, process mining, throughput time, bottleneck, line balance rate.

PROSES MADENCİLİĞİ VE YALIN ÜRETİM TEKNİKLERİ KULLANILARAK DEĞER AKIŞ HARİTALAMA: ARMATÜR FİRMASINDA UYGULAMA (DÖKÜM VE TALAŞLI İMALAT)

ÖZ

Son yıllarda, üretim sistemleri giderek daha karmaşık hale gelmektedir. Bunun temel nedeni, üretim lotlarının azalırken ürün çeşitliliğinin artmasıdır. Bilgi sistemlerinin hızla gelişmesi, şirketlerin varlıklarını sürdürebilmeleri için hayati öneme sahip olan değerli veriler elde etmelerini sağlamıştır. Ancak, bu kıymetli verileri değerlendirmek için kullanılan geleneksel yöntemler, örneğin Değer Akış Haritalama, genellikle manuel olup, süreç akışlarını zaman içinde dinamik olarak modelleyebilme yeteneğinden yoksundur. Bu nedenle, süreç akışlarının nasıl değiştiğini haritalayamazlar. Bu tez, Değer Akış Haritalama ile proses madenciliğini bir arada kullanarak, Değer akış haritalamanın statik yapısını ortadan kaldırmayı ve sistemi daha geniş zaman aralıklarında inceleyerek performans analizi yapmayı mümkün kılan yenilikçi bir prosedür önermektedir. Bu tezin temel amacı, bir musluk fabrikasının akış süresini azaltmak ve verimliliğini artırmaktır. Bu da ara operasyon stoklarının ve işletme sermayesinin azalmasını sağlayacaktır. Bunu başarmak için, değer akış haritalama ve proses madenciliği yöntemlerini bir arada kullanarak akış analizi yapılmıştır. Proses madenciliği performans analizi sonuçlarına göre Mevcut Durum Haritası yeniden düzenlenmiştir. Ardından, güncellenmiş Mevcut Durum Haritasına dayanarak bazı süreç iyileştirme önerileri sunulmuştur. Sonuç olarak, akış süresi %27 oranında iyileştirilmiş, 48,8 günden 35,6 güne düşürülmüştür.

Anahtar Kelimeler: Değer akış haritalama, proses madenciliği, akış süresi, darboğaz, hat denge oranı.

CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

ABBREVIATIONS

CNC	: Computer Numerical Control
EPC	: Event-Driven Process Chain
ERP	: Enterprise Resource Planning
MES	: Manufacturing Execution System
MRP	: Material Requirements Planning
OEE	: Overall Equipment Effectiveness
PVD	: Physical Vapor Deposition
VSM	: Value Stream Mapping

CHAPTER ONE

INTRODUCTION

Productivity, defined as achieving greater output from a fixed set of inputs, has long been a central focus in economic history and development, and remains a key topic closely related to operations management. The productivity of a process is influenced by various factors, many of which are tied to its engineering. These include the technology employed, the amount of capital equipment utilized, the quality of materials, the effectiveness of the process itself, product design, efficient resource allocation and scheduling, workforce education and training, employee effort, and management practices. Nevertheless, the exact effects of these factors remain unclear, and there are widespread misconceptions about the true sources of productivity gains (Schmenner, 2015).

Productivity continues to be vital to our standard of living today, with many factors influencing it, but effective operations management plays a key role. Incorporating productivity issues into the framework of swift, even flow is considered a good management practice. Swift, even flow clarifies why the major innovations throughout business history were highly effective across various sectors, and this remains just as relevant today. Regardless of the economic sector or industry, asking questions about variation and throughput time can provide valuable insights into the most effective path to productivity increase. Alternative paths to productivity gains are often less certain and less rewarding (Schmenner, 2015).

The theory of swift, even flow asserts that only two factors are critical to achieving productivity gains, regardless of how it is measured. The first key factor is reducing variation, which can occur in three forms: quality, quantity, and timing. The second key factor is measuring the throughput time -the total time required to produce something from start to finish- and minimizing it as much as possible. Swift, even flow concentrates on the movement of materials through a process, emphasizing the smooth and uninterrupted flow of these materials. Also asks people to take the viewpoint of the materials moving through a process. By reducing the variation and throughput time of the materials, we can eliminate the non-value-added factors of production, where costs and inefficiencies are most common (Schmenner, 2015).

According to Little's law, the throughput rate is inversely proportional to the average time a part spends in a system (Little, 1961). So, to optimize the throughput of a manufacturing system, it is necessary to minimize the average throughput time. Throughput time in a manufacturing system is affected by bottlenecks, process variation, and non-value-adding activities. A clear understanding of how individual parts flow through the manufacturing system is necessary for reducing throughput times. The central objective of productivity improvement in manufacturing is to increase the rate at which parts flow through a manufacturing system (Lorenz et al., 2021).

Lean manufacturing is one of the leading philosophies for companies to increase their productivity and sustain their existence. VSM (Value Stream Mapping) is one of the most important tools of lean manufacturing that enables flow analysis. However, the biggest disadvantage of VSM is the tedious and potentially inaccurate data collection, as well as its static nature (Sullivan et al., 2022).

Process mining, on the other hand, is a very new method compared to VSM as a lean manufacturing technique, which allows big data to be automatically evaluated with statistical methods, process models developed, and performance analysis performed on them with the development of information systems.

The main objective of this thesis is to analyze the flow of components within a faucet manufacturing facility with the aim of reducing throughput time. To achieve this reduction in throughput time, it is essential to meticulously analyze the movement of individual components through the production process. This thesis employs a novel integration of VSM and process mining to dissect how components navigate the manufacturing flow, an approach that is relatively uncommon in the existing literature.

VSM serves as a pivotal tool for assessing and summarizing the flow of parts, providing a framework for discussing the production process. However, VSM traditionally demands substantial manual effort and lacks the capability to dynamically analyze the system's behavior over time. It typically offers insights based on averages, such as cycle time and uptime, or on instantaneous data points, like work in process (WIP) stock levels.

Conversely, process mining enables a more dynamic analysis by replaying and examining the system's operations through detailed event logs. This method allows for a granular, time-based investigation into how processes unfold within the production system. By merging these two methodologies, this thesis leverages the structured analytical framework of VSM while simultaneously overcoming its inherent limitations of static analysis. The combined approach thus provides a more comprehensive and dynamic understanding of the manufacturing flow, facilitating more effective strategies for throughput time reduction.

In this thesis, the innovative integration of two methodologies, VSM and process mining, was employed to perform a comprehensive performance analysis of a Faucet Factory. This dual approach facilitated a thorough evaluation of the factory's operations and identified key areas for improvement based on the performance data obtained from the combined analysis.

By leveraging the strengths of both VSM and process mining, the study meticulously assessed how the production processes could be optimized. VSM provided a systematic framework to visualize and analyze the current state of the factory's operations, highlighting inefficiencies and bottlenecks. Meanwhile, process mining enabled a dynamic exploration of the actual process flows, revealing temporal variations and operational anomalies that static analysis might overlook.

Following the identification of the improvement opportunities, a future state map was developed using the VSM approach, aligned with the specific targets set for enhancing the factory's performance. This future state map serves as a strategic blueprint for guiding the implementation of process enhancements, aiming to streamline operations and achieve the desired efficiency gains.

The main contributions of this thesis are stated as follows.

- A contribution to academic discourse is made through the application of process mining in manufacturing,

- The synergistic integration of VSM and process mining in manufacturing is demonstrated,

- Practical guidance is offered for practitioners, enabling the effective implementation of these methods to achieve substantial performance improvements.

This thesis is organized as follows. A comprehensive literature review on the VSM, digitalization of VSM that aims to overcome the manual nature of VSM and to automate and expedite its time-consuming processes and process mining and the application of process mining in manufacturing is presented in chapter 2. Chapter 3 offers an in-depth background on lean manufacturing and the principles of VSM. Chapter 4 delves into the methodology of process mining. In Chapter 5, a case study is presented, illustrating the combined use of VSM and process mining within a faucet factory. Chapter 6 discusses the challenges of using VSM and process mining. Finally, Chapter 7 concludes the thesis, presenting the key findings and offering recommendations for future research in this area.

CHAPTER TWO

LITERATURE REVIEW

This literature review discusses the use of VSM and process mining in manufacturing processes. It explores the application of VSM and process mining individually and together. The review also examines the digitalization of VSM, the methodological advancements in process mining, and their combined use to optimize production processes.

2.1 Literature Review on VSM

There is extensive literature on the use of VSM in manufacturing, we examined only VSM in casting like our faucet factory and a comprehensive study about VSM disadvantages and challenges. Chao et al. (2022) proposed a production line balance optimization model, which is similar to what we used in our research, for a sand-casting workshop, where VSM was employed to diagnose the problems within the production process. An optimization framework was formulated by leveraging principles of elimination, combination, rearrangement, simplification, and increase theory along with kanban management methods. Additionally, a simulation study was conducted to compare the current state VSM with the future state VSM, resulting in a 44.7% increase in line balance.

In the literature, there are studies like Forno et al. (2014) that address the challenges and disadvantages encountered in the application of VSM. They identified several key issues, including the oversimplification of complex processes, the significant manual effort required for data collection, and the static nature of traditional VSM which does not account for dynamic changes in the production environment. Additionally, they highlighted the difficulties in accurately mapping and analyzing processes in highly variable and customized production settings. Their insights were invaluable in our explanation of the challenges associated with VSM, particularly in understanding the limitations of VSM in capturing real-time data and adapting to continuous improvements. Forno et al. also discussed potential solutions and best practices for overcoming these challenges, such as integrating VSM with digital tools to enhance its effectiveness and accuracy.

Sá Ribeiro et al. (2023) present a literature review about VSM Approach to Manufacturing Systems in Industry 4.0. They have analyzed 37 records about VSM within the framework of Industry 4.0. They aim to contribute to the academic understanding of VSM in modern manufacturing and provide a foundation for future research in this area.

2.2 Literature Review on VSM Digitalization

With the transition to Industry 4.0, the digitalization of manufacturing and the storage of data on servers, various studies have been published on the digitalization and automatic implementation of VSM, which traditionally is performed manually and requires significant effort. The following are some studies related to the digitalization of VSM.

Teriete et al. (2022), noted that the digitalization of VSM faces challenges related to heterogeneous and incomplete data landscapes. They proposed an event-based framework to address these issues. Although they provided a detailed explanation of this framework, their study did not include an application of it.

Ferreira et al. (2022), conducted a study on the integration of lean manufacturing with Industry 4.0. In their research, they discussed the integration of the lean practice of VSM with a hybrid simulation that combines discrete event and agent-based modeling. Their objective was to extend the scope of VSM within the context of Industry 4.0, thereby supporting Industry 4.0 initiatives in manufacturing companies.

Sullivan et al. (2022), noted that VSM aids in the visualization of product and material flows as a snapshot, thereby facilitating a better understanding of production behavior. However, they also pointed out that VSM can result in oversimplification, tedious and inaccurate data collection, and a static nature. To address these issues, they proposed a semi-automated VSM solution utilizing a dynamic real-time location system to track and map processes, equipment, people, and materials.

Hortshofer-Rauch et al. (2022), conducted a comprehensive study on digitalized VSM, providing both a review and outlook. In this study, they addressed VSM and its integration with process mining and simulation, reviewing the existing literature on these subjects. They emphasized the significance of process mining for the digitalization of VSM but also highlighted several challenges, foremost among them being the lack of appropriate software.

2.3 Literature Review on Process Mining

Process mining serves a similar purpose to VSM. It is a tool that extracts models from process records, automatically generating current state maps through algorithms. This allows for in-depth examination of the process, performance analysis, and the identification of bottlenecks. Additionally, process mining enables continuous monitoring and replaying of the process flow to observe variations and deviations, thus facilitating problem detection. One of its key advantages is that it requires relatively little effort to implement. However, process mining lacks the flexibility of VSM; while VSM maps can be supplemented with necessary data, process mining is constrained by the limitations of its software and algorithms.

Process mining, introduced by Wil van der Aalst (2004) serves as a pivotal link between data mining and business process modeling. In one of the most comprehensive books on this subject, W. van der Aalst (2011) asserts that process mining bridges the gap between these two fields. Data mining focuses on extracting value from data to inform decision-making in areas such as marketing, sales, and production. Similarly, process mining seeks to derive value from business process transaction data, using this information to enhance business process modeling.

Process mining is also employed in quality control within manufacturing. A notable application of process mining in quality control is the study by Rozinat et al. (2009)

They implemented process mining in the test process of a wafer scanner manufacturer. Rather than focusing on quality control itself, their study emphasized the quality control process, investigating how quality control is actually conducted. Rozinat et al. (2009), assert that their study provides valuable insights; however, they also highlight the need for further research to develop process mining techniques for analyzing less structured processes. Such processes, like those in healthcare, often result in complex and confusing models, making it challenging to derive meaningful insights.

Hadžiosmanović et al. (2012), conducted a study for application of process mining in manufacturing, utilizing SCADA and MES software. However, they encountered difficulties in creating an event log from the data obtained from SCADA and MES systems due to missing information. To address this issue, they manually analyzed and artificially generated the event log, which they then replayed to analyze the ongoing processes.

Bettacchi et al. (2016), had studied about process mining with a big data that includes 450.000 events about 32 different products of a coffee machine producer in a six years' timeline. They compared the process mining techniques, Alpha Miner, Heuristic Miner, Integer Linear Programming Miner, Inductive Miner and Evolutionary Tree Miner and they stated that the Inductive Miner and Evolutionary Tree Miner gave successful results.

Meincheim et al. (2017), introduced a study about combining process mining with trace clustering. They made an application in manufacturing. They mentioned that the process mining results were difficult to understand. They proposed trace clustering to reduce the complexity of process models. They examine the records that make up 90% of the process types in the process model and look at their lead-times. Then, by replaying the model, they realized the queues in front of the machines and found the bottlenecks. As a result of the study, they identified the bottlenecks and supported process improvement.

Er et al. (2018), analyzed the production planning process of a global manufacturing company using process mining techniques. They used the Heuristic Miner algorithm to discover and analyze the production planning process. Their research specifically focused on understanding the effects of changes in the production plan on material allocation priorities.

In their extensive review, Garcia et al. (2019), delved into the various techniques and applications of process mining. They detailed the methodology of process mining and trace its historical development. Their review encompassed a wide array of studies, revealing that 77 out of 578 studies (13%) are applied in the manufacturing sector.

There are several studies exploring the use of process mining in real-time, automated environments. Nagy et al. (2019), conducted a study aimed at reducing the number of faulty products through the real-time application of process mining methods. By executing these methods in real time, they proposed that deviations in events and potential error sources could be detected more quickly, allowing for faster resolution and a subsequent reduction in faulty products. Although they provided a thorough explanation of their proposal, it is important to note that their application was conducted using test data, rather than in a real-world environment.

Another literature review on process mining and its industrial applications was conducted by Corallo et al. (2020). It is very useful for us that they have mentioned that %67 of the studies about process mining uses ProM application and %67 of them uses Heuristic Miner. 18 articles they analyzed, focus only on process mining, just 2 of the 18 articles' case study is in manufacturing company. The use of process mining in the manufacturing sector has been accelerating in recent years, though studies focusing on this application remain relatively few.

Farooqui et al. (2020), introduced an event-based data architecture designed to collect data from the factory floor using process mining. They utilized the Sequence Planner, a tool for modeling and analyzing production systems, to develop this architecture. Their application was implemented in an automotive company that is closely aligned with Industry 4.0 principles, incorporating robotics and capable of collecting extensive data from the factory floor.

One of the most inspiring studies is the Lorenz et al.'s (2021). They made a study about process mining to increase productivity in make-to-stock manufacturing. They mentioned about the VSM challenges and used process mining in Manufacturing. They explained how we can use process mining in manufacturing very well. The study is made at a sanitary product manufacturer and make-to-stock manufacturing as in our study. Their aim was reducing the throughput time as in our study. But they didn't employ VSM and process mining together.

Jonas Friederich et al. (2022), introduced an architecture aiming to generate dynamic models for production systems using process mining with real-time, continuously updated data streams. In their research, they have presented a joint approach for both material flow and machine behavior. They have detailed the necessary data for this architecture and explained from which sources these data can be obtained.

2.4 Literature Review on VSM and Process Mining

There are few studies which is using VSM and process mining in manufacturing. There are three studies below but one of them is in interior logistics. Knoll et al. (2019), introduced a study about VSM for internal logistics using multidimensional process mining. İt's different from our study because they didn't use process mining and VSM together. They used process mining for internal logistics, and they mentioned that the discovered process model is equal to VSM current state map and at the end they used lean manufacturing techniques.

Nawcki et al. (2021), made a study in an automative company's quality process by using Process Mining and Value Stream Mapping. They used two techniques separately and they have compared the two techniques. They also mentioned the problems of using Value Stream Mapping in their study. Disco process mining software is used in the study.

Rudnitckaia et al. (2022), made a study about process mining and value stream techniques is industrial manufacturing processes. The application was made at a gas meter manufacturer's assembly shop. The manufacturing was digitalized, and all the process data was recorded. They used process mining and VSM separately and they focused process mining and bottleneck detection.

A comprehensive study on the combined use of VSM and process mining was published yet by Julia Horsthofer-Rauch et al. (2024). Traditional VSM, requiring a high degree of manual effort, is noted to be inefficient in volatile and high-variance production environments. To address this, the researchers integrated process mining with VSM to digitize the process and incorporate sustainability elements. The implementation at HAWE Hydraulics SE demonstrated how sustainability-focused VSM can be practically applied through systematic recording and analysis of production data. The study involved collaboration with Celonis, where enhancements to the program were made for application. The process map derived from process discovery was used as the VSM current state map. The results indicated that digitized and sustainability-focused VSM provides more effective and efficient analysis and improvement of production processes. Furthermore, the integration of predictive and prescriptive analytics, along with simulations, was highlighted as a means to achieve a more effective value stream design. The study underscores that while digitized VSM requires significant effort and detailed process knowledge, it can yield substantial benefits when properly implemented. Additionally, it was noted that concepts for automated value stream design based on digitized value stream analysis were not addressed and should be explored in future research.

A review of the literature reveals only three studies where VSM and process mining were applied together in a production setting. However, in the Nawcki et al.'s (2021) and Rudnitckaia et al.'s (2022) study, VSM and process mining were used separately. In the Horsthofer-Rauch et al.'s (2024) study, VSM and process mining used together by Celonis process mining software and does not contain some of VSM steps like future state map or values for process enhancement or future state map. In contrast, our research employs VSM and process mining interactively, offering a novel approach that integrates these methodologies. Related literature is summarized in Table 2.1.

Table 2.1 Literature review about VSM, VSM digitalization and process mining

Authors and Year	Application Area	Method	Addressed Problem
Rozinat et al. (2009)	Quality control	Process mining	Analysis of quality control processes.
Hadžiosmanović et al. (2012)	Manufacturing	Process mining	Usage of SCADA and MES software for process mining
Forno et al. (2014)	-	VSM	Challenges in VSM implementation.
Bettacchi et al. (2016)	Manufacturing	Process mining	Big data analysis for a coffee machine producer.
Meincheim et al. (2017)	Manufacturing	Process mining	Reducing complexity in process models.
Er et al. (2018)	Production Planning	Process mining	Analysis of production planning process.
Garcia et al. (2019)	-	Process mining	Techniques and applications of process mining.
Nagy et al. (2019)	Manufacturing	Process mining	Reducing the number of faulty products.
Knoll et al. (2019)	Internal Logistics	Process mining & VSM	Analysis of internal logistics processes.
Corallo et al. (2020)	-	Process mining	Industrial applications of process mining.
Farooqui et al. (2020)	Manufacturing	Process mining	Data collection from production systems
Nawcki et al. (2021)	Quality Control	Process mining & VSM	Analysis of quality control processes by using two methods
Lorenz et al. (2021)	Manufacturing	Process mining	Enhancing productivity.
Chao et al. (2022)	Manufacturing	VSM	Production line balancing optimization.
Teriete et al. (2022)	-	VSM digitalization	Issues with heterogeneous and incomplete data.
Ferreira et al. (2022)	Manufacturing	VSM & hybrid simulation	Integration of VSM and industry 4.0.
Sullivan et al. (2022)	Manufacturing	VSM digitalization	Static nature and data collection challenges.
Friederich et al. (2022)	-	Process mining	Data requirements of production systems.
Horsthofer-Rauch et al. (2022)	-	Process mining & VSM	Digitalization of VSM and simulation.
Rudnitckaia et al. (2022)	Manufacturing	Process mining & VSM	Identifying bottlenecks.
Sá Ribeiro et al. (2023)	-	VSM	VSM within the framework of industry 4.0
Horsthofer-Rauch et al. (2024)	Manufacturing	Process mining & VSM	Digitalization of VSM and Using VSM, Process mining

CHAPTER THREE

LEAN MANUFACTURING AND VALUE STREAM MAPPING

Manufacturing in the twenty-first century is characterized by the increasing demand for customized products. This shift has required the development of more complex production planning and control systems, leading to traditional mass production methods becoming more challenging. Many organizations have struggled to adapt to the new customer-driven and globally competitive markets. These factors pose a significant challenge, compelling organizations to seek out new tools and methods to continue advancing in the transformed market landscape. While some organizations continued to grow by leveraging economic stability, others struggled due to their insufficient understanding of shifting customer mindsets and evolving cost practices. To address these challenges and enhance profitability, many manufacturers adopted "Lean Manufacturing." The primary goal of lean manufacturing is to be highly responsive to customer demand by minimizing waste. It focuses on producing products and services at the lowest possible cost while meeting the speed and efficiency required by customers (Bhamu & Singh Sangwan, 2014).

3.1 Definition of Lean Manufacturing

According to Krafcik (1988), Compared to mass production lean manufacturing uses less of everything, less human effort, less manufacturing space, less investment, less engineering hours to develop a new product. Additionally, lean manufacturing requires maintaining significantly less than half the necessary inventory on site, results in far fewer defects, and supports the production of a greater and continually expanding variety of products.

According to Womack (1990), Lean is a dynamic process of change, guided by a systematic set of principles and best practices focused on continuous improvement. Lean manufacturing integrates the most effective elements of both mass production and craft production, offering a balanced approach to productivity and quality.

According to Lean Enterprise Institute (2023): “Lean is a way of thinking about creating needed value with fewer resources and less waste. And lean is a practice consisting of continuous experimentation to achieve perfect value with zero waste. Lean thinking and practice occur together”.

Several researchers have provided variety of definitions for lean manufacturing. Lean manufacturing is;

- A Way,
- A Process,
- A Set of Principles,
- A Set of Tools and Techniques,
- An Approach,
- A Concept,
- A Philosophy,
- A Practice,
- A Program,
- A Manufacturing Paradigm (Bhamu & Singh Sangwan, 2014).

Key concepts of Lean Production, such as reducing non-value-adding activity, continuously improving processes, and shifting production control towards demand-oriented practices, enable a swift response to changing market demands (Busert & Fay, 2021).

3.2 History Of Lean Manufacturing

The two atomic bombs dropped on Japan during World War II had a profound impact on the country, both materially and spiritually. After the war, the United States banned Japan from forming an army, and the country made efforts to revitalize its economy. The material resources that had been allocated to the army were invested in industrialization after the army was disbanded. The need to manage scarce resources in the best possible way due to the war also created a favorable environment for the birth of lean production in Japan. The invitation of Dr. William Edwards Deming to

Japan by the Union of Japanese Scientists and Engineers in the 1950s and Dr. Deming's awareness of Japan about Total Quality Management were the first steps in the birth of lean production (Timeline - The W. Edwards Deming Institute, 2023).

On the other hand, Toyota, which was called "The Machine That Changed the World" by James P. Womack, Daniel T. Jones, and Daniel Ross, is moving forward with confidence (J.P.Womack et al., 1990). The engineer who is Toyota's production genius is Taiichi Ohno examined Ford's production system, also analyzed his own existing production systems and came up with concepts such as eliminating production losses, continuous improvement, and balanced production with small batch sizes. These concepts expanded and multiplied to give birth to the Toyota Production System.

Western countries turned their attention to Japan in 1974 with the oil crisis. The rise in oil prices and the economic crisis turned the attention of US automotive companies, which produced large-volume vehicles and had car models that did not take fuel consumption into account, to Toyota, which was capable of producing the same amount of production with only 4% of their workforce. Afterwards, the Toyota Production System spread all over the world (J.P.Womack et al., 1990).

The term "LEAN" was first used by John Krafcik in his 1988 article "The Triumph of the Lean Production System". Krafcik based this article on his master's thesis, which he wrote at the MIT (Massachusetts Institute of Technology) Sloan School of Management (Krafcik, 1988).

3.3 Lean Manufacturing Tools, Techniques

Since the beginning of the new century, many organizations have been working to adopt lean practices. This has led to the development and identification of numerous lean manufacturing tools, techniques, and methodologies, with new ones being proposed regularly. Lean manufacturing has evolved into an integrated system composed of highly interrelated elements and numerous management practices. There is a plethora of tools and techniques available for various purposes, all aimed at eliminating waste. However, the lean manufacturing tools and techniques often have multiple names; some overlap with others, and certain tools and techniques may even be implemented in different ways. Many of these tools and techniques should be integrated to achieve optimal results. Some of these techniques are given below.

- VSM,
- Kanban/Pull,
- JIT,
- TPM,
- 5S,
- Cellular Manufacturing,
- Continuous İmprovement,
- TQM,
- Kaizen,
- SMED,
- Multifunctional Teams/Employee İnvolvement,
- Production Smoothing,
- Visual Control,
- Supplier Relationship,
- Poke Yoke,
- Standardized Work,
- Simulation,
- Automation (Jidoka) (Bhamu & Singh Sangwan, 2014).

3.3.1 VSM – Value Stream Mapping

VSM is originally developed as a method for the Toyota Production System and later introduced as a distinct methodology. VSM is an easy and effective technique for gaining a detailed overview of the condition of an organization's value streams. Following the analysis of the current condition, flow-oriented target value streams are designed and then implemented (Sunk et al., 2017).

The term ‘value stream’ includes two descriptive elements – ‘value’ and ‘stream’. ‘Value’ refers to the inherent value creation in the production of goods, which represents the fundamental purpose of production: the transformation of raw materials into a product considered to be of higher value. The term ‘stream’ refers to the essential characteristic of production, which involves the spatial movement and qualitative transformation of parts and products throughout the production process. Due to machine utilization and work-sharing specialization, however, the various process steps cannot all be carried out in the same location or simultaneously (Erlach, 2013).

VSM is a widely applied lean management methodology that aids in analyzing value streams and identifying opportunities for optimization. Essential process steps and key metrics, particularly throughput time, are visualized in a value stream map, which enhances understanding of the current as-is process and serves as a tool for communication. Traditionally, data is gathered manually on the shop floor using pen and paper, leading in a time consuming process that offers only a limited snapshot of reality (Horsthofer-Rauch et al., 2022).

A value stream includes all activities value adding, non-value adding, and value preserving (supporting) that are necessary for creating a product or providing a service for the customer. Operational processes, flow of materials between processes, all control and steering activities, as the flow of information are included. To identify potential areas for improvement, VSM particularly focuses on the comparison between the total operating time and the overall lead time. The larger the gap between operating time and lead time, the greater the potential for improvement.

The value stream of a factory is modeled using six basic elements, each of them is described by specific parameters and further categorized by type, as shown in the Figure 3.1 below (Erlach, 2013).

- 'Production Process' represents the directly productive activities performed in the factory, as well as any external process,

- 'Business Process' describes order processing activities, encompassing tasks such as production planning and production control,
- 'Material Flow' refers to the movement of materials both within and between production processes, including those held in inventory,
- 'Information Flow' represents the transmission of data and documents between individual business processes and production processes, including the frequency of data exchanges,
- 'Customer' represents the demand that production must meet, thereby modeling the system's load.
- 'Supplier' represents the supply of raw materials and components to the production process (Erlach, 2013).

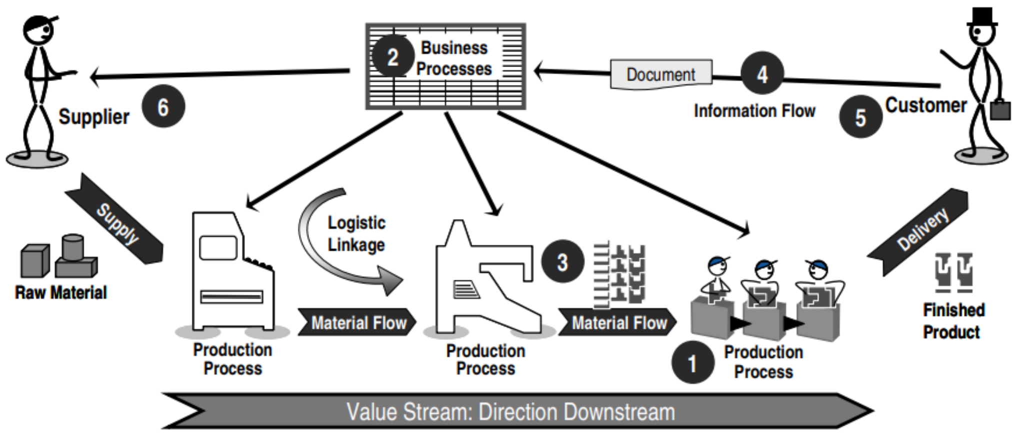


Figure 3.1 Value stream (Erlach, 2013)

The value stream flows from suppliers, passing through the factory, and ultimately reaching the customer, depicted from left to right in the illustration above (Figure 3.1). Accordingly, downstream production processes are positioned closer to the customer compared to upstream processes. The business processes related to order processing, the actual material flow within the factory, and the entire information flow across all production processes collectively form the complete production logistics of a factory. The logistical connection between two production processes represents the material flow between them and their respective control logic. At a higher level, we can differentiate between production and logistics.

VSM consists of 4 steps as shown in the figure below.

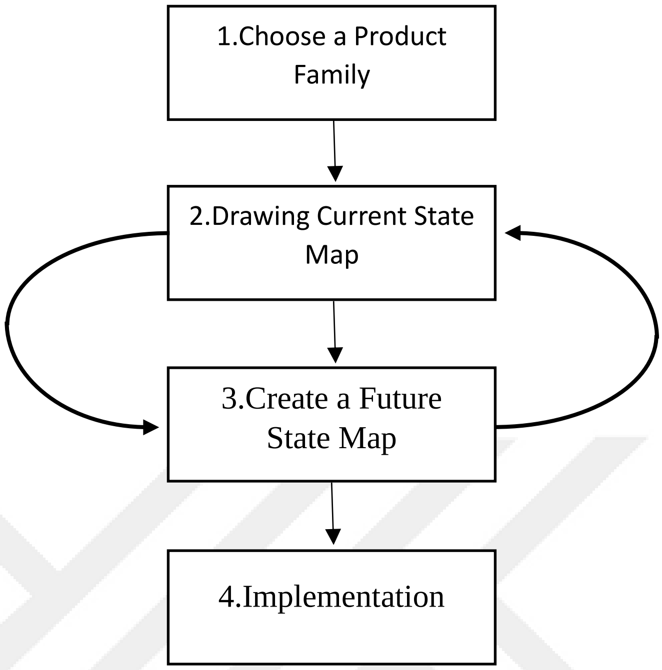


Figure 3.2 The steps of VSM. (Duggan, 2013; Rother & Shook, 1998)

3.3.1.1 Choose a Product Family

It's a rule that VSM applies to the combination of all production processes for a specific product. Therefore, we must determine the product family before the starting the Drawing the VSM. Every different product may follow its own path through production or may require different materials on the same production processes and so it creates its own value stream. If we draw a map of complex productions of products that have different properties, we can see less details and this hinders us from identifying areas that can be improved.

For mass producers dealing with a limited range of products, a value stream analysis may involve examining each product separately. Conversely, in the context of assembly-to-order scenarios, individually mapping products might not be the most effective approach. Similar products, produced through comparable production processes and largely using the same raw materials, could be better consolidated into a single value stream. Hence, the initiation of any value stream analysis entails breaking down the entire spectrum of products manufactured in a facility based on production-related criteria. This leads to the categorization of product families – a distinction from the product groups defined in sales. Subsequently, each of these product families is represented in a distinct value stream analysis.

The first step of a value stream analysis involves identifying product families. In a sense, a product family represents a distinct segment isolated from the rest of the factory for individual observation. A proficient analyst knows placing division lines to simplify the complex production processes within a factory. The goal is to minimize complexity without separating processes, components, or products that are inherently interconnected. A simple method can be used, known as the product family matrix (Erlach, 2013).

Table 3.1 Product family matrix

		Processes							
		Casting	Forging	Multi Spindle Machining	Machining	Polishing	Chrome Plating	Assembly	
Products	A	X		X	X	X	X	X	Product Family 1
	B	X		X	X	X	X	X	
	C	X		X	X	X	X	X	
	D	X		X	X	X	X	X	
	E		X	X	X	X	X	X	Product Family 2
	F		X	X	X	X	X	X	
	G		X	X	X	X	X	X	

3.3.1.2 *Drawing Current State Map*

The current state map should be completed quickly by a single person, by going to the field, personally making all measurements from end to end, and taking notes with a pencil. The baseline map is a snapshot of the flow. For this reason, it should be completed within two days at most. All data needed during drawing should be measured during this time. There is a possibility that previously measured and noted data may have changed or may be incorrect. Preparation of the map by more than one person causes a disconnection between the areas under the responsibility of the individuals and the entire flow may not be understandable (Rother & Shook, 1998).

Value stream maps are created by conducting a walkthrough of the shop floor, where we identify processes, identifying the start and end process where flow stops and inventory accumulates. Information about each process is then recorded in a data box. The typical data required includes the following.

- Cycle time (C/T)

Cycle time is a critical metric in the value stream method, reflecting the amount of time required to complete a product or component within the production process (Erlach, 2013).

- Setup time (Change Over (C/O))

The time interval needed to modify the settings on a machine to prepare it for processing a job. Reducing setup time is very important for enabling short production runs, which allows a business to more effectively implement just-in-time production. By reducing setup time, a business can profitably run smaller batches of products, as the setup cost associated with each unit is minimized.

- Uptime (Ut)

It's a parameter of machine performance and can be defined as the frequency with which the machine is in good working order when needed. Uptime can be formulated as follows (Chao et al., 2022):

$$Ut_t = \frac{Actual\ Production\ Time}{Total\ Available\ Time} \tag{3.1}$$

- Number of operators

The number of working operators at the process.

- Number of shifts

The number of working shifts at the process. It is used when processes that are illustrated at the VSM map working different shift numbers.

- Inventory

We also record inventory for all-process steps and show it by using an inventory -warning- triangle. The inventory quantity is illustrated below that triangle.

- Information flow

We must show the information flow between customers and suppliers. (forecasts, orders, production orders, purchasing order etc.) (Duggan, 2013). Material and Information Flow are the two sides of the coin. Two of them must be shown on the map (Rother & Shook, 1998).

- Scrap rate

It's the rate of defect parts/products in the process that cannot be send to the next process. If the scrap rate effects the stream, it must be shown on the map.

3.3.1.3 Value Stream Mapping Symbols

A series of symbols are used to represent processes and stream on the map. In addition to these, other symbols can be used, and each map drawer can use his own symbols. A common language should be created only within company employees who use these maps, and everyone should understand the same things from the symbols. This way, everyone will know how to draw and read the maps needed to move to lean production (Rother & Shook, 1998).

3.3.1.4 Create a Future State Map

The key to successfully drawing the future state map is having experience. The more future state maps you draw, it is more likely to draw a successful map in the new study. For this reason, working with a person experienced in lean production and VSM will bring the work closer to success.

During the birth of VSM, a guide for drawing the future state map was created at Toyota. Below are the steps of the manual used by Toyota (Rother & Shook, 1998).

1. First, takt time is calculated. Takt time; It is found by dividing the available work time in a shift to customer demand. In this way, the answer to what the cycle time should be to meet the daily demand of the customer is found. After the takt time is found, the aim should be to balance the production speed according to the takt time. Balancing production speed with takt time will prevent problems such as excessive stock or delayed delivery of customer orders.

2. Secondly, continuous flow should be ensured in all applicable areas. In a continuous flow, the philosophy of one produce and one convey is valid. Instead of separate processes, the parts produced by lining up the processes are delivered to the next process. In this way, semi-finished goods stocks are eliminated and the time it takes for the first incoming part to turn into a finished product is shortened. It is the necessity of swift, productivity and stream.

3. Third, supermarkets are installed in areas where continuous flow is not feasible. In this way, the customer process gets the product it needs from the supermarket when it needs it, and the supplier process produces to replace the products withdrawn from the supermarket.

4. Fourthly, production planning is aimed to be done for a production unit. Production scheduling is required for the next process after the last supermarket is established. The processes before this process should produce according to the pull system, and the processes after this process should produce according to the push system.

5. In the fifth step, the production mix level is determined in the production scheduling process. This means that all products that need to be produced in the process are produced in small batches alternately in the same time period.

6. In the sixth step, the production volume level is determined by sending work in small and consistent increments to the process where production scheduling is done. Generally, this process involves ordering as many products as a case or ceiling used. The process is ordered to produce this quantity of product and this quantity of product is picked. In this way, initial traction occurs.

7. In the seventh and final step, the production capabilities of the processes preceding the process for which production scheduling is made are improved. Model conversion times are shortened, and non-value creating times are reduced. In this way, previous transactions can adapt more quickly to changes in the demand of the next transaction.

When all these steps are followed, a future state map is drawn in the light of the information obtained in the current state map. Examining the current state map will reveal waste in the system, making it easier to eliminate it (Rother & Shook, 1998).

The future state serves as a guide for defining or specifying the various target conditions for the processes. It is specified by characteristics like;

- Customer takt,
- 100% added value,
- Continuous one-piece-flow,
- Zero defects,
- Lack of impairment for employees (Sunk et al., 2017).

The orientation towards the ideal-state, avoiding waste and stockless production system are critical factors in the design and rationalization paradigms of the ‘ideal value stream’ (Sunk et al., 2017).

3.3.1.5 Implementation

Problems that hinder the flow are identified from the current state map and an action plan is prepared to solve them. The action plan is a guide that will lead us to the future state. Additionally, an action plan is needed to monitor the implementation and to outline the measures (what, by whom, and by when) required to enhance the stream.

CHAPTER FOUR

PROCESS MINING

In the present day, particularly in the realm of industrial enterprises, the significance of concepts such as Industry 4.0 and the internet of things become more important. In the face of an increasingly competitive environment, these enterprises are compelled to prioritize activities that enhance productivity in order to sustain their existence.

Based on the philosophy that you cannot improve what you cannot measure, industrial enterprises aim to analyze all kinds of data, interpret the results of their analyses, and take actions to reduce product and process cost losses while increasing quality. Enterprises capable of achieving this effectively steer their direction, ensuring sustainability in the face of environmental and economic challenges posed by the competitive environment. On the contrary, those unable to meet these conditions find their existence threatened either due to rising costs or the inability to compete with rivals in terms of product and service quality, ultimately failing to sustain themselves.

The transition to Industry 4.0 has provided businesses with data registered on their own servers, which are essentially the untapped, most valuable treasures. In this context, the importance of data collection and analyzing big data resulting from collected data is once again highlighted.

Process mining is a recent research that introduced in 2004 by Wil van der Aalst (W. M. P. Van der Aalst & Weijters, 2004). Process mining is a discipline between data mining and process modeling and analysis (Corallo et al., 2020), but while classical data mining techniques are primarily data-centric, process mining is process-centric (W. M. P. Van der Aalst, 2011).

As a research discipline, process mining models process flows by using event logs (W. M. P. Van der Aalst, 2011). In contrast to manual process mapping, process mining allows analyzing process flows dynamically and automatically identifying non-value-adding activities (Schuh et al., 2020). Process mining includes data analytics techniques aimed at extracting process-related information (Yasmin et al., 2019).

In process mining there are three main terms:

1. Processes: The purpose of operation and enable the generation of a large volume of data.
2. Event Logs: Collects activities and actions performed by resources
3. Process Models: Represent processes and are created by events (Roldán et al., 2019).

Process mining works with various types of models as below;

- Transition systems,
- Petri Nets,
- Business Process,
- Modeling Notation,
- Causal Nets (Roldán et al., 2019).

Business process mining, or simply process mining, aims to automatically generate models that explain the behavior that observed in event logs. It can also be interpreted as the process of obtaining information about processes from event logs. Process mining has become a popular technique for business process management, especially after 2010. It establishes a significant bridge between data mining, process modeling, and process analysis.

4.1 Event Logs as a Starting Point for Process Mining

Digital event data is present in every sector, economy, organization, and household, and its growth is exponential. The presence of this data opens up new opportunities for new types of process analysis, based on observed behaviors rather than handmade models. The starting point of process mining is an event log (W. Van der Aalst, 2012).

Process miners require an event log that includes sequentially ordered events signifying the occurrence of specific activities. Additionally, these events must be associated with a case identifier, indicating a single instance of the process, such as an individual making an online purchase. The sequential arrangement of events is crucial for recognizing causal dependencies within process models (Garcia et al., 2019).

An event log contains of activities and actions performed by resources. Depending on our model and aim, different event logs may need to be produced. We will explain below the data requirements for process mining and the sources from which this data can be obtained.

Figure 4.1 shows the framework concerning the data requirements. Starting with a generic manufacturing system, the system's data is captured by sensors and distributed among the enterprise information system. Various tools within the company may store information about the physical system, such as Supervisory Control And Data Acquisition (SCADA), Manufacturing Execution System (MES), or Enterprise Resource Planning (ERP). The data held by such tools and needs aggregation and collection in event logs.

Depending on our model and aim, different event logs may need to be produced. For example, a model concentrating on material flow will utilize event logs containing information on the physical movements of parts within the system. A model concentrating the state of resources needs such records and the availability of resources (e.g., machines, tools, operators) can be retrieved through records of their deployment in the system (Friederich et al., 2022).

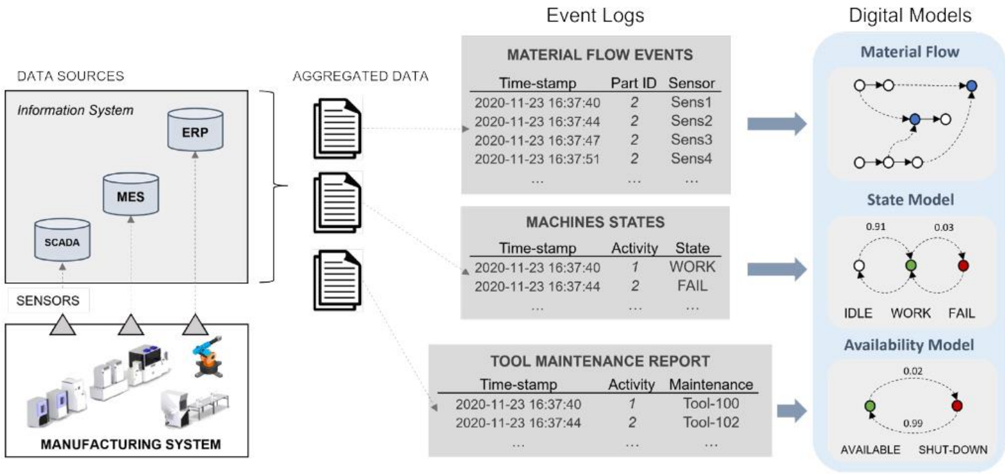


Figure 4.1 Data requirements framework for manufacturing systems (Friederich et al., 2022)

In process mining terminology, there are events for each case. Each event is related to a case. An event is specified with different properties, for example, an event has a timestamp that shows when the event occurred, and a resource that indicates the processor. If all the events of a case are chronologically ordered, we have a trace (a finite and non-empty series of events, and each event appears once, and time is not decreasing). It is possible to have different events following the same trace, but each case is different. An event log is a set of traces. In theory, an event log database can be recorded for every process that has a time dimension. A production order includes all activities related to a product (necessary machines, raw materials, time consumed) (Choueiri et al., 2020). The data of a production order for a product is used in our study.

4.2 Process Mining Software

Various software tools have been developed to facilitate process mining tasks. Some noncommercial process mining software's are ProM, PMLAB, RapidProM, CoBeFra and some commercial Software are; Celonis process mining (Celonis), Disco (Disco), Enterprise Discovery Suite (EDS), Interstage Business Process Manager Analytics (Fujitsu), Minit (Minit), myInvenio (myInvenio), Perceptive process mining (Perceptive), QPR ProcessAnalyzer (QPR), Rialto Process (Rialto), SNP Business Process Analysis (SNP), and webMethods Process Performance Manager (PPM) (W.

Van der Aalst, 2016). ProM is an open-source framework providing a variety of plugins for different process mining techniques. Disco is known for its user-friendly interface and visualization capabilities. Celonis offers advanced process analytics and is widely used for business process improvement. Fluxicon's tools, such as Fluxicon's Disco and Fluxicon's SiMiL, are designed to support process discovery and analysis. These tools assist organizations in gaining insights into their processes, identifying bottlenecks, and enhancing overall efficiency.

According to the literature review conducted by Angelo Corallo and others, ProM is the most widely used program, accounting for 67% of the studies in the literature related to process mining. In 28% of the studies, DISCO and ProM programs have been used together (Corallo et al., 2020). The most significant factors in this situation can be considered as the software being open-source and available for free use, as well as being developed by Wil van der Aalst, who is considered as the father of process mining.

In ProM, plugins can be easily added to the source code without the need for knowledge or recompilation, showcasing one of the user-friendly features of the ProM tool. ProM's architecture allows for 5 different types of plugins. These include mining plugins, where mining algorithms generating a Petri net based on certain event logs are applied; export plugins, facilitating the export of extensions; import plugins, enabling the import of extensions; analysis plugins, applying feature analyses on mining results; and finally, conversion plugins, implementing transformations between different data formats such as EPC and Petri nets (W. Van der Aalst, 2016).

4.3 Process Mining Methodology

Fig. 4.2 provides an overview of the process mining methodology, include six stages that relate to various input and output objects falling into three categories:

1. goal-related objects,
2. data objects,
3. models.

The 4 goal-related objects are:

1. Research questions, derived from project objectives, that are answered by,
2. Performance findings,
3. Compliance findings, leading to,
4. Improvement ideas, to achieve the goals (Van Eck et al., 2015).

The data objects represent 3 different representations of process-related data:

1. Information systems contain dynamic process data in varied formats, extractable and connectable to discrete events to shape,
2. Event data, the transformation of event data into,
3. Event logs, characterized by a case notion and event classes (Van Eck et al., 2015).

There are 2 types of models:

1- Process models describe the sequence of activities in a process, potentially enhanced with additional information such as temporal constraints, resource usage, or data usage,

2- Analytic models, any other type of models that give insight into the process as decision trees (Van Eck et al., 2015).

The initial two stages of the methodology are (1) planning and (2) extraction, wherein preliminary research questions are defined, and event data is extracted. Subsequently, one or more analysis iterations are executed, potentially concurrently. Generally, each analysis iteration involves the following stages, performed one or more times: (3) data processing, (4) mining & analysis, and (5) evaluation. An analysis iteration concentrates on addressing a particular research question by applying process mining activities and evaluating the discovered process models and other findings. The duration of such an iteration can range from minutes to days, depending on the complexity of the mining & analysis. If the findings are satisfactory, they can be applied to (6) process improvement & support (Van Eck et al., 2015).

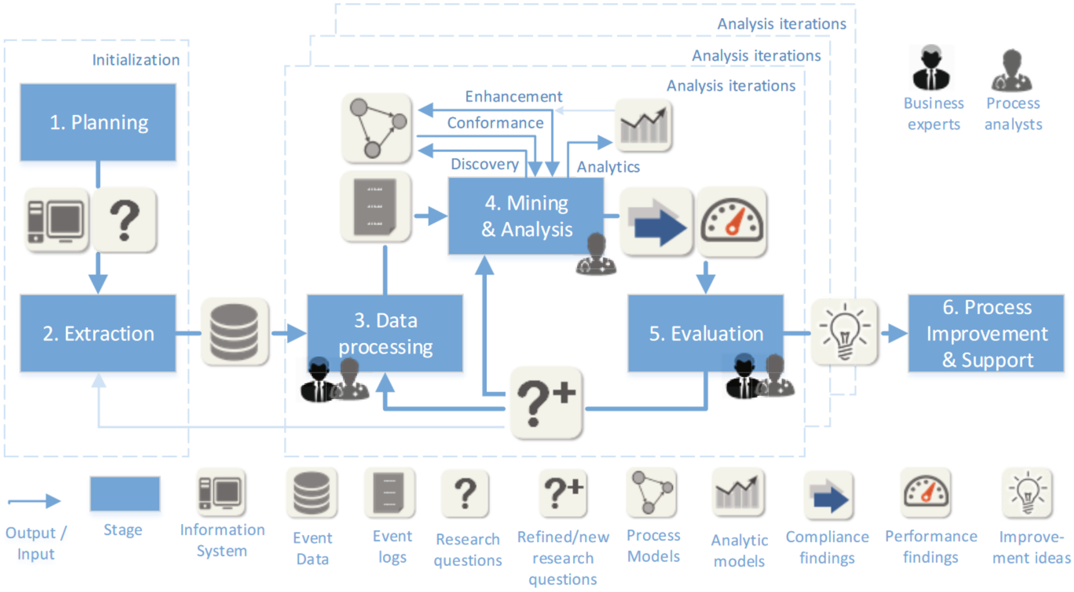


Figure 4.2 An overview of process mining methodology (van Eck et al., 2015)

4.3.1 *Process Mining Types*

Three main types of process mining are introduced by Wil van der Aalst (2016). These are named as;

- Process discovery:
- Process conformance checking,
- Process enhancement.

Event logs are utilized for the application of various process mining approaches. The analysis iteration stage holds a distinct significance in the process mining methodology. The step of mining and analyzing, where process mining techniques are applied, plays a crucial role within the analysis iteration cycle. It can be argued that a process mining methodology cannot be considered well-structured if it lacks a well-executed mining and analysis step where process mining techniques are applied.

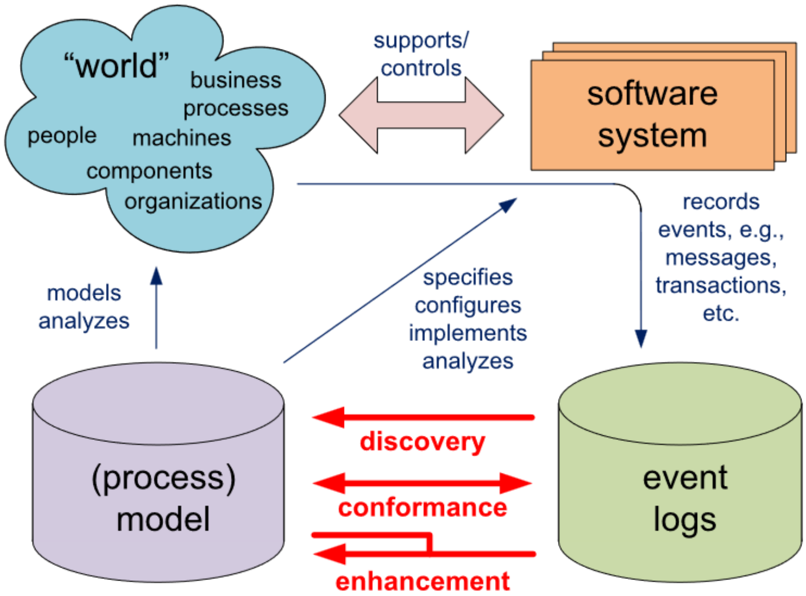


Figure 4.3 Place of 3 main types of process mining: discovery, conformance, enhancement (W. van der Aalst, 2016)

After that, we will explain the event log, that we use as an input for process mining then we will explain the process discovery, which is the first type of process mining, in section 4.3.3 we will talk about the second type of process mining, process conformance checking, then we will explain the third type of process mining, enhancement.

4.3.2 Process Discovery

Organizations use protocols to manage cases, often enforced by information systems. These protocols are frequently informal and may not have been documented. Additionally, documented procedures may not accurately reflect the real-world scenario. Hence, it becomes crucial to discover the actual processes by using event data. A discovery technique uses an event log to produce a model without relying on any pre-existing information.

The discovered process models can be used.

- To discuss problems with stakeholders (to achieve consensus; it is a necessity to have a shared perspective of the actual process),

- To generate ideas for process improvements (observing the real process and its problems force re-engineering efforts),
- for model enhancement, (like bottleneck analysis)
- to configure Workforce Management System or Business Process Management System (discovered model can be used as a template)

Since the mid-nineties, various research groups have been developing techniques for automated process discovery using event logs. Many classical approaches face challenges related to concurrency. At this point, the Alpha Algorithm serves as a simple technical example that takes concurrency as a starting point. Effectively addressing noise and incompleteness is crucial for practical applications of process discovery, leading several discovery algorithms such as Heuristic Mining, Fuzzy mining, and Genetic Mining to focus on tackling these issues (W. Van der Aalst, 2016). Furthermore, within this category, recently developed techniques such as Inductive Mining and Inductive Visual Mining, capable of creating a process tree by applying noise filtering parameters to paths, as suggested by Leemans et al (Leemans & Fahland, 2014), have gained prominence. These process mining algorithms will be explained below.

4.3.2.1 Alpha Miner Algorithm

The Alpha Algorithm is a technique that takes concurrency as its starting point. It scans event logs only for specific patterns. For example, if activity “a” is followed by “b” but “b” is never followed by “a”, then it is assumed that there is a causal dependency between “a” and “b”. The corresponding Petri net should include a place connecting “a” to “b” to represent this dependency. While the Alpha Algorithm is simple and effective, it encounters challenges with complex routing structures and noise (W. Van der Aalst, 2016).

The Alpha Miner Algorithm was not flexible in dealing with noise and incomplete event logs. It lacked the capability to identify short loops, map non-local dependencies, and handle restrictive choice structures. In other words, the alpha algorithm was unable to define main pathways and distinguish outliers because it represented all process steps and sequences without any numerical or frequency information (Garcia et al., 2019).

4.3.2.2 Heuristic Miner Algorithm

Lots of researchers have worked to improve the Alpha algorithm, proposing various algorithms to overcome its limitations. Heuristic mining, initially introduced by Weijters et al. in 2004, emerges as a readily executable algorithm capable of handling with noise, excluding exceptions and details, and discover common patterns. (W. M. P. Van der Aalst & Weijters, 2004) They propose a new approach including reverse engineering process, so instead of starting with process model design, they propose collecting event logs and subsequently applying a new process mining technique to obtain the actual execution flow. The proposed technique extends the Alpha algorithm, eliminating noise by calculating dependency and frequency tables to acquire a heuristic net and the Miner is known as Heuristic Miner. It is frequency-based, considering the frequency of the dependency/frequency table relationship, allowing for comparisons between manually devised models and executed processes. Although the Heuristic algorithm is the most used and customized process mining algorithms, ensuring good fitness, but it cannot provide total soundness due to the infrequent paths are not incorporated to the model (Garcia et al., 2019).

Noise can be filtered against outliers and the frequency of remaining paths can be shown by the Heuristic Algorithm. However, due to this filtering, outliers, erroneous behavior and process failures cannot be examined (Natschläger et al., 2017).

According to Mannhardt et al., most heuristics are based on frequency-based measures that evaluate the strength of dependencies between any two events based on observations in the event log (Mannhardt et al., 2017).

4.3.2.3 Fuzzy Miner Algorithm

Traditional process mining approaches encounter challenges when dealing with unstructured processes. The primary issue with existing algorithms is their assumptions about logs. Many algorithms presume the existence of a single, exact process model and assume that the log data is entirely accurate. These assumptions often lead to unrealistic and overly complex models. The discovered models are often “spaghetti-like”, showing all details without identifying what is important and what is not. The fuzzy algorithm is introduced by Güntner and Van der Aalst in 2007 to solve these problems. Fuzzy algorithm has been proposed as a new approach for process discovery, aiming to address the complexity issues of other process algorithms. Fuzzy mining algorithm aims to discover processes, even when they are unstructured and complex, by filtering activities only according to their importance defined by the user (Günther & van der Aalst, 2007).

Significance is defined as the relative importance of an action. The degree of interest lies in the frequency of the activities or the order in which they occur. If the user increases the significance level, higher filtering is applied, resulting in simpler models. Conversely, if the significance level is decreased, a more detailed model is produced. This is the most important advantage of this algorithm that the user can adjust the significance level according to process characteristics or business needs. If the process is structured and the event log is complete, the significance level can be decreased by the user, and a detailed process model can be analyzed and if the process is unstructured and the event log is incomplete, the significance level can be increased by the user and only common points can be highlighted (Yılmaz, 2019).

The Fuzzy Miner algorithm does not give a Petri Net as an output. So, the generated process model cannot be directly used in process conformance algorithms.

4.3.2.4 Inductive Miner Algorithm

The inductive algorithm is functional, and user-friendly process discovery tool, was first developed by Leemans et al. in 2013 (Leemans et al., 2013). The algorithm, presented with an extended framework and also called the B algorithm, is known as the inductive algorithm. With the inductive mining algorithm, the goal is to discover block-structured process models that are suitable and robust to the behaviors observed in the event logs. The algorithm characterizes the minimum information to find the process model. However, the inductive mining algorithm provides a polynomial-time complexity that provides a feasible computational cost. This algorithm has high reliability and is considered as a suitable method for resolving the variability of event records to simplify complex models (Garcia et al., 2019).

This mining algorithm also includes various criteria such as frequency analysis, clustering, detection of deviations and irregularities, analysis of time and bottlenecks, process understanding and evaluation of the overall appearance and values according to the result, all of which are included in the same solution (Garcia et al., 2019).

Current commercially available process discovery tools offer many options for adjusting the scope of discovery, but they generally do not produce models with executable semantics, so they cannot be used for automatic evaluation or further exploitation. There are lots of academic tools available for adjusting the scope of exploration, discovering the process model and evaluate it. Because of the nature of process exploration, using them iteratively is boring.

This time in 2014, Leemans et al. aimed to close this gap between commercial and academic tools with the Inductive Visual Mining Algorithm. Inductive Visual Miner supports the steps of process exploration by chaining existing academic tools and streamlining their use. The Inductive Mining algorithm improves the evaluation process with a new representation, animation and fast node selection filtering. Inductive Visual Miner has been developed as a plug-in for the ProM framework (Leemans & Fahland, 2014).

Inductive Visual Miner also generates a Petri net model, allowing the output of this miner to be utilized by other conformance checking and enhancement applications. Although Inductive Visual Miner offers many advantages, its execution time is significantly longer than that of other mining algorithms, particularly when dealing with large event logs. Nevertheless, The Inductive Visual Miner is the most valuable plug-in among the miners discussed in this section (Yılmaz, 2019).

4.3.3 Conformance Checking

Conformance checking compares events in the log within the model. Events are displayed in the petri net, so that the observed behavior of the system can be compared with the modeled behavior, and deviations and differences can be detected. Conformance checking can be used;

- For measuring the quality of the documented process,
- Identifying deviations and differences and identifying common points of these differences,
- Identifying process particles where most deviations occur,
- For auditing,
- To measure the quality of the discovered process model,
- To guide evolutionary process discovery algorithms (e.g. conformance checking is used checking for continuously evaluate the quality of each generation produced within genetic algorithm),
- As a starting point for process enhancement (W. Van der Aalst, 2012).

This list demonstrates that conformance checking can be utilized for various purposes, including evaluating process discovery algorithms, auditing, and compliance monitoring. It is important to note that auditors are responsible for verifying information about organizations by assessing whether business processes are being conducted within the boundaries established by managers, governments, and other stakeholders. Event logs serve as valuable input for this verification process (W. Van der Aalst, 2012).

Four quality metrics are used when comparing model and log. These are;

1. Fitness
2. Simplicity
3. Precision
4. Generalization

A model with good fitness accurately reflects most of the behaviors observed in the event log. A model achieves perfect fitness if all traces in the event log can be reproduced from start to finish by the model. Typically, fitness is quantified by a value between 0 (very poor fitness) and 1 (excellent fitness). Naturally, the simplest model that can explain the observed behavior in the event log is considered the best model.

Fitness and simplicity alone are not sufficient to measure the quality of a discovered process model. For example, it is relatively easy to create an extremely simple Petri net that can replay all traces from an event log (as well as other event logs that involve the same set of activities). Similarly, having a model that only replicates exact behavior is often undesirable. It is important to recognize that the log contains only a sample of behaviors, and many possible traces may not have been observed yet.

A model is precise if it does not allow "too much" behavior. A model which is not precise has an “underfitting”. Underfitting is a problem where the model overgeneralizes sample behavior in the log (i.e., the model allows for behavior that is very different from what is seen in the log) (W. Van der Aalst, 2012).

A model should also generalize behavior and not limit it to just the examples seen in the log. A model that does not generalize sufficiently is said to be “overfitting”. Overfitting is a problem of building a very specific model, whereas it is clear that the log only accommodates sample behavior (i.e., the model describes the specific sample log, the subsequent sample log of the same process should not generate a completely different process model (W. Van der Aalst, 2012).

As a summary; If L is recorded and M is modelled behavior (Kurniati & Wisudiawan, 2022).

$$Fittnes = \frac{|L \cap M|}{|L|}$$

$$Precision = \frac{|L \cap M|}{|M|}$$

$$Q_G = 1 - \frac{\sum nodes\left(\sqrt{\#\ executions}\right)^{-1}}{\#\ nodes\ in\ tree} \tag{4.1}$$

The discovered process and the filters applied must balance quality perspectives: fitness, precision, generalization, and simplicity. The most prominent observed metric is fitness, which results in a value between 0 (lack of capacity to support all event traces) and 1 (perfect fitness). Three other quality dimensions were studied by Rozinat, de Jong et al. (2009), proposed conformity by focusing on the idea of Occam's razor (Rozinat et al., 2009) (Garcia et al., 2019).

In this study, the fitness quality metric is calculated and used because of the most dominant metric is fitness.

4.3.4 Process Enhancement

Process enhancement is the last of the process mining types. This type of process mining activity focuses on extending the process model using relevant information. Process mining can serve a similar function to GPS applications that highlight congested streets by utilizing timestamping on event logs and integrating this data with the process model for predictions using statistical methods or machine learning. These extended process models are very useful for providing operational support, the most ambitious form of process mining (Garcia et al., 2019).

The performance of a process can be evaluated from three dimensions: time, cost, and quality. In this context, we focus on analyzing the time dimension of a process. The time perspective is concerned with the timing and frequency of events. By replaying traces on a process model, time information for different steps of the process becomes available. When events are timestamped, it becomes possible to identify bottlenecks, measure service levels, monitor resource usage, and estimate the remaining processing time of ongoing processes (Yasmin et al., 2019).

This process mining type supports answering questions such as the average processing time of cases, what the transition and decision probabilities are, which transitions are time-based, what are the critical activities and resources, how much is the total processing time, and how much is the waiting time between two activities.

According to García et al. (2019), this process mining technique allows the visualization of all process instances within defined date time periods, performs each case presentation on the process model, and provides time acceleration to obtain information about what will happen over weeks in a few minutes.

In the process enhancement step, all visual highlights have a well-defined meaning, and these can be exemplified as follows:

- The activity size represents the number of events or metrics associated with costs or resource usage.
- The color of the activity highlights the work time or process.
- The width of a connection/arc reflects the importance of it.
- The color of the connections highlights the waiting time to perform the next activity and reveals bottlenecks.

- The position of activities may have an indirect meaning, such as identifying hub or support activities (Garcia et al., 2019).

The enhancement type of process mining can be described as improving a process model by using information about the actual process recorded in an event log. The inputs of this type of mining are; the process model, performance and compliance findings obtained from the process discovery and conformance checking steps. The outputs are development ideas and research questions (Van Eck et al., 2015).

An example of this would be extending the process model with performance information to repair the process model based on time or cost or based on current executions shown in the relevant event log. The results of the enhancement step are process models, if we consider this from another perspective, although the results of the conformance check can be considered without any process model, this is not the case for the enhancement step (Van Eck et al., 2015).

CHAPTER FIVE

CASE STUDY

5.1 Problem Statement

In this study, we examine the applicability of integrating process mining and Value Stream Mapping (VSM) with lean production techniques. The focus is accelerating the material flow from raw material to final product, increasing efficiency, minimizing throughput time, improving responsiveness to customer orders, and reducing costs in the construction materials sector.

5.2 About The Faucet Factory

The factory produces water faucets for the building materials industry. Established in 1957, the company has designated its central production location as the Aegean Region since 1990. Operating as a foundation company, it employs a total of 125 white-collar workers and 350 blue-collar workers across two separate production facilities. Annually, it processes 4,980 tons of brass to produce 4,200,000 pieces of various faucets, siphons, and accessories within a closed area of 20,547 square meters, out of a total area of 62,281 square meters.

The company's product range includes mixers, taps, and accessories such as shower and bathroom mixers, sink mixers, bidet mixers, infrared mixers, self-closing taps, thermostatic mixers, concealed bath and shower mixers, as well as siphons and other accessories compatible with their faucets.

Only the brass components of the products in the range are manufactured in the factory. The factory's processes include casting, hot forging, machining, multi-spindle machining, polishing (surface polishing), and chrome plating (surface coating). These processes are conducted in batch-type manufacturing workshops. The casting, hot forging, machining, and multi-spindle machining workshops have long setup times. These processes use batch-type and make-to-stock manufacturing methods, with parts moving to the next process using a push system.

After production, the parts are stored in a warehouse along with other procured components. The products are then assembled using these procured parts. The assembly workshop operates on a make-to-order basis and employs a pull system. Necessary parts for each order are retrieved from the warehouse, and assembly operations are performed accordingly.

In the casting workshop, sand cores are first prepared to create internal cavities in the casting process and are allowed to dry for one day. Next, the sand cores are placed into casting molds by an operator, and melted brass is injected into the molds using low-pressure casting machines. The cast parts are then sent to a sandblasting machine to remove any residual sand from their interiors. After sandblasting, the parts are transferred to an automatic runner cutting machine to trim the runners. If the casting mold contains multiple cavities, the parts are separated into individual units during the runner cutting process.

The parts, initially shaped in the casting workshop, are sent to the machining workshop for further processing. In this workshop, holes are drilled, and threads are created to prepare the parts for assembly with their subcomponents. Computer Numerical Control (CNC) transfer machines, which are specifically designed and manufactured to meet the business's machining requirements, are used in this process. These machines are equipped with cutting tools tailored to the design specifications of the parts, enabling them to machine multiple features in a single operation. After machining, the parts are cleaned of cutting oil and brass chips using a parts washer.

The processed and functionally completed parts are sent to an outsourced surface polishing service to meet cosmetic requirements. At this stage, the parts first undergo a strip sanding process, followed by polishing with rotating fabric discs on polishing motors. A final polishing is then performed using a finer paste. Once polished, the parts are returned to the factory for plating.

After polishing, the parts are plated according to the specifications of the work order. The company's current product range includes chrome plating, gold-colored physical vapor deposition (PVD) coating, gold plating, antique matte lacquer coating, and various color painting methods. However, approximately 95% of production involves chrome plating.

After the coating process, the parts are sent to the assembly workshop for integration with their subcomponents. The assembly is conducted using machines specifically designed to meet the business's requirements. Once assembled, the parts are automatically boxed, labeled, and sent to the Product Warehouse.

Quality control procedures are applied at various stages throughout these processes. Initial manufacturing approval is granted after machine setup. During production, critical dimensions and process parameters are checked twice daily. Before the parts leave the workshop, a final quality control inspection is conducted. Parts sent to outsourced polishing workshops are subject to sampling inspections by quality control personnel on-site. Upon their return, the parts undergo a 100% inspection in the incoming quality control department

5.3 Data Collection Process

A Manufacturing Execution System (MES) is used in the selected factory for Overall Equipment Effectiveness (OEE) calculations by collecting real-time signals from all machines and transmitting them to the main computer for processing. This system enables tracking of downtime details, cycle time analyses, performance monitoring, scrap rates, and other relevant metrics for specified date ranges. The required data for OEE calculations has been obtained through this efficiency program. Additionally, for part-specific information such as stock levels, the Enterprise Resource Planning (ERP) system used by the business has been leveraged. The accuracy of the data has been verified through on-site manual counts.

MES and ERP programs are considered as fundamental components of Industry 4.0. "These systems help create smarter and more connected manufacturing environments by supporting digitalization, data integration, and the use of advanced technologies in business processes. Our study aims to utilize the data stored on servers in a way that minimizes effort while maximizing utility. Additionally, in later stages, automating these processes will enable the automatic generation of a process mining model and VSM after each production order. This will allow analysts to analyze all processes efficiently."

5.4 Value Stream Mapping Study

In this section we will draw a VSM current state map for our case study. The first step of VSM process is choosing a product family as mentioned before.

5.4.1 Choose a Product Family

A product family is a group of products that undergo similar process steps, particularly sharing common hardware in the final stages of production processes (Rother & Shook, 1998). In the VSM method, a diagram is created for a single product family to maintain the simplicity and prevent the loss of clarity in the map. Including the entire structure in the map could complicate it and hinder the visibility of improvement opportunities.

There are two types of mixer bodies in the factory. The first type, the cast mixer body, is used much more frequently than the other. The second type, the hot-forged mixer body, accounts for 20% of the total mixer body consumption. Both types follow the same processes, except for the first operation as seen in Table 5.1.

Table 5.1 The operations of two types of body

	Processes					
	Casting	Hot Forging	Machining	Polishing	Chrome Plating	Assembly
Casted Body	X		X	X	X	X
Hot Forged Body		X	X	X	X	X

This study includes value stream mapping and process mining analysis for the production of a mixer body, which sells 65,000 pieces per year but involves casting approximately 82,000 pieces due to a high scrap rate in the casting process. Cast mixer bodies, made of brass, represent a process where 70% of the Work in Process stock is held, while the other cast mixer bodies follow the same process route.

5.4.2 Drawing Current State Map

The factory generates 80% of its sales domestically, which are managed by a sales company affiliated with the same holding company. Similarly, export sales are handled by an export company also affiliated with the holding company. Export orders are sent to the planning department, where delivery dates are determined collaboratively with the export company. Production is then customized based on these orders.

Our VSM study is conducted for domestic sales, which constitute 80% of the total sales. Domestic sales are operated through the domestic sales company using a consignment system. The company sends the manufactured products to the consignment warehouse. Although the inventory of the products in the consignment warehouse belongs to the manufacturing company, the management is under the control of the sales company. The sales company makes shipments from the consignment warehouse based on dealer demands. As shipments are made, notifications are sent to the manufacturing company, and the products in the inventory are invoiced to the sales company.

Additionally, 60% of consignment sales are made in the last three days of the month. This is due to the sales company pressuring dealers to make sales using various sales strategies. Because 60% of sales are made in the last three days of the month, the consignment warehouse does not accept product shipments during the first week of the following month. The sales company physically ships the bulk sales to dealers. The manufacturing company holds the products produced in the first week until they are accepted for shipment to the consignment warehouse.

At the beginning of each month, the sales company informs the Planning department of the products and quantities it wants to be shipped to the consignment warehouse for that month. The manufacturing company is obligated to ship these products to the consignment warehouse by the end of the month. Therefore, the production cycle of the manufacturing company is monthly.

The Planning department creates forecasts based on past months' sales data, in addition to the orders received at the beginning of the month. Both the forecasts generated by the Planning department and the orders received at the beginning of the month are used to run Material Requirements Planning (MRP), and production orders are created and released to the workshops. However, the priority in production plans is consistently given to orders from the sales company. When the workload in the workshops decreases, production is carried out according to the forecast made by the Planning department.

The Planning department makes a schedule for the casting process at the beginning of the month, reviewing the plan daily based on changing conditions and revising it if necessary. A similar approach is applied to the machining process. Cast parts move forward through the push system until completion. For the assembly workshop, the Planning department creates weekly plans based on the condition of parts in the machining, polishing, and plating processes. The reason for planning this way for the assembly workshop is the high scrap and rework rates, making it challenging to achieve the planned production for the initial workshops.

A single type of brass ingot is used as raw material in the casting workshop. The brass ingots are supplied by the raw material company affiliated with the same holding. The advantage of being affiliated with the same holding, along with the proximity of the factory, allows for timely ingot shipments when needed, reducing the need for high stock levels at the factory. In the current state, a truckload of 17 tons of raw material is shipped every 2 days. This quantity is adjusted by the Planning department based on the number of active casting furnaces in the casting workshop.

The working duration for a shift is 8 hours (480 minutes). Within this working period, there is a 30 minutes break for meals and two tea breaks, each lasting 10 minutes. The net working time, calculated by deducting planned downtime, amounts to 430 minutes (25,800 seconds). The shift schedule varies across workshops based on the order quantity for the respective month and the number of human resources available.

At the time the current state map is drawn, the processes of sand core machines, casting, sandblasting, runner cutting, machining, and part washing operate on a 3-shift basis. Meanwhile, chrome-plating processes operate on a 2-shift basis, and outsourced surface polishing and assembly processes operate on a 1-shift basis.

5.4.2.1 Current State Process Definitions

The brass mixer body is processed as follows; sand core process, casting, sandblasting, runner cutting, machining, washing, surface polishing, chrome plating and assembly.

The first process is sand core preparation. The factory has four identical sand core machines. The cycle time for these machines is 44 seconds, and the setup time is 2 hours, equivalent to 7,200 seconds. Each machine is operated by one operator. The responsibilities of the machine operator include performing touch-ups on the produced sand cores and arranging them in the casings.

The factory has six casting machines. The average cycle time for the low-pressure casting machines, which represent the initial step in the emergence of the main product, is 41 seconds. The selected component for mapping can be produced on any of the six casting machines. The average model changeover time for the casting machines is 2 hours (7,200 seconds). One operator is assigned to each casting machine. The operator places the sand cores into the casting mold and initiates the operation by pressing the start button. Once the casting is completed, the finished components are automatically removed from the mold. At the time of mapping the current state, there are 998 sand cores awaiting use in the casting process.

The cast bodies must wait for complete cooling for one shift before sandblasting. If sandblasting is performed before sufficient cooling, it results in deformities in the cast bodies. This waiting period is a mandatory waste for the factory.

The casted bodies, amounting to an average of 200 pieces per casing, are filled into casings and transported to the sandblasting process. Processing time for one casing in the sandblasting machine is 31 minutes (1,860 seconds). So, cycle time for a body is (31 x 60) / 200 = 9.3 seconds nearly 9 seconds. And the processing time is mentioned at the current state map is 31 x 60 = 1,860 seconds. Before loading each casing, there is a setup time of 2 minutes (120 seconds). There is no dedicated operator responsible for the sandblasting machine, and there is no continuous operator working on the sandblasting machine. When the sandblasting process is completed, an available operator goes to the machine to handle the loading process. At the time of mapping the current state, there are 520 pieces that have been casted and are awaiting sandblasting.

The bodies, after being sandblasted and separated from their sand cores, are sent to the automatic runner cutting machine to have the runners, which facilitate the flow of molten metal into the cavity where the shape of the body is located in the casting mold, cut. There are two runner cutting machines, each with two stations. The average cycle time for these machines is 41 seconds, and the model changeover time is 0.25 hours (900 seconds). An operator is responsible for both stations on the runner cutting machine. Consequently, the operator is engaged in the production of the relevant body for half of their time. At the time of mapping the current state, there are 2,770 pieces that have been sandblasted and are awaiting runner cutting.

Cast brass bodies, after having their runners cut, proceed to the vertical machining center for further processing. A body is produced at the machining center every 175 seconds. The average model changeover time is 2 hours (7,200 seconds). One operator at the vertical machining center is responsible for two of the machines, placing the piece into the machining chuck, initiating the machine, and retrieving the processed body from the chuck to fill the casing. At the time of mapping the current state, there are 1,747 pieces with cut runners awaiting machining.

The machined bodies are directed to the part washer to be cleaned of cutting oil and brass chips. These machined bodies are loaded into the washing station with baskets containing 50 pieces each, and every 15 minutes, a basket is removed from the station. Therefore, the cycle time for a body is (15 x 60) / 50 = 18 seconds, and the processing time mentioned in the current state map is 15 x 60 = 900 seconds. There is no specified setup time for the washing process. A single operator at the washing station is responsible for loading baskets into the machine, counting the parts in the removed baskets, loading them into casings, and providing production confirmations in the ERP program for completed operations. At the time the current state map is drawn, there are 151 pieces of processed bodies awaiting washing.

The washed bodies are outsourced for surface polishing outside the factory. Parts sent during the day are processed the next day and return to the factory in batches. The return of a finished part to the factory takes approximately 24 hours. The factory can outsource 10,000 pieces of bodies in a day. Therefore, the cycle time for a body is (24 x 3,600) / 10,000 = 8.64 seconds, and the processing time mentioned in the current state map is 24 x 3,600 = 86,400 seconds. Shipments for outsourcing for the surface polishing process occur twice a day. At the time the current state map is drawn, there are 310 pieces awaiting shipment for surface polishing.

The parts coming from the outsourced surface polishing process are sent to the plating center for chrome plating. To facilitate the plating process, the parts are initially hung on specific hangers tailored to each part. Subsequently, these hangers, bearing the parts, are placed onto platforms known as bars, which enter and exit the plating baths, before being sent to the plating center. Each bar accommodates 3 hangers, with approximately 20 pieces hung on each hanger. A bar is removed from the plating center every 4 minutes. The processing time for the facility is 2 hours, meaning that it takes 2 hours from the moment the facility is initiated for the first bar sent to and exit the plating baths. So, cycle time is 4 minutes for 20 x 3 = 60 bodies and so on the cycle time is (4 x 60) / 60 = 4 seconds. Therefore, the plating center works 2 shifts for about 16 hours and the first piece can only be retrieved 2 hours after the initiation, the uptime is mentioned as 14 hours /16 hours = %87.5. Two operators are responsible for loading and unloading hangers into and from the baths. At the current state map, there are 55 pieces that have completed the surface polishing process and are awaiting the chrome plating process.

The final step for the transformation of the parts into a finished product as a mixer is the assembly process. The assembly process for the selected body is semi-automatic. An assembly machine constitutes one station of the 8-station assembly, and the machine determines the system's speed. Other stations involve preparation for the machine and subsequent surface quality control operations. Each station has one operator, and at the station with the machine, the operator loads and unloads the body onto/from the machine. The same task is performed by 14 operators with a longer cycle time without the machine. The machine's cycle time is 10 seconds, and the setup time is 900 seconds (15 minutes). At the current state map, there are 1,220 plated parts awaiting assembly.

The assembled products are sent for shipment. The shipment department dispatches the products to the consignment warehouse. At the current state map, there are 2,400 pieces awaiting shipment. The production progresses in batches from start to finish, preventing the emergence of mixed products from the assembly.

While drawing the VSM Current State Map given below, cycle Time (CT), change over time (CO), uptime (Ut), number of shifts (Shifts), scrap rate (Scrap), number of operator (Operator) were included in the databoxes created for each process. Definitions about these have been given before.

WIP stock values received instantly from SAP is written between each process.

The raw material supplier has been added to the map and the shipment frequency is indicated. Similarly, the customer warehouse has been added to the map and the shipment frequency has been specified.

Production planning and control department has also been added to the map. The relevant department directs the raw material supplier according to the information from the customer, and schedules are given to some points in the production, thus directing the production.

At the bottom of the map, the cycle time of that process is stated under each process. The sum of the cycle times gives us the Total Processing Time. This value corresponds to value added time. Production lead times, which are found by dividing the average WIP (Work In Process) stocks to the customer's daily demand, have been added under the average WIP stocks. The sum of these gives us the total throughput time

According to the VSM current status map given below in Figure 5.1, there is a value added time of 96,671 seconds versus a production lead time of 48.8 days. This means that only 2.2% of the total 48.8 days of throughput time is value added time.

Our aim is to reduce throughput time and increase the value-added time rate. As a result, we aim to reduce stocks and the money tied to stock, achieve lower lead times, reduce the total product cost, and strengthen the company's position in the market.

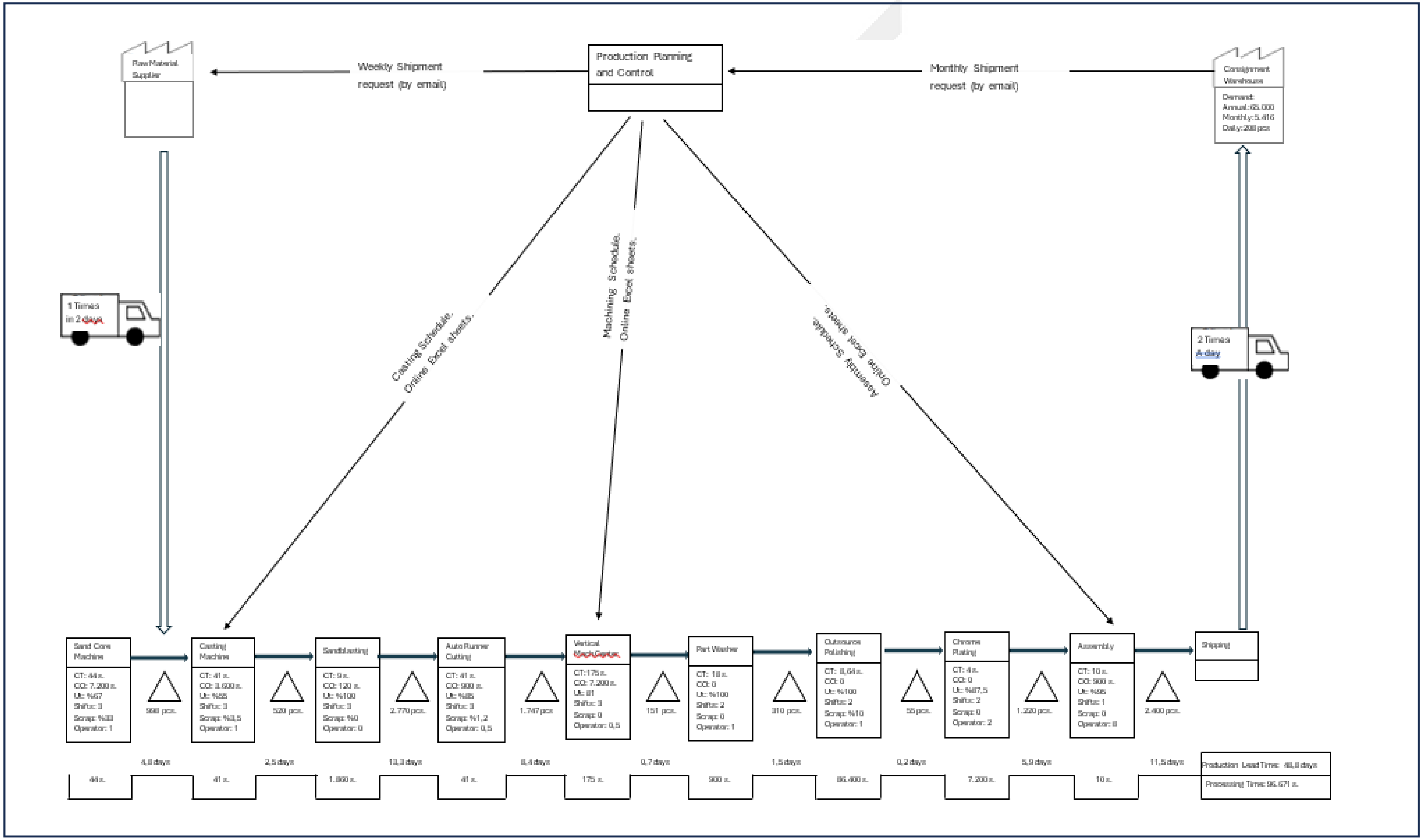


Figure 5.1 VSM current state map

5.5 Process Mining Application

In this section, we will employ three process mining techniques. We will begin by creating an event log and performing process discovery. Next, we will conduct conformance checking of the discovered process model. Finally, we will proceed with the process enhancement phase to improve the process.

Process mining is a data mining technique that specifically focuses on processes rather than just data. Unlike traditional data mining methods that primarily analyze data, process mining aims to automatically generate models that describe process behavior using data from event logs (W. van der Aalst, 2016). Unlike manually created models with Value Stream Mapping (VSM), which require significant effort and data collection, process mining tools can automatically generate and analyze models based on the data loaded into these applications. This enables the creation and examination of automatic models for various processes and tasks repeatedly and efficiently.

There are few studies on production in the literature, but their implementation requires systems that collect data from production machines. In other words, real-time signals must be received from the machines and recorded in a database. Factories that have adopted or are transitioning to Industry 4.0 already have this data available in databases. This availability is one of the challenges of implementing process mining. However, since the factory is at the beginning stages of Industry 4.0 applications, the necessary information was manually retrieved from servers and processed. Additionally, a review of the literature revealed a lack of studies on the use of process mining in enterprises engaged in batch manufacturing, outsourced manufacturing, or those that produce numerous parts for later assembly.

In our current VSM analysis, it was previously identified that the casting and machining workshops were bottlenecks. Therefore, the process mining study will focus on these workshops.

5.5.1 Creating Event Log

The machines from which signals are received are the sand core machine, casting, automatic runner cutting, and machining, respectively. As we recall from the VSM study, we can’t see the sandblasting machine here because no signal is received from it. The production data of the sandblasting machine is not digitalized and cannot be held on servers.

In addition, since the parts produced after the processes in the machining shop are shipped to outsource suppliers for polishing. The issue encountered in data collection should be addressed as a separate project.

The first task is to create the event log. There is a MES system in the casting and machining workshops, and the signals are stored in the database. The data accessible in the database belongs to the moment of each part's process ending on each machine. In other words, the timestamp at the end of each process is taken from this database. However, this database does not contain information about which order or part the signal corresponds to. Therefore, a report of which machine processes which order and in which time interval is obtained from another database file of the same program. These two reports are combined in Excel VBA, and order number information is added next to the timestamp and machine information.

In addition, as a result of a process, a single part may not be produced each time a signal is received. If the mold connected to the machine is a two-cavity mold, two pieces of the same part are produced each time a signal is received. This situation is addressed by adding another record with the same timestamp to the event log if double production was performed, using the Excel VBA application.

In batch-type manufacturing, even production confirmations are handled in batch format. For instance, the sand core machine may produce 1,000 units, while the casting machine produces 500 units. However, it is not explicitly clear which 500 units produced by the casting machine were sourced from the sand core machine and which specific units were processed. Here, data is derived by applying the First In First Out (FIFO) assumption.

When the process mining application first emerged, it was primarily utilized in service sector applications, such as hospitals and insurance. To illustrate with an example from hospital practice: a patient arrives, a record is opened, followed by a preliminary examination, tests are conducted, the treatment method is decided, and finally, the patient is discharged, completing the relevant process. process mining applications involve maintaining consecutive records of events carried out by individual tasks and analyzing these records.

In batch-type manufacturing, there is only one production order number for the entire batch. However, when preparing the process mining event log, we must assign a production order number to each individual part passing through all machines. This is because we consider each produced part similar to a patient in a hospital scenario. That is, a part enters the system with a production order, is processed at the sand core machine, then the same number is used as it is processed at the casting machine. Subsequently, the part, still has the same production number, passes through the auto runner cutting and vertical machining finally exiting the system. Hence, for the VSM map we draw, each part produced as part of a production order for 7,500 units, is assigned a unique production order number in the event log. This is done by appending a "-" and the produced part number to the end of the 7,500-unit production order number. For instance, the first part passing through the system is given the production order number 2184788-1, and the second part produced is assigned the number 2184788-2, etc.

The production records for each unit in each machine are arranged in chronological order, from the earliest to the latest, and corresponding records are created. For example, the initial production of the sand core machine on 24th February 2023 at 22:01:08 is recorded and assigned the first order number. Subsequently, for the casting process, the production on 28th February 2023 at 15:18:59, marking the moment when the first casting process is performed, is recorded as the next step for the first order number. This method is consistently applied to prepare records for other production processes.

We need to keep in mind that the time range of all records is between 22nd February 2023 and 26th March 2023, and these records do not encompass all productions made within this period. The records only cover the production of the specific part, for which the VSM was conducted, to be worked on in the relevant machines and workshops.

Another issue to consider when preparing the event log is the scrap rates, which vary between machines. The scrap rate for the sand core machine is approximately 30%, while the scrap rate for the casting machine is 3.5%. Scrap productions must also be included in the event log to ensure that, upon completion of our process, no products remain in the system, and all have exited the system. For this purpose, the scrap rate for the relevant operation is calculated for the entire job. As each record is added to our event log in VBA, the scrap rate for the specific operation in that record is compared to the cumulative scrap rate recorded up to that moment. If the current scrap rate is lower than the total scrap rate, a scrap event is added for the subsequent operation of the relevant process, thereby completing the corresponding job.

The process for the relevant order started on 24th February 2023 at 22:01:08 on the sand core machine and was completed on the vertical machining center on 16th March 2023 at 19:07. Our event log contains a total of 40,679 events. The event log includes columns for Machine (Event), Machine exit time, Order number, and Quantity.

The "Machine (Event)" field specifies the machine where a particular process occurred, while the "Machine exit time" represents the time when each job is completed. The "Order number" serves as the identifier for the respective job. This number is generated by appending a hyphen to the end of the order number, followed by a sequentially assigned number starting from 1. In total, 12,248 parts were produced for this job, as previously explained.

The "Quantity" field indicates the quantity produced for the relevant job in the specific event. In this dataset, all rows indicate a quantity of 1 piece, as the production is transformed into single pieces when double production is made in 2-piece molds. Consequently, following the logic of process mining, a unique number is assigned to each produced part, making each part a distinct order within the system. The operations of each produced part are recorded sequentially, each with the same timestamp, and then the next part's record is added to the event log. The first rows of our event log as follows:

Table 5.2 Some records of event log

Machine (Event)	Machine Exit Time	Order Number	Quantity
Sand Core 4	24.02.2023 22:01	2184788-1	1
Casting 4	28.02.2023 15:18	2184788-1	1
Runner Cutting3	3.03.2023 22:29	2184788-1	1
Vertical Machining7	4.03.2023 23:00	2184788-1	1
Sand Core 4	24.02.2023 22:04	2184788-2	1
CoreScrap	24.02.2023 22:04	2184788-2	1
Sand Core 4	24.02.2023 22:06	2184788-3	1
Casting 4	28.02.2023 15:18	2184788-3	1
Runner Cutting3	3.03.2023 22:29	2184788-3	1
Vertical Machining7	4.03.2023 23:11	2184788-3	1
Sand Core 4	24.02.2023 22:08	2184788-4	1
Casting 4	28.02.2023 15:20	2184788-4	1
Runner Cutting3	3.03.2023 22:34	2184788-4	1
Vertical Machining7	4.03.2023 23:20	2184788-4	1
Sand Core 4	24.02.2023 22:08	2184788-5	1
CoreScrap	24.02.2023 22:08	2184788-5	1
Sand Core 4	24.02.2023 22:10	2184788-6	1
Casting 4	28.02.2023 15:20	2184788-6	1
Runner Cutting3	3.03.2023 22:34	2184788-6	1
Vertical Machining7	4.03.2023 23:25	2184788-6	1
Sand Core 4	24.02.2023 22:15	2184788-7	1
Casting 4	28.02.2023 15:22	2184788-7	1
Runner Cutting3	3.03.2023 22:36	2184788-7	1
Vertical Machining7	4.03.2023 23:30	2184788-7	1

Our process mining application will consist of three stages. First, we will make a process discovery, and then we will check the conformance of the discovered process. Then we will apply the enhancement phase for process improvement.

5.5.2 Process Discovery

In our process mining study, we utilize the ProM application. Initially, we create our 40,679-line event log using Excel VBA and save this log in an Excel file, which is then saved as a CSV file. We import this CSV file into the ProM application. The event log to be uploaded to the ProM application must be in XES format. Fortunately, the ProM application includes a plugin that converts CSV files to XES files. Therefore, we import the CSV file into the ProM application for conversion and further analysis.

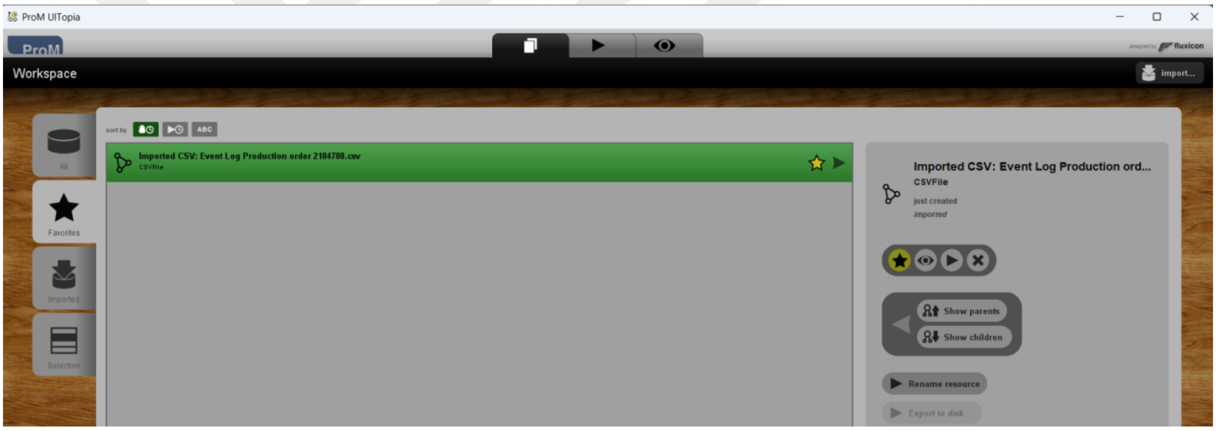


Figure 5.2 ProM workspace view

After this step, our file appears in the Workspace tab of the ProM application. By clicking the "Use Resource" option, we can view the plug-ins available for the uploaded file in the “Actions” tab.

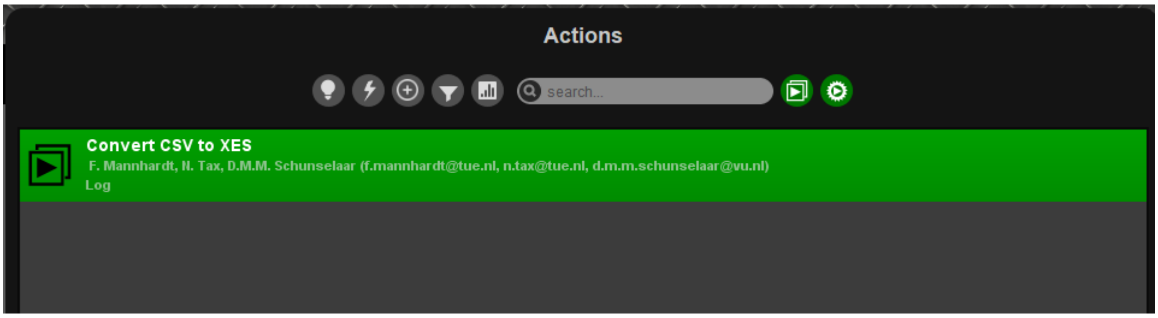


Figure 5.3 ProM actions, convert CSV to XES

The plug-ins that can be used with the uploaded file in the “Actions” tab are indicated in green. Here, we click on the "Convert CSV to XES" option and then press the "Start" button located at the bottom of the screen.

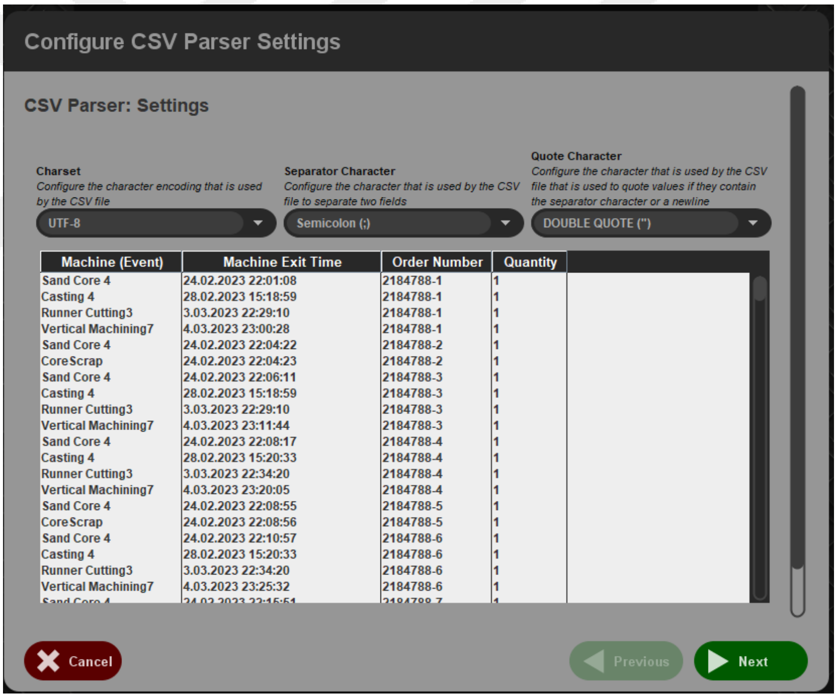


Figure 5.4 ProM event log convert to XES

Now we can see our imported event log and we continue by clicking next.

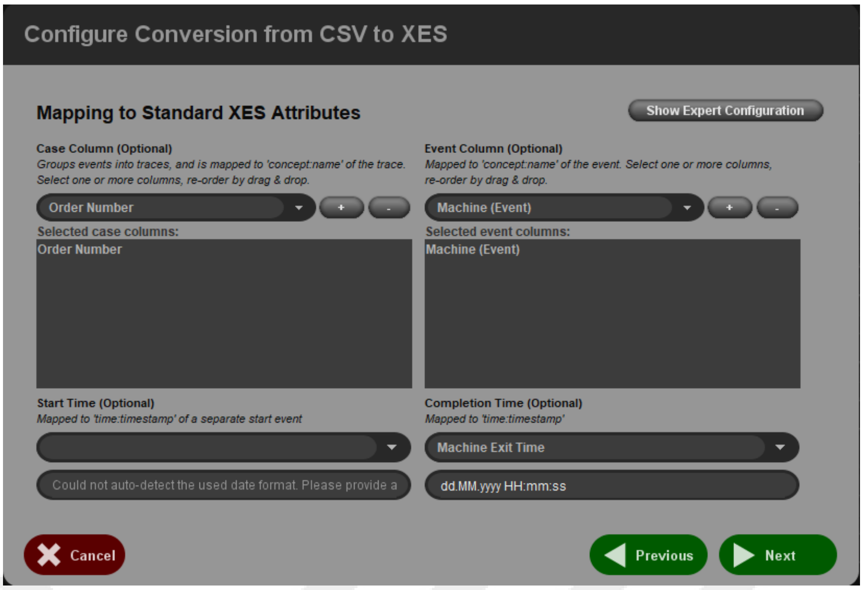


Figure 5.5 ProM convert to XES column configuration

Next, we proceed with the Mapping settings. In this step, we define the purpose of each column. We select the Order number as "Case," the machine column as "Event," and the time column as "Completion time." It is important to note that only the end time is recorded for each event, and the start time is not maintained in the database. After making these selections, we complete the process by clicking "Next" and then "Finish."

The dashboard screen that opens provides summary information about our event log. On the right is the time period in which the events took place is displayed (24.02.2023 22:01:08 - 16.03.2023 19:03:29). In our event log, there were 12,248 cases (indicating that 12,248 items were produced on the Core Press machine, which is our first operation), and a total of 40,678 events occurred.

In each case, a maximum of 4 events, an average of 3 events, and a minimum of 2 events occurred. When examining the operations for the relevant part (sand core, casting, runner cutting, and machining), we can see that there are a total of 4 operations. However, as explained previously, the scrapping process has also been added to the event log. Since some parts are scrapped after the core press machine, they exit the system after passing through the sand core and core scrap events.

Consequently, in these instances, the relevant case exits the system after only 2 events. Therefore, the number of events per case is a maximum of 4, a minimum of 2, and an average of 3. The width of the bars on the graph indicates the frequency of occurrence for the number of events.

The histogram below shows the event class, and in our records, the event and event class display the same values. The number of event classes is 10, which corresponds to the number of machines producing the part, including the scrap events added for each operation. Each part passes through 4 machines until production is completed. The core operation is performed on a single machine (Sand Core 4) until the order is completed, the casting operation is conducted on 2 different machines at different times (Casting 1 and Casting 4) until the order is completed, and runner cutting is done using 2 different machines (Runner Cutting 1 and Runner Cutting 3). Machining is completed using Vertical Machining 7 and Vertical Machining 8 machines. The total number of machines used in the order is 7. Additionally, if we include the artificial operations—Core scrap, casting scrap, and cutting scrap, which are added to the event log for the scrapped parts—we obtain a total of 10 event classes.

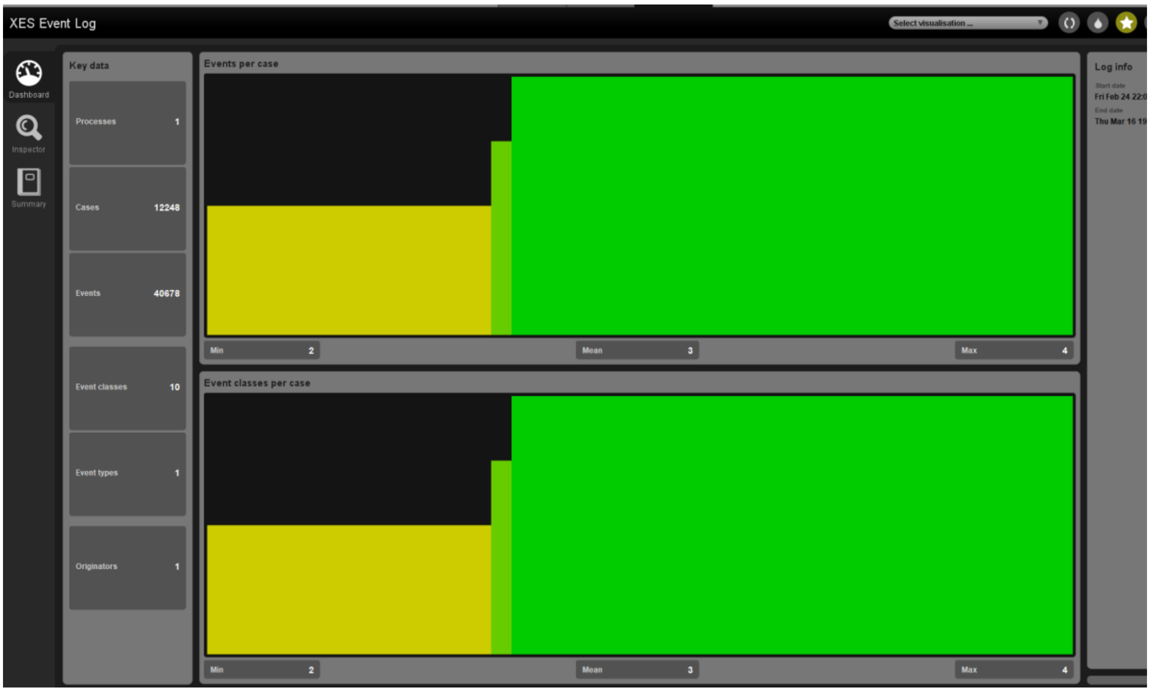


Figure 5.6 ProM dashboard view

On the left side of the screen, we can access the Inspector tab, which contains information about our records. There are three tabs at the top. In the first tab, the Browser tab, the cases from the event log are listed in the Instances section. As previously explained, a hyphen ("-") is added after the production order for the relevant part in the factory, and a new number is incremented by 1 for each case. When we click on each case, the events related to that case appear, along with the timestamps indicating when these events occurred.

On the screen below, the part that entered our system in the 10th place, numbered 2184788-10, is first produced on the Sand Core 4 Machine on 24th February 2023 at 22:29:08. Subsequently, the casting process begins 4 days later on the Casting 4 Machine on 28th February 2023 at 15:24:15. The runner cutting is performed on the Runner Cutting 3 Machine on 3rd March 2023 at 22:37:43, and finally, the machining process is completed on the Vertical Machining 7 Machine on 4th March 2023 at 22:36:40.

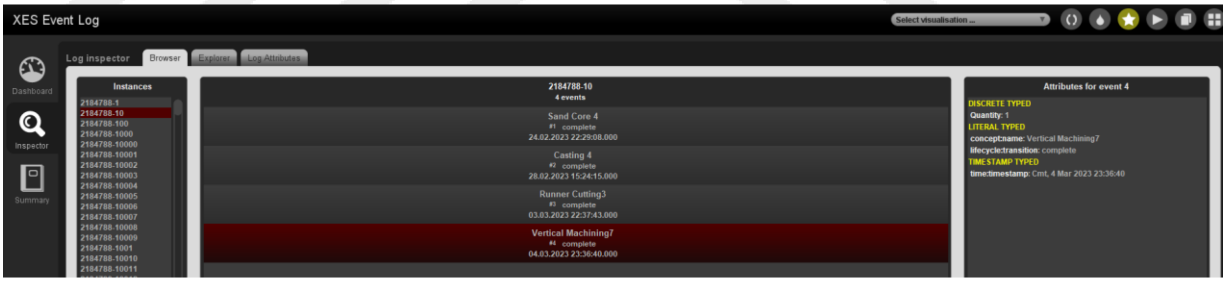


Figure 5.7 ProM inspector-browser view

When we click on the Explorer tab, we can see each case number, the number of events occurred in the relevant case, and the order in which the events occurred. Additionally, each event is colored on a scale according to its frequency of occurrence. Here, the ones that occur frequently are shown in green, and the ones that occur infrequently are shown in red.

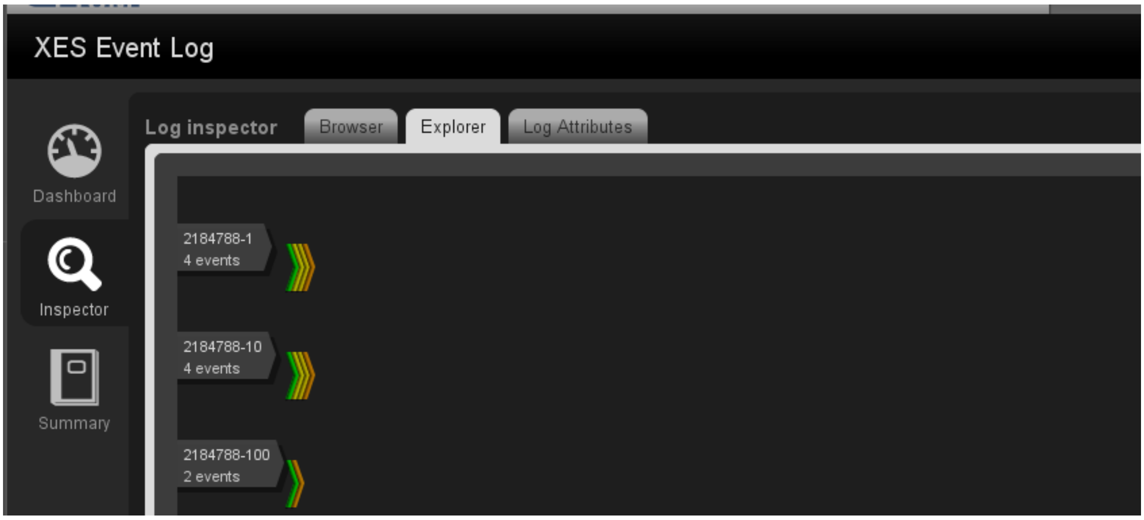
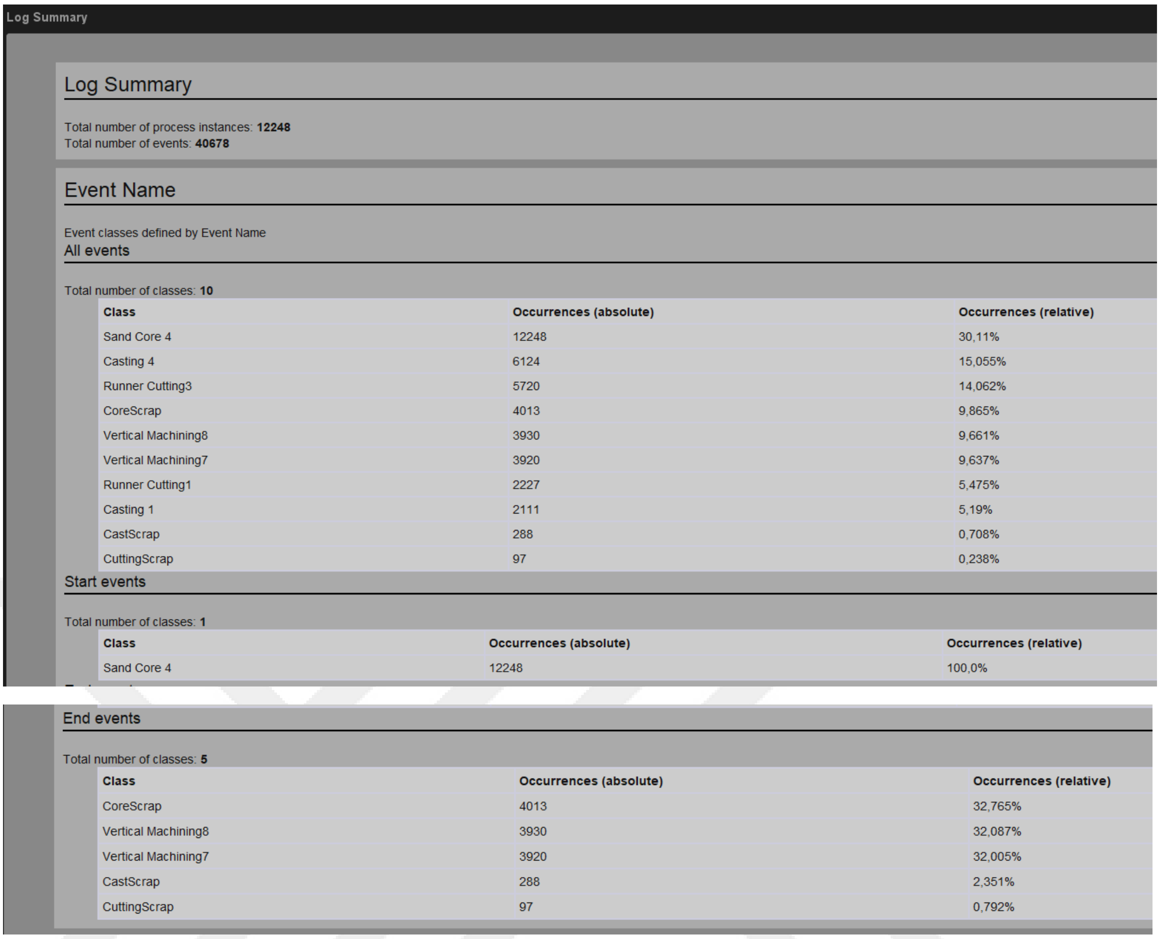


Figure 5.8 ProM inspector-explorer view

When we switch to the Summary tab in the ProM program, we can view summary information about our event log. Here, we first note that our records contain 12,248 cases, indicating that 12,248 parts entered our system, which corresponds to the total number of sand core parts produced for this order. We observe that a total of 40,678 events occurred and there are 10 different event classes. These event classes are listed below, along with their occurrence numbers and percentages

The Sand Core 4 Machine is the unique starting event, meaning that all parts were introduced into our system by the Sand Core 4 machine. Regarding the end events, which are the events where the parts leave the system, we see the artificially added scrap events. In addition, we observe the Vertical Machining 7 and Vertical Machining 8 events, where the machining operations are completed, and the parts actually exit the system in a finished state.


Log Summary

Log Summary

Total number of process instances: **12248**
Total number of events: **40678**

Event Name

Event classes defined by Event Name
All events

Total number of classes: **10**

Class	Occurrences (absolute)	Occurrences (relative)
Sand Core 4	12248	30,11%
Casting 4	6124	15,055%
Runner Cutting3	5720	14,062%
CoreScrap	4013	9,865%
Vertical Machining8	3930	9,661%
Vertical Machining7	3920	9,637%
Runner Cutting1	2227	5,475%
Casting 1	2111	5,19%
CastScrap	288	0,708%
CuttingScrap	97	0,238%

Start events

Total number of classes: **1**

Class	Occurrences (absolute)	Occurrences (relative)
Sand Core 4	12248	100,0%

End events

Total number of classes: **5**

Class	Occurrences (absolute)	Occurrences (relative)
CoreScrap	4013	32,765%
Vertical Machining8	3930	32,087%
Vertical Machining7	3920	32,005%
CastScrap	288	2,351%
CuttingScrap	97	0,792%

Figure 5.9 ProM log summary

When we select the Explore Event Log option in the select visualization section of the ProM program, we are presented with a summary of the paths followed by the cases. Here we can sort the cases in descending and ascending order according to the number of occurrences. Cases can also be sorted by number of events. There are various options like this available.

Below is a part of the screenshot of the same screen. Here, cases are grouped according to events, that is, according to the path they follow, and listed according to the number of occurrences. Accordingly, the most repeated path of events is the Sand Core 4 and Core Scrap path. The reason for this is the high scrap rate of operation of the sand core. While 12,248 parts enter our system, 4,013 of them, or 32.76%, are scrapped. The following path of events is Sand Core 4, Casting 4, Runner Cutting 3 and Vertical Machining 7.

Figure 5.10 ProM event log summary

First of all, we will do the process discovery process mining step with different algorithms and look at the results.

5.5.2.1 Alpha Miner Algorithm

When we examine the Petri net using the Alpha Miner algorithm in ProM, we observe that it accurately represents the general flow of parts and creates a realistic flow. The Petri net view is shown below. The differences between this model and the current state VSM map will be explained later. First, we will conduct the process discovery step of process mining using different algorithms and analyze the results.

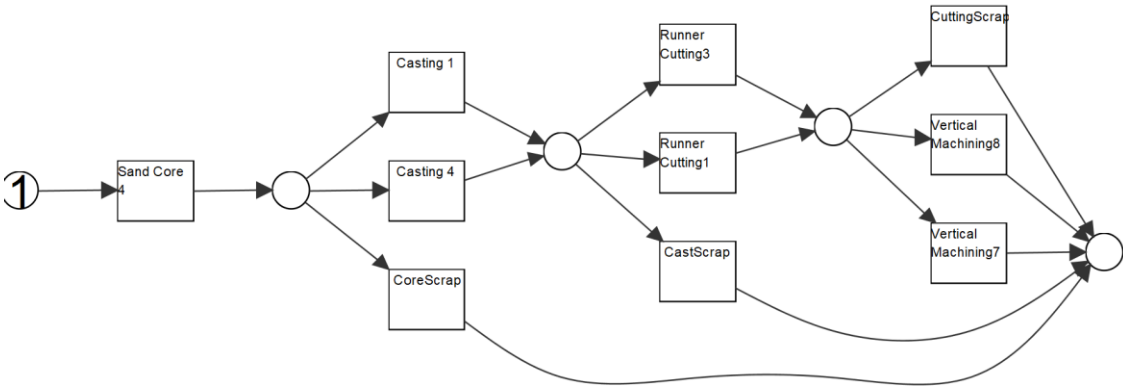


Figure 5.11 Alpha miner petri net model

5.5.2.2 *Heuristic Miner Algorithm*

When we look at the resulting model of the Heuristic Miner Algorithm, the input and output numbers for each event indicate the connections between events. Events are represented in dark or light colors depending on the number of occurrences. The model is similar to the Petri net of the Alpha Algorithm but provides additional information about our process, such as the number of occurrences and color coding based on event frequency.

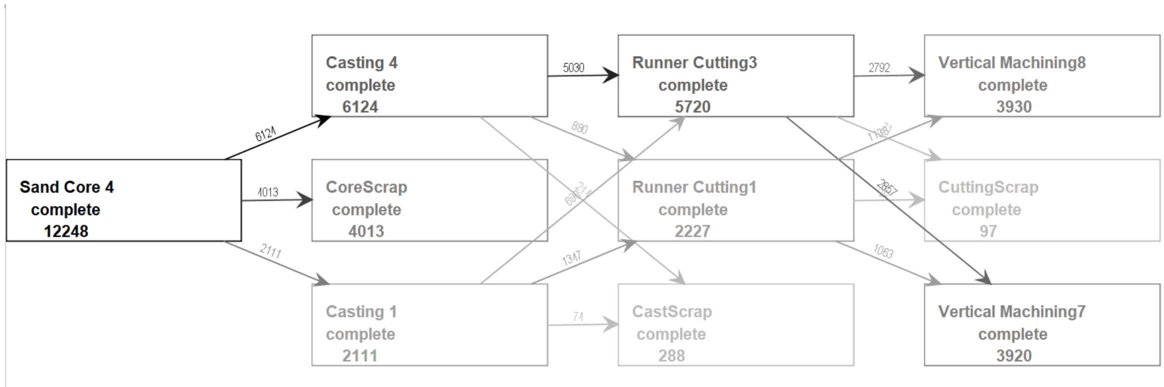


Figure 5.12 Heuristic miner model

5.5.2.3 *Fuzzy Miner Algorithm*

The resulting Fuzzy Model is shown below. Here, our process is represented as a tree and leveled. For each level, the number of cases that went through each event is indicated in the boxes. Additionally, the significance level of events is provided instead of the number of occurrences. These significance level values are used in the model to filter important and unimportant events, allowing us to focus on the significant events as mentioned before.

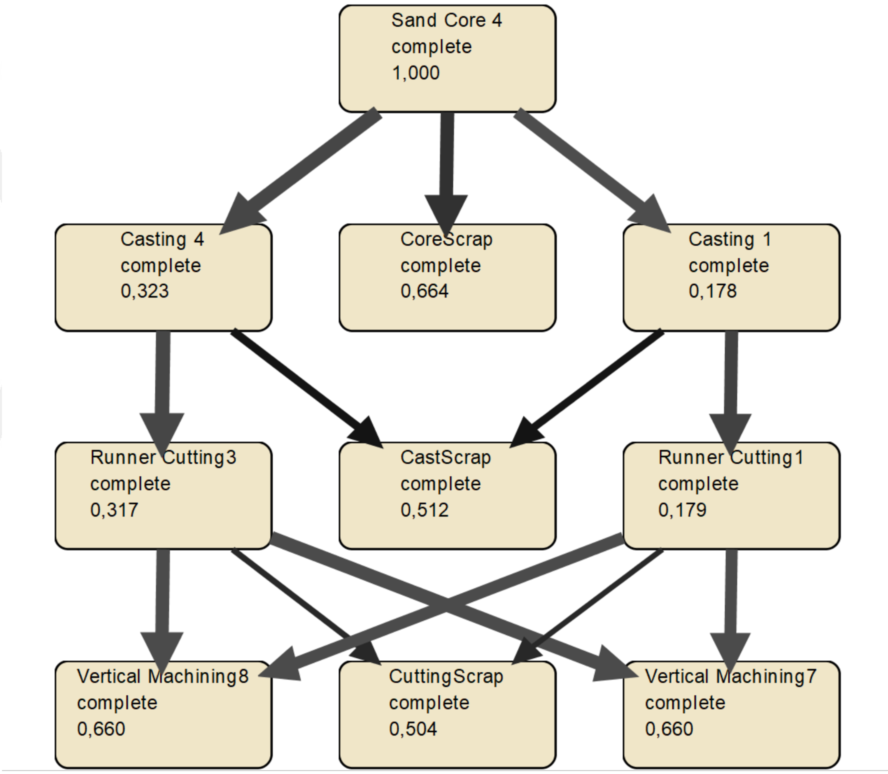


Figure 5.13 Fuzzy miner model

Below, the importance levels of events, here it is our machines, are shown on the Event Class Inspector page in the Fuzzy Model. Since all parts start from sand core production, the most important machine is sand core machine. In addition, the Core Scrap Event, which is the artificial event we created due to the high scrap rate, is the second most important event.

Figure 5.14 Fuzzy miner event class inspector

On another page of the Fuzzy Model Algorithm, the importance levels of the events are shown on a graphic. Three different significance metrics are displayed here. The first is the Frequency Significance Metric, based on the number of occurrences. The second is the Routing Significance Metric, which indicates the importance of the event in terms of routing. The third is the Aggregate Unary Metric, which combines these two metrics and averages them. Since the ProM screenshot is not readable in this text, the data is presented in the table below, along with a readable graph.

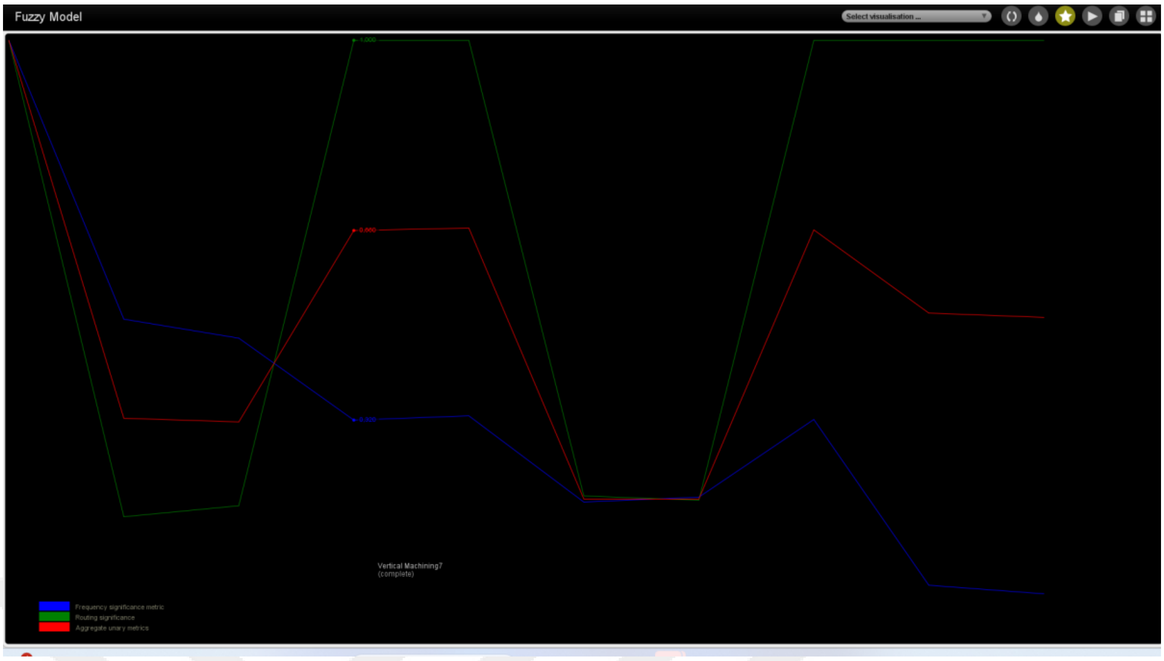


Figure 5.15 ProM fuzzy miner importance levels of events

If we analyze this table and graph excluding the artificial events, that is, the Scrap events.

- The Sand Core 4 Event is the most important event across all three metrics, because all cases start from this event.

- Since the Aggregate Unary Metric combines two metrics, other important events are Vertical Machining 7 and Vertical Machining 8. As we recall from the VSM current state map, these machines are the slowest. This indicates that they are the most critical machines that we must concentrate on.

Table 5.3 Fuzzy miner importance levels of events

	Sand Core 4	Casting 4	Runner Cutting 3	Vertical Machining 7	Core Scrap	Casting 1	Runner Cutting 1	Vertical Machining 8	Cast Scrap	Cutting Scrap
Frequency Significance metric	1	0,5	0,467	0,32	0,328	0,172	0,18	0,321	0,024	0,008
Routing Significance	1	0,146	0,166	1	1	0,184	0,178	1	1	1
Aggregate Unary metric	1	0,323	0,317	0,66	0,664	0,178	0,179	0,66	0,512	0,504

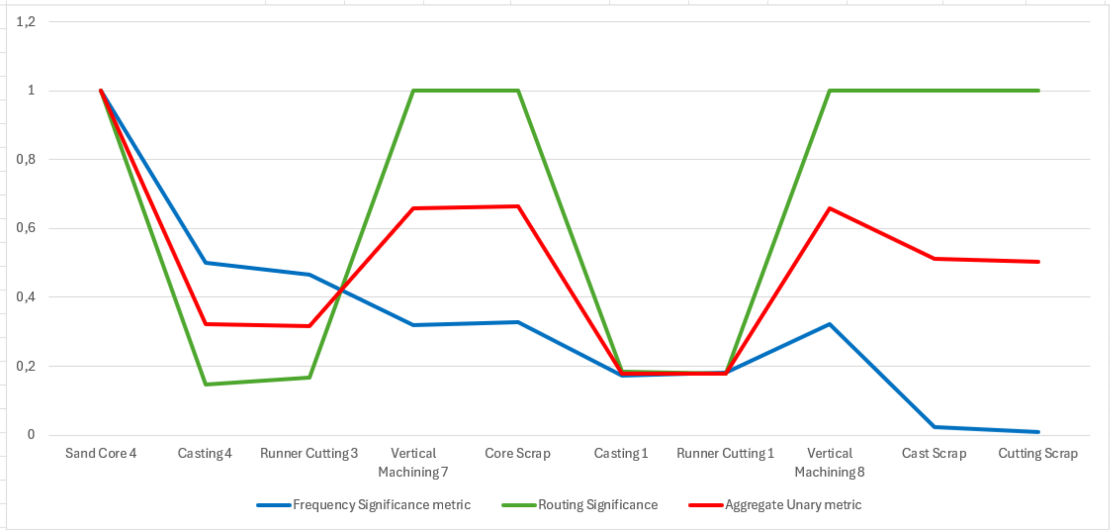


Figure 5.16 Fuzzy miner event importance levels of events

5.5.2.4 Inductive Miner Algorithm

When we run the Inductive Visual Miner Algorithm in ProM, our model appears as follows. In this model, there are two filters on the right: Activities and Paths. These filters allow us to filter activities and events based on their occurrence rate. In our model, both filters are set to 1, meaning all activities and events are displayed.

In addition, in our model, parts (cases) flow between machines (events) according to time changes, and our event log shows the flow of production for the entire time interval. The Inductive Visual Miner provides us replay capability. Events (machines) are colored according to the number of occurrences, with the number of occurrences of each event indicated inside the boxes. Furthermore, the cases (parts) are colored according to their waiting time in the system.

A detailed analysis of the results of this algorithm will be explained later.

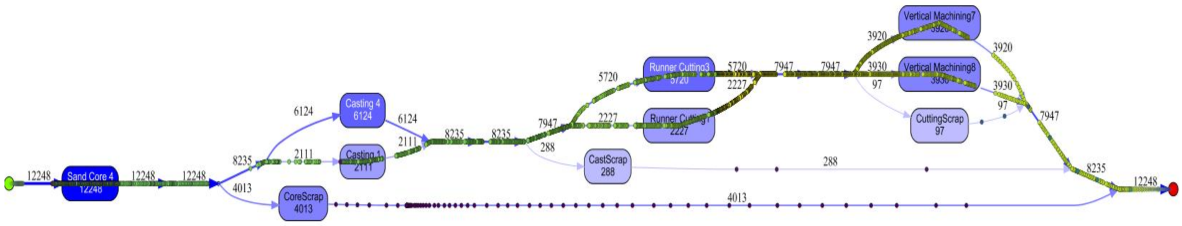


Figure 5.17 Inductive miner model

5.5.3 Conformance Checking

Process mining aims to produce a "representative" process model. The representativeness of a process model can be operationalized by ensuring that the model can replay all behaviors recorded in the log. This capability is a prerequisite for "fitness," a quality dimension often regarded as crucial for process models. Fitness assesses how well the process model aligns with the observed behavior in the event log and also highlights where the actual process deviates from the model (La Rosa & Soffer, 2013).

Although Alpha Miner and Inductive Miner algorithms give Petri Net as output, it is not available in Fuzzy Miner and Heuristic Miner algorithms.

There are many plug-in algorithms for conformance checking in the ProM application. Here, we used the “Multi-perspective Process Explorer” algorithm. Our process model is shown again in this plug-in below. Additionally, this screen contains summary data about our event log. The information panel in the lower right corner indicates that 12,248 cases occurred, meaning that 12,248 sand cores were produced on the sand core machine and introduced into the system. A total of 40,678 events occurred, with 12,248 pieces passing through these 40,678 events in 10 different event classes. This signifies that a total of 10 machines are included in our model, including artificially added machines. The first event occurred on 24th February 2023, and the last event occurred on 16th March 2023.

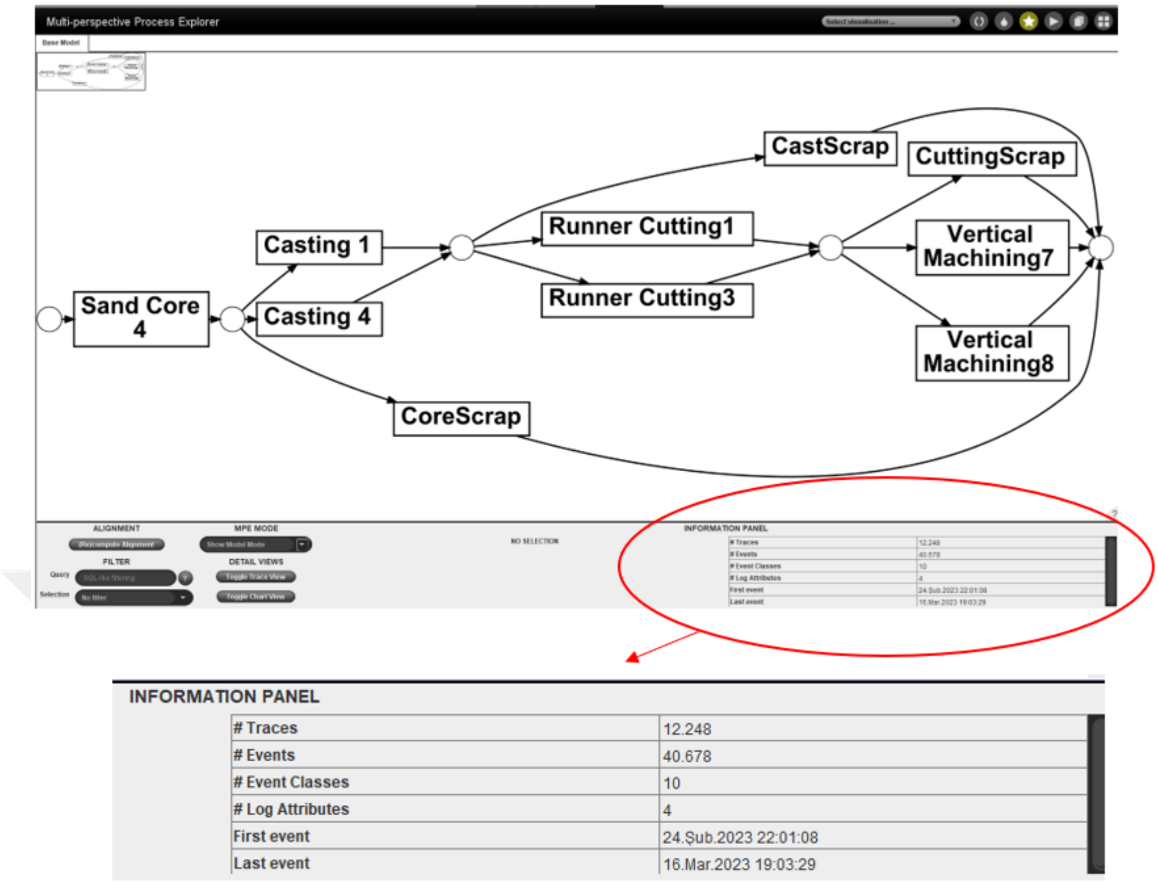


Figure 5.18 Multi-perspective process explorer

We can see that the Precision and Fitness value of our model is 100% (1) in the Information Panel section, when we select Show Fitness Mode from the MPE mode section on the panel in the lower left corner of this screen.

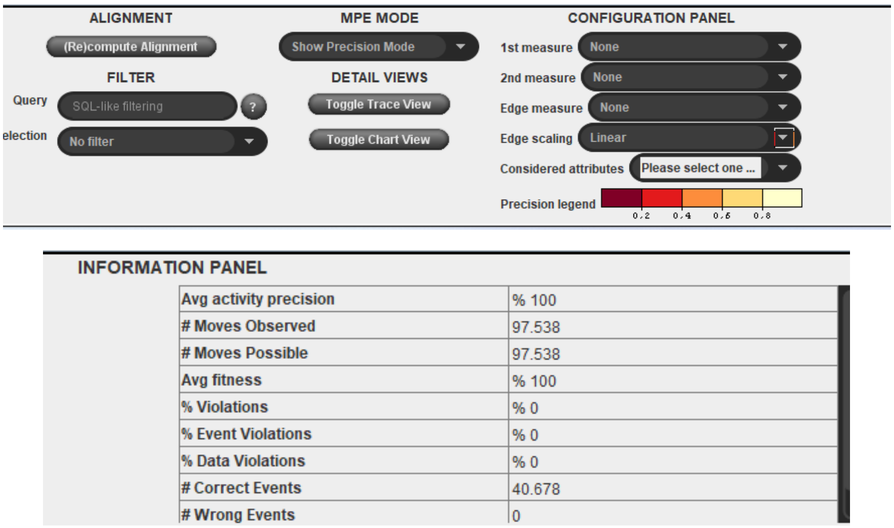


Figure 5.19 Multi-perspective process explorer information panel

So, the same pug-in is used to find the fitness value of the Inductive Miner algorithm and it is found as 1 again. Although we cannot use the same plug-in to find the fitness value of the Heuristic Miner Algorithm, it gives the fitness value of the model it produces within the Heuristic Miner algorithm, and it has been seen that the fitness value is calculated as 1, that is, 100%. In other words, the model produced completely reflects the data in the event log.

The same plug-in is used to find the fitness value of the Inductive Miner algorithm, and it is found to be 1 again. Although we cannot use the same plug-in to find the fitness value of the Heuristic Miner Algorithm, the Heuristic Miner Algorithm provides the fitness value of the model it produces within itself. It has been observed that the fitness value is calculated as 1, or 100%. In other words, the model produced completely reflects the data in the event log.

The ProM algorithm does not provide the fitness value of the model produced with the Fuzzy algorithm. Additionally, since this algorithm does not produce a Petri Net model as output, we cannot use the Multi-Perspective Process Explorer plug-in that we have used before.

The conclusions are;

- The process models we produced using the four algorithms are the same.
- We calculated both the fitness and precision values for the Alpha algorithm and the Inductive Miner algorithm. The fitness and precision values of the models we found are both 1. This means that the produced models completely reflect the data in the event log.
- Fitness value is calculated as 1 for Heuristic Miner Algorithm.
- The fitness or precision value of the model produced with the Fuzzy algorithm cannot be calculated.

We can explain why all produced models fully reflect the event log for the following reasons:

  - Although the number of cases is high in our model, the number of event class (machines) is low,
  - High-Quality Event Log; the event log is meticulously maintained, reducing errors and increasing the reliability of the models.
  - Processing the event log with the help of Excel VBA before uploading to the ProM,
  - Adding artificial machines, to the model for scrapped parts that will reduce the quality of the data and models.

Since the models we produced with all algorithms reflect the event log, we will now analyze our process using the performance analysis and replay features of the ProM application. Our goal is to identify problems and areas that need improvement.

5.5.3.1 Process Mining Effect on VSM Current State Map

Our study up to this point includes;

- The VSM current state map is drawn. This map contains all operations of the casted mixer body, from the supply of raw materials to the delivery of the product to the customer.

- Cycle time, changeover time, uptime, scrap rate, and WIP stocks for all processes are indicated on this map. While drawing the current state map, workshops are visited to observe which machines process the casted mixer body, and the map is created based on these observations.

- The event logs for an entire production order of the casted mixer body, specifically for casting and machining operations, are then prepared and uploaded to ProM for examination of the created models.

The models created by the ProM for casting and machining and VSM current state maps were compared.

- Although only one casting machine is shown on the current state map, two casting machines (Casting 1, Casting 4) are depicted in the ProM model for the casting operation. However, the two casting machines in the model do not work in parallel. Due to a fault in Casting 4, casting was stopped on the Casting 4 Machine and started on the Casting 1 machine. Therefore, it is not necessary to show two casting machines on the current state map.

- The sandblasting operation shown on the current state map isn't depicted in the ProM model. This is because our event log doesn't contain any data for the sandblasting operation, as the sandblasting machine is not digitalized and its data is not stored on servers. The sandblasting machine is not considered a critical operation for the faucet factory, which is evident when we examine the databox on the current state map.

- Similarly, although one auto runner cutting machine is shown on the current state map, two auto runner cutting machines are depicted in the ProM model. The runner cutting operation started on Runner Cutting 3, and on 08 March 2023 at 03:01:47, Runner Cutting 1 began to perform the same job, meaning they started to work in parallel. However, towards the end of the work, on 09 March 2023 at 09:45:01, Runner Cutting 3 stopped working, and the job was completed by only Runner Cutting

1 until 09 March 2023 at 22:14:23. Therefore, it is not necessary to show two auto runner cutting machines on the current state map.

- Finally, although one vertical machining center is shown on the current state map, two vertical machining centers are depicted in the ProM model. However, at the beginning of the work (and when the workshop visited for the VSM map), the machining operation started in Vertical Machining 7. The same operation was later started in Vertical Machining 8 and the two machines worked in parallel to complete the job. Therefore, both machines should be shown on the current state map.

- As a result, the current state map revised using the process mining model by adding the number of machines to the process boxes. We write 1 except for the vertical machining center and outsource polishing. Two machines are used for machining the parts. We do not write the number for outsource polishing because the number of machines and workers varies according to the number of parts that need to be polished.

- The Revised VSM current state map is as below in Figure 5.20.

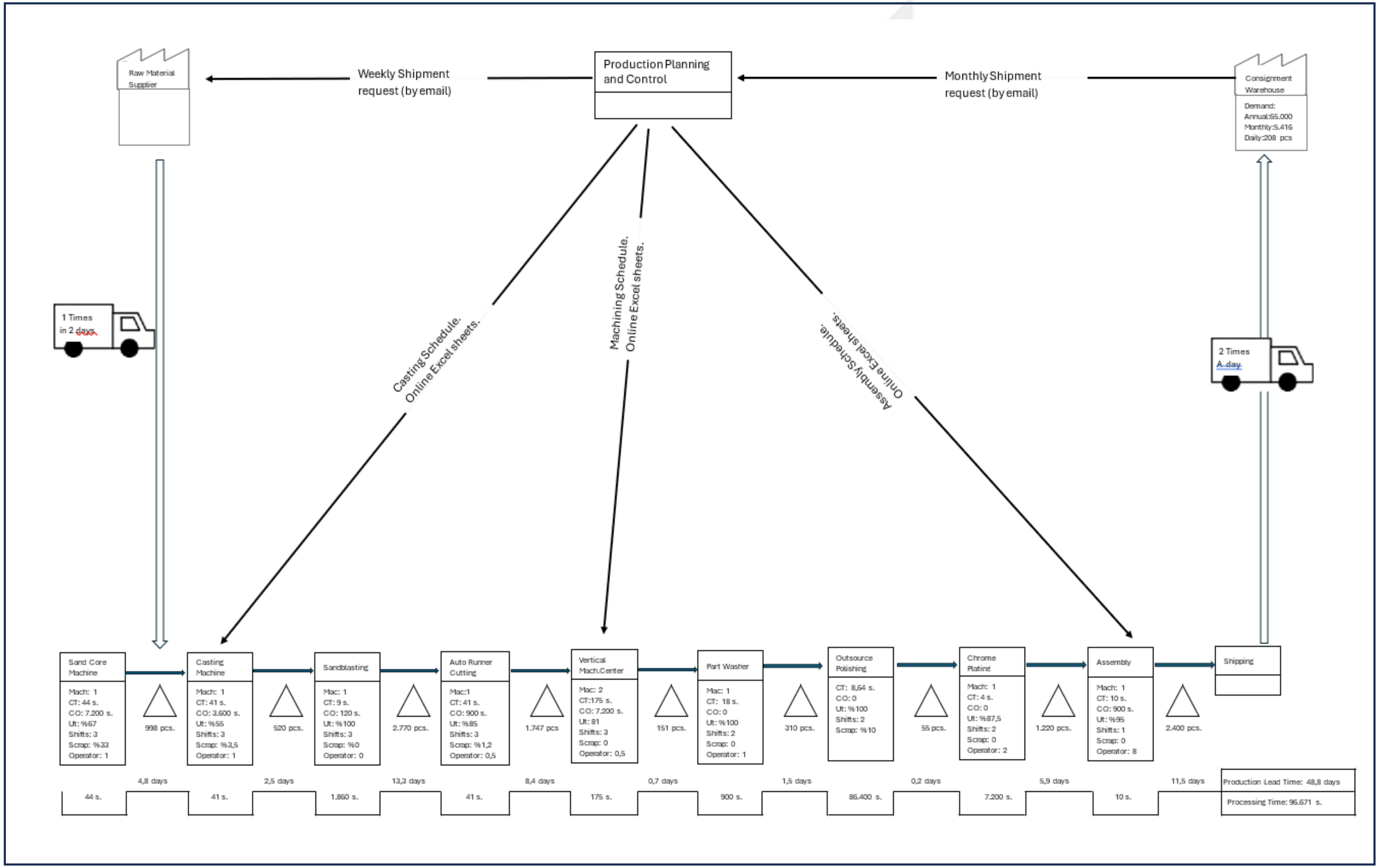


Figure 5.20 The revised VSM current state map

5.5.4 *Process Enhancement*

We will use the Visual Inductive Miner and Multi-Perspective Process Explorer algorithms, previously employed in Conformance Checking, to analyze performance. We will leverage the performance view section, which was demonstrated on this screen earlier.

We found it highly beneficial to use the Inductive Visual Miner algorithm along with the Multi-Perspective Process Explorer for performance analysis. This algorithm enables us to replay all the events in our event log from start to finish, providing a bird's eye view of our process—something not feasible in real life due to the physical distance between machines and their locations in different workshops. Additionally, this algorithm provided us with information about which machine operated during specific time intervals.

While the Visual Inductive Miner algorithm typically performs this replay, it also shows live queue lengths for each machine. Initially, we could not view the queue lengths in the ProM 6.12 version we were using. To address this, we contacted Sander Leemans, the author of the algorithm. He informed us that using an older version of ProM would allow us to see the queue lengths. Consequently, we used ProM Lite 1.4. Although this version did display queue lengths for each machine temporarily, the high number of events in our model caused the computer to interrupt, as its capacity was insufficient. Therefore, queue length information could not be obtained.

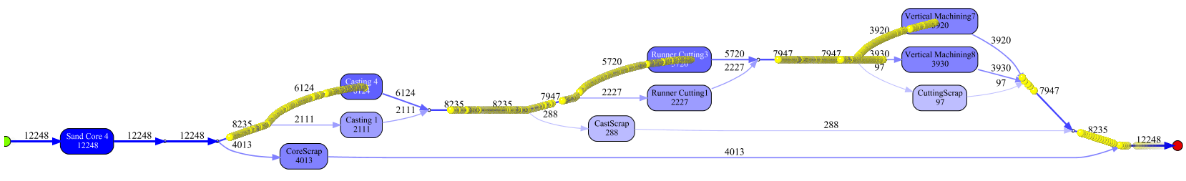


Figure 5.21 Inductive visual miner replay screenshot

The Multi-Perspective Process Explorer provides four different views for throughput time in the performance view: maximum, minimum, average and median throughput times. For each process specified as throughput time, the sum of the waiting and processing times of the parts for that process is given. Additionally, the number of parts passing through each process is provided. The boxes indicating each process are colored according to the number of passing parts, with the darkest color representing the process with the highest number of parts. This helps identify which operations require the most attention. Furthermore, the data obtained from the performance views are summarized in Table 5.

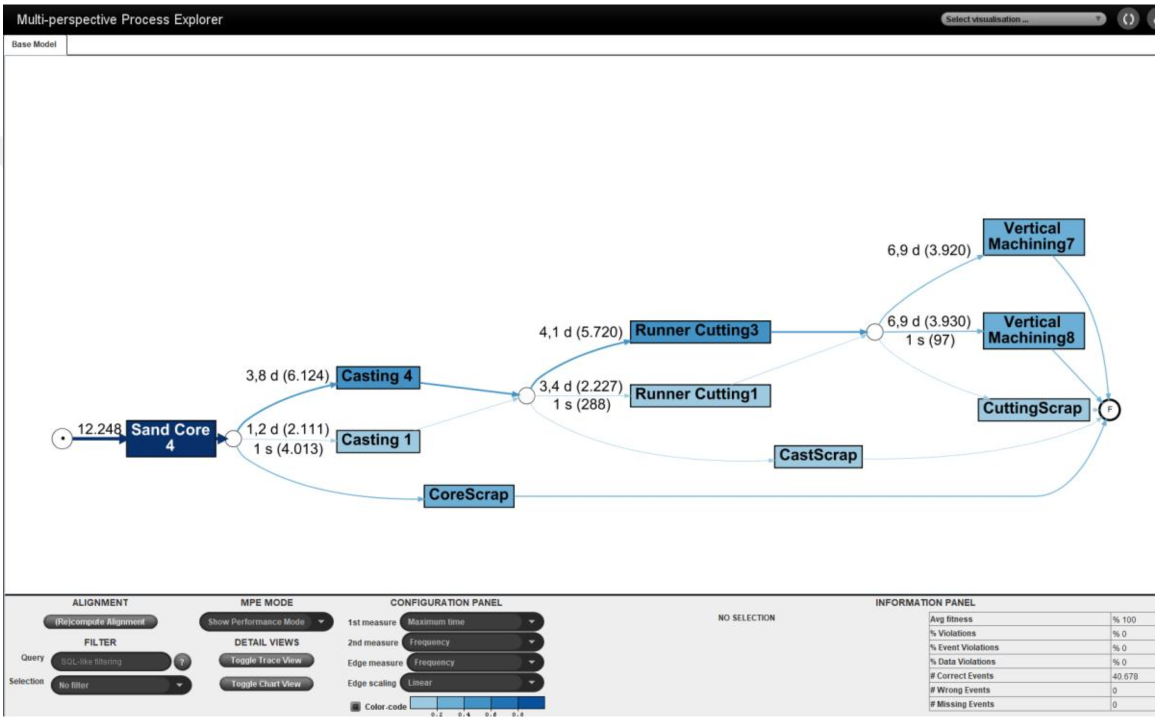


Figure 5.22 Performance view - maximum throughput time

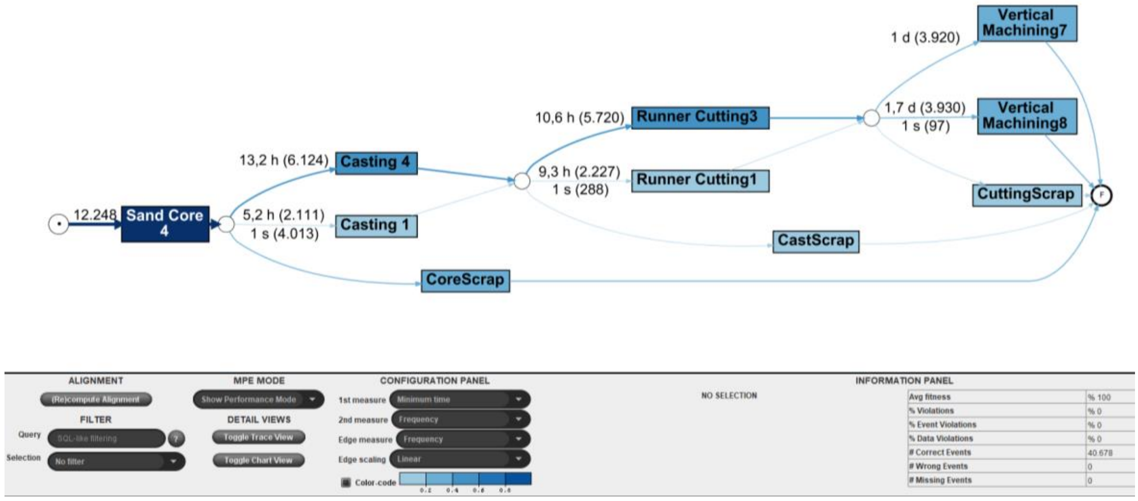


Figure 5.23 Performance view - minimum throughput time

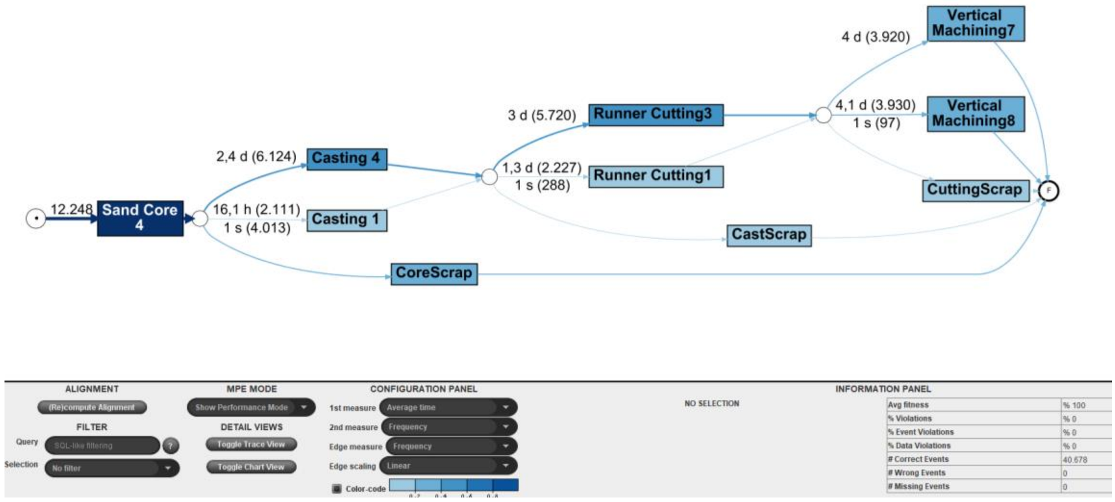


Figure 5.24 Performance view - average throughput time

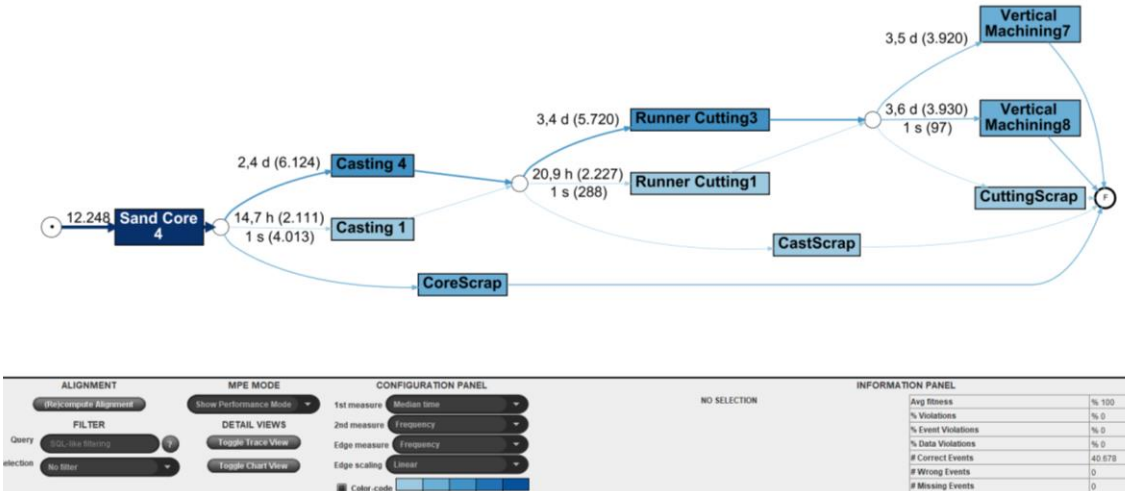


Figure 5.25 Performance view – median throughput time

Table 5.4 Summarized performance view data

2184788 - 24.02.2023 - 16.03.2023	Sand Core 4	Casting 4	Casting 1	Runner Cutting 3	Runner Cutting 1	Vertical Machining 7	Vertical Machining 8
Number of Parts	12.248	6.124	2.111	5.720	2.227	3.920	3.930
Total Number of Parts	12.248	8.235		7.947		7.850	
%	100%	74,37%	25,63%	71,98%	28,02%	49,94%	50,06%
Scrap Rate	33%	3,50%		1,22%			
Start Time	24.02.2023 22:01:08	28.02.2023 15:18:59	7.03.2023 15:30:19	3.03.2023 22:29:10	8.03.2023 03:01:47	4.03.2023 23:00:28	6.03.202309:40:16
Finish Time	9.03.2023 06:47:07	7.03.2023 01:12:55	9.03.2023 12:54:56	9.03.2023 09:45:01	9.03.2023 22:14:23	16.03.2023 18:45:10	16.03.2023 19:03:29
Maximum Waiting Time	0	3,8 d	1,2 d	4,1 d	3,4 d	6,9 d	6,9 d
Minimum Waiting Time	0	13,2 h	5,2h	10,6 h	9,3 h	1,7 d	1 d
Average Waiting Time	0	2,4 d	16,1 h	3 d	1,3 d	4 d	4,1 d
Median Waiting Time	0	2,4 d	14,7 h	3,4 d	20,9 h	3,5 d	3,6 d

The line balance rate is a crucial indicator for evaluating the performance of a production line, as it reflects the uniformity of the load across each process. Generally, the higher the line balance rate, the less waiting time there is between processes, leading to reduced waste and fewer work-in-process (WIP) stocks. The equation 5.1 for this variable is given by (Xiao & Shao, 2018) :

$$Line\ Balance\ Rate = \frac{\sum_{t=1}^{n} C/T_t}{C/T_{max} * N} \tag{5.1}$$

Here C/T_t is represents the cycle time of the t'th operation, C/T_{max} is the maximum cycle time and N is the number of machines. The cycle times of each operation are shown below and on the VSM current state map. The line balance rate is calculated as 61%. However, after considering the high scrap rates and low uptime, we decided to incorporate these data into the cycle time. Consequently, the line balance rate for scrap-added cycle times is 68%, and for scrap and uptime-added cycle times, it is 77%. The impact of the line balance rate can also be observed in the graphics below. The data used here are taken from the VSM map.

Table 5.5 Line balance rate table

	Sand Core	Casting	Runner Cutting	Vertical Machining	Max Cycle Time	Line Balance Rate
Scrap Rate	33%	3,5%	1,2%	0,0%		
Uptime	67%	55%	85%	81%		
CT (seconds)	44	41	41	88	88	61%
CT (seconds) - Scrap Added	66	42	41	88	88	68%
CT (seconds) - Scrap and Uptime Added	98	77	49	108	108	77%

Figure 5.26 Line balance rate graphics

The findings from our performance analysis of the model, using all the information in conjunction with the VSM current state map, are as follows:

When we examine the scrap rates: The scrap rate in the sand core is 33%. While the scrap rates in other processes do not significantly impact the flow, the high scrap rate in the core press disrupts the flow.

When we examine the uptimes: The process with the lowest uptime is casting, followed by the sand core operation.

When we examine the line balance rate graphics, we can easily see that the first bottleneck operation is vertical machining. The second bottleneck operation is the sand core due to the high scrap rate and low uptime.

When we examine the Number of produced parts:

- Casting 4 processes 75% of the casting, while Casting 1 handles 25%. When Casting 1 breaks down, the team transfers the mold from Casting 4 to Casting 1, continuing and completing production on Casting 1.

- Runner Cutting 3 handles 72% of the runner cutting process. When the cast pieces accumulate, Runner Cutting 1 also begins cutting the same pieces. Thus, both machines perform cutting for a while, and then Runner Cutting 1 completes the order.

- Vertical Machining 7 and Vertical Machining 8 completed the machining process of the order simultaneously. Therefore, we revised our VSM current state map accordingly.

When we examine the waiting times:

- Due to the absence of an additional drying process, sand cores are required to wait approximately one shift before casting. This waiting period allows the sand cores to dry naturally.

- The minimum waiting time for Casting 4, which initiates the casting operation first, is 13.2 hours—exceeding the approximate one-shift waiting time required for the core. In contrast, the minimum waiting time for Casting 1, to which the casting mold is subsequently transferred, is 5.2 hours, closely aligning with the core's waiting time. Therefore, there is no significant additional waiting time or excess WIP stock for the casting process.

- According to our event log, the first core operation began on February 24, 2023, and the first casting operation began on February 28, 2023, four days after the core operation started. During this period, core stock is accumulated.

- According to our event log, both the core operation and the casting operation are completed on March 9, 2023. The core stock accumulated over the initial four days is consumed by the time the casting process is completed.

- To accelerate the flow, we need to reduce the scrap rate or decrease the cycle time of the core press, allowing the processes to proceed simultaneously.

- Casted parts must wait for 1 shift for cooling before going through the runner cutting process.

- The minimum waiting time is 10.6 hours for Runner Cutting 3 and 9.3 hours for Runner Cutting 1. This duration slightly exceeds the time required for the parts to cool.

- Although the casting operation began on February 28, 2023, the runner cutting operation did not start until March 3, 2023. In other words, the cutting process started five days after the casting operation began. Consequently, the runner cutting process was performed on two machines simultaneously on March 8 and 9, 2023.

- When we examine the line balance rate graphic, we see that runner cutting has the shortest cycle time. To improve uptime for runner cutting, the casting workshop manager ensures there is no part shortage due to the casting operation. However, this approach ultimately disrupts the entire system.

- To accelerate the flow, the runner cutting process must begin one shift after the casting process starts. One or two changeovers can be performed, as the changeover time for runner cutting is 15 minutes (900 seconds). Starting the cutting process earlier will also allow the subsequent process to begin sooner.

- After the runner cutting process started on March 3, 2023, the machining process began first on vertical machining 7 on March 4, 2023. Vertical machining 8 started processing two days later.

- To accelerate the flow, both machines must start the machining process simultaneously and immediately after the runner cutting process begins.

- Additionally, despite the simultaneous operation of the two machines, parts are kept waiting for an average of four days, indicating that the cycle times of the machines are long. Vertical machining 7 and vertical machining 8 are single unit machines. It is suggested that using different machines with shorter cycle times available in the factory may yield better results.

5.6 Create Future State Map

Our aim is to reduce throughput time and increase flow while drawing the future state map. To achieve this, we will focus on the problems identified in the previous section, particularly those related to the bottleneck machines. The scope of the study includes only the processes analyzed through process mining; we didn't change other processes in the Future State VSM because we didn't analyze them.

In the previous section, we identified the first bottleneck as the vertical machining Process. This bottleneck is primarily due to the high cycle time of the machine. The recommended solution is using another multi-unit machine with a lower cycle time.

The current market demands a wide variety of special products in low quantities, leading to a decrease in the lot sizes produced by the factory. In the production line where we examined value stream and process mining, there are 20 types of parts with monthly demands of 1,000 pieces or more, while there are 159 types of parts with lower demands.

We based our study on a part that was machined on a single-unit machine. These single-unit machines were specifically purchased for machining parts with low lot sizes. Therefore, they should be used to process the 159 types of parts mentioned earlier. However, since the part we focused on in our study has an order size of 7,500 pieces, it needs to be processed on a multi-unit machine with low cycle times.

In the previous section, we identified the sand core Machine as the second bottleneck. This machine's bottleneck issue is not related to cycle time, as illustrated by the line balance rate graphs on the left, which display only cycle times. Despite its low cycle time, the machine's scrap rate is 33% and its uptime is 67%, causing it to be a bottleneck.

To address this, we initiated a KAIZEN project aimed at reducing the scrap rate. The goal of this KAIZEN is to lower the sand core scrap rate from 33% to 16%, achieving a 50% improvement.

Additionally, we examined machine downtimes to improve machine uptime. KAIZEN was initiated for tracking the life of core molds, addressing the mold revision downtimes that caused the most time loss. We also started KAIZEN projects to reduce mold and machine cleaning times, which are significant sources of downtime. Furthermore, a SMED (Single-Minute Exchange of Dies) initiative was launched to reduce the changeover time of the sand core machine from 7,200 seconds to 3,600 seconds, aiming for a 50% improvement.

As a result of these activities, our goals are to increase machine uptime to 80% and reduce the setup time for the sand core machine to 3,600 seconds. Additionally, a FIFO (First In, First Out) queue has been added between the sand core machine and the casting machine, with a WIP (Work in Process) stock of 400 pieces due to the waiting time of sand cores before casting. Consequently, our line balance rate graph is as follows:

Table 5.6 Revised line balance rate table

	Sand Core	Casting	Runner Cutting	Vertical Machining	Max Cycle Time	Line Balance Rate
Scrap Rate	16%	3,5%	1,2%	0,0%		
Uptime	80%	55%	85%	81%		
CT (seconds)	44	41	41	88	88	61%
CT (seconds) - Scrap Added	52	42	41	88	88	64%
CT (seconds) - Scrap and Uptime Added	65	77	49	108	108	69%

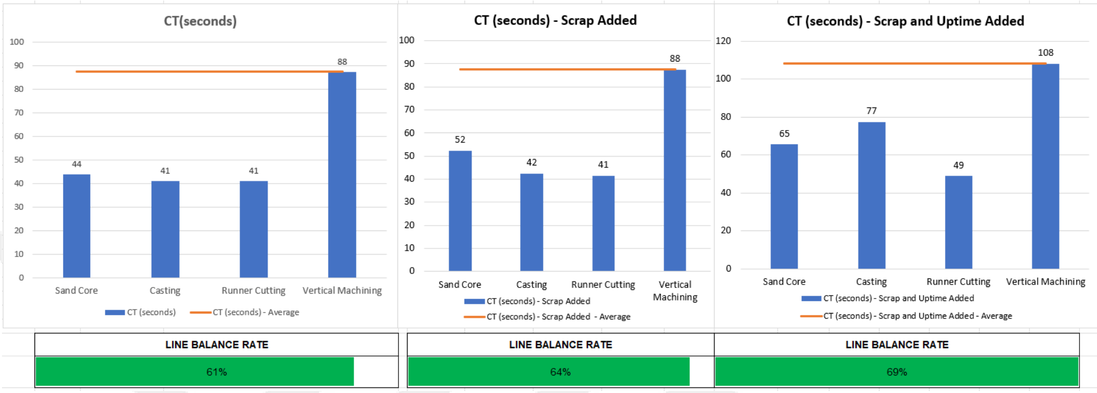


Figure 5.27 Revised line balance rate graphics

The second bottleneck operation is now casting. Although the scrap rate for casting is low, the uptime is insufficient, making it a bottleneck. To improve machine uptime, we examined casting downtimes and initiated a KAIZEN project for tracking the life of casting molds, addressing the mold revision downtimes that caused the most time loss. Additionally, we started KAIZEN projects to reduce mold and machine cleaning times, which are also significant sources of downtime.

As a result of these efforts, we aim to increase machine uptime to 65%. Consequently, the revised line balance rate graph is as follows. We also added a supermarket buffer between casting and sandblasting. Sandblasting will begin once the supermarket is full and will continue until it is empty. The supermarket consists of three cases, each containing 200 pieces. The casted parts must wait for a shift before sandblasting. Since 400 pieces can be cast in a shift, the average WIP (Work In Progress) stock between casting and sandblasting is 500 pieces.

Table 5.7 Second revised line balance rate table

	Sand Core	Casting	Runner Cutting	Vertical Machining	Max Cycle Time	Line Balance Rate
Scrap Rate	16%	3,5%	1,2%	0,0%		
Uptime	80%	65%	85%	81%		
CT (seconds)	44	41	41	88	88	61%
CT (seconds) - Scrap Added	52	42	41	88	88	64%
CT (seconds) - Scrap and Uptime Added	65	65	49	108	108	67%

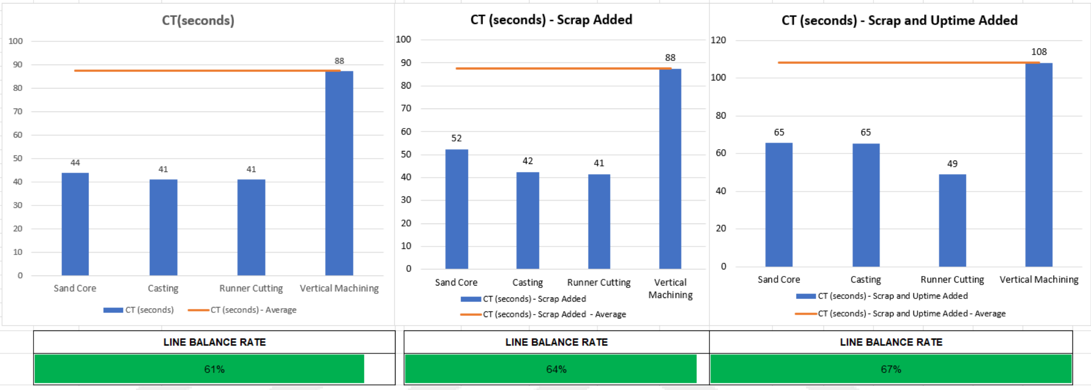


Figure 5.28 Second revised line balance rate graphics

As previously mentioned, runner cutting starts 5 days after the casting operation begins to avoid part shortages. However, this delays the start of the machining process. To prevent this, we need to start runner cutting after 3 shifts of casting and perform changeovers when parts are finished. Runner cutting will resume after accumulating parts for 3 shifts. Therefore, a supermarket buffer is added before runner cutting, with a WIP stock averaging between 400 pieces (one shift) and 1,200 pieces (three shifts).

When examining the scrap and uptime added line balance rate, it is 77% in Figure 5.26, 69% in the revised graph Figure 5.27, and 67% in the second revised graph Figure 5.28. Although it appears to worsen, the real reason is the long cycle time of the vertical machining process. As we improve uptime and scrap rates in sand core and casting, and add cycle time improvements, the line balance rate decreases because the difference between them, and the vertical machining cycle time widens.

To solve this problem, as mentioned before, we switched from vertical machining to a transfer machining center with a lower cycle time. The transfer machine center has a cycle time of 34 seconds for this part and an average uptime of 61%.

Table 5.8 Final revised line balance rate table

	Sand Core	Casting	Runner Cutting	Transfer Machining		Max Cycle Time	Line Balance Rate
Scrap Rate	16%	3,5%	1,2%	0,0%			
Uptime	80%	65%	85%	61%			
CT (seconds)	44	41	41	34		44	91%
CT (seconds) - Scrap Added	52	42	41	34		52	81%
CT (seconds) - Scrap and Uptime Added	65	65	49	56		65	90%

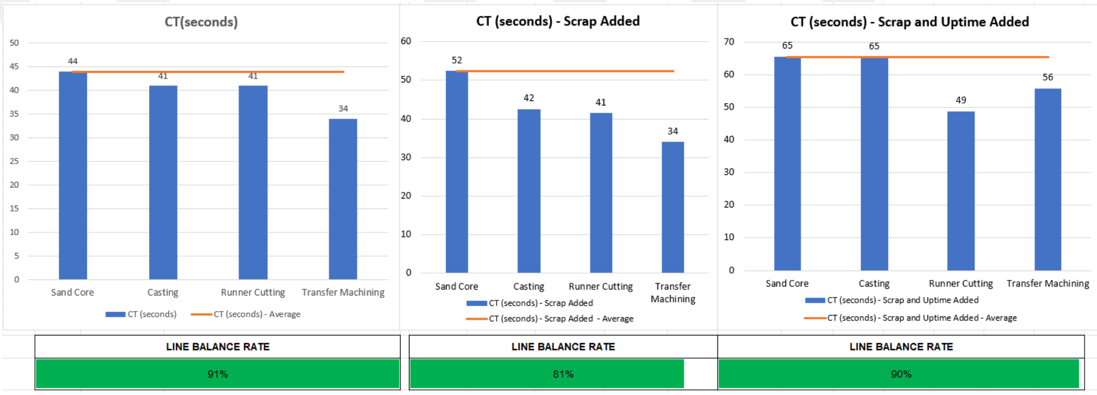


Figure 5.29 Final revised line balance rate graphics

In the final state, the scrap rate and uptime added cycle times of the sand core and casting operations are equal, resulting in an improved line balance rate of 90%. The future state VSM Map is shown below. The production lead time has improved by 27%, decreasing from 48.8 days to 35.6 days.

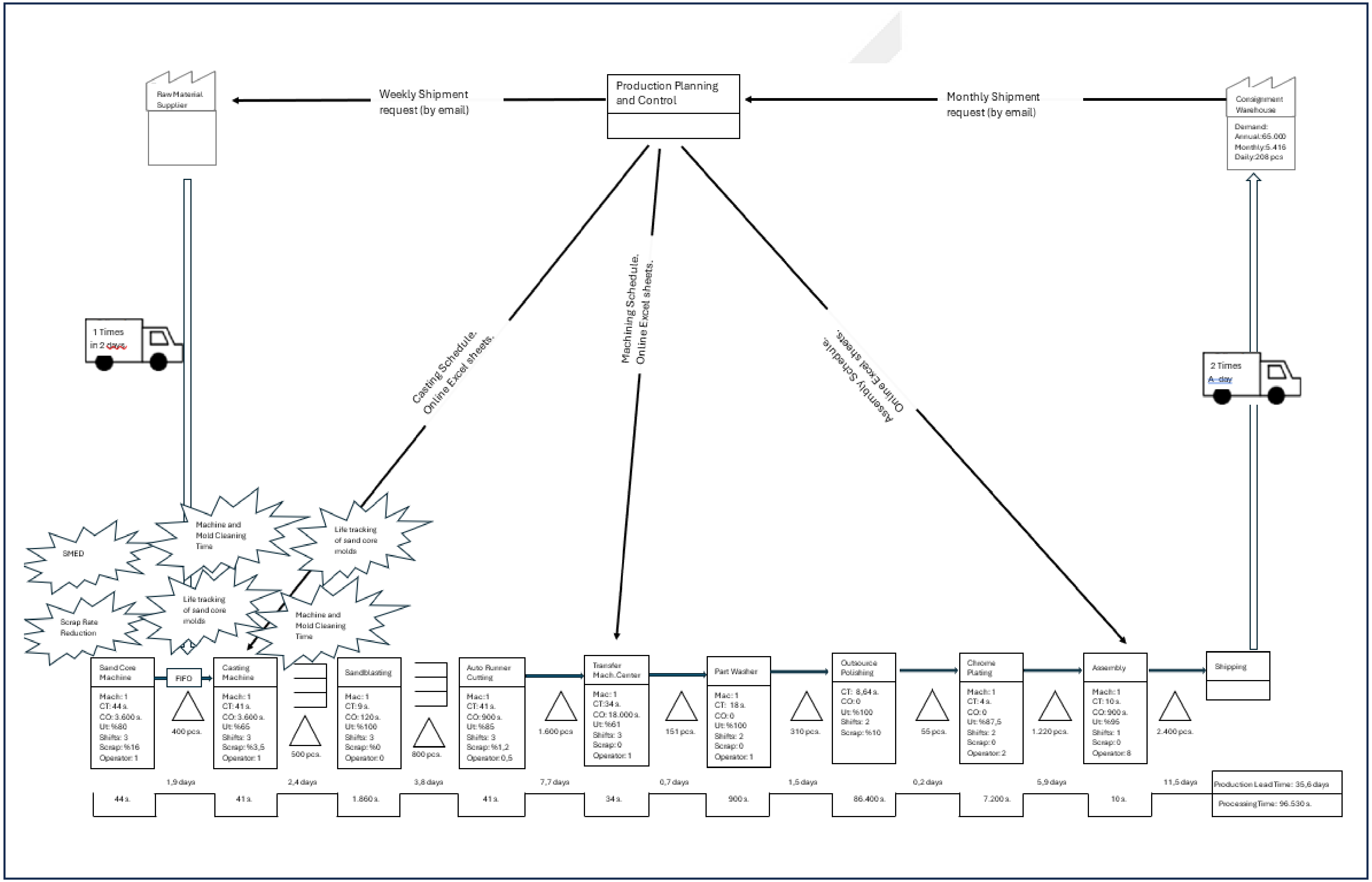


Figure 5.30 Future state VSM

CHAPTER SIX

COMPARISON AND CHALLENGES OF USING VSM AND PROCESS MINING

This chapter first presents comparison of the using of VSM and process mining. It then outlines the challenges encountered in both our study and the literature regarding the using of VSM and process mining.

6.1 Comparison of Using Value Stream Mapping and Process Mining

The comparison of VSM and process mining reveals significant differences in their application and effectiveness. VSM requires the involvement of a larger team, consuming substantially more time, while process mining is more efficient, involving fewer people and requiring less time. VSM's reliance on individuals' abstract reasoning limits the level of detail and reliability of the information, often resulting in difficulties in process repetition and disagreements regarding activity durations. Conversely, process mining provides detailed logs directly from the system, ensuring higher accuracy and ease in repeating the process whenever needed. While VSM offers a static snapshot of the process, process mining allows for the analysis of different process moments, including the ability to assess system compliance by identifying all occurrences, even those that deviate from the required process. This comparison underscores the advantages of process mining in terms of efficiency, accuracy, and the dynamic analysis of manufacturing processes. The comparison table derived from Nawcki et al. (2021) was adapted to fit the context of our case study (Nawcki et al., 2021).

Table 6.1 Comparison of VSM and process mining (Nawcki et al., 2021)

Comparison	**VSM**	**Process mining**
Numbers of people involved	22 – the entire Production and Planning team	2 engineers, 1 planning and 1 IT engineer
Hours used	6 meetings of 5 h × 10 people = 300 h	To extract the event log from the system, uploading to ProM and analysis: 20 h (process structure is established, future process mining analyses can be completed within 1-2 hours.)
Level of detail	Relation to people's capacity for abstraction	Logs written to the system
Ability to repeat the process	Difficult - difficulty in bringing together all involved	Easy - whenever needed
Reliability of information	Estimated times – People often have varying perspectives on the process, leading to disagreements about the duration of activities	Times collected from the system – The data align with the process
Period analyzed	It is a static photograph of the process	It is possible to replay the process

6.2 Challenges of Using Value Stream Mapping in Manufacturing Systems

- VSM is done manually, time consuming and requires high effort. It requires people that have knowledge about the process together to draw the map and comment. This situation causes to analyze on a static map in a dynamic environment (Horsthofer-Rauch et al., 2022).

- Low/lack of integration between processes: Difficulties or absence of integration among processes makes difficult to analyze the flow,

- Low/lack of clarity of processes: The production processes are not well defined. Materials and parts follow different paths within the production line,

- Low/lack of product modularity: Modular design is not achieved for the products, which complicates their manufacturing and assembly,

- Low-skilled people: The people with inadequate skills impede understanding and tool utilization,

- Poor/lack of process stability: The processes which have a lack of standardization and process stability,

- Problems/difficulties in measuring data in processes: Layout constraints, product complexity, or process type make time data and quantity measurements impractical,

- Obsolescence of the current state map: Processes have evolved without documentation,

- Small batches with highly mixed production: Compromised VSM application due to the assembly of diverse product types under the same infrastructure and an unregulated production schedule,

- Production too flexible: The production line is highly flexible, frequently adapting to market demand,

- Process too intuitive: The process flow heavily relies on operator’s real time decisions. (Forno et al., 2014).

6.3 Challenges of Using Process Mining in Manufacturing Systems

The presented work holds promise for widespread applicability in manufacturing environments characterized by machine executing systems and unknown or dynamic capacity constraints. Despite demonstrating the potential of the proposed procedure, its implementation is accompanied by several challenges.

- The first and the primary challenge, for companies lacking pre-installed machine data collection systems, manufacturing executing systems (MES), leading to potential obstacles in using process mining due to setup costs and the need for specialized software (Lorenz et al., 2021).

- The second challenge is associated with the incompleteness and disparate formats of manufacturing data. The accuracy of derived process models depends on the completeness of event logs, and incomplete observations can yield incorrect conclusions. Despite the growing volume of generated manufacturing data, accessibility in suitable formats remains a hurdle, necessitating considerable manual effort for preprocessing. Additionally, retrofitting older machines with limited sensory capabilities for automated data capture can be costly. While some manufacturers, have gradually embraced automation, challenges persist, such as encountering production lines lacking readers, limiting the scope of process mining analysis. The sandblasting has no sensor, no data and so there is no data in event log and we can’t show it in the process mining model.

- The third challenge arises when analyzing the flow of multiple products assembled together, necessitating the merging of case IDs. This complexity poses a current research area for advancing process mining algorithms, particularly challenging in manufacturing processes with variable entity levels (Lorenz et al., 2021).

- The fourth challenge; outsourced process cannot be shown in process mining model. The outsourced polishing operation cannot be shown in our process mining model. This can be further research.

- The fifth challenge, the processes that the parts are processed in batches cannot be shown in process mining model. The sandblasting and chrome plating processes cannot be shown in our process mining model. This can be further research.

- The sixth challenge involves the extensive coordination effort required between process experts, especially in the case of long value streams. Process understanding is crucial for suggesting improvement actions, requiring manual analyses and domain expertise in addressing deviations (Van Eck et al., 2015).

- The seventh challenge emphasizes maintaining data management discipline. While process mining aids in clarifying data requirements, effective communication and guidance for relevant but missing data on the shop floor is a must (Lorenz et al., 2021).

- The eighth challenge, the process may be changing while being analyzed. Understanding such concept drifts is of prime importance for the management of processes (W. Van der Aalst, 2012).

- The ninth challenge is Cross-Organizational Mining. There are various use cases where event logs from multiple organizations are available for analysis. In some cases, organizations collaborate to manage process instances (e.g., supply chain partners), or they may execute essentially the same process while sharing experiences, knowledge, or a common infrastructure. However, traditional process mining techniques typically focus on analyzing a single event log within one organization (W. Van der Aalst, 2012).

CHAPTER SEVEN

CONCLUSION

In the twenty-first century, manufacturing is characterized by the trend towards customized products, which presents a challenge to traditional mass production methods. The complexity of production planning and control systems has increased in response to this demand for tailored goods, making mass production more difficult. Industries like automotive, historically reliant on standardized assembly lines, are particularly feeling the strain of this adaptation. Companies in this segment are driven by two key forces: the ever-evolving preferences of a globalized customer base and the fierce competition that comes with it. These challenges prompt organizations to seek new tools and methods to remain competitive. While some have thrived due to economic stability, others have faced setbacks due to a lack of understanding of evolving customer needs and cost management practices. In response, many manufacturers have turned to lean manufacturing as a solution. The primary goal of lean manufacturing is to meet customer demand efficiently and responsively by minimizing waste. By implementing Lean principles, manufacturers aim to produce goods and services at the lowest cost and in the shortest time possible to meet customer expectations (Bhamu & Singh Sangwan, 2014).

VSM serves as a widely utilized lean management approach aimed at analyzing value streams and identification of optimization possibilities. A value stream map visually highlights essential process steps and key metrics, with a particular emphasis on throughput time. This visualization aids in comprehending the current state of processes and serves as a communication tool. Traditionally, data collection occurs manually on the shop floor, typically with pen and paper. However, this method is time-consuming and only provides a snapshot of reality (Horsthofer-Rauch et al., 2022).

The shift to Industry 4.0 has given businesses access to a wealth of data stored on their servers, which are essentially untapped but incredibly valuable resources. In this context, the importance of collecting data and analyzing the resulting big data is once again highlighted.

Since manual processing of data as in the VSM is time-consuming, it will ensure that less data, that is, the most valuable thing for businesses, is processed. For this reason, there are studies in the literature to automatically analyze and improve the processes examined and improved in VSM, or to examine and improve the processes in a more productive way by using different and automatic techniques such as process mining.

Process mining was introduced in 2004 by Wil van der Aalst(W. M. P. Van der Aalst & Weijters, 2004) aims to automatically create models that explain the behavior observed in event logs. It can also be seen as the process of extracting information about processes from event logs. Process mining has become a popular technique for business process management, especially after 2010. It serves as an important link between data mining, process modeling, and process analysis.

At the faucet factory The MES and ERP programs are considered fundamental components of Industry 4.0. These systems contribute to the creation of smarter and connected manufacturing environments by supporting digitalization, data integration, and the use of intelligent technologies in business processes. Our study aims to leverage the data stored on servers in a manner that minimizes effort while maximizing utility. Furthermore, in subsequent stages, the automation of these processes will enable the automatic generation of a process mining model and VSM for analysts after each production order to analyze through all processes.

Process mining is a tool that reveals the model from the process records (automatically reveals the current state map with algorithms), enables the process to be examined, enables performance analysis, enables bottlenecks to be revealed, allows us to monitor and replay the flow of the process over and over again, to see process variations and deviations, and to see the problem. Moreover, it does not require much effort to use it. However, it is not as flexible as VSM. Although we can add the data, we need to the VSM map, we are limited by the software and algorithms used in process mining.

The aim of this thesis is to improve the throughput time by examining the processes of a produced part in our sample faucet factory. For this purpose, firstly the current state VSM map is drawn for our process. Later, process mining is carried out for a part of the VSM map. Process mining study includes production data belonging to a production order of the part whose VSM map is drawn. This data was taken from the MES and ERP system and processed with Excel VBA and converted into event log data, which is the basic input of process mining.

The event log was uploaded to ProM software, that is an open-source process mining software that we used, and the process discovery step is carried out using various algorithms. Then, the quality of these discovered process models is examined (conformance checking) and then performance analysis is performed. VSM current state map is updated according to data obtained from the performance analysis. In addition, these data and the data obtained from the revised VSM map are analyzed using the line balance rate technique and the bottlenecks are searched. Then, with the results obtained here, the points requiring improvement in our process were determined and the future state VSM map was drawn according to the determined improvement targets.

As a result, it is predicted that Production Lead Time that we can say it Throughput Time will decrease from 48.8 days to 35.6 days, that is, there would be a 27% improvement, and a future state VSM Map is drawn. Hard-to-achieve targets have not been determined for the improvement work to be carried out, and the future situation map has not been drawn utopic.

In this thesis, two techniques, VSM and process mining, were integrated and applied to a real-world industrial case using real data. This integration resulted in improvements in throughput time. Finally, a comparison of the two methods is conducted, highlighting the challenges associated with their implementation.

In the future studies, It would be beneficial for process experts to develop software that can automatically use process mining and VSM together. Also, It is thought that it would be useful to conduct studies on how to use process mining for processes where parts are processed in batches (sandblasting, Chrome plating operations in our study), how to use process mining for processes where multiple produced parts are assembled (assembly process in our study), and how to use process mining for outsourced processes (the outsourced polishing process in our study).

REFERENCES

Bettacchi, A., Polzonetti, A., & Re, B. (2016). Understanding Production Chain Business Process Using Process Mining: A Case Study in the Manufacturing Scenario. In J. Krogstie, H. Mouratidis, & J. Su (Eds.), Advanced Information Systems Engineering Workshops (Vol. 249, pp. 193–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-39564-7_19

Bhamu, J., & Singh Sangwan, K. (2014). Lean manufacturing: Literature review and research issues. International Journal of Operations & Production Management, 34(7), 876–940. https://doi.org/10.1108/IJOPM-08-2012-0315

Busert, T., & Fay, A. (2021). Information quality focused value stream mapping for the coordination and control of production processes. International Journal of Production Research, 59(15), 4559–4578. https://doi.org/10.1080/00207543.2020.1766720

Chao, B., Zhang, C., Zhang, Y., Guo, H., Ren, Y., & Zhang, H. (2022). Research on optimization and simulation of sand casting production line based on VSM. International Journal of Lean Six Sigma, 13(6), 1185–1199. https://doi.org/10.1108/IJLSS-10-2021-0183

Choueiri, A. C., Sato, D. M. V., Scalabrin, E. E., & Santos, E. A. P. (2020). An extended model for remaining time prediction in manufacturing systems using process mining. Journal of Manufacturing Systems, 56, 188–201. https://doi.org/10.1016/j.jmsy.2020.06.003

Corallo, A., Lazoi, M., & Striani, F. (2020). Process mining and industrial applications: A systematic literature review. Knowledge and Process Management, 27(3), 225–233. https://doi.org/10.1002/kpm.1630

Duggan, K. J. (2013). Creating Mixed Model Value Streams (2nd Edition). C.R.C.Pres.

Er, M., Arsad, N., Astuti, H. M., Kusumawardani, R. P., & Utami, R. A. (2018). Analysis of production planning in a global manufacturing company with process mining. Journal of Enterprise Information Management, 31(2), 317–337. https://doi.org/10.1108/JEIM-01-2017-0003

Erlach, K. (2013). Value Stream Design: The Way Towards a Lean Factory. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12569-0

Farooqui, A., Bengtsson, K., Falkman, P., & Fabian, M. (2020). Towards data-driven approaches in manufacturing: An architecture to collect sequences of operations. International Journal of Production Research, 58(16), 4947–4963. https://doi.org/10.1080/00207543.2020.1735660

Ferreira, W. de P., Armellini, F., Santa-Eulalia, L. A. de, & Thomasset-Laperrière, V. (2022). Extending the lean value stream mapping to the context of Industry 4.0: An agent-based technology approach. Journal of Manufacturing Systems, 63, 1–14. https://doi.org/10.1016/j.jmsy.2022.02.002

Forno, A. J. D., Pereira, F. A., Forcellini, F. A., & Kipper, L. M. (2014). Value Stream Mapping: A study about the problems and challenges found in the literature from the past 15 years about application of Lean tools. The International Journal of Advanced Manufacturing Technology, 72(5–8), 779–790. https://doi.org/10.1007/s00170-014-5712-z

Friederich, J., Lugaresi, G., Lazarova-Molnar, S., & Matta, A. (2022). Process Mining for Dynamic Modeling of Smart Manufacturing Systems: Data Requirements. Procedia CIRP, 107, 546–551. https://doi.org/10.1016/j.procir.2022.05.023

Garcia, C. dos S., Meincheim, A., Faria Junior, E. R., Dallagassa, M. R., Sato, D. M. V., Carvalho, D. R., Santos, E. A. P., & Scalabrin, E. E. (2019). Process mining techniques and applications – A systematic mapping study. Expert Systems with Applications, 133, 260–295. https://doi.org/10.1016/j.eswa.2019.05.003

Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. In Business Process Management (Vol. 4714, pp. 328–343). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75183-0_24

Hadžiosmanović, D., Bolzoni, D., & Hartel, P. H. (2012). A log mining approach for process monitoring in SCADA. International Journal of Information Security, 11(4), 231–251. https://doi.org/10.1007/s10207-012-0163-8

Horsthofer-Rauch, J., Guesken, S. R., Weich, J., Rauch, A., Bittner, M., Schulz, J., & Zaeh, M. F. (2024). Sustainability-integrated value stream mapping with process mining. Production & Manufacturing Research, 12(1), 2334294. https://doi.org/10.1080/21693277.2024.2334294

Horsthofer-Rauch, J., Schumann, M., Milde, M., Vernim, S., & Reinhart, G. (2022). Digitalized value stream mapping: Review and outlook. Procedia CIRP, 112, 244–249. https://doi.org/10.1016/j.procir.2022.09.079

J.P.Womack, D.T.Jones, & D.Ross. (1990). The Machine That Change The World. Macmillian Publishing Company.

Knoll, D., Reinhart, G., & Prüglmeier, M. (2019). Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Systems with Applications, 124, 130–142. https://doi.org/10.1016/j.eswa.2019.01.026

Krafcik, J. F. (1988). Triumph of the Lean Production System. MIT SLOAN Management Review.

Kurniati, A. P., & Wisudiawan, G. A. A. (2022). Process Mining using Inductive Miner Algorithm to Determine the actual Business Process Model. JURIKOM (Jurnal Riset Komputer), 9(4), 1128. https://doi.org/10.30865/jurikom.v9i4.4769

La Rosa, M., & Soffer, P. (Eds.). (2013). Business Process Management Workshops: BPM 2012 International Workshops, Tallinn, Estonia, September 3, 2012. Revised Papers (Vol. 132). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36285-9

Leemans, S. J. J., & Fahland, D. (2014). Process and Deviation Exploration with Inductive visual Miner.

Leemans, S. J. J., Fahland, D., & van der Aalst, W. M. P. (2013). Discovering Block-Structured Process Models from Event Logs—A Constructive Approach. In J.-M. Colom & J. Desel (Eds.), Application and Theory of Petri Nets and Concurrency (Vol. 7927, pp. 311–329). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38697-8_17

Little, J. D. C. (1961). A Proof for the Queuing Formula: L = λ W. Operations Research, 9(3), 383–387. https://doi.org/10.1287/opre.9.3.383

Lorenz, R., Senoner, J., Sihn, W., & Netland, T. (2021). Using process mining to improve productivity in make-to-stock manufacturing. International Journal of Production Research, 59(16), 4869–4880. https://doi.org/10.1080/00207543.2021.1906460

Mannhardt, F., de Leoni, M., & Reijers, H. A. (2017). Heuristic Mining Revamped: An Interactive, Data-aware, and Conformance-aware Miner. 15th International Conference on Business Process Management.

Meincheim, A., Garcia, C. dos S., Nievola, J. C., & Scalabrin, E. E. (2017). Combining Process Mining with Trace Clustering: Manufacturing Shop Floor Process - An Applied Case. 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), 498–505. https://doi.org/10.1109/ICTAI.2017.00082

Nagy, Z., Werner-Stark, A., & Dulai, T. (2019). Using Process Mining in Real-Time to Reduce the Number of Faulty Products. In T. Welzer, J. Eder, V. Podgorelec, & A. Kamišalić Latifić (Eds.), Advances in Databases and Information Systems (Vol. 11695, pp. 89–104). Springer International Publishing. https://doi.org/10.1007/978-3-030-28730-6_6

Natschläger, C., Kossak, F., Lettner, C., Geist, V., Denkmayr, A., & Käferböck, B. (2017). A Practical Approach for Process Mining in Production Processes. In F. Piazolo, V. Geist, L. Brehm, & R. Schmidt (Eds.), Innovations in Enterprise Information Systems Management and Engineering (Vol. 285, pp. 87–95). Springer International Publishing. https://doi.org/10.1007/978-3-319-58801-8_8

Nawcki, M. L., Zanon, G. N., de Paula Santos, L. C., Santos, E. A. P., Szejka, A. L., & de Lima, E. P. (2021). Process Mining and Value Stream Mapping: An Incremental Approach. In D. A. Rossit, F. Tohmé, & G. Mejía Delgadillo (Eds.), Production Research (Vol. 1408, pp. 171–183). Springer International Publishing. https://doi.org/10.1007/978-3-030-76310-7_14

Roldán, J. J., Crespo, E., Martín-Barrio, A., Peña-Tapia, E., & Barrientos, A. (2019). A training system for Industry 4.0 operators in complex assemblies based on virtual reality and process mining. Robotics and Computer-Integrated Manufacturing, 59, 305–316. https://doi.org/10.1016/j.rcim.2019.05.004

Rother, M., & Shook, J. (1998). Görmeyi öğrenmek.

Rozinat, A., de Jong, I. S. M., Gunther, C. W., & van der Aalst, W. M. P. (2009). Process Mining Applied to the Test Process of Wafer Scanners in ASML. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 39(4), 474–479. https://doi.org/10.1109/TSMCC.2009.2014169

Rudnitckaia, J., Venkatachalam, H. S., Essmann, R., Hruska, T., & Colombo, A. W. (2022). Screening Process Mining and Value Stream Techniques on Industrial Manufacturing Processes: Process Modelling and Bottleneck Analysis. IEEE Access, 10, 24203–24214. https://doi.org/10.1109/ACCESS.2022.3152211

Sá Ribeiro, D. R., Mendes, L. G., Forcellini, F. A., Frazzon, E. M., & Neumann, C. (2023). VSM Approach to Manufacturing Systems in Industry 4.0: From Literature Review to a Conceptual Framework. In F. Deschamps, E. Pinheiro de Lima, S. E. Gouvêa da Costa, & M. G. Trentin (Eds.), Proceedings of the 11th International Conference on Production Research – Americas (pp. 538–547). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36121-0_68

Saka, G., & Bilgen, B. (2024). Process Mining and Value Stream Mapping with Lean Manufacturing Techniques: A Case Study at a Faucet Factory (Casting and Machining). 13th International Congress on Engineering, Architecture And Design, Istanbul, TURKEY, page 657. http://www.muhendislikmimarliktasarimkongresi.org/eng/

Schmenner, R. W. (2015). The Pursuit of Productivity. Production and Operations Management, 24(2), 341–350. https://doi.org/10.1111/poms.12230

Schuh, G., Gützlaff, A., Schmitz, S., & van der Aalst, W. M. P. (2020). Data-based description of process performance in end-to-end order processing. CIRP Annals, 69(1), 381–384. https://doi.org/10.1016/j.cirp.2020.03.013

Sullivan, B. P., Yazdi, P. G., Suresh, A., & Thiede, S. (2022). Digital Value Stream Mapping: Application of UWB Real Time Location Systems. Procedia CIRP, 107, 1186–1191. https://doi.org/10.1016/j.procir.2022.05.129

Sunk, A., Kuhlang, P., Edtmayr, T., & Sihn, W. (2017). Developments of traditional value stream mapping to enhance personal and organisational system and methods competencies. International Journal of Production Research, 55(13), 3732–3746. https://doi.org/10.1080/00207543.2016.1272764

Teriete, T., Böhm, M., Sai, B. K., Erlach, K., & Bauernhansl, T. (2022). Event-based Framework for Digitalization of Value Stream Mapping. Procedia CIRP, 107, 481–486. https://doi.org/10.1016/j.procir.2022.05.012

Timeline—The W. Edwards Deming Institute. (2023, December 20). https://deming.org/timeline/

Van der Aalst, W. (2012). Process Mining: Overview and Opportunities. ACM Transactions on Management Information Systems, 3(2), 1–17. https://doi.org/10.1145/2229156.2229157

Van der Aalst, W. (2016). Process Mining. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49851-4

Van der Aalst, W. M. P. (2011). Process mining: Discovery, conformance and enhancement of business processes. Springer. https://doi.org/10.1007/978-3-642-19345-3

Van der Aalst, W. M. P., & Weijters, A. J. M. M. (2004). Process mining: A research agenda. Computers in Industry, 53(3), 231–244. https://doi.org/10.1016/j.compind.2003.10.001

Van Eck, M. L., Lu, X., Leemans, S. J. J., & Van der Aalst, W. M. P. (2015). A Process Mining Project Methodology. In J. Zdravkovic, M. Kirikova, & P. Johannesson (Eds.), Advanced Information Systems Engineering (Vol. 9097, pp. 297–313). Springer International Publishing. https://doi.org/10.1007/978-3-319-19069-3_19

What is Lean? | Lean Thinking. (2023, December 20). Lean Enterprise Institute. https://www.lean.org/explore-lean/what-is-lean/

Xiao, F., & Shao, L. (2018). Optimizing Production Line Balance Based on Witness Simulation. 2018 8th International Conference on Logistics, Informatics and Service Sciences (LISS), 1–5. https://doi.org/10.1109/LISS.2018.8593241

Yasmin, F. A., Bemthuis, R., Elhagaly, M., & Bukhsh, F. A. (2019). A Process Mining Starting Guideline for Process Analysts and Process Owners: A Practical Process Analytics Guide using ProM.

Yılmaz, Y. (2019). Business Process Reengineering using Process Mining. Boğaziçi University, Institute of Social Sciences.