Universidade do Minho
Escola de Engenharia

Vishad Viralbhai Vyas

Stochastic approach to time-driven activity-based costing in an automotive company

february 2024

Stochastic approach to time-driven activity-based costing in an automotive company
Vishad Viralbhai Vyas
UMinho | 2024

Universidade do Minho
Escola de Engenharia

Vishad Viralbhai Vyas

Stochastic approach to time-driven activity-based costing in an automotive company

Doctoral Thesis

Doctoral Program in Industrial and Systems Engineering (DPISE)

Work performed under the supervision of

Professor Paulo Sérgio Lima Pereira Afonso

Professor Lino António Antunes Fernandes da Costa

February, 2024

Acknowledgements

I am immensely grateful to everyone who has contributed to the completion of this PhD thesis. Their support, encouragement, and guidance have been invaluable, and I could not have achieved this milestone without them.

Firstly, I would like to express my deepest gratitude to my supervisor, Professor Paulo Afonso, for his unwavering support and guidance throughout my PhD journey. His extensive knowledge, critical insights, and constructive feedback have significantly improved the quality of my work. I am particularly thankful for his patience, flexibility, and kindness during challenging times.

I would also like to thank my co-supervisor, Professor Lino Costa, for his valuable input and advice, which have greatly enriched my research. His expertise in Statistics and Stochastic models has been instrumental in shaping my research. I am grateful for his commitment to my academic development and his willingness to go above and beyond to help me.

I would like to acknowledge the administrative staff of the Department of Production systems, for their assistance and support during my candidature. They have facilitated my progress by providing me with administrative guidance, organizing events and workshops, and responding to my inquiries promptly.

I would also like to thank my family and friends for their unwavering love, encouragement, and motivation. Their emotional support and belief in me have kept me going through the ups and downs of this journey. I appreciate their sacrifices, understanding, and patience during my long hours of work.

Finally, I would like to acknowledge the company of my study, without whom my research would not have been possible. Their willingness to share their data, experiences and insights has enriched my understanding of their manufacturing processes and cost models. I am grateful for their time, effort, and trust in me.

In conclusion, I express my heartfelt gratitude to everyone who has supported and guided me throughout my PhD journey. Their contributions have been significant, and I am deeply grateful for their presence in my life. Thank you all.

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, 28th February 2023

Full name: Vishad Viralbhai Vyas

Signature:

Vishad

Stochastic approach to time-driven activity-based costing in an automotive company

Abstract

Traditional costing methods based on volume measures are not efficient in modern times characterized by product diversity, production complexity and market volatility and uncertainty. In such situations, traditional costing systems based on deterministic cost models are not effective. Indeed, decision making must be based on sophisticated, timely and accurate costing systems. Activity-based costing (ABC) models were introduced in the 1980s to help to deal with such complexity. Furthermore, Time-driven ABC (TDABC) was proposed in the 2000s to overcome some of the limitations of the previous cost models based on activities. These models are deterministic but due to the existence of variability and uncertainty, product costing should be approached stochastically. This research focuses on developing a stochastic approach to costing systems that considers the variability in the process cycle time of the different activities that compose business and production processes, e.g., workstations in the assembly lines. This approach provides a range of values for the product costs, allowing for a better perception of the risk associated with costs instead of providing a single value of the cost. The confidence interval for the mean and the use of quartiles one and three as lower and upper estimates are proposed to include variability and risk in costing systems. Furthermore, a stochastic activity-based cost model has been developed to take into consideration variability in manufacturing processes. The stochastic cost model is supported on a set of time equations structured around a cost hierarchy that highlights and differentiates the costs in the production line, namely, workstation specific cost, cost induced by the line bottleneck, planned unused capacity cost and unplanned unused capacity cost. The model supports and can be extended for cost prediction purposes and for cost optimization under conditions of variability in resources, processes and operations conditions, and demand changes imposed by the market. This research also presents and discusses the use of prediction models, based on simple and sophisticated methods such as linear regression for forecasting production quantities, to predict deviant behaviour of each specific production line. The model developed enabled us to draw attention to the importance of production control and planning. With greater control on production quantities, consequently, costs will also be more controlled and less subject to large variations. Finally, optimization models can be used to extend the application of the proposed stochastic cost model. In this case, for instance, such models can help to identify the number of operators required, their allocation to the workstation along with the cost associated with them and also the optimum number of equipment required at each workstation based on the target cycle time and cost of the product. The models and applications presented in this thesis can be further extended by considering more variables that might be relevant in other type of products and industries.

Keywords: Costing systems, Variability, Stochastic models, Time driven activity-based costing, Prediction of costs, Optimization models

Abordagem estocástica para custeio baseado em atividades e tempo numa empresa do setor automóvel

Resumo

Os métodos tradicionais de custeio baseados em volume não são eficientes nos tempos modernos caracterizados pela diversidade de produtos, complexidade da produção e volatilidade e incerteza do mercado. Neste contexto, os sistemas de custeio tradicionais baseados em modelos de custos determinísticos não são eficazes. De facto, a tomada de decisão deve ser baseada em sistemas de custeio sofisticados, que providenciam informação oportuna e precisa. Os modelos de custeio baseado em atividades (ABC) foram introduzidos na década de 1980 para ajudar a lidar com essa complexidade. Além disso, o modelo de custeio baseado em atividades e tempo (TDABC) foi proposto na década de 2000 para superar algumas das limitações dos modelos de custos anteriores baseados em atividades. Esses modelos são determinísticos, mas devido à existência de variabilidade e incerteza, o custeio do produto deve ser abordado estocasticamente. Esta pesquisa foca-se no desenvolvimento de uma abordagem estocástica para sistemas de custeio que considere a variabilidade no tempo de ciclo do processo das diferentes atividades que compõem os processos de negócio e de produção, por exemplo, postos de trabalho nas linhas de montagem. Esta abordagem fornece uma gama de valores para os custos do produto, permitindo uma melhor perceção do risco associado aos custos em vez de fornecer um único valor de custo. O intervalo de confiança para a média e a utilização do primeiro e terceiro quartis como estimativas inferior e superior são propostos para incluir a variabilidade e o risco nos sistemas de custeio. Além disso, um modelo de custos estocástico baseado em atividades foi desenvolvido para levar em consideração a variabilidade nos processos de fabricação. O modelo de custo estocástico é suportado por um conjunto de equações de tempo estruturadas em torno de uma hierarquia de custos que destaca e diferencia os custos na linha de produção, ou seja, custo específico de cada posto de trabalho, o custo induzido pelo gargalo da linha, o custo da capacidade não utilizada planeada e o custo da capacidade não planeada não utilizada . O modelo suporta e pode ser estendido para fins de previsão de custos e otimização de custos sob condições de variabilidade de recursos, processos e condições de operação e mudanças na procura impostas pelo mercado. Esta pesquisa também apresenta e discute o uso de modelos de previsão baseados em métodos simples e sofisticados, tais como regressão linear para previsão de quantidades de produção, para prever o comportamento desviante de cada linha de produção específica. O modelo desenvolvido permitiu evidenciar a importância do controlo e planeamento da produção. Com maior controlo nas quantidades de produção, consequentemente, os custos também poderão ser mais controlados e estar menos sujeitos a grandes variações. Finalmente, os modelos de otimização podem ser usados para estender a aplicação do modelo de custo estocástico proposto. Neste caso, por exemplo, tais modelos podem ajudar a identificar o número de operadores necessários, a sua alocação ao posto de trabalho juntamente com o custo a eles associado e também o número ótimo de equipamentos necessários em cada posto de trabalho com base no tempo de ciclo e o custo alvo do produto. Os modelos e aplicações apresentados nesta tese podem ser estendidos considerando mais variáveis que podem ser relevantes em outros tipos de produtos e indústrias.

Palavras Chave: Sistemas de custeio, Variabilidade, Modelos estocásticos, Custeio baseado em atividades e tempo, Previsão de custos, Modelos de otimização

Table of contents

List of tables

List of figures

List of abbreviations and acronyms

ABC - Activity-Based Costing

AFIA - Portuguese Automotive Suppliers Association

AI - Artificial intelligence

AR - Autoregressive models

ERP - Enterprise Resource Planning

LSTM - Long Short-Term Memory

MAE – Mean Absolute Error

MA – Moving Average

MAPE- Mean Absolute Percentage Error

OEE – Overall Equipment Effectiveness

RNNs - Recurrent Neural Networks

TDABC – Time driven activity-based costing

UEP – Production Effort Unit

UVA - Unité de Valeur Ajoutée

VAR – Vector Autoregression

Chapter 1

1 Introduction and thesis overview

This chapter focuses on the following aspects: the context of the research project, fundamentals, concepts about product costing, opportunity and motivation of the research, objectives and research questions, research methodology, general characterization of the automotive components industry. Finally, the thesis structure is presented.

1.1 Context

Currently, the world is strongly driven by technology and innovation and in such context the importance of product costing is growing. Companies are forced to produce low-cost and high-quality products in order to maintain their competitiveness. The business world is characterized by intense global competition and therefore effective cost management is a crucial strategic advantage for companies. The high competitiveness among companies means that those with the best strategies and levels of internal and external integration stand out positively above the others. The best strategies allow us to create more competitive advantages for the company. Good internal integration allows companies to achieve excellent operational levels at low cost and external integration allows companies to approach the business beyond their internal operations. In this context, cost management, which includes the accounting, control, management, and cost reduction is seen, increasingly, as a condition of survival in today's market, where companies define strategies in order to achieve the highest possible quality and functionality at the lowest total cost.

Costs are a relevant criterion that influences decision making and their estimation plays an important role in management. The key to a rising enterprise in the 21st century is product quality, competitive cost, fast delivery, and flexibility. The quality-functionality-price paradigm is a fundamental element of modern cost management that influences design changes that can be made collaboratively involving suppliers or clients towards the reduction of the costs of new products.

A good costing system helps managers to understand the detailed cost of different short and long-term activities and processes which involves the cost of services provided to customers, raw material costs to supplier and other costs, but it needs appropriate and consistent information to be successful (Alessandroni et al., 2002). Manufacturing product costs are important in costing systems, and they can be explained through the typical three components: direct materials, direct labor, and indirect costs (such as energy, amortization, indirect labor). Direct labor plus indirect costs represent the conversion costs. Costs that can vary particularly with the production are called variable costs and that cannot vary are called fixed costs. The unitary product cost can also include non-manufacturing costs (e.g., logistics costs, sales, and administrative costs, etc.), typically allocated on a volume basis, for instance, the amount of sales. A complete cost can be compared to the price in order to evaluate the profitability of the product. But a first analysis of the margins must be based on the manufacturing cost, from which several actions can be taken on the shop floor, e.g., optimization of processes, changes in raw materials, waste reduction plans, changes in the product mix, among others.

In costing, the main focus of researchers is to solve problems related to the allocation of costs to products. Deterministic cost models play a pivotal role in understanding the product costing and performing profitability analysis and pricing strategies. On the other hand, cost estimation can also be used to get important information about budgeting and to support the design of business plans. Hence, to reduce the cost of the product or to optimize the cost, the use of cost estimation models can be implemented (Goh et al., 2010).

Traditional costing systems have been used for many years to assign costs to products or services. However, they have several limitations, including: (1) Overhead allocation: as traditional costing systems allocate overhead costs based on a single predetermined rate. This method assumes that all products or services consume overhead costs in the same proportion, which is not always true. This can result in inaccurate cost estimates for individual products or services. (2) Complexities in tracing costs: traditional costing systems rely on the allocation of indirect costs, which can make it difficult to trace costs to specific products or services. This is particularly true in organizations with multiple product lines or service offerings. (3) Limited usefulness for decision-making: traditional costing systems provide historical cost information and do not take into account changes in the market, technology, or other external factors. This makes it difficult to use the information generated by traditional costing systems to make informed decisions about pricing, product mix, or resource allocation. Overall, traditional costing systems are limited in their ability to provide accurate and relevant cost information for decision-making in today's complex business environment (Banker et al., 2008). As a result, many organizations are turning to alternative costing methods such as activity-based costing (ABC) and lean accounting to better capture the true cost of their products and services.

Deterministic costing systems use a predetermined standard to estimate the cost of a product. This system has several limitations that can affect its accuracy and usefulness such as: (1) Ignoring variability: deterministic costing systems assume that inputs and outputs are consistent and do not vary. However, in reality, there may exist variability in the quality and quantity of inputs, and the outputs may vary due to changes in the production process or customer demand. (2) Limited flexibility: deterministic costing systems are inflexible and cannot adapt to changes in the production process, materials, or labor costs. This can lead to inaccurate cost estimates and make it difficult for managers to make informed decisions. (3) Cannot handle complex processes: deterministic costing systems are suitable for simple processes with straightforward inputs and outputs. However, for complex processes that involve multiple inputs and outputs, deterministic costing systems may not provide accurate cost estimates. (4) Ignores non-value-added activities: deterministic costing systems do not differentiate between value-added and non-value-added activities. This means that activities that do not contribute to the production process, such as inspections and quality checks, are included in the cost estimates (Drury, 2013).

Nevertheless, variability is a crucial aspect in costing models as it can have a significant impact on the accuracy of cost estimates and the effectiveness of decision-making processes. Costing models that fail to consider variability may result in over or under-estimation of costs, leading to inefficient resource allocation, poor profitability, and ultimately, business failure. It is important to understand the relevance of variability in the cost models because : (1) To account for fluctuations in demand: costing models must be able to account for variability in demand, as this directly affects the quantity and cost of resources required to fulfill customer orders. This can be achieved by implementing a cost driver-based approach, where costs are assigned to specific activities or processes that drive demand. (2) To manage uncertainty: variability is inherent in business, and as such, costing models must be able to manage uncertainty. By analyzing past trends and considering potential risks, costing models can provide more accurate estimates of costs and assist in making informed decisions. (3) To identify areas of improvement: costing models can help identify areas of inefficiency and opportunities for improvement by analyzing the causes of variability in costs. By understanding the root causes of variability, organizations can implement measures to reduce or eliminate unnecessary costs, increase profitability, and enhance customer satisfaction. (4) To support strategic decision-making: costing models can provide critical information to support strategic decision-making, such as whether to invest in new product lines, expand operations, or enter new markets. By providing accurate cost estimates that reflect the variability in demand and other factors, costing models can help organizations make informed decisions that align with their business goals (Cooper & Kaplan, 1988).

Nowadays, manufacturing analytics are important to derive insights about the impacts on the organization of internal and external changes and variability (Meister et al., 2019). Thus, statistical methods are an important tool since they can be used to deal with the variability in observed data. Moreover, data can be organized and summarized to understand the information available. Descriptive statistics are widely used to identify the important features of the data, for example, the mean, standard deviation, quartiles, minimum and maximum values, range, and coefficient of variation.

In this context, descriptive statistics such as the minimum, maximum, mean, and standard deviation can be used to analyze variability and risk on product cost data (Okoye et al., 2013). Another study was made to analyze the impact of strategic costing techniques, where the descriptive statistics mean, and standard deviation were also applied to verify if the new strategy achieved a successful performance when comparing with prior years (Alsoboa et al., 2015). Furthermore, the first quartile was also used for measuring machinery usage (Sobczyk & Koch, 2008). The coefficient of variation is another metric used in previous studies, for example, operations cycle times were used to analyze the optimal allocation of storage space in production lines (Hillier & So, 1991). According to Montgomery & Runger, (2018) estimations using the mean could be close or far from the true mean, and in order to avoid this, it can be used, instead, a range of the potential values, like a confidence interval.

Thus, a stochastic analysis is fundamental to include variability and risk within costing systems. Besides the variability in the production processes, measured, for instance, considering the production cycle times, we can also have variability caused by changes in the demand and variability in the value of the resources used. Process variability is particularly relevant in costing systems and for optimization purposes and it is the focus of this research work. This analysis is important for decision making in general and engineering and manufacturing, in particular (Montgomery & Runger, 2018). For example, the detection of deviant behavior (large or small variations) in costs can be detected with measures such as standard deviation and coefficient of variation.

1.2 Research gap and motivation

Product costing is a crucial aspect of financial and operations management for any manufacturing organization. It involves identifying and calculating the total cost incurred in producing a particular product or service. The cost of production plays a critical role in determining the price of the product or service, which directly impacts the profitability of the company. Therefore, accurate product costing is essential for any organization to remain competitive in the market (Pember & Lemon, 2015).

Traditionally, companies have used a top-down approach to determine the cost of products. This approach involves allocating the total cost of production to the products based on predetermined factors such as the number of units produced, or the time spent in production. While this approach is relatively simple to implement, it does not accurately reflect the cost of each individual manufacturing process involved in production. As a result, it is difficult to identify areas where cost savings can be made or to predict the impact of changes in the production process (Cooper & Kaplan, 1988).

Moreover, the traditional costing approach is not effective in accommodating the variability that persists in most production processes. For instance, customer requirements may change over time, resulting in fluctuations in the quantity of products to be produced. Similarly, the cycle time of each operation may vary due to factors such as equipment breakdowns, worker absenteeism, or quality issues. These variations have a significant impact on the cost of production and can result in a lot of variation in the profitability of the company (Cooper & Kaplan, 1988).

To address these challenges, there is a need for stochastic cost models. In this research, a model based on time-driven activity-based costing (TDABC) has been developed. TDABC is a bottom-up approach that involves identifying and measuring the time required to perform each activity in the manufacturing process. The model then assigns a cost to each activity based on the cost of the resources used and the time required to perform the activity. By applying the TDABC, the calculation of the cost of production starts from the basic operations of the manufacturing process, enabling a more accurate and detailed cost analysis (Kaplan & Anderson, 2007).

A stochastic cost model can accommodate the variability that persists in most production processes. For example, it can incorporate the differences between the planned quantities and the actual quantities produced. By considering the production cycle time, bottleneck, and other factors, the model can generate a range of probabilistic values for the cost of production. This range can be used to calculate optimistic and pessimistic costs, which can aid in managerial decision-making.

Furthermore, a prescriptive model is also developed to forecast production quantities, which helps in predicting the cost of production accurately. By forecasting the production quantities, such models can predict the resources required to manufacture the product, enabling the optimization of resources used in production. In this research, an optimization model has been used to determine the cost of labor and equipment, and the optimal number of operators and workstation equipment required. This model allows the allocation of the operators to workstation but also helps to understand the value-added and non-value-added activities and cost.

In addition to aiding in managerial decision-making, a stochastic TDABC model can be used to evaluate the performance of the different manufacturing processes. By comparing the actual cost of production with the estimated cost of production, the model can identify areas where the production process can be improved. For example, if the actual cost of production is higher than the estimated cost of production, it may indicate that there are inefficiencies in the manufacturing process that need to be addressed.

Stochastic cost models can also aid in identifying the profitability of specific product lines. By analyzing the cost of production for each product line, such models can identify which assembly lines are the most profitable and which assembly lines are not generating enough revenue. This information can be used to make strategic decisions about which product lines to focus on and which product lines to discontinue.

Since, the traditional top-down approach to product costing is insufficient in accommodating the variability that persists in most production processes, it gives the motivation to develop stochastic cost models based on TDABC which can provide a more accurate and detailed cost analysis, which aids in managerial decision-making and performance evaluation. By incorporating the variability of the production process, these models can generate a range of probabilistic values for the cost of the product. This range can be used to determine optimistic and pessimistic cost estimates, allowing companies to better plan and manage their resources. For example, if the pessimistic cost estimate is significantly higher than the optimistic estimate, companies can allocate more resources to reduce the likelihood of the pessimistic estimate being realized.

1.3 Research questions and objectives

Product cost management and control is an emerging field of research with a multidisciplinary nature and with growing importance in most industries and companies, particularly complex, and globalized supply chains, such as automotive ones. Thus, calculating the cost correctly is considered to be extremely important from both academic and industry perspectives in order to identify, develop and promote the application of best practices, tools, and management approaches.

The objective of this research was to develop novel cost models for financial and operational management in manufacturing organizations with emphasis of inclusion of variability in the product costing. The traditional top-down approach to product costing has limitations in accurately reflecting the cost of each individual manufacturing process involved in production. Thus, this research introduces the concept of stochastic cost model using time-driven activity-based costing (TDABC) as a bottom-up approach that provides a more accurate and detailed cost analysis by identifying and measuring the time required to perform each activity in the manufacturing process. The stochastic cost model can accommodate the variability that persists in most production processes, and it can be used to evaluate the performance of the manufacturing process and identify the profitability of specific product assembly lines. Additionally, this research proposes the use of prescriptive model to forecast production quantities, which helps in predicting changes in the cost of products more accurately, enabling the optimization of resources used in production and an increased profitability. Overall, the objective of this thesis is to propose a novel cost model and to provide insights into how the proposed stochastic TDABC model can aid in managerial decision-making, performance evaluation, and operational and strategic planning in manufacturing organizations.

The specific objectives of this research were the development of:

1. A methodology to calculate the cost of products by performing the descriptive statistical analysis of the production cycle time: This methodology involves analyzing the production cycle time data to understand the distribution of the cycle time and to identify any trends or patterns in the data. Descriptive statistics such as mean, median, mode, standard deviation, range, and interquartile range are calculated to summarize the data. Once the data is analyzed, the cost of the product can be calculated by assigning a cost per unit of time and multiplying it by the cycle time for each product.

2. A stochastic cost model which is an extended version of time-driven activity-based costing model to include the variability in the cost of the product: This stochastic cost model considers the variability in the cost of the product due to uncertainties in the production process. It uses probability distributions to model the uncertain variables, such as the cycle time, the bottleneck, and the quantities produced.

3. A proposal for the creation of an additive regression model, for a set of real data relating to industrial production. After collecting information regarding the production lines, over different months, and different weeks, they were analysed, extracting the correlations between the respective variables. As such, an additive regression model is proposed to predict the actual quantities as a function of all other variables. This model is based on the M5 Tree classifier. The goal is to study which variables can be considered for a forecasting the production quantities considering the measures R^2 (correlation coefficient), MAE (Mean absolute error), RMSE (Root mean squared error) and MAPE (Mean absolute percentage error).

4. Cost optimization using the concepts of optimization and the proposed stochastic models: The cost optimization models use optimization techniques to minimize the cost of the product subject to various constraints, such as target bottleneck, number of operators, and resource availability. The stochastic models can be incorporated into the optimization models to account for the variability in the cost factors. The goal is to calculate the value-added and non-value-added cost and to find the optimal allocation of operators and the equipment needed in the assembly that minimizes the cost of the product while satisfying the constraints.

The research question that has been addressed in each chapter starting from Chapter 2 to Chapter 5 can be found in the table 1.1.

Table 1.1 Research questions

#	**Chapters**	**Research questions**
2	A stochastic approach for product costing in manufacturing processes	How variability in cost can be calculated using the descriptive statistical analysis?
3	A stochastic costing model for manufacturing environment	How to calculate the cost of the product using TDABC considering the variability that prevails during manufacturing?
4	An Additive Regression Approach for Industrial Production Forecasting: A Comprehensive Analysis of Relevant Variables	How to predict production quantities using additive regression model?
5	Optimization of labour and workstation costs	How can labor and workstation costs be calculated using a stochastic cost model?

The methodology developed to calculate cost by descriptive analysis helps to clearly establish the existence of variability in the production processes and its impact on the cost of the product. A better understanding of product variability is necessary for better control over the cost. A proper risk assessment asks for a good understanding of the impact of uncertainty and variability on costs (Afonso & Jiménez, 2015). Since, TDABC has limitations accommodating the variability and to properly address the complexity of process (Vedernikova et al., 2020), there was a need to extend the TDABC model. Hence, a Stochastic TDABC model is developed considering variability in the cycle time, bottleneck time, and production quantities. Prescription aids in the understanding of variability and the risks they provide, particularly when forecasting is challenging due to the unpredictability and volatility of the business environment. The research intends to contribute to forecast model design, specifically to forecast the production quantities and associated costs. Based on the prediction of the quantity, it is essential to understand the optimum need and allocation of operators and equipment on the workstation along with controlling the cost of the product and identifying the value-added activities and non-value-added activities.

1.4 Research methodology

This thesis followed a multi-methodological approach, where three different types of research methods were followed: 1) quantitative research which involved collecting numerical data and using statistical analysis to draw conclusions; 2) action research which involved conducting research in collaboration with stakeholders in order to solve practical problems or improve a process or system; 3) experimental research which involved manipulating variables in a controlled setting to test hypotheses and establish cause-and-effect relationships.

Chapter 2: “A Stochastic Approach for Product Costing in Manufacturing Processes” follows the quantitative research methodology. As the statistical analysis was performed in this chapter. Mean, Standard deviation, range, and Confidence intervals were calculated along with other statistical tests including Mann-Witney tests. With the help of the methodology followed it enabled to identify the existence of variability in the manufacturing process and deterministic model was created from it. Four assembly lines were considered. A sufficiently large sample of 38000 observations were taken into the consideration from the period of 18th December 2020 to 24th December 2020. The data collected was focused on the quantities produced in each assembly line and cycle time on each workstation for each unit were gathered. The bottleneck of the assembly lines was identified, and tests were made on that workstation. The cycle times of each workstation were extracted from company´s ERP system to SPSS software where the various tests and analysis were made.

An action research methodology has been implemented in Chapter 3: “A Stochastic Costing Model for Manufacturing Environment”. A novel stochastic cost model has been developed in this chapter.

The company where this research was carried out had a Manufacturing Execution system (MES) and with the help of enhanced Internet of things (IoT) features it was possible to record every activity that was happening on each workstation. The activities were recorded as the event along with the start time and end time of each event. Hence, giving a clear idea about the number of activities taking place in the workstation and time taken by each activity was identified with the help of stakeholders from the company where the research was performed. After identifying the activity, the time taken for that activity and cost calculation was done. Overall, the improved method to calculate the cost in stochastic way was developed which solved their problem as they were using Traditional top-down approach. The analysis of product costs was made over the course of 5 months. A single product was studied over different assembly lines, as well as different products were studied over one specific line.

In Chapter 4: "An Additive Regression Approach for Industrial Production Forecasting: A Comprehensive Analysis of Relevant Variables", taking a step further in the development of an additive regression model using real data on industrial production. After gathering data on the manufacturing lines across a variety of weeks and months, it was analyzed to determine the correlations between the various factors. In order to forecast the actual numbers as a function of all other factors, an additive regression model is therefore developed. The M5 Tree classifier serves as the model's foundation. In light of the measures R2 (correlation coefficient), MAE (Mean absolute error), RMSE (Root mean squared error), and MAPE (Mean absolute percentage error), the results demonstrate that the models with the variables "Bottleneck value," "Amort Specific cost," "Specific other cost," and/or "Bottleneck Workstation" have a good forecasting quality.

In the last stage of the research, presented in Chapter 5 "Optimization of labour and workstation costs", an action research methodology has been followed. In the complex manufacturing system, there are a lot of production processes involved with many activities. It is important to understand, and bifurcate which activities are adding value to the customer, and which are non-value-added activities. To achieve this goal the equations have been developed. Also, it is important to control the cost of the labor and cost of equipment on the workstation. Based on the target cycle time the optimum allocation of operators is calculated. If the bottleneck is too big, it can be further split by adding another equipment to the workstation and the cost can be optimized.

1.5 The automotive industry

The field work of this research was performed in an automotive company which is tier-1 supplier of automotive components with several thousands of employees in several factories in Portugal. Since this research is performed in the automotive industry it is relatively important to understand the importance of automotive companies in the context of the manufacturing industry.

Automotive companies have a significant impact on the manufacturing industry, as they are responsible for producing a vast array of vehicles that are used in transportation across the world. The importance of automotive companies can be seen in their contribution to the global economy, their impact on technological advancements, and their role in promoting sustainable development.

Firstly, automotive companies are major players in the global economy. According to a report by the International Organization of Motor Vehicle Manufacturers (OICA), the automotive industry contributes to around 5% of the world's GDP and employs over 60 million people globally. In addition, automotive companies are responsible for a significant proportion of the world's trade and investment, with many countries relying on the industry for their economic growth.

Secondly, automotive companies are driving technological advancements in the manufacturing industry. The development of electric vehicles, for example, has revolutionized the way we think about transportation, and many automotive companies are investing heavily in research and development to improve the efficiency and sustainability of their vehicles. This has led to the emergence of new technologies such as autonomous vehicles, which have the potential to transform the way we travel in the future.

Finally, automotive companies play an important role in promoting sustainable development. The industry is under increasing pressure to reduce its carbon footprint and adopt more sustainable manufacturing practices. Many companies are implementing initiatives to reduce their environmental impact, such as using recycled materials and developing more efficient production processes. In addition, the development of electric and hybrid vehicles is helping to reduce greenhouse gas emissions and promote a shift towards cleaner forms of transportation.

In conclusion, the importance of automotive companies in the manufacturing industry cannot be overstated. Their contribution to the global economy, their role in driving technological advancements, and their commitment to sustainable development make them essential players in the manufacturing sector. As we look to the future, it is clear that the automotive companies will continue to play a significant role in shaping the way we live and work.

Considering the complexity of processes and products in this industry, models developed here can be replicated and applied in other industries using similar models or simplified ones. The maturity level of information and communication systems in this industry also allows a better application of complex cost models which required a considerable amount of information on the process and products. The models developed here can also be used in smaller companies, probably, through adapted and simplified versions.

1.6 Manufacturing industry and process overview

The manufacturing industry involves the production of goods and products using various complex processes, sophisticated technologies, and machines. This sector covers a wide range of industries, including automotive, aerospace, electronics, food and beverage, pharmaceuticals, and more. The industrial sector has undergone considerable changes recently as a result of technological improvements, globalization, and heightened rivalry. Manufacturers, specifically in automotive industry are putting more effort into implementing continuous improvement processes, which refers to the constant endeavor to discover and remove waste, decrease costs, increase efficiency, and enhance quality, in order to stay competitive and relevant. With this strategy, processes are continuously assessed and improved in order to maximize performance and produce superior results. After the implementation of these improvement tactics, still exist the gap to understand its effect on the cost of the product. Therefore, a proper costing model is essential particularly in the challenging automotive industry.

The automotive electronic component manufacturing industry involves the production of a wide range of components used in vehicles such as sensors, controllers, and displays. The manufacturing processes for these components often take place on an assembly line, which is a series of interconnected workstations where components are assembled into finished products.

One common manufacturing process used in the automotive electronic component industry is surface mount technology (SMT), where electronic components are mounted onto printed circuit boards (PCBs) using automated equipment. Other processes include through-hole technology (THT), where components are inserted into drilled holes on PCBs, and wire bonding, where wires are attached to semiconductor chips. The manufacturing processes for electronic components in an assembly line typically involve the following steps.

1. Surface Mount Technology (SMT): This process involves the placement of electronic components on printed circuit boards (PCBs) using automated pick and place machines. SMT allows for the efficient placement of small components with high precision.

2. Reflow Soldering: After the components are placed on the PCBs, the boards go through a reflow soldering process. This involves heating the board to a specific temperature to melt the solder, which then solidifies and bonds the components to the board.

3. Through-Hole Assembly: For larger components that cannot be placed using SMT, through-hole assembly is used. This involves drilling holes in the PCB and manually inserting the components through the holes. The components are then soldered to the PCB using wave soldering.

4. Inspection: After assembly, the PCBs are inspected for any defects or faults using automated optical inspection (AOI) machines. This helps to ensure that the components are properly placed and soldered.

5. Testing: Once the PCBs pass inspection, they are tested to ensure that they function correctly. This involves powering up the board and testing the various electronic components to ensure that they are working properly.

6. Final Assembly: After testing, the PCBs are assembled into the final electronic component, such as a control module or sensor. The final assembly may involve the addition of connectors, enclosures, or other components.

7. Quality Control: Throughout the manufacturing process, quality control measures are in place to ensure that the components meet the required specifications and standards.

8. Packaging and Shipping: After the components are manufactured, they are packaged and shipped to customers or to other manufacturing facilities for further assembly.

Even if the assembly line is well designed and efficient, variability can still exist due to several factors. Assembly lines are complex systems involving several interdependent processes, and even small variations in one process can have a significant impact on the overall output of the line. One of the primary factors that can contribute to variability in assembly lines is the variability in the machines and equipment used in the process. Machines can experience wear and tear, which can result in changes in performance and output. Additionally, different machines may have slightly different specifications, which can result in variations in the output from one machine to another. The variability in the skills and experience of the operators is another factor that can contribute to variability in assembly lines. Even with rigorous training programs, different operators can have different levels of proficiency in performing certain tasks, which can result in variations in the output. This variability can result in rework, scrap, and delays in the assembly line. Finally, environmental factors such as temperature, humidity, and air quality can also affect the performance of assembly lines. For example, high humidity can affect the adhesion of glue, while low temperatures can affect the viscosity of liquids.

Variability can have a significant impact on costs, depending on the context in which it occurs. In terms of production variability can lead to increased costs due to increase in cycle time, rework, scrap, and lost productivity. Variability in the supply chain can impact costs in a number of ways, such as increased inventory carrying costs due to uncertainty in demand, increased transportation costs due to expedited shipping, and increased sourcing costs due to changes in supplier availability and pricing. Variability in labor costs can occur due to factors such as absenteeism, turnover, and overtime. These can lead to increased costs in the form of decreased productivity, higher recruitment and training costs, and increased labor expenses. Overall, variability can lead to increased costs and decreased efficiency in a variety of contexts, making it an important factor to consider in cost management and planning.

1.7 Structure of the thesis

This thesis is structured as follows. The second chapter is about approaching manufacturing processes with a stochastic cost model. In this chapter, costs are calculated based on the variability present in the cycle time of the workstation after the statistical analysis of the data. In the third chapter, a stochastic novel model based on TDABC is developed to calculate the manufacturing cost of the product. After developing this model, which helps in understanding and calculating the cost per unit of the product, in the fourth chapter a prescriptive analysis has been done to forecast production quantities and the impact of other supply and demand conditions in the assembly lines. The fifth chapter focuses on the optimization of the cost using the proposed cost model, particularly to compute and optimize labor costs and workstation costs and to identify the value added and non-value-added activities. The sixth chapter, which is the last chapter, covers the conclusions derived from this study along with identifying the main contributions, limitations, and opportunities for future research.

Chapter 2

2 A stochastic approach for product costing in manufacturing processes

2.1 Introduction

To control the cost of products and services, companies should have accurate information of the relevant cost objects (Cooper & Kaplan, 1988). Therefore, it is very important to have a proper costing system and controlling of the cost (Pember & Lemon, 2015). Furthermore, companies now are facing competition globally, pushing the development of new products using different and more complex production processes which enforce even more sophisticated costing systems (Drury & Tayles, 1994). Major technical decisions taken in the manufacturing industry related to new products and processes must be supported by complete, accurate, and timely information about costs and profitability (Kaplan, 1984).

Nevertheless, many companies continue to perform the costing of the product in a traditional way, e.g., allocating costs to products proportionally to the quantities produced. Nowadays, companies are characterized by complex systems with multiple products being manufactured in multiple assembly lines. In such situations, traditional costing systems cannot be used. It is very often that cost gets distorted in a traditional costing system as accounting decisions are made as years ago (Cooper & Kaplan, 1987), when the range of products in the company was narrow.

Actual manufacturing systems are characterized by a high level of variability and uncertainty which can drastically affect the cost of the product. To control and monitor the process of manufacturing it is important to consider the uncertainty to achieve the required levels of consistency, quality, and economy (Orellana et al., 2021). The results can be affected by the uncertainty incorporated by the unknown variables that are not modelled. Uncertainty factors in manufacturing processes are demand, cycle time, resources, etc. The uncertain scenarios in the stochastic approach are based on the probability distribution (Wang et al., 2021). Such scenarios can be created for uncertainty based on optimistic, pessimist and neutral forecast values and their probability (Mohseni et al., 2011). For the probability, the mean values of the uncertain parameters can be used for the processing of the stochastic model (Zanjani et al., 2013).

Thus, the motivation to develop this work is, as there is space for improvement in the costing of the product in practice (approaches and tools used by companies) and conceptually (concepts and models). It is indeed necessary to accommodate all the variabilities and uncertainty that prevails during the manufacturing process. This chapter focuses on developing a stochastic approach to costing systems which considers the variability in the process cycle time of the different workstations in the assembly line.

Such an approach will provide a range of values for the product costs allowing a better perception of the risk associated with these costs, instead of providing a single value of the cost. The stochastic multiscale approach is a useful tool to solve the engineering problems considering uncertainties, but it is computationally intensive as the number of uncertainties increase (Tang et al., 2022).

Stochastic analysis is associated with the analysis of events in which at least one of the components of the process is random. It is usually applied in situations that face a set of random variables over time. A Markov process is described as such when only the current value of the variable is taken into account to predict its future value. Process simulations are associated with the random generation of values, assuming that they may be in a given interval or in accordance with, for example, the variable's mean value and respective standard deviation (Abounadi et al., 2002). Usually, values are generated following a normal distribution of mean x and standard deviation s. The stochastic frontier analysis can be used to identify the efficiency and productivity of manufacturing processes. A model can be designed for each plant, process or production line and these models can be combined to achieve overall performance. This technique can be used to bridge the gap between a plant´s current productivity and a set target to be achieved (Oh & Shin, 2021).

Thus, in order to achieve that it is essential to use appropriate approaches to understand, in the first instance, the variability and then to identify a range of values for the product cost applying statistical methods. With applied mathematics, a better perception of the product process can be achieved, and better decisions can be made.

This chapter extends and complements recent research on applied cost models in the industrial engineering setting where stochastic modelling and new models for the allocation of process costs to products were used to improve performance evaluation, optimization of maintenance costs, supply chain costs, investment decisions and production inventory with flexible manufacturing (Singh & Prasher, 2014)(Noh et al., 2019)(Singh & Singh, 2020). Recent work by (Wang et al., 2021) shows the use of the stochastic approach in production planning and to optimize the profit as uncertainties are involved. By using the methodology proposed in this chapter, it will showcase the historical statistical data of a defined period so the variability in the process can be observed, and the real time cost of the product can be calculated. This can contribute significantly to support continuous improvement in production systems. Moreover, it can also bridge the gap between financial and production departments, integrating production and accounting information. The comparison between the standard and the real cost can be done in a better way which will facilitate both operational and strategic decision making.

In this chapter, the developed six-step methodology is presented and explained. Firstly, a data analysis has been performed to obtain relevant descriptive statistics and identify the outliers which must be removed from the product cost analysis. After removing the outliers, the descriptive analysis must be done again to understand the data. The third step is to perform the hypothesis test to compare cycle times across the different assembly lines. The next step is to obtain the confidence interval for the mean which gives the information about the potential risk and variability in the cost of the product. Also, quartiles Q1 and Q3 will give a range of potential values for the computation of product costs highlighting the inherent risk of such cost.

Specifically, one product was taken into consideration on which analysis has been performed. To manufacture this product, it needs 18 workstations. Only the bottleneck workstation of the assembly line was considered. Since it represents the cycle time of the assembly line. So, any variability and uncertainty in the bottleneck will affect the cycle time of the assembly line resulting in changes in the cost. The cycle time of the bottleneck workstation has been gathered and after analyzing and removing the outliers, a descriptive analysis was performed on it. After performing the statistical tests, it was evident the existence of variability and uncertainty. Following the stochastic approach, the range of cost has been calculated using the quartiles and confidence interval for the mean. This stochastic range of cost accommodates the variability and uncertainty that can prevail during the manufacturing process in a specific period for a certain type of product.

In the next section, a literature review on uncertainty and variability, and stochastic approaches in costing systems is presented. To counter the impact of uncertainty and variability on cost computation, the use of a stochastic approach is proposed. The methodology is explained in section 3 and the results of its application are presented in section 4. The computation of product costs and relevance of the proposed methodology are discussed in section 5. Finally, the final remarks and opportunities for further research are presented in the last section.

2.2 Literature review

2.2.1 Uncertainty and variability in cost models

Uncertainty can be defined as a potential deficiency in any phase or activity of the modelling process that is due to the lack of knowledge which causes model-based predictions to differ from reality (Oberkampf et al., 1999). At the time of cost estimation, there are significant sources of uncertainty associated with it. During the estimation of the cost, the information of the cost available are usually the historic product costs and sometimes this information has a high degree of uncertainty for instance, first in the design stage, and later, in the manufacturing process, e.g., the dimensions of the product, process cycle times (Scanlan et al., 2006). The uncertainty in the cost estimation can be reduced by providing more complete information about the manufacturing process, product design, product support, reliability, and requirements of disposal (Corbett & Crookall, 1986).

When an event is modelled, a decision must be made by the decision maker if it is necessary to include the uncertainty that prevails with the event. In some cases, it is enough to approximate the uncertain events by a deterministic model. But if uncertainty happens to be significant and can impact the results then it is important to include it in the model and uncertain events should be studied. In this kind of situation, it is necessary to choose a proper method to model the uncertainty (Zimmermann, 2000). Uncertainty and variability are present in many real-life events, e.g., variation in the cycle time of assembly lines. It is possible to occur from several reasons such as lack of information, ambiguity, complexity in the information, errors in measurements, and beliefs instead of real information, etc. (Zimmermann, 2000). The level of uncertainty is related to the information available and the complexity of the event. Uncertainty can be related with the possibility of error for not having ample information about the event and the surrounding environment (Nachtmann & Needy, 2003).

There is no definitive definition for uncertainty. Perhaps there are some agreements in very specific fields (Goh et al., 2010). For example, in the financial domain, there has been a fine distinction made between uncertainty and risk. Indeed, it is normal to talk of risk when the future is unknown, but the probability distribution of likely future is known. On the other hand, uncertainty occurs when the probability distribution is itself unknown (Miller, 1977). (Zimmermann, 2000) proposes the following definition: "uncertainty implies that in a certain situation a person does not dispose about information which quantitatively and qualitatively is appropriate to describe, prescribe or predict deterministically and numerically a system, its behavior or other characteristic".

It is evident from different sources that uncertainties and variabilities can impact cost estimation. If more information is available, it is possible to reduce the uncertainty in the estimation of the cost specifically details about manufacturing and design of the product (Goh et al., 2010). In the field of cost calculation and costing systems many methods can be used. In the past, sensitivity analysis was used to estimate the cost and uncertainty by measuring the level of risk. It can give brief information to the decision-maker about what can happen if the input changes considering the best-case and worst-case situation. Although this method has its limitations, for example when there is a considerable number of variables it is more difficult to estimate what can happen and understand how the interaction among the variables can affect results. To lessen the gravity of this limitation, various probabilistic methods are developed. In these methods, the uncertainty in the cost information is represented by probability functions. Probabilistic methods help decision makers with more information, but they require a larger amount of information and greater statistical handling than sensitivity analysis (Datta & Roy, 2010).

Management of the cost is important to understand the cost behavior as the variability in the market, prices, work methodology affect the cost. This variability can affect the cost of the product; hence the efficiency of the company can be affected. A better understanding of product variability is necessary for better control over the cost. A proper risk assessment asks for a good understanding of the impact of uncertainty and variability on costs (Afonso & Jiménez, 2015)

Several methods including activity-based costing have been proposed but they have been unable to efficiently incorporate uncertainty and variability. It has been proposed to use the fuzzy approach which is based on the rationality of uncertainty because of factors like vagueness and inaccuracy (Zadeh, 1965). In cost-volume-profit analysis various factors such as risk and uncertainty are ignored, thus it severely limits the usefulness of this methodology, by using the fuzzy set concept it is possible to handle imprecision, and that gives access to evaluate the cost-volume- profit decision making process (Yuan, 2009). Monte Carlo simulation and fuzzy set theory are better methods to handle the uncertainty and inaccuracy in the data for cost models (Nachtmann & Needy, 2003). When uncertainties are partially or totally random, they are usually represented by probability density functions which can be used also in costing systems (Byrne, 1997). When the decision maker faces an uncertain problem, it can be expressed as uncertain ratios also known as fuzzy numbers (Kishk & Al-Hajj, 2002). In recent times, techniques based on the fuzzy numbers have emerged as an excellent tool to manage uncertainty in models used to calculate costs (Nachtmann & Needy, 2003). The uncertainty of the parameters in linear planning models has been also approached in terms of fuzzy numbers (Wang et al., 2021). For example, fuzzy based activity-based costing can help to compute the cost (Durán & Durán, 2018). Additionally, it can allow managers to control the cost and implement useful indicators for better quantitative information on the performance and behavior of important drivers, thus improving decision making (Durán & Durán, 2018).

2.2.2 Stochastic approaches in cost models

A stochastic model can be defined as the collection of random variables which are arranged in a specific mathematical set, which is associated with an element of the set (Anderson et al., 1985). A stochastic model is a tool to estimate the probability distribution of the possible results by allowing the random variable in one or more inputs over time. The random variables are based on the variability and uncertainty found in the historical data for the selected time using standard time series methods. The distribution of the probable results is derived from the number of stochastic projections which reflect the random variable in the inputs (Ioannou et al., 2017).

A stochastic cost model will set up a projection model which allows the analysis of single products, several products in the same assembly line and all products in the entire company. It can use random variation to understand what conditions (risk, variability, and uncertainty) can affect the cost of the product. In the end, the distribution of the outcome can portray not only the most likely cost of the product but the whole possible range of product costs from a reasonable set of assumptions and constraints. The most likely estimate is given by the distribution curve also known as the probability density function which is typically also the mode of the curve (Ioannou et al., 2017).

The advantage of a stochastic approach is unlike the deterministic approach which only gives a value to the product cost, this method of cost modelling gives the entire range of possible cost values of the product considering the related variability, risk, and uncertainty. It can reflect better real-world situations, providing different ranges of possible results. Also, by running several rounds of calculations, and using many different estimates of future economic situations, the model will show the range of costs showing the potential upside and downside of each.

Stochastic frontier analysis models are being used as a statistical benchmark to provide an overview of the industrial sector. It can also be adjusted to understand plant productivity. But the general form of stochastic frontier analysis is difficult to implement in a complex manufacturing industry because of the problems of multicollinearity (Oh & Shin, 2021). It is important to have accurate information and with modern complex systems, deterministic models are impractical as it cannot represent the disturbance and uncertainty adequately. So, it is necessary to apply stochastic models (Orellana et al., 2021).

The deterministic cost model was extended by (Ioannou et al., 2017) to systematically account for the stochastic input. They performed the Monte Carlo simulation to derive the joint probability distributions which allows to estimate the probabilities of exceeding a set of thresholds and determining the confidence interval. It stressed the importance of the appropriate statistical modelling of stochastic variables so that it can reduce modelling uncertainties and can contribute for a better-informed decision making to make the investments. An advanced stochastic model was applied to identify the most relevant parameters which influence cost and uncertainty. The stochastic model was implemented by using a @risk software extension. It also investigates the variables which behave as the most important cost drivers, which are behind the effective reduction of the cost thus, providing information on where additional efforts are required and can effectively reduce the costs (Mora et al., 2021).

There are stochastic cost models addressing the influence of uncertainties in wind generation on the optimal operation of power systems (Swaroop et al., 2009). In this particular case of wind generation, there are a lot of possible scenarios, so a scenario reduction algorithm has been applied. The stochastic model will give a near possible optimal solution considering all possible scenarios. That particular solution cannot be optimal for one particular scenario, but it is robust over all possible realization of the uncertainties. An adaptive particle swarm optimization algorithm was proposed to solve the stochastic cost model which overcomes the traditional drawbacks such as penalty coefficients, and parameter tuning. (Sobu & Wu, 2012) developed stochastic scenarios by using observed mean-values and standard deviation from data, and then based on these stochastic scenario data, stochastic operation cost optimization models for minimizing operation cost were formulated. Then, they used a particle swarm optimization approach.

Stochastic approaches were also applied in the context of life cycle costing to study the economic value of energy-efficient building retrofitting investments (Baldoni et al., 2019). It also investigates the effect of interdependent stochastic variables such as explicit evaluation. The economic evaluation is itself stochastic so it can express both the expected value and the value inherent to uncertainty and risks. The actual validity of stochastic costing depends on the reliability of the conclusions that can be drawn. This reliability consists of the robustness of the results and accuracy with respect to the real system which the model aims to represent. The first aspect in particular not only generates the stochastic life cycle cost outcomes, but also allows to assess the variability of the outcome given by the dataset.

Furthermore, (Baldoni et al., 2021) developed a novel software tool for the evaluation of life cycle impacts and costs assessments, which is aimed to support decision making. This tool allows evaluating the long-term trade-off between economic and environmental performance of investment projects, while accounting for uncertainties in input parameters. The software also includes several tools for sensitivity analysis.

Thus, by using a stochastic approach, the range of cost showing the potential upside and downside can be obtained (see Figure 2.1).

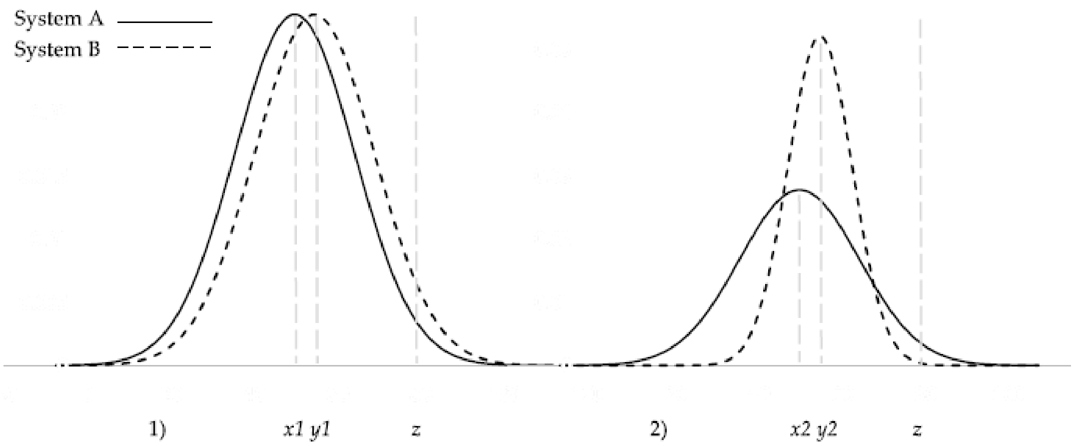

Figure 2.1 Comparison between deterministic and stochastic values in different situations

In the figure above, based on (Park et al., 2010), it can be observed that from a deterministic perspective Systems A and B are far below risk *z*, but that analysis is insufficient, and a probabilistic approach should be taken into consideration. Situations 1 and 2 show the range of possible values for each system under different conditions. In situation 2, System A, even if presenting a lower average (x_2), has a higher risk of reaching *z*. Nevertheless, the risk of System B in situation 1 is higher. Thus, we should go beyond the simple deterministic analysis and understand the behavior of the phenomenon under study. This reasoning can be applied to products, activities and processes and a stochastic cost model will turn possible to measure such variability and its impact on the cost of the relevant cost objects. The range of costs, along with the measurement of predictions should be provided. Such predictions should be based on the calculated risk for the relevant variables.

2.3 Materials and methods

In this chapter, to analyze the product development process variability, it is proposed a methodology based on the following six main steps.

1. Firstly, data analysis must be done to identify the descriptive statistics and outliers by activity and process. Namely, mean, standard deviation, quartiles, minimum and maximum values, range, and coefficient of variation. In the case study, the analysis was made by workstation and line. If the outliers are caused due to external causes, then they must be removed.

2. After this removal, a new descriptive analysis must be performed to do a critical assessment. In this step it is intended to identify what is happening and if it is possible to find differences between lines for a specific workstation and product. The analysis of the outliers is important to analyze the efficiency of the process and to identify opportunities to improve the process.

3. The third step is related to performing hypothesis tests, where it is intended to compare the workstation in different lines and identify in each one if there are significant differences.

4. The next step is to perform the confidence interval for the mean, considering the level of confidence as 95%. Note that instead of considering the mean value to compute the product cost, the confidence interval gives information on the variability and the potential risk of the cost. In this case, the variability is being analyzed since it is a range of values.

5. Furthermore, the quartiles values can be another possibility to take into consideration, also giving an idea of the risk associated with product cost. The interval for the values can be computed using the first and the third quartile thus, focusing on the 50% of the values around the median. The value associated with Q3 is a measure for product cost risk because 25% of the produced units will have a higher cost than such value. This is a conservative approach for cost risk analysis and the 90-percentile can be used to signalize the risk of a too high cost. On the other hand, Q1 represents a reference value for quotations because prices lower than this value will push the margin to negative values. Again, alternatively, a less conservative approach can be used considering, in this case, the 10-percentile. Thus, both the lower and upper limits can be used as risk measures.

6. Finally, the values achieved in the confidence interval and the quartiles (first and third) can be used to compute the product cost considering each workstation or aggregated by line (usually, considering the bottleneck of the line).

Note that the proposed approach can be applied differently, depending on what is intended to be achieved. The six steps presented before can be simplified or developed, if necessary. For example, it can be only necessary to identify the confidence interval for the mean instead of the quartiles or vice-versa.

To compare two independent samples, one can use parametric and non-parametric tests. A parametric test must be used when the population follows a normal distribution, have equal variance and it is continuous. However, non-parametric tests are applied when at least one of the parametric assumptions is not validated. Note that parametric tests are more robust than non-parametric and consider less information to make stronger conclusions. Therefore, the student t-test is a parametric test commonly used to identify if the mean of one sample is different from a known mean or to identify if there are differences between the mean of two identical samples. Furthermore, Mann-Whitney test is a non-parametric alternative to evaluate if two samples are from the same population (Kaur & Kumar, 2015). In order to check if the sample follows a normal distribution, Shapiro-Wilk and Kolmogorov-Smirnov are the well-known tests to do it. Where Shapiro-Wilk is commonly used in small samples ($n \leq 50$) and Kolmogorov-Smirnov is applied in the other cases (Yazici & Yolacan, 2007).

Moreover, the Wilcoxon Sign Rank Test was used for the validation and analysis of the computed costs. It is a t-test alternative since it is a non-parametric test. Thus, this test intends to evaluate if the median of one sample is different from a known value, instead of using the mean value (Montgomery & Runger, 2018). In this case, the test was used to assess whether the calculated values of costs present significant variations over the weeks in relation to the planned/standard cost. The Wilcoxon test was performed, paired and with the assumption of two-tailed distributions and for a significance level of 5%. This test allows us to evaluate the differences or disparities of the median values of the data, being useful to understand whether observations or values of the same variable, recorded at different times, present significant variations or not. This way, it will allow us to evaluate the adequacy of the cost model and evidence the existence of cost variability justifying the stochastic analysis of costs in detriment of the traditional deterministic approach.

2.4 Analysis of results

The proposed methodology to include variability in costing systems was applied following the six steps explained before. A data sample was considered corresponding to a weekly period (seven consecutive days), considered normal (they were not considered holiday periods, breaks or other), referring to all production lines (A, B, C and D) where the product is produced. These are data obtained from the production information system, where the cycle time values for each day, per line and workstation, are recorded. These, in turn, were processed, removing the outliers from the cycle times per line and workstation. Next, the mean, quartile, extreme values (maximum and minimum), standard deviation and coefficient of variation were computed.

In each line, composed of several workstations, the bottleneck (workstation with the highest cycle time) was identified, and its frequency (count) was identified corresponding to the number of units produced. In total, it is a sufficiently large sample, with about 38,000 observations, which are distributed by all the lines, in a variable way, but also with considerable frequency values. That is, we have frequencies much higher than 50 (even the minimum value exceeds 3,000).

Nowadays, for a company to be able to respond to customer demand and bring value through its products, it must be able to produce with great flexibility and diversity. To do so, it is necessary an enormous complexity in the production process. Having complex processes in the assembly lines causes variation in the cycle time of the workstations which will consequently affect the cost of the product. Thus, if a company wants to be competitive it must understand and control the variation of the several activities that compose the production process. This asks for a stochastic approach in controlling activities of productions processes.

It was selected a product produced in 4 different production lines. These lines are considered semi-automatic lines since they require manual assembling (performed by operators) and automatic assembling. All products pass through different tests, most of them automatic, but also tests with human intervention. Before arriving at these 4 lines the product had already gone through other processes in the factory being the studied process the final one before shipping the product to the client. Small lines A and B produce fewer quantities and therefore have fewer operators allocated. Lines A and B have a different number of machines per workstation. Big lines C and D are considered large lines because their production volume is much higher than small lines.

In order to analyze the variability between lines (A, B, C, and D), the bottleneck's cycle time (workstation 17) was analyzed. Figure 2.2 presents the cycle time of each piece produced in workstation 17 in the 4 production lines in one week. The data (i.e., cycle times and daily produced quantities) were collected for the period between the 18th and the 24th of December 2020. Both the quantities and process times were different in each production line, so this data clearly highlights the variability that exists in the production process. So, the cycle time at each workstation was recorded for each product unit manufactured during that week, what helps to check the correct bottleneck of the production line. Once confirmed that the 17th workstation represented the bottleneck, the tests were made on that workstation as it will define the production line cycle time. All the recorded cycle times from each assembly line were extracted from the company's ERP system to the SPSS software where various tests were done on the data.

As mentioned earlier, there are four different assembly lines involved in producing the product under scrutiny. Line A and B are considered small lines and line C and D are considered as the big lines. The difference between small and big lines is the amount of equipment at each workstation. Big lines have more equipment compared to small lines. As they have more equipment in the workstations, big lines can process more parts in parallel. Hence, big lines produce faster and in greater quantity. Big lines produce around 15,000 parts per week whereas small lines produce around 3,000 parts. Production planning and scheduling prioritizes big lines, and small lines complement the big ones.

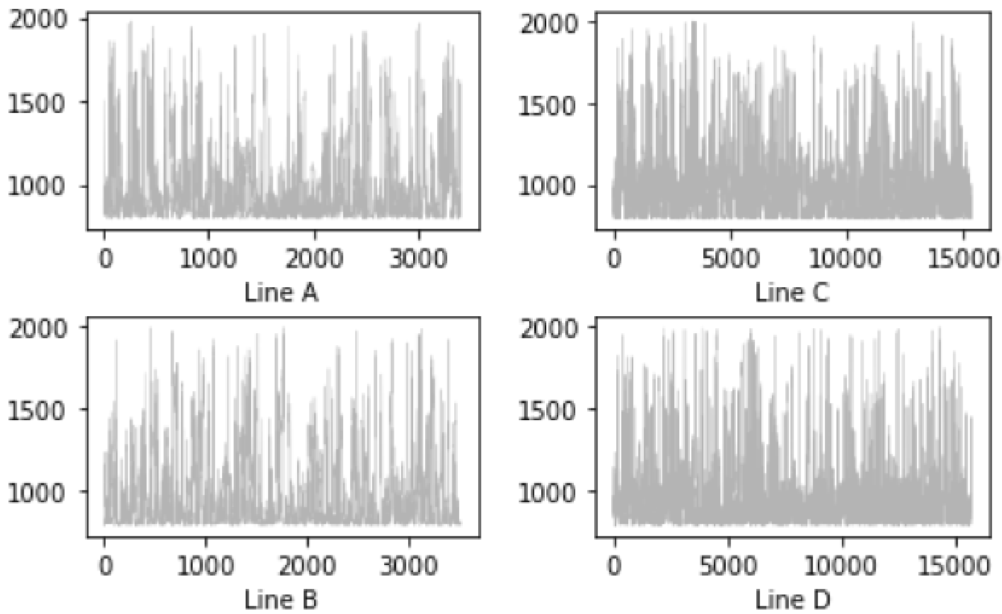

Figure 2.2 Cycle time (in seconds), per line.

A descriptive analysis was conducted. Table 2.1 presents the number of observations, mean, standard deviation, minimum, first, second and third quartile (Q1, Q2 and Q3), maximum, range and the coefficient of variation, for each line. The coefficient of variation is commonly used to identify if the mean is representative. When this metric is less than 50%, then the mean is representative. Otherwise, if it is not, it is preferable to use a median instead.

Table 2.1 Descriptive statistics (in seconds), per line

	Line A	**Line B**	**Line C**	**Line D**
Count	3,407.00	3,510.00	15,314.00	15,664.00
Mean	984.28	992.15	993.30	952.42

Standard deviation	240.33	257.90	217.50	210.91
Minimum	789.00	786.00	786.00	784.00
Q1	823.00	807.00	843.00	815.00
Q2	886.00	868.00	938.00	888.00
Q3	1,036.00	1,089.75	1,049.00	988.00
Maximum	1,981.00	1,997.00	1,999.00	1,998.00
Range	1,192.00	1,211.00	1,213.00	1,214.00
Coefficient of variation	24.42	26.00	21.90	22.14

According to the results obtained, lines A and B have almost the same number of observations. The same conclusion can be drawn for lines C and D. Furthermore, according to the mean, line B has a longer cycle time when compared to line A. Besides that, line C is the one with the longer cycle time, when compared to line D. The minimum and maximum cycle time is nearly the same for all the lines. Another conclusion is that the mean is representative in all the lines, although there is variability since the range values (the difference between the maximum and minimum value) is high. Thus, it is important to understand the cause of these high values to avoid wrong conclusions.

According to what was observed, lines A and B present very close mean values, being the difference 7.87 seconds while in lines C and D the difference between the mean values is 40.88 seconds. In terms of standard deviation values, the difference is greater between lines A and B than between lines C and D, being respectively 17.57 and 6.59 seconds. That is, the small lines show a tendency for greater variations in cycle times, around the mean. The interquartile range is 213 and 282 seconds for lines A and B, respectively, and 206 and 173 seconds for lines C and D, respectively. There is a greater difference in the small lines compared to the big lines.

For the coefficient of variation, the values on the small lines are close (differential of 1.58 seconds) but on lines C and D they are even more similar (differential of only 0.24 seconds). All lines show variation, although the highest values are observed in the small lines. That is, in general, the pairs of lines ((A, B); (C, D)) have characteristics that resemble each other, namely, Count, Mean, Standard Deviation and

Coefficient of Variation and, at the same time, allow the distinction between the two types of lines (Small and Big lines).

The available capacity and cycle times of the machines is fundamental to allocating the cost of resources used to the cost objects. The variability in cycle times will also give us information on the variability of the cost. Therefore, it is necessary to study the variability of the cycle time and the average confidence interval can be a way to do it. Hence, in Table 2.2 it is shown the confidence interval for the mean cycle time in each line (given by the cycle time of the line's bottleneck which is workstation 17).

Table 2.2 Confidence interval for the mean cycle time (in seconds), per line

	Line A	**Line B**	**Line C**	**Line D**
Lower bound	976.21	983.62	989.86	949.11
Upper bound	992.35	1,000.68	996.75	955.72

We can see that line D has the smallest values and line B and C have the higher ones. With these results, there is a suspicion that there are differences between lines A and B, and between lines C and D. Differences between lines should be identified and analyzed because they can result from different and not optimized planning, efficiency, demand requirements, etc. Considering the high variability in internal processes and external demand these differences must be monitored on a weekly or monthly basis to support effective and timely action plans from a continuous improvement philosophy.

To analyze these differences and trigger eventual action plans, non-parametric tests were performed since the lines do not follow a normal distribution. To evaluate the differences between lines, the analysis was conducted considering line pairs A and B, C and D. Thus, the Mann-Whitney test was performed to assess differences between lines and the hypotheses to take into consideration were:

H_0: There are no significant differences between lines in terms of the average execution (cycle) time.

H_1: There are significant differences between lines in terms of the average execution (cycle) time.

Table 2.3 presents the p-value for the Mann-Whitney test and the mean value for each line pair. According to these results, the hypothesis H_0 is rejected since the p-value is less than the level of significance (α=0.05). Therefore, there are significant differences between the cycle time in lines A and B.

Where the same conclusion can be drawn for lines C and D. According to the mean, lines B and C have higher cycle times than lines A and D, respectively. This variation between the small and big lines can influence the product cost and represent opportunities for improvement in process costs. In other words, if computed by line, it is expected that the product cost will be higher in line B than in A.

Table 2.3 Results of Mann-Whitney test

Lines	Mann-Whitney	Mean (seconds)
A & B	<0.001	984.28 & 992.15
C & D	<0.001	993.3 & 952.42

After this analysis, it is important to verify if there are outliers. Thus, Figure 3 presents the boxplot to visualize the cycle time variation in each line. With this visualization it is possible to identify outliers and, since there are too many, they contribute to a very high variability. Hence, it is essential to understand why these values are happening to reduce such variability.

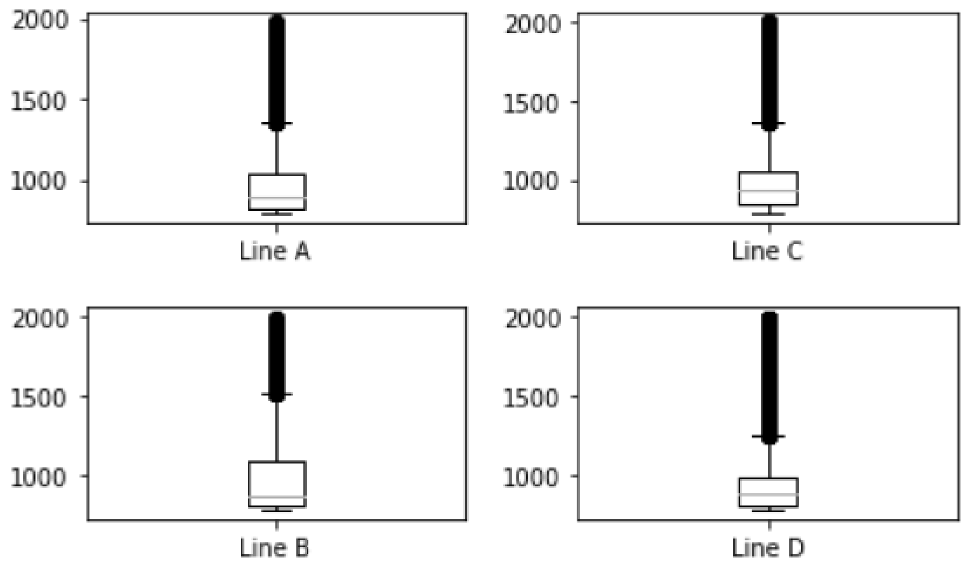

Figure 2.3 Boxplot for the cycle time (in seconds), per line.

The mean cycle time is higher for the small lines compared to the big lines. When the demand is lower than the total capacity given by the four lines, the company chooses to produce in the big lines at full capacity, complemented by the small lines. This causes those small lines to produce below their capacity and also reduces the performance of the small lines and their cycle times are higher than in big lines.

In terms of product cost, if the cycle time has a higher variability, then the variability on cost will be higher. Minimizing the final cost is important to increase the margin and minimizing variability contributes to decreasing cost risk. Outliers are caused by internal and external factors to the process, which should be managed differently, namely in the context of continuous improvement or within the costing system. Then, a new analysis was conducted without the outliers to reduce the variability which can be managed within the costing system. Figure 2.4 presents the cycle time, per line, for workstation 17 without the outliers. In the first analysis, it can be observed that the maximum value has decreased in all lines.

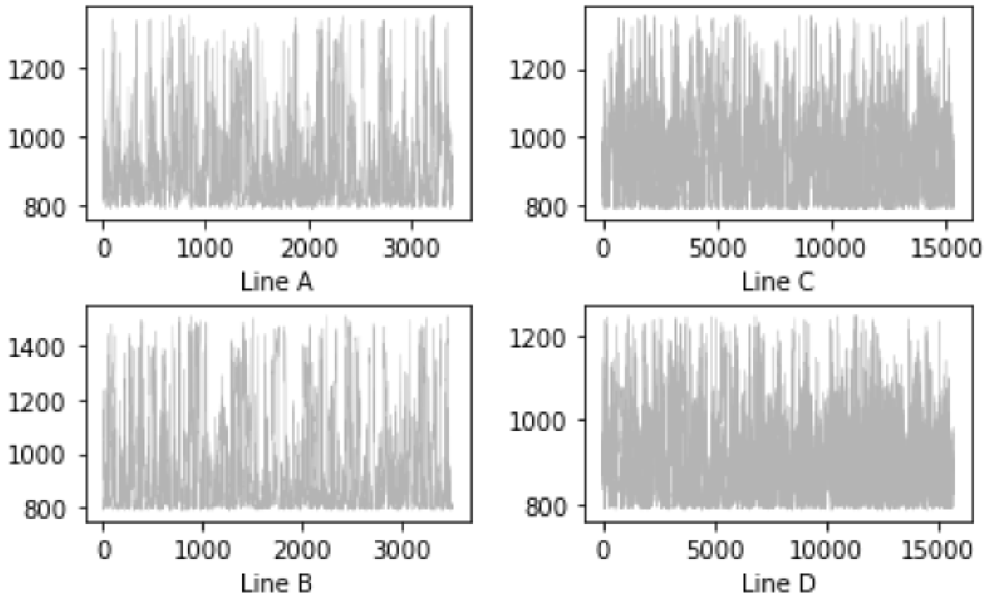

Figure 2.4 Cycle time per line (in seconds), without outliers

The next step of the proposed methodology is to perform the descriptive statistics to identify which metrics change when the outliers are removed. Thereby, Table 4 presents the descriptive statistics, and we can see that most statistics have decreased. Except for the minimum, which remained the same. Besides that, the range decreased considerably, as was expected and, according to the coefficient of variation, the mean is still representative. Moreover, there is more evidence that the cycle time is different in the small and big lines since the means are slightly different.

Table 2.4 Descriptive statistics of cycle time (in seconds) without outliers, per line

	Line A	Line B	Line C	Line D
Count	3070.00	3305.00	14096.00	14286.00
Mean	917.85	948.39	940.65	896.19
Standard deviation	129.38	191.53	121.74	96.41
Minimum	789.00	786.00	786.00	784.00
Q1	820.00	805.00	835.00	811.00
Q2	869.00	854.00	923.00	873.00
Q3	979.75	1040.00	1015.00	956.00
Maximum	1355.00	1513.00	1358.00	1247.00
Range	566.00	727.00	572.00	463.00
Coefficient of variation	14.10	20.19	12.94	10.76

Furthermore, the confidence interval for the mean cycle time is presented in Table 2.5, considering the confidence level as 95%. These values also decreased, and the amplitude is, also, smaller. With these results, it is expected that there are differences between the cycle times per line.

Regarding the analysis of the measures without the presence of outliers, the count values are very similar when analyzing the pairs of lines, A and B and C and D. There is a greater difference in the means between these two pairs of lines and the respective standard deviation values. The values are lower compared to those obtained with the presence of outliers but more differentiated between lines of the same type. The coefficients of variation are also lower for all lines, but there is a greater difference between them when analyzing pairs of lines, A and B, C and D.

Table 2.5 Confidence interval for the mean per line, without outliers

	Line A	Line B	Line C	Line D
Lower bound	913.27	941.86	938.64	894.61
Upper bound	922.43	954.92	942.66	897.77

To verify if there are differences in the cycle times per line, the Mann-Whitney test was performed, and Table 2.6 presents the results achieved. According to the p-value, in the Mann-Whitney test, there are significant differences between lines A and B. The same conclusion can be drawn for lines C and D. Lines B and C have a higher cycle time when compared with lines A and D, respectively. Thus, the conclusions are the same when it was used all the available information. However, it is important to remember that we intend to analyze the variability within product cost, where extreme values can lead to wrong conclusions.

Table 2.6 Mann Whitney results without outliers

Lines	**Mann-Whitney**	**Mean**
A & B	<0.001	917.85 & 948.39
C & D	<0.001	940.65 & 896.19

After these analyses, the last step of the proposed methodology is to identify how many values are in each quartile to provide optimist and pessimist estimations for the product cost instead of a deterministic cost. Therefore, Table 2.7 presents the number of observations in each quartile, where the first count is the first 25% of data, the second for the values of 25% to 50% of data and the last one for 50% to 75% of data. For example, in line A, there are 795 observations with the cycle time less than or equal to 820. The product cost for these cycle times will be the lowest when compared with the other quartiles because there is a reduced consumption of the resources. Thus, using these values it is possible to propose a range for product cost and measure cost risk, particularly, using the cycle time achieved in Q1 and Q3, respectively.

Table 2.7 Quartiles and number of observations for each line

Line	**Q1**	**Count**	**Q2**	**Count**	**Q3**	**Count**
A	820	795	869	743	979.75	764
B	805	838	854	819	1,040	823
C	835	3,536	923	3,527	1,015	3,515
D	811	3,588	873	3,572	956	3,572

Taking into consideration the results achieved and presented in Table 2.7, the boxplot (Figure 2.5) was performed to visualize the variability of the data. Thus, there are new outliers, which are included in the variability of the process that is intended to be allocated to product cost. Initial outliers are supposed to be removed or, if not, to be allocated to the product as general costs not specific to the process/line. Identifying the different levels of cost and understanding their behavior is so important as allocating them to products. Costs can be specific to each produced unit, to the batch, the process, general costs of the product or general cost of the company/business. High variability in cycle times can be explained by reasons related to all these different levels.

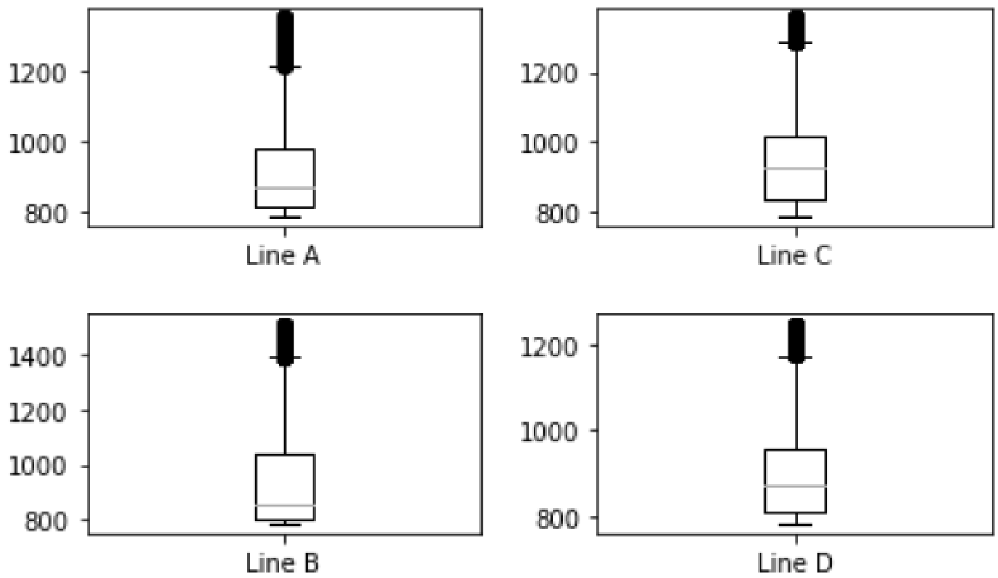

Figure 2.5 Boxplot for the cycle time per line, without outliers

Thus, for the inclusion of variability in the computation of product costs they will be used the cycle times associated to Q1 and Q3 and the confidence interval for the mean. Both can be calculated or estimated for each workstation or just considering the bottleneck of the line (in this case, workstation 17). The analysis made was to compare production lines thus, it was centered on the bottleneck which defines the production speed of the line. After this high-level approach to optimize production lines, a detailed analysis within each line should be made to analyze and optimize workstations.

2.5 Discussion

2.5.1 Main assumptions

The cost analysis made in this case is focused on the manufacturing cost and on process variability. Further work can be done to extend it to the other dimensions of variability and the non-manufacturing costs. Thus, to calculate the cost of the product, these are the key inputs of the cost model, namely:

- Quantities demanded by the client.
- Available time to produce the product in the line: (n° of days x shifts per day x minutes per shift x 60).
- Workstations - the stations where the work associated with each process is carried out.
- Number of equipment per workstation and respective investment costs (i.e., depreciation).
- Cycle times per unit produced.
- Tariffs for the different resources used (e.g., area, energy, maintenance).

A general expectation is to have an Overall Equipment Effectiveness (OEE) of 90%, taking into consideration the possible losses while production. This efficiency of 90% multiplied by the takt time will lead the real time expected by the production line. The main resources are related to labor, depreciation, maintenance, auxiliary material, energy, area, other internal costs, tooling, etc. Considering the planned quantities and the budgeted costs, a specific tariff for each category of resources can be calculated. Summing all those tariffs, it can be obtained the general tariff for the line. Table 2.8 below shows the general tariff for each line.

Table 2.8 General tariff (in euros) per line

Line	General tariff
A	0.1178
B	0.1243
C	0.1470

D	0.1470

Tariffs are different if resources used and/or available capacities are different. Lines C and D use similar resources and offer identical capacity levels. Having the values of the tariffs we can calculate the cost of the product in the different lines. To calculate the cost of the product, one must multiply the general tariff by the cycle time.

According to the statistical analysis performed and presented in the previous section, we can obtain the range for product cost considering process variability in each line. The first quartile of the cycle time is considered as the lower range and the third quartile is considered as the upper range value giving us an interval of expected variation and allowing us to estimate the risk of cost. The lower range helps with budgeting exercises, quotations, and the development of new products because it represents the potential for lower costs of the product. The upper range gives an alert that margins can be compromised if the efficiency of the line is not improved. In this case, a conservative approach was followed taking the values for Q1 and Q3, but these limits could be calculated using the 10th and the 90th percentile.

2.5.2 Computation of the costs

Table 2.9 presents the range for product costs considering the values for the first and third quartiles which cover 50% of the values around the median considering one week of analysis. It is important to note that in line A and line B there are 18 parts that are produced in parallel and in line C and line D there are 36 parts.

Table 2.9 Range of product cost considering Q1 and Q3, per line

		Process Time (seconds)			**Cost (euros)**	
Line	**Tariff**	**Lower range**	**Upper range**	**Parts in parallel**	**Lower range**	**Upper range**
A	0.1178	820	979.75	18	5.366	6.411
B	0.1243	805	1040	18	5.558	7.181
C	0.1470	835	1015	36	3.409	4.144

D	0.1470	811	956	36	3.311	3.903

In Table 2.10 the range of the cost can be observed based on the mean value of process time with a confidence interval of 95%. By using this smaller amplitude, cost variability is reduced and results more related to the standard efficiency of the process.

Table 2.10 Range of product cost considering the confidence interval for the mean, per line

		Process time (seconds)			**Cost (euros)**	
Line	**Tariff**	**Lower range**	**Upper range**	**Parts in parallel**	**Lower range**	**Upper range**
A	0.1178	913.22	922.43	18	5.976	6.036
B	0.1243	941.86	954.92	18	6.504	6.594
C	0.1470	938.64	942.66	36	3.832	3.849
D	0.1470	894.61	897.77	36	3.652	3.665

It can be observed that the tariff for each line is different because the number of equipment is different except for big lines (C and D). In line B there is more equipment than in line A, but planned quantities are almost the same. Thus, the amortization cost per product unit in line B increases. Hence, the parts produced in line B are costlier. Line C and D are a replication of each other, so they have similar tariffs. The amount invested in equipment in big lines is bigger than in small lines but at the same time, the quantities produced in these lines are way higher. Therefore, the product produced in the big lines has a smaller cost although having more equipment in each workstation. The cycle time of the assembly line decreases with the increase in the amount of equipment in the workstations.

The methodology purposed here is related to some work done in the recent past. For example, (Zanjani et al., 2013) developed a stochastic model using the mean values of the uncertain parameters for the probability distribution. The developed model was applied in the milling industry with the purpose of supporting the production planning. Also, (Sobu & Wu, 2012) developed stochastic scenarios by using observed mean-values and standard deviation from data, and then based on these stochastic scenario data, stochastic operation cost optimization models for minimizing operation cost were formulated. These models were used to measure uncertainty in power generation and renewable energy.

By using this methodology firstly, it is easy to understand how the assembly line is performing. It can be identified that how many parts are produced falls under the standard cycle time allocated for the production. With a stochastic approach, the range of the cost is available which can help the manager to make decisions about the planning of the production as this approach can facilitate the understating about the real-time cost of the product along with the allocation of production quantities for each assembly line. It is important to note that even though the assembly lines are replicated to each other, there may be some variability in them. Line C and D although being exactly like each other have a significant difference between them.

2.5.3 Cost analysis per line and product

For a better understanding of the variation in the data and analysis of their unpredictability, the mean values of cycle times corresponding to 12 weeks of three consecutive months were extracted and analyzed. Each of these weeks corresponds equally to a period of seven days.

For each of the lines, a confidence interval for the mean of 95% was calculated, as well as the first and third quartiles. To better understand the variation in the final costs per line, these were calculated according to the values presented in Table 8, that is, multiplying the cycle times obtained by the respective tariff.

Thus, the cost variation intervals were found when considering the confidence interval for the mean and the interquartile range. The values are shown in Figure 2.6.

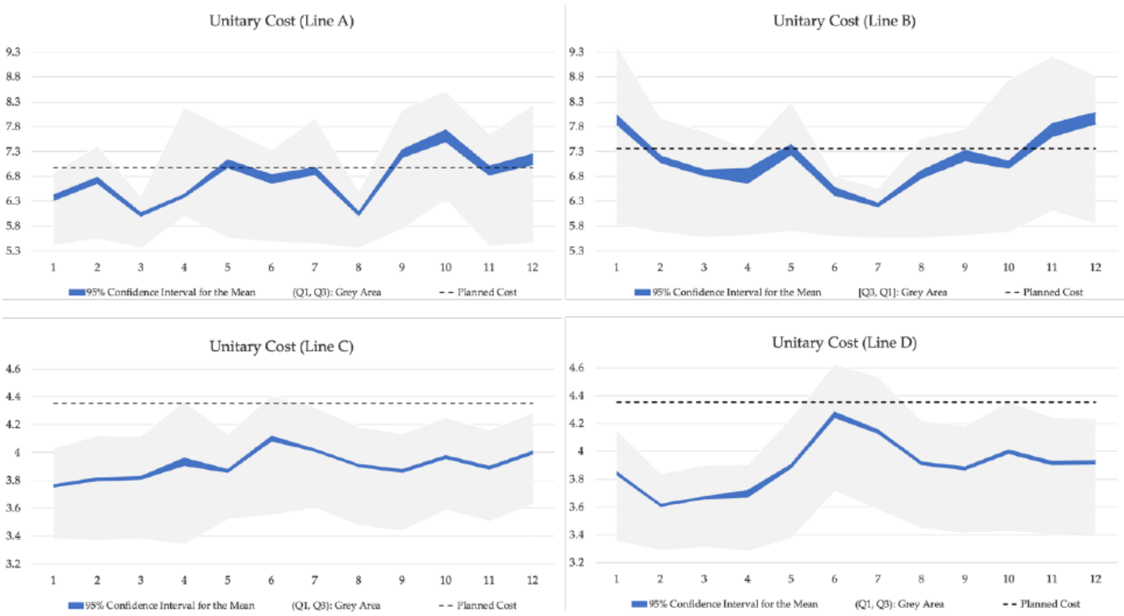

Figure 2.6 Stochastic cost analysis, per line. Values in euros for each of the 12 weeks

Starting with lines A and B, we see that costs vary over time, above or below what had been planned. Lines C and D tend to present their mean costs lower than planned. The upper limits (i.e., the 3rd quartile values) for the Small Lines are almost twice as high as for the Big Lines. In the Big Lines, values vary between 3.3 and 4.7 euros, while in the Small Lines they can reach maximum values around 9.4 euros.

Furthermore, in most cases, the planned and expected cost value is above the value of Q3, that is, the planning presupposed obtaining a higher cost than the reality. This is not necessarily positive because could represent excessive pressure in the product development and quotation stages.

Considering all lines combined, the cost of the product did not exceed the planned cost, globally. The risk of higher costs given by the values related to the third quartile is not significant and it is higher in the last weeks. Nevertheless, the average cost has been increasing consistently in alternating weeks of increases and decreases in cost, as we can see in Figure 2.7.

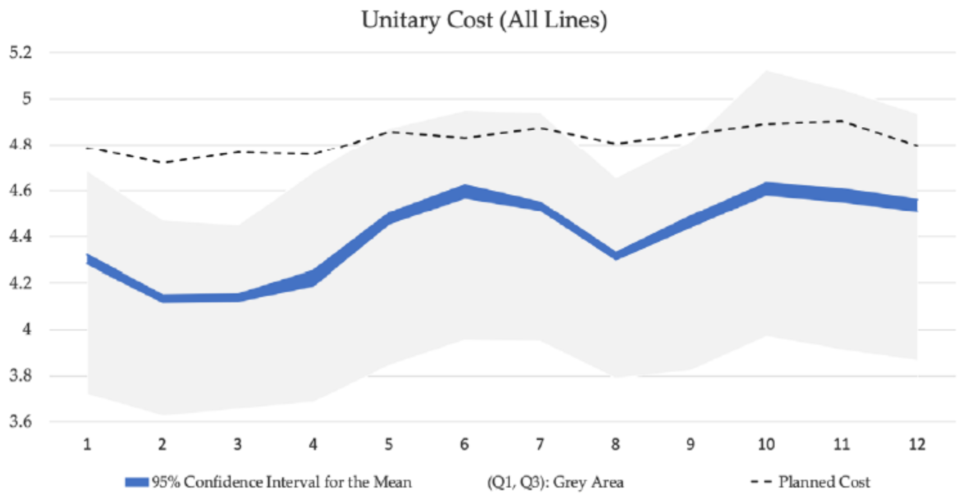

Figure 2.7 Stochastic product cost analysis. Values in euros for each of the 12 weeks

The Wilcoxon sign rank test was used for the validation and analysis of the proposed methodology and the computed costs. The test was used to assess whether the calculated values of costs present significant variations over the weeks in relation to the planned/standard cost.

Table 2.11 Results of the test for the comparison of planned and real costs

Line	p-value	Difference between real and planned cost
A	0.167	3.9%
B	0.129	-3.7%
C	<0.001	-10.0%
D	<0.001	-15.4%
All lines	<0.001	-10.1%

As we can see from the results obtained, Lines A and B present p- values of 0.167 and 0.130, respectively. We are led to conclude that these do not present significant differences between the planned value and the observed values. It means a greater tendency for the cost values, actual and planned, to come closer together. If we notice, in absolute terms, the median values of the observed costs present a variation of around 4% compared to the planned values.

As regards Lines C and D, the test values obtained p-values are <0.001. In other words, there are significant differences between the values planned and those obtained. In real terms, this means that these lines are more sensitive to having cost values significantly different from the planned ones due to different levels of productivity, planning efficiency and process variability. The median values are indeed different, with the actual values being 10% and 15% lower than planned.

The results obtained are consistent with the company's situation that allowed the product cost to be lower than planned. Namely because big lines work very efficiently and much below the target cycle time. The majority of production is done in big lines (around 15000 parts per week compared to 3000 parts per week in small lines). Also, the amount of equipment is more in big lines, so more parts are produced in parallel. Thus, the cost of the product tends to be much lower than the standard cost decided by the finance department.

In general, the Wilcoxon test shows that the average real costs, considering all lines, tend to be lower than the planned ones. On the other hand, this difference is significant for big Lines (Lines C and D). Small Lines (Lines A and B) do not present significant differences between planned and observed values. However, given the influence of the big Lines on total production, performing the same test and under the same conditions, we observe that there are significant differences between planned product costs and actual observed costs. In this case, we obtained a test value of <0.001 with actual median values 10% lower than planned, considering all production lines.

We are led to infer that big Lines considerably influence the variation of the final cost of the product and that the explanation for real costs lower than planned lies in the operating conditions of these lines. The Wilcoxon test allows us to deduce that the median values, in total, are very different (47341.423 and 42580.857, comparing planned and observed, respectively), and that the actual average values are about 10% lower than the planned value.

As we can see, the Wilcoxon test reinforces the scientific validity of the significance of the cost variation, if any, and furthermore shows us that the proposed methodology is able to present and describe that same variation and its effects on the product cost.

In this analysis, the workstation that represents the bottleneck, per line, was considered for the computation of the cycle time, allowing a view of the minimum time required to produce an article in the production line. Further work can be developed to support a much more detailed analysis considering all the stations that compose the line and, consequently, of the respective specific costs associated with each workstation. With the methodology adopted here, it will be interesting to notice the variation, along the same lines, of the different costs by workstation and by line and its consequent variations over time. Moreover, an intensive outliers' analysis must be performed since it increases the variability in the process, and it is important to understand these occurrences to reduce them.

By following this methodology, it is possible to know the real-time cost of the product which can facilitate controlling the cost of the product in a timely manner. The manager can decide better the allocation of quantities to each line as each line can provide different margins of profit based on the quantities produced and variability of process time in each assembly line. Investment in new equipment can also be verified by this approach, whether it will be profitable or not. Thus, investment appraisal exercises will also benefit from the use of a stochastic approach in product cost calculations.

2.6 Final remarks

The proposed methodology was applied in a real context and the main remarks to take into consideration is that the presence of outliers can lead to wrong misperceptions. That is, in Table 2.1 it was not clear that there are differences between the lines A and B, considering the mean values. Thereby, when the outliers were removed, that is more perceptible (Table 2.4) and, in terms of variability, the range was halved. Note that the removal of the outliers was done for those caused by external factors. Despite the hypothesis test having the same conclusions with (Table 2.2) and without outliers (Table 2.5), the range between the lower and upper bound also decreased considerably. This means that part of the variability was removed.

The lines use different resources and also have different capacity levels. Small line B has the highest costs and, incidentally, the highest fare. The big lines, C and D, have very similar costs and, by the way, the same fare. Furthermore, it is also expected that costs have their own variability, which is not studied here. The combination of cycle times variability, costs variability and demand or planning variability will turn the model too complex. But all these variabilities should be taken into consideration.

Statistics, namely the study of averages and respective variations in values by lines, allows us to have a broad view of production time and, consequently, respective costs. It is possible to verify that, depending on the type of line, these values differ, allowing inferences about different trends and variation intervals for the average cycle times and cost.

With the presentation of the confidence intervals for the mean, it is possible to obtain a notion of the expected variation, in terms of production times, helping to forecast costs. Comparing actual and planned results, we confirm the existence of cost variability, which in some cases may be above and in others below the expected value. In other words, the clarity of the uncertainty in cost forecasting is expressed. With a confidence interval of 95%, it is possible to predict that big lines have a greater tendency to lower-than-expected mean cost values, compared to small lines.

The study presented also allowed us to conclude on the importance of studying outlier elements, and when these are extracted, we are led to a clearer analysis, closer to what we may consider common (without major variations and differences in values). In other words, it became easier to see that the non-consideration of aberrant elements helps in predicting results within something that can be considered as expected.

With the approach presented here, an important contribution is made to estimate something that is uncertain and that can vary greatly over time. Even though the cycle time considered only referring to the bottleneck station, the results show the existence of divergences in costs in the two types of lines.

Despite the results obtained, it should be noted that the cycle time considered was related only to the bottleneck station. In other words, although other workstations are operating simultaneously, the specific production time of those workstations was not considered. For further work, it may be considered the times of all workstations that make up the line and understand how this influences the final cost. In addition, it will be important to extend the study to the relationship between cycle times, if any can be found. It opens doors to the analysis of the average cycle times and respective variations of each workstation in each of the lines, to estimate and predict the respective costs with high confidence.

In addition, outliers can be analyzed carefully to understand which and what types of lines are more sensitive to large variations in time cycles. That is, if lines with differing cycle times lead to different costs and/or large cost variations.

The main obstacles and difficulties faced in the implementation of this methodology were related to the access and integration of financial and production information. This process takes some time as data must be extensively collected and various tests must be performed on it. The presentation and visualization of the results in an automated and simplified manner must also be improved, which will contribute to the routinization and institutionalization of the entire process. Business intelligence and analytics tools are particularly useful in this context. Company managers are experienced with such tools (e.g., Tableau software) but there is still needed a better integration among databases, reporting models, routines, and procedures.

Chapter 3

3. A Stochastic costing model based on TDABC for manufacturing environments

3.1 Introduction

Most manufacturing companies compete in global markets, in which competitors are offering similar products with almost the same quality and price. Due to these reasons the focus then turns towards the cost of the product. To survive in the market, it is mandatory to provide a competitive price of the product, which in turn asks for the use of accurate costing systems (Barth et al., 2008). Costing systems help to control the cost of the product and provide accurate information to managers (Cooper & Kaplan, 1988). There can be severe effects if the costing is not done well, namely, the company might lose money if the selling price of the product is underpriced or might not achieve proper sales target if the product is overpriced. Therefore, it is important to have a proper costing system and good control over the cost (Pember & Lemon, 2015).

Traditionally, cost drivers have been based on volume measures, such as quantities of product, labor time spent in production, etc., which, nowadays, do not correctly reflect the consumption of resources in manufacturing companies. So, in the 1980s, Kaplan and Cooper proposed a new methodology named activity-based costing (ABC) which links resources to the activities needed to produce products and services and other relevant cost objects. Its goal is to accurately allocate overhead costs using the best cost drivers of the process. The allocation of the cost of the product is based on the related activities and it contributes to solve the problem of increasing indirect costs and its influence on the product cost structure. ABC models can support a more accurate costing of the relevant cost objects (e.g., products, job orders, clients, processes) leveraging the profitability of the company (Kaplan & Anderson, 2007).

The successful implementations showed that ABC models had a clear impact on the performance of the company (Banker et al., 2008). However, many companies faced some difficulties in the implementation and use of these models and eventually abandoned their ABC projects (Gosselin, 1997). As the collection of the information can be very time consuming, the implementation becomes costly. Furthermore, ABC ignores the potential of unused capacity. To overcome these drawbacks, Kaplan and Andersen developed a new methodology which was termed as time-driven activity-based costing (TDABC) (Kaplan & Anderson, 2007).

TDABC focuses on two main factors, cost, and time, and it also measures the available capacity of the resources and considers how it can be included in the cost of the product. With the use of TDABC it is possible to identify the resources used, their costs and the effective used capacity of the different resources and this allows to calculate the cost per unit more accurately (Hoozée et al., 2012). For each resource, the cost is allocated using two sets of estimation 1) Capacity cost rate and 2) Process time. The cost of all the resources used, divided by the time available to perform the task gives the capacity cost rate. The process time reflects the time needed for each cost object (Hoozée & Hansen, 2018). With the increasing complexity in production process, TDABC can help by representing complex processes using time equations (Gervais et al., 2010). All the time required for sub-activities affecting the costs are incorporated by the time equations in TDABC which helps to reduce the complexity that is typically not well approached in ABC models (Zamrud et al., 2020).

The costing of the product in practice (methods and tools used by companies) and the concepts that support such practice can be improved. In particular, the costing models based on activities, which inherently follow a deterministic approach can be extended to include the risk and uncertainty related to the production and its main cost objects (i.e., products, activities, processes).

Uncertainty occurs in the costing because of the insufficient information or knowledge which affects the predictions as they will differ from the reality (Oberkampf et al., 1999). To deal with uncertainty, fuzzy logic can be included in the cost model providing more accurate results when compared to the standard models (Esmalifalak et al., 2015). Particularly, a stochastic approach can be used as an extension of TDABC models which can accommodate variability that prevails during the production process (Afonso et al., 2021). Stochastic models can be built using the probability distribution of possible results by allowing random variables in one or more inputs over time. The random variables are based on the uncertainty and variability found in the historical data (Ioannou et al., 2017).

Afonso et al., 2018 compared ABC, TDABC and the Production Effort Unit (UEP) method (developed and used in Brazil but based on the GP concept proposed by the French engineer Georges Perrin in the 1940s), in order to use them to manage the unused capacity properly. To facilitate the comparative analysis, an equivalence was made among the operational workstations (in UEP), activities as defined in the ABC and resources as considered in TDABC. The results of TDABC differ significantly when compared to the other two costing methods. To a large extent, this difference is explained by the way as each method deal with idleness. While in TDABC only the capacity cost effectively used is allocated to products, in the UEP and ABC methods this does not occur. Under ABC and the UEP method, the monthly expenses are fully allocated to the production of the period (measured in UEP), according to the concept of full absorption cost, regardless of whether or not the production potential of the company is used. In TDABC, the production cost is based on the number of minutes actually consumed by the capacity used, which tends to be smaller than the installed practical capacity, which entails factory idleness.

In France, the GP method gradually evolved and changed the focus on production costs to include other relevant costs of the products sold, namely, selling and customer-related costs, and in 1995, changed its name to UVA (Unité de Valeur Ajoutée). The UVA method asks for less human and IT resources than ABC (Levant & de La Villarmois, 2011).

For Gervais et al. (2010), the UVA method as other equivalent unit methods, face a number of challenges, namely, consistency and reliability. But, it should be noted that these technical issues are neither specific to the UVA method, nor to the TDABC; indeed, they arise regardless of the costing method used. For example, the measurement of unused capacity can be made using standards, as in the standard costing method. Both, UVA and TDABC may be considered advanced equivalence methods (Levant & Zimnovitch, 2013). But the focus on capacity costs may led to restore hourly burden rates, turn the methods simple instead of advanced.

Due to the existence of variability, product costing should be approached stochastically. Thus, in this chapter, a stochastic TDABC is proposed to extend the model proposed by Kaplan and Anderson. It was extended, namely, to take into consideration the variability in the computation of the cost of the activities. This stochastic cost model is supported on a set of time equations structured around a cost hierarchy that highlights and differentiates the workstation specific cost, cost induced by the line bottleneck, planned unused capacity cost and unplanned unused capacity cost.

This cost model is an effective tool for companies to gain a better understanding of their production costs and profitability. By dividing the cost into optimistic and pessimistic costs, companies can identify which products and production lines are more profitable and which ones are driving costs. The breakdown of costs by products and production lines provides a transparent and accessible view of the cost structure, allowing for better decision making. The model also allows companies to simulate the impact of changes on costs, enabling them to determine the most effective approach. This approach helps to optimize the manufacturing process by identifying areas for improvement and increasing efficiency. The production manager can gain valuable insights into each manufacturing process, which can help to identify areas for improvement and reduce costs.

The model developed has been applied in a first-tier supplier of the automotive industry which produces products like driver control systems, sensor systems and control units for cars and other vehicles. The company also develops and produces innovative solutions for electric and hybrid vehicles. There are multiple assembly lines capable of producing multiple types of products from the same product family. Specifically, the studied product can be produced in four different assembly lines. Each assembly line consists of four similar workstations, characterized by a huge variability among them. Due to the existence of such variability, the cost of the product can change considerably. The proposed model facilitates the understanding of the stochastic range of the cost, along with that it will also be possible to compare the cost of the product across different assembly lines which can facilitate to take managerial decisions such as which assembly line is more profitable, or to decide about the production quantity by production line in order to optimize the cost of the product. Using this model, it is possible to calculate the value-added cost and non-value-added cost by operation, and machine.

The remainder of the chapter is organized as follows, a literature review on activity-based cost models is presented, particularly, TDABC and stochastic cost models. The literature review is followed by the model development section where the novel stochastic model is presented. The application of the model to a real situation showcases the stochastic range of product costs and a comparison of the cost was made when it is produced in different assembly lines over a period. Also, the cost of different products from the same product family that are produced in the same assembly line, was computed, and compared. The last section highlights the conclusions and opportunities for further research.

3.2 Literature review

3.2.1 Activity-based cost models

In the last few decades, ABC models have offered a solution for allocating more accurately indirect costs using adequate cost drivers that reflect the consumption of resources by activities and of those by cost objects (Barth et al., 2008). Activity-based costing links resources to the activities needed to produce the relevant cost objects (e.g., products). There are two main stages in activity-based costing, in the first one the resources are allocated to activities by resource drivers, and in the second stage, the cost of each activity is allocated to the cost objects by activity drivers. The goal of activity-based costing is to allocate more accurately overhead costs using the best cost drivers of the process. It has been assumed that the resources are consumed by the activities which are needed to produce the product (Kaplan & Anderson, 2007). The allocation of the cost to the product by means of allocation bases related to activities results in more accurate costs of the cost objects (Cooper & Kaplan, 1988)). ABC was developed to solve the problem of the increasing indirect costs of the product and its influence on the product cost structure which are derived from the process of industrialization and automation.

Nevertheless, there are some drawbacks of implementing ABC models namely because it can be very time-consuming resulting in costly to implement and use. Also, it does not contribute to understand and optimize the cost of unused capacity (Gilbert, 2007). Hence, a lot of companies avoided implementing or dropped using ABC models. To resolve these problems, Kaplan and Anderson proposed the TDABC approach. It considers time as the main driver of capacity and cost measure, helping to understand how much capacity is required for each product along with the capacity cost rate to allocate resource costs to the products and other relevant cost objects (e.g., clients) (Pember & Lemon, 2015). TDABC uses time to drive the cost directly from the resources to the cost objects, moving over from allocating departmental resources to the multiple activities that the department performs. For each department or pool of resources, the cost of resources is directly assigned to cost objects using two sets of estimates: the cost per unit of time and the process time which can be calculated using time equations (Hoozée, 2013). The cost based on time is calculated by dividing the cost of all the resources by the available time or practical capacity of the resources to perform the operation (Hoozée & Hansen, 2018). Because time drivers can reflect conditions of efficiency or standard costs, in TDABC it is not needed to perform regular surveys to know the work time of the different activities which makes the maintenance of TDABC approach easier (Gervais et al., 2010).

In this context, some extended activity-based cost models have been proposed by several authors. Namely, Feature Based Cost Management (Filomena et al., 2011), Efficiency Based Absorption Costing (Benjamin et al., 2009), Fuzzy Activity-Based Costing (Chansaad et al., 2012; Nachtmann & Needy, 2003), Fuzzy Performance Focused Activity based Costing (PFABC) (Sarokolaei et al., 2013), Activity-Based Life-Cycle Costing (Durán et al., 2020).

Feature Based Cost Management (Filomena et al., 2011) proposes a feature-based approach to cost management that aims to improve cost estimation accuracy and better align costs with the features or functionalities of a product. The approach involves identifying the cost drivers of each feature, analyzing their impact on overall costs, and allocating costs accordingly. The paper presents a case study demonstrating the effectiveness of the proposed approach. Efficiency Based Absorption Costing (Benjamin et al., 2009) proposes a modification to the traditional absorption costing method, called Efficiency-Based Absorption Costing (EBAC). EBAC aims to address the limitations of traditional costing methods by incorporating efficiency measures into the cost allocation process. The paper presents a case study comparing EBAC with traditional absorption costing and demonstrates its effectiveness in improving the accuracy of cost allocation.

Fuzzy Performance Focused Activity-based Costing (PFABC) (Sarokolaei et al., 2013) is an extension of Fuzzy Activity-Based Costing that incorporates performance measurement into the cost allocation process. The approach involves identifying the performance metrics that are relevant to each activity and using them to adjust the cost allocations. The paper presents a case study demonstrating the effectiveness of PFABC in improving the accuracy of cost estimation. Activity-Based Life-Cycle Costing (Durán et al., 2020) Activity-Based Life-Cycle Costing (AB-LCC) is an approach that combines Activity-Based Costing with Life-Cycle Costing to provide a more comprehensive view of costs over a product's life cycle. The approach involves identifying the activities that drive costs at each stage of the product's life cycle and allocating costs accordingly. The paper presents a case study demonstrating the effectiveness of AB-LCC in providing more accurate cost estimates and supporting decision-making.

The extended TDABC model has been applied in several manufacturing contexts, including in the automotive industry, where it has been used to determine the costs associated with various manufacturing processes such as stamping, welding, and assembly. Other manufacturing industries where the extended TDABC model has been applied include the aerospace industry, where it has been used to analyze the costs associated with aircraft assembly processes. Khademian et al., (2018) extended TDABC model in a manufacturing context. The study analyzed the costs associated with the production of a wind turbine generator in Iran. The results of the study showed that the extended TDABC model provided a more accurate and reliable estimation of the manufacturing costs compared to traditional costing methods. The study also identified several areas where cost reductions could be made by optimizing manufacturing processes.

Ganorkar et al. (2019) implemented a TDABC using the Maynard operation sequence technique (MOST) to improve productivity and profitability in manufacturing. The proposed approach was suitable for fast implementation and for validating the existing costing system. It also helped to identify the opportunities for productivity and profitability improvement. Furthermore, the analysis supported on the TDABC model allowed to identify the bottleneck and using the MOST analysis time required is less than the actual time required.

Vedernikova et al. (2020) made an analysis by implementing the TDABC model to the assembly industry and compared the cost of the product with the traditional costing system. It was found that the cost was lower in the TDABC model, and it was found that not all resources were working at the maximum efficiency. The authors also highlighted the use of TDABC to understand the production process in detail which can facilitate process optimization, cost reduction, and the planning of the master budget.

But there are also some problems that arise relating to the use of TDABC. For instance, documenting the flow can be a time-consuming process and uncomfortable for the staff (Park et al., 2019). Furthermore, there are problems in terms of measurement of the error, it relies on homogeneous and repetitive activities, TDABC can also be biased by the information. Also, the time equation can be hard to design due to the complexity of the process, it asks for structured information systems to supply the cost model and it also might imply the need for constant review. There is also a high probability that error occurs mainly in measuring the time as there are large volumes of data that are needed to estimate time equations. TDABC demands structured information systems and robust databases. Large companies have powerful Enterprise Resource Planning (ERP) systems, and the data is updated periodically. Nevertheless, in small and medium-sized companies, the process of data recording and analysis is much more complicated and time consuming. Vedernikova et al., (2020) suggest that TDABC has some drawbacks in the analysis of strategic and support processes due to irregularity and low standardization. It also confirms the need of the ERP system to implement TDABC so more detailed and accurate information can be obtained.

Thus, measurement of errors, subjectivity and dependence on an assumed homogeneity are the most cited problems related to the use of TDABC (Sobu & Wu, 2012). They still reflect issues and problems which were found previously in the ABC model. For example, it is understood that both ABC and TDABC must be extended and improved to deal with the complexity that comes along with production and business processes, particularly, the uncertainty and variability that prevails during production which can have a great impact on the cost of the product.

3.2.2 Stochastic cost models

According to Anderson et al. (1985), a stochastic model is a group of random variables structured in a particular mathematical set and linked to set members. By incorporating the random variable into one or more inputs throughout time, it can be utilized to calculate the probability distribution of the potential outcomes. The variability in the historical data for a certain period of analysis utilizing time series methods serves as the foundation for the random variables. The distribution curve's most likely estimate is the mode of the curve, commonly known as the probability density function(Ioannou et al., 2017).

By providing a more complete information about the production process, the uncertainty in the estimation of the cost can be reduced. Several reasons like variation in the cycle time of the production resources, variability in the demand of the product, production planning and scheduling, and other unexpected and unpredictable events previals the uncertainty and variability in production costs. Also the missing information, errors in the measurement, complexity of the information, etc. Contribute for the variability in the cost of the product (Zimmermann, 2000). In order to estimate product costs under variability and uncertainty, a sensitivity analysis can be performed. Probabilistic methods can provide more accurate and reliable information, but they require a significant amount of data and sophisticated statisitcs than sensitvity analysis (Datta & Roy, 2010).

A stochastic approach enables to set up a projection model which can further help to analyze product costs in the assembly line along with all the activities and processes used to get each specific product. Presenting costs in the form of distributions can showcase the possible range of product costs from a reasonable set of constraints and assumptioms. Ioannou et al., (2020) performed a Monte Carlo simulation to get the probability distributions which allow the estimation of the probabilities of exceeding a set of thresholds and determining a confidence interval. It was also stressed on the importance of the statistical modelling of the stochastic variables so it can contribute to reduce variability and uncertainty. It can also facilitate a better decision-making process.

Kottas & Lau, (1981) presented a heuristic procedure to design the production lines with stochastic task times. The explicit target was to reduce labor costs and incompletion cost, for that the first stage of the procedure uses a probabilistic procedure to create assembly line designs. The second procedure was used to evaluate operation costs generated line design and to identify which is more economical. Tests with a computerized procedure using randomly generated test problem showcased that the procedure could produce very good designs of line with modest time. Sarin & Erel, (1990) developed a cost model for the single-model stochastic assembly line to solve the problem of reducing labor costs in each workstation of the assembly line along with the costs which arises from the activities not finished within the prescribed cycle time. It is demonstrated in a dynamic programming procedure which is implemented using a bounding strategy to curtail storage and computational requirements. The solutions obtained were compared with those obtained using the procedure of Kottas & Lau, (1981).

Wang et al., (2020) developed a stochastic model for additive manufacturing processes. The model used a normal distribution to represent the variability in labor productivity and a lognormal distribution to represent the variability in material costs. The results showed that the stochastic model provided more accurate cost estimates compared to a deterministic model. Tayyab et al., (2020) developed a model to estimate the cost of electricity in sustainable manufacturing. The model used a lognormal distribution to represent the variability in the cost of electricity. The results showed that the stochastic model provided a more realistic estimate of the cost of electricity and its variability compared to a deterministic model.

The cost model presented by (Vrat & Virani, 1976) for balancing the stochastic assembly line in which the incompletion tasks are handled based on modular assembly concept, is similar to the approach proposed by Kottas and Lau in 1973, and it incorporates the activity time being greater than the cycle time which brings the structural variation in the cost of the product. The assembly line was simulated by using the model proposed to estimate process-storage capacity for the physical layout and design of the assembly line.

Furthermore, Jha (1992) developed a mathematical model for stochastic cost optimization. Jha mentions that to get an exact solution, a two-stage stochastic geometric program needs to be solved which needs a significant computational effort and it is tedious to perform. But managers are concerned about making the decision as it can be based on the probable lower and upper limit of the stochastic cost function. It also facilitates estimating the probable cost range and calculating the exact expected cost. The probability level on the lower bound of cost can was calculated using the theory of error propagation. To find the precise cost under real-world constraints, a decomposition algorithm was used.

Anderson & Merwe, (2021) observed stochastic variation in additive manufacturing and proposed a method of TDABC integrated with digital to optimize the use of capacity and time available for production. In the case of stochastic variation, the time can be allocated to each and every element with a certain confidence level. The model proposed by (Anderson & Merwe, 2021)uses a digit twin, based on the statistical data available and stochastic variation, to predict the time required for each process element.

Stochastic cost models assume that manufacturing costs follow a normal distribution. However, recent research has shown that this assumption may not be accurate in all cases. An extension of the Stochastic model is the Bayesian methods which is used to estimate the parameters of the cost distribution. Bayesian methods allow for the incorporation of prior knowledge about the distribution of costs, which can improve the accuracy of the model. Bayesian stochastic cost modeling methods have been applied to a variety of manufacturing contexts, including semiconductor manufacturing, aerospace manufacturing, and pharmaceutical manufacturing (Lee & Chen, 2012).

3.3 Model development

In the manufacturing process, there are a variety of products which require multiple operations to transform them from raw material to the final product. So, it is important to calculate the time and cost required for these operations which can be done using the TDABC model and time equations. Time equations are an important building block of TDABC (Hoozée & Bruggeman, 2010), resulting in a powerful tool for both operational and strategic management decision making. Particularly, in organizations characterized by complex production and business processes and those operating in very dynamic, competitive, and uncertain markets.

Using time equations, complex activities and processes can be easily modelled, making the costing process much easier, accurate and cheaper. The simplicity of the model results from the fact that it needs only two parameters which are the capacity cost rate and the cost of used capacity performed during the manufacturing of the product (Kaplan & Anderson, 2007). Generically, the time consumed by the event E in the activity A can be expressed as a function of the different characteristics called time drivers (Bruggeman et al., 2005), as presented through Equation 2.

$$T_{EA} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \ldots + \beta_p * X_p \tag{1}$$

T_{EA} - Time required for executing the event E in terms of activity A

β_0- Constant amount of time for activity A

β_1 - Time required per unit of time driver 1

β_p - Time required per unit of time driver p

X_1 - Time driver 1

X_p - Time driver p

p - Number of drivers related to the activity A

Furthermore, the cost of used and unused capacity is based on the capacity cost rate and used/unused capacity.

Capacity Cost Rate = Cost of capacity supplied / Practical capacity of resources supplied

Unused Capacity = Practical capacity − Time spent in production

Cost of Unused Capacity = Unused Capacity × Capacity cost rate

In a production system, the activities can be represented by the work done by the workstations which have a specific capacity and capacity cost rate. Time equations can be designed for each workstation considering both the cycle time of the workstation and the cycle time of the line given by the bottleneck. Each workstation will have its capacity cost rate considering specific costs (e.g., depreciation, direct labor costs, specific maintenance costs, energy consumed) and allocated indirect and general costs. In one specific production line, all workstations will have the same practical capacity given by the available or planned production time measured typically in available production minutes or seconds in manufacturing companies.

Multiplying the different elements of the time equation by the capacity cost rate we have what we can call a "cost equation" (Afonso & Santana, 2016). Using dummy variables, cost equation can give information on the cost per workstation considering both the workstation´s specific cycle time and the bottleneck.

For a specific product what matters is the cycle time of the production line given by the bottleneck workstation which defines the speed that products can be produced and how possible is to fulfil the takt time. But each product will use the workstations differently and the analysis of the cost per workstation is also important to identify the opportunity and impact of improvements if there are alternatives to the existing equipment, if different technologies can be used, etc.

In such a model, the impact on the costs of changes in the bottleneck can be calculated and the costs more directly related to each workstation can be distinguished. This understanding of the different layers of the conversion cost (i.e., workstation specific, line specific, general costs) and (planned and unplanned) unused capacity costs, is important to support continuous improvement and the optimization of the processes. The optimization of the production line will impact on the bottleneck and the optimization of the operations will impact on the different times of each workstation: cycle time, balancing of the line and idle capacity. The analysis can be from the product's perspective but also from the workstation or subprocess perspective if costs are computed for several products and other relevant cost objects.

Thus, a general formula for both time and cost were developed that considers the cycle time cost by workstation, the bottleneck of the production line, the planned and real units produced, and the respective cost rates.

$$\mathbf{T_{i,p}} = \beta_i \times X_{i,1} + (\beta_b - \beta_i) \times X_{i,2} + (\beta_L - \beta_b) \times X_{i,3} + \left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,4} + \varepsilon_i \qquad (2)$$

$\beta_i \times X_{i,1}$: Workstation cycle time

$(\beta_b - \beta_i) \times X_{i,2}$: Bottleneck time

$(\beta_L - \beta_b) \times X_{i,3}$:Unused planned time (flexibility)

$\left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,4}$: Unused unplanned time

$$\mathbf{C_p} = \sum_i^n \beta_i \times X_{i,1} \times \delta_i + \sum_i^n (\beta_b - \beta_i) \times X_{i,2} \times \delta_i + \sum_i^n (\beta_L - \beta_b) \times X_{i,3} \times \delta_i + \sum_i^n \left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,4} \times \delta_i + \varepsilon_p \qquad (3)$$

$\sum_i^n \beta_i \times X_{i,1} \times \delta_i$: Workstation specific cost

$\sum_i^n (\beta_b - \beta_i) \times X_{i,2} \times \delta_i$: Production cost

$\sum_i^n (\beta_L - \beta_b) \times X_{i,3} \times \delta_i$: Unused planned cost

$\sum_i^n \left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,4} \times \delta_i$: Unused unplanned cost

$T_{i,p}$ – product p production time per unit or total units (respectively, X = 1 or X = units produced) at workstation i

C_p – product p cost per unit or total units (respectively, X = 1 or X = units produced)

δ_i - cost rate of the workstation i

$X_{i,j}$ – 1 or 0 if time/cost component j is to be included or not, respectively

$X_{i,1}$ – related to the workstation

$X_{i,2}$ – related to the production line

$X_{i,3}$ – related to the planned flexibility/unused capacity of the production line

$X_{i,4}$ – related to the unplanned unused capacity

β_i – average cycle time of the workstation I

β_L – takt time

β_b – bottleneck time of the production line

$(\beta_b - \beta_i)$ – difference between the bottleneck time and the cycle time of the workstation

$(\beta_L - \beta_b)$ – difference between the takt time and the bottleneck time of the production line

Q_1 – real quantities

Q_2 – planned quantities

ε_i – residual and error measurement time

ε_p – residual and error measurement cost

Equation 3 presents the novel equation the time is bifurcated based on the cycle time of the workstation, bottleneck of the assembly line, planned unused capacity and unplanned unused capacity time. To calculate the cost for each of the components, the respective cost rate has to be multiplied by each component. Equation 4 sums up all the costs of the available workstations in the assembly line. The four components of the time and cost equations are important elements of this model and are described in detail below.

Workstation specific costs

This component basically represents the cycle time of each workstation. When this cycle time is further multiplied by the workstation cost rate it results in workstation specific costs.

Production line costs

In all the assembly lines there is always a bottleneck, so this component represents the cost incurred by the time other resources are waiting due to the bottleneck. This component calculates the cost as the difference between the bottleneck and the cycle time of the workstation which is then multiplied by the cost rates. Also, when the bottleneck workstation is considered this cost component will be 0.

Flexibility or unused planned capacity costs

Due to fluctuations in the demand from the market or the customer, while designing the line there is the possibility to invest in additional capacity in order to have flexibility in production to react in case of an increase in the demand from the customer or changes in scheduling and production planning. So, on usual days the line is not working at full capacity. Hence, there is a cost for that, and this component considers that cost by calculating the difference between the takt time and the planned bottleneck time, which when multiplied by the cost rate gives us the cost of flexibility or the unused planned capacity cost.

Unplanned unused capacity

There is a planned value for the quantities that need to be produced. If the real quantities are less than the planned ones, there will be additional unused time with non-value added by the workstations since the production line is installed with the purpose of producing the planned quantities required by the client and not less. It might happen due to various factors such as breakdowns of machines, insufficient supply of raw material or exceptional situations like disruptions in the supply chains.

Stochastic variables

β_i represents the cycle time of the workstation which will change with each product produced. β_b which represents the bottleneck workstation also changes. Sometimes the value of the bottleneck changes or the bottleneck workstation might also change. ε_k is constant for a particular activity, but it is a random variable within the activities carried out, as they change stochastically from activity to activity. These variables give the range of probable values capable enough to make a probability distribution, so they are considered to be stochastic variables.

They represent a novel cost hierarchy for production costs composed by these different layers of cost allowing different cost calculations depending the perspective (e.g., focused on the specific cost of the workstation, including the costs induced by the bottleneck, with or without product and structure or sustaining-facilities costs. Thus, this cost hierarchy can complement and extend the one proposed by (Robin Cooper, 1990, 1995) unit and batch-level, and product and facility-sustaining.

3.4 Application of the model

There are multiple assembly lines capable of producing multiple types of products from the same product family. Specifically, the studied product can be produced in four different assembly lines. Each assembly line consists of four similar workstations, characterized by a huge variability among them. Using equation 4 the cost for each component can be calculated. The total cost of each workstation including all the four dimensions of the cost as can be observed in Table 3.1 that represents the cost for day 1 of month 1 in line A.

Table 3.1 Component cost by workstation in line A

WS number	Specific cost	Line cost	Flexibility cost	Unplanned unused capacity cost	Total cost
1	0.29 €	0.07 €	0.21 €	0.13 €	0.71 €
2	0.45 €	- €	0.29 €	0.16 €	0.90 €
3	0.37 €	0.03 €	0.29 €	0.27 €	0.97 €
4	0.39 €	0.06 €	0.28 €	0.16 €	0.91 €

Workstation 2 is the bottleneck so the line cost of that workstation will be 0. So, each component of the cost can be obtained along with the total cost of the workstation. The cost obtained consists of all the fixed and variable costs involved but only variable costs can be considered, or overhead costs can be partially allocated. Using this method, the cost of the entire assembly line can be known and compared during various weeks to understand the variability that persists in production. Table 3.2 showcases the total cost of the assembly line (combining all the 4 cost dimensions for 4 different workstations). The analysis was made over 4 different months.

Table 3.2 Monthly comparison of total cost of line A

Month	Specific cost	Line cost	Flexibility cost	Unplanned unused capacity cost	Total cost
1	2.21 €	0.35 €	2.20 €	1.24 €	6.00 €
2	1.96 €	0.19 €	1.50 €	1.10 €	4.76 €
3	2.34 €	0.46 €	2.52 €	1.396 €	6.72 €
4	2.04 €	0.27 €	1.58 €	1.122 €	5.01 €

It can be observed in Table 3.2 that total cost varies each month. It is clear that there exists variability in the cycle time of the different workstations which results in different costs in the four months into consideration. It can be observed that in month 2 the total cost is at its minimum level compared to the other months. It signifies that cycle times and bottlenecks were lower in month 2, which resulted in a lower cost. Whereas month 3 has the highest total cost among all studied months, which resulted from higher cycle times and bottleneck. As cycle time is stochastic in nature, further analysis was performed on it to understand the variability involved.

This approach was applied to the 4 different production lines which were used to produce the same product. A comparison of the product cost for each line was made over 5 months. Figure 3.1 provides the aggregated cost of the product across the four production lines for 5 months. The variation in the cost can be seen evidently.

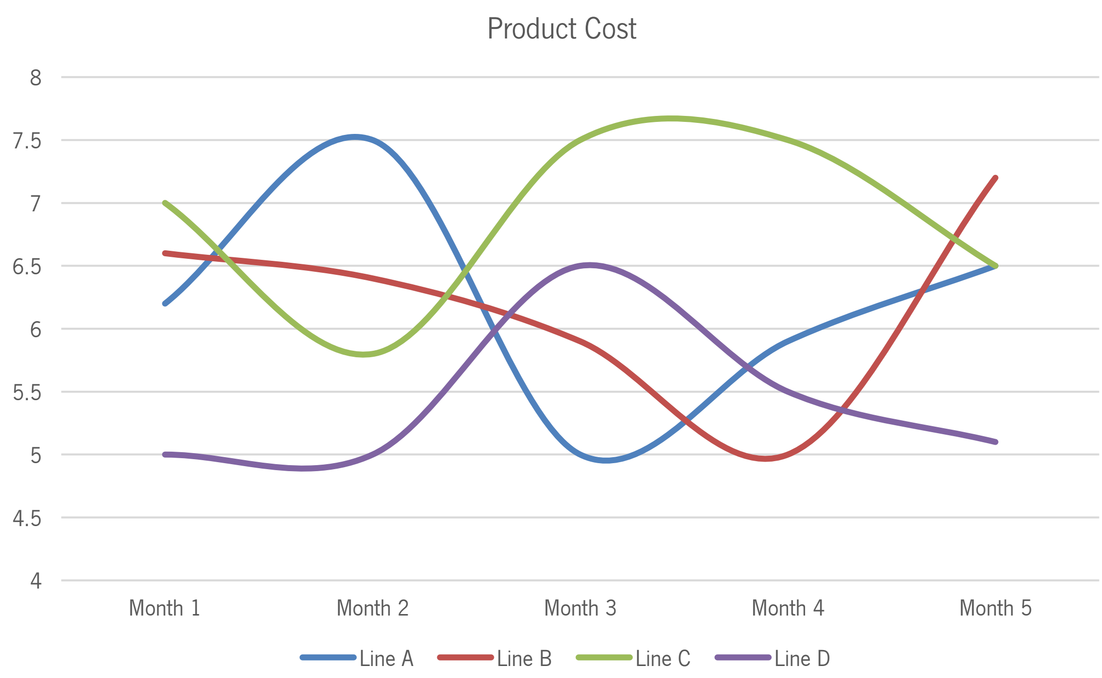

Figure 3.1 Monthly product cost across different production lines

In this case we have a significant monthly cost variability between production lines. The product unitary cost varies considerably (by line, between approximately 5 and 7,5 euros) throughout the months studied across all production lines.

Table 3.3 showcases the standard deviation of average total cycle time of all workstations. Based on those values the range of the cost can be obtained by multiplying it with the tariff cost.

Table 3.3 Cycle time, standard deviation, and interval of values

Month	Total Cycle time (s)	Standard deviation	Lower limit cycle time	Upper limit cycle time	Lower limit cost (€)	Upper limit cost (€)
1	109	19.17	89.83	128.17	4.67	6.66
2	117	23.49	93.51	140.49	4.86	7.31
3	103	13.83	89.17	116.83	4.64	6.08
4	110	13.08	96.92	123.08	5.04	6.40
5	115	21.45	93.55	136.45	4.86	7.10

After calculating the lower and upper limit for all workstations and multiplying it with the tariff enables us to obtain the range of cost which are considered as optimistic and pessimist cost of the product. Using this information, it can be easy for managers to understand which week was more profitable and it can also help to identify the risks involved if the target cycle time is not achieved. Also, the impact of changes in external conditions can be assessed or estimated. This is important to manage the different production lines and evaluate how they perform under different internal and external conditions.

Furthermore, the model was implemented on different products produced in the same line to understand the variability of the cost. The cost of the products was then compared across the period of 5 months. During these 5 months, the analysis was done on a monthly basis and the range of cost of product for each month was recorded, which can be observed below.

Table 3.4 Cost comparison of different products

Product	Month 1		Month 2		Month 3		Month 4		Month 5	
	Lower limit cost	Upper limit cost	Lower limit cost	Upper limit cost	Lower limit cost	Upper limit cost	Lower limit cost	Upper limit cost	Lower limit cost	Upper limit cost

Product A	4.9	5.6	6.7	8.2	5.9	6.3	7.0	7.4	6.2	6.7
Product B	6.2	6.8	5.1	5.5	6.8	7.2	4.8	5.5	6.9	7.4
Product C	7.8	8.4	5.8	6.3	8.8	9.2	7.1	7.9	7.8	8.3
Product D	4.2	4.7	4.9	5.4	4.1	4.6	4.1	4.5	5.0	5.6

These results demonstrate that the different cost drivers (e.g., quantities produced, cycle times) impact significantly on product cost and that there are opportunities to optimize the process and reduce costs. This variability can be explained by a combination of factors internal and external to the company. Some of these factors might be interrelated (e.g., market demand and production planning) and others are independent (e.g., cycle times and bottlenecks in the different production lines). Understanding the behavior of these factors is fundamental for better manufacturing management and control. The variability studied in this chapter is based on cycle time and bottleneck variations which are largely caused due to reasons such as time required for product setup on the machine, time to load/unload product on the machine, operator walking time between the machines etc. This affects the cycle time to a large scale, so overseeing some of these factors can have a very bad effect on the product costing. The stochastic analysis of production cycle times is fundamental to include variability and risk within costing systems. Besides the variability in the production processes, we can also have variability caused by changes in the demand and variability in the value of the resources used. Process variability is particularly relevant in costing systems and for optimization purposes.

Furthermore, to visualize the cost clearly and with transparency, a dynamic dashboard was developed. It helps to understand the breakdown of the costs into Its component. It makes It possible to filter through the dates, products, assembly line, and the product family to get the costs accordingly. Real production data was used in the development and implementation of the dashboard which can be used for the computation and analysis of product costs; as well as, for the design of more effective cost management practices and cost reduction and optimization plans. The real production data refers to real quantities and real cycle times of the machines used in the production lines in a certain period of time. The extraction of the data is the first step. The information related to the quantities and real cycle time for a certain product that passes in X lines and Y workstation is in Hadoop cluster. To extract the exact information from Hadoop that matches the cost model it is necessary to respect some conditions and to follow some steps. Thereby, to facilitate this process all the steps were written in an accessible file that functions as a link between Hadoop data and the cost model. The architecture of the dashboard can be observed in Figure 3.2.

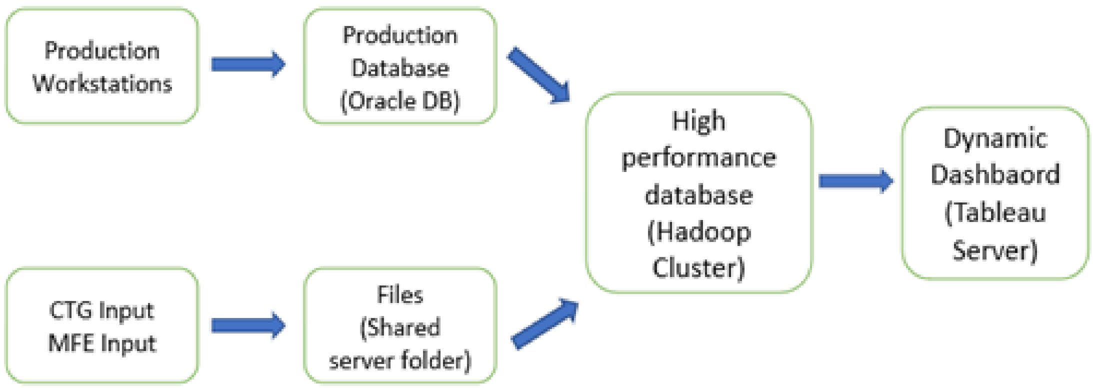

Figure 3.2 Dashboard architecture

The figure below shows the product cost along with the bifurcation among all the four components of the costs. Figure 3.3 shows the dashboard view.

Figure 3.3 Dashboard view for cost for product A

Figure 3.3 shows the cost component which includes specific cost, line cost, flexibility cost, and unplanned unused capacity cost. By using the filters on the left panel, it is possible to change the period of the search, product, as well as the assembly line.

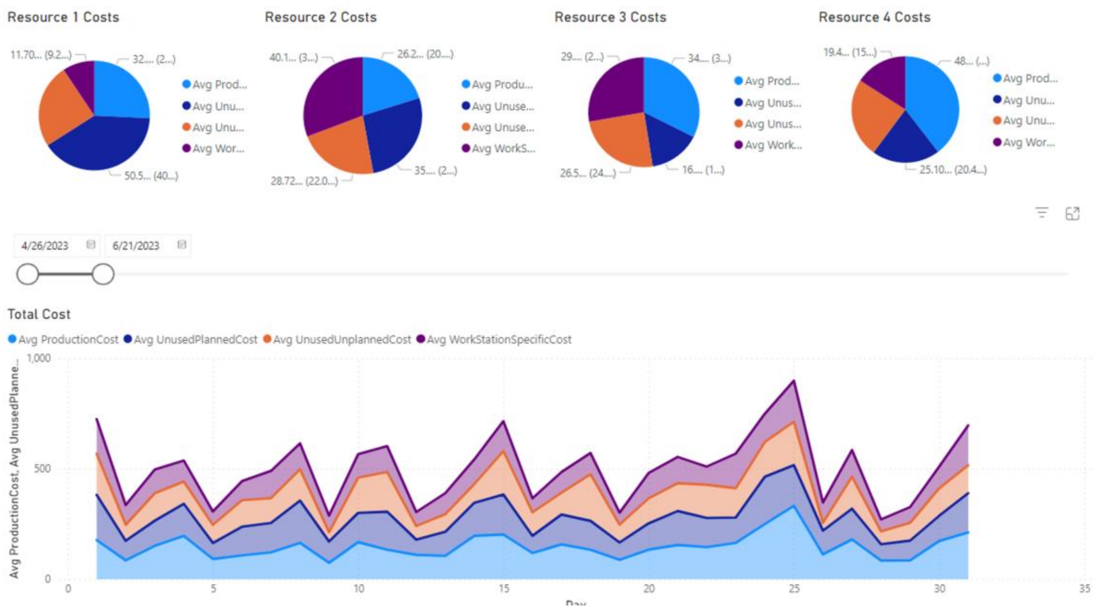

Figure 3.4 Dashboard view of resource cost

In figure 3.4, it can be observed the cost for different resources. With the use of this dashboard a clear visual can be provided about the product cost and variability can be observed in a transparent manner.

3.5 Final remarks

In today's competitive environment, manufacturing organizations are always seeking to improve their productivity and profitability. For this reason, it is crucial for manufacturers to have a detailed understanding of the cost of producing their products. Traditional cost accounting methods have been used for many years, but they may not provide an accurate representation of the real cost of production. The traditional costing method uses direct labor, direct material, and overhead cost allocation to calculate the product cost. However, it does not take into account the variability of the quantities produced, which affects the cost to a great extent, as required time and bottleneck may vary. Thus, a stochastic costing approach has been developed based on activity-based cost models.

The use of stochastic costing approaches, particularly activity-based cost models, has become increasingly popular in manufacturing systems. In such systems, the variability of the quantities produced can have a significant impact on costs. This is due to the fact that the required cycle time and bottleneck may vary, resulting in a range of possible costs rather than a single value. Therefore, it is important to calculate a complete range of possible costs to provide a better understanding of the workstation-specific cost, bottleneck cost, unused capacity cost, and unplanned unused capacity cost.

By using the equations developed in this chapter, it is possible to obtain past, real-time, or predictive costs of a product, allowing companies to recognize which products and production lines are more profitable and which are the drivers of costs and profitability. Moreover, this cost model enables organizations to visualize and understand the cost breakdown of their products and production lines in a transparent and accessible way. The outputs provided by the model can help to acquire a new cost breakdown of the product, which is more relevant for decision making and grants an intelligible understanding and visualization of the cost from different perspectives.

The model also allows us to simulate the impact of certain changes on costs and identify the most valuable approach to adopt. By breaking down the cost into four components, we can distinguish the value-added and non-value-added costs of a product. For instance, it is possible to visualize the impact of the bottleneck workstation on product cost and what occurs when the company is producing more or less than the planned quantities.

The production manager can use this methodology to gain clearer insights into each manufacturing process, identifying value-added and non-value-added activities in the process. This can open opportunities to optimize the process as each process is minutely observed. By using machine learning techniques, the cost of the product for each assembly line can be predicted, benefiting the scheduling of production planning accordingly to achieve maximum possible profit by incurring the minimum achievable cost of the product.

Furthermore, the stochastic costing approach can be extended to other areas of business such as service operations, healthcare, and transportation, where it can help to identify the most efficient processes and improve operational efficiency. For example, in healthcare, it can help to determine the cost of treating patients with specific illnesses, enabling healthcare providers to allocate their resources more effectively.

The use of stochastic costing can also facilitate budgeting and resource allocation decisions. By providing a range of possible costs, the organization can make more informed decisions about budgeting and resource allocation, as they can consider both the optimistic and pessimistic scenarios. It can also help in the identification of the most profitable products, enabling the company to allocate their resources accordingly.

Additionally, the approach can help in identifying and addressing inefficiencies in the production process, which can ultimately lead to cost savings. By identifying the non-value-added activities in the process, the organization can find ways to eliminate them or reduce their impact, leading to improved efficiency and reduced costs.

Furthermore, this methodology can be used to support decision-making related to outsourcing or insourcing. By comparing the costs of producing a product in-house versus outsourcing it, the organization can determine the most cost-effective option. This can help in making informed decisions that balance costs and revenue.

The stochastic costing approach can also help organizations to evaluate and compare the costs of different products. By breaking down the cost of each product, the organization can identify the most profitable products and focus their resources on them. It can also help in identifying the factors that drive the cost of a product, enabling the organization to optimize these factors to reduce costs.

Chapter 4

4 A Regression Approach for Industrial Production Forecasting

4.1 Introduction

Companies are very interested in forecasting production costs. To help companies in planning and decision-making, new analytical tools have been emerging to enable the transformation and representation of the data, namely mathematical modeling. Good forecasting models will, above all, allow a better planning and control of the production. Thus, prescriptive analysis emerges as an opportunity, as it learns from the past to suggest options that can be used in decision-making and risk reduction. Predictive analysis allows estimating and predicting the value of a variable, such as quantity and cost. It involves operations, mathematical and statistical techniques that allow studying and discovering trends, predicting failures (or anomalous situations) and variations that tend to exist in production systems. Thus, for a better understanding of behaviors over time, mathematical models that can describe present data and help in future prediction have emerged and been implemented in several companies. The idea is to assist with decision-making, preventing, and being prepared to deal with serious problems that are predictable and may have a high probability of occurring. Nowadays, companies face higher levels of volatility, uncertainty, complexity, and ambiguity (VUCA). Thus, important decision-making in the industry such as introducing new products into the market, continuation, or termination of the existing products, must be supported by reliable information based on industrial cost and profitability. No less important is the contribution of this type of analysis to a consequent risk assessment associated with the behavior and trend of events.

There are several ways to build predictive models. Some popular methods to build predictive models are:

1. Linear regression: A linear regression model is used to predict a continuous variable based on one or more predictors. It assumes a linear relationship between the dependent and independent variables.

2. Decision tree: Decision trees are used to create models that predict a target variable based on a set of input variables. Models are built by recursively partitioning the data into subsets based on the input variables.

3. Random Forest: Random Forest is an ensemble learning technique that creates multiple decision trees and combines their results to make predictions. It is commonly used for classification and regression problems.

4. Support Vector Machine: A support vector machine (SVM) is a machine learning algorithm used for classification and regression problems. It tries to find the best boundaries separating the data into different

classes. SVM builds models that can model non-linear relationships between dependents and one or more independent variables.

5. Artificial Neural Networks: Artificial Neural Networks (ANNs) are a set of algorithms that model the structure and function of the human brain. They are used for various tasks such as image recognition, speech recognition, and natural language processing.

6. Deep learning: Deep learning is a subset of artificial neural networks that use multiple interconnected layers of neurons to learn and predict. It is often used for tasks such as image and speech recognition.

Linear regression is a statistical technique commonly used for forecasting that helps understand the relationship between two variables. Independent variables are used to predict dependent variables. The goal of linear regression is to find the best-fit line representing the relationship between these variables. This can be used to predict future outcomes. Advantages of linear regression in forecasting are as follows:

1. Simple and straightforward: Linear regression is a relatively simple statistical technique that is easy to understand and requires only basic knowledge of mathematics and statistics. It is widely used because it can be applied to a wide range of problems with minimal assumptions and gives immediate results.

2. Provide a clear relationship: Linear regression provides a clear relationship between dependent and independent variables and is easy to visualize and interpret. This makes it an ideal technique for forecasting as it helps identify the most important variables in predicting future outcomes.

3. The model can be interpreted: Linear regression models are very easy to interpret and provide a clear understanding of the relationships between the analyzed variables. This makes the model easier to interpret and the results of the analysis can be easily communicated to stakeholders.

4. Provides Accurate Forecasts: A linear regression model can provide accurate predictions when the underlying assumptions are correct. Applied to problems with linear relationships between variables, it is very effective at predicting future outcomes.

Prediction together with Artificial Intelligence (AI) techniques help in the analysis of variations and risks associated with them, namely, in situations where forecasting is complex due to the uncertainty and instability that surrounds the company. The work presented here aims to contribute to the design of forecast models, namely, to predict the quantities to be produced. Multiple regression models allow approximations to the real data, even under high degrees of unpredictability. In this work, a set of forecasting models were developed and implemented, using linear regression techniques, namely, multiple linear regression. These are models that, given a set of variables as input, calculate a value for the output variable, which is the variable to be predicted. Data from September 2020 to June 2021 of a first-tier supplier of the automotive industry was used considering two different training periods, in two groups of production lines of a specific product.

For computing the models, two distinct time periods were considered, corresponding to the average values of machines cycle times in twelve consecutive weeks. They were tested for the subsequent periods. The assumptions underlying the creation of a multiple linear regression model were analysed. In addition, measures of goodness of fit of the model were also analysed, namely, R^2, and Mean Absolute Error (MAE). To test the robustness of the models, the percentage errors of the observed values in relation to the predicted ones were analysed. On the other hand, based on the 95% confidence intervals for the model coefficients, the risk associated with the respective component was evaluated. Next section presents some studies on industrial contexts using from traditional statistical techniques to artificial intelligence approaches.

4.2 Literature review

In general terms, among the more traditional techniques found in the literature are those related to autoregressive models (AR), moving average (MA), simple exponential smoothing (SES), and Autoregressive Integrated Moving Average (ARIMA). These are techniques that are based on time series and use the independent variable masking time to predict other variables. Autoregressive is related with autocorrelation, that is, correlation of certain past periods with the current period. Autoregressive time series are commonly seen as an approximation to univariate time series modelling. The moving average, in turn, is an average calculated for a certain period. Regression errors result from a linear combination of error terms that occurred at past times. However, other approaches have emerged with potential for the development of forecasting techniques. These arise to, together with specialists, contribute to the identification of weaknesses and behavioural deviations in the business. Thus, allowing to take, in advance, decisions adapted to the changes and circumstances. Among others, they allow estimating production costs and volumes, forecasting sales and profit.

Galli et al. (2021) use machine learning in conjunction with stochastic optimization to minimize drug stocks. They present a set of scenarios and over these a stochastic optimization approach to provide good decisions, both in quantity and quality.

Sengewald & Lackes (2022) use the Bayesian technique to learn a stochastically optimal sourcing strategy directly from quote data. They did not focus on making great predictions but on making optimal decisions. They present a significant improvement in the costs of the purchase and acquisition process. The model turns out to be also more robust to forecast errors.

Bertsimas & Kallus (2020) combine machine learning, operational research and management science for decision-making. They focused on a conditional stochastic optimization problem, with data related to costs/revenues and on other associated data. They end up showing that the developed methods are applicable and computationally tractable in several situations, even when the data are not independent and identically distributed. They extended the technique to cases of some decision variables, such as price, that may affect the uncertainty of the future and whose causes are unknown. They also developed a prescriptive coefficient P to measure the prescriptive content of the data and operational effectiveness. In the manipulated data, the authors found an improvement in the order of 88%, according to the P coefficient.

Lily Koops (2020) presents a prescribing solution based on the probabilistic approach to cost-benefit analysis and the definition of relevant metrics. She combines a wiener model, to repair imperfections, and simulations by the Monte Carlo method, to support the decision. The probabilistic approach helps to detect the best decision options for greater profit, in addition to the associated risk and potential cost. In addition, they allow to reduce the uncertainty of the results. According to the author, this approach can give the industry guidelines to identify and optimize business cases for prescriptive maintenance, pointing to the sources of valuable data or information and, therefore, justify investments.

Poornima & Pushpalatha (2020) present a technique to improve decision-making support in big data, prescriptive analytics, which offers an improved advance in predicting consequences and their outcome. The best result is achieved using optimization techniques in prescriptive analysis that identify uncertainties in making better decisions. Since optimization improves the effectiveness of prescriptive analytics with varied applications.

Devriendt et al. (2018) provide a structured and detailed literature search on elevation modelling, identifying, and contrasting various groups of approaches. In addition, evaluation metrics for evaluating the performance of elevation models are reviewed. Random forests are among the best performing techniques in terms of Qini and Gini measures, although considerable variability in performance is observed between the various datasets of the experiments. In addition, elevation models are often observed to be unstable and exhibit strong variability in terms of performance. Furthermore, it appears that the available evaluation metrics do not provide an intuitively understandable indication of the actual use and performance of a model. Specifically, existing evaluation metrics do not make it easy to compare elevation models and predictive models and evaluate performance at an arbitrary cut or across the full spectrum of potential cuts. In conclusion, we highlight the instability of uplift models and the need for an application-oriented approach to evaluating uplift models as key topics for future research.

Berk et al., (2019) used real data simulation, also to create a planning model for human resources. They applied robust optimization and maximum flexibility which, according to the authors, leads to increased profit. Seyedan & Mafakheri (2020) analysed algorithms related to K-neighbors, neural networks, regression, and vector support machines. They mention that, although neural networks and regression analysis are the most common approaches used, they can be improved by optimization or simulations. Chen et al. (2021) forecast the costs and duration of a project through the added value management theory, through the application of the Monte Carlo method, to simulate large amounts of data, together with statistics. Punia et al. (2020) propose a machine learning model, based on quantile regression, to determine order quantities. They suggest a distribution-free and Z-constraint-oriented solution approach. Something important to mention, and according to the author, despite all the techniques that can be applied, external variables play a significant role in the accuracy of a model and in the estimates and, these, in turn, have a strong impact on the cost of inventory management.

Additive models emerge as a proposal for estimating multiple non-parametric regression models and can be seen as a generalization of linear regression models (Hastie & Tibshirani, 1990). Additive models are considered semi-parametric, in which the response variable (variable to predict) can assume different probability distributions (Gaussian, among others), and the data can be binary or categorical, ordinal, or non-ordinal [2].

An additive multiple linear regression model, with p>1 regressors, is defined by:

$$Y = \alpha + \sum_{d=1}^{p} m_d \,(x_d) + \varepsilon \qquad (1)$$

The ε errors can verify the classic validation assumptions of a linear regression model (Heteroscedasticity and Normality). The functions m_d (designated by additive components) are univariate functions and result from smoothing processes, being assumed that they verify the identification condition:

$$E\big(m_d(x_d)\big) = 0, \forall d \qquad (2)$$

The condition above allows us to state that the expectation of the variable to be predicted, Y, is zero, that is, E(Y) = 0. Functions resulting from an additive model resemble linear regression coefficients. Hence, the assumptions associated with linear regression must be analysed. Namely, to analyse the significance of the correlation between the variables.

It is advisable to select, include or remove variables in stages until you realize which variables are important for the model. An analysis of the adjusted functions is recommended, knowing that all of them are necessary for the construction of the whole (Hastie & Tibshirani, 1990).

Gradient Boosting is a technique for building additive regression models in which a simple parameterized function (baseline learning) is sequentially adjusted to the "pseudo-residuals" by means of the least squares algorithm, in each iteration. Pseudo-residuals are the minimized gradient of the loss function, in relation to the model values at each point of the training data evaluated in a current step of the algorithm (Friedman, 2002).

The stochastic gradient boosting incorporates randomness in the procedures and execution of the respective algorithm. Briefly, in each iteration, a subsample of the training data is randomly extracted (without replacement). This subsample is then used, instead of the full sample, to adjust baseline learning and calculate the model update for the current iteration.

It is shown that both the accuracy of the approximation and the speed of execution of the gradient increase can be substantially improved by incorporating randomization in the procedure (Friedman, 2002).

Let $\{y_i, X_i\}_i^N$ be the entire training data sample and $\{\pi(i)\}_i^N$ be a random permutation of the integers $\{1, \dots, N\}$. Then a random subsample of size $\widetilde{N} < N$ is given by $\{y_{\pi(i)}, X_{\pi(i)}\}_1^{\widetilde{N}}$. L is terminal node regression tree; Ψ *is* general loss criterion. One starts with an initial guess F_0(x), and then for m=1, 2,..,M

The stochastic gradient boosting algorithm is then (Friedman, 2002)

$$F_0(x) = argmin_\gamma \sum_{i=1}^{N} \Psi(y_i, \gamma).$$

$For\ m = 1\ to\ M\ do$:

1. $\{\pi(i)\}_1^N = randperm\{i\}_1^N$

2. $\tilde{y}_{\pi(i)m} = -\left[\frac{\delta\Psi(y_{\pi(i)}, F(x_{\pi(i)}))}{\delta F(x_{\pi(i)})}\right]_{F(x)=F_{m-1}(x)}, i = 1, \tilde{N}$

3. $\{R_{lm}\}_1^L = L - terminal\ node\ tree\left(\{\tilde{y}_{\pi(i)m}, x_{\pi(i)}\}_1^{\tilde{N}}\right)$

4. $\gamma_{lm} = argmin_{\gamma} \sum_{x\_\pi(i)\epsilon R_{lm}} \Psi\left(y_{\pi(i)}, F_{m-1}(x_{\pi(i)}) + \gamma\right)$

5. $F_m(x) = F_{m-1}(x) + v.\gamma_{lm} 1(x\epsilon R_{lm})$

$endFor.$

Model trees are data-driven and can be classified as modular models. To train them, a set of data may be split into subsets corresponding to a particular sub-process to be modelled, and then each module is trained on these non-intersecting subsets. After the model is trained, each new input vector is first classified to one of the regions for which the modules were trained, and then only one module is run to produce the prediction (Haykin, 1998).

Classes of such methods employing consecutive progressive splits are referred to as "trees", for example decision trees, regression trees (Solomatine & M. B. Siek, 2003)and M5 model trees (Quinlan, 1992).

The model tree is a machine learning technique that works with continuous numerical values. The M5 model tree is one of the algorithms available in some software, such as Weka 1 through the M5P function (Y. Wang & Ian H. Witten, 1997).

According to (Quinlan, 1992), the M5 model tree technique combines any regular decision tree model with the probability of linear regression at the decision tree's leaf nodes.

The M5 model tree is built in two stages. In first is a traditional decision tree and, in second, a linear regression function is generated. One regression tree is constructed using the decision tree induction procedure. The standard deviation at each node will be determined to assess the predicted reduction in error for the splitting criterion. This node splitting in M5 will continue until there are very few instances left. After constructing the normal regression tree, internal sub nodes are pruned and replaced with the regression plane rather than constant values. Pruning estimates the predicted error at each node (Quinlan, 1992).

Let S be an input set, splitting divides it into smaller subsets, $S_1, S_2, \ldots, S_n$. It is a recursive technique in which each sub-split is sub-split into offspring, and the process continues until there are very few instances remaining. At each internal node, this M5 employs a greedy approach to error minimization, with standard deviation reduction (SDR) determined one node at a time and provided by equation 3.

A SDR is calculated at each node for splitting, and then cost complexity pruning (CCP), illustrated below, is applied at each leaf node to remove areas that may not be beneficial for the final tree, lowering the number of rules (Landwehr et al., 2005).

$$SDR = \frac{SD(T) - \sum SD(T_i)|T_i|}{T} \tag{3}$$

$$CCP = \frac{err(prune(T,t),S) - err(T,S)}{|leaves(T)| - |leaves(prune(T,t))|} \tag{4}$$

Smoothing is also applied after pruning on the created Model Tree, which is used to prevent the harsh discontinuities of the sub trees. Flattening the sharp nodes of nearby models is a typical way to enhance prediction accuracy (Sheikh Amir et al., 2022; Sidiq et al., 2019).

The first step, when building a model tree, is to compute the standard deviation of the target values of cases in S. Unless S contains very few cases or their values vary only slightly, S is split on the outcomes of a test. Potential test is evaluated by determining the subset of cases associated with each outcome.

Let S_i the subset of cases that have the $i-th$ outcome of the potential test, and $sd(S_i)$ the standard deviation of the values in S_i as a measure of error, the expected reduction in error test can be written as

$$\varDelta error = sd(S) - \sum_i \frac{|S_i|}{S} sd(S_i) \tag{5}$$

After examining all possible tests, M5 chooses one that maximises this expected error reduction.

The M5 tree algorithm can translate better results using pruning and smoothing. When the pruning option is turned on, every non-leaf node in the tree is examined, starting near the bottom. M5 selects for the final model for the node the simplified linear model above or the subtree of the model, according to which presents the smallest estimated error. If the linear model is selected, the subtree at the respective node is pruned to one leaf.

Smoothing has most effect on a case when the models along the path predict very different values and when some models were constructed from few training cases. When smoothing is turned on, if the value of a case is predicted by a model tree, the value given by the model at the appropriate leaf is adjusted to reflect the predicted values at the nodes along the path from the root to that leaf (Quinlan, 1992). The smoothed predicted value of M5 is copied from the leaf to the root according to the following definition:

$$PV(S) = \frac{n_i \times PV(S_i) + k \times M(S)}{n_i + k} \qquad (6)$$

where k is a smoothing constant.

4.3 Model development & description

One of the biggest problems in building an additive regression model is the selection of regressors. To do so, let's start by looking at the correlation matrix and the respective tests of significance. Remember that in a correlation test, the null hypothesis to be tested is that there is no correlation (or is null) between two variables. As an alternative hypothesis, it is assumed that there is some (cor)relation between the variables. Normally, as in this work, the statistical test is performed at a significance level of 5% (0.05). In the case under study, we are dealing with qualitative variables ("Year", "Month", "Week", "Bottleneck value", "Real qty", "Amort Specific cost", "Specific other cost", and "total cost") and ordinal ("Line", "Bottleneck Workstation"). As such, Spearman's coefficient is adjusted for the study of correlations between variables. The results are shown in the following table.

Table 4.1 Correlation between Variables

Variable	Variable	Correlation value	p value
Year	Month		<0.005
Year	Line	0.00	1.00
Year	Week	-0.04	0.72
Year	Bottleneck value	-0.19	0.06

Year	Bottleneck Workstation	0.48	<0.005
Year	Real qty	-0.01	0.89
Year	Amort Specific cost	-0.20	0.05
Year	Specific other cost	-0.16	0.11
Year	total cost	0.19	0.05
Year	Index	-0.10	0.35
Month	Line	0.00	1.00
Month	Week	-0.06	0.55
Month	Bottleneck value	0.03	0.75
Month	Bottleneck Workstation	0.08	0.43
Month	Real qty	0.17	0.09
Month	Amort Specific cost	0.23	0.02
Month	Specific other cost	0.23	0.02
Month	total cost	-0.01	0.92
Month	Index	0.32	0.00
Line	Week	0.00	1.00
Line	Bottleneck value	-0.58	<0.005
Line	Bottleneck Workstation	0.07	0.47
Line	Real qty	0.26	0.01
Line	Amort Specific cost	-0.07	0.51
Line	Specific other cost	-0.04	0.71
Line	total cost	-0.33	<0.005
Line	Index	0.00	1.00
Week	Bottleneck value	-0.03	0.75
Week	Bottleneck Workstation	-0.02	0.88
Week	Real qty	0.02	0.88
Week	Amort Specific cost	0.07	0.48
Week	Specific other cost	0.02	0.85
Week	total cost	-0.07	0.51
Week	Index	0.14	0.18

Bottleneck value	Bottleneck Workstation	-0.10	0.31
Bottleneck value	Real qty	-0.68	<0.005
Bottleneck value	Amort Specific cost	-0.26	0.01
Bottleneck value	Specific other cost	-0.26	0.01
Bottleneck value	total cost	0.50	<0.005
Bottleneck value	Index	0.04	0.68
Bottleneck Workstation	Real qty	-0.15	0.13
Bottleneck Workstation	Amort Specific cost	-0.35	<0.005
Bottleneck Workstation	Specific other cost	-0.33	<0.005
Bottleneck Workstation	total cost	0.03	0.78
Bottleneck Workstation	Index	0.25	0.01
Real qty	Amort Specific cost	0.83	<0.005
Real qty	Specific other cost	0.85	<0.005
Real qty	total cost	-0.40	<0.005
Real qty	Index	0.01	0.94
Amort Specific cost	Specific other cost	0.98	<0.005
Amort Specific cost	total cost	-0.21	0.04
Amort Specific cost	Index	0.00	0.98
Specific other cost	total cost	-0.18	0.08
Specific other cost	Index	0.01	0.90
total cost	Index	0.25	0.01

Let's start by building an additive regression model with the "Amort Specific cost", "Specific other cost" and "Bottleneck value" regressors, considering the variable "Real qty" as a predictor (variable to be predicted).

The additive regression model was run using the M5P tree classifier, which implements base routines for generating M5 model trees and rules. The generated trees/rules are pruned, and the predictions are smoothed.

For the creation of the Additive Regression model, the shrinkage rate value is considered, by default (that is, 1.0).

To evaluate the additive regression models, the measures R^2, MAE (Mean absolute error), RMSE (Root mean squared error) and MAPE (Mean absolute percentage error) were considered.

R is the value of the correlation between the variables “Real qty” and “Real qty predicted”, which is obtained by applying the created model. Its value varies between 0 and 1 (since R is between -1 and 1) and the closer to 1, the better the model's goodness of fit.

MAE corresponds to the average of the observed absolute errors, that is, MAE = mean (abs("Real qty"-"Real qty predicted")). Obviously, the smaller its value, the better the approximation of predicted values to observed values.

RMSE corresponds to the standard deviation of the residuals (difference between prediction and observed values). Residuals are a measure in allowing us to assess how far the data points are from the regression line of the relationship between predicted and observed values.

Table 4.2 Evaluation of Additive Regression Model

Variable to predict	**Regressive Variables**	**Measures**
Real qty	Bottleneck value Amort Specific cost Specific other cost	R^2 =0.998 MAE=158.856 Root mean squared error=212.683 Mean absolute percentage error=0.094
Real qty	Bottleneck value Amort Specific cost Specific other cost Bottleneck Workstation	R^2 =1 MAE=56.137 Root mean squared error=87.18 Mean absolute percentage error=0.03
Real qty	Bottleneck value Amort Specific cost Specific other cost Bottleneck Workstation	R^2 =1 MAE=68.183 Root mean squared error=99.329 Mean absolute percentage error=0.046

	Line	
Real qty	Bottleneck value Amort Specific cost Specific other cost Bottleneck Workstation Week	R^2 =1 MAE=73.14 Root mean squared error=106.921 Mean absolute percentage error=0.042
Real qty	Bottleneck value Amort Specific cost Specific other cost Bottleneck Workstation Month	R^2 =1 MAE=73.14 Root mean squared error=99.329 Mean absolute percentage error=0.046

In fact, if we consider the regression variables “Bottleneck value”, “Amort Specific cost”, “Specific other cost” and “Workstation that is bottleneck”, the additive regression presents the best results.

The method presented aims at building an additive regression model, considering “Real qty” as the variable to be predicted. For this, a first selection of the regressive variables is necessary. This was based on the (cor)relation between the variable “Real qty” and each of the others. The results above (see table of correlations) show us that cor(“Real qty”, “Bottleneck value”)=-0.68, cor(“Real qty”, “Amort Specific cost”)=0.83 and cor(“Real qty ”,” Specific other cost”)=0.85 with results of p-value<0.005. This means that the (cor)relation between the referred pairs of variables is significant. To calculate the correlation between pairs of variables, Spearman's coefficient was used, as it is one of the most appropriate in the presence of quantitative and ordinal qualitative variables. The results are shown in the table above.

An additive regression model was applied, based on the M5 tree classifier. Through the generated model, the values estimated by it for the variable “Real qty” are presented. Bearing in mind that the models will be evaluated by comparing the actual values and the predicted values, namely, in terms of the measures of R^2, MAE, RMSE and MAPE. The ideal is to consider the best that presents a higher value of $R^2$and lower possible values in terms of MAE, RMSE and MAPE.

Starting from this initial model, additive models are generated, that is, to the initial model formula with and only the regressive variables "Bottleneck value", "Amort Specific cost" and "Specific other cost" new regressive variables are added, successively, as shown in the table above, Analogously, the predicted values are compared with the observed values of the variable "Actual qty" and the measured values R^2, MAE, RMSE and MAPE are also recorded.

4.4 Result and Discussion

In today's data-driven world, accurate forecasting is crucial for making informed decisions in various industries. Predicting quantities accurately can help organizations optimize resources, improve production efficiency, and meet customer demand effectively. Additive regression models are powerful tools used in predictive analytics, and in this discussion, we explore how they can be employed to forecast production quantities. Specifically, we focus on an iterative approach that enhances the model's performance through the incorporation of key variables.

At first we examine the foundational additive regression model used for forecasting production quantities. This model was based on regressive variables and employed the M5 classifier in conjunction with a decision tree. The initial model achieved impressive results, with an R-squared (R^2) value of 0.998, a mean absolute error (MAE) of approximately 159 units, and a mean absolute percentage error (MAPE) of 0.094. These metrics indicate a high level of accuracy in predicting production quantities.

One of the notable aspects of this study is the iterative nature of model improvement with eagerness to enhance forecasting accuracy, decided to introduce a new variable into the mix: "Bottleneck Workstation" This addition proved to be a game-changer, as it significantly improved the model's performance across multiple evaluation measures.

With the inclusion of the " Bottleneck Workstation " variable, the R^2 value reached an ideal 1, indicating a perfect fit between the model and the data. The MAE plummeted to a mere 56 units, and the MAPE decreased to an impressive 0.03. Additionally, the root mean square error (RMSE), a critical metric in regression analysis, exhibited a considerable decrease when compared to the initial model. This improvement in RMSE underscores the model's increased accuracy in capturing the underlying patterns in the data.

The enhanced additive regression model relies on several key variables, including "Bottleneck," "Amort Specific Cost," and "Specific Other Cost." These variables, when combined with the " Bottleneck Workstation" form a robust set of regressive variables that provide superior forecasting results. The "Bottleneck" variable plays a crucial role in the model, as evidenced by its presence in the variable set. It represents a condition that can have a significant impact on production quantities. By considering the value of this variable, the model gains insight into potential production bottlenecks, allowing for better resource allocation and production planning.

Similarly, "Amort Specific Cost" and "Specific Other Cost" are essential components of the model. These variables provide insights into cost structures and financial considerations that can influence production quantities. By incorporating these variables into the regression model, organizations can make more informed decisions regarding cost management and production optimization.

While the "Bottleneck," "Amort Specific Cost," and "Specific Other Cost" variables have proven to be instrumental in improving forecasting accuracy, it is worth considering the impact of other variables, such as "Line," "Month," and "Week."

In the context of this study, we found that these variables, while not necessarily weak in terms of predictive power, did not contribute to significant improvements in MAE, RMSE, and MAPE. This suggests that, for the specific production quantities being forecasted, the variations in "Line," "Month," and "Week" do not play a substantial role. It's essential to recognize that the relative importance of variables can vary depending on the context of the forecasting problem.

The findings of this study have several important implications for organizations seeking to enhance their production quantity forecasting such as Variable Selection: The choice of regressive variables is critical in building accurate additive regression models. Variables such as "Bottleneck," "Amort Specific Cost," and "Specific Other Cost" have demonstrated their importance in this context and should be considered in similar forecasting endeavors. Iterative Model Improvement: Model improvement should be an iterative process. As demonstrated in this study, continually assessing and enhancing the model with additional variables can lead to substantial improvements in forecasting accuracy. Context Matters: The relative importance of variables can vary depending on the specific forecasting problem. It is essential to evaluate the impact of each variable in the context of the data and the production process. Cost Considerations: The inclusion of cost-related variables, such as "Amort Specific Cost" and "Specific

Other Cost," can provide valuable insights into the financial aspects of production forecasting, aiding in cost-effective decision-making.

In conclusion, the study presented an insightful journey into the world of additive regression models for forecasting production quantities. By employing an iterative approach and incorporating key variables like "Workstation that is bottleneck," "Bottleneck," "Amort Specific Cost," and "Specific Other Cost," we were able to develop a highly accurate forecasting model. This model not only achieved an ideal R^2 value of 1 but also demonstrated significantly improved MAE, RMSE, and MAPE metrics.

The findings highlight the importance of variable selection and continuous model improvement in predictive analytics. Moreover, they emphasize that not all variables carry equal weight in forecasting, and their impact can vary depending on the specific production process. Organizations should consider these insights when developing and refining their forecasting models, ultimately leading to better decision-making and improved operational efficiency.

For future work, more sophisticated models can be developed such as Vector Autoregression (VAR) models that are data-driven and focus on the historical relationships among variables. They use past observations to trace the impact of shocks without relying on explicit theoretical structures. Vector Autoregression models are flexible and can capture complex interactions among variables, making them valuable for short to medium-term forecasting. The future of VAR modeling promises exciting developments driven by advancements in computational power, data availability, and methodological innovations. One significant avenue of progress lies in the integration of machine learning techniques with traditional VAR frameworks. Hybrid approaches combining VAR with deep learning, reinforcement learning, or Bayesian methods hold the potential to enhance forecasting accuracy, capture nonlinear relationships, and handle big data challenges more effectively.

Furthermore, the expansion of VAR models to accommodate high-dimensional data is another frontier. Traditional VAR models are limited in their ability to handle a large number of variables due to the curse of dimensionality. Future work may focus on developing sparse VAR models, leveraging techniques such as regularization or variable selection to address this issue and improve model interpretability.

Other methods such as, Recurrent Neural Networks (RNNs) are well-suited for sequential data forecasting tasks where the order of data points matters, such as time series forecasting. They have loops within their architecture, allowing them to maintain a state or memory of previous inputs. Long Short-

Term Memory (LSTM) and Gated Recurrent Unit (GRU) are popular variants of RNNs that are capable of capturing long-term dependencies and mitigating the vanishing gradient problem, which makes them particularly effective for time series forecasting. Some forecasting tasks may benefit from hybrid architectures that combine multiple deep learning techniques or incorporate domain knowledge into the model. For example, a model may combine Convolutional Neural Networks (CNNs) for feature extraction from input data with LSTM layers for sequence modeling and forecasting. Such hybrid architectures can leverage the strengths of different techniques to improve overall forecasting performance.

Chapter 5

5 Optimization of labour and workstation costs

5.1 Introduction

Many companies are now including human resources as a core component of their business strategies due to the quickly evolving and fiercely competitive environment. Managers are looking for more effective tools to optimize the use and allocation of their available resources among the various services or systems in an effort to maximize or minimize certain functions related to performance and productivity. They are aware that human resources can play a significant role in the success of an organization. The issue arises in a number of real-world contexts, and research papers in the literature demonstrate a persistent interest in the human resource allocation problem and cover several useful applications.

Manufacturing systems are designed and controlled to satisfy the customer's orders and demands in a timely and cost-effectively manner; therefore, it is important to deliver the manufactured products at the exact time that was agreed upon with the customer. This practice encouraged to develop methods to determine the allocation of cross-trained operators to handle more than one machine. Keeping with the high production rate and performance, assigning the correct and optimal number of machines to each operator is essential, it will assure the utilization of the operators and not overload them and keep the machine running with minimum downtime due to changeover or loading/unloading activities. Therefore, knowing how many machines and how many employees are needed to accomplish a specific manufacturing job is vital.

It is not profitable to have more employees attending the production machines (fewer machines assigned to one operator) because this will act as a burden on the company as there will be extensive operator's idle time and will increase the usage/workload on the machines which require extra maintenance over the production period. On the other hand, having fewer employees (more machines assigned to one operator) will result in overloading the operators and decrease their work moral standards, and machines will be idling for the operator availability. Both scenarios will cause a delivery delay. The purpose in both cases is to accomplish the task of the production or the service for all the customers within the time and cost constraints. The number of machines should be higher than the calculated numbers because there is a possibility that some of these machines will fail down.

This chapter focuses on development of the operator cost model considering the activities performed by them on the workstation taking into consideration the variability. Also, traditional optimization model is applied to understand the optimum number of operators and workstation equipment in the assembly line which are further compared with cost.

5.2 Literature review

Production lines managers are working non-stop to maximize profits by increasing yields and reducing the cost of manufacturing at the same time (Tirkel & Rabinowitz, 2014). An important approach to achieve that goal is to determine the optimum number of operators needed to run the production machines. Using the fuzzy logic controller provide an easy tool where management can determine the optimal number of operators to be assigned, and the number of workstations that should be used by the operator in terms of controlling, or operation to reach the production goals (Keren & Hadad, 2016) .

Assigning several workstations to one operator may not increase the system's overall performance (Chien et al., 2014). Although assigning the proper number of workstations to an operator is a critical and non-trivial decision. Trade-offs will take place and manger will need to select tradeoffs the best scenario (Stecke & Aronson, 1985). Too many workstations assigned to one operator may increase overworked operator occurrences, idle workstation, defects and failures, safety, and health problems, and so on. On the other hand, few workstations assignment may cause unnecessary labor costs associated with idle or bored operators (Stecke, 1982).

The operator-workstation assignment affects both workstation and operator utilization and production yields' cost (Haque & Armstrong, 2007; Stafford, 1988). The number of operators assigned to a given number of workstations, and the number of machines that will be controlled by each operator, must both be optimized. Additionally, different objective functions such as minimizing cost, maximizing profit, minimum idle time, and/or minimum overload, may require different assignments. Many papers in the literature dealt with similar problems assigning the operations to workstation/ operators, both in job-shops or flexible manufacturing systems (Chen & Ho, 2005; Park et al., 2014).

The notion of a bottleneck is important in many planning methods. A bottleneck is the group of similar machines that limits the production rate (Johri, 1993). As a result of this, the importance of an efficient machines capacity plan for this tool is required (Mönch et al., 2013). Capacity planning problems appear in many forms and have attracted thousands of research papers. In order to structure this large body of research, comprehensive literature reviews by Costa et al., (2014) and Wu et al., (2005) have been published. As stated by Costa et al., (2014), deterministic models have got much attention in solving capacity allocation models. To illustrate, Huang et al., (2014) considered the problem associated with the decentralized allocation of the finite capacity of a single facility among different business organizations with fuzzy demand. Simultaneously, game theoretic approaches have been widely applied to capacity planning problems in strategic and tactical levels of manufacturing organizations (Renna & Argoneto, 2010) . Renna & Argoneto, (2010) developed a distributed approach, for a network of independent enterprises, able to facilitate the capacity process by using a multi-agent architecture and a cooperative protocol. In another application of game theory concepts in solving capacity allocation problems. Seok & Nof, (2014) proposed an adaptive collaborative demand and capacity sharing (CDCS) protocol based on dynamic contract mechanism. Liu et al., (2015) proposed a model of cloud manufacturing resource service sharing based on the Gale-Shapley algorithm and analyzed it in the context of fluctuating resource service supply and demand.

With regard to machine capacity allocation models in photolithography, one of the earliest studies on both capacity allocation problem and machine capability is the study by (Leachman & Carmon, 1992), in which they defined a production plan by presenting a linear programming (LP) model in order to maximize total profit. A similar production plan with an LP model was also presented by Hung & Cheng, (2002). The former study scrutinized machine process capability constraints by introducing 'alternative machine set' that is defined to represent a capability for a particular operation, and the capacity limitations of these machine sets are indicated by proposed models (i.e., step-separated, workload allocation and direct mix formulations) with the assumption of identical or proportional processing times (Leachman & Carmon, 1992). For LP formulations, the number of decision variables increases due to revisits of products to process areas because of the number of alternative machine types to the power of the number of re-entrant visits (Leachman & Carmon, 1992).

With regard to a latter study, Hung & Cheng, (2002) presented the capacity partition generation procedure (CPGP) in which the uniformity assumption is relaxed in the direct mix formulation provided by Leachman & Carmon, (1992) with the capacity set generation procedure (CSGP). Toktay & Uzsoy, (1998) transformed the capacity allocation problem into a maximum flow network problem for maximizing throughput. Their mathematical formulation includes not only machine capabilities but also tooling and set-up constraints together with integer side constraints. They compared results of the problem by two proposed heuristics, i.e., greedy, and extended heuristics. Toktay & Uzsoy, (1998) decomposed the shift-scheduling problem into two sub-problems which are capacity allocation and lot sequencing, in order to analyze them sequentially. To solve the problem, capacity allocation routine (CAR) was applied by the greedy heuristic defined by Toktay & Uzsoy, (1998), and embedded in a simulation model. Also, two different sets of capabilities (i.e., operation-stepper matrices) were defined as fully flexible and nested sets. That is, the fully flexible set was defined for processing capability of all operations, and the nested set was defined for the processing capability of a partial set of operations. Their simulation experiments included analyses of stepper capabilities, reticle, and setup constraints.

The optimization problem is an important issue that has been studied by many types of research throughout the last decade in various manufacturing fields such as structural design (Yildiz, 2013), cell formation (Anbumalar & Sekar, 2015; Balakrishnan & Cheng, 2007; Brauner & Finke, 2001), U-shaped manufacturing lines (Nakade & Nishiwaki, 2008; Sirovetnukul & Chutima, 2009) and others. The research is related to applying Fuzzy logic to the machine operator allocation problem in the cellular manufacturing setting. Few researchers did utilize fuzzy logic to analyze facility layouts, working conditions, lighting effect on employee behavior, dynamic layout, and even the selection of the layout and machine allocations to estimate the downtime and Fuzzy Inference System (FIS) had been utilized to determine the machine criticality levels for maintenance activities (Kunduracı & Kazanasmaz, 2019; Osuch et al., 2020; Torun & Çetinkaya, 2019; Zha et al., 2020).

5.3 Model development

5.3.1 Value added activities identification model

There are several activities performed during the production. But from a managerial perspective it is important to understand which activities are actually adding value to the product, and which are non-value added. The model developed helps to calculate the value-added time, non-value-added time, and unused unplanned time.

$$\mathbf{T_{i,p}} = \beta_i \times X_{i,1} + (\beta_b - \beta_i) \times X_{i,2} + \left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,3} + \varepsilon_i \qquad (1)$$

$\beta_i \times X_{i,1}$: Value added time

$(\beta_b - \beta_i) \times X_{i,2}$: Non-value-added time

$\left(\frac{Q_2 - Q_1}{Q_1}\right) \times \beta_L \times X_{i,3}$: Unused unplanned time

$T_{i,p}$ – Total time taken by the operator to perform all manual operations at workstation i per unit or total units (respectively, X = 1 or X = units produced)

β_i – Value added time by operator on the workstation i

β_L – takt time for manual operations

$X_{i,j}$ - 1 or 0 if time component j is to be included or not, respectively

Q_1 - real quantities

Q_2 - planned quantities

ε_i- residual and error measurement time

The three components of the time equations are important elements of this model and are described below.

Value added time:

This is related to the time required by the operator to perform certain manual operations, which adds value to the product. By multiplying the labor cost rate by the time required, the costs for value added operations can be obtained.

Non-value-added time:

This represents the cost of all the non-value-added operation costs such as setup, travel time between workstations, allowances etc. These operations are needed for the production, but they don´t add any value to the product so the cost incurred by these operations will be considered as non-value-added costs.

Unplanned unused capacity time:

There is a planned value for the quantities that need to be produced. If the real quantities are less than the planned ones, there will be additional unused time with non-value added by the operators.

In this case there is no component related to planned unused capacity, as the target cycle time of the line is defined by the production manager based on the availability of the operators and it is assumed here that operators are the bottleneck. So, it is important to understand the optimum number of operators required, so an optimization model is implemented to fulfil this requirement and along with that it will also give an overview of on which workstation the operator should work.

5.3.2 Optimization model

5.3.2.1 Labour cost optimization model

The objective of the optimization model was to understand how many operators are required and to minimize the cost of the labor. But there are certain conditions that work done by the operator shouldn´t exceed the target cycle time. The work on each workstation has to be done by just one operator.

Objective function: To calculate and allocate the optimum number of operators to the workstation to achieve target cycle time of assembly line

$$\sum_{i \in Op} \sum_{j \in WS} c_{ij} x_{ij} \qquad (2)$$

where, c_{ij} - time needed on workstation by operator i on workstation j

x_{ij} – allocation of operator i on workstation j

Constraints:

The work in each WS is done by just one Operator

$$\sum_{i \in Op} x_{ij} = 1, \ \ \forall j \in WS \qquad (3)$$

Each Operator can work on zero or more WS

$$\sum_{j \in WS} x_{ij} \geq 0, \ \ \forall i \in Op \qquad (4)$$

Each Operator works no more than the bottleneck

$$\sum_{j \in WS} c_{ij} x_{ij} \leq Bottleneck, \ \ \forall i \in Op \qquad (5)$$

5.3.2.2 Workstation specific cost optimization

Traditionally, the number of workstations required to achieve desired cycle time of the line can be calculated as follows:

$$N_W = \frac{t_p}{t_c} \tag{6}$$

N_W = Number of workstations required

t_p = Time of bottleneck workstation

t_c = target cycle time

Increasing the number of workstations brings the advantage of achieving the target cycle time and reduce the bottleneck time of the assembly line, but it also comes with the disadvantage that having more workstations will result in the higher fixed cost of the resources and the amortization cost which will directly result in higher cost of the product as the cost tariff will be increasing.

So, it is important to find a balance between the required number of workstations, bottleneck, and the cost of the product. Because the focus of the production manager will be to decrease the bottleneck time but also controlling the cost of the product. Using the Stochastic cost model developed in Chapter 3, cost of the product will be calculated which will help us to understand the effect of increased cost tariff. Some iterations of the cost calculation can be made by increasing the number of workstation and effects of that on the cost can be observed. Reducing the bottleneck time will help reduce the cost and increasing the workstation will increase the cost so the optimised scenario must be chosen that how much bottleneck time should be reduced and increase the workstation in order not to have drastic increase in the cost of the product.

5.4 Analysis of results

It is important to understand the optimum allocation of the operator as the variability persists even in their operation time. Also, it is important to understand the bifurcation between the value-added time and the non-value-added time by the operators along with the optimum number of operators required in the assembly line. Table 5.1 mentions the time spent in all the 3 components of Equation 13.

Table 5.1 Time of operator tasks for line A

WS number	Value added time	Non-value-added time	Unplanned unused time	Total time	Standard deviation	Lower limit time	Upper limit time

1	2	3	1	6	1.2	4.8	7.2
2	5	6	2	13	2.5	10.5	15.5
3	5	7	2	14	2.1	11.9	16.1
4	3	6	2	11	1.9	9.1	12.9
5	5	7	3	15	1.5	13.5	16.5
6	4	6	2	12	1.8	10.2	13.8
7	3	4	1	8	1.4	6.6	9.4
8	4	4	2	10	2.1	7.9	12.1
9	3	3	1	7	1.9	5.1	8.9
10	5	5	2	12	2.3	9.7	14.3
11	2	2	1	5	1.4	3.6	6.4
12	3	4	3	10	1.9	8.1	11.9
13	3	4	1	8	1.7	6.3	9.7
14	3	4	2	9	1.5	7.5	10.5
15	2	3	2	7	1.6	5.4	8.6
16	4	8	4	16	2.4	13.6	18.4
17	6	7	5	18	2.6	15.4	20.6
18	2	2	1	5	1.8	3.2	6.8

The target cycle time for the operator was 34 seconds for line A. The total time spent by the operator on 18 workstations is around 186 seconds, as in the model it was defined in the constraint that an operator cannot work more than the target cycle time so based on that we got the solution that there is requirement of 7 operators. Based on the standard deviation considering the worst-case scenario i.e., the upper limit time which totals to 219 seconds is still less than 238 seconds (34 seconds x 7 operators). Hence, even considering the extreme situation 7 operators will be sufficient enough to work on these workstations. Using the optimization model, we can solve the allocation problem to understand what the best possible workstation for each operator is. Solving the problem, results can be seen in table 5.2.

Table 5.2 Operator allocation to workstation for line A

Workstation	**Operator**
WS1	Operator 1

WS2	Operator 1
WS3	Operator 2
WS4	Operator 1
WS5	Operator 2
WS6	Operator 3
WS7	Operator 4
WS8	Operator 3
WS9	Operator 3
WS10	Operator 4
WS11	Operator 5
WS12	Operator 5
WS13	Operator 5
WS14	Operator 5
WS15	Operator 6
WS16	Operator 6
WS17	Operator 7
WS18	Operator 7

As it can be observed in the table 5.2, operator 1 will be working on workstation 1, workstation 2, and workstation 4. Operator 2 will be working on workstation 3 and workstation 5. Operator 3 will be working on workstation 6,8, and 9. Operator 4 on workstation 7 and 10. Operator 5 on workstation 11,12,13, and 14. Operator 6 on workstation 15 and 16. Lastly, operator 7 on workstation 17 and 18. The details on the total allocation for each operator can be found in table 5.3

Table 5.3 Total allocation of each operator for line A

Operator	**Total allocation (s)**	**Allocation (%)**
OP1	30	88.24
OP2	29	85.29
OP3	29	85.29
OP4	20	58.82
OP5	32	94.12

OP6	23	67.65
OP7	23	67.65

The target cycle time for the assembly line was 34 out of which operator 1 will be working for 30s which accounts for 88.24% of allocation. Operator 2 and 3 will be working for 29s which account for 85.29% of the total capacity. Operator 4 has comparatively less allocation of 20s which is 58.82%. Operator 5 will be working for 32s out of 34s with allocation of 94.12%. Operator 6 and 7 will be working for 23s with 67.65% of the allocation. Furthermore, the hourly cost of the operator is 13€/hr so based on the allocation and time spent by the operator it is important to calculate the cost of the labor for each unit. Table 5.4 presents the range of cost based on the tariff.

Table 5.4 Cost of operator tasks for line A

WS number	**Lower limit time**	**Upper limit time**	**Lower limit cost**	**Upper limit cost**
1	4.8	7.2	0.017	0.026
2	10.5	15.5	0.038	0.056
3	11.9	16.1	0.043	0.058
4	9.1	12.9	0.033	0.047
5	13.5	16.5	0.049	0.060
6	10.2	13.8	0.037	0.050
7	6.6	9.4	0.024	0.034
8	7.9	12.1	0.029	0.044
9	5.1	8.9	0.018	0.032
10	9.7	14.3	0.035	0.052
11	3.6	6.4	0.013	0.023
12	8.1	11.9	0.029	0.043
13	6.3	9.7	0.023	0.035
14	7.5	10.5	0.027	0.038
15	5.4	8.6	0.020	0.031
16	13.6	18.4	0.049	0.066

17	15.4	20.6	0.056	0.074
18	3.2	6.8	0.012	0.025

In table 5.4, it can be observed the total labor cost on each workstation per piece of the product which ranges from 0.55€ to 0.79€ per unit which includes all the three components of value-added cost, non-value added cost and unused unplanned.

Similarly, table 5.5 shows the cost components for line B.

Table 5.5 Time of operator tasks on line B

WS number	Value added time	Non-value-added time	Unplanned unused time	Total time	Standard deviation	Lower limit time	Upper limit time
1	2	2	1	5	1.2	3.8	6.2
2	7	8	2	17	2.5	14.5	19.5
3	5	6	1	12	2.1	9.9	14.1
4	2	5	2	9	1.9	7.1	10.9
5	6	7	2	15	1.5	13.5	16.5
6	5	6	3	14	1.8	12.2	15.8
7	3	5	2	10	1.4	8.6	11.4
8	2	4	1	7	2.1	4.9	9.1
9	2	3	1	6	1.9	4.1	7.9
10	5	5	3	13	2.3	10.7	15.3
11	3	4	1	8	1.4	6.6	9.4
12	3	3	1	7	1.9	5.1	8.9
13	2	3	1	6	1.7	4.3	7.7
14	3	5	2	10	1.5	8.5	11.5
15	3	4	2	9	1.6	7.4	10.6
16	4	8	3	15	2.4	12.6	17.4
17	7	8	4	19	2.6	16.4	21.6
18	2	3	2	7	1.8	5.2	8.8

The target cycle time for the operator was 34 seconds for line B. The total time spent by the operator on 18 workstations is around 189 seconds, as in the model it was defined in the constraint that an operator cannot work more than the target cycle time so based on that we got the solution that there is requirement of 6 operators. Solving the problem, results can be seen in table 5.6.

Table 5.6 Operator allocation to workstation for line B

Workstation	Operator
WS1	Operator 1
WS2	Operator 1
WS3	Operator 1
WS4	Operator 2
WS5	Operator 2
WS6	Operator 3
WS7	Operator 3
WS8	Operator 3
WS9	Operator 4
WS10	Operator 4
WS11	Operator 4
WS12	Operator 4
WS13	Operator 5
WS14	Operator 5
WS15	Operator 5
WS16	Operator 6
WS17	Operator 6
WS18	Operator 5

As it can be observed in the table 5.6, operator 1 will be working on workstation 1, workstation 2, and workstation 3. Operator 2 will be working on workstation 4 and workstation 5. Operator 3 will be working on workstation 6,7, and 8. Operator 4 on workstation 9,10,11 and 12. Operator 5 on workstation

13,14,15 and 18. Operator 6 on workstation 17 and 1618. The details on the total allocation for each operator can be found in table 5.7

Table 5.7 Total allocation of each operator for line B

Operator	Total allocation (s)	Allocation (%)
OP1	34	100
OP2	24	70.59
OP3	31	91.18
OP4	34	100
OP5	32	94.12
OP6	34	100

The target cycle time for the assembly line was 34 out of which operator 1 will be working for 34s which accounts for 100% of allocation. Operator 2 will be working for 24s which account for 70.59% of the total capacity. Operator 3 has allocation of 31s which is 91.18%. Operator 4 and 6 will be working for 34s with allocation of 100%. Operator 5 will be working for 32s with 94.12% of the allocation. Furthermore, the hourly cost of the operator is 13€/hr so based on the allocation and time spent by the operator it is important to calculate the cost of the labor for each unit. Table 5.8 presents the range of cost based on the tariff.

Table 5.8 Cost of operator tasks for line B

WS number	Lower limit time	Upper limit time	Lower limit cost	Upper limit cost
1	3.8	6.2	0.01	0.02
2	14.5	19.5	0.05	0.07
3	9.9	14.1	0.04	0.05
4	7.1	10.9	0.03	0.04
5	13.5	16.5	0.05	0.06
6	12.2	15.8	0.04	0.06
7	8.6	11.4	0.03	0.04

8	4.9	9.1	0.02	0.03
9	4.1	7.9	0.01	0.03
10	10.7	15.3	0.04	0.06
11	6.6	9.4	0.02	0.03
12	5.1	8.9	0.02	0.03
13	4.3	7.7	0.02	0.03
14	8.5	11.5	0.03	0.04
15	7.4	10.6	0.03	0.04
16	12.6	17.4	0.05	0.06
17	16.4	21.6	0.06	0.08
18	5.2	8.8	0.02	0.03

In table 5.8, it can be observed the total labor cost on each workstation per piece of the product which ranges from 0.56€ to 0.81€ per unit

Similarly, table 5.9 shows the cost components for line C.

Table 5.9 Time of operator tasks on line C

WS number	Value added time	Non-value-added time	Unplanned unused time	Total time	Standard deviation	Lower limit time	Upper limit time
1	3	4	1	8	1.2	6.8	9.2
2	4	5	2	11	2.5	8.5	13.5
3	6	7	2	15	2.1	12.9	17.1
4	4	6	3	13	1.9	11.1	14.9
5	6	8	2	16	1.5	14.5	17.5
6	5	6	2	13	1.8	11.2	14.8
7	3	4	2	9	1.4	7.6	10.4
8	6	7	2	15	2.1	12.9	17.1
9	4	5	3	12	1.9	10.1	13.9

10	5	7	2	14	2.3	11.7	16.3
11	4	5	3	12	1.4	10.6	13.4
12	5	7	3	15	1.9	13.1	16.9
13	4	7	2	13	1.7	11.3	14.7
14	6	6	2	14	1.5	12.5	15.5
15	3	5	3	11	1.6	9.4	12.6
16	4	8	4	16	2.4	13.6	18.4
17	6	8	5	19	2.6	16.4	21.6
18	4	6	4	14	1.8	12.2	15.8

The target cycle time for the operator was 34 seconds for line C. The total time spent by the operator on 18 workstations is around 240 seconds, as in the model it was defined in the constraint that an operator cannot work more than the target cycle time so based on that we got the solution that there is requirement of 8 operators. Solving the problem, results can be seen in table 5.10.

Table 5.10 Operator allocation to workstation for line C

Workstation	Operator
WS1	Operator 1
WS2	Operator 1
WS3	Operator 2
WS4	Operator 1
WS5	Operator 2
WS6	Operator 3
WS7	Operator 4
WS8	Operator 3
WS9	Operator 4
WS10	Operator 5
WS11	Operator 4
WS12	Operator 5
WS13	Operator 6
WS14	Operator 7

WS15	Operator 6
WS16	Operator 7
WS17	Operator 8
WS18	Operator 8

As it can be observed in the table 5.10, operator 1 will be working on workstation 1, workstation 2, and workstation 4. Operator 2 will be working on workstation 3 and workstation 5. Operator 3 will be working on workstation 6 and 8. Operator 4 on workstation 7,9 and 11. Operator 5 on workstation 10 and 12. Operator 6 on workstation 13 and 15. Operator 7 on workstation 14 and 16. Lastly, operator 8 on workstation 17 and workstation 18. The details on the total allocation for each operator can be found in table 5.11

Table 5.11 Total allocation of each operator for line C

Operator	**Total allocation (s)**	**Allocation (%)**
OP1	32	94.12
OP2	30	88.24
OP3	28	82.35
OP4	33	97.06
OP5	29	85.29
OP6	24	70.59
OP7	30	88.24
OP8	33	97.06

Operator 1 will be working for 32s which accounts for 94.12% of allocation. Operator 2 will be working for 30s which account for 88.24% of the total capacity. Operator 3 has allocation of 28s which is 82.35%. Operator 4 will be working for 33s with allocation of 97%. Operator 5 will be working for 29s with 85.29% of the allocation. Operator 6 will be working for 24s with 70.59% of the allocation. Operator 7 will be working for 30s with 88.24% and operator 8 with 33s and 97.06% of allocation. Table 5.12 presents the range of cost based on the tariff.

Table 5.12 Cost of operator tasks for line C

WS number	Lower limit time	Upper limit time	Lower limit cost	Upper limit cost
1	6.8	9.2	0.02	0.03
2	8.5	13.5	0.03	0.05
3	12.9	17.1	0.05	0.06
4	11.1	14.9	0.04	0.05
5	14.5	17.5	0.05	0.06
6	11.2	14.8	0.04	0.05
7	7.6	10.4	0.03	0.04
8	12.9	17.1	0.05	0.06
9	10.1	13.9	0.04	0.05
10	11.7	16.3	0.04	0.06
11	10.6	13.4	0.04	0.05
12	13.1	16.9	0.05	0.06
13	11.3	14.7	0.04	0.05
14	12.5	15.5	0.05	0.06
15	9.4	12.6	0.03	0.05
16	13.6	18.4	0.05	0.07
17	16.4	21.6	0.06	0.08
18	12.2	15.8	0.04	0.06

In table 5.12, it can be observed the total labor cost on each workstation per piece of the product which ranges from 0.75€ to 0.99€ per unit

Similarly, table 5.13 shows the cost components for line D.

Table 5.13 Time comparison of operator tasks on line D

WS number	Value added time	Non-value-added time	Unplanned unused time	Total time	Standard deviation	Lower limit time	Upper limit time
1	3	3	1	7	1.2	4.8	7.2
2	5	4	1	10	2.5	10.5	15.5
3	6	7	3	16	2.1	11.9	16.1
4	4	6	2	12	1.9	9.1	12.9
5	5	6	2	13	1.5	13.5	16.5
6	4	6	3	13	1.8	10.2	13.8
7	3	4	2	9	1.4	6.6	9.4
8	4	5	2	11	2.1	7.9	12.1
9	3	5	1	9	1.9	5.1	8.9
10	3	5	2	10	2.3	9.7	14.3
11	2	2	1	5	1.4	3.6	6.4
12	3	3	2	8	1.9	8.1	11.9
13	3	4	2	9	1.7	6.3	9.7
14	4	5	1	10	1.5	7.5	10.5
15	2	2	1	5	1.6	5.4	8.6
16	5	6	3	14	2.4	13.6	18.4
17	5	8	3	16	2.6	15.4	20.6
18	2	3	1	6	1.8	3.2	6.8

The target cycle time for the operator was 34 seconds for line D. The total time spent by the operator on 18 workstations is around 183 seconds, so based on that we got the solution that there is requirement of 6 operators. Solving the problem, results can be seen in table 5.14.

Table 5.14 Operator allocation to workstation for line D

Workstation	Operator
WS1	Operator 1

WS2	Operator 1
WS3	Operator 2
WS4	Operator 2
WS5	Operator 1
WS6	Operator 3
WS7	Operator 3
WS8	Operator 3
WS9	Operator 4
WS10	Operator 4
WS11	Operator 4
WS12	Operator 4
WS13	Operator 5
WS14	Operator 5
WS15	Operator 5
WS16	Operator 6
WS17	Operator 6
WS18	Operator 5

As it can be observed in the table 5.14, operator 1 will be working on workstation 1, workstation 2, and workstation 5. Operator 2 will be working on workstation 3 and workstation 4. Operator 3 will be working on workstation 6,7 and 8. Operator 4 on workstation 9, 10, 11 and 12. Operator 5 on workstation 13, 14, 15 and 18. Operator 6 on workstation 16 and 17. The details on the total allocation for each operator can be found in table 5.15

Table 5.15 Total allocation of each operator for line D

Operator	Total allocation (s)	Allocation (%)
OP1	30	88.23
OP2	28	82.35
OP3	33	97.06

OP4	32	94
OP5	30	88.24
OP6	30	88.24

Operator 1 will be working for 30s which accounts for 88.23% of allocation. Operator 2 will be working for 28s which account for 82.35% of the total capacity. Operator 3 has allocation of 33s which is 97.06%. Operator 4 will be working for 32s with allocation of 94%. Operator 5 will be working for 30s with 88.24% of the allocation. Operator 6 will be working for 30s with 88.24% of the allocation. Table 5.16 presents the range of cost based on the tariff.

Table 5.16 Cost of operator tasks for line D

WS number	Lower limit time	Upper limit time	Lower limit cost	Upper limit cost
1	4.8	7.2	0.02	0.03
2	10.5	15.5	0.04	0.06
3	11.9	16.1	0.04	0.06
4	9.1	12.9	0.03	0.05
5	13.5	16.5	0.05	0.06
6	10.2	13.8	0.04	0.05
7	6.6	9.4	0.02	0.03
8	7.9	12.1	0.03	0.04
9	5.1	8.9	0.02	0.03
10	9.7	14.3	0.04	0.05
11	3.6	6.4	0.01	0.02
12	8.1	11.9	0.03	0.04
13	6.3	9.7	0.02	0.04
14	7.5	10.5	0.03	0.04
15	5.4	8.6	0.02	0.03
16	13.6	18.4	0.05	0.07
17	15.4	20.6	0.06	0.07
18	3.2	6.8	0.01	0.02

In table 5.16, it can be observed the total labor cost on each workstation per piece of the product which ranges from 0.55€ to 0.79€ per unit.

Table 5.17 Operator cost comparison across lines

Line	Lower limit cost (€)	Upper limit cost (€)
Line A	0.55	0.79
Line B	0.56	0.81
Line C	0.75	0.99
Line D	0.55	0.79

As can be observed in table 5.17, the range of labor cost varies across the lines to manufacture the same product.

From a total of 18 workstations in the assembly line, four workstations with the highest cycle time were chosen. Workstation 17 is the bottleneck workstation among all the other workstations in the assembly line. As the bottleneck time is also stochastic in nature, the range of the bottleneck was from 58 seconds to 62 seconds. Since the target cycle time of the assembly line is 34 seconds, there will be a need for 2 equipment on Workstation 17 to achieve the desired cycle time of the assembly line. But this increase in the equipment comes with the disadvantage of increased cost tariff which directly results in an increase in the cost of the final product.

Table 5.18 Component cost after installing 2nd equipment on WS 17

WS number	Specific cost	Line cost	Flexibility cost	Unplanned unused capacity cost	Total cost
5	0.68 €	0.062 €	0.392 €	0.240 €	1.374 €
17	0.82 €	0.004 €	0.40 €	0.212 €	1.436 €
16	0.85 €	- €	0.391 €	0.274 €	1.515 €
3	0.63 €	0.067 €	0.389 €	0.268 €	1.354 €

As it can be observed in table 5.3 the cost of the product for each component, since this cost was measured after installation of 2nd equipment the bottleneck of the assembly line was shifted to workstation 16 so the line cost is zero for that particular workstation. It is important to notice that line cost component decreased as the waiting period for all the workstation decreased which is somehow compensating the increased cost tariff. But considering the bottleneck time is stochastic in nature it is important to calculate the range of the cost. The cost is calculated for the period of 5 months for four different products which are produced in the same assembly line and the table can be seen below:

Table 5.19 Monthly cost comparison between products with extra equipment on WS 17

Product	**Month 1**		**Month 2**		**Month 3**		**Month 4**		**Month 5**	
	Lower limit cost	**Upper limit cost**	**Lower limit cost**	**Upper limit cost**	**Lower limit cost**	**Upper limit cost**	**Lower limit cost**	**Upper limit cost**	**Lower limit cost**	**Upper limit cost**
Product A	5.1	5.7	6.3	6.4	6.1	6.5	5.8	6.3	5.6	6.1
Product B	5.9	6.4	5.4	5.8	6.0	6.6	5.6	5.9	6.2	6.7
Product C	5.7	6.3	6.0	6.9	5.6	6.1	6.2	6.5	6.3	6.5
Product D	5.4	5.9	5.9	6.4	4.9	5.6	6.1	6.8	5.6	5.9

As it can be observed from the table 5.19 the range of the cost of the product follows the range between 5.1€ to 6.6€ for almost all the products in the period of 5 months and the accumulation of the cost is between 5.5€ to 6.4€. So, adding the equipment lowered the cost of product, even though the cost tariff increased but due to reduction the line cost it had positive impact on the cost of the product.

After installing the 2nd equipment in the workstation 17 and performing this analysis, the new bottleneck of the assembly line is workstation 16. So, one more iteration was made to test if adding 2nd equipment on workstation 16 will be still advantageous to reduce the cost of the product. So, following this model a simulation was made assuming to add an equipment to this workstation which resulted to the following results:

Table 5.20 Component cost after installing 2nd equipment on WS 16

WS number	Specific cost	Line cost	Flexibility cost	Unplanned unused capacity cost	Total cost
5	0.96 €	0.059 €	0.413 €	0.448 €	1.88 €
17	1.09 €	- €	0.40 €	0.432 €	1.923 €
16	1.12 €	0.003 €	0.423 €	0.512 €	2.058 €
3	1.10 €	0.063 €	0.441 €	0.491 €	2.095 €

As it can be observed in table 5.7 that by adding an equipment on workstation 16 it increases the cost tariff drastically which results into significant increase in the cost of the product and the bottleneck than is shifted back to workstation 17. The decrease in the bottleneck time and cost is so negligible that it cannot manage to balance the cost of equipment. So, it is not worth to add an additional equipment to any workstation as it will only result into increased cost of the product.

5.5 Final remarks

It is evident that there exist huge variations in the demand from the customer. So based on that changes it is required to adjust the allocation of the resources accordingly. Over allocation or under allocation of the resources can have a huge impact on the production abilities and eventually on the profitability of the company. To avoid these situations, it is very important to design this allocation of resources carefully. But for allocation of resources, it is important to understand how much time is spent on each activity and variability that comes along and the calculation of the cost.

In this chapter, the optimization of resource allocation was divided into two parts namely Operators and Equipment of Workstation. Using the time equations, the variability in the operator tasks were measured and costs were calculated. As the assembly line had the target cycle time, it was necessary that the total time spent by the operator should be less than the bottleneck or target cycle time. By using the traditional linear programming resource allocation problem to obtain the optimum number of operators required and best possible workstations that should be allocated to them was decided. With the equations developed it was possible to understand the activities performed by operator which were value added and non-value added to the product. So, it was possible to bifurcate between them and calculate the cost for it to clearly understand how much it costs for value added activities and non-value-added activities along with the inclusion of the variability.

As it was showcased in equation 22, based on the target cycle time it is possible to calculate the number of equipment needed in order to achieve that target. Following that, various iterations were performed to find a balance between the cost and adding an equipment to the workstation, as adding an equipment had direct effect in the increase of the cost of the product. But in some cases, adding the equipment was advantageous as the waiting time for other workstations decreased along with the bottleneck of the assembly line. Therefore, the cost component of bottleneck decreased (might as well change to some other workstation) which compensated the increased cost tariff, so the cost of the product was reduced.

But it is not always the case, as it was observed in table 5.5 that when the extra equipment was added it didn't have much effect on the bottleneck, so the cost of bottleneck component didn't reduce much but the tariff cost increased a lot which resulted into significant increase in the final cost of the product.

Hence, it can be concluded that using this methodology in order to calculate value added time and non-value added time of the operator along with the information about the optimum number of operators needed and their allocation of the workstation will help a lot to take some managerial decisions as they will have clear picture about the cost and focus can be then shifted on the optimization of the non-value added activities. Also, in terms of decision making for the designing of the line this approach can be helpful to understand better the effect on the cost of the product by adding the equipment to the workstation.

Chapter 6

6 Conclusions

All manufacturing companies must consider product costing as a critical component of its financial and operational management. It entails figuring out and totaling up all expenses incurred in the production of a specific product. The cost of a product which has an immediate effect on a company's profitability is heavily influenced by the cost of manufacturing. As a result, precise product costing is crucial for any business to be competitive in the marketplace. The cost of products has typically been decided by businesses using a top-down strategy. According to preset criteria like the quantity produced or the time spent in production, the total cost of production is divided among the products using this strategy. While this approach is relatively simple to implement, it does not accurately reflect the cost of each individual manufacturing process involved in production. As a result, it is difficult to identify areas where cost savings can be made or to predict the impact of changes in the production process

The variability still exists in most industrial processes which cannot be accommodated by the traditional costing approach. For instance, the quantity of products to be produced may fluctuate over time depending on changes in customer demands. Similar to this, the cycle time of each operation may change as a result of things like equipment failures, or quality problems. These differences have a large impact on the cost of production and can greatly fluctuate the company's profitability.

This research introduces the concept of stochastic cost model using time-driven activity-based costing as a bottom-up approach that provides a more accurate and detailed cost analysis by identifying and measuring the time required to perform each activity in the manufacturing process. The stochastic cost model can be used to assess the efficiency of the manufacturing process and determine the profitability of particular product assembly lines, can take into account the variability that is present in the production processes. The research also suggests using a prescriptive model to estimate production quantity, which aids in more correctly anticipating changes in product costs, allowing for the optimization of resource utilization throughout production and an increase in profitability. The stochastic models can be incorporated into the optimization models to account for the variability in the cost factors. The goal is to calculate the value-added and non-value-added cost and to find the optimal allocation of operators and the equipment needed in the assembly that minimizes the cost of the product while satisfying the constraints.

6.1 Contributions

Nowadays, traditional costing systems based on deterministic cost models are not enough to support adequately decision making. As it was identified in previous studies, the mean and standard deviation are widely used to analyze the product cost. Although, making estimations using the mean, in some cases, can lead to wrong conclusions (Zanjani et al., 2013)(Zanjani et al., 2013)(Zanjani et al., 2013), since the estimation can be far from the true value. Thus, the confidence interval can be used instead of the mean. Other alternatives are to consider the quartiles or the coefficient of variance.

This research proposes a stochastic approach to costing systems which considers the variability in the process cycle time. This approach provides a better perception of the risk associated to product costs. The confidence interval for the mean and the use of quartiles 1 and 3 as lower and upper estimates are proposed to include variability and risk in costing systems. The developed six-step methodology was applied in chapter 2. Only the bottleneck workstation of the assembly line was considered since it represents the cycle time of the line, but the analysis can be extended to all workstations in the production line. The allocation of the cost to line can be enough for product costing and to compute margins and compare costs and revenues supporting profitability analysis. Nevertheless, the optimization of production processes asks for a deeper analysis where costs should be highlighted by activities, workstations, or individual machines.

With this approach it opened new opportunities to perform the analysis of the cost variations over time with an awareness of the associated risks. The use of descriptive statistics gives the ability to understand and evaluate the behavior of cycle times and their influence on costs. In particular, the study of a confidence interval for the mean and the interquartile range gives us insights into what we can expect and the risk of getting values much higher and also lower than what was predicted.

The proposed methodology was extended by developing stochastic cost models based on activities or other more sophisticated costing systems. Indeed, by using a time driven activity-based costing approach, the variability for each activity can be recorded and more accurate product costs are obtained. The equations developed for costing can be used to determine the product's historical, present-day, or predictive cost, as well as to understand workstation-specific, bottleneck, unused capacity, and unanticipated unused capacity costs. The cost and its split by products and production lines can also be seen and understood using this cost model, which is simple and easy to use. The model's outputs made it possible to obtain a new cost breakdown of the product that is more pertinent for decision-making and provides an understandable comprehension and visualization of the cost from many angles. The model also enables us to simulate the effects of various adjustments on costs and the most advantageous course of action. With this division, we may have a deeper understanding of the cost by process, activity, operation, and particular workstation in addition to the cost of the product. The value added and non-value-added costs of the product can be distinguished based on this cost breakdown into four parts. One might imagine, for instance, how the bottleneck workstation affects product costs and what happens when a corporation produces more or less than anticipated. The dashboard developed gave a good visual aid to observe always changing costs because of variability across the assembly lines and different products.

The creation of an additive regression model, for a set of real data relating to industrial production. After collecting information regarding the production lines, over different months, and different weeks, they were analysed, extracting the (cor)relations between the respective variables. As such, an additive regression model is proposed to predict the actual quantities as a function of all other variables. This model is based on the M5 Tree classifier. The results show that the models with the variables “Bottleneck value”, “Amort Specific cost”, “Specific other cost” and/or “Bottleneck Workstation” have a good forecasting quality considering the measures R^2 (correlation coefficient), MAE (Mean absolute error), RMSE (Root mean squared error) and MAPE (Mean absolute percentage error). The linear regression model also enables the definition of a confidence interval for the corresponding predicted values, both in terms of costs and quantities. Understanding the variation in production over time is aided by the analysis of various time periods. They make clear various production-related behaviors, patterns, and trends. In addition, the prediction model enables the detection and forecasting of anomalous circumstances that may arise during the course of production (breaks, a lack of supplies, etc.). It is possible to specify the upper and lower bounds for acceptable variation using confidence intervals for the expected values. In other words, the model enables the evaluation of the risk associated with each of the lines produced outside the expected limits.

The proposed models enable the prediction of each unique manufacturing line's irrational behavior in addition to forecasts of production quantities. Since almost all production lines are more sensitive to significant differences in output, the work that was developed allowed us to draw attention to the caution that we must exercise in the control and planning of production. As a result of better productivity management, costs will also be better managed and less vulnerable to wide fluctuations. The method here helps in the study of the associated product costs and the risk of such costs by detecting abnormal behavior in production quantities. This risk relates to the possibility that a particular manufacturing line would produce quantities that are excessive and/or contrary to what would be wanted or expected given the production and market conditions. This has a significant impact on product costs and associated margins, necessary regular revisions of standard pricing based on accurate information about the state of the market and manufacturing.

Therefore, from a managerial perspective it is very advantageous as these models give information in real time about the cost of the product including the inherent variability. The pace of production can also be understood as the different and possible bottlenecks of the assembly line can be taken into analysis. Also, this will help to bridge the gap between the finance and production departments and bring them to common ground.

6.2 Limitations and opportunities for future research

In this research the focus was made on the variability in factors such as cycle time of the workstation, bottleneck of the assembly line, and production quantities. Depending on the type of industry there might exist more variables that have high impact on the cost of the product. One such variable is quality, variability in the quality can have impact on the cost. if a product is of inconsistent quality, it may result in additional costs associated with rework, repairs, or even recalls which can be costly in terms of time, materials, and labor.

Since certain manufacturing industry such as the aerospace industry and medical devices follow the high-quality standards and level of precision is very high, margin of error is very low. It will be important to consider quality as one of the factors of variability and stochastic cost model based on TDABC can be extended to this type of industries. Also, there are certain type of industry where the production happens continuously such as the oil and gas industry.

In the oil and gas industry the model might have a limited applicability as most of the processes happens automatically without any stoppage between the batches. It makes it difficult to keep the track on the quantities produced, takt time and eventually the process of production. It might be difficult to follow the bottom-up approach and to calculate the price per unit.

In other industries which still follows mostly the manual process, or the level of automation is very limited, the gathering of the data might be a big challenge specially when TDABC needs well structured data coming from ERP and MES. It is advisable to use IoT equipment integration in the industry so the resource can communicate independently and record time for each activity, that can be further used for the analysis. So, it is preferred to have automation as much as possible that can make the gathering of the data easier for the activities done by the machine as well as activities done by the operator. But implementing IoT equipment integration can be expensive, which small and medium industry might not be able to afford.

More complex models, such data-driven Vector Autoregression (VAR) models that concentrate on the historical correlations between variables, can be created. They do not rely on explicit theoretical structures; instead, they trace the impact of shocks using historical facts. Vector Autoregression models are useful for short- to medium-term forecasting because they are adaptable and can capture intricate relationships among variables. Exciting increases in computer capacity, data availability, and methodological innovations are expected to drive the future of VAR modeling. The amalgamation of machine learning methodologies with conventional VAR setups represents a noteworthy trajectory of advancement. Hybrid systems that combine VAR with Bayesian, deep learning, or reinforcement learning techniques have the ability to improve forecasting accuracy, capture nonlinear relationships, and more skillfully manage huge data difficulties.

Another frontier is the extension of VAR models to handle high-dimensional data. The curse of dimensionality limits the handling capacity of traditional VAR models for high numbers of variables. In order to solve this problem and enhance model interpretability, future study may concentrate on creating sparse VAR models by utilizing strategies like variable selection or regularization.

For sequential data forecasting jobs like time series forecasting, where the order of data points counts, other techniques like Recurrent Neural Networks (RNNs) are well suited. They can retain a state or memory of past inputs because of loops in their construction. Popular RNN variations like Long Short-Term Memory (LSTM) and Gated Recurrent Unit are particularly useful for time series forecasting because they can capture long-term relationships and mitigate the vanishing gradient problem. Hybrid architectures, which integrate various deep learning approaches or include domain knowledge into the model, may prove advantageous for some forecasting applications. For instance, a model might incorporate Long Short-Term Memory (LSTM) layers for sequence modeling and forecasting along with Convolutional Neural Networks for feature extraction from input data.

Multi-objective optimization for product costing is a strategic approach aimed at simultaneously achieving multiple goals while determining the cost of manufacturing a product. Traditionally, cost estimation in manufacturing has been focused primarily on minimizing production costs. However, in today's competitive landscape, companies must consider various objectives beyond mere cost reduction to stay competitive and sustainable. These objectives may include quality improvement, lead time reduction, resource utilization enhancement, and environmental impact minimization, among others. Multi-objective optimization provides a framework to balance these competing goals and find optimal solutions that satisfy multiple criteria simultaneously.

One key aspect of multi-objective optimization for product costing is the identification and prioritization of relevant objectives. This involves understanding the specific needs and constraints of the organization, as well as considering broader industry trends and customer expectations. For example, a company may prioritize cost reduction while also aiming to improve product quality and reduce environmental impact. By clearly defining these objectives, decision-makers can guide the optimization process towards generating solutions that align with the company's strategic goals.

Another important consideration in multi-objective optimization is the selection of appropriate optimization techniques and algorithms. Various mathematical optimization methods, such as linear programming, nonlinear programming, genetic algorithms, and simulated annealing, can be employed to explore the trade-offs between different objectives and identify Pareto-optimal solutions. These solutions represent the best possible compromises between conflicting objectives, where no single objective can be improved without sacrificing the performance of another.

Furthermore, integrating advanced modeling and simulation techniques into the optimization process can enhance decision-making and facilitate scenario analysis. By developing accurate cost models and simulating different production scenarios, companies can evaluate the impact of various decisions on cost, quality, and other relevant objectives. This enables them to make informed decisions and identify opportunities for improvement across multiple dimensions.

Moreover, multi-objective optimization for product costing requires collaboration and alignment across different functions within the organization. Cross-functional teams involving professionals from finance, operations, engineering, and supply chain management can provide diverse perspectives and expertise, leading to more robust and effective solutions. Additionally, involving stakeholders such as suppliers, customers, and regulatory agencies can help ensure that the optimization process considers broader societal and environmental implications.

Bibliography

A.B. Ganorkar, R.R. Lakhe, & K.N. Agrawal. (2019). COST AND PRODUCTIVITY ANALYSIS OF THE MANUFACTURING INDUSTRY USING TDABC & MOST. *South African Journal of Industrial Engineering*, *30*(1), 196–208.

Abounadi, J., Bertsekas, D., & Borkar, V. S. (2002). Learning algorithms for Markov decision processes with average cost. *SIAM Journal on Control and Optimization*. https://doi.org/10.1137/S0363012999361974

Afonso, P., & Jiménez, V. J. (2015). *Costing Systems for Decision Making under Uncertainty Using Probabilistic Models*. https://doi.org/10.4018/978-1-4666-9458-3.ch010

Afonso, P., & Santana, A. (2016). Application of the TDABC model in the logistics process using different capacity cost rates. *Journal of Industrial Engineering and Management*. https://doi.org/10.3926/jiem.2086

Afonso, P., Vyas, V., Antunes, A., Silva, S., & Bret, B. P. J. (2021). A stochastic approach for product costing in manufacturing processes. *Mathematics*, *9*(18). https://doi.org/10.3390/math9182238

Alessandroni, S., Dell'Isola, F., & Porfiri, M. (2002). A revival of electric analogs for vibrating mechanical systems aimed to their efficient control by PZT actuators. *International Journal of Solids and Structures*. https://doi.org/10.1016/S0020-7683(02)00402-X

Alsoboa, Sliman , AL-Ghazzwi, Ali, J. A. (2015). The Impact of Strategic Costing Techniques on the Performance of Jordanian Listed Manufacturing Companies. *Research Journal of Finance and Accounting*, *6*(10).

Anbumalar, V., & Sekar, R. C. (2015). Methods for solving cell formation, static layout and dynamic layout cellular manufacturing system problems: A review. *Asian Journal of Science and Technology*, *6*(12), 2107–2112.

Anderson, A. M., & van der Merwe, A. (2021). Time-driven activity-based costing related to digital twinning in additive manufacturing. *South African Journal of Industrial Engineering*. https://doi.org/10.7166/32-1-2271

Anderson, O. D., Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1985). The Forecasting Accuracy of Major Time Series Methods. *The Statistician*. https://doi.org/10.2307/2988175

Balakrishnan, J., & Cheng, C. H. (2007). Multi-period planning and uncertainty issues in cellular manufacturing: A review and future directions. *European Journal of Operational Research*, *177*(1), 281–309. https://doi.org/10.1016/j.ejor.2005.08.027

Balakrishnan, J., & Hung Cheng, C. (2005). Dynamic cellular manufacturing under multiperiod planning horizons. *Journal of Manufacturing Technology Management*, *16*(5), 516–530. https://doi.org/10.1108/17410380510600491

Baldoni, E., Coderoni, S., D'Orazio, M., Di Giuseppe, E., & Esposti, R. (2019). The role of economic and policy variables in energy-efficient retrofitting assessment. A stochastic Life Cycle Costing methodology. *Energy Policy*. https://doi.org/10.1016/j.enpol.2019.03.018

Baldoni, M., Bergenti, F., El Fallah Seghrouchni, A., & Winikoff, M. (2021). Special issue on Current trends in research on software agents and agent-based software systems. In *Autonomous Agents and Multi-Agent Systems*. https://doi.org/10.1007/s10458-021-09510-0

Banker, R. D., Bardhan, I. R., & Chen, T. Y. (2008). The role of manufacturing practices in mediating the impact of activity-based costing on plant performance. *Accounting, Organizations and Society*. https://doi.org/10.1016/j.aos.2006.12.001

Barth, M., Livet, A., & de Guio, R. (2008). Effective activity-based costing for manufacturing enterprises using a shop floor reference model. *International Journal of Production Research*, *46*(3), 621–646. https://doi.org/10.1080/00207540600845750

Benjamin, S. J., Muthaiyah, S., & Marathamuthu, M. S. (2009). An improved methodology for absorption costing: Efficiency Based Absorption Costing (EBAC). *Journal of Applied Business Research*, *25*(6). https://doi.org/10.19030/jabr.v25i6.998

Berk, L., Bertsimas, D., Weinstein, A. M., & Yan, J. (2019). Prescriptive analytics for human resource planning in the professional services industry. *European Journal of Operational Research*, *272*(2), 636–641. https://doi.org/10.1016/j.ejor.2018.06.035

Bertsimas, D., & Kallus, N. (2020). From Predictive to Prescriptive Analytics. *Management Science*, *66*(3), 1025–1044. https://doi.org/10.1287/mnsc.2018.3253

Borràs Mora, E., Spelling, J., & van der Weijde, A. H. (2021). Global sensitivity analysis for offshore wind cost modelling. *Wind Energy*. https://doi.org/10.1002/we.2612

Brauner, N., & Finke, G. (2001). Optimal moves of the material handling system in a robotic cell. *International Journal of Production Economics*, *74*(1–3), 269–277. https://doi.org/10.1016/S0925-5273(01)00132-3

Bruggeman, Werner; Everaert, Patricia; Anderson, Steven R. Anderson; Levant, Y. (2005). Modeling Logistics Costs using Time-Driven ABC : A Case in a Distribution Company. *Conceptual Paper and Case Study*.

Byrne, P. (1997). Fuzzy DCF: a contradiction in terms, or a way to better investment appraisal. *Cutting Edge 1997*.

Chansaad, A. P., Rattanamanee, W., Chaiprapat, S., & Yenradee, P. (2012). A Fuzzy Time-Driven Activity-Based Costing Model in an Uncertain Manufacturing Environment. *Proceedings of the Asia Pacific Industrial Engineering & Management Systems Conference 2012*.

Chen, J.-H., & Ho, S.-Y. (2005). A novel approach to production planning of flexible manufacturing systems using an efficient multi-objective genetic algorithm. *International Journal of Machine Tools and Manufacture*, *45*(7–8), 949–957. https://doi.org/10.1016/j.ijmachtools.2004.10.010

Chen, X., Cheng, L., Deng, G., Guan, S., & Hu, L. (2021). Project duration-cost-quality prediction model based on Monte Carlo simulation. *Journal of Physics: Conference Series*, *1978*(1), 012048. https://doi.org/10.1088/1742-6596/1978/1/012048

Chien, C.-F., Zheng, J.-N., & Lin, Y.-J. (2014). Determining the operator-machine assignment for machine interference problem and an empirical study in semiconductor test facility. *Journal of Intelligent Manufacturing*, *25*(5), 899–911. https://doi.org/10.1007/s10845-013-0777-3

Chukwubuikem Okoye, P. V., Chinedu Egbunike, F., & Mofolusho Meduoye, O. (2013). Product Cost Management via the Kaizen Costing System: Perception of Accountants. *Journal of Management and Sustainability*, *3*(4). https://doi.org/10.5539/jms.v3n4p114

Cılasun Kunduracı, A., & Kazanasmaz, Z. T. (2019). Fuzzy logic model for the categorization of manual lighting control behaviour patterns based on daylight illuminance and interior layout. *Indoor and Built Environment*, *28*(5), 584–598. https://doi.org/10.1177/1420326X17703772

Cooper, R., & Kaplan, R. (1988). Measure costs right: make the right decisions. *Harvard Business Review*, *66*(5), 96–103.

Cooper, R., & Kaplan, R. S. (1987). How Cost Accounting Systematically Distorts Product Costs. *Accounting and Management Field Study Perspectives*, 49–72.

Corbett, J., & Crookall, P. R. (1986). Design for Economic Manufacture. *CIRP Annals - Manufacturing Technology*. https://doi.org/10.1016/S0007-8506(07)61846-0

Datta, P. P., & Roy, R. (2010). Cost modelling techniques for availability type service support contracts: A literature review and empirical study. *CIRP Journal of Manufacturing Science and Technology*. https://doi.org/10.1016/j.cirpj.2010.07.003

de Iuliis, M., Kammouh, O., Cimellaro, G. P., & Tesfamariam, S. (2019). Resilience of the Built Environment: A Methodology to Estimate the Downtime of Building Structures Using Fuzzy Logic. In *Resilient Structures and Infrastructure* (pp. 47–76). Springer Singapore. https://doi.org/10.1007/978-981-13-7446-3_2

Devriendt, F., Moldovan, D., & Verbeke, W. (2018). A Literature Survey and Experimental Evaluation of the State-of-the-Art in Uplift Modeling: A Stepping Stone Toward the Development of Prescriptive Analytics. *Big Data*, *6*(1), 13–41. https://doi.org/10.1089/big.2017.0104

Drury, C. (2013). *Management and cost accounting* (8th ed.).

Drury, C., & Tayles, M. (1994). Product costing in uk manufacturing organizations. *European Accounting Review*. https://doi.org/10.1080/09638189400000031

Durán, O., Afonso, P., & Minatogawa, V. (2020). Analysis of long-term impact of maintenance policy on maintenance capacity using a time-driven activity-based life-cycle costing. *Mathematics*. https://doi.org/10.3390/math8122208

Durán, O., & Durán, P. A. (2018). Activity Based Costing for wastewater treatment and reuse under uncertainty: A fuzzy approach. *Sustainability (Switzerland)*. https://doi.org/10.3390/su10072260

Esmalifalak, H., Albin, M. S., & Behzadpoor, M. (2015). A comparative study on the activity based costing systems: Traditional, fuzzy and Monte Carlo approaches. *Health Policy and Technology*. https://doi.org/10.1016/j.hlpt.2014.10.010

Filomena, T. P., Anzanello, M. J., Neto, F. J. K., Duffey, M., & Campos-Nanez, E. (2011). Manufacturing feature-based cost management system: A case study in Brazil. *Production Planning and Control*, *22*(4). https://doi.org/10.1080/09537287.2010.497505

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Galli, L., Levato, T., Schoen, F., & Tigli, L. (2021). Prescriptive analytics for inventory management in health care. *Journal of the Operational Research Society*, *72*(10), 2211–2224. https://doi.org/10.1080/01605682.2020.1776167

Gervais, M., Levant, Y., & Ducrocq, C. (2010). Time-Driven Activity-Based Costing (TDABC): An Initial Appraisal through a Longitudinal Case Study. *Journal of Applied Management Accounting Research*, *8*(2), 1.

Gilbert, S. J. (2007). Adding Time to Activity-Based Costing. *Harvard Business School*.

Goh, Y. M., Newnes, L. B., Mileham, A. R., McMahon, C. A., & Saravi, M. E. (2010). Uncertainty in through-life costing-review and perspectives. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2010.2040745

Gosselin, M. (1997). The effect of strategy and organizational structure on the adoption and implementation of activity-based costing. *Accounting, Organizations and Society*. https://doi.org/10.1016/S0361-3682(96)00031-1

GUPTA, Y., GUPTA, M., KUMAR, A., & SUNDARAM, C. (1996). A genetic algorithm-based approach to cell composition and layout design problems. *International Journal of Production Research*, *34*(2), 447–482. https://doi.org/10.1080/00207549608904913

Haque, L., & Armstrong, M. J. (2007). A survey of the machine interference problem. *European Journal of Operational Research*, *179*(2), 469–482. https://doi.org/10.1016/j.ejor.2006.02.036

Hastie, T. J., & R. J. Tibshirani. (1990). *Generalized additive models London chapman and hall*.

Haykin, S. (1998). *Neural networks: a comprehensive foundation*.

Hillier, F. S., & So, K. C. (1991). The effect of the coefficient of variation of operation times on the allocation of storage space in production line systems. *IIE Transactions (Institute of Industrial Engineers)*, *23*(2). https://doi.org/10.1080/07408179108963854

Hoozée, S. (2013). Designing time-driven activity-based costing systems: A review and future research directions. In *The Routledge Companion to Cost Management*. https://doi.org/10.4324/9780203101261-16

Hoozée, S., & Bruggeman, W. (2010). Identifying operational improvements during the design process of a time-driven ABC system: The role of collective worker participation and leadership style. *Management Accounting Research*. https://doi.org/10.1016/j.mar.2010.01.003

Hoozée, S., & Hansen, S. C. (2018). A comparison of activity-based costing and time-driven activity-based costing. *Journal of Management Accounting Research*. https://doi.org/10.2308/jmar-51686

Hoozée, S., Vermeire, L., & Bruggeman, W. (2012). The Impact of Refinement on the Accuracy of Time-driven ABC. *Abacus*, *48*(4), 439–472. https://doi.org/10.1111/j.1467-6281.2012.00366.x

Huang, M., Song, M., Leon, V. J., & Wang, X. (2014). Decentralized capacity allocation of a single-facility with fuzzy demand. *Journal of Manufacturing Systems*, *33*(1), 7–15. https://doi.org/10.1016/j.jmsy.2013.09.004

Hung, Y., & Cheng, G. (2002). Hybrid capacity modeling for alternative machine types in linear programming production planning. *IIE Transactions*, *34*(2), 157–165. https://doi.org/10.1023/A:1011943930465

Ioannou, A., Angus, A., & Brennan, F. (2017). Stochastic Prediction of Offshore Wind Farm LCOE through an Integrated Cost Model. *Energy Procedia*. https://doi.org/10.1016/j.egypro.2016.12.180

Ioannou, A., Angus, A., & Brennan, F. (2020). Stochastic financial appraisal of offshore wind farms. *Renewable Energy*. https://doi.org/10.1016/j.renene.2019.06.111

JHA, N. K. (1992). Stochastic mathematical modelling and manufacturing cost estimation in uncertain industrial environment. *International Journal of Production Research*, *30*(12), 2755–2771. https://doi.org/10.1080/00207549208948189

Johri, P. K. (1993). Practical issues in scheduling and dispatching in semiconductor wafer fabrication. *Journal of Manufacturing Systems*, *12*(6), 474–485. https://doi.org/10.1016/0278-6125(93)90344-S

Kaplan, R. S. (1984). The evolution of management accounting. In *Readings in Accounting for Management Control*. https://doi.org/10.1007/978-1-4899-7138-8_27

Kaplan, R. S., & Anderson, S. R. (2007). Time-driven activity-based costing: a simpler and more powerful path to higher profits. *Harvard Business School Press Books*.

Kaur, A., & Kumar, R. (2015). Comparative Analysis of Parametric and Non-Parametric Tests. *Journal of Computer and Mathematical Sciences*, *6*(6).

Keren, B., & Hadad, Y. (2016). ABC Inventory Classification Using AHP and Ranking Methods via DEA. *2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO)*, 495–501. https://doi.org/10.1109/SMRLO.2016.87

Khademian, M., Tavana, M., Di Caprio, & D., & S.-A. (2018). An extended time-driven activity-based costing model for cost estimation in manufacturing. *Journal of Industrial and Production Engineering*, *35*(7), 413–424.

Kishk, M., & Al-Hajj, A. (2002). A Fuzzy Model And Algorithm to Handle Subjectivity In Life Cycle Costing Based Decision-Making. *Architecture & Built Environment*.

KOTTAS, J. F., & LAU, H.-S. (1981). A stochastic line balancing procedure. *International Journal of Production Research*, *19*(2), 177–193. https://doi.org/10.1080/00207548108956640

Landwehr, N., Mark Hall, & Eibe Frank. (2005). Logistic model trees. *Machine Learning*, *59*, 161–205.

Lars Mönch, Scott J. Mason, & John W. Fowler. (2013). *Production planning and control for semiconductor wafer fabrication facilities: modeling, analysis, and systems*.

Leachman, R. C., & Carmon, T. F. (1992). ON CAPACITY MODELING FOR PRODUCTION PLANNING WITH ALTERNATIVE MACHINE TYPES. *IIE Transactions*, *24*(4), 62–72. https://doi.org/10.1080/07408179208964234

Lee, Y. H., & Chen, H. (2012). Bayesian Stochastic Cost Modeling for Semiconductor Manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, *25*(3), 452–463.

Levant, Y., & Zimnovitch, H. (2013). Contemporary evolutions in costing methods: Understanding these trends through the use of equivalence methods in France. *Accounting History*, *18*(1), 51–75. https://doi.org/10.1177/1032373212463793

Lily Koops. (2020). *Optimized Maintenance Decision-Making – A Simulation-supported Prescriptive Analytics Approach based on Probabilistic Cost-Benefit Analysis* (Bauhaus Luftfahrt e.V & Willy-Messerschmitt-Str, Eds.).

Liu, Y., Zhang, L., Tao, F., & Wang, L. (2015). Resource service sharing in cloud manufacturing based on the Gale–Shapley algorithm: advantages and challenge. *International Journal of Computer Integrated Manufacturing*, 1–13. https://doi.org/10.1080/0951192X.2015.1067916

Martínez-Costa, C., Mas-Machuca, M., Benedito, E., & Corominas, A. (2014). A review of mathematical programming models for strategic capacity planning in manufacturing. *International Journal of Production Economics*, *153*, 66–85. https://doi.org/10.1016/j.ijpe.2014.03.011

Meister, M., Beßle, J., Cviko, A., Böing, T., & Metternich, J. (2019). Manufacturing analytics for problem-solving processes in production. *Procedia CIRP*, *81*. https://doi.org/10.1016/j.procir.2019.03.001

Miller, E. M. (1977). RISK, UNCERTAINTY, AND DIVERGENCE OF OPINION. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Montgomery, D. C., & Runger, G. C. (2018). *Applied statistics and probability for engineers*. Wiley Hoboken, NJ.

Nachtmann, H., & Needy, K. L. (2003). Methods for handling uncertainty in activity based costing systems. *Engineering Economist*. https://doi.org/10.1080/00137910308965065

Nakade, K., & Nishiwaki, R. (2008). Optimal allocation of heterogeneous workers in a U-shaped production line. *Computers & Industrial Engineering*, *54*(3), 432–440. https://doi.org/10.1016/j.cie.2007.08.007

Nakade, K., & Ohno, K. (1999). An optimal worker allocation problem for a U-shaped production line. *International Journal of Production Economics*, *60–61*, 353–358. https://doi.org/10.1016/S0925-5273(98)00145-5

Noh, J. S., Kim, J. S., & Sarkar, B. (2019). Stochastic joint replenishment problem with quantity discounts and minimum order constraints. *Operational Research*. https://doi.org/10.1007/s12351-016-0281-6

Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V., & Alvin, K. F. (1999). New methodology for the estimation of total uncertainty in computational simulation. *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. https://doi.org/10.2514/6.1999-1612

Oh, S. C., & Shin, J. (2021). The assessment of car making plants with an integrated stochastic frontier analysis model. *Mathematics*. https://doi.org/10.3390/math9111296

Orellana, R., Carvajal, R., Escárate, P., & Agüero, J. C. (2021). On the uncertainty identification for linear dynamic systems using stochastic embedding approach with gaussian mixture models. *Sensors*. https://doi.org/10.3390/s21113837

Osuch, A., Osuch, E., Rybacki, P., Przygodziński, P., Kozłowski, R., & Przybylak, A. (2020). A Decision Support Method for Choosing an Agricultural Machinery Service Workshop Based on Fuzzy Logic. *Agriculture*, *10*(3), 76. https://doi.org/10.3390/agriculture10030076

Park, C., Park, P. K., Mane, P. P., Hyung, H., Gandhi, V., Kim, S. H., & Kim, J. H. (2010). Stochastic cost estimation approach for full-scale reverse osmosis desalination plants. *Journal of Membrane Science*. https://doi.org/10.1016/j.memsci.2010.07.055

Park, J., Bae, H., Dinh, T.-C., & Ryu, K. (2014). Operator allocation in cellular manufacturing systems by integrated genetic algorithm and fuzzy data envelopment analysis. *The International Journal of Advanced Manufacturing Technology*, *75*(1–4), 465–477. https://doi.org/10.1007/s00170-014-6103-1

Park, Y., Jung, S., & Jahmani, Y. (2019). Time-Driven Activity-Based Costing Systems for Marketing Decisions. *Studies in Business and Economics*. https://doi.org/10.2478/sbe-2019-0015

Paulo Afonso, Rodney Wernke, & Antonio Zanin. (2018). Managing the cost of unused capacity: an integrative and comparative analysis of the ABC, TDABC and UEP methods. *Revista Del Instituto Internacional de Costos*, *1*, 150–163.

Pember, A., & Lemon, M. (2015). Measuring and managing environmental sustainability: Using activity-based costing/management (ABC/M)1. *2nd Environmental Considerations in Energy Production Conference*.

Poornima, S., & Pushpalatha, M. (2020). A survey on various applications of prescriptive analytics. *International Journal of Intelligent Networks*, *1*, 76–84. https://doi.org/10.1016/j.ijin.2020.07.001

Punia, S., Singh, S. P., & Madaan, J. K. (2020). From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. *Decision Support Systems*, *136*, 113340. https://doi.org/10.1016/j.dss.2020.113340

Quinlan, J. (1992). Learning with continuous classes. *5th Australian Joint Conference on Artificial Intelligence. Vol. 92*, 343–348.

Renna, P., & Argoneto, P. (2010). A game theoretic coordination for trading capacity in multisite factory environment. *The International Journal of Advanced Manufacturing Technology*, *47*(9–12), 1241–1252. https://doi.org/10.1007/s00170-009-2254-x

Robin Cooper. (1990). Cost classification in unit-based and activity-based manufacturing cost systems. *Journal of Cost Management*, *Fall*, 4–14.

Robin Cooper. (1995). Activity-based costing: Theory and practice. In B. J. Brinker (Ed.), *Cost Management*.

SARIN, S. C., & EREL, E. (1990). Development of cost model for the single-model stochastic assembly line balancing problem. *International Journal of Production Research*, *28*(7), 1305–1316. https://doi.org/10.1080/00207549008942794

Sarokolaei, M. A., Bahreini, M., & Bezenjani, F. P. (2013). Fuzzy Performance Focused Activity based Costing (PFABC). *Procedia - Social and Behavioral Sciences*, *75*. https://doi.org/10.1016/j.sbspro.2013.04.039

Scanlan, J., Rao, A., Bru, C., Hale, P., & Marsh, R. (2006). DATUM project: Cost estimating environment for support of aerospace design decision making. *Journal of Aircraft*. https://doi.org/10.2514/1.17362

Sengewald, J. , & Lackes, R. (2022). *Prescriptive Analytics in Procurement: Reducing Process Costs*.

Seok, H., & Nof, S. Y. (2014). Dynamic coalition reformation for adaptive demand and capacity sharing. *International Journal of Production Economics*, *147*, 136–146. https://doi.org/10.1016/j.ijpe.2013.09.003

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), 53. https://doi.org/10.1186/s40537-020-00329-2

Sheikh Amir, Majid Zaman, & Muheet Ahmed Butt. (2022). Knowledge discovery in geographical sciences—A systematic survey of various machine learning algorithms for rainfall prediction. *International Conference on Innovative Computing and Communications : Proceedings of ICICC 2021.Springer Singapore.*, 593–608.

Sidiq, S. J., Majid Zaman, & Muheet Ahmed. (2019). How machine learning is redefining geographical science: A review of literature. *International Journal of Emerging Technologies and Innovative Research*, *6*, 1731–1746.

Singh, C., & Singh, S. R. (2020). Supply chain model for expiring items following ramp-type demand with stochastic lead time under crisp and fuzzy environment. *International Journal of Fuzzy System Applications*. https://doi.org/10.4018/IJFSA.2020010103

Singh, S. R., & Prasher, L. (2014). A production inventory model with flexible manufacturing, random machine breakdown and stochastic repair time. *International Journal of Industrial Engineering Computations*. https://doi.org/10.5267/j.ijiec.2014.7.003

Sirovetnukul, R., & Chutima, P. (2009). Worker allocation in U-shaped assembly lines with multiple objectives. *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, 105–109. https://doi.org/10.1109/IEEM.2009.5373414

Sirovetnukul, R., & Chutima, P. (2010). The Impact of Walking Time on U-Shaped Assembly Line Worker Allocation Problems. *Engineering Journal*, *14*(2), 53–78. https://doi.org/10.4186/ej.2010.14.2.53

Sobczyk, T., & Koch, T. (2008). A method for measuring operational and financial performance of a production value stream. *IFIP International Federation for Information Processing*, *257*. https://doi.org/10.1007/978-0-387-77249-3_16

Sobu, A., & Wu, G. (2012). Optimal operation planning method for isolated micro grid considering uncertainties of renewable power generations and load demand. *2012 IEEE Innovative Smart Grid Technologies - Asia, ISGT Asia 2012*. https://doi.org/10.1109/ISGT-Asia.2012.6303271

Solomatine, D. P., & M. B. Siek. (2003). Flexibility and optimality in M5 model trees. *3rd Int. Conf. on Hybrid Intelligent Systems*.

Sparling, D., & Miltenburg, J. (1998). The mixed-model U-line balancing problem. *International Journal of Production Research*, *36*(2), 485–501. https://doi.org/10.1080/002075498193859

Stafford, E. F. (1988). On the Development of a Mixed-Integer Linear Programming Model for the Flowshop Sequencing Problem. *Journal of the Operational Research Society*, *39*(12), 1163–1174. https://doi.org/10.1057/jors.1988.193

Stecke, K. E. (1982). *Machine interference: The assignment of machines to operators. I* (John Wiley and Sons, Ed.; G. Salvendy).

STECKE, K. E., & ARONSON, J. E. (1985). Review of operator/machine interference models. *International Journal of Production Research*, *23*(1), 129–151. https://doi.org/10.1080/00207548508904696

Swaroop, P. V., Erlich, I., Rohrig, K., & Dobschinski, J. (2009). A stochastic model for the optimal operation of a wind-thermal power system. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2009.2016504

Tang, Z., Guo, L., Zheng, T., Li, Z., Sun, R., & Huang, K. (2022). A combined machine learning and numerical approach for evaluating the uncertainty of 3D angle-interlock woven composites. *Composite Structures*, *294*, 115726. https://doi.org/10.1016/j.compstruct.2022.115726

Tayyab, M., Jemai, J., Lim, H., & Sarkar, B. (2020). A sustainable development framework for a cleaner multi-item multi-stage textile production system with a process improvement initiative. *Journal of Cleaner Production*, *246*, 119055. https://doi.org/10.1016/j.jclepro.2019.119055

Tirkel, I., & Rabinowitz, G. (2014). Modeling cost benefit analysis of inspection in a production line. *International Journal of Production Economics*, *147*, 38–45. https://doi.org/10.1016/j.ijpe.2013.05.012

Toktay, L. B., & Uzsoy, R. (1998). A capacity allocation problem with integer side constraints. *European Journal of Operational Research*, *109*(1), 170–182. https://doi.org/10.1016/S0377-2217(98)80011-8

Torun, H., & Çetinkaya, S. (2019). Machine criticality level assignment with fuzzy inference system for RCM. *International Conference on Intelligent and Fuzzy Systems*, 1363–1370.

Vedernikova, O., Siguenza-Guzman, L., Pesantez, J., & Arcentales-Carrion, R. (2020). Time-Driven Activity-Based Costing in the Assembly Industry. *Australasian Business, Accounting & Finance Journal*, *14*(4), 3–23. https://doi.org/10.14453/aabfj.v14i4.2

Venugopal, V., & Narendran, T. T. (1992). Cell formation in manufacturing systems through simulated annealing: An experimental evaluation. *European Journal of Operational Research*, *63*(3), 409–422. https://doi.org/10.1016/0377-2217(92)90158-6

VRAT, P., & VIRANI, A. (1976). A cost model for optimal mix of balanced stochastic assembly line and the modular assembly system for a customer oriented production system. *International Journal of Production Research*, *14*(4), 445–463. https://doi.org/10.1080/00207547608956618

Wang, C. N., Nhieu, N. L., & Tran, T. T. T. (2021). Stochastic chebyshev goal programming mixed integer linear model for sustainable global production planning. *Mathematics*. https://doi.org/10.3390/math9050483

Wang, C., Tan, X. P., Tor, S. B., & Lim, C. S. (2020). Machine learning in additive manufacturing: State-of-the-art and perspectives. *Additive Manufacturing*, *36*, 101538. https://doi.org/10.1016/j.addma.2020.101538

Wang, Y., & Ian H. Witten. (1997). Induction of model trees for predicting continuous classes. *9th European Conference on Machine Learning, 1997*.

Wu, S. D., Erkoc, M., & Karabuk, S. (2005). Managing Capacity in the High-Tech Industry: A Review of Literature. *The Engineering Economist*, *50*(2), 125–158. https://doi.org/10.1080/00137910590949887

Yazici, B., & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, *77*(2). https://doi.org/10.1080/10629360600678310

Yildiz, A. R. (2013). Comparison of evolutionary-based optimization algorithms for structural design optimization. *Engineering Applications of Artificial Intelligence*, *26*(1), 327–333. https://doi.org/10.1016/j.engappai.2012.05.014

Yuan, F. C. (2009). The use of a fuzzy logic-based system in cost-volume-profit analysis under uncertainty. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2007.11.025

Yves Levant, & Olivier de La Villarmois. (2011). From adoption to use of a management control tool: Case study evidence of a costing method. *Journal of Applied Accounting Research*, *12*(3), 234–259.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*. https://doi.org/10.1016/S0019-9958(65)90241-X

Zamrud, N. F., Abu, M. Y., Nik Mohd Kamil, N. N., & Safeiee, F. L. M. (2020). *The Impact of Capacity Cost Rate and Time Equation of Time-Driven Activity-Based Costing (TDABC) on Electric Component*. https://doi.org/10.1007/978-981-15-0950-6_13

Zanjani, M. K., Ait-Kadi, D., & Nourelfath, M. (2013). A stochastic programming approach for sawmill production planning. *International Journal of Mathematics in Operational Research*. https://doi.org/10.1504/IJMOR.2013.050604

Zha, S., Guo, Y., Huang, S., & Tang, P. (2020). A hybrid optimization approach for unequal-sized dynamic facility layout problems under fuzzy random demands. *Journal of Engineering Manufacture*, *234*(3), 381–399.

Zhang, X., Prajapati, M., & Peden, E. (2011). A stochastic production planning model under uncertain seasonal demand and market growth. *International Journal of Production Research*. https://doi.org/10.1080/00207541003690074

Zimmermann, H. J. (2000). Application-oriented view of modeling uncertainty. *European Journal of Operational Research*. https://doi.org/10.1016/S0377-2217(99)00228-3