applied sciences

Special Issue Reprint

Design and Optimization of Manufacturing Systems, 2nd Edition

Edited by
Zoran Jurković, David Ištoković and Janez Gotlih

mdpi.com/journal/applsci

MDPI

Design and Optimization of Manufacturing Systems, 2nd Edition

Design and Optimization of Manufacturing Systems, 2nd Edition

Guest Editors

Zoran Jurković
David Ištoković
Janez Gotlih

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Zoran Jurković
Faculty of Engineering
University of Rijeka
Rijeka
Croatia

David Ištoković
Faculty of Engineering
University of Rijeka
Rijeka
Croatia

Janez Gotlih
Faculty of Mechanical Engineering
University of Maribor
Maribor
Slovenia

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Applied Sciences* (ISSN 2076-3417), freely accessible at: https://www.mdpi.com/journal/applsci/special_issues/A7TG1WZEKW.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-4405-0 (Hbk)
ISBN 978-3-7258-4406-7 (PDF)
https://doi.org/10.3390/books978-3-7258-4406-7

Contents

About the Editors

Zoran Jurković

Zoran Jurković is a Full Professor at the Faculty of Engineering, University of Rijeka, Croatia, where he works in the Department of Industrial Engineering and Management. His scientific and professional activities focus on industrial engineering, production management, and the digital transformation of production systems. He is particularly interested in developing and applying modern optimization and simulation methods to improve complex production processes. Prof. Jurković is currently involved in several scientific and professional projects to improve production efficiency and sustainability and to strengthen cooperation between the academic and industrial sectors. He is the author and co-author of numerous scientific and professional papers published in international journals and proceedings. He actively contributes, as a Guest Editor, to international journals in the field of mechanical engineering. His contribution has been recognized through awards for scientific and professional work as well as for mentoring students and young researchers.

David Ištoković

David Ištoković is an Associate Professor at the Faculty of Engineering, University of Rijeka, Croatia, where he works in the Department of Industrial Engineering and Management. He obtained his Ph.D. in Mechanical Engineering in 2021, with a research focus on production planning and scheduling. His scientific and professional activities focus on the digital transformation of production systems, the application of digital twins, Industry 4.0, and the modeling and optimization of production processes. Dr. Ištoković is currently involved in several scientific research and professional projects dealing with the simulation, visualization, and optimization of complex systems in production. His scientific interests include the application of advanced optimization algorithms, the digitalization of processes and the introduction of modern engineering tools into production and operational practice. He is the author and co-author of several scientific and professional papers and an active collaborator in industrial projects with the aim of transferring knowledge and technology from academia to industry, in order to optimize production and increase the competitiveness of companies.

Janez Gotlih

Janez Gotlih is an Assistant Professor at the Faculty of Mechanical Engineering, University of Maribor, Slovenia. He obtained his Ph.D. in Mechanical Engineering in 2020, with a research focus on robotic machining. He is actively engaged in student, industrial, and research projects at both national and international levels. Dr. Gotlih collaborates closely with academic and industrial partners and contributes to the wider scientific community through participation in professional societies and as a Guest Editor for international journals in the field of mechanical engineering. His research interests lie in the development of intelligent manufacturing systems, with a particular focus on robotics, the optimization of manufacturing processes, advanced automation, and the integration of artificial intelligence into production environments.

Preface

This reprint brings together a selection of peer-reviewed articles originally published in the Special Issue "Design and Optimization of Manufacturing Systems, 2nd Edition", which aimed to showcase state-of-the-art research and development in the field of modern manufacturing. The scope of the Special Issue encompassed diverse aspects of manufacturing system design, including intelligent scheduling, simulation-based optimization, artificial intelligence techniques, industrial and mobile robotics, sustainability, and human-centered approaches. The motivation behind this reprint was to provide a consolidated platform for disseminating innovative methodologies and practical solutions that support the transition toward more adaptive, efficient, and resilient manufacturing systems.

The reprint is intended for researchers, engineers, and practitioners working in manufacturing science, industrial engineering, operations research, and related disciplines. The included works reflect the efforts of an international community of authors committed to advancing knowledge and application in the field. As Guest Editors, we are grateful to all the contributing authors for their high-quality submissions and to the reviewers for their thoughtful evaluations. We also acknowledge the editorial team at *Applied Sciences* for their support throughout the publication process.

Zoran Jurković, David Ištoković, and Janez Gotlih
Guest Editors

applied sciences

MDPI

Article

Research on Quality Prediction for Thermal Printing Using a Particle Swarm Optimization with Back Propagation (PSO-BP) Neural Network

Chun-Ling Ho [1,*], Zhiyun Wu [2], Tung-Chiung Chang [3] and Shenjun Qi [2]

[1] Key Laboratory for Intelligent Infrastructure and Monitoring of Fujian Province (Huaqiao University), College of Civil Engineering, Huaqiao University, Xiamen 361021, China
[2] College of Civil Engineering, Huaqiao University, Xiamen 361021, China; 23024125023@stu.hqu.edu.cn (Z.W.); qisj972@163.com (S.Q.)
[3] Department of Civil Engineering and Geomatics, Cheng Shiu University, Kaohsiung 833301, Taiwan; 4946@gcloud.csu.edu.tw
* Correspondence: holling0712@gmail.com

Abstract: Thermal printing is a prevalent method due to its advantages of rapid output, cost effectiveness, and ease of use. However, the quality of thermal printing is influenced by the printing speed, the temperature, and the concentration and characteristics of the materials. This research employs a BP neural network to forecast print quality, utilizing two activation functions. The findings indicate that a dual-layer hidden configuration utilizing the GeLU activation function yields a lower root-mean-square error (RMSE). The optimal configuration identified consists of six neurons in the first hidden layer and three neurons in the second hidden layer. To enhance the predictive performance, a PSO algorithm was integrated with the PSO-BP model to refine the parameter selection, which included ambient temperature, printing speed, and printing concentration, with iterative training and validation conducted via the gradient descent algorithm. The PSO-BP network achieved an MAE of 0.1108, an RMSE of 0.145, an MSE of 0.021, and an R^2 value of 0.9916 in predicting print quality. These results substantiate the stability and reliability of the neural network model developed with the PSO algorithm. Further validation with ten sets of test samples demonstrated that the model attained an average absolute error of 2.77% in print quality predictions, indicating robust generalization capabilities and precise forecasting.

Keywords: thermal printing; quality prediction; optimization; BP neural network; particle swarm optimization (PSO)

1. Introduction

1.1. Traditional Printing Quality

Thermal barcode printing technology is an economical, fast, and environmentally friendly printing method that has been widely used in various fields, such as logistics, retail, and healthcare, in recent years. Its working principle involves heating thermal paper with a thermal print head, causing a chemical reaction in the thermal coating on the paper to form images or text. However, the quality of thermal printing is influenced by various factors, such as the temperature control of the print head, printing speed, and paper material [1]. Small changes in these factors can lead to fluctuations in printing quality, which, in turn, affect the readability of the barcode and the overall printing effect.

In practical applications, the readability of barcodes is directly related to the accurate transmission of logistics information and the efficiency of inventory management, while fluctuations in print quality may result in barcode scanners failing to correctly identify the barcodes, which, in turn, can trigger a series of chain reactions, such as logistics delays and inventory errors. Therefore, the instability of thermal barcode print quality is an urgent issue that needs to be addressed, and how to accurately predict and control the print quality of thermal barcodes has become a current research hotspot and challenge.

Conventional quality control methods in printing predominantly depend on empirical parameter adjustments and manual inspections [2]. These approaches are characterized by inefficiency and struggle to adapt to the complexities and variability in printing environments [3]. Consequently, the integration of contemporary artificial intelligence and machine learning technologies, particularly neural network-based predictive models, presents a promising avenue for addressing these challenges [4]. Neural networks possess robust nonlinear mapping capabilities and can analyze extensive datasets to predict and optimize printing parameters, thereby enhancing the stability and precision of printing quality [5].

1.2. The Quality Prediction of Printing with BP Neural Networks and PSO

Concerning pertinent studies and predictions, a back-propagation artificial neural network (BP neural network) model has been used to predict the performance of gel printing [2,6]. Therefore, a BP neural network is a multilayer feed-forward neural network that uses error back-propagation, which has the characteristics of a simple structure and strong fault tolerance [4], and a BP neural network also has powerful generalization and nonlinear mapping capabilities [7]. The methods focus on data normalization in order to improve the maximum speed limit, the inertia constant, and the fitness function by using a BP neural network for color prediction [8], predicting the flow of ink during 3D bioprinting [9], and using the BP neural network to optimize parameters on 3D printing [10]. To optimize the selection of parameters, it is common to combine particle swarm optimization (PSO) algorithms with BP networks to improve model performance [11–14]. For 3D printing, artificial neural networks (ANNs) are also applied in potential solutions [15,16], and both multiple regression analysis (MRA) and ANNs are used to analyze the influence of machining parameters [17].

1.3. The Optimal Parameters of Printing Quality Using Machine Learning

Other related prediction methods include the application of convolutional neural networks (CNNs) to predict printing errors 17 and an AI-driven approach to identify and forecast printing errors using CNNs and the Cox proportional hazards (CPH) approach [18]. Other research used learning algorithms to successfully predict the failure mode of printed thermal resistance [19,20] and to predict the relationships between measured color and four control factors by a multilayer perceptron (MLP) neural network model [5]. Another study also employed eight distinct classification algorithms, including K-nearest neighbors, support vector machines, and Gaussian naive Bayes, to assess print quality [21]. About the quality of printing, automated algorithms were adopted to analyze and compare differences in original and printed documents [22], real-time programming methods were chosen to detect the level of prints [23], shape and color classification algorithms were used to classify printing defects [24], and image segmentation and color image analysis techniques were proposed to detect the quality of printing [3]. By introducing advanced machine learning algorithms, such as support vector machine (SVM) and Random Forest, the high-precision classification of different printing materials and technologies was achieved using printing pattern features, demonstrating more than 92% classification accuracy [25]. Another study

also used machine learning with optimal input parameters [26]. In terms of printing parameter adjustment, a fast self-calibration method for laser displacement sensors was proposed to complete parameter calibration at one time [27], and, in addition, iterative function calculations were found to improve the measurement accuracy of printing [28]. Expert experience was also used to design and formulate multiple sets of parameters to analyze the rules of influence [29,30]. Regarding 3D printing, some studies applied machine learning for the determination of optimal parameters. Using metadata and video data from the 3D printing process, a unique image dataset was constructed, and a regression model was trained to accurately predict and adjust the flow of printing material [31].

1.4. The Purpose of This Research

To enhance the accuracy of models within a constrained data range, this research employs the well-established PSO-BP to leverage the distinctive benefits for the optimization of thermal parameters via gradient descent while also considering factors such as the volume of training data and the interpretability of parameters. The objective is to facilitate the effective adjustment of parameters in industrial printers, thereby ensuring their operational availability and supporting informed decision making. Therefore, this article discusses the prediction effect of the BP neural network on printing parameters and uses a simple and fast solution PSO algorithm to solve the optimal weight and threshold of the BP neural network, thus improving the convergence speed and optimization accuracy of the BP neural network.

2. The Principle of Printing and Key Parameters

Hot barcode printing is a method for generating barcodes by heating. When the print head is heated, the thermal coating reacts chemically to display the barcode information. Thermal printing primarily involves controlling the thermal print head to generate heat, thereby heating the thermally sensitive medium to enable information printing. The core of the thermal printer is the thermal print head, which is an electronic component that uses Joule heating effects to print information. As shown in Figure 1, it is mainly composed of resistors, ceramic substrates, driver chips, circuit boards, socket interfaces, radiators, and other components. The quality of barcode printing is influenced by several key parameters, including print head temperature, print speed, print density, and print quality.

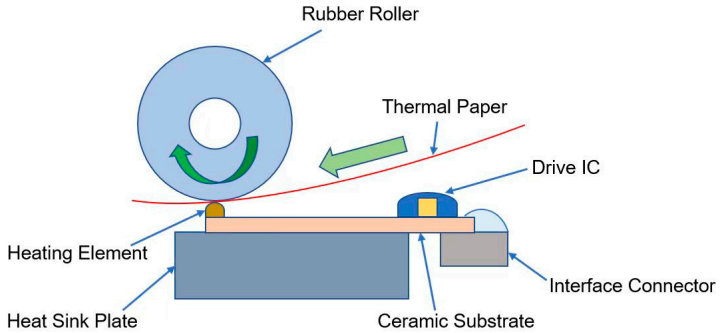
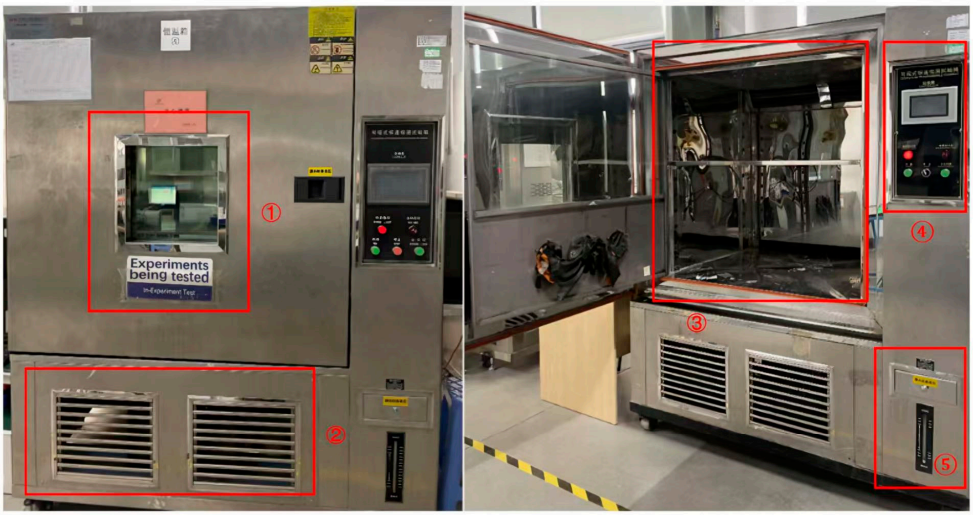


Figure 1. The principle of thermal printing.

The printer firmware manages the printing speed by adjusting the rotational speed of the stepping motor, thereby allowing for various speed settings. Additionally, the firmware controls the printing concentration by regulating the heating duration of the printer head to achieve different levels of concentration. Temperature is modeled to reflect the ambient temperature of the natural environment.

The temperature of the print head plays a vital role in determining the quality of the thermal barcode printing, as shown in Figure 2. If the temperature is too high or too low, unsatisfactory printing results can be obtained, including unclear or damaged barcodes. Different countries and regions experience diverse climate conditions, which can notably affect printing quality, particularly in colder environments. In this study, the temperature examined variable pertains to the ambient temperature, which was simulated in a temperature-controlled chamber to mimic the actual conditions. Print density affects the darkness of barcodes; if the density is too high, the barcode may become overly dark and difficult to scan, whereas if it is too low, the barcode might be too light and prone to blurring. Furthermore, the type of thermal paper used and the conditions under which it is stored can affect the print density of thermal barcodes. This density is mainly adjusted through software that offers 15 different settings.



① Observation window
② Heat dissipation port
③ The printer is placed at chamber, and the printing paper is placed inside the printer.
④ Temperature controller
⑤ Water level filling port and water level indicator

Figure 2. Temperature-controlled chamber.

Printing at excessively high speeds can lead to inadequate contact time between the print head and thermal paper, which can negatively impact the print quality. However, printing too slowly may cause the print head to overheat, damaging the thermal coating. The speed variable discussed in this study was categorized into nine levels. The primary equipment used for testing barcodes is the WebScan Barcode Verifier (Barcode Detection Scanning Sensor, Model USB-DC71), and its accompanying barcode detection software is the Barcode Recognition Prediction System V1.0 (1D and 2D Barcode Recognition and Detection). The specific evaluation metrics of the scan reflectance profile offer important insights into quality control in the barcode printing process. Consequently, when assessing the quality grades of the barcode symbols, such as grades A, B, C, D, and E, we can compute the average values of the individual evaluation metrics from each scan reflectance profile and determine the arithmetic mean of the grades for each parameter across the profiles. To enhance information regarding the quality of the barcode symbol, the average values of these evaluation metrics or their grade averages can be included in the test report data, as illustrated in Table 1 and Figure 3.

Table 1. The grades of letters (ANSI) corresponding to the level of the numeric symbol.

Level of Numeric Symbol	Grade of Letter
[3.5–4.0)	A
[2.5–3.5)	B
[1.5–2.5)	C
[0.5–1.5)	D
Less than 0.5	E

Figure 3. The quality of printed test results.

3. Methodology

3.1. The Experimental Technique

This research employed an experimental technique known as Latin Hypercube Sampling (LHS) to train and validate the data required for the BP neural network. The primary operational features of the LHS in this study consisted of two main steps: random sampling and probability distribution reconstruction. Random sampling involves selecting samples randomly from a multivariate parameter distribution to analyze a small yet representative set of samples. Probability distribution reconstruction is the process of creating a probability distribution from a limited number of samples, thereby yielding a representative dataset. This comprehensive experimental study comprised seven phases, ranging from data preparation to evaluating the performance of the models, as illustrated in Figure 4.

3.2. Experimental Samples

The prediction of thermal printing quality is a common situation involving a limited number of samples and many dimensions. Considering factors such as production costs, equipment lifespan management, and time efficiency, this study employed simplified parameters to improve sample representativeness. To ensure comprehensive feature coverage, an enhanced Latin Hypercube Sampling (LHS) method was utilized to select 30 key samples with the aim of achieving the best balance between data volume, cost, and accuracy.

Random sampling was approximated from the multivariate parameter distribution using LHS, and the probability distribution was re-established with fewer samples. A set of thirty data points was selected as the sample, including the printing speed, printing

concentration, printing temperature, and printing quality. The experimental samples are listed in Table 2. Then, these data were used to train and improve the BP neural network model to verify its predictive performance.

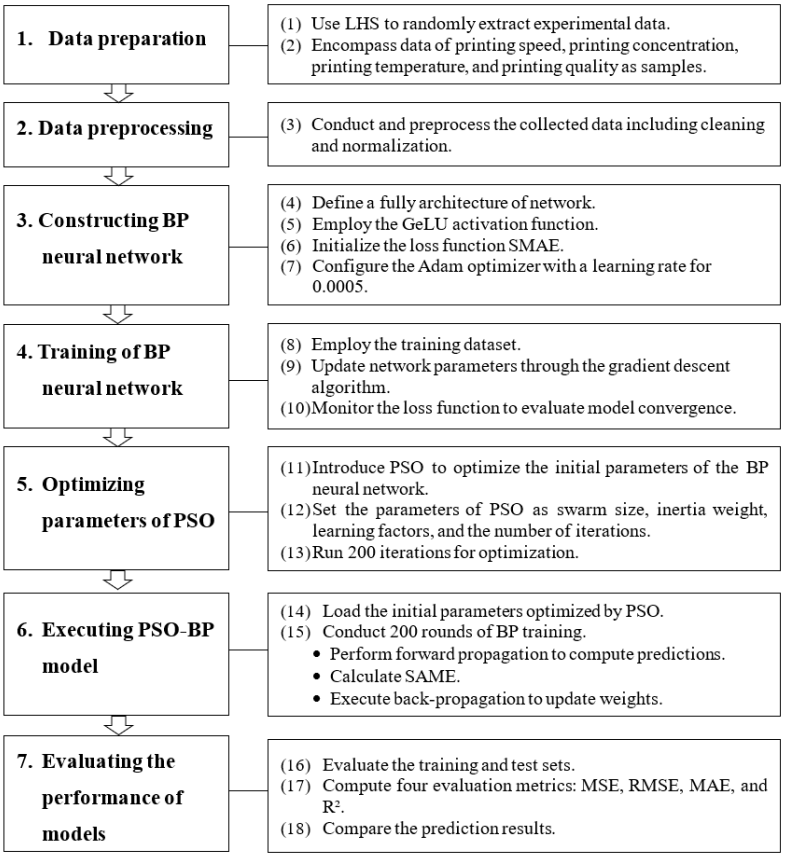


Figure 4. The experimental process of the PSO-BP model.

Table 2. The data of the samples.

No. Sample	Printing Temperature (°C)	Printing Speed (Index)	Printing Concentration (Index)	Printing Quality (Index)
1	25	1	1	0.1
2	25	1	2	0.2
3	25	1	3	0.7
4	25	1	4	2.2
5	25	1	5	2.7
6	25	1	6	2.9
7	25	1	7	3.5
8	25	1	8	3.7
9	25	1	9	3.8
10	25	1	10	3.9
11	25	1	11	3.9
12	25	1	12	4
13	25	1	13	4
14	25	1	14	3.9
15	25	1	15	3.9
16	−25	1	1	0
17	−25	1	2	0
18	−25	1	3	0

Table 2. *Cont.*

No. Sample	Printing Temperature (°C)	Printing Speed (Index)	Printing Concentration (Index)	Printing Quality (Index)
19	−25	1	4	0
20	−25	1	5	2.2
21	−25	1	6	2.3
22	−25	1	7	3.1
23	−25	1	8	3.2
24	−25	1	9	3.3
25	−25	1	10	3.7
26	−25	1	11	3.7
27	−25	1	12	3.8
28	−25	1	13	3.8
29	−25	1	14	3.7
30	−25	1	15	3.7

3.3. Modeling of the BP Neural Network

This study focused on enhancing a BP neural network by employing suitable optimization techniques. These include modifying the network architecture (e.g., adding or removing hidden layers and changing the number of nodes), refining learning algorithms (e.g., incorporating momentum terms, adjusting the learning rate, and selecting different optimizers), and choosing or creating more effective activation and loss functions. The goal was to boost the generalization capability of the model. By evaluating and analyzing various aspects, such as prediction accuracy, stability, and computational efficiency, we aimed to demonstrate the benefits of the improved BP neural network in predicting the quality of thermal barcode printing.

In this study, a neural network model was built in PyTorch 2.7 to accurately forecast printing quality by inputting essential parameters, such as temperature, printing speed, and printing concentration. Initially, it efficiently loaded data from a CSV file and conducted thorough preprocessing to ensure the quality and consistency of the data. The data were then transformed into a PyTorch tensor, and a data loader was established to facilitate an efficient data stream for training and evaluating the model.

In the construction phase of the model, a neural network featuring two fully connected layers was created, as shown in Figure 5 and Table 3, employing the GeLU activation function to improve the model's ability to express nonlinear relationships. The loss function was based on the Standardized Mean Absolute Error (SAME) average absolute error to maintain the sensitivity to prediction errors during training. The Adam optimizer was chosen to enhance the model convergence through its effective parameter update method. Throughout the training process, the model underwent continuous iterations, adjusting weights to minimize the loss function, outputting the training loss by 10 iterations, and tracking the learning progress.

Table 3. Parameters of the BP neural network.

Parameters	Value
Number of input layer nodes	3
Number of nodes in the 1st hidden layer	6
Number of nodes in the 2nd hidden layer	3
Number of output layer nodes	1
Activation function	GeLU

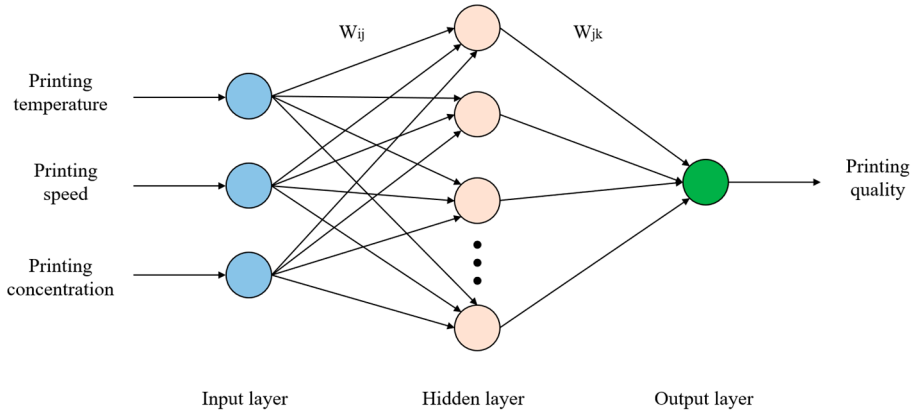


Figure 5. Printing quality with the BP neural network.

During the evaluation stage, the model's performance was assessed using a test dataset, and the average test loss was computed to measure the generalization ability of the model. Finally, predictions were made using new data, and the root-mean-square error (RMSE) was calculated to objectively evaluate the accuracy of the model's predictions. The entire process represents a comprehensive BP neural network workflow encompassing data processing, model construction, training, evaluation, and prediction. The main steps include the following:

(1) Data reading and preprocessing.
(2) Conversion of data into a PyTorch tensor and creation of a data loader.
(3) Definition and initialization of neural network models.
(4) Specification of the loss function and optimizer.
(5) Model training.
(6) Model evaluation.
(7) Prediction of new data and RMSE calculation.

3.4. Optimization of the PSO-BP Model

BP neural networks are highly influenced by the choice of initial weights and biases. If initialization is not performed correctly, it can result in reduced prediction accuracy. To enhance the prediction accuracy, a particle swarm optimization (PSO) algorithm was introduced to optimize the parameter selection of the BP neural network, thereby improving the model's predictive capabilities. Additionally, the GeLU function was utilized as the activation function in the hidden layer, whereas the loss function employed the SAME average smoothing absolute error. By optimizing the BP neural network with PSO, we explored the weight and bias search space more effectively, leading to better solutions and improved performance and generalization of the neural network. This approach helped avoid local optima and aimed to determine global or near-global optimal weight configurations.

The fitness function used in the PSO algorithm aligns with the loss function of the BP neural network, utilizing both mean square error (MSE) functions. This article outlines the process of training a BP neural network using the PSO algorithm to optimize its initial parameter settings, which includes the following steps.

(1) Data preparation: Split the dataset into training and testing sets and normalize the samples.
(2) Network architecture setup: Determine the number of layers and nodes in the hidden layers of the BP neural network.

(3) PSO optimization of initial parameters: Use the MSE regression loss function as the fitness function for the PSO algorithm to iteratively update and obtain the initial weights and biases for the neurons in each layer of the BP neural network.
(4) Training the BP neural network: Input the training set into the BP neural network for training while calculating the fitness value and historical best position for each particle.
(5) Parameter updating and model output: Incorporate the parameters optimized by the PSO algorithm into the neural network, apply the gradient descent algorithm to update the network parameters, and ultimately produce the optimal network model.

4. Results

4.1. The Framework of the BP Neural Network

This study aimed to assess model performance in the context of network structure by utilizing the sigmoid activation function and examining the configurations of single-layer and double-layer hidden layers. In the case of the single-layer hidden layer, the model achieved its lowest root-mean-square error when configured with six neurons, indicating optimal prediction accuracy, as illustrated in Table 4. Conversely, the double-layer hidden-layer configuration generally resulted in a reduced root-mean-square error. For the task of predicting printing quality, the most effective configuration was identified as having 6 neurons in the first hidden layer and 10 neurons in the second hidden layer. This structural modification led to a significant enhancement in the model performance, as detailed in Table 5.

Table 4. The printing quality of the hidden layer with a single layer (using the sigmoid activation function).

Number of Neurons	1	2	3	4	5	6	7	8	9	10
	1.5728	1.5673	1.596	1.5695	1.5728	1.5831	1.5834	1.5794	1.6098	1.583
	1.617	1.5868	1.6071	1.5928	1.5799	1.5757	1.5799	1.5743	1.578	1.5924
RMSE	1.5549	1.5843	1.6055	1.5667	1.5727	1.5825	1.5866	1.5832	1.5791	1.5891
	1.6043	1.6135	1.5849	1.5976	1.5903	1.575	1.5752	1.6022	1.5948	1.5971
	1.5998	1.6235	1.59	1.5992	1.5945	1.5816	1.5947	1.5800	1.5807	1.593
Average	1.5898	1.5951	1.5967	1.5852	1.582	1.5796	1.584	1.5838	1.5885	1.5909

Table 5. The printing quality of the hidden layer with a double layer (using the sigmoid activation function).

Number of Neurons	1	2	3	4	5	6	7	8	9	10
	0.5944	0.6968	0.5893	0.5772	0.5773	0.6113	0.5953	0.5991	0.634	0.548
	0.6342	0.659	0.6292	0.6898	0.5481	0.6293	0.6028	0.6308	0.5627	0.5267
RMSE	0.6457	0.601	0.6278	0.6334	0.5303	0.5732	0.5823	0.5299	0.5512	0.5843
	0.5713	0.5977	0.6674	0.5405	0.6267	0.5727	0.6125	0.6486	0.561	0.5179
	0.6161	0.614	0.6357	0.6880	0.6232	0.6226	0.5951	0.5532	0.6433	0.5081
Average	0.6123	0.6337	0.6299	0.6258	0.5811	0.6018	0.5976	0.5923	0.5904	0.5370

The Gaussian error linear unit (GeLU) activation function offers significant advantages over the sigmoid activation function in terms of smoothness, ability to express nonlinearity, computational efficiency, resource utilization, generalization, and applicability. These benefits have led GeLU to become a popular choice in contemporary deep learning. This study aimed to investigate more optimized activation functions by comparing them with the sigmoid function. When using GeLU as the activation function in a hidden layer with six neurons, the model achieved the lowest root-mean-square error and highest prediction

accuracy. In comparison to a model with a sigmoid activation function in a two-layer hidden layer setup, the GeLU model demonstrated superior prediction accuracy, as shown in Table 6. In addition, in a two-layer hidden layer configuration utilizing GeLU, the model consistently exhibited lower root-mean-square error values, as indicated in Table 7. This research identified that the best configuration for predicting printing quality consists of six neurons in the first hidden layer and three neurons in the second hidden layer. Furthermore, this study examined three-layer configurations to analyze the arrangement of hidden layer neurons. The findings revealed that the three-layer hidden layer model suffers from overfitting and issues with gradient disappearance or explosion, leading to subpar performance. Consequently, we chose a two-layer hidden layer model architecture for the experiments. Tables 4–7 provide detailed information regarding the number of hidden layers and activation functions. Figure 6 illustrates the accuracy across different parameters, with the hidden layer of the double layer using GeLU yielding the best results.

Table 6. The printing quality of the hidden layer with a single layer (using the GeLU activation function).

Number of Neurons	1	2	3	4	5	6	7	8	9	10
	0.2789	0.2467	0.2157	0.2601	0.3174	0.2413	0.3102	0.2974	0.2246	0.3401
	0.3147	0.2501	0.236	0.2089	0.2054	0.2817	0.3354	0.2573	0.3413	0.3017
RMSE	0.2284	0.2011	0.2744	0.2422	0.2622	0.205	0.3212	0.3011	0.2987	0.2988
	0.3056	0.269	0.2563	0.2563	0.2543	0.2652	0.2687	0.2411	0.2655	0.2341
	0.3121	0.3003	0.2822	0.2516	0.2931	0.2733	0.2914	0.2831	0.3199	0.2561
Average	0.2879	0.2534	0.2529	0.2438	0.2665	0.2533	0.3054	0.276	0.29	0.2992

Table 7. The printing quality of the hidden layer with a double layer (using the GeLU activation function).

Number of Neurons	1	2	3	4	5	6	7	8	9	10
	0.1512	0.2119	0.1166	0.1211	0.1214	0.1321	0.1124	0.1311	0.1543	0.1917
	0.1193	0.1792	0.1022	0.0981	0.1143	0.1298	0.1346	0.1456	0.1672	0.1876
RMSE	0.1283	0.1273	0.1232	0.0966	0.1052	0.1143	0.1532	0.1515	0.1811	0.1654
	0.1537	0.1345	0.1561	0.1073	0.1543	0.1515	0.1321	0.1321	0.1765	0.1346
	0.1394	0.1294	0.1313	0.0954	0.1176	0.0991	0.1264	0.1129	0.1298	0.1576
Average	0.1384	0.1565	0.1259	0.1037	0.1226	0.1254	0.1317	0.1346	0.1618	0.1674

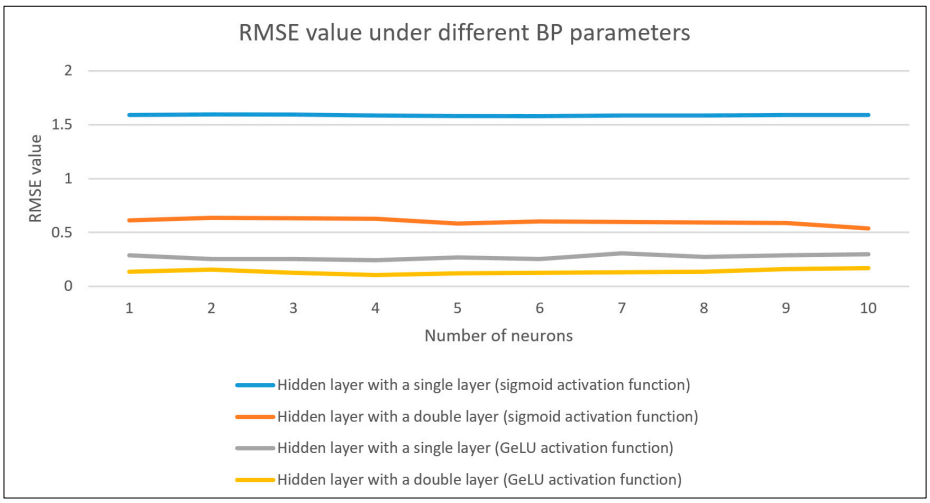


Figure 6. The value of RMSE with different BP parameters.

The ideal fitting performance of printing quality is shown in Figure 7, where the target value perfectly matches the predicted value. The solid line represents the optimal fitting line for practical use, while the R value indicates the correlation between the output and predicted values. A value closer to 1 signifies a better fitting performance. The blue dots represent the actual and predicted values, whereas the red line denotes the ideally matched trend line. The proximity of these points to the red line suggests a strong alignment between the actual and predicted values. The notation "Test: R = 0.9477" indicates that the correlation coefficient (R value) for the test set is 0.9477, which is close to 1, signifying a stronger correlation between the predicted and the actual values and reflecting improved model performance.

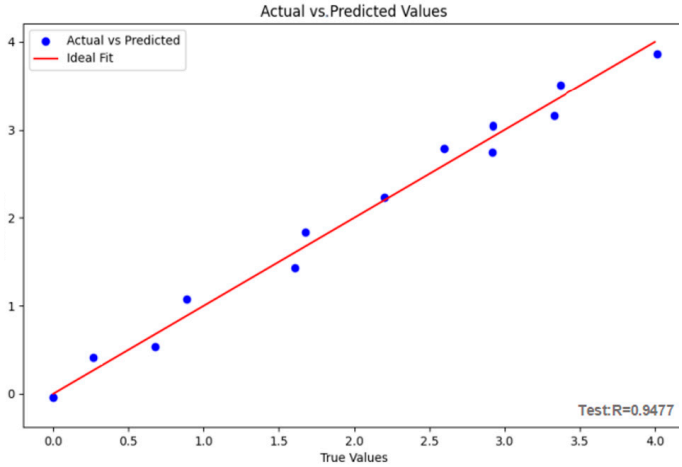


Figure 7. The fitting effect of printing quality.

This relationship is illustrated in Figure 8, which depicts the effects of various parameters on the overall printing quality. This 3D diagram highlights the intricate connections among temperature, velocity, concentration, and mass, using different colors to represent specific numerical values related to print quality. As the temperature, speed, and concentration fluctuate, the quality also demonstrates corresponding trends, providing valuable insights for parameter optimization and enhancing quality.

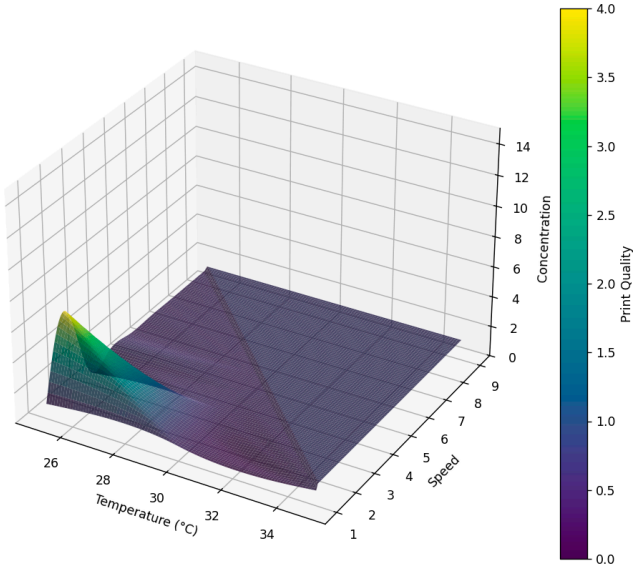


Figure 8. Printing quality generated by different parameters.

The printing quality assessment demonstrates strong results with the optimal network structure, further validating the model's predictive accuracy. Moreover, when the ambient temperature is held constant, the impact of printing speed on quality is minimal, while the effect of the printing concentration on printing quality exhibits considerable variability. Conversely, when the printing speed is constant, the influences of the printing concentration and ambient temperature on quality become more pronounced. Additionally, under consistent conditions of ambient temperature and printing concentration, printing quality tends to decline as the printing speed increases.

4.2. Training and Prediction of the PSO-BP Model

The quality prediction model was constructed using a PSO-BP neural network architecture. This model incorporates input variables such as ambient temperature, printing speed, and printing concentration, with printing quality serving as the output variable. The model undergoes iterative training facilitated by the gradient descent algorithm, followed by validation. The parameters for the PSO were established as follows: a population size of 50, an inertia weight (ω) of 0.8, and learning factors (c1 and c2) both set at 2.0, with a total of 100 iterations. Upon completion of parameter configuration, 60 datasets were selected from a single experimental group, with 50 sets designated as training samples and 10 sets allocated for testing the model's generalization performance. Figures 9 and 10 illustrate the comparison between the actual values of the training samples and the corresponding predicted values.

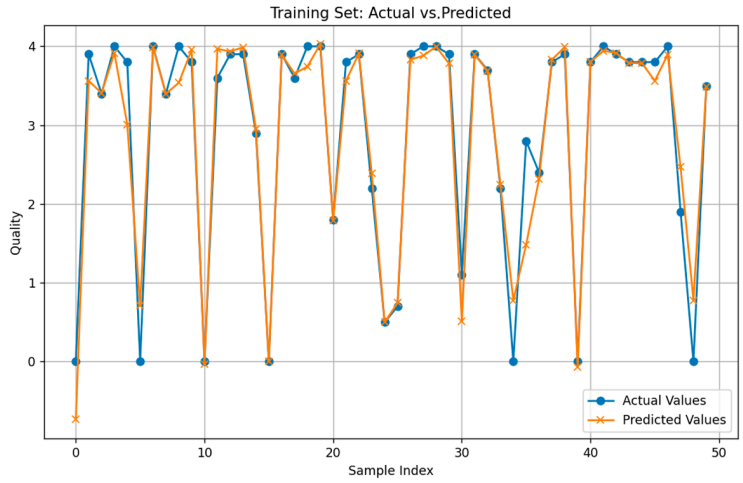


Figure 9. The training of the PSO-BP model for printing quality.

Figure 9 displays the training outcomes of the PSO-BP model for predicting quality. This research indicates that the predicted values are generally closer to the actual values at most points, demonstrating that the model's strong predictive capability on the training dataset. Figure 10 illustrates how the PSO-BP model predicts print quality on the test dataset, with actual and predicted values closely aligning in certain samples. This alignment suggests that the model provides more accurate predictions at those points.

4.3. The Optimization of Models with Prediction

This study employed four distinct metrics to conduct a comprehensive assessment of the model's predictive accuracy, namely, the R^2 (coefficient of determination), MSE, RMSE, and Mean Absolute Error (MAE), as detailed in Table 8. Regarding the prediction of printing quality, the MAE is recorded at 0.1108, the RMSE at 0.145, the MSE at 0.021,

and the R^2 at 0.9916. The outcomes of the network training substantiate the stability and reliability of the neural network model developed using the PSO algorithm.

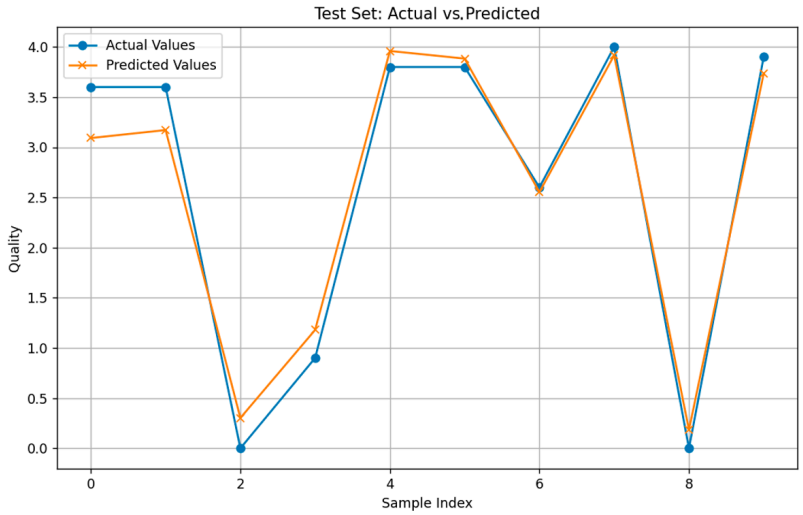


Figure 10. The prediction of the PSO-BP model for printing quality.

To evaluate the predictive efficacy of the PSO-BP neural network model, this study selected ten sets of test samples for experimentation and validation. The experimental data, illustrated in Figure 11, indicate that the model attained an average absolute error of 2.77% in predicting printing quality. These findings substantiate the capability of the developed PSO-BP model to effectively forecast the nonlinear relationships among various influencing parameters on printing quality. The experimental results demonstrate that the model not only possesses a high degree of predictive accuracy but also exhibits robust generalization capabilities, allowing it to accurately predict printing quality.

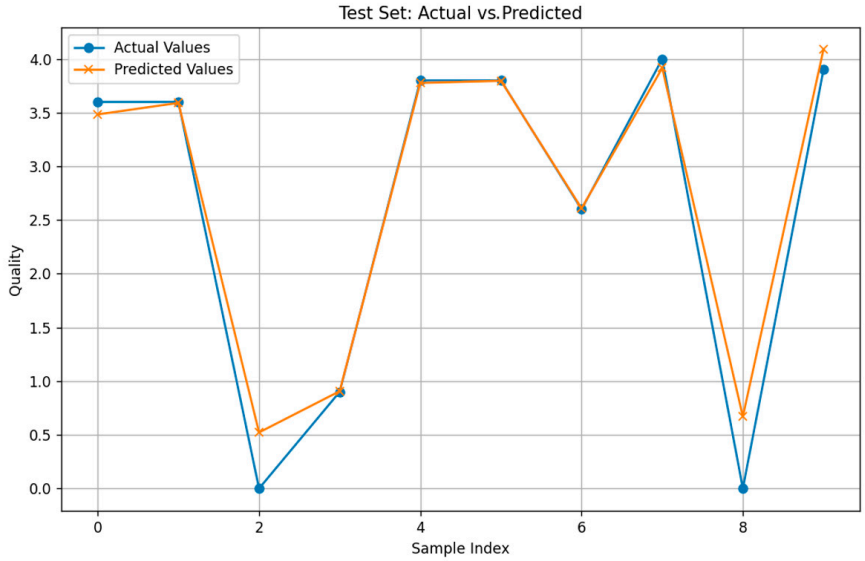


Figure 11. Ten sets of data for experiments and verification.

Table 8. The result of the PSO-BP model.

Evaluation Index	Value
MAE	0.1108
RMSE	0.145
MSE	0.021
R^2	0.9916

4.4. Summary

Regarding comparisons with various algorithms, considering the actual needs and feasibility of the current situation, we will expand this section in future research to gain a deeper understanding of the advantages of various algorithms. First, this study used traditional BP for analysis, followed by comparison with BP-PSO to enhance the model's accuracy. The accuracy comparison of BP-PSO is shown in Figure 12. The findings indicate that the BP-PSO model is more suitable for predicting printing quality.

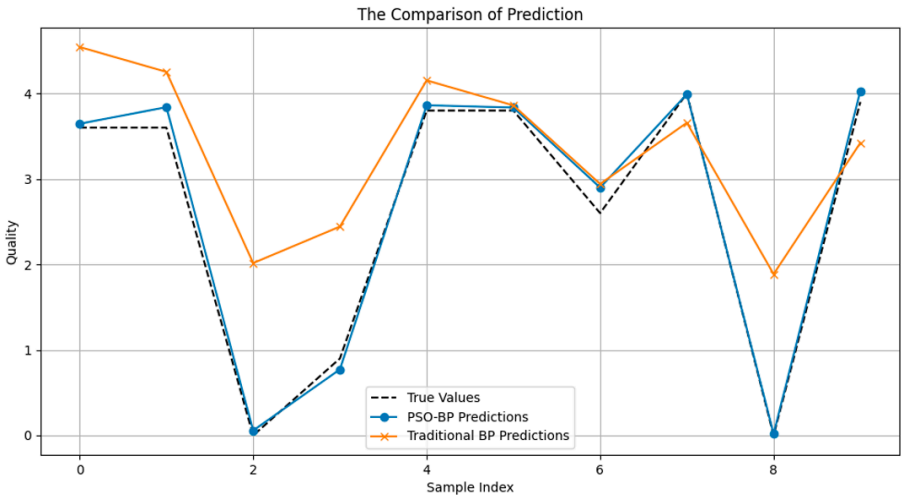


Figure 12. A comparison of the accuracy between BP and BP-PSO.

This research indicates that the PSO-BP model for thermal barcode printing demonstrates significant advantages in parameter application within specific domains, as well as the practical viability of the algorithms in real-world scenarios. In terms of parameter application in specialized fields, the PSO-BP model exhibits the ability to adapt to dynamic thermal fluctuations in real time and is compatible with hardware suited for existing production environments. Furthermore, it shows robust convergence even with limited training data, a common characteristic in industrial settings. The model's capabilities minimize the response time, reduce the carbon footprint of production, and facilitate interpretable parameter adjustments, such as production debugging. This will enable the rapid establishment of an optimized printing environment.

In terms of algorithmic feasibility, the PSO approach requires only an objective function (fractional values that reflect printing quality) and does not depend on precise mathematical models. This method facilitates the implementation of predictive conditions that are crucial for real-world industrial applications, especially in situations where accurately simulating thermal printing parameters is challenging. The penalty function of the model can directly incorporate maximum temperature and pressure limits, allowing for flexible constraint management without the need for reprocessing. Furthermore, the concurrent investigation of multiple parameter combinations by various particles can significantly reduce the time

needed for experimental trials and lead to a faster convergence rate compared to sequential techniques, such as grid search.

Following the training of the model and its subsequent validation in thermal printing, the PSO-BP model introduced in this study demonstrates significant applicability and practical importance. Consequently, this paper presents three key findings.

(1) Rapid debugging and optimization settings.

In the stage of production and commissioning for thermal printing, the PSO-BP model can efficiently determine the optimal combination of printing parameters. By inputting various factors, such as printing speed, temperature, and concentration, the model can predict the resulting printing quality. This assists operators in quickly identifying parameter configurations that meet standards of quality, reducing debugging time, and enhancing overall production efficiency.

(2) Real-time observation and feedback.

While thermal printers are in operation, the PSO-BP model can be employed to monitor print quality in real time. By consistently gathering data on printing parameters and print quality, the model can make real-time adjustments to these parameters to maintain high printing quality. Additionally, the model provides immediate feedback on printing quality, assisting operators in identifying and addressing potential quality issues.

(3) Quality control and assessment.

The PSO-BP model is applicable in the quality control of thermal printers for assessing the quality of printed barcodes. By inputting data on various influencing factors, it can effectively predict the printing quality of the barcode and assign an appropriate quality rating. This enables production staff to quickly identify and resolve quality issues, thereby ensuring product reliability and consistency while enhancing customer satisfaction.

(4) Dynamic Parameter Optimization.

As production settings and material properties change, the parameters of thermal printers must be adjusted accordingly. The PSO-BP model can re-optimize these parameters in response to new production conditions and material characteristics, ensuring that printing quality consistently remains at an optimal level. This ability enables thermal printers to adapt to various complex and changing environments, thereby maintaining stable printing quality.

In summary, the PSO-BP model has extensive applications in thermal printing. It enhances the accuracy of predicted printing quality and production efficiency while also offering strong support for quality control and advancements within the printing industry.

5. Conclusions

Thermal-sensitive barcodes are widely used in areas such as commodity distribution and logistics management, where the quality of their printing significantly impacts information accuracy and circulation efficiency. Improving prediction accuracy and refining the printing process can help minimize errors and losses associated with barcode quality issues, thereby enhancing overall operational efficiency. As neural network technology continues to advance, research focused on predicting the printing quality of thermal barcodes through advanced PSO-BP neural networks aligns with technological trends and provides innovative solutions to the practical challenges of barcode printing quality. This study investigates the structure, algorithms, and parameter designs of BP neural networks to conduct comprehensive experiments aimed at achieving optimal results under various

conditions. The model considers multiple influencing factors to accurately forecast the printing quality of thermal-sensitive barcodes, which can lead to increased production efficiency and higher customer satisfaction. Furthermore, this research can serve as a valuable reference for the thermal printing industry regarding material selection and equipment maintenance strategies.

This research demonstrates promising outcomes in predicting printing quality by using BP-PSO. However, the inclusion of data volumes and multiple parameters presents some challenges for industrial applications. In the future, we will aim to identify superior models to expand the range of data collected from manufacturers, including substrate materials and ambient humidity. The goal is not only to enhance but also to optimize energy efficiency and minimize material waste. Consequently, our research will focus on multi-objective algorithms to derive production parameters that align with low-carbon and energy-saving principles, making it more beneficial for industrial applications.

Author Contributions: Conceptualization, C.-L.H.; methodology, Z.W. and S.Q.; validation, C.-L.H. and T.-C.C.; formal analysis, Z.W.; resources, Z.W.; data curation, T.-C.C. and S.Q.; writing—original draft preparation, Z.W.; writing—review and editing, C.-L.H. and S.Q.; visualization, T.-C.C.; supervision, C.-L.H.; funding acquisition, C.-L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xu, J.; Ma, Y.; Han, Z.; Wang, Q.; Ma, T. Thermal design of printed circuit heat exchanger used for lead-bismuth fast reactor. *Appl. Therm. Eng.* **2023**, *226*, 120343. [CrossRef]
2. Jiao, X.; Ren, G.; Law, C.L.; Li, L.; Cao, W.; Luo, Z.; Pan, L.; Duan, X.; Chen, J.; Liu, W. Novel strategy for optimizing of corn starch-based ink food 3D printing process: Printability prediction based on BP-ANN model. *Int. J. Biol. Macromol.* **2024**, *276*, 133921. [CrossRef] [PubMed]
3. Bian, Y.; Zhong, Z.; Fu, C. Research of on-line printing quality detection based on mathematical morphology. *Packag. J.* **2013**, *5*, 35–39. [CrossRef]
4. Sun, H.W.; Huang, Y.C.; Tang, T.H.; Shen, J.; Gu, L.J.; Wang, J.S. A prediction model for the replication quality of nanoimprinting patterns based on BP neural network. *Microsyst. Technol.* **2025**. [CrossRef]
5. Wei, X.J.; Zou, N.; Li, Z. PolyJet 3D printing: Predicting color by multilayer perceptron neural network. *Ann. 3D Print. Med.* **2022**, *5*, 100049. [CrossRef]
6. Bao, Y.; Li, L.; Chen, J.; Weiwei Cao Liu, W.; Ren, G.; Luo, Z.; Pan, L.; Duan, X. Prediction of 3D printability and rheological properties of different pineapple gel formulations based on LF-NMR. *Food Chem. X* **2024**, *24*, 101906. [CrossRef] [PubMed]
7. Jun, J. A train F-TR lock anti-lifting detection method based on improved BP neural network Journal of Measurements in Engineering. *J. Meas. Eng.* **2024**, *12*, 149–161.
8. Miao, H.T.; Zhang, L.Z. The color characteristic model based on optimized BP neural network. In *Advanced Graphic Communications, Packaging Technology and Materials*; Springer: Singapore, 2016; Volume 369, pp. 55–63.
9. Wu, X.; Wang, S. Prediction of ink flow for 3D bioprinting of tubular tissue based on a back propagation neural network. *J. Comput. Methods Sci. Eng.* **2023**, *23*, 3071–3080. [CrossRef]
10. Li, Y.; Ding, F.; Tian, W. Optimization of 3D printing parameters on deformation by BP neural network algorithm. *Metals* **2022**, *12*, 1559. [CrossRef]
11. Lin, Z.; Fan, Y.; Tan, J.; Li, Z.; Yang, P.; Wang, H.; Duan, W. Tool wear prediction based on XGBoost feature selection combined with PSO-BP network. *Sci. Rep.* **2025**, *15*, 3096. [CrossRef]

12. Liu, Y.; Fan, L.; Wang, L. Urban virtual environment landscape design and system based on PSO-BP neural network. *Sci. Rep.* **2024**, *14*, 13747. [CrossRef] [PubMed]
13. Liu, L.; Yang, D.; Chen, R.J.; Li, K.; Zhou, X. Research on quality prediction of POI based on KPCA-GA-BP neural network. *J. Telecommun. Sci.* **2023**, *39*, 108–116.
14. Wei, H.; Tang, L.; Qin, H.; Wang, H.; Chen, C.; Li, Y.; Wang, C. Optimizing FDM 3D printing parameters for improved tensile strength using the Takagi–Sugeno fuzzy neural network. *Mater. Today Commun.* **2024**, *38*, 108268. [CrossRef]
15. Mahmood, M.A.; Visan, A.I.; Ristoscu, C.; Mihailescu, I.N. Artificial neural network algorithms for 3D printing. *Materials* **2020**, *14*, 163. [CrossRef]
16. Janmanee, P.; Ratchapakdee, P. Comparison of artificial neural network and response surface methodology in predicting the tensile strength and optimization of 3D printed objects. *Eng. Appl. Sci. Res.* **2024**, *51*, 704–715.
17. Djurović, S.; Lazarević, D.; Ćirković, B.; Mišić, M.; Ivković, M.; Stojčetović, B.; Petković, M.; Ašonja, A. Modeling and prediction of surface roughness in hybrid manufacturing–milling after FDM using artificial neural networks. *Appl. Sci.* **2024**, *14*, 5980. [CrossRef]
18. Bauer, M.; Augenstein, C.; Schäfer, M.; Theile, O. Artificial Intelligence in Laser Powder Bed Fusion Procedures–Neural Networks for Live-Detection and Forecasting of Printing Failures. *Procedia CIRP* **2022**, *107*, 1367–1372. [CrossRef]
19. Phillip, L. Predicting the Failure of a Thermal Print-Head Resistor. Ph.D. Thesis, University of Portsmouth, Portsmouth, UK, 2019.
20. Rafieyan, S.; Ansari, E.; Vasheghani-Farahani, E. A practical machine learning approach for predicting the quality of 3D (bio)printed scaffolds. *Biofabrication* **2024**, *16*, 045014. [CrossRef]
21. Olowe, M.; Ogunsanya, M.; Best; Hanif, Y.; Bajaj, S.; Vakkalagadda, V.; Fatoki, O.; Desai, S. Spectral features analysis for print quality prediction in additive manufacturing: An acoustics-based approach. *Sensors* **2024**, *24*, 4864. [CrossRef]
22. Liu, B.; Chen, Y.; Xie, J.; Chen, B. Industrial printing image defect detection using multi-edge feature fusion algorithm. *Complexity* **2021**, *2021*, 036466. [CrossRef]
23. Bai, W.G.; Qiu, C.L.; Zhang, J. Research on evaluation and detection method of print quality based on the CCD information. *Appl. Mech. Mater.* **2013**, *303–306*, 543–549. [CrossRef]
24. Karina, Z.; Hoeppener, S. Schubert, Inkjet printing and 3D printing strategies for biosensing, analytical, and diagnostic applications. *Adv. Mater.* **2022**, *34*, 2105015.
25. Nguyen, Q.T.; Mai, A.; Chagas, L.; Nadège, R.B. Microscopic printing analysis and application for classification of source printer. *Comput. Secur.* **2021**, *108*, 102320. [CrossRef]
26. Ratnavel, R.; Viswanath, S.; Subramanian, J.; Selvaraj, V.K.; Prahasam, V.; Siddharth, S. Predicting the optimal input parameters for the desired print quality using machine learning. *Micromachines* **2022**, *13*, 2231. [CrossRef] [PubMed]
27. Qin, Y.; Kang, R.; Sun, J.S. A fast self-calibration method of line laser sensors for on-machine measurement of honeycomb cores. *Opt. Lasers Eng.* **2022**, *152*, 106981. [CrossRef]
28. Jie, L.; Cai, Z.Q.; Yao, B. Error compensation and accuracy analysis of laser measurement system based on laser-beam calibration. *Optik* **2020**, *200*, 163272.
29. Yang, L.X.; Zhou, W.M. Study on precision of PS/ABS composite powders formed parts based on selective laser sintering. *China Plast. Ind.* **2017**, *45*, 35–38.
30. Bo, F.X.; He, B.; Zong, X.M. Experimental study on selective laser sintering of coated sand. *Laser Optoelectron. Prog.* **2017**, *54*, 247–253.
31. Brion, D.A.J.; Pattinson, S.W. Quantitative and real-time control of 3D printing material flow through deep learning. *Adv. Intell. Syst.* **2022**, *4*, 2200153. [CrossRef]

applied sciences

MDPI

Article

Module Partition of Mechatronic Products Based on Core Part Hierarchical Clustering and Non-Core Part Association Analysis

Shuai Wang [1], Yi-Fei Song [1], Guang-Yu Zou [2] and Jia-Xiang Man [3,*]

[1] School of Environmental Engineering, Xuzhou University of Technology, Xuzhou 221018, China; ws626@xzit.edu.cn (S.W.); syf_xzit@163.com (Y.-F.S.)
[2] School of Mechatronics Engineering, China University of Mining and Technology, Xuzhou 221116, China; tb20050029b4@cumt.edu.cn
[3] School of Mechatronics Engineering, Xuzhou University of Technology, Xuzhou 221018, China
* Correspondence: manjx@xzit.cn; Tel.: +86-18761434991

Abstract: Production using modular architecture can not only shorten the product development cycle and improve the efficiency of product development, but also facilitate the upgrading of a product's main functions and the recycling of materials. However, mechatronic products are plagued by various problems, such as greater difficulty in development and longer product development cycles due to their large numbers of parts with intricate internal relationships. However, the existing modular design method still faces problems when dealing with the modular design of mechatronic products. The structure of mechanical and electrical products is very complex, which is not conducive to the establishment of a model, and complex structural models lead to low efficiency and poor accuracy of module identification. Therefore, we propose an integrated module division method for mechatronic products based on core part hierarchical clustering and non-core part association analysis. Firstly, the core part screening method is used to simplify the structural model of mechatronic products and reduce the difficulty of modeling. Then, based on the core parts, the corresponding product design structural matrix (DSM) model is established. Secondly, the hierarchical clustering algorithm is used to obtain the module division scheme of different levels of mechatronic products, and the optimal modular scheme is obtained through an evaluation of modularity and a rationality analysis of module structure. Finally, based on the analysis of the association strength between the non-core parts and the existing modules, the non-core parts are classified into the corresponding product modules, and the final modularization scheme is obtained. A case study demonstrates the feasibility of the proposed method through the modular design of an electric bicycle.

Keywords: module partition; mechatronic products; hierarchical clustering; association analysis

1. Introduction

There are many types of mechatronic products used in daily life and industrial production which occupy a very important position in the national economy. With the increasingly fierce international competition and the gradual awakening of environmental protection awareness, the design of mechatronic products should not only meet the dynamic market demand, but also strive for better environmentally friendly attributes [1–3]. Current research suggests that the modular design concept can effectively balance the relationship between design elements, such as product development, market demand, product structure, and environmental protection [4,5]. The advantages of modular architecture include

the following aspects: (1) product development costs can be effectively controlled; (2) the product development cycle can be shortened; and (3) the negative impacts of the product life cycle on the environment can be reduced. The primary feature of modular design is to distribute the functions of products into different modules to the greatest extent possible, while minimizing the association between modules and maximizing the association within modules [6]. Many products are currently developed using modular structures, such as smart phones, household appliances, new energy vehicles, industrial steam turbines [7], and wind turbines [8].

Product structure modeling and module identification are the basis of product modular design [9]. Common mechatronic products usually comprise mechanical systems, electrical systems, and hydraulic systems, with each system containing a certain number of parts. However, due to the large number of mechanical product parts and the intricate relationships between them, there are some important problems that must be urgently solved to modularize the mechatronic products. Firstly, how can we reduce the difficulty of modeling mechanical product structures while assuring the accuracy of structural models? Secondly, when the mechanical product structure is complex, how can we ensure the rationality of the product module division scheme? Motivated by these questions, we propose an integrated modular design methodology for mechatronic products based on core part hierarchical clustering and non-core part association analysis. In this paper, we describe how the difficulty of mechanical product modeling can be reduced by selecting the core parts of the product. At the same time, the credibility of the established product model is ensured. Then, the remaining product parts are divided into product modules using the correlation analysis method to guarantee the rationality of the product modularization scheme.

2. Related Work

2.1. Product Structure Modeling

Product structure modeling involves quantitatively analyzing the relationship between the parts after splitting the product and displaying them in an intuitive and concise way. Researchers have proposed various methods and tools to achieve this, such as the design structure matrix (DSM) [10], node-link diagrams [11], Petri nets [12], and other analytical tools [13].

For more than a decade, the DSM has been a widely used modeling framework across many areas of research because of its conciseness in representation and ease of analysis [14]. In Browning's review, the DSM can be divided into four types according to different application objects: the component-based DSM, people-based DSM, activity-based DSM, and parameter-based DSM [15,16]. The use of the DSM to model products or system architectures can be traced back to the works of Simon and Alexander in the early 1960s, but it was not until 1981 that Steward formally proposed the concept of the DSM [17]. The product DSM is challenging to build because of the large amount of relevant professional knowledge required and the varied interpretations of a product's structural relationships.

Product structure modeling usually needs to consider two factors. One is the decomposition granularity of the product, which determines the size of the product structure model. The second is the quantitative evaluation of the correlations among parts, which determines the complexity of the product structure model. For small products with a simple structure, the smallest particle size is often used to decompose them to the part level. For example, Han et al. completely decomposed a gear oil pump to the part level when establishing a DSM model of a gear oil pump; then, the community identification algorithm was used to realize the module division of the gear oil pump [18]. Li et al. realized the extraction of 3D model information through the secondary development of 3D modeling software, and established the DSM model based on the extracted product information [19].

For most mechatronic products, the complex product structure results in a huge number of parts. If product structure modeling is decomposed to the part level, the size of the established model will be huge, which is not conducive to the subsequent product module identification. Because of this, the relevant research literature began to focus on reducing the difficulty of modeling and improving the efficiency of modeling. Ma et al. proposed a method for selecting product parts according to the market value and established a structural model of a coffee machine using this method [20]. Wang et al. proposed a method of module division of complex products based on core parts, but the screening method of core parts in this paper is too cumbersome [6]. Li et al. proposed a pre-modular approach to the structural modeling of mechatronic products. This method divides a product into several small modules in advance, based on an analysis of the physical structure of the product; a structural model of the product is then built on the basis of these small modules [21]. Zhang et al. solved the problem of excessive parts by improving the granularity of product decomposition; they established a DSM model of a wind turbine using this method [8]. However, there is no specific principle that can be used to control the decomposition size of products in the literature, only relying on the design experience of engineers, so this method cannot be widely promoted. Some researchers proposed a method of building a product structure model based on product 3D model information extraction [22]. Such methods have great advantages in modeling efficiency and model data objectivity, but the limited product information that can be extracted cannot meet the modeling requirements of diversity.

Based on the above analysis, we can divide the mainstream methods used for the structural modeling of mechatronic products into two types. The first type is the pre-modular method; that is, an expert divides a product's parts into several small modules based on their own knowledge and experience. Then, a structural model of electromechanical products is established on the basis of these small modules. The advantage of this method is that it is easy to implement, but the disadvantage is that the modeling results are easily affected by the subjective opinions of experts. The second type is the automatic modeling method. Generally, the basic information needed for product modeling is obtained by programming means, and then the product DSM model is built based on this information. The advantage of this method is its efficiency, but the disadvantage is that it relies heavily on the normativity of the product model, and the product information that can be extracted is limited.

Therefore, we selected the approach using the screening of core parts to establish a mechanical and electrical product structure model. Considering the ease of application of the core part screening method, we propose a method for determining core parts based on an assessment of the functional and structural importance of the parts. For the remaining non-core parts, correlation analysis is used to classify them into the existing product modules.

2.2. Module Partition Method

Module partition involves the identification of the module structure in a product by technical means, based on the product structure model. In the past two decades, module partitions based on function analysis [23], physical architecture [24], manufacturing [25], disassembling [26], and other aspects of the product life cycle have been proposed. Based on the various principles of the module partition algorithm, it can be divided into heuristic algorithm-based methods and cluster algorithm-based methods.

Modular methods driven by the heuristic algorithm usually establish the corresponding objective function according to different modular requirements and then use the appropriate heuristic algorithm to obtain the optimal modular solution. Guo and Sosale took

the number of interfaces between product modules as the objective function of module division and then used the simulated annealing algorithm to obtain the optimal product modularization scheme [27]. To reduce the environmental impact of product recycling, Ji et al. established a comprehensive model based on a material reuse and technology system and used a bi-level optimization model based on a constraint genetic algorithm for the solution [28]. Zheng et al. solved a multi-objective optimization model based on minimizing maintenance costs and maximizing product modularity by using an improved-strength Pareto evolutionary algorithm 2 to obtain the optimal Pareto set of product modularity schemes [29]. Wang et al. established a comprehensive DSM structure model and used an improved genetic algorithm to obtain the optimal scheme for the modular design of a bicycle [22]. The modular method based on the clustering algorithm is to use quantitative tools to express the association between parts, and then realize the module division of products according to the intimate relationship between parts. Beek et al. established a DSM model of the shifting system of an equation car based on the function–behavior–state model, and then used the k-means clustering algorithm to obtain its modular scheme [9]. Yang et al. established a DSM model of a gear reducer based on function- and sustainability-related factors, and then determined the optimal module number and the final module division scheme using a genetic algorithm and a kernel-based fuzzy C-means algorithm [30]. On the basis of establishing the function–behavior-structure mapping model of 3D printer, Li et al. used the K-means clustering algorithm to realize the module division of the printer [31]. AlGeddawy and ElMaraghy obtained a clustering tree that reflected the internal structure of a product by using the hierarchical clustering algorithm; they then selected the optimal separation granularity to obtain the final modular results [32]. Li et al. used a modularization index based on product hierarchy clustering to obtain the optimal product modularization scheme and then demonstrated the feasibility of this method for a concrete-spraying machine product [33]. Li and Wei proposed an improved elbow evaluation method to analyze the results of product hierarchical clustering and applied it to the modular design of jaw crusher products [34].

Researchers have also put forward some other algorithms to realize module division. For example, the community identification algorithm in complex networks was introduced by Li and Zhang into the modular design of a product [8,35]. Zhang et al. proposed a multi-granularity module partition approach for complex mechanical products based on a complex network, and proved the effectiveness and superiority of the method through the module division of elevator products [36]. Liu et al. proposed a module partition method for complex products based on stable overlapping community detection and overlapping component allocation, and used this method to obtain a modularization scheme for CNC grinding machines [37]. The community recognition algorithm can effectively identify the community structure in complex networks, but the solution efficiency is generally low when dealing with products with complex structure and cannot show the hierarchy of the internal structure of products. In addition, Hossain et al. built a two-layer collaborative optimization model of modular product design and supply chain architecture and designed a nested bi-level particle swarm optimization algorithm (NBL-PSO) to solve the problem [38]. Also for the problem of module partition, Tian et al. took the maximization of green attributes and recycling attributes as the optimization goal and introduced an innovative social engineering optimizer (SEO) to solve the problem [39].

From the above analysis, it can be seen that the heuristic algorithm and the clustering algorithm remain the mainstream methods used in the modular design of products. The heuristic algorithm has problems such as a long solution time and unstable solution results when dealing with mechatronic products [6,8]. Therefore, considering the relatively complex structure of mechanical and electrical products, we used the hierarchical clustering

algorithm to divide the modules and combined this with modular evaluation to obtain the optimal modular scheme.

3. Research Gaps and Challenges

Based on a review of the literature regarding the modular design of mechatronic products in recent years, we identified several knowledge gaps and challenges:

- The first is the structural modeling of mechatronic products. How can the complexity of the structural modeling of mechatronic products be reduced and the accuracy and rationality of the established model be ensured?
- The second is the acquisition of an optimal modularization scheme for mechatronic products. After obtaining the basic modular framework of mechatronic products based on the core parts, how do we deal with the module division of the remaining non-core parts?

4. Method

The research framework we used was divided into four steps, as shown in Figure 1. These were (1) mechatronic product core part screening based on expert scoring; (2) mechatronic product structure modeling with core parts; (3) mechatronic product module division by hierarchical clustering; and (4) non-core part module division based on association analysis.

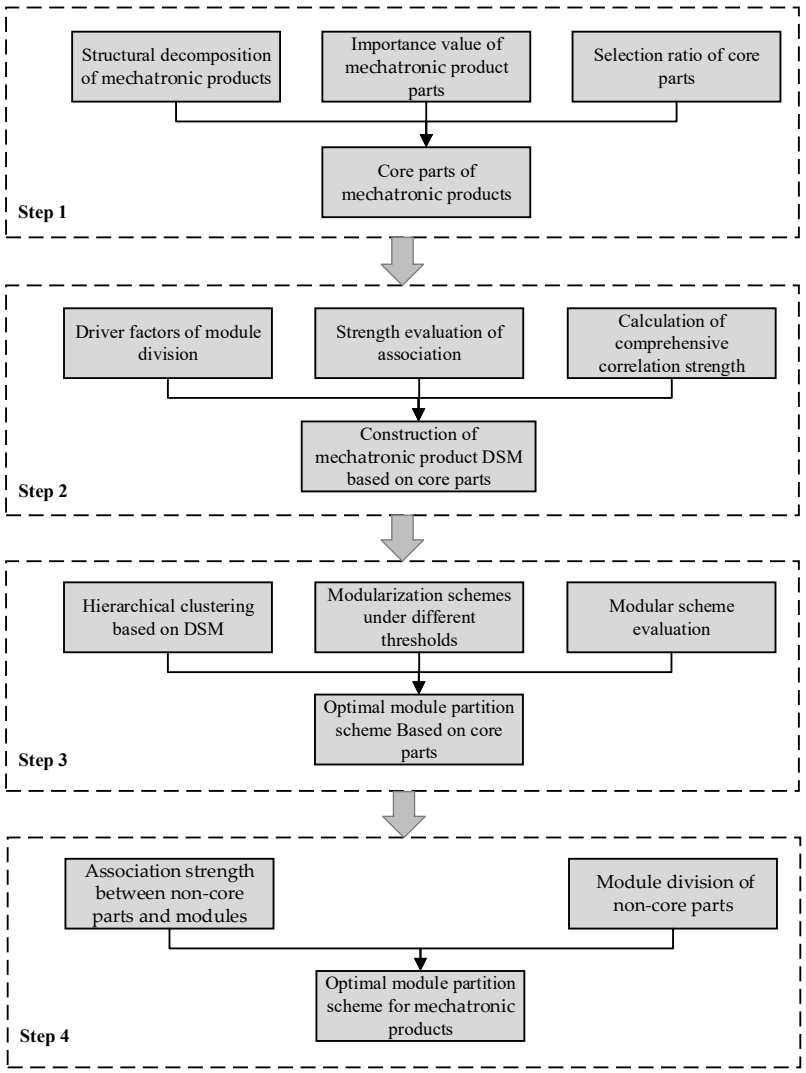


Figure 1. The research framework used in this study.

Step 1: Mechatronic product core part screening based on expert scoring. A number of product design engineers were selected to score the functional importance and structural importance of product parts. The product parts were sorted according to the weighted average of their importance, and the core parts of the product were selected by choosing a reasonable threshold.

Step 2: Establishment of a DSM model based on core parts. Based on the existing literature, the rules of different correlation strength values between core parts were determined. We then established a DSM model for mechatronic products.

Step 3: Mechatronic product module division by hierarchical clustering. We used the hierarchical clustering method because it can reveal the modular structure of products at different levels without specifying the number of modules.

Step 4: Non-core part module division based on association analysis. According to the strength of the relationship between the non-core parts and the existing modules of the product, the non-core parts were divided into the existing modules of the mechatronic product.

4.1. Mechatronic Product Core Parts Screening Based on Expert Scoring

A structural model of mechatronic products is difficult to establish, mainly due to the large number of product parts and the intricate interactions between them. To reduce the difficulty of the structural modeling of mechatronic products, some researchers have adopted the method of using pre-modular product parts. This requires product designers to have a high level of professional knowledge and inevitably involves some subjectivity, which may influence the results of complex product modularization. Therefore, we propose a means of screening the core parts of mechatronic products to reduce the complexity of product modeling to a reasonable level, while ensuring that the product model established can reflect the structural characteristics of the mechatronic product.

To ensure the rationality and credibility of the results of core parts screening, a given number of product design engineers were selected to score the functional and structural importance of product parts. The scoring criteria are shown in Table 1. After being scored by the engineers, the importance ranking of complex product parts was obtained by calculating the weighted average. The comprehensive importance of part i can be calculated as follows:

$$V_i = \frac{1}{k}\left(\sum_{j=1}^{k} w_1 F_{ij} + \sum_{j=1}^{k} w_2 S_{ij}\right) \quad (1)$$

where V_i is the comprehensive importance value of product part I, k is the number of engineers participating in the scoring, w_1 and w_2 are function importance weight and structure importance weight, respectively, and F_{ij} and S_{ij} represent the functional importance and structural importance points assigned by engineer j to part I, respectively. After obtaining the ranking of product parts, we filtered the core parts by determining an appropriate threshold. For example, when the threshold value was 60, the comprehensive importance of the part was greater than 60, so it was considered to be a core part.

Table 1. Scoring criteria for the importance of different parts.

Score	Functional	Structural
81–100	Very important	Very complex
61–80	Important	Complex
41–60	General	General
21–40	Unimportant	Simple
1–20	Very unimportant	Very simple

4.2. DSM Model of Mechatronic Products Based on Core Parts

This section covers two steps. The first is to evaluate the correlation strength value between the core parts according to the evaluation criteria. The other is to obtain the comprehensive correlation strength value between the core parts through weighted summation and build the DSM model of the product.

4.2.1. Correlation Between the Core Parts

The calculation of the correlation between parts is the basis of the establishment of a complex product structure model. In the process of analyzing the correlation between parts, the multi-source correlation information (structure, function, flow, etc.) is taken into account to assess the strength of correlation between the parts. The multi-source correlation information is also considered by many researchers to be the module partition driver factors. Different module partition driver factors will generate different module structures. In the modular design process of mechatronic products, it is not only the influence of structure, function, and flows that should be considered, but also the influence of life span and recovery methods. Therefore, we present the key driver factors of module partition in this paper, which include structural correlation, functional correlation, flow correlation, life span correlation, and recovery method correlation.

- Structural correlation: The connection form and fitting type are the main factors determining the strength of the structural correlation between parts. The acquisition of these data relies mainly on engineers' design knowledge and relevant engineering information.
- Functional correlation: The functions to be realized by mechatronic products are generally complex, so the total functions are often decomposed into simple subfunctions in the design process. It is beneficial for product design and subsequent function upgrading to categorize parts that have the same function into one module.
- Flow correlation: Flow correlation is analyzed by determining whether energy flow, material flow, or signal flow exists between parts. The flow correlation strength between assembled parts is also evaluated according to energy flow, material flow, and signal flow.
- Life span correlation: Mechatronic products usually comprise a large number of parts, while the service times and maintenance periods vary. Parts that have similar service times and maintenance periods can be grouped into the same module to reduce the use costs of a mechatronic product.
- Recovery method correlation: Increasingly serious environmental issues have led government agencies to focus more on the recycling of renewable resources. Mechatronic products have a large number of parts with high recycling values. However, the complexity of mechatronic product structures results in higher costs for recycling. Therefore, it is beneficial to reduce the recycling cost by grouping parts that can be recycled using the same method into the same module.

Based on this, the engineers can use these factors as an assessment index of correlations between product parts, and then estimate the correlation strengths between parts based on the corresponding evaluation criteria. The correlation strengths of linguistic variables can be divided into four levels: strong, medium, weak, and none; the corresponding interval values are shown in Table 2.

Table 2. Standard interval values for different correlation strengths.

Strength of c_k	Interval Value	Description of Correlation Strength Between Parts m and n				
		Structure	**Functional**	**Flow**	**Life Span**	**Recovery**
strong	0.7~0.9	Impartible connection	Bear the same core function	Energy flow between them	Frequent repairs or maintenance	Reuse
medium	0.4~0.6	Connection difficult to disassemble	Bear the same general function	Signal flow between them	Regular repairs or maintenance	Recycle
weak	0.1~0.3	Connection easy to disassemble	Bear the same additional function	Material flow between them	Almost no repairs or maintenance	Discard
none	0	No physical connection	No functional relationship	No flow between them	Different repairs or maintenance cycle	Different recovery methods

4.2.2. Construction of Weighted DSM Model

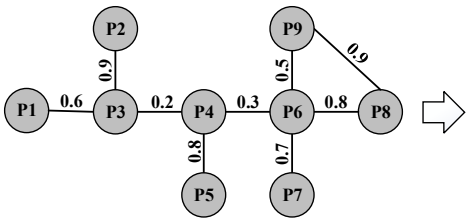
By analyzing the structural characteristics of mechatronic products, we can determine what kind of correlation exists between the core parts. The values of the corresponding correlation strengths are then obtained based on the standard values table. At this time, the weighted summation method is adopted to reflect the degree of influence of various modularization driving factors on the modularization plan, and the value of weight can be changed to meet variant modularization requirements. Among them, structure correlation, function correlation, flow correlation, life span correlation, and recovery method correlation are represented by c1, c2, c3, c4, and c5, and the corresponding weights are w1, w2, w3, w4, and w5, respectively. The comprehensive correlation strength between core parts m and n is obtained as follows.

By analyzing the structural characteristics of mechatronic products, the type of correlation that exists between the core parts can be determined. The values of the corresponding correlation strengths can then be obtained based on the standard values table. At this time, the weighted summation method is adopted to reflect the degree of influence of various modularization driving factors on the modularization plan, and the value of weighting can be changed to meet variant modularization requirements. Structure correlation, function correlation, flow correlation, life span correlation, and recovery method correlation are represented by c1, c2, c3, c4, and c5, respectively, and the corresponding weights are w1, w2, w3, w4, and w5, respectively. The comprehensive correlation strength between core parts m and n can be obtained from the following equation:

$$PC_{mn} = \sum_{l=1}^{5} w_l pc_l \tag{2}$$

where pc_l is the strength value of the l-th correlation between core parts, m and n, w_l represents the weight of the l-th correlation relationship, and PC_{mn} refers to the comprehensive correlation strength value between the core parts, m and n. The value of weight w_l is generally determined using a mathematical weighting method after obtaining the opinions of experts via a questionnaire survey.

Based on the evaluation value of the comprehensive correlation strength between the mechatronic product parts, a DSM model of the mechatronic product can be constructed. Figure 2a shows the structural model of a virtual product, and the corresponding DSM model is shown in Figure 2b. In Figure 2b, the names of product parts correspond to the first row and the first column of cells of the DSM, while the values in the other cells represent the comprehensive correlation strength of the corresponding two parts.



(a)

	P1	P2	P3	P4	P5	P6	P7	P8	P9
P1	0	0	0.6	0	0	0	0	0	0
P2	0	0	0.9	0	0	0	0	0	0
P3	0.6	0.9	0	0.2	0	0	0	0	0
P4	0	0	0.2	0	0.8	0.3	0	0	0
P5	0	0	0	0.8	0	0	0	0	0
P6	0	0	0	0.3	0	0	0.7	0.8	0.5
P7	0	0	0	0	0	0.7	0	0	0
P8	0	0	0	0	0	0.8	0	0	0.9
P9	0	0	0	0	0	0.5	0	0.9	0

(b)

Figure 2. (**a**) A product structure model; (**b**) a product DSM model.

4.3. Module Division of Mechatronic Products Based on Hierarchical Clustering

The clustering algorithm is simple, efficient, and has good data adaptability, so it is widely used in product modularization. There are many types of clustering algorithms, including X-means, fuzzy C-means, and hierarchical clustering [19,40]. Of these, the hierarchical clustering algorithm can effectively display a product's hierarchical structure under various module granularities; that is, it can provide modular solutions with different modular granularities. Therefore, in this study, we used hierarchical clustering to obtain the multi-level module structure of mechatronic products, and we used module degree evaluation to obtain the optimal modular scheme. The hierarchical clustering algorithm generally computes the tightness of the elements, merges nodes in turn according to the tightness value, and finally expresses the merging process in the form of a cluster tree. To reduce unnecessary programming work and to achieve the above hierarchical clustering algorithm, we used the "pdist", "linkage", and "dendrogram" functions in MATLAB software (2016a) to implement the clustering operation of DSM elements.

We used the DSM model of the virtual product shown in Figure 2 as an example to further explain the specific process of hierarchical clustering, and the final clustering result is shown in Figure 3. The coordinate value of the abscissa of Figure 3 expresses the nine parts of the virtual product, and the value of the ordinate refers to the distance between clustering parts. Nodes (parts) to be clustered are merged successively along the positive direction of the Y-axis according to the distance between them, and they finally intersect at a point to form an inverted Y-shaped tree structure. The corresponding module division scheme can be obtained by horizontally cutting the clustering tree with different thresholds. When the threshold value is 2, as shown by the gray horizontal dotted line in the figure, the entire product model is divided into five modules; as the threshold increases to 2.5, 3, 3.5, and 4, the number of modules in the product decreases to 4, 3, 3, and 2, respectively.

To help design engineers choose the optimal product modularization scheme, researchers have put forward a variety of indices to evaluate the advantages and disadvantages of modularization schemes. Among the many modularity indices, the modularity index Q, proposed by Newman et al., has been widely used in the solution and evaluation of product module division schemes due to its excellent module recognition ability and balanced module evaluation effect [41]. The calculation formula is as follows:

$$Q = \sum_{i=1}^{k} \left(e_{ii} - a_i^2 \right). \tag{3}$$

where e_{ii} represents the fraction of edges with both end vertices in the same module i, and a_i represents the fraction of edges with at least one end vertex inside module i. k refers to the number of modules in the modular scheme. The modularity index Q is a real number of $[-1, 1]$; the larger the value, the more reasonable the module partition scheme, and vice versa. The calculation of the modularity index Q involves knowledge of complex network theory; however, given the limited space in this paper, we will not describe this in detail. For further details of the specific calculation method, please refer to the relevant literature [8,41]. As shown in Figure 3, when the threshold values are 3.5 and 4, the modularity index Q reaches the maximum value of 0.51, and the module division scheme obtained at this time is the optimal result.

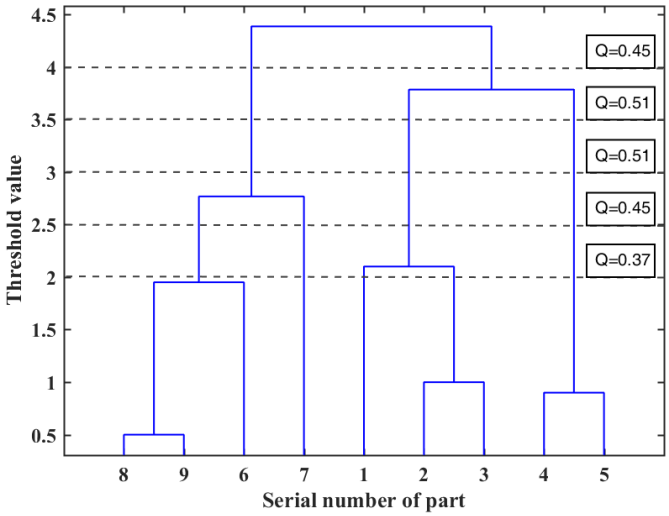


Figure 3. Hierarchical clustering of the DSM model.

4.4. Module Division of Non-Core Parts Based on Association Analysis

Based on the hierarchical clustering of core parts, a preliminary module division scheme of mechatronic products can be obtained. The following describes how to divide the remaining non-core parts into existing modules, to complete the final module division of mechatronic products. The module division process of non-core parts can be divided into two main parts: firstly, the correlation strength between non-core parts and modules is calculated; secondly, the module division of non-core parts is completed according to the strength of the relationship.

4.4.1. Analysis of the Association Strength Between Non-Core Parts and Modules

The correlation strength between the non-core parts and the existing modules of a product is the decisive factor in determining which module the non-core parts belong to. As modules generally contain multiple parts, the relationship between parts and modules is not as clear and easy to judge as the relationship between two parts. For example, suppose that part 1 and module A (including three parts, part 2, part 3, and part 4) have a structural association, and there are seven different situations in the structural association relationship between the two, as shown in Figure 4. This still does not take into account the changes in the strength of structural correlation.

Through the analysis of the diversity and complexity of the relationship between parts and modules, we used the method of "taking the maximum" to assess the correlation strength between them. According to the evaluation standard of the correlation strength between the parts as outlined in Section 4.2, the correlation strength between the non-core parts and each core part in the module can be obtained. Then, the strength of the correlation

between the non-core part and the module can be determined by taking the maximum. The formula is as follows:

$$MC_{xy} = \sum_{k=1}^{5} w_k \max\{mc_{kz,}\}, z = 1, 2, \cdots u. \quad (4)$$

where MC_{xy} is the comprehensive correlation strength between the non-core part x and the module y, w_k refers to the weight of the k-th correlation relationship, mc_{kz} is the strength of the k-th correlation strength between the non-core part x and the core part z (belonging to module y), and u represents the total number of parts in module y.

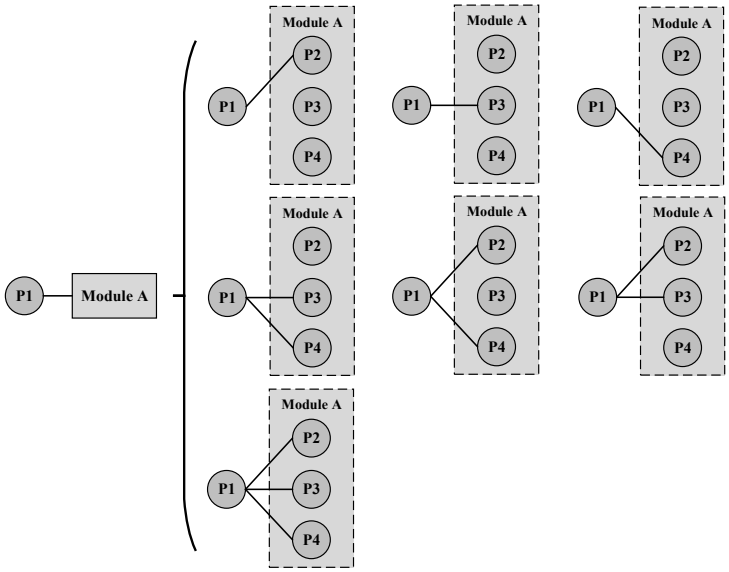


Figure 4. Complexity analysis of the association between non-parts and modules.

4.4.2. Module Division of Non-Core Parts

The strength of different types of correlation relationships between non-core parts and core parts of each module can be obtained by expert evaluation. The comprehensive correlation strength of non-core parts and modules is then calculated using Formula (3). The final comprehensive correlation matrix is shown in Figure 5. By comparing the strength of the relationship between the non-core parts and different modules of the product, the non-core parts can be divided into the corresponding modules with the strongest association relationship, ultimately completing the module division of the entire complex product.

$$MC = \begin{bmatrix} mc_{11} & mc_{12} & \cdots & mc_{1b} \\ mc_{21} & mc_{22} & \cdots & mc_{2b} \\ \vdots & \vdots & \ddots & \vdots \\ mc_{a1} & mc_{a2} & \cdots & mc_{ab} \end{bmatrix}$$

Figure 5. The correlation matrix between non-core parts and modules.

The quantity of data that must be processed during the process of module division of non-core parts is very large. Therefore, we compiled a module division program of non-core parts using MATLAB software (2016a) to reduce the workload of module division. Figure 6 shows a flow chart of the module division of non-core parts of mechatronic products.

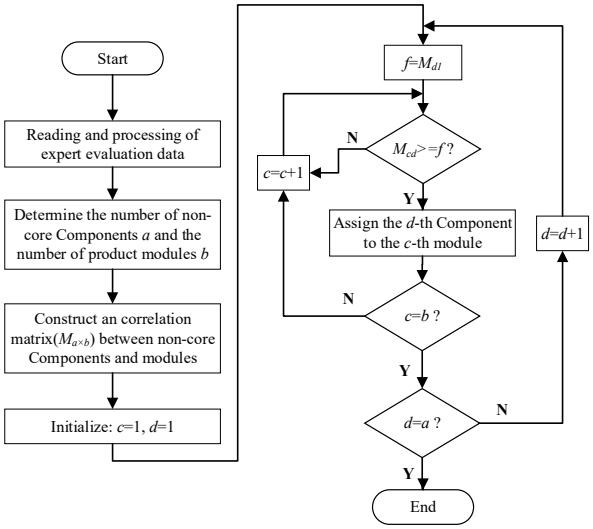


Figure 6. The module division process of non-core parts.

5. Case Study

We applied the proposed method to the modular design of a complex mechanical product family in a factory. The module partition of electric bicycle production was carried out to demonstrate the effectiveness of the approach proposed in this paper. The purpose of this was to improve customer satisfaction and product development efficiency while reducing enterprise production costs and the impact of product life cycles on the environment.

5.1. Core Parts Screening of an Electric Bicycle

An electric bicycle (shown in Figure 7) consists of approximately 100 main parts. Obviously, it is very difficult to establish the DSM structure model of the product based on these 100 main parts, so the size of the product DSM model can be reduced by screening the core parts. Five product engineers involved in the design of electric bicycles evaluated the functional importance and structural complexity of the parts according to the scoring standards shown in Table 1. Through the weighted sum of the scores given by the designers, the comprehensive importance values and ranking of electric bicycle parts were obtained, as shown in Table 3. When the threshold was 60, the top 28 parts were selected as the core parts of the electric bicycle.

Figure 7. Illustration of the electric bicycle structure.

Table 3. Importance value and ranking of electric bicycle parts.

Ranking	Part Name	Score	Ranking	Part Name	Score	Ranking	Part Name	Score
1	Hub motor	96.8	32	Trumpet	57.8	63	Backrest strut	33.5
2	Battery	92.9	33	Speed transmitter	56.4	64	Front brake reset spring	33.3
3	Controller	89.8	34	Speed control handle	55.9	65	Back brake reset spring	32.6
4	Headset	89.3	35	Back mudguard	55.8	66	Saddle	32.6
5	Front hub	88.5	36	Front mudguard	55.8	67	Charging mouth	32.4
6	Back axle	87.5	37	Battery holder	55.7	68	Back saddle	31.6
7	Back brake drum	87.4	38	Right crankset arm	55	69	Seatpost collar	30.9
8	Front brake drum	86.6	39	Lift crankset arm	55	70	Head lamp pedestal	27.1
9	Fork	84.9	40	Right crankset	54.9	71	Kickstand	26.7
10	Seat stay	83	41	Lift crankset	54.9	72	Back brake handle bracket	26.2
11	Chain stay	82.9	42	Control cable	54.9	73	Front brake handle bracket	26.2
12	Front axel	80.9	43	Lower leg	54.8	74	Front brake lever	25.3
13	Screem	78.2	44	Head lamp	54.4	75	Back brake lever	25.3
14	Cassette	77.5	45	Power cable	51.6	76	Backrest scale board	25.1
15	Backseat frame	76	46	Supporting leg	50.4	77	Screen cover	24.8
16	Back tire	74.8	47	Seatpost	50.2	78	Basket studdle	23.7
17	Front tire	74.6	48	Lift pedal	50	79	Lift damping spring	22.8
18	Middle axle	73.3	49	Right pedal	49.8	80	Right damping spring	22.8
19	Chain	72.7	50	Seat tube	47.3	81	Lampshade	22.1
20	Back tube	72.5	51	Backrest	45.6	82	Right guard plate	21.6
21	Chain ring	68.8	52	Battery holder cover	44.8	83	Lift guard plate	21.6
22	Back inner tube	66.1	53	Crown	44.3	84	Back mudguard studdle	20.3
23	Front inner tube	66.1	54	Starting switch	43.6	85	Front mudguard studdle	20.3
24	Middle tube	65.1	55	Bike lock	43.4	86	Grip	19.9
25	Upright rod	65	56	Upright rod collar	41.9	87	License plate buckle	17.4
26	Bike basket	62.1	57	Battery holder's lock	38.8	88	Headset cap	16
27	Head tube	61.1	58	Front brake cable	38.8	89	Reflector	13.6
28	Basket cover	60.4	59	Back brake cable	38.8	90	Support leg return spring	11.8
29	Rail	59.6	60	Back brake handle	35.2	91	Basket buckle	11.7
30	Upper fork	59.3	61	Front brake handle	35.2	92	Back mudguard buckle	7.5
31	Handlebar	59	62	Trumpet button	34.7	93	Front mudguard buckle	7.5

5.2. DSM Model of Electric Bicycles Based on Core Parts

After the selection threshold of the core parts was set to 60, 28 core components of electric bicycles were screened out, thus determining the size of the DSM specifications. Then, the values of the different correlation strengths between the core parts were evaluated based on the evaluation criteria shown in Table 2. Finally, based on the weights of five correlations, where structure correlation, function correlation, flow correlation, life span correlation, and recovery method correlation are 0.25, 0.3, 0.25, 0.1, and 0.1, the comprehensive correlation strength value was calculated, and the DSM model of an electric bicycle was obtained, as shown in Figure 8. In addition, the weights were obtained by a mathematical weighting calculation on the basis of the questionnaire survey conducted among five senior engineers in this professional field. The weight value is not fixed, and enterprises can adjust it according to the characteristics and design needs of their products. For example, if an enterprise wanted to reduce the environmental impact of product obsolescence, the weight of the recovery attribute could be increased.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Hub motor	1	0																											
Battery	2	0.69	0																										
Controller	3	0.32	0.74	0																									
Headset	4	0.02	0	0	0																								
Front hub	5	0.02	0	0	0.07	0																							
Back axle	6	0.32	0	0	0.07	0.07	0																						
Back brake drum	7	0.12	0.05	0.05	0.05	0.05	0.53	0																					
Front brake drum	8	0	0.05	0.05	0.05	0.33	0.05	0.1	0																				
Fork	9	0.02	0	0	0.46	0.22	0.07	0.05	0.18	0																			
Seat stay	10	0.02	0	0	0.07	0.07	0.2	0.05	0.05	0.07	0																		
Chain stay	11	0.02	0	0	0.07	0.07	0.22	0.18	0.05	0.07	0.16	0																	
Front axel	12	0.02	0	0	0.07	0.52	0.07	0.05	0.53	0.2	0.07	0.07	0																
Screem	13	0.23	0.65	0.55	0	0	0	0.05	0.05	0	0	0	0	0															
Cassette	14	0	0	0	0.05	0.05	0.33	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0														
Backseat frame	15	0.02	0	0.23	0.07	0.07	0.07	0.05	0.05	0.07	0.55	0.07	0.07	0	0.05	0													
Back tyre	16	0.53	0	0	0.07	0.07	0.25	0.23	0.05	0.07	0.07	0.07	0.07	0	0.05	0.07	0												
Front tyre	17	0.02	0	0	0.07	0.55	0.07	0.05	0.23	0.07	0.07	0.07	0.25	0	0.05	0.07	0.07	0											
Middle axle	18	0.02	0	0	0.07	0.07	0.07	0.05	0.05	0.07	0.2	0.07	0.07	0	0.56	0.07	0.07	0.07	0										
Chain	19	0	0	0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0	0.55	0.05	0.05	0.05	0.23	0									
Back tube	20	0.02	0	0	0.07	0.07	0.07	0.05	0.05	0.07	0.07	0.67	0.05	0	0.05	0.07	0.07	0.07	0.25	0.05	0								
Chain ring	21	0	0	0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0	0.3	0.05	0.05	0.05	0.38	0.55	0.05	0							
Back inner tube	22	0.28	0.05	0.05	0.05	0.05	0.2	0.28	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.23	0.05	0.05	0.05	0.05	0.05	0						
Front inner tube	23	0	0.05	0.05	0.05	0.33	0.05	0.1	0.28	0.05	0.05	0.05	0.23	0.05	0.05	0.05	0.05	0.23	0.05	0.05	0.05	0.05	0.1	0					
Middle tube	24	0.02	0.08	0	0.07	0.07	0.07	0.05	0.05	0.07	0.07	0.34	0.07	0	0.05	0.07	0.07	0.07	0.07	0.05	0.54	0.05	0.05	0.05	0				
Upright rod	25	0.02	0	0	0.64	0.07	0.07	0.05	0.05	0.64	0.07	0.07	0.07	0	0.05	0.07	0.07	0.07	0.07	0.05	0.07	0.05	0.05	0.05	0.07	0			
Bike basket	26	0.02	0	0	0.02	0.02	0.02	0	0	0.02	0.11	0.02	0.02	0.02	0	0.11	0.02	0.02	0.02	0	0.02	0	0	0	0.02	0.15	0		
Head tube	27	0.02	0	0	0.22	0.07	0.07	0.05	0.05	0.25	0.07	0.34	0.07	0	0.05	0.07	0.07	0.07	0.07	0.05	0.34	0.05	0.05	0.05	0.54	0.25	0.02	0	
Basket cover	28	0.02	0	0	0.02	0.02	0.02	0	0	0.02	0.11	0.02	0.02	0.02	0	0.11	0.02	0.02	0.02	0	0.02	0	0	0	0.02	0.02	0.49	0.02	0

Figure 8. DSM model of an electric bicycle based on core parts.

5.3. Hierarchical Clustering of Core Parts of an Electric Bicycle Based on a DSM Model

The tree diagram in Figure 9 shows the clustering results of the core components of the electric bicycle, where the horizontal coordinate is the serial number of the part and the vertical coordinate is the threshold value. To obtain a more useful module division scheme, the threshold values are 2.5, 3, 3.5, 4, and 4.5, and the corresponding modularity index Q value and module division scheme are shown in Figure 9 and Table 4, respectively.

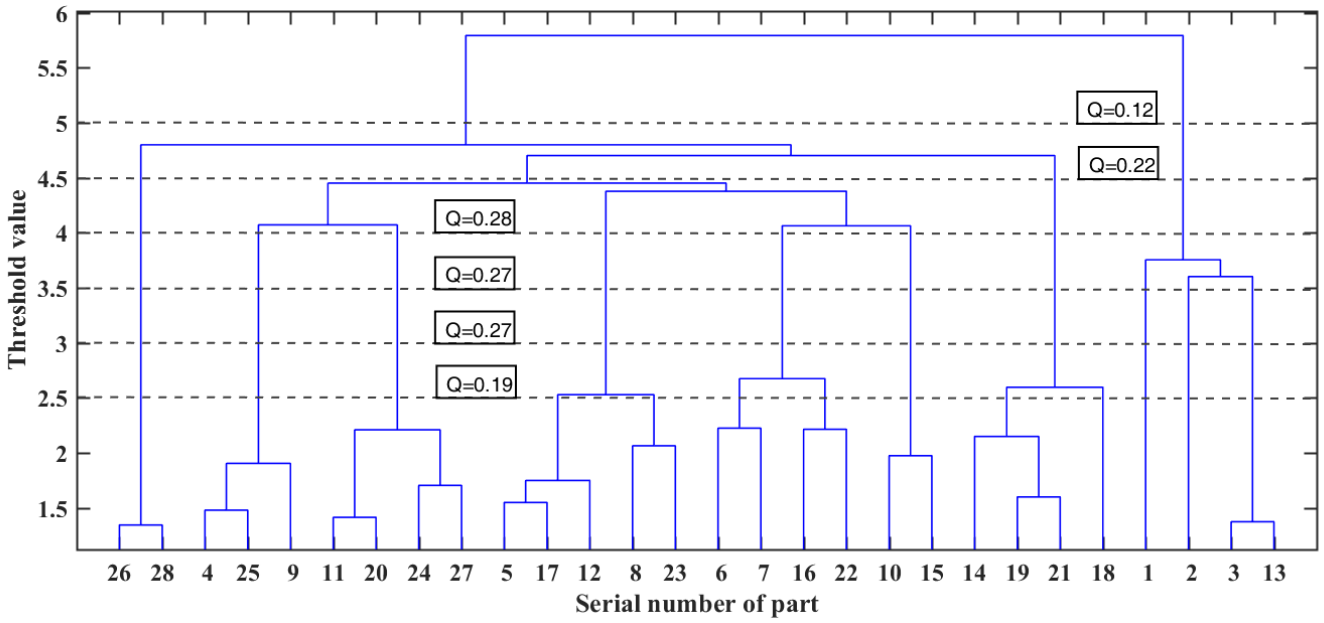


Figure 9. Hierarchical clustering of core parts of an electric bicycle.

Table 4. Modularity results under different thresholds.

Threshold	Number of Modules	Modularity Results
2.5	13	{1} {2} {3 13} {4 9 25} {5 12 17} {8 23} {6 7 16 22} {10 15} {11 24} {14 21} {18 19} {20 27} {26 28}
3 or 3.5	10	{1} {2} {3 13} {4 9 25} {5 8 12 17 23} {6 7 16 22} {10 15} {11 20 24 27} {14 18 19 21} {26 28}
4	7	{12 3 13} {4 9 25} {5 8 12 17 23} {6 7 10 15 16 22} {11 20 24 27} {14 18 19 21} {26 28}
4.5	4	{12 3 13} {26 28} {14 18 19 21} {4 5 6 7 8 9 10 11 12 15 16 17 20 22 23 24 25 27}
5	2	{1 2 3 13} {4 5 6 7 8 9 10 11 12 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28}

When the module partition scheme is evaluated according to the modularity index Q, the maximum value corresponding to the threshold value is 0.28 when the threshold value is 4. At this time, the electric bicycle is divided into seven modules, and the parts contained in each module are shown in Table 4. Only the core parts of the electric bicycle are considered in the hierarchical clustering, and the proportion of the core parts in the total number of parts is limited. Therefore, other factors should be referred to on the basis of the maximum modular index when selecting the optimal modular scheme. Among them, the number of modules and module structure, as two core contents of modular design, should be used as important references for selecting the optimal modular scheme. For example, when the total number of modules is small, the addition of non-core parts will cause the structure of some modules to be very large.

When only the modularity index Q is considered in the optimal modularity scheme, as shown in Figure 9, the maximum value of the modularity index Q is 0.28. At this time, the electric bicycle is divided into seven modules, with each module containing the parts information shown in Table 4. If the optimal number of modules of the product is calculated based on the total number of product parts in the relevant literature, the optimal number of modules of the electric bicycle should be 9 or 10. In Figure 9, when the threshold is 3 or 3.5, the corresponding number of modules is exactly 10. The corresponding modularity index Q here has a value of 0.27. Through the comparison of the two modular schemes, it can be seen that the modularity index Q of the two is just 0.01, and the difference between the two module division schemes is only in parts 1, 2, 3, and 13. After analyzing the results of module division, it seems unreasonable to divide parts 1 and 2 into two separate modules. However, it is necessary to consider that a large number of non-core parts will be added in the future. In addition, part 1 and part 2 are the hub motor and battery of the electric bicycle, respectively. Different models of electric bicycles also have different requirements for the power of the motor and the capacity of the battery. Therefore, dividing the motor and battery into separate modules is conducive to the development and upgrade of electric bicycles. Through the above analysis, the modularization scheme of 10 modules is regarded as the optimal solution for the modularization of electric bicycles. In the following work, non-core parts will be added to each module, and the final modular scheme of the electric bicycle will be obtained.

5.4. Module Division of Non-Core Parts of Electric Bicycles Based on Correlation Analysis

In light of the standards shown in Table 2, the strength of association between non-core parts and core parts in each module is evaluated in turn. Then, the association strength

between non-core parts and modules is calculated according to Formula (4). To ensure the consistency of driving factors in the process of module division, the weights used in Section 5.2 are still used for the five types of association relationships. The correlation matrix established between non-core parts and modules is shown in Figure 10.

$$MC_{65,10} = \begin{bmatrix} 0.02 & 0 & 0 & 0.07 & 0.07 & 0.22 & 0.25 & 0.28 & 0.07 & 0.02 \\ 0.02 & 0 & 0 & 0.66 & 0.22 & 0.07 & 0.07 & 0.32 & 0.07 & 0.02 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.02 & 0 & 0 & 0.145 & 0.16 & 0.07 & 0.07 & 0.07 & 0.07 & 0.02 \end{bmatrix}$$

Figure 10. Correlation matrix between non-core parts and modules.

Based on the correlation matrix between non-core parts and modules, the module division procedure of non-core parts is executed to gain the final product modularization scheme. The final product modularization scheme for an electric bicycle is shown in Table 5. An electric bicycle is divided into 10 modules, with the largest number of parts for the front wheel module and the rear wheel module. Each module is composed of 13 parts. The modules with the lowest number of parts are the motor module and the basket module, with each module containing four parts. The modular scheme is very balanced, both in terms of the number of module parts and module structure. In particular, the motor module and battery module initially contain only one part, and the structure of the modules tends to be reasonable even with the addition of non-core parts.

Table 5. The ultimate modular solution for electric bicycles.

No	Module	Module Parts
1	Motor	{1 33 34 45}
2	Battery	{2 37 52 57 67 82 83}
3	Control system	{3 13 32 42 44 54 62 70 77 81}
4	Steering system	{4 9 25 30 31 43 53 56 86 88}
5	Front wheel	{5 8 12 17 23 36 58 61 64 73 74 85 93}
6	Back wheel	{6 7 16 22 35 55 59 60 65 72 75 84 92}
7	Back seat	{10 15 51 63 6876 79 80 87 89}
8	Frame of bicycle	{11 20 24 27 29 46 47 50 66 69 71 90}
9	Pedal system	{14 18 19 21 38 39 40 41 48 49}
10	Bicycle basket	{26 28 78 91}

5.5. Compared with Community Detection Algorithm

The community identification algorithm of complex weighted networks is often used to divide complex product modules, and researchers who adopt this method use the pre-modular method to reduce the difficulty of complex product modeling. The complex weighted network model of an electric bicycle is shown in Figure 11, and the modularization scheme is obtained through the modified GN algorithm [42]. To facilitate this comparative study, the modular scheme with ten modules was selected as the optimal one, as shown in Table 6.

Among the ten modules, the structure of the steering system module, front wheel module, pedal system module, back wheel module, back seat module, and bicycle basket module is the same in the two modular schemes. The differences between the two modular schemes are shown in Figure 12, and the rationality of the modules is analyzed as follows:

1. The community recognition algorithm divides the support leg module from the frame of the bicycle module, and the support leg module only contains two parts and is closely connected with the frame of the bicycle module. Therefore, it is unreasonable for the support leg to be a module alone.

2. The community recognition algorithm reconstructs the motor module, battery module, and control system into two modules: the display system and the control and power system. However, with electric bicycle manufacturers, the motor and battery are usually purchased parts and have different types of customer requirements. It is conducive to product configuration and new product development to use them as separate modules.
3. In the pre-modularization stage, the community identification algorithm categorizes the left guard plate and the right guard plate into the frame module, while the non-core parts association analysis method in this paper categorizes the left guard plate and the right guard plate into the battery module. The guard plate is close to the frame module in structure and has a strong functional correlation with the battery module, but it is a plastic product and is not compatible with the frame module in terms of material.

Table 6. Modular results of an electric bicycle based on modified GN algorithm.

No	Module	Pro-Module	Module Parts
1	Steering system	Handlebar Steering	{25 31 56 86} {4 9 30 43 53 88}
2	Display system	Trumpet Head lamp Screen	{32 62} {44 70 81} {13 77}
3	Front wheel	Front wheel Front brake	{5 12 17 23 36 85 93} {8 58 61 64 73 74}
4	Frame of bicycle	Seat Frame of bicycle	{29 47 66 69} {11 20 24 27 50 71 82 83}
5	Pedal system	Pedal system	{14 18 19 21 38 39 40 41 48 49}
6	Control and power system	Motor Battery Controller	{1 33 34 45} {2 37 52 57 67} {3 42 54}
7	Back wheel	Back brake Lock Back wheel	{7 59 60 65 72 75} {55} {6 16 22 35 84 92}
8	Back seat	Back seat	{10 15 51 63 68 76 79 80 87 89}
9	Support leg	Support leg	{46 90}
10	Bicycle basket	Bicycle basket	{26 28 78 91}

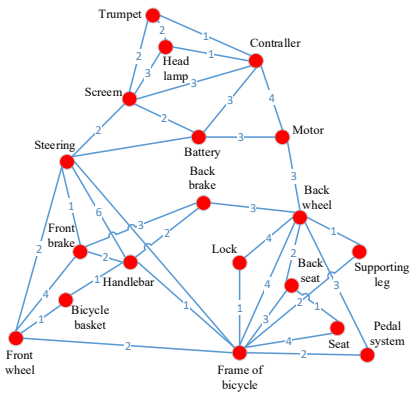


Figure 11. Complex weighted network model of an electric bicycle.

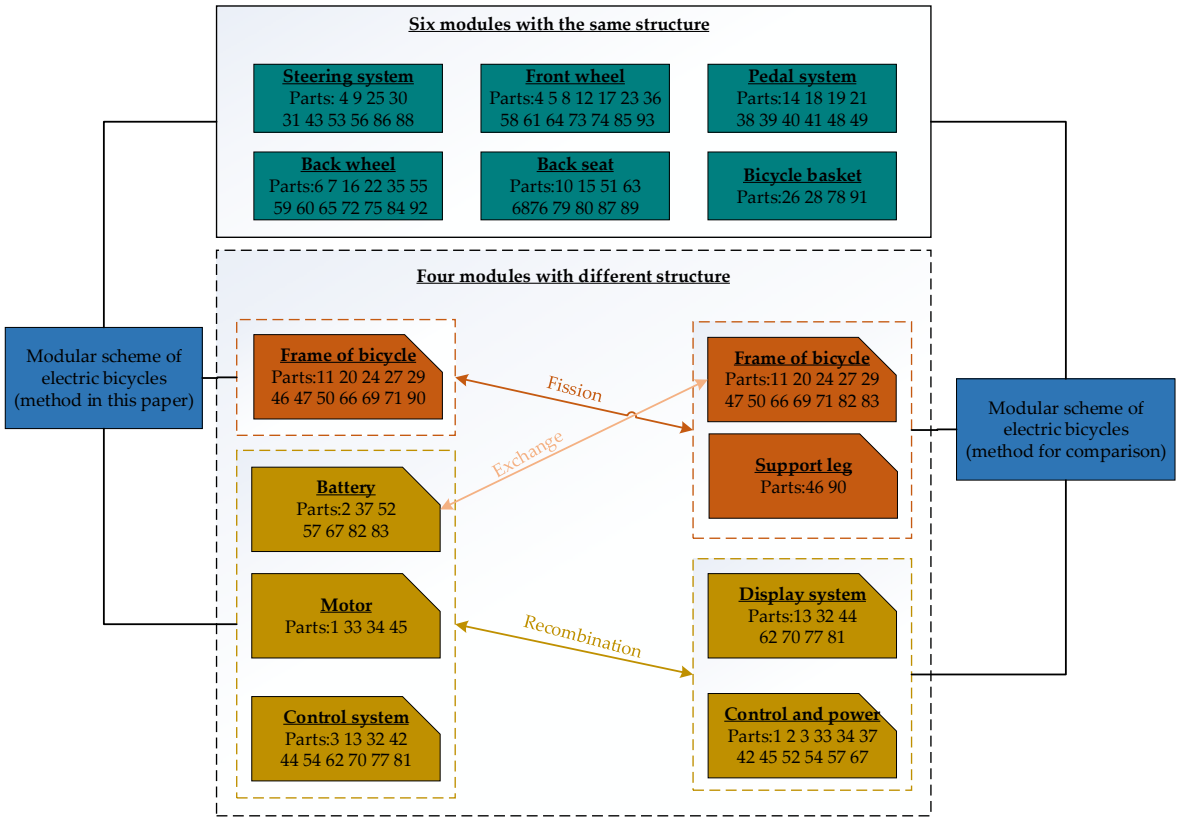


Figure 12. The differences between the two modular schemes.

In general, the pre-modularization part of the community identification algorithm relies heavily on the subjective experience of engineers, resulting in the module division of some parts being unreasonable. In addition, the strength of the correlation between modules after pre-modularization is difficult to accurately assess. As for the algorithm itself, the efficiency of the hierarchical clustering algorithm is slightly better than that of the community recognition algorithm, especially when dealing with complex product models.

5.6. Sensitivity Analysis of Screening Threshold for Core Parts

The selection ratio of core parts is related to the difficulty of modeling and the rationality of the model. An appropriate screening ratio can effectively reduce the product modeling workload and improve the solution efficiency of modular schemes. Therefore, it is necessary to conduct sensitivity analysis on the selection ratio of core parts.

After we reduce the number of core parts from 28 to 26, the hierarchical clustering results of the electric bicycle are shown in Figure 13. When Q values are 2.7 and 2.8, we obtain two relatively good modularization schemes for the electric bicycle. The modular scheme with a Q value of 0.7 is consistent with the scheme structure adopted in this paper.

We continue to reduce the number of core parts from 26 to 24, and the hierarchical clustering results of electric bicycles are shown in Figure 14. When Q is 0.27 and 0.28, we obtain two better modularization schemes for e-bikes. However, it is a pity that these three schemes are different from the schemes adopted in this paper. When Q is 2.7, we obtain a result that is very close to the scheme adopted in this paper. The structure of nine of the modules is exactly the same, only one bicycle basket module is missing. The reason for this phenomenon is that in the process of reducing the number of core parts, the two parts ranked 28 and 26 in the bicycle basket module are successively excluded, which eventually leads to the disappearance of the bicycle basket module. Therefore, this also shows that

too small a core part screening ratio may lead to the disappearance of some unimportant product modules.

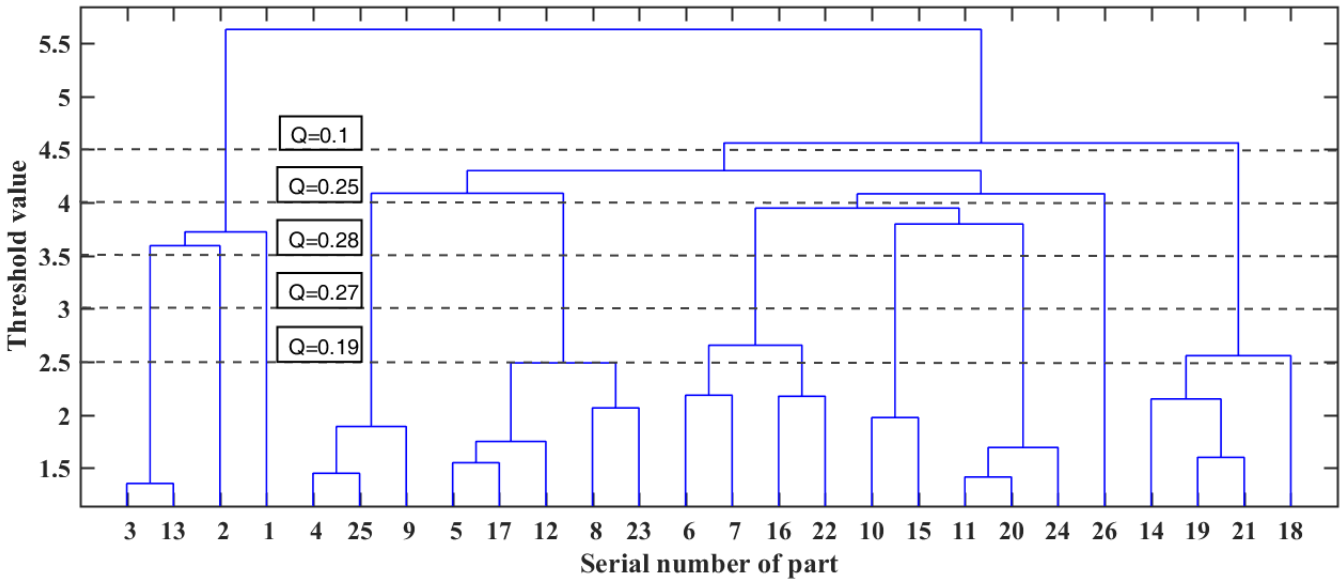


Figure 13. Hierarchical clustering of core parts of an electric bicycle (26 core parts).

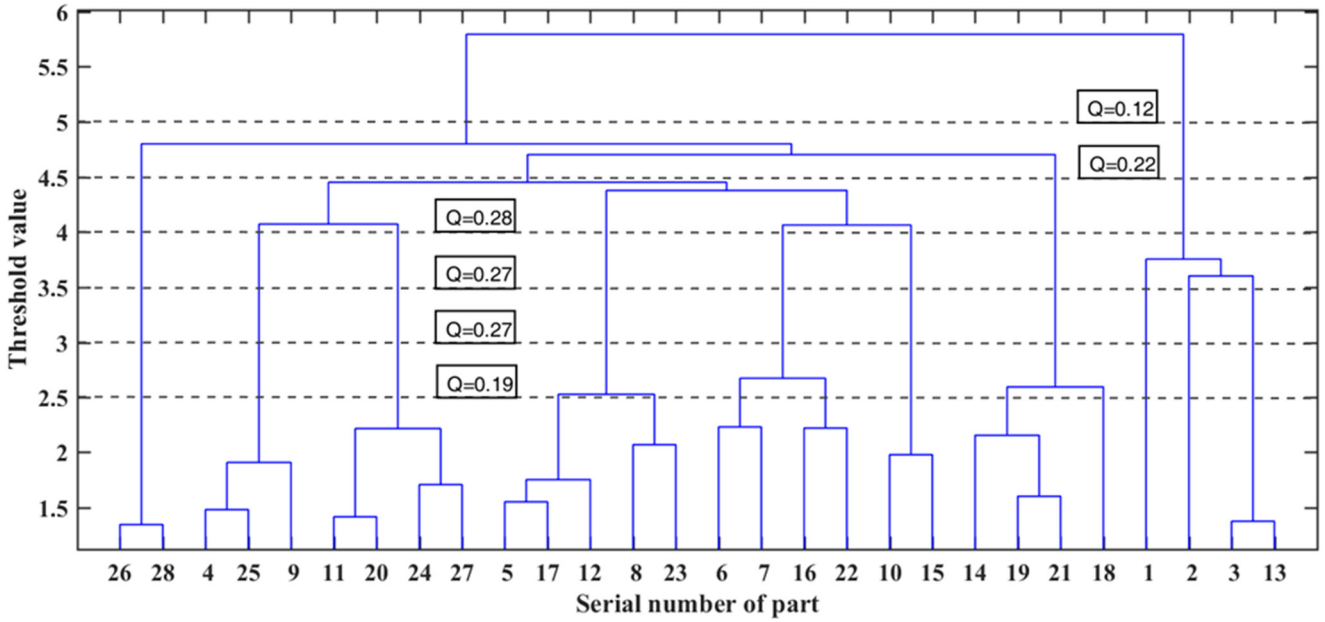


Figure 14. Hierarchical clustering of core parts of an electric bicycle (24 core parts).

6. Discussion

The difficulty of the module division of mechatronic products lies in the establishment of a product structure model, and the emphasis is on the rationality of the module division scheme. In this paper, we used the core parts selection method to reduce the difficulty of mechatronic product structure modeling, while hierarchical clustering and non-core part association analysis were used to ensure the rationality of module partition results.

Product pre-modularization is often used to reduce the difficulty of modeling mechatronic products; that is, the product components are divided into several sub-modules according to their function or structure. The advantage of this method is that it is relatively simple; however, it can be easily affected by the subjective opinions of product designers,

and pre-modularity lacks a clear operating process. In contrast, the selection process of core components is very objective, even if it is affected by some subjective factors of the design engineer, but it can be mitigated or eliminated by weighted methods. In addition, we also tried to use the part information extracted from the three-dimensional assembly model of the product to automatically screen the core components, so as to improve the efficiency of the existing method and further reduce the influence of objective factors, but the experimental results need to be improved.

7. Conclusions

In this paper, we proposed an integrated module partition method to deal with the modularization of mechatronic products. Based on the multi-source relationship between core parts, the product structure model was established, and then hierarchical clustering and association analysis were used to complete the product module division. The specific research contributions are summarized as follows:

1. The DSM was constructed to visually represent the physical structure of a complex product, including the core parts and the correlation strengths among them. This method effectively reduces the difficulty of complex product modeling and ensures the accuracy of the model.
2. The hierarchical clustering algorithm was used to obtain the initial modular scheme of mechatronic products. Through association analysis, the non-core parts were divided into modules with the strongest comprehensive association relationships.
3. We selected an electric bicycle as a case study to explain the methods described. The effectiveness of these methods in engineering applications was demonstrated.

Author Contributions: Conceptualization, S.W.; methodology, S.W.; software, G.-Y.Z.; validation, Y.-F.S. and G.-Y.Z.; formal analysis, S.W.; writing—review and editing, S.W. and Y.-F.S.; visualization, S.W.; supervision, S.W.; project administration, J.-X.M.; funding acquisition, J.-X.M. All authors have read and agreed to the published version of the manuscript.

Funding: The research reported in this article was funded by National Natural Science Foundation of China (52305071) and the Xuzhou Science and Technology Program Project (KC23010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the first author upon request.

Acknowledgments: The authors would like to thank Xuzhou University of Technology and China University of Mining and Technology for its support. The authors are sincerely grateful to the reviewers for their valuable review comments, which substantially improved the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gershenson, J.K.; Prasad, G.J.; Zhang, Y. Product modularity: Definitions and benefits. *J. Eng. Des.* **2003**, *14*, 295–313. [CrossRef]
2. Zhang, J.; Li, B.; Peng, Q.; Gu, P. Product Specification Analysis for Modular Product Design Using Big Sales Data. *Chin. J. Mech. Eng.* **2023**, *36*, 17. [CrossRef]
3. Modrak, V.; Soltysova, Z. Influence of Manufacturing Process Modularity on Lead Time Performances and Complexity. *Appl. Sci.* **2023**, *13*, 7196. [CrossRef]
4. Kremer, G.E.O.; Gupta, S. Analysis of modularity implementation methods from an assembly and variety viewpoints. *Int. J. Adv. Manuf. Tech.* **2013**, *66*, 1959–1976. [CrossRef]
5. Lai, X.; Wang, N.; Jiang, B.; Jia, T. Choosing Recovery Strategies for Waste Electronics: How Product Modularity Influences Cooperation and Competition. *Sustainability* **2024**, *16*, 9035. [CrossRef]

6. Wang, S.; Li, Z.K.; He, C.; Liu, D.Z.; Zou, G.Y. Core components-oriented modularisation methodology for complex products. *J. Eng. Des.* **2022**, *33*, 691–715. [CrossRef]
7. Fan, B.; Qi, G.; Hu, X.; Yu, T. A network methodology for structure-oriented modular product platform planning. *J. Intell. Manuf.* **2015**, *26*, 553–570. [CrossRef]
8. Zhang, N.; Yang, Y.; Zheng, Y.; Su, J. Module partition of complex mechanical products based on weighted complex networks. *J. Intell. Manuf.* **2019**, *30*, 1973–1998. [CrossRef]
9. Beek, T.; Erden, M.S.; Tomiyama, T. Modular design of mechatronic systems with function modeling. *Mechatronics* **2010**, *20*, 850–863. [CrossRef]
10. Li, S.; Mirhosseini, M. A matrix-based modularization approach for supporting secure collaboration in parametric design. *Comput. Ind.* **2012**, *63*, 619–631. [CrossRef]
11. Gelan, Y. Module partition method based on complex network theory. *J. Graph.* **2012**, *33*, 69–75.
12. Hong, Z.; Feng, Y.; Li, Z.; Tian, G.; Tan, J. Reliability-based and cost-oriented product optimization integrating fuzzy reasoning Petri Nets, interval expert evaluation and cultural-based DMOPSO using crowding distance sorting. *Appl. Sci.* **2017**, *7*, 791. [CrossRef]
13. Li, J.; Zhang, H.C.; Gonzalez, M.A.; Yu, S. A multi-objective fuzzy graph approach for modular formulation considering end-of-life issues. *Int. J. Prod. Res.* **2008**, *46*, 4011–4033. [CrossRef]
14. Xiao, R.; Chen, T. Research on design structure matrix and its applications in product development and innovation: An overview. *Int. J. Comput. Appl. Technol.* **2010**, *37*, 218–229. [CrossRef]
15. Browning, T.R. Design Structure Matrix Extensions and Innovations: A Survey and New Opportunities. *IEEE Trans. Eng. Manag.* **2016**, *63*, 27–52. [CrossRef]
16. Browning, T.R.; Eppinger, S.D. Modeling impacts of process architecture on cost and schedule risk in product development. *Eng. Manag. IEEE Trans.* **2000**, *49*, 428–442. [CrossRef]
17. Steward, D.V. The design structure system: A method for managing the design of complex systems. *IEEE Trans. Eng. Manag.* **1981**, *EM-28*, 71–74. [CrossRef]
18. Han, Z.; Mo, R.; Yang, H.; Hao, L. Module partition for mechanical CAD assembly model based on multi-source correlation information and community detection. *J. Adv. Mech. Des. Syst. Manuf.* **2018**, *12*, 1–15. [CrossRef]
19. Li, B.M.; Xie, S.Q. Module partition for 3D CAD assembly models: A hierarchical clustering method based on component dependencies. *Int. J. Prod. Res.* **2015**, *53*, 5224–5240. [CrossRef]
20. Ma, J.F.; Kremer, G.E.O. A sustainable modular product design approach with key components and uncertain end-of-life strategy consideration. *Int. J. Adv. Manuf. Technol.* **2016**, *85*, 741–763. [CrossRef]
21. Li, Y.; Chu, X.; Chu, D.; Liu, Q. An integrated module partition approach for complex products and systems based on weighted complex networks. *Int. J. Prod. Res.* **2014**, *52*, 4608–4622. [CrossRef]
22. Wang, S.; Li, Z.; He, C.; Liu, D.; Zou, G. An Integrated Method for Modular Design Based on Auto-Generated Multi-Attribute DSM and Improved Genetic Algorithm. *Symmetry* **2022**, *14*, 48. [CrossRef]
23. Kreng, V.B.; Lee, T.P. QFD-based modular product design with linear integer programming—A case study. *J. Eng. Des.* **2004**, *15*, 261–284. [CrossRef]
24. Kamrani, A.K.; Gonzalez, R. A genetic algorithm-based solution methodology for modular design. *J. Intell. Manuf.* **2003**, *14*, 599–616. [CrossRef]
25. Pandremenos, J.; Chryssolouris, G. A neural network approach for the development of modular product architectures. *Int. J. Comput. Integr. Manuf.* **2011**, *24*, 879–887. [CrossRef]
26. Kim, S.; Moon, S.K. Disassembly Complexity-Driven Module Identification for Additive Manufacturing. In Proceedings of the 24th ISPE Inc. International Conference on Transdisciplinary Engineering, Singapore, 10–14 July 2017.
27. Gu, P.; Sosale, S. Product modularization for life cycle engineering. *Robot. Comput.-Integr. Manuf.* **1999**, *15*, 387–401. [CrossRef]
28. Ji, Y.J.; Jiao, R.J.; Chen, L.; Wu, C.L. Green modular design for material efficiency: A leader follower joint optimization model. *J. Clean. Prod.* **2013**, *41*, 187–201. [CrossRef]
29. Zheng, H.; Feng, Y.X.; Tan, J.R.; Zhang, Z.X. An integrated modular design methodology based on maintenance performance consideration. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2017**, *231*, 313–328. [CrossRef]
30. Yan, J.; Feng, C.; Cheng, K. Sustainability-oriented product modular design using kernel-based fuzzy c-means clustering and genetic algorithm. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2012**, *226*, 1635–1647. [CrossRef]
31. You, Y.; Liu, Z.; Liu, Y.; Peng, N.; Wang, J.; Huang, Y.; Huang, Q. K-Means Module Division Method of FDM3D Printer-Based Function-Behavior-Structure Mapping. *Appl. Sci.* **2023**, *13*, 7453. [CrossRef]
32. AlGeddawy, T.; ElMaraghy, H. Optimum granularity level of modular product design architecture. *Cirp Ann.-Manuf. Technol.* **2013**, *62*, 151–154. [CrossRef]
33. Li, Z.k.; Wang, S.; Yin, W.w. Determining optimal granularity level of modular product with hierarchical clustering and modularity assessment. *J. Braz. Soc. Mech. Sci. Eng.* **2019**, *41*, 342. [CrossRef]

34. Li, Z.; Wei, W. Modular design for optimum granularity with auto-generated DSM and improved elbow assessment method. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2022**, *236*, 413–426. [CrossRef]
35. Li, Y.; Wang, Z.; Zhong, X.; Zou, F. Identification of influential function modules within complex products and systems based on weighted and directed complex networks. *J. Intell. Manuf.* **2019**, *30*, 2375–2390. [CrossRef]
36. Zhang, Z.J.; Lu, B.T.; Xu, X.B.; Shen, X.F.; Feng, J.; Brunauer, G. CN-MgMP: A multi-granularity module partition approach for complex mechanical products based on complex network. *Appl. Intell.* **2023**, *53*, 17679–17692. [CrossRef]
37. Liu, Z.Y.; Zhong, P.C.; Liu, H.; Jia, W.Q.; Sa, G.; Tan, J.R. Module partition for complex products based on stable overlapping community detection and overlapping component allocation. *Res. Eng. Des.* **2024**, *35*, 269–288. [CrossRef]
38. Hossain, M.S.; Chakrabortty, R.K.; Elsawah, S.; Ryan, M.J. Hierarchical joint optimization of modular product family and supply chain architectures considering sustainability. *Sustain. Prod. Consum.* **2023**, *43*, 15–33. [CrossRef]
39. Tian, G.; Sheng, H.; Zhang, L.; Zhang, H.; Fathollahi-Fard, A.M.; Zhang, X.; Feng, Y. Enhancing end-of-life product recyclability through modular design and social engineering optimiser. *Int. J. Prod. Res.* **2024**, 1–9. [CrossRef]
40. Tucker, C.S.; Kim, H.M.; Barker, D.E.; Zhang, Y. A ReliefF attribute weighting and X-means clustering methodology for top-down product family optimization. *Eng. Optim.* **2010**, *42*, 593–616. [CrossRef]
41. Newman, M.E.J.; Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E-Stat. Nonlinear Soft Matter Phys.* **2004**, *69*, 026113. [CrossRef] [PubMed]
42. Li, Y.; Wang, Z.; Zhang, L.; Chu, X.; Xue, D. Function Module Partition for Complex Products and Systems Based on Weighted and Directed Complex Networks. *J. Mech. Des. Trans. ASME* **2017**, *139*, 021101. [CrossRef]

applied sciences

MDPI

Article

Modal Analysis and Optimization of Tractor Exhaust System

Ayla Tekin [1,*] and Halil Şamlı [2]

[1] Department of Machinery and Metal Technologies, Soma Vocational School, Manisa Celal Bayar University, Manisa 45500, Türkiye
[2] Department of Mechanical Engineering, Manisa Celal Bayar University Yunusemre, Manisa 45140, Türkiye; samlihalil@gmail.com
* Correspondence: ayla.tekin@cbu.edu.tr

Abstract: Excessive vibrations in exhaust systems can significantly reduce a vehicle's lifespan and compromise performance. These vibrations, caused by factors such as engine operation and road conditions, lead to wear and tear. To address this issue, a finite element analysis (FEA) was conducted on a 90-horsepower tractor's exhaust system. Using ANSYS WB®, a 3D model was created and modal analysis was performed to determine the system's natural frequencies and mode shapes. Based on the results, geometric modifications were made to the exhaust system, increasing its stiffness and shifting vibration frequencies to higher values. Consequently, vibration levels, noise, and the risk of component failure were significantly reduced. The redesigned exhaust system was successfully implemented in production. This study demonstrates the effectiveness of FEA in analyzing exhaust system vibrations and facilitating design improvements. By extending vehicle lifespan and providing a quieter, more comfortable driving experience, this research offers valuable insights for automotive and mechanical engineers.

Keywords: tractor exhaust system; natural frequency; vibration; modal analysis; finite element method

1. Introduction

The agricultural sector plays a vital role in meeting the basic needs of human life. The use of modern production methods in agriculture is a crucial requirement to increase production efficiency. Tractor powertrain systems are essential for performing fundamental agricultural operations such as seeding, fertilizing, spraying, and harvesting. As technology advances, new methods have been employed in tractor production, and research in this field has deepened. Engines, the primary source of vibration in internal combustion engines, subject exhaust systems to intense vibrations [1]. Vibration can be defined as the periodic or random occurrence of mechanical oscillations around an equilibrium point [2]. These oscillations are usually caused by rapid linear or rotational movements of components in a system. Factors such as speed, position, load conditions, roughness of moving surfaces, and operating environment of components are important parameters that affect the vibration level [3–5]. Since rotating machinery and powertrains are considered to be the primary sources of vibration, these components as well as assembled and thin sheet parts are highly affected and damaged by vibration [6–8]. Therefore, the exhaust system is a critical subsystem that filters harmful gasses produced by the engine and reduces noise levels before they are released into the environment. Understanding the vibration characteristics of the exhaust system and taking precautions against these vibrations is essential [9]. The dynamic characterization of vehicle structures can be examined in detail through

simulations performed in a virtual environment before vehicle prototypes are produced. In this process, numerical methods such as finite element analysis can be used to analyze important characteristics such as vehicle dynamics, durability, and NVH (Noise, Vibration, and Harshness). In this way, potential problems can be controlled, and vehicle performance, life, and passenger comfort can be improved [10]. In addition, severe vibrations generated during vehicle operation can seriously threaten the health of operators, leading to various health problems [11,12].

Exhaust systems consist of multiple components that trap harmful gasses and particulates to protect the environment and reduce vehicle noise [13]. This study focuses on the design and analysis methods of exhaust systems. Studies in the literature form the reference points of the current study.

Nefske et al. [14] analyzed the exhaust system by modeling it in one dimension. They showed that the finite element model gives accurate results in predicting the low-frequency rigid body motion of the automobile exhaust system. Li et al. [15] investigated the vertical vibration characteristics of the exhaust system and developed a simplified model of the system. This study was validated by simulation and experimental comparisons. Lupea [16] observed the dynamics of an automobile exhaust system and developed a finite element model. In another study, Lupea [17] evaluated the vertical vibrations of a car exhaust system. This study stated that simplified models can be used to quickly predict system vibrations in the vertical direction. Englund et al. [18] presented the dynamic behavior of an exhaust system containing pipes, mufflers, and a catalyst with a nonlinear flexible connection. The study demonstrated the validity of the modeling by showing a strong agreement between theoretical and experimental results. Çiplak et al. [19] analyzed the NVH characteristics of a tractor exhaust system using the finite element method. The natural frequencies and mode shapes of the exhaust system were determined, and the spectral density of random vibrations occurring under different operating conditions of the tractor was examined. Avcu et al. [20] created three-dimensional models using ANSYS Workbench for the design of exhaust systems for diesel engines.

Modal analysis methods of vibrating structures include test modal analysis and finite element modal analysis [21,22]. Modal testing is usually used to verify the accuracy of finite element analysis. Considering the economy and accuracy of test analysis, how to determine the optimization object and design plan is always an important issue. Before the finite element model is updated, a modal test is performed to extract the modal parameters from the real structure [23]. The modal test is an important step in the updating process of the finite element model. Sometimes, researchers and engineers encounter difficulties in directly drawing the model in finite element analysis software due to the complexity of the structure's geometry. There are many studies in the literature that report such an approach. In these studies, exhaust systems were modeled using CATIA V-5 software and transferred to finite element analysis software [24,25]. Fouzi et al. [26] aimed to perform finite element modeling of the exhaust structure and use a model updating approach to improve its dynamic behavior. Xu et al. [27] analyzed the vibration characteristics of an automobile exhaust system using finite element software. Vibration mode diagrams were obtained in three different ways: natural, partially constrained, and fully constrained.

One of the most common problems encountered in tractors with vertical exhaust systems is the breakage and deformation of the exhaust system due to heavy working conditions. Since experimental studies to investigate the origin of such failures are costly and time-consuming, computer-aided engineering analyses are often preferred. Manda and his colleagues [28] performed modal analysis of the connecting rod, which is an important part of the automobile engine dynamic system, with the ANSYS program. They determined six points on the connecting rod and determined the magnitudes of the deformations and

which frequency values caused these deformations. They also created and displayed graphical representations of natural frequencies and deformation values. Jin et al. [29] used three-dimensional modeling software to construct a chassis model of a rice planter, and ANSYS software was used for stress analysis. Then, the modal eigenfrequency and vibration mode of the planter chassis were obtained. The obtained data were compared with the modal test results to demonstrate the effectiveness and accuracy of the finite element modal analysis method.

The aim of this study is to solve the vibration problem of the exhaust system of a 90-horsepower tractor with a vertical exhaust system. The exhaust system was designed with PTC Creo V8.0.6.0 and subjected to computer simulations using ANSYS WB® software. The vibration behavior of the exhaust system at different frequencies was investigated during the simulation process. Based on the obtained results, necessary improvements were made to the exhaust system and a new design was created. This new design was verified both by computer simulations and experimental measurements. The new exhaust system prototype was produced and the success of the design was tested in practice.

2. Materials and Methods

Exhaust systems are complex structures designed to prevent the release of harmful gasses and particulates produced by the engine into the atmosphere, generate energy, and reduce engine noise. Modal analysis is commonly used to investigate the vibration characteristics of exhaust systems. Modal analysis is a type of analysis used to determine the natural vibration frequencies, vibration modes (mode shapes), and damping parameters of a system. These analyses can be performed experimentally and numerically. In experimental modal analysis, modal parameters are obtained by conducting vibration tests on a real exhaust system. In numerical modal analysis, the vibration behavior of the exhaust system is investigated using a model created with numerical methods such as the finite element method.

This study consists of two main parts. In the first part, the tractor's vertical exhaust system was modeled in three dimensions; in the second part, the accuracy of this model was compared with the results of the modal analysis. Comprehensive investigations were carried out in both parts of the study, and the obtained results are presented in detail.

2.1. Technical Specifications of the Tractor and Its Equipment

The tractor, which forms the basis of the study, is a standard model with a double axle, rear-wheel drive, and rubber tires. The tractor has an engine power of 90 horsepower, a maximum torque of 289 Newton meters, and a maximum lifting capacity of 2200 kg for its hydraulic system.

Motor vehicles are subjected to structural vibrations induced by various sources such as road irregularities, engine/transmission mass oscillation, and exhaust system connections. Exhaust systems, in particular, are exposed to a variety of stress sources, including significant vibrations. The deformed state of the existing exhaust system used in this study is shown in Figure 1.

To identify the problem, the previously used exhaust system was examined, and the dimensions of the tractor's vertical exhaust were measured to create a suitable geometry. Based on these measurements, a new exhaust model was designed, and modal analysis was performed on this model using the ANSYS program. The components and lengths of the exhaust system are detailed in Table 1.

Figure 1. Existing Exhaust Deformation.

Table 1. Components and Lengths of the Exhaust System.

Part Name	Length
Inlet pipe (m)	0.410
Radius of the rear elbow (m) and slope angle (°)	R46/90
Radius of the outlet elbow (m) and slope angle (°)	R127/60
Outlet pipe (m)	0.537
Muffler (m)	0.572
Complete system (m)	1.596

2.2. Characterization of Materials

The material properties used in the modal analysis were determined based on the structural steel presented in Table 2.

Table 2. Material properties of a vertical exhaust.

Density	Elastic Modulus	Poisson's Ratio
7.85×10^{-6} kg/mm^3	200,000 MPa	0.3

2.3. Software Packages

A new vertical exhaust model design for a tractor was created using PTC Creo 3D software. The finite element method was applied in the ANSYS WB® R19.1 package program to create the mesh structure and perform a modal analysis of this model.

2.4. Data Acquisition System for Experimental Modal Analysis

Vibration measurements of the tractor exhaust system were conducted at the Hattat Tractor Co., Inc. R&D (Gaziosman Paşa Neighborhood, Hema Street No:6, Çerkezköy/Tekirdağ, Turkey) department during the experimental modal analysis phase. The experimental measurements were performed by connecting the accelerometer to the Hydrotechnik Multisystem 5060 Plus device via a cable (Figure 2). The equipment here was supplied by Hattat Tractor Inc. The acquired signals were transferred to a computer using the HYDROcom 6 program. The interface of this program is shown in Figure 3.

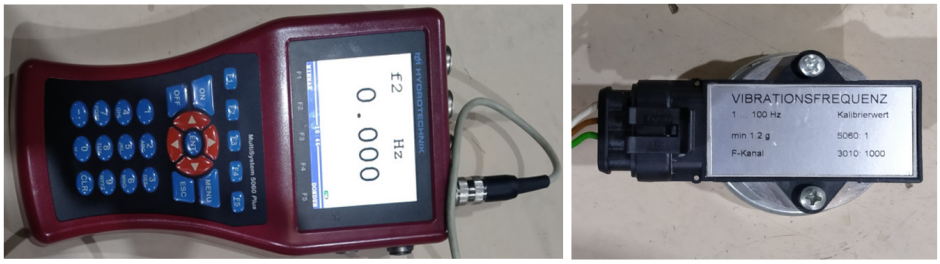


Figure 2. Hydrotechnik Multisystem 5060 Plus device and accelerometer.

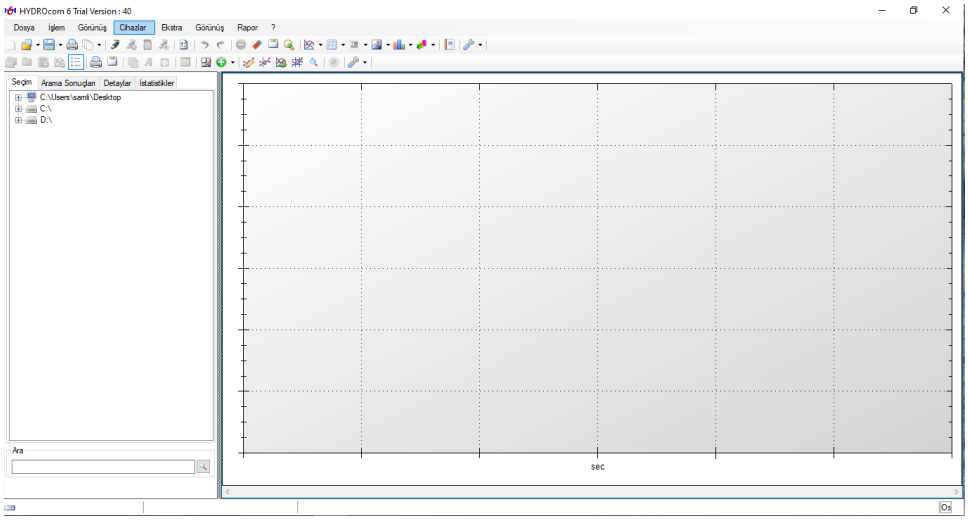


Figure 3. HYDROcom 6 software.

3. Results and Discussion

In this study, a new vertical exhaust model was designed to improve the vibration behavior of the existing exhaust system and was compared to the current model. The primary objective of these modifications was to increase the natural frequencies of the system and thus enhance its resistance to external excitation forces. The first ten natural frequencies for both models were determined using modal analysis performed in ANSYS Workbench. The obtained results confirmed the effects of the modifications on the system. Therefore, both the accuracy of the model and detailed information about the vibration behavior of the exhaust system were obtained.

3.1. Exhaust Modeling

Detailed field inspections revealed significant deformations in specific areas of the existing exhaust system (Figure 1). These deformations have been observed to shorten the system's lifespan and negatively impact its performance. The primary causes of these deformations were identified as sharp force transitions, irregular flow, and vibrations. To address these issues and extend the system's lifespan, various design alternatives were evaluated. Comprehensive analyses and simulations determined that a new design, achieved by optimizing the existing geometry, was the most suitable solution. The newly designed exhaust system was meticulously modeled using PTC Creo 3D software. Subsequently, the model was transferred to ANSYS Workbench for a comprehensive finite element analysis (FEA). The results indicated that the new design is significantly more robust than the existing one and substantially mitigates the deformation problem. Figure 4 provides detailed dimensional measurements of the newly designed exhaust model in millimeters, offering the precise information required for production. Figure 5 visually compares the geometric

differences between the existing and newly designed exhaust systems, providing a more comprehensive understanding.

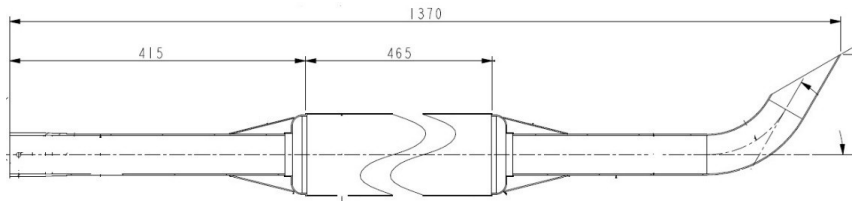


Figure 4. Modified exhaust.

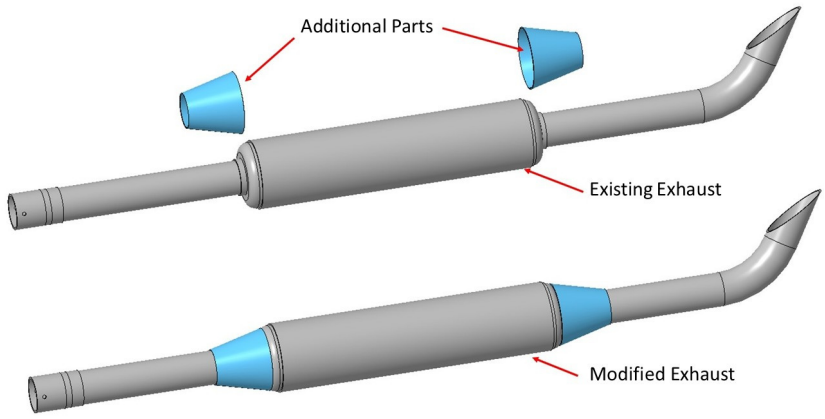


Figure 5. Three-dimensional model of vertical exhaust system.

In the modeling process, a solid model of the exhaust system was created, and this model was divided into smaller parts to form a mesh structure. Subsequently, the physical properties of the material were defined, and a modal analysis was performed. The natural frequencies and mode shapes obtained from the modal analysis are of critical importance in determining the potential for the structure to resonate [30].

3.2. Analysis of the Modified Exhaust System

ANSYS software was used to investigate the stresses experienced by the tractor's vertical exhaust system. The flowchart outlining the analysis steps performed in ANSYS is presented in Figure 6.

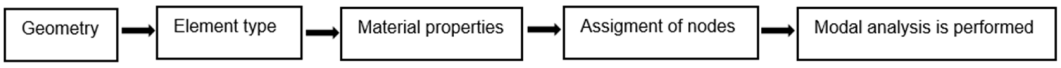


Figure 6. Flowchart of analysis performed in ANSYS.

The creation of the mesh structure, a crucial step in finite element analysis, is among the factors affecting the accuracy of the solution. In this study, the mesh structure of the exhaust system was created in the ANSYS Workbench environment and the solution parameters were defined (Figure 7). The mesh structure of the model consists of 514,382 elements and 1,018,427 nodes. Meshing of the models was carried out with the free method. The free mesh method allowed the software to automate element generation, significantly speeding up the modeling phase.

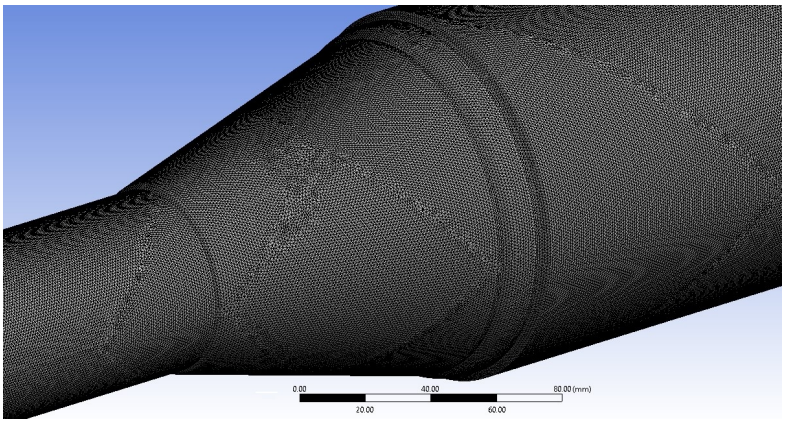


Figure 7. Mesh structure of vertical exhaust.

3.3. Modal Analysis of the Exhaust

To improve the reliability of the tractor exhaust system, a modal analysis method was used. The first ten natural frequencies were determined for both the current and modified designs using ANSYS 2022 R2 (Tables 3 and 4). The results presented in Tables 3 and 4 illustrate the vibration modes and amplitudes of the exhaust system at different frequencies. This enables us to determine at which frequencies the system is more sensitive and how design changes affect this sensitivity. Comparisons made across different frequency ranges demonstrate how the system's vibration behavior changes with frequency, allowing for a more comprehensive analysis. This has enabled us to more clearly identify the system's weak points.

Table 3. Natural frequencies of the existing (unmodified) exhaust system.

Mode	1. Mode	2. Mode	3. Mode	4. Mode	5. Mode
Frequency (Hz)	20.535	20.549	125.32	126.09	271.96
Mode	6. Mode	7. Mode	8. Mode	9. Mode	10. Mode
Frequency (Hz)	272.61	348.32	401.88	510.06	510.08

Table 4. Natural frequencies obtained after changes made to the exhaust system.

Mode	1. Mode	2. Mode	3. Mode	4. Mode	5. Mode
Frequency (Hz)	31.971	31.978	256.25	262.69	363.6
Mode	6. Mode	7. Mode	8. Mode	9. Mode	10. Mode
Frequency (Hz)	486.96	514.96	570.82	570.83	792.57

Excessive vibrations observed in the exhaust system were caused by resonance, which occurred when the forces generated during engine operation and external influences from road conditions were close to the system's natural frequencies. This condition resulted in excessive deformations and potential fractures in certain areas of the exhaust system. Design modifications made to solve the problem increased the system's natural frequencies, reducing the likelihood of resonance. As a result, the exhaust system became more resistant to external forces. The modal analysis results presented in Tables 3 and 4 and Figure 8 demonstrate the successful implementation of the design changes and significant improvements in the system's vibration behavior. These comparisons at different frequency ranges have contributed to a better understanding of the exhaust system's dynamic behavior and enabled the development of more effective solutions for future designs.

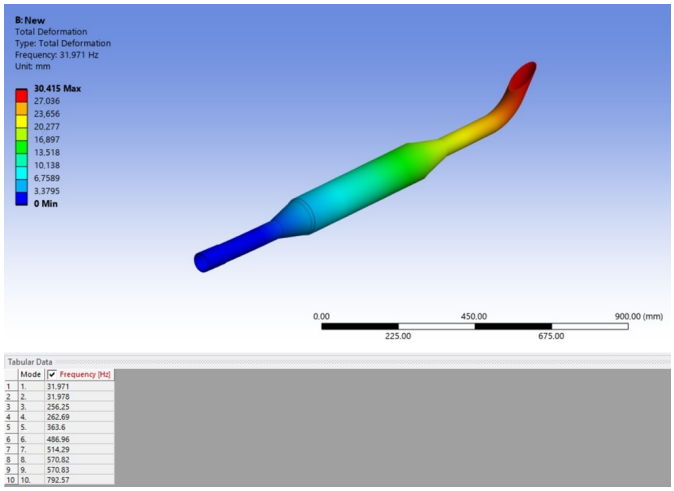


Figure 8. Modified exhaust system modal analysis, ten mod.

Figure 9 clearly demonstrates how the design modifications have affected the vibration behavior of the exhaust system. This graph clearly demonstrates how the design modifications have affected the vibration behavior of the exhaust system by comparing the natural frequencies before and after the changes. The most significant finding is that the design changes have significantly altered the natural frequencies of the system. As a result of these changes, the natural frequencies have generally shifted to higher values. This has moved the system further away from the operating frequencies of the engine, reducing the likelihood of resonance. In other words, the exhaust system is now less prone to vibration while the engine is running.

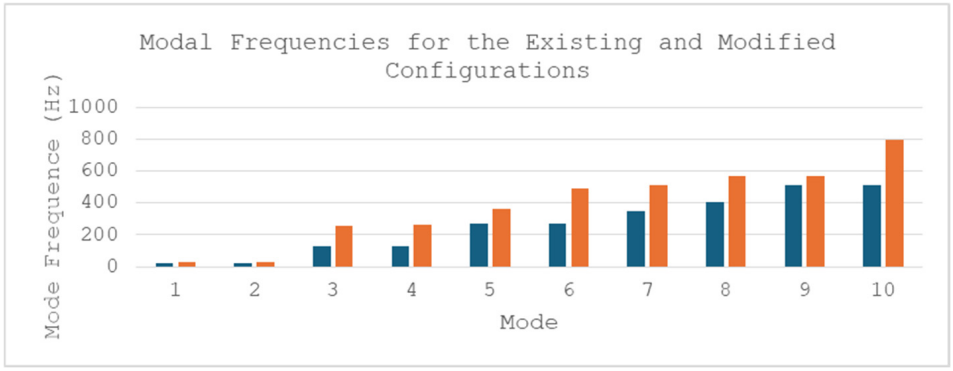


Figure 9. Modal frequency graph for existing and modified exhaust.

The primary objective of our study was to adjust the exhaust system's natural frequencies to the desired level and thereby eliminate vibration-related problems. Figure 10 visually presents the first ten vibration modes of the exhaust system according to the modal analysis results. Deformations shown in different colors reveal different levels of deformation in different parts of the system. Red areas represent maximum deformation, while blue areas represent minimum deformation. These visualizations have been valuable tools both in diagnosing the problem and in making design improvements.

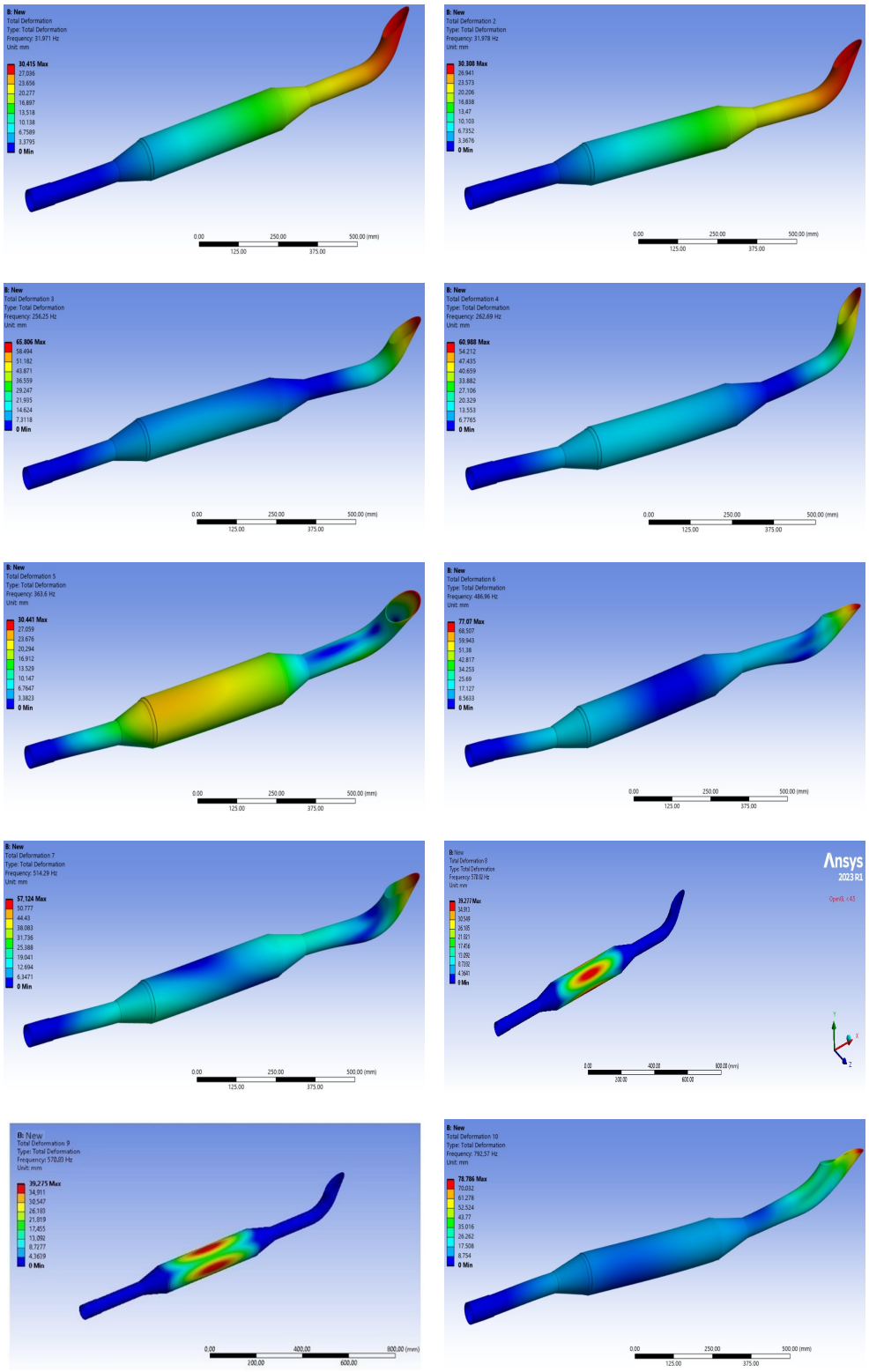


Figure 10. Modified exhaust system modal analysis graphs, ten mod.

The modal analysis results obtained from ANSYS indicated that the maximum deformation occurred at the fifth mode in the outlet section of the exhaust pipe, and this area was found to be the most prone to fracture (Figure 11). In ANSYS, red represents the maximum deformation, while blue represents the minimum deformation. The measured deformation of 32.573 mm in this area revealed the weakness in this region of the exhaust system. As a result of the improvements made by changing the geometry of the last part of

the exhaust pipe and adding additional supports, an increase in the natural frequency of the fifth mode was observed, and the total deformation in the fracture area was reduced by approximately 92% (Figure 12). Geometric modifications are design changes implemented to enhance system performance. However, these enhancements typically lead to increased system weight and cost. The design modifications implemented in this study resulted in a 13%(580 g) increase in part weight and a 5%(USD 1.5) increase in cost. Despite this minor weight gain, the system's vibration characteristics were significantly enhanced. A performance evaluation of the existing exhaust system revealed a total of 35 product returns from the field. The implemented improvements shifted the exhaust system's natural frequency away from the hazardous range, thereby increasing the system's durability and extending its lifespan, ultimately enhancing reliability. Consequently, customer complaints related to fractures were entirely eliminated, and return rates decreased to zero. These findings confirm the accuracy and effectiveness of the conducted analyses and implemented improvement processes.

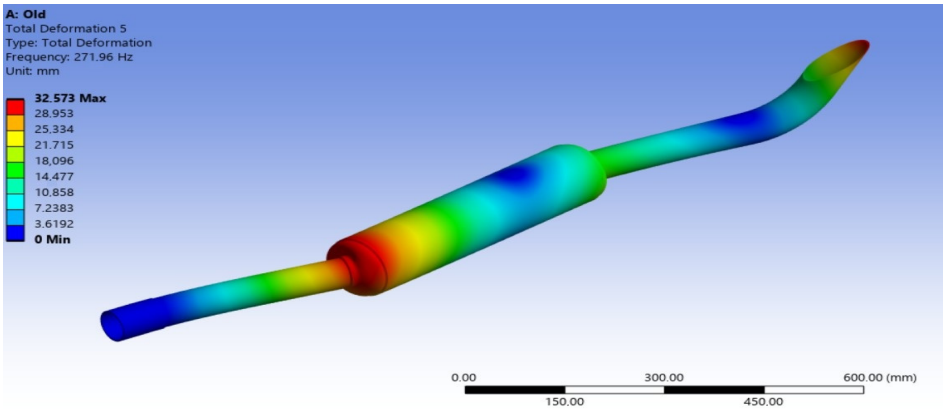


Figure 11. The total deformation of the existing exhaust in the 5th mode.

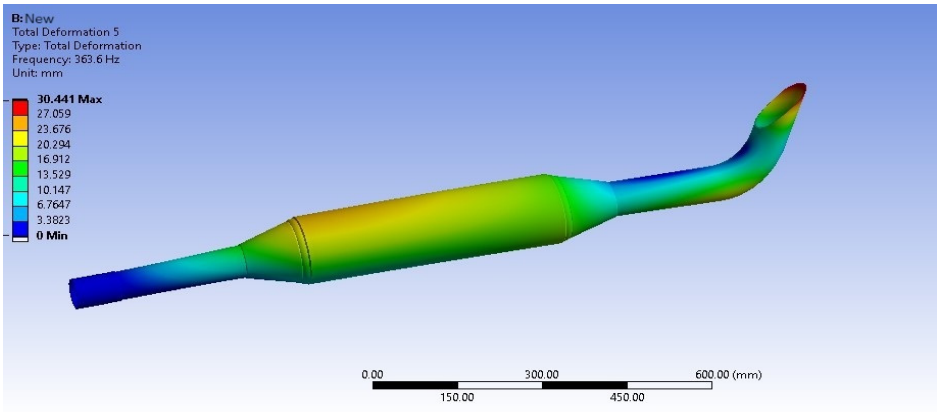


Figure 12. The total deformation of the modified exhaust in the 5th mode.

3.4. Experimental Modal Analysis

Vibration measurements were conducted using accelerometer sensors placed on the vertical exhaust system of the tractor. The data obtained from the sensors were recorded for 60 s in the 0–200 Hz frequency range. According to previous studies and based on the excitation of the exhaust system, the most important modes to induce vibration and consider are between 20 Hz and 200 Hz [31]. This frequency range is a typical range where vibrations in tractor exhaust systems are observed. The location of the sensors on the exhaust is shown in Figure 13. Differences in stiffness across various sections of the exhaust system can lead to the emergence of multiple modes and uncertain vibration

patterns. Since this situation complicates vibration analysis, a sensor was placed near the fracture point, which is the most critical region and where vibrations are most intense, to conduct measurements. This allowed for a clearer identification of the vibration source. The obtained data provided detailed information about the magnitude, frequency, and distribution of vibrations in the exhaust system.

Figure 13. A view of the accelerometer placed on the exhaust.

The vibration measurement graphs depicted in Figures 14 and 15 provide a detailed analysis of the time-varying vibration characteristics and frequency spectra of the original and modified exhaust systems operating at 1200 rpm. Taking measurements at 1200 rpm engine speed was preferred because it is in a speed range in which tractors are frequently operated, and data appropriate to real operating conditions were obtained. The modified exhaust system exhibits the emergence of new frequency components and the attenuation of certain existing ones. The peak amplitudes in Figure 14 are noticeably lower and exhibit a distinct distribution compared to Figure 15, suggesting that the modified exhaust system has resulted in reduced vibration levels or altered vibration frequencies.

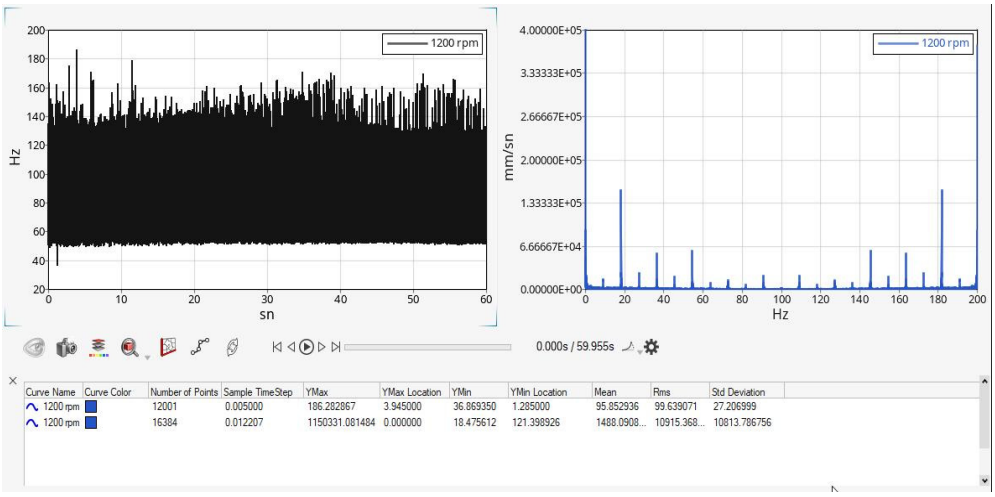


Figure 14. Experimental vibration measurement graph for existing exhaust at 1200 rpm.

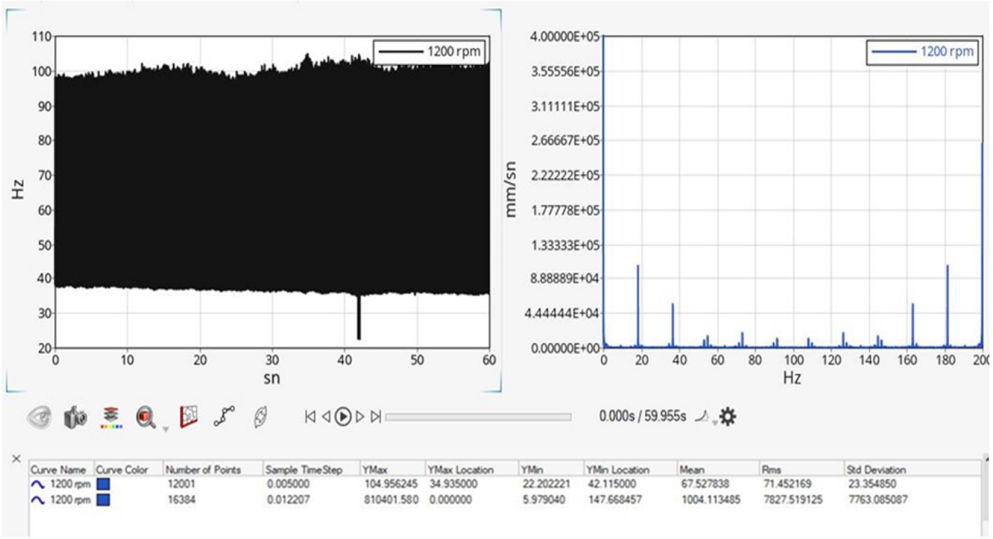


Figure 15. Experimental vibration measurement graph for modified exhaust at 1200 rpm.

4. Conclusions

In this study, the vibration characteristics of the vertical exhaust system of a 90 HP tractor were analyzed using the finite element method. The main steps of the study are as follows: First, vibration measurements were performed on the existing exhaust system using an accelerometer, and regions with permanent deformations were identified. Based on these findings, a 3D model of the exhaust system was created using PTC Creo 3D software and transferred to the ANSYS Workbench environment for finite element modeling. As a result of the modal analysis performed in ANSYS, the natural frequencies and mode shapes of the system were determined. The analyses revealed the following:

Effect of Geometric Changes: It was determined that changes in the geometry of the exhaust system (pipe diameter, length, bending angles, and connection points) affected the vibration behavior. These changes caused the system to vibrate at higher frequencies and increased its resistance to external forces. As Xu et al. [27] also stated in their research, vibrations in exhaust systems are generally concentrated at low frequencies. This situation is consistent with the results of our study, once again confirming the effect of the geometric properties of the exhaust system on vibration behavior.

Mode Shapes: Mode shapes provide information to the designer about which areas of the system undergo more deformation, indicating where to focus attention. In this way, the overall performance of the exhaust system has been improved.

Modal Analysis Results: Modal analyses performed with ANSYS simulations showed that the natural frequency increased, especially in the fifth mode, and the deformation in critical regions decreased by 92%. These findings indicate that the exhaust system has been moved away from dangerous resonance frequencies.

Experimental Verification: Measurements taken at 1200 rpm showed that some old frequencies disappeared and new frequency components emerged in the modified exhaust system. This situation indicates that the change in the exhaust system has changed the vibration characteristics of the vehicle. Modifications have reduced some vibrations but created new vibration sources. This also shows that the changes in the exhaust system have affected the vibrations of the vehicle.

The ANSYS 2023 R1 Workbench software, as verified by experimental studies, provides high accuracy, cost effectiveness, and time savings in the design of exhaust systems. Natural frequency analyses, in particular, are critical for optimizing system performance and preventing failures caused by unwanted vibrations. For these analyses to provide

reliable results, a realistic numerical model of the exhaust system is essential. The structural modifications made were verified by computer simulations and experimental measurements on the actual exhaust system. These verifications show that the design changes made are effective not only theoretically but also in practical application.

As in the study of Jin et al. [29], a high agreement was observed between the data obtained as a result of two different modal analyses. This result shows that both experimental and numerical methods provide reliable and consistent results for preventing vibration-induced damage in tractor exhaust systems. These findings constitute an important reference point for solutions to vibration problems in other engineering systems with similar structural dynamics. In this context, the findings of the study once again demonstrate the critical importance of vibration analysis and structural improvements in the design of agricultural machinery and other heavy vehicles. It is considered that future studies will contribute to obtaining more comprehensive results by applying similar methods under different exhaust system and vibration conditions.

Author Contributions: Conceptualization, Writing—Review and Editing, Supervision, A.T.; Methodology, Resources, Analysis, H.Ş. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Acknowledgments: We would like to thank Hattat Tractor Inc. for the technical support they provided throughout our study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Abdelrahman, A.A.; Esen, I.; Özarpa, C.; Shaltout, R.; Eltaher, M.A.; Assie, A.E. Dynamics of perforated higher order nanobeams subject to moving load using the nonlocal strain gradient theory. *Smart Struct. Syst.* **2021**, *28*, 515–533. [CrossRef]
2. Anegundi, R.; Tegginmani, V.; Battepati, P.; Tavargeri, A.; Patil, S.; Trasad, V.; Jain, G. Prevalence and characteristics of supernumerary teeth in a non-syndromic South Indian pediatric population. *J. Indian Soc. Pedod. Prev. Dent.* **2014**, *32*, 9–12. [CrossRef] [PubMed]
3. Gao, Z.; Xu, L.; Li, Y.; Wang, Y.; Sun, P. Vibration measure and analysis of crawler-type rice and wheat combine harvester in field harvesting condition. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 48–55. [CrossRef]
4. Xu, L.; Li, Y.; Sun, P.; Pang, J. Vibration measurement and analysis of tracked-whole feeding rice combine harvester. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 49–55. [CrossRef]
5. Rabbani, M.A.; Tsujimoto, T.; Mitsuoka, M.; Inoue, E.; Okayasu, T. Prediction of the vibration characteristics of half-track tractor considering a three-dimensional dynamic model. *Biosyst. Eng.* **2011**, *110*, 178–188. [CrossRef]
6. Jin, X.; Chen, K.; Ji, J.; Zhao, K.; Du, X.; Ma, H. Intelligent vibration detection and control system of agricultural machinery engine. *Meas. J. Int. Meas. Confed.* **2019**, *145*, 503–510. [CrossRef]
7. Liu, Z.; Yuan, S.; Xiao, S.; Du, S.; Zhang, Y.; Lu, C. Full Vehicle Vibration and Noise Analysis Based on Substructure Power Flow. *Shock Vib.* **2017**, *2017*, 8725346. [CrossRef]
8. Chowdhury, M.; Islam, M.N.; Iqbal, M.Z.; Islam, S.; Lee, D.H.; Kim, D.G.; Jun, H.J.; Chung, S.O. Analysis of overturning and vibration during field operation of a tractor-mounted 4-row radish collector toward ensuring user safety. *Machines* **2020**, *8*, 77. [CrossRef]
9. Gedikli, H.; Heyal, Y.; Bayraktar, A.; Türker, T. Bir egzoz sisteminin doğal frekans ve mod şekillerinin deneysel ve sayısal yöntemle belirlenmesi. In Proceedings of the XVI Ulusal Mekanik Kongresi, Kayseri, Türkiye, 22–26 June 2009.
10. Hanouf, Z.; Faris, W.F.; Nor, M.J.M. Dynamic characterisation of vehicle structural panels. *Int. J. Veh. Noise Vib.* **2015**, *11*, 199–209. [CrossRef]
11. Kabir, M.S.N.; Chung, S.O.; Kim, Y.J.; Sung, N.S.; Hong, S.J. Measurement and evaluation of whole body vibration of agricultural tractor operator. *Int. J. Agric. Biol. Eng.* **2017**, *10*, 248–255. [CrossRef]

12. Martínez-Aires, M.D.; Quirós-Priego, J.; López-Alonso, M. Analysing worker exposure to WBV at the doñana biological reserve (Spain). A case study. In *Advances in Safety Management and Human Factors*; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2019.
13. Parr, O.; Krüger, J.; Vilser, L. Intake and exhaust systems. In *Handbook of Diesel Engines*; Mollenhauer, K., Tschöke, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 387–400.
14. Nefske, D.J.; Sung, S.H.; Feldmaier, D.A. *Correlation of a Beam-Type Exhaust System Finite-Element Model for Vibration Analysis*; SAE Technical Papers; SAE International: Warrendale, PA, USA, 2003.
15. Li, S.B.; Guan, X.Q.; Lu, T.L.; Zhang, J.W. Vehicle modeling with the exhaust system and experimental validation to investigate the exhaust vertical vibration characteristics excited by road surface inputs. *Int. J. Automot. Technol.* **2008**, *9*, 483–491. [CrossRef]
16. Lupea, I. Updating of an exhaust system model by using test data from ema. *Proc. Rom. Acad. Ser. A—Math. Phys. Tech. Sci. Inf. Sci.* **2013**, *14*, 326–334.
17. Iulian, L. An Exhaust System Lumped Model—Identification and Simulation. *Acta Tech. Napoc.-Ser. Appl. Math. Mech. Eng.* **2016**, *59*, 359–364.
18. Englund, T.L.; Wall, J.E.; Ahlin, K.A.; Broman, G.I. Significance of non-linearity and component, Internal Vibrations in an Exhaust System. In Proceedings of the 2nd WSEAS International Conference on Simulation, Modelling and Optimization, Skiathos, Greece, 25–28 September 2002.
19. Çıplak, A.Y.; Bilgin, A.S.; Kullukçu, A. Egzoz sisteminin nvh karakteristiğinin sonlu elemanlar benzetimi ile çıkartılması. In *Teknolojik Gelişmeler ve Mühendislik Uygulamaları*; Güven Plus Grup A.Ş. Yayınları: İstanbul, Türkiye, 2022; pp. 89–104. Available online: https://www.guvenplus.com.tr/imagesbuyuk/b7f89TEKNO.pdf (accessed on 13 February 2025).
20. Avcu, M.; Teke, M.; Kopuz, Ş. Mtu 16V 4000 M90 Brand/Model Diesel Engine Exhaust System Design. *J. Nav. Sci. Eng.* **2010**, *6*, 39–58.
21. Zhang, J.; Yang, C.; Zhang, L.; Jiang, Y.; Wang, C. Analysis and experiment on strength and vibration characteristics of corn stubble plucking mechanism. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 72–78. [CrossRef]
22. Gao, F.; Han, L.J.; Liu, X. Vibration spectroscopic technique for species identification based on lipid characteristics. *Int. J. Agric. Biol. Eng.* **2017**, *10*, 255–268. [CrossRef]
23. Avitabile, P. Experimental modal analysis a simple non-mathematical presentation. *Sound Vib.* **2001**, *35*, 20–31.
24. Shojaeifard, M.H.; Ebrahimi-Nejad R, S.; Kamarkhani, S. Optimization of Exhaust System Hangers for Reduction of Vehicle Cabin Vibrations. *Int. J. Automot. Eng.* **2017**, *7*, 2314–2325.
25. Thomson, W.T. *Theory of Vibration with Applications*, 4th ed.; CRC Press: London, UK, 2018.
26. Fouzi, M.S.M.; Sani, M.S.M.; Muchlis, Y. Finite Element Modelling and updating of welded joint for dynamic study of exhaust structure. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *469*, 012099. [CrossRef]
27. Xu, J.M.; Zhou, S.; Chen, S.X. An analysis of the vibration characteristics of automotive exhaust systems and optimization of suspension points. *Open Mech. Eng. J.* **2014**, *8*, 574–580. [CrossRef]
28. Manda, M.; Kola, R.; Karunakarreddy, K. Modal analysis of a connecting rod using ANSYS. *Int. J. Mech. Eng.* **2017**, *4*, 30–35. [CrossRef]
29. Jin, X.; Cheng, Q.; Tang, Q.; Wu, J.; Jiang, L.; Wu, C.; Wang, H. Research on vibration reduction test and frame modal analysis of rice transplanter based on vibration evaluation. *Int. J. Agric. Biol. Eng.* **2022**, *15*, 116–122. [CrossRef]
30. Fu, Z.F.; He, J. *Modal Analysis*; Elsevier: Oxford, UK, 2001; pp. 1–11.
31. Noorazizi, M.S.; Aminudin, B.A.; Zetty, M.I. Systematic fea study of vehicle exhaust system hanger location using addofd method. *Appl. Mech. Mater.* **2013**, *392*, 161–164. [CrossRef]

applied sciences

MDPI

Article

Robotic Cell Layout Optimization Using a Genetic Algorithm

Raúl-Alberto Sánchez-Sosa [1] and Ernesto Chavero-Navarrete [2,*]

[1] Posgrado CIATEQ AC, Centro de Tecnología Avanzada, Querétaro 76150, Mexico; raul.sanchez@meau.com
[2] CIATEQ AC, Centro de Tecnología Avanzada, Querétaro 76150, Mexico
* Correspondence: ernesto.chavero@ciateq.mx; Tel.: +52-442-196-15-00

Featured Application: This work contributes to different industrial sectors through a computer tool that optimizes the layout of a robotic cell, which allows reducing both production times and energy consumption, mainly.

Abstract: The design of the work area of a robotic cell is currently an iterative process of trial and improvement, where, in the best cases, the user places the workstations and robotic manipulators in a 3D virtual environment to then semi-automatically verify variables such as the robot's reach, cycle time, geometric interferences, and collisions. This article suggests using an evolutionary computation algorithm (genetic algorithm) as a tool to solve this optimization problem. Using information about the work areas and the robot's reach, the algorithm generates an equipment configuration that minimizes the cell area without interference between the stations and, therefore, reduces the distances the robotic manipulator must travel. The objective is to obtain an optimized layout of the workstations and to validate this optimization by comparing the transfer times between stations with the actual times of an existing screwdriving cell. As a result, the transfer time was reduced by 9%. It is concluded that the algorithm can optimize the layout of a robotic cell, which can lead to significant improvements in efficiency, quality, and flexibility.

Keywords: genetic algorithm; robotics; optimization

1. Introduction

An assembly line is a manufacturing process where different parts are assembled in a specific sequence, creating a final product at each workstation until the final operation is completed. Operations performed at workstations can be manual, automatic, or semi-automatic, and the type of operation to be used at the station will depend on the manufacturer's strategy [1].

Industry tries to increase its competitiveness by implementing strategies of greater flexibility at a low cost, obtaining greater productivity and efficiency. The use of robots in automated material handling boosts productivity and enhances automation. Robotic cells, which combine robots equipped with grippers or other specialized capabilities along with CNC machines and complementary systems, are an integral part of these manufacturing environments. By integrating industrial robots with other technologies, these automation systems provide productive flexibility, allowing manufacturers to efficiently adapt to short- and medium-term demand fluctuations [2]. On a global scale, there is no standardized process for designing the work area when implementing robotic manufacturing cell solutions. The increasing demand for flexible cells with high productivity and product quality necessitates faster and more efficient design and planning methods to create an appropriate work area layout for production tasks [3].

Having a computer program that assists in the location of workstations will allow the industry to design robotic manufacturing cells with the shortest transit time between stations, choose the ideal robot for the tasks, and maximize the productivity rate by minimizing production cycle time. Additionally, it will enable seamless integration of different

systems, improve operational efficiency, reduce setup and reconfiguration times, enhance adaptability to changes in production demands, and ensure optimal use of available space. This will not only increase throughput but also reduce energy consumption and operational costs, thereby contributing to a more sustainable and cost-effective manufacturing process [4].

Optimizing a robotic cell is a critical process in modern manufacturing that seeks to maximize efficiency and productivity. It consists of designing and adjusting the components and their arrangement within the cell to minimize cycle time, reduce energy consumption, and improve product quality. Traditionally, robotic cell design involved virtual prototyping, with physics-based simulations playing a crucial role in obtaining accurate models. These simulations allowed the behavior and interaction of components to be predicted in a controlled environment, making it easier to identify potential failures and areas for improvement before physical implementation [5,6]. The use of advanced algorithms is now included, for example, in [7] mixed programming techniques, which are presented for small instances and a genetic algorithm for large instances. It is modeled as a production flow problem with blocking constraints, a single transport robot, and controllable processing times. This research addressed the case where processing times vary linearly according to the allocated resources and focuses on obtaining maximum performance. However, comparative results with real solutions are not presented. In [8], hierarchical optimization is proposed, being more important for first the posture optimization and then the motion optimization. The pose optimization is solved with a genetic algorithm, and an objective function considers the design constraints. It is confirmed that the proposed method can solve the optimization problem quickly by experiments. A reference robotic cell was used, and only the execution time of the algorithm was compared against the manual design time of the cell, but no improvements in cycle time were documented. In [9], design criteria for a robotic cell are defined, design candidates are represented by a sequence pair scheme to avoid interference between components of the assembly system, and the use of dummy components is proposed to represent design areas where components are scarce. Objective functions are formulated, and optimal solutions are obtained using a genetic algorithm. Numerical evaluations are performed to illustrate the effectiveness of the proposed method. The existence of a relationship between design area and manipulability is concluded. That is, too small a design area reduces manipulability, since the robot arm must undergo more radical movements during assembly operations. In [10], the location of the work centers and the robot are parameterized as the homogeneous transformation matrix with respect to the environment. Each component is assigned a cylindrical envelope, and the interference is modeled with the circular projection. The robot joints are also projected to resolve collisions in 2D space. The deepest collision point is removed outside the cylindrical envelope along the vertical or radial direction. And the trajectory is segmentally replanned according to the relocated point. The optimization is achieved in a cascade flow, and the heuristic algorithm is applied. Experiments performed on a robotic ultrasonic shot peening work cell validate the effectiveness of the proposed method. No improvements in process times are reported. In [11], five nature-inspired algorithms, a genetic algorithm, differential evolution, an artificial bee colony, a charge search system, and particle swarm optimization are proposed. Design area criteria, operation time, and robot manipulability are simultaneously optimized. Numerical examples are provided to illustrate the effectiveness and usefulness of the proposed methods. It is observed that swarm optimization algorithms perform better than the other algorithms in terms of solution quality. Although the three criteria are optimized simultaneously, this optimization is independent, the optimal solution of the system is still pending, and only the best solutions are proposed. The use of evolutionary algorithms is an area of research for robotic optimization, and the following table shows some other research that has been developed in the last decade, each with different optimization objectives and, therefore, different algorithm proposals for their solution.

Based on the advancements and state-of-the-are approaches presented in Table 1, the research question is as follows: Can genetic algorithms optimize the workspace and reduce cycle time in a robotic cell? Additionally, do they offer any advantages due to their ability to handle large search spaces, adapt to complex nonlinear solutions, and find near-global optimum solutions efficiently, even in dynamic and highly variable environments?

Table 1. Robotic cell optimization research (2014–2022).

Year	Author	Objective	Solution	Restrictions
2014	Daoud et al. [12]	Maximize line efficiency and balance tasks between robotic equipment.	Three proposed hybrid evolutionary algorithms	A single product in the robotic cell
2015	Mukund Nilakantan et al. [13]	Minimize cycle time and energy consumption	Integer programming model 0–1	A single product in the robotic cell
2016	Cil et al. [14]	Minimize the total cost of robot use, number of stations, and cycle time.	Hierarchical objective preventive algorithm	Various products in the robotic cell and fixed robots.
2017	Nilakantan et al. [15]	Maximize line efficiency and minimize carbon footprint	Multi-objective coevolutionary algorithm, artificial bee colony, random simulated annealing, and fast elitist non-dominated sorting	A single product on the line, fixed stations and robots.
2018	Pereira et al. [16]	Minimize Fixed and Variable Costs	Memetic elitism algorithm, genetic algorithm, multiple start algorithm, and random search algorithm.	A single product in the robotic cell, fixed stations and robots.
2019	Weckenborg et al. [17]	Maximizing manual labor efficiency and productivity	Multi-integer programming model with hybrid genetic algorithm.	A single product in the robotic cell with deterministic cycle times
2020	Zhou and Wu [18]	Minimize fixed and variable costs	Hybrid particle swarm combined with dynamic programming.	A single product in the robotic cell, fixed stations and robots.
2021	Mehmet Pinarbasi et al. [19]	Minimize number of stations and cycle time	Constraint programming model with mixed integer programming and ABSALOM software.	A single product in the robotic cell
2022	Yuanying Chi et al. [20]	Minimize number of stations and energy consumption	Mixed integer linear programming model with a cross-station design.	A single product in the robotic cell

In this article, a multi-objective design optimization method for the layout of workstations in a robotic cell is proposed. The robot base is considered the center of the work area, and the 2D dimensions of both the workstations and the robot base are specified. A genetic algorithm is employed to evaluate numerous possible configurations of components, selecting the most efficient ones. The objective function aims to minimize the work area while considering the interference of the stations and the robot base as constraints. The algorithm is adjusted by modifying the mutation probability.

This approach enables the algorithm to minimize the work area, reduce cycle times, lower energy consumption, and maximize productivity, which would be difficult to achieve with traditional methods. Therefore, this algorithm is proposed as a design tool for production engineers and automation consultants in the industry seeking to implement robotic solutions and optimize robotic cells cost-effectively and efficiently.

2. Materials and Methods

The methodological approach of this research is based on the analysis of a production process involving three robotic cells already installed in the industry. Each cell is treated as an independent study object to experiment with the proposed algorithm. The goal is to obtain an optimized distribution of the workstations and validate this optimization by comparing the transfer times between stations with the current times.

For each robotic cell, a specific number of workstations of different sizes, an industrial robot, and a limited workspace are considered, as illustrated in Figure 1. A genetic algorithm is employed to minimize the workspace of each cell, also incorporating a penalty function to evaluate the presence of interferences between the workstations and the robot's base. In the design of the robotic cell, the robot's position must allow it to reach all the workstations, so the algorithm considers the robot's base as the center of the operational area.

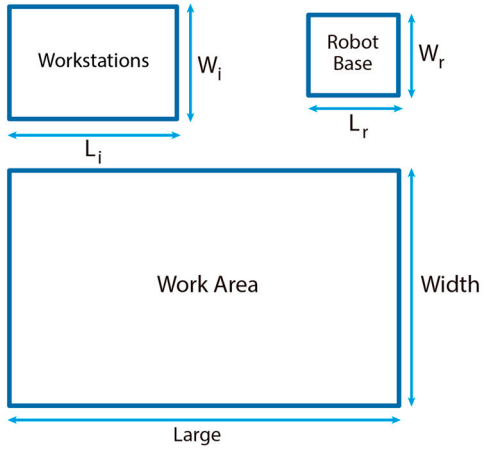


Figure 1. Input variables considered in the genetic algorithm for work area optimization.

The algorithm development was carried out using Python V3.5 [21] in the Visual Studio Code V1.91 environment [22]. Unlike previous works [7–20], this study introduces a significant novelty by integrating Mitsubishi's RT TOOLBOX3 PRO software [23] into the simulation stage. Instead of merely mathematically modeling the robot's trajectories and cycle times, the results generated by the algorithm are directly imported into RT TOOLBOX3 PRO, allowing for a realistic simulation of the robot's movements. This not only provides a more accurate calculation of cycle times but also offers practical visualization and validation of the robot's behavior in the simulated environment, adding a level of robustness and precision not found in similar approaches.

Furthermore, the integration of RT TOOLBOX3 PRO enables better solution comparison, optimizing trajectories and times in simulated real-world scenarios, which strengthens the applicability of the results in industrial settings.

2.1. Robotic Cell

A robotic cell is an integrated system used in manufacturing and industrial automation that consists of one or more robots, along with other equipment and tools, designed to perform specific tasks such as assembly, welding, painting, material handling, and inspection. These cells are engineered to improve efficiency, precision, and safety in production processes. A robotic cell is an integrated system used in manufacturing and industrial automation that consists of one or more robots, along with other equipment and tools, designed to perform specific tasks such as assembly, welding, painting, material handling, and inspection. These cells are engineered to improve efficiency, precision, and safety in production processes [24]. Typically, a work cell can only process one workstation at a time, so the robotic cell is considered as a mass production system with blocking. And according to their layout, they can be classified into the following [25]:

- Linear or Semicircular: The robot passes through each of the workstations sequentially moving from the input to the output of the process and back, as shown in Figure 2.

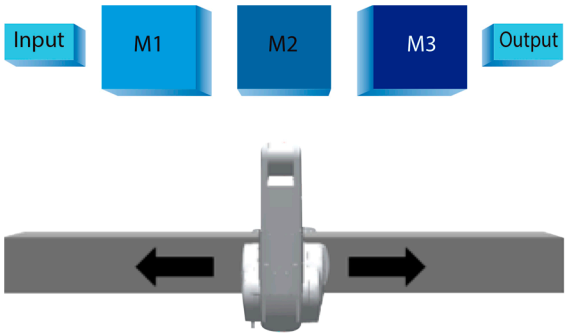


Figure 2. Linear or semicircular robotic cell.

- Circular: In this configuration, the robot is required to pass through each of the workstations sequentially, as shown in Figure 3. Having greater flexibility in the sequence of movements results in higher productivity.

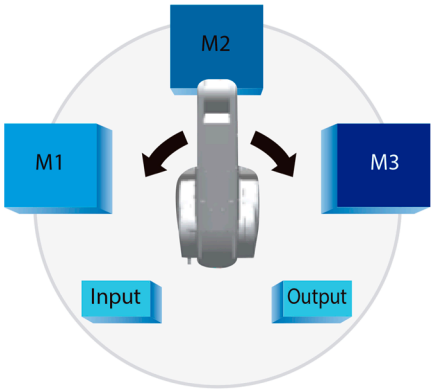


Figure 3. Circular robotic cell.

2.2. Work Area Optimization

Optimizing a work area is a research problem that not only applies to manufacturing systems but also to the design of integrated circuits, assemblies, etc. A search for works was carried out to define the optimization criteria.

One of the first investigations to address the optimization of the area of an industrial robot was [26], where an automatic system in three dimensions was presented to generate collision-free trajectories using conventional algorithms of flexible manufacturing systems. For this work, 3 degrees of freedom were considered for the robot. Later in [27], a genetic algorithm was used. The number of workstations in defined areas within the robot's range was defined, and the system determined the best order of the workstations to be executed, while another algorithm adjusted the location of the workstations with respect to the efficiency of the robot.

On the other hand, in [28], a genetic algorithm was used to minimize the cycle time of a series of operations, which was achieved by determining the relative positions of the machines or workstations around the industrial robot.

2.3. Genetic Algorithm

A genetic algorithm is a search heuristic inspired by the principles of natural selection and genetics. It is commonly used to find approximate solutions to optimization and search problems and was used in this research to optimize the work cell area through functions that minimize the number of interferences between the workstations.

Here is an overview of how genetic algorithms work [26]:

- Population initialization: A set of candidate solutions (called individuals) is generated. Everyone represents a potential solution to the problem.
- Selection: Each individual will be evaluated using a fitness function, which measures how well it solves the problem. Individuals are selected for reproduction based on their fitness scores. Higher fitness individuals have a higher probability of being selected. Common selection methods include roulette wheel selection, tournament selection, and rank-based selection.
- Crossover: Selected individuals are paired to create offspring. The genetic material from two parent individuals is exchanged at the crossover point, producing two new offspring.
- Mutation: With a low probability, individual bits in the offspring's chromosomes are flipped or altered. Mutation introduces genetic diversity and helps prevent premature convergence to local optima.
- Replacement: The offspring replace some or all the old population, creating a new generation. The process of selection, crossover, and mutation is repeated for many generations.

Figure 4 shows the typical cycle of the evolution of a genetic algorithm.

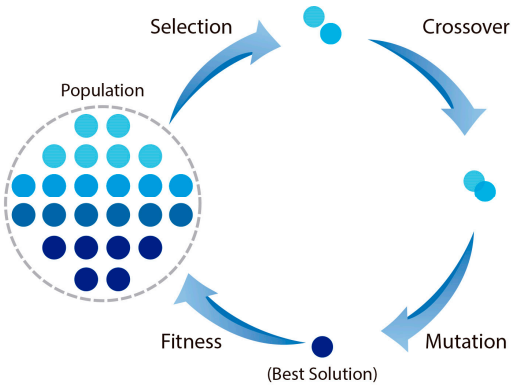


Figure 4. Cycle of the genetic algorithm.

Some recent examples of the use of the genetic algorithm in optimization applications for industrial robot problems are optimal movement trajectories in industrial robots, aiming to minimize the objective function of the manipulator's velocity rate, producing the highest possible speed at the end-effector while keeping the axis speeds at a minimum [27], with stochastic multi-modal processing times with multiple parallel-working robots per workstation. The objective is to minimize the number of workstations at a given production rate and the probability limit of violating the cycle time [28].

The proposed algorithm uses an uninformed initialization based on assigning random values to the genes of everyone. In this case, the representation is real, so each gene will take values in a defined interval with a uniform probability.

$$\text{Station/Gen} = [(0, \text{Working area in long}), (0, \text{Working area in width})] \quad (1)$$

The solution is a combinatorial optimization problem whereby each station location is a gene and, together with the other station locations, together form the "genetic code of the individual".

$$Individual = \{Gen_1, Gen_2, Gen_3, \ldots, Gen_n\} \; where \; each \; Gen_n \; is \{x, y\} \quad (2)$$

$$Robot \; base = [\{x_1, y_1\}, \{x_2, y_2\}] \quad (3)$$

With the initial population, the code generates the points that correspond to the corner or diagonal of the stations to have the complete area of each one of them, as shown in Figure 5.

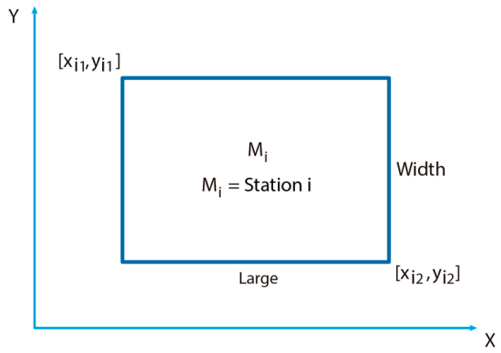


Figure 5. Generation of workstation areas.

Obtaining the minimum and maximum values [{X_{min}, Y_{min}}, {X_{max}, Y_{max}}], the area of each combination generated is calculated and then evaluated in the first fitness function to be minimized. The second fitness equation is to minimize the number of interferences between stations, since this would be considered a physical constraint for the solution. The fitness equations to be minimized in the genetic algorithm are shown below.

$$f(area) = (X_{max} - X_{min}) * (Y_{max} - Y_{min}) \quad (4)$$

$$f(M_i) = \sum(x1_{max} \geq x2_{min} \wedge x2_{max} \geq x1_{min} \wedge y1_{max} \geq y2_{min} \wedge y2_{max} \geq y1_{min}) \quad (5)$$

where M_i is a working station. With two objective functions to be minimized, it is considered a multi-objective problem. The advantage of this multi-objective approach is that it allows adding more functions or constraints that help to find new solutions with different conditions in the optimization of the work area. Our objective function to minimize will be the sum of both Equations (4) and (5):

$$f(Comb) = (area\_weight) * f(area) + (interference\_weight) * f(M_T) \quad (6)$$

where *area_weight* is a value between 1 to 10, $f(area)$ is a total area of the combination of stations in the working area, *interference_weight* is a value between 100 to 1000, and $f(M_T)$ is the total number of interferences between stations. Two weight variables are used to give preference to interference reduction over area reduction.

With the evaluation of each combination by area and total number of interferences, a selection by tournament is made, which consists in the realization of λ tournaments, where the individuals of the current population are randomly selected by sampling with a uniform probability in such a way that the individuals with the minimum value are selected to be the parents in the next stage.

For the crossing of parents, we use the point crossover where we randomly generate a number *n* between 1 and the length of the vector (genes) that represents the individual and P_c (crossover probability). The first child receives the first genes of the first identical parent up to *n*, and after this number it receives the segment of the second parent. For the second child, it is also generated by this crossing, and it is conducted in the same way by inverting the order. Now, the second parent transfers its segment up to a number n to the child and the remaining one for the segment of the first parent.

Once the new population of offspring is generated, the mutation stage is performed. For this case, we will use the mutation for the real representation, where each gene is modified with probability P_m (probability of mutation). This consists of randomly selecting

a gene and assigning it a new value from a uniform probability distribution over the range in which it is defined. The new generation becomes the parents for the next cross, repeating the process from the evaluation of the population.

In the genetic algorithm, the variables of number of generations, crossover probability, mutation probability, population size, and number of stations will directly affect the processing time of the algorithm. Algorithm 1 shows the genetic algorithm developed.

Algorithm 1: Robotic cell layout optimization algorithm

Initialize. Number of generations
  while $generatios < counter$:
  Random starting positions of stations
    for $index\ in\ range\ (population_{size})$
    $\{Gen_1, Gen_2, Gen_3, \ldots, Gen_n\}$
  Diagonal points of each station
    while $flag == 0$
      $each\ Gen_n$ is$\{x, y\}$
  Maximum and minimum points for calculating the area
    for $rows\ in\ range\ (population\_size)$
      print (X min, $X\ max$, $Y\ min$, $Y\ max$)
    Maximum area of each combination of stations
      $$f(area) = (X_{max} - X_{min}) * (Y_{max} - Y_{min})$$
    Interference counting between stations
      if (x1_max >= x2_min and x2_max >= x1_min and y1_max >= y2_min and y2_max >= y1_min): return interference
      $$f(M_i) = \sum Number\ of\ interferences\ between\ stations$$
    Fitness function
      $$f_n(Comb) = (area\_w) * f(area) + (interference\_w) * f(M_T)$$
    Best solution selection
      if $f_n(comb) < f_{n-1}(comb)$ then
  New generation of solution
    for $rows\ in\ range\ (population\_size/\ 2)$
Best solution
  return: $f_{min}(area)$

3. Results

The genetic algorithm was applied to three robotic cells that are part of a production line, with each cell representing a different case study. Results from different case studies help ensure that the conclusions are more robust, generalizable, and applicable across a wide range of contexts.

The algorithm is parameterized to perform 1000 iterations to find the best solution, with a convergence criterion applied if the minimum area for the robotic cell does not change after 50 iterations. A mutation probability of 0.20 is provided to ensure genetic diversity and that the algorithm sufficiently explores the solution space. This parameter was obtained through experimentation and adjustment of the algorithm, allowing for a balance between exploring new solutions and exploiting the best solutions found so far.

The experimentation was conducted with an 11th Gen Intel® Core™ i7-11800H @ 2.30 GHz processor and 64 GB of RAM (Santa Clara, CA, USA). With this hardware, the average time per solution was 20 min, considering that, as a combinatorial optimization algorithm, it is considered a polynomial time problem.

3.1. Screwing Cell—Case Study 1

The first robotic cell performs a screwing function using a 4-axis robot (RH-6CRH6020-D SCARA, Mitsubishi Electric, Tokyo, Japan). There are six pieces that are placed within a nest of positions. The location of the pieces and the screws are currently defined, and each piece to be screwed are represented with a rectangle of different color to be clearly visualized, as shown in Figure 6. Table 2 provides the dimensions of each workstation.

The algorithm is executed to obtain an optimal distribution of the stations in the screwing cell. Figure 7 shows the part distribution proposed by the algorithm, revealing a vertical matrix arrangement. The parts are more compact and vertically aligned in relation to the robot's position. Table 3 presents the current travel times performed by the robot and the times obtained through simulation based on the proposed solution.

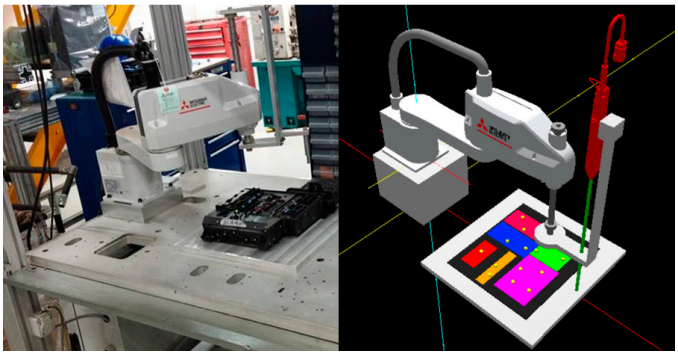


Figure 6. Robotic screwing cell.

Table 2. Dimensions of screwing stations.

Stations	Length (mm)	Width (mm)
Workstation	600	500
Station 1	175	80
Station 2	170	85
Station 3	140	120
Station 4	195	180
Station 5	120	60
Station 6	170	34

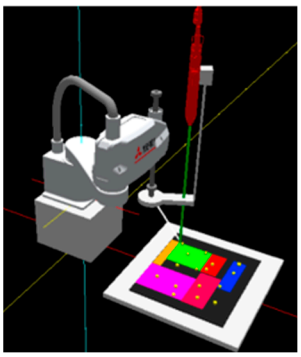


Figure 7. Screwing stations distribution proposed by the algorithm.

Table 3. Trajectory time, screwing cell.

Trajectory	Current Transfer Time (s)	Algorithm Transfer Time (s)
1 to 2	3.12	2.91
2 to 3	1.92	1.53
3 to 4	2.15	2.22
4 to 5	1.98	1.71
5 to 6	2.48	1.88
Total time	11.65	10.25

Once the simulation was carried out, a reduction of 1.4 s was obtained, which represents 13.17% of the transfer time implemented in the currently installed robotic cell compared to the optimized cell proposed by the algorithm.

3.2. Machining Cell—Case Study 2

The second cell is responsible for loading and unloading parts in a machining center, which produces three types of shafts of different lengths and diameters. A 6-axis robot (RV8CRLD-S15 M, Mitsubishi Electric, Tokyo, Japan) loads the material to be processed and then unloads the finished product, placing it in a storage station corresponding to the produced model. Figure 8 shows the current layout of the machining center; station 1 is the loading (input) station, and stations 2, 3, and 4 are the storage (output) stations. It shows how the CNC is in opposite position to the input and output stations. Table 4 provides the dimensions of each workstation.

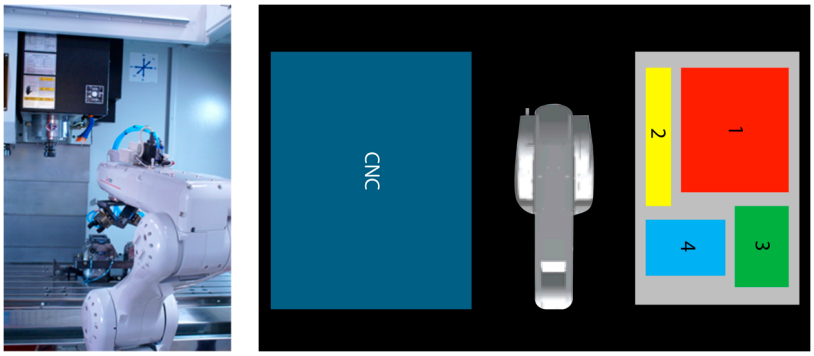


Figure 8. Robotic machining cell.

Table 4. Dimensions of machining stations.

Stations	Length (mm)	Width (mm)
CNC	750	600
Station 1	350	325
Station 2	400	150
Station 3	275	225
Station 4	260	197

Figure 9 shows the solution proposed by the algorithm, where the storage stations are grouped on the left side of the nest, and the loading station is positioned on the right. Additionally, it can be observed that the loading station involved in each cycle is closer to the robot, which will reduce the cycle time for each loading operation. Table 5 presents the current travel times performed by the robot and the times obtained through simulation based on the proposed solution. A reduction of 1.06 s was achieved, representing a 6.8% decrease in the transfer time of the currently installed robotic cell compared to the optimized cell proposed by the algorithm.

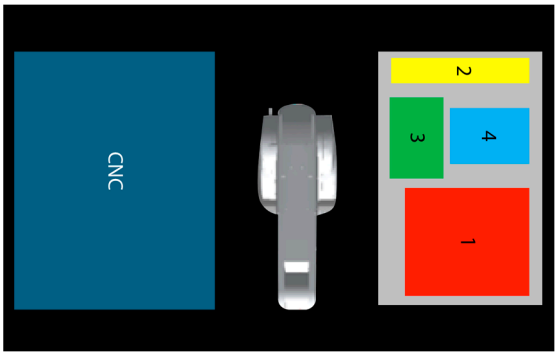


Figure 9. Machining station distribution proposed by the algorithm.

Table 5. Trajectory time, machining cell.

Trajectory	Current Transfer Time (s)	Algorithm Transfer Time (s)
CNC to 1	4.56	4.10
2 to CNC	3.45	7.37
3 to CNC	3.89	3.21
4 to CNC	3.71	3.49
Total time	15.61	14.55

3.3. Assembly Cell—Case Study 3

The assembly of the parts is carried out in the third cell, where the material stations are distributed in such a way that a 6-axis robot (RV-7FRD, Mitsubishi Electric, Tokyo, Japan) takes the material from each station, places the parts in the assembly station, and removes the finished product. Figure 10 shows the actual layout of the assembly cell, a circular cell where the stations with the individual components are arranged around the base of the robot. Table 6 provides the dimensions of each workstation.

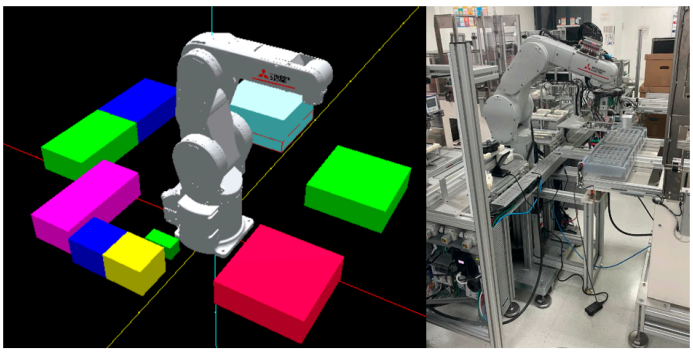


Figure 10. Robotic assembly cell. Case study 3.

Table 6. Dimensions of assembly stations.

Stations	Length (mm)	Width (mm)
Assembly 3	750	600
Station 1	350	325
Station 2	400	150
Station 4	275	225
Station 5	260	197

In Figure 11, the proposed solution by the algorithm is shown. In Figure 11, the proposed solution by the algorithm is shown. Table 7 presents the current travel times performed by the robot and the times obtained through simulation based on the proposed solution.

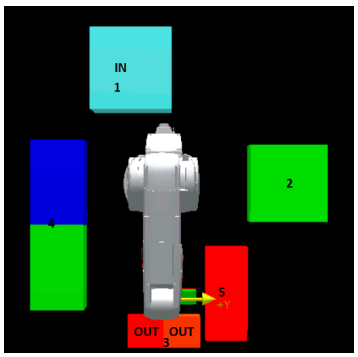


Figure 11. Assembly station distribution proposed by the algorithm.

Table 7. Trajectory time, assembly cell.

Trajectory	Current Transfer Time (s)	Algorithm Transfer Time (s)
1 to 3	2.78	2.40
2 to 3	2.61	2.38
4 to 3	2.90	2.10
5 to 3	2.41	2.84
Total time	10.70	9.72

A reduction of 0.98 s in transfer time was achieved, representing a 9.15% decrease compared to the currently operating robotic cell. This improvement was made possible through the optimization performed by the algorithm, which adjusted the station layout and the robot's movement paths within the cell. Table 8 summarizes the results obtained for each of the robotic cells evaluated, highlighting the superiority of the proposed method with a 3.44 s improvement in production time.

Table 8. Trajectory time, production line.

Robotic Cell	Current Transfer Time (s)	Algorithm Transfer Time (s)
Screwing	11.65	10.25
Machining	15.61	14.55
Assembly	10.70	9.72
Total time	37.96	34.52

The results demonstrate a significant improvement in the operational efficiency of the robotic cell due to the optimization of the workstation layout using a genetic algorithm. The new arrangement allowed for a notable reduction in transfer times, achieving a decrease in operational cycles by 9.06%. This optimization not only speeds up the production process but also maximizes the utilization of available resources, resulting in increased productivity and greater competitiveness on the production line. The use of the genetic algorithm has been key to exploring multiple possible configurations and finding the optimal solution that would otherwise be difficult to identify manually.

4. Discussion

The optimization of the robotic cell's area using the genetic algorithm led to a more compact and efficient design. The algorithm was able to identify station layouts that reduced the required space without compromising the robot's mobility or the cell's functionality, which is especially beneficial in manufacturing environments with space constraints. However, the algorithm does not account for spaces needed for operator movement, so it is recommended to consider these factors when designing a robotic cell. A limitation of the use of the algorithm is that it doesn't consider the possibility of the rotation of the workstations as a possible solution. This approach of no rotation is very useful for the algorithm, as, with the intersections, it is more difficult to count with rotation workstations, which only a few researchers have performed, like Zhang and Li [29]. Some further research its need it to evaluate if rotation of workstations needs to be considered.

Optimizing the travel times and the area of the cell can result in lower operational costs, as both energy consumption and robot wear due to unnecessary movements were reduced. This translates into increased equipment lifespan and greater efficiency in resource utilization. This algorithm of this work only considers as a valid solution that the location of the workstation is within the reach of the robot, while other works considered the Denavit–Hartenberg equations [30] of a proposed robot; this gives an advantage, since the solution is evaluated as reachable or not from the same algorithm. For future work, we will seek to introduce the D–H equations of various robot models so that the user has the ability to evaluate their solutions without having to simulate them.

Regarding the limitation of the algorithm using only one robot per robotic cell, this significantly reduces the complexity of the problem but also limits the optimization potential. In more advanced industrial environments, where multiple robots can work simultaneously in the same cell or in adjacent cells, the interactions between the robots offer opportunities to improve efficiency and reduce cycle times. However, optimizing such systems with multiple robots introduces a new level of complexity, as it is necessary to manage the coordination between robots to avoid collisions and minimize downtime.

The implementation of the genetic algorithm can enhance the overall productivity of the robotic cell. By minimizing travel times and maximizing space utilization, the number of production cycles completed within a given period increased, which benefits the overall production capacity. This means that it is not necessary to continue performing the traditional approach as the only way to design and determine the location of the workstations, which now can be assisted with an algorithm to receive some location proposals for the workstations. Also, it is important to evaluate the use of some other evolutionary algorithms, as, according to some research, Lim [11] found that Particle Swarm Optimization could lead to better results in robotic cell optimization.

5. Conclusions

The proposed genetic algorithm effectively optimized the travel times between stations, resulting in a significant reduction in cycle times. This is made possible by the inherent ability of genetic algorithms to explore multiple configurations and find solutions that minimize the robot's downtime, thereby improving operational efficiency. In the case studies conducted, the algorithm showed great adaptability to different configurations and design requirements, allowing for the optimization of both simple and complex robotic cells, which is crucial for production lines that need to quickly adjust to changes in demand or product specifications.

However, despite its advantages, the use of genetic algorithms also presents certain limitations. One of the main drawbacks is that, although genetic algorithms are effective in exploring large search spaces, they do not guarantee finding the globally optimal solution. Since their search process is based on the evolution and mutation of potential solutions, they often become trapped in local optima, especially in highly complex problems. This can lead to solutions that, while sufficiently good, do not always represent the best possible configuration. Furthermore, the performance of the algorithm heavily depends on the quality of the fitness function and the parameters selected, such as population size or mutation and crossover rates, which must be carefully tuned to avoid both premature convergence and excessively prolonged exploration of the solution space.

Another limitation is that genetic algorithms often require a considerable amount of computational time, particularly when applied to more complex systems, such as production lines with multiple robots or irregularly shaped workstations. Although the technique has proven to be applicable to various industrial scenarios, from small robotic cells to complex production lines, scalability can be a significant challenge. As the number of variables and possible configurations increases, the time needed to evaluate each generation of solutions grows exponentially, which can make real-time optimization impractical in certain cases.

Additionally, the genetic algorithm not only identified the optimal configuration but also generated a series of sub-optimal solutions that offer a good balance between space and time. These alternative solutions can be useful for robotic cell designers, as they provide options that allow balancing different design criteria according to the specific needs of production. However, the availability of multiple solutions also presents the challenge of selecting the most suitable one for a particular scenario, which may require further analysis outside the algorithm's framework.

Future research could focus on addressing these limitations by developing hybrid techniques that combine genetic algorithms with other optimization methods, such as gradient-based or local search algorithms. Moreover, increasing the complexity of the

robotic cells included in the studies, incorporating more robots and diverse shapes of workstations, would be beneficial. This would allow for the generation of more comprehensive layout templates, helping robotic cell designers consider different arrangements based on the number of robots and the specific geometry of the workstations, further improving efficiency in complex industrial production environments.

Author Contributions: Conceptualization, R.-A.S.-S. and E.C.-N.; methodology, R.-A.S.-S. and E.C.-N.; software, R.-A.S.-S.; validation, R.-A.S.-S. and E.C.-N.; formal analysis, R.-A.S.-S. and E.C.-N.; investigation, R.-A.S.-S.; resources, E.C.-N.; data curation, R.-A.S.-S.; writing—original draft preparation, R.-A.S.-S.; writing—review and editing, E.C.-N.; visualization, R.-A.S.-S.; supervision, E.C.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank the CIATEQ graduate program for the support provided in carrying out this research work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chutima, P. A comprehensive review of robotic assembly line balancing problem. *J. Intell. Manuf.* **2022**, *33*, 1–34. [CrossRef]
2. Vaisi, B. A review of optimization models and applications in robotic manufacturing systems: Industry 4.0 and beyond. *Decis. Anal. J.* **2022**, *2*, 100031. [CrossRef]
3. Raffaeli, R.; Bilancia, P.; Neri, F.; Peruzzini, M.; Pellicciari, M. Engineering Method and Tool for the Complete Virtual Commissioning of Robotic Cells. *Appl. Sci.* **2022**, *12*, 3164. [CrossRef]
4. Ribeiro, F.M.; Pires, J.N.; Azar, A.S. Implementation of a robot control architecture for additive manufacturing applications. *Ind. Robot* **2019**, *46*, 73–82. [CrossRef]
5. Raffaeli, R.; Neri, F.; Peruzzini, M.; Berselli, G.; Pellicciari, M. Advanced virtual prototyping of robotic cells using physics-based simulation. *Int. J. Interact. Des. Manuf.* **2024**, *18*, 981–996. [CrossRef]
6. Holubek, R.; Delgado-Sobrino, D.R.; Košťál, P.; Ružarovský, R. Offline Programming of an ABB Robot Using Imported CAD Models in the RobotStudio Software Environment. *Appl. Mech. Mater.* **2014**, *693*, 62–67. [CrossRef]
7. Al-Salem, M.; Kharbeche, M. Throughput optimization for the Robotic Cell Problem with Controllable Processing Times. *Oper. Res.* **2017**, *51*, 805–818. [CrossRef]
8. Kuratani, R.; Kojima, T.; Fujii, H.; Matoba, S.; Saitoh, Y.; Takanishi, K. Hierarchical Optimization for Robotic Cell Systems. *Trans. Inst. Syst. Control. Inf. Eng.* **2022**, *35*, 118–125. [CrossRef]
9. Izui, K.; Murakumo, Y.; Suemitsu, I.; Nishiwaki, S.; Noda, A.; Nagatani, T. Multiobjective layout optimization of robotic cellular manufacturing systems. *Comput. Ind. Eng.* **2013**, *64*, 537–544. [CrossRef]
10. Qiu, B.; Chen, S.; Gu, Y.; Zhang, C.; Yang, G. Concurrent layout and trajectory optimization for robot workcell toward energy-efficient and collision-free automation. *Int. J. Adv. Manuf. Technol.* **2022**, *122*, 263–275. [CrossRef]
11. Lim, Z.Y.; Ponnambalam, S.G.; Kazuhiro, I. Nature inspired algorithms to optimize robot workcell layouts. *Appl. Soft Comput.* **2016**, *49*, 570–589. [CrossRef]
12. Daoud, S.; Chehade, H.; Yalaoui, F.; Amodeo, L. Solving a robotic assembly line balancing problem using efficient hybrid methods. *J. Heuristics Syst.* **2014**, *20*, 235–259. [CrossRef]
13. Nilakantan, M.J.; Ponnambalam, S.G.; Jawahar, N.; Kanagaraj, G. Bio-inspired search algorithms to solve robotic assembly line balancing problems. *Neural Comput. Appl.* **2015**, *26*, 1379–1393. [CrossRef]
14. Çil, Z.A.; Mete, S.; Ağpak, K. A goal programming approach for robotic assembly line balancing problem. *IFAC-PapersOnLine* **2016**, *49*, 938–942. [CrossRef]
15. Nilakantan, J.M.; Li, Z.; Tang, Q.; Nielsen, P. Multi-objective co-operative co-evolutionary algorithm for minimizing carbon footprint and maximizing line efficiency assembly line systems. *J. Clean. Prod.* **2017**, *156*, 124–136. [CrossRef]
16. Pereira, J.; Ritt, M.; Vásquez, O.C. A memetic algorithm for the cost-oriented robotic assembly line balancing problem. *Comput. Oper. Res.* **2018**, *99*, 249–261. [CrossRef]
17. Weckenborg, C.; Kieckhäfer, K.; Müller, C.; Grunewald, M.; Spengler, T.S. Balancing of assembly lines with collaborative robots. *Bus. Res.* **2019**, *13*, 93–132. [CrossRef]

18. Zhou, B.; Wu, Q. Decomposition-based-bi-objective optimization for sustainable robotic assembly line balancing problems. *J. Manuf. Syst.* **2020**, *55*, 30–43. [CrossRef]
19. Pinarbasi, M.; Alakas, H.M. Balancing stochastic type-II assembly lines: Chance-constrained mixed integer and constraint programming models. *Eng. Optim.* **2020**, *52*, 2146–2163. [CrossRef]
20. Chi, Y.; Qiao, Z.; Li, Y.; Li, M.; Zou, Y. Type-1 Robotic Assembly Line Balancing Problem That Considers Energy Consumption and Cross-Station Design. *Systems* **2022**, *10*, 218. [CrossRef]
21. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, USA, 2009. Available online: https://www.python.org/ (accessed on 25 August 2024).
22. Microsoft. Visual Studio Code. 2024. Available online: https://code.visualstudio.com/updates/v1_91 (accessed on 25 August 2024).
23. Mitsubishi Electric Corporation. RT ToolBox3 Pro. 2024. Available online: https://www.mitsubishielectric.com/fa/products/rbt/robot/smerit/rt3/index.html (accessed on 25 August 2024).
24. Geismar, H.N.; Sriskandarajah, C.; Ramanan, N. Increasing throughput for robotic cells with parallel Machines and multiple robots. *IEEE Trans. Autom. Sci. Eng.* **2004**, *1*, 84–89. [CrossRef]
25. Groover, M.P. *Automation, Production Systems, and Computer-Integrated Manufacturing*, 4th ed.; Pearson Higher Education, Inc.: Hoboken, NJ, USA, 2015; pp. 366–389.
26. Yang, X.S. (Ed.) Genetic Algorithms. In *Nature-Inspired Optim*, 2nd ed.; Elsevier: London, UK; Middlesex University London: London, UK, 2021; pp. 91–100. [CrossRef]
27. Valsamos, H.; Nektarios, T.; Aspragathos, N.A. Optimal placement of path following robot task using genetic algorithms. *IFAC Proc. Vol.* **2006**, *39*, 13–137. [CrossRef]
28. Stade, D.; Manns, M. Robotic Assembly Line Balancing with Multimodal Stochastic Processing Times. In *Advances in Automotive Production Technology—Towards Software-Defined Manufacturing and Resilient Supply Chains. Stuttgart Conference on Automotive Production SCAP 2022. ARENA2036*; Kiefl, N., Wulle, F., Ackermann, C., Holder, D., Eds.; Springer: Cham, Switzerland, 2022. [CrossRef]
29. Zhang, J.; Li, A. Genetic Algorithm for Robot Workcell Layout Problem. In Proceedings of the 2009 WRI World Congress on Software Engineering, Xiamen, China, 19–21 May 2009; Volume 4, pp. 460–464. [CrossRef]
30. Barral, D. Optimization Tool for Assembly Workcell Layout. Google Patents US6470301B1 United States. Available online: https://patents.google.com/patent/US6470301B1/en (accessed on 25 August 2024).

applied sciences

MDPI

Article

Development of a Simulation Model to Improve the Functioning of Production Processes Using the FlexSim Tool

Wojciech Lewicki [1,*], Mariusz Niekurzak [2,*] and Jacek Wróbel [3]

[1] Faculty of Economics, West Pomeranian University of Technology in Szczecin, 71-210 Szczecin, Poland
[2] Faculty of Management, AGH University of Krakow, 30-067 Krakow, Poland
[3] Department of Bioengineering, West Pomeranian University of Technology in Szczecin, 71-210 Szczecin, Poland; jacek.wrobel@zut.edu.pl
* Correspondence: wojciech.lewicki@zut.edu.pl (W.L.); niekurz@agh.edu.pl (M.N.)

Abstract: One of the goals of Industry 4.0 is to increase the transparency of the value chain through modern tools in production processes. This article aims to discuss the possibility of increasing the efficiency of a production system by modernizing it with the use of computer modelling tools. This article describes a method for the simulation modelling of a selected production system using the specialized FlexSim 2023 software in a 3D environment. The results and benefits of the practical application of the object-oriented modelling are presented, as well as the possibilities of collecting simulation data used to optimize production processes. The analyses were conducted at a selected production plant in a case study. The research assessed the effectiveness of the existing system and determined the impact of process changes in the event of the introduction of a new design solution. The simulation identified bottlenecks in the material flow. The basis for creating the simulation model was the analysis of the technological process. A simulation model for a real situation was created, and a simulation model was designed to identify and indicate a solution to eliminate the detection of the bottleneck. The problem area identified using visualization in the technological process slowed down the entire production process and contributed to time and economic losses. Thus, the authors confirmed the thesis that the simulation modelling of production systems using the FlexSim program can help eliminate bottlenecks and increase the efficiency of human resource use. At the same time, the use of this tool can lead to increased efficiency, reduced costs and improved sustainability and other performance indicators important for modern production environments as part of the promoted Industry 4.0 idea. A noticeable result of these changes was an increase in production from about 80–90 units. In addition, it was noticed that the condition of the machines preceding the stand changed.

Keywords: production systems; Industry 4.0; FlexSim; production planning; 3D model; optimization; simulation

1. Introduction

As observations of production environments indicate, one common phenomenon is the constant pursuit of the continuous improvement of production processes [1]. There is a constant demand for the development of new solutions to improve the efficiency and harmonization of production systems [2]. In particular, this practice has intensified in the era of challenges posed by the increasingly popular idea of Industry 4.0 [3]. In practice, the use of modern optimization tools is expected to lead to an increase in efficiency while reducing costs [4]. When browsing the offers of simulation software manufacturers, it can be assumed that they have very advanced functions and extensive object libraries that enable the very precise construction of the desired model and its actual operation. In the case of production systems, these tools can be used to conduct experiments on existing production lines and to examine the possibility of undertaking optimization and

harmonization projects of already operating production lines [5]. With the use of a properly selected tool, it is possible to model these processes without the need to have physical devices constituting their elements. As indicated in the literature on the subject, the obtained results can illustrate the operation of the system over a selected time course, showing its behavior in the event of a failure, and they also enable the analysis of areas defined as problematic [6]. A very common subject of research on functioning production systems is the location of the so-called bottleneck of production processes. In the available literature, many managers and engineers of manufacturing companies are looking for effective methods to eliminate this phenomenon. By addressing this issue, it is possible to streamline production processes and increase their system efficiency while reducing costs. This optimization and harmonization can be achieved with the right simulation tools. Their proper use allows for manufacturing companies to conduct a thorough process analysis and implement beneficial modifications [7]. However, it is worth bearing in mind that using the above-mentioned solution also has disadvantages. The purchase of appropriate software is associated with relatively high financial outlays, especially in the case of detailed models, where creating a correct model can be very time-consuming and cost-consuming [8,9].There is also a risk of making a mistake at the design stage of the model due to a lack of real data, resulting in erroneous results. It is worth noting that there are also very advanced production environments for which it may be impossible to build a correct simulation model [10,11]. However, in most cases, a correct approach to modelling and optimizing production processes based mainly on the use of these tools and real data for the analysis of production systems can bring many systemic benefits in terms of the harmonization and optimization of production. The course and logic of the simulation modelling procedure and its relationship to a real system are shown in Figure 1.

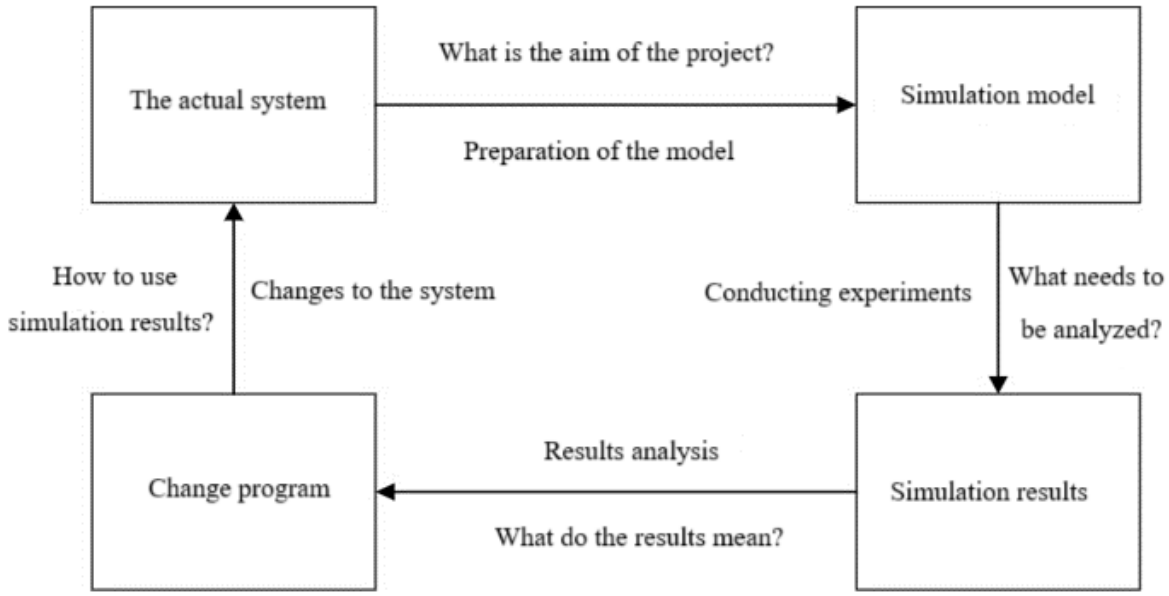


Figure 1. General diagram of simulation modelling in production processes [12].

At this stage of consideration, it is important to understand that every manufacturing company may have many sources of waste, e.g., overproduction, defective products, unnecessary inventory, failures, excessive storage, etc. Each of these phenomena has a major impact on the efficiency of the production line and, consequently, on sales revenues. This is an important argument that motivates the use of modelling and simulation tools to eliminate waste from areas at risk. The ideal tool for solving this type of problem seems to be the specialized FlexSim software. FlexSim is an analytical tool used in practice in process design and decision-making [13]. In the FlexSim simulation software, three-dimensional models can be created that map a real system, e.g., a production system.

The considerations presented in this paper focus on the development of a simulation model of the production process of a given product—lighting poles—based on real data obtained from one of the main manufacturers of these structures in Europe. Achieving this goal required the use of the specialized FlexSim software [14]. After preparing the model and analyzing the results obtained, an attempt was made to improve the process by increasing one of the key parameters, i.e., the production efficiency. The research process was used to confirm the hypothesis that the simulation software is an effective tool for investigating potential solutions to improving existing production systems. In addition, the use of computer simulations can lead to a reduction in the time it takes to conduct experiments on production line models. The simulation itself can lead to a reduction in the costs necessary in the process of testing solutions for already functioning systems in the production environment.

To verify the research hypotheses, it was decided to experiment with a production system operating in one of the largest companies in Europe dealing with, among others, the production of lighting poles. After defining the research problem, its implementation and verification of the results obtained were carried out. During the work, the specialized FlexSim software was used, which is a tool for building models and simulating the functioning of systems. The final stage of the research and analysis was to try to improve the process by proposing changes to the bottleneck and examining their impact on the system.

After reviewing the literature from the last 5 years, the authors noticed that researchers had not undertaken the use of simulation tools to optimize and harmonize the production processes of lighting poles. This is a diversified production process, which is why individual methods of approaching this type of project have always been used. This solution is very time-consuming and economically unjustified. Therefore, this article undertook a case study and complements research and analysis of the practical use of FlexSim tools to solve complex problems and design systems of analyzed cases.

This article is divided into the following sections: Section 1 provides an introduction to the topic. Section 2 contains a review of the literature on the development of simulation models for the production process of the analyzed product. Section 3 describes the research methodology used to analyze the production process. Section 4 contains the results of the research and numerical experiment and an extensive commentary on them. Section 5 contains the research conclusions and perspectives for further development for analysis and research on this topic.

2. Literature Review of the Analyzed Problem

In the available literature on process modelling and simulation, many items that explain this issue and indicate its various aspects can be found [15]. Many papers present the application of a particular method or software along with the results obtained by a specific researcher. The first research focused on the theoretical aspects of simulation and modelling. The paper [16] contains definitions of modelling and illustrates the advantages and disadvantages of using this tool to investigate potential solutions. Reference [17] focuses on the issue of model construction. The authors indicate the purpose of modelling, methods of implementation and stages of its preparation [18]. The next analyzed items present the stages of building a simulation model with the use of computer software [19]. This is an extremely important issue because in today's production environment, this solution is widely used due to the significant complexity of the systems operating in enterprises. In [20], the authors developed one of the most important elements of model building, which is validation. It is of key importance in system analysis because it is used to check the correctness of the implemented project, which translates into the reliability of the results obtained [21]. In addition, this study provides information on the areas of application of simulation modelling and then presents the issue of using this tool in the aspect of production process management [22]. In Ref. [23], simulation models are divided according to their characteristics. Other studies [24] focused on the use of this tool in companies producing a particular product. They present individual production areas in

which simulation modelling can be used and describe the stages of production preparation based on simulation models. They describe the methods of analysis of the functioning system through the selection of an appropriate tool and the implications of the results obtained in the real production environment.

In a narrower approach to simulation tools, researchers focus on issues related to the modelling and simulation of production processes themselves. In some articles [6], the authors present the course of the work and the analysis of the results obtained using the ARENA software © 2024 Rockwell Automation [7]. However, reference [8] focuses the authors' attention on the mathematical aspects of building models. Simulation methods are presented here as well as an experiment illustrating how the results obtained should be considered. Production processes are complex and multi-stage activities, consisting of various tasks and phases [25]. However, the organization of processes in the conditions of dynamic changes occurring in a company's environment is very difficult and problematic, because in order to maintain the continuity of the production process, work must be properly designed and rationally organized [26]. Another researcher emphasizes that, currently, the availability of new digital technologies shapes the business environment of enterprises and creates new opportunities for development by introducing innovations to the structures within Industry 4.0 [27]. These innovations refer to the possibility of blurring the boundaries between physical and virtual entities [28]. Digitalization is forcing traditional industrial organizations to rethink and develop existing business models [29]. Modelling connections between operations and analyzing real scenarios for a given process, taking into account the internal and external environment of the enterprise, is possible using computer simulations and specialized tools [30]. Solving various technical problems using simulation modelling is effective, and its effectiveness has been confirmed, among others [31], in the research of one of the authors [32]. Other studies [33] show that a properly designed production process can reduce operating costs in a company by up to 50%. The effect of modelling processes, objects and phenomena is to obtain results in the form of temporal dependencies of statistical data [34]. In serial production, simulation modelling can be used to determine the most important parameters that will make the technological process more efficient and effective [35]. Another researcher emphasizes that the modelling of production processes is possible by using computer software to conduct experiments and tests. The author's research focuses primarily on modelling operations, information visualization and data analysis, thus allowing for the analysis of various scenarios of production activities without the need to disrupt them in reality [36]. Design is conducted by learning and using knowledge about past events. Analysis of available solutions may result from analogous cases or known external potential improvements [37].

At this stage of consideration, it should be emphasized that the production process analyzed by the authors is a complicated one, which may involve many unnecessary or improper activities, leading to production stoppages. Production analysis using a simulation model allows for reducing costs and optimizing the production time. Popular tools for managing and controlling production are known as Manufacturing Execution Systems (MES) [38], which is used in the operational area. However, these systems have significant drawbacks compared to the possibilities of material flow simulation. MES do not allow, for example, production planning or storage based on designated areas, which is important in the process of manufacturing large-size steel structures [39]. Therefore, the identification of complex relationships between material flows in the long term can only be achieved through simulation modelling. FEM allows for short and reasonable process cycle planning whilst only basing the simulation of the material flow on cooperation with MES to optimize production planning [40].

The authors in their works [41–43] show how the fourth industrial revolution (Industry 4.0) contributes to the development of the process of the technological and organizational transformation of enterprises, which includes the integration of the value chain, the introduction of new business models and the digitization of products and services. The implementation of these solutions is possible by using new digital technologies and data

resources and ensuring communication in the cooperation network of machines, devices and people. The factors driving the transformation towards Industry 4.0 are the increasingly individualized needs of customers and the growing trend of the personalization of products and services. The transformation towards Industry 4.0, according to the authors of works [41–43], includes seven stages. The first concerns technological advancement, which takes into account flexible production systems that facilitate quick adaptation to changes in the number or category of products. The next step is to share information about the manufacturing process by people, machines and products. The third phase is about taking into account the principles of the circular economy to fully use raw materials and reduce emissions. The process of the comprehensive implementation of customer expectations towards products, i.e., end-to-end customer-focused engineering, is the fourth stage towards Industry 4.0. Next, it is crucial to focus on people, including using individual differences to strengthen the organization and build a meaningful work environment. The sixth stage, smart manufacturing, involves the use of integrated systems that respond to changing conditions in real time. In this context, storing and sharing large data sets (big data) is of great importance. The final step is an open factory that understands the needs of all participants in the value chain. All these principles make it possible to achieve the most effective production and quality of the final products [44].

In the lighting pole manufacturing industry, material flow simulation should therefore be used as a production planning tool [45]. Through simulation modelling, the authors in [46] analyzed the production process in terms of resource allocation, data collection necessary for decision-making and full capacity utilization [47]. The authors in [48–51] analyzed the role of the use of simulation modelling in the production of steel structures for large offshore structures. The functionality was presented in a practical example, which allowed us to conclude that simulation modelling contributed to the improvement in the quality of planning in the production of this type of structure.

Modern production systems are very complex and difficult to analyze; therefore, many methods (including mathematical modelling, combinator optimization, Petri net and scenario analysis) are used to solve the above-analyzed problem [52]. Computer simulation methods, especially discrete event simulations (DES), are the most universal and are widely used [53]. There are many DES software tools dedicated to simulating manufacturing processes, such as ARENA, Enterprise Dynamics, FlexSim, Plant Simulation, SIMIO, Witness and others [54]. The main advantage of DES is the ability to conduct many simulation experiments in a short time. Building a simulation model helps in gaining knowledge that can lead to improvements in a real system [55]. The disadvantage of DES is the randomness of some simulation parameters, and it is therefore sometimes difficult to distinguish whether an observation is the result of intersystemic relationships or randomness [56]. Designing complex production systems requires integrating various aspects, including production strategies, system architecture, capacity planning, management techniques, performance assessment, scenario analysis and risk assessment [57]. The beginning of the production system design process is conceptual modelling [58–60]. A summary of the most important information from the literature on the subject is presented in Table 1.

When analyzing issues related to the research area described above, it should be emphasized that despite there being a large number of articles on simulation modelling, no analyses were found regarding the use of the FlexSim tool to improve the manufacturing process of a specific product, such as lighting poles. Therefore, it was decided to test the production line model using FlexSim, and then the critical areas were identified, and an experiment was undertaken to improve this process.

Table 1. Summary of the most important information from the literature regarding the research problem.

Authors	The Most Important Information from the Literature
Toczyńska, 2016 [16]	Avoiding production errors and thus optimizing production requires the use of simulation modelling methods.
Łatuszczyńska, 2015 [21]	Using a simulation model, represent key system features can be represented and data can be collected and used to optimize processes.
Gołda et al., 2015 [23]	A simulation is an imitation of the operation of a real-world process or system over time.
Leminen et al., 2020 [28]	One of the tools that allows for mapping real production and logistics systems in a 3D environment to analyze and optimize their operation based on collected simulation data whilst additionally cooperating with virtual reality technologies is the FlexSim software.
Roháč et al., 2020 [31]	This software can be successfully implemented in the areas of supply, production, warehousing, transport systems and many others.
Qin et al., 2016 [34]	Organizing processes in the conditions of dynamic changes taking place in the enterprise environment is, however, very problematic, because to maintain the continuity of the production process, work must be properly designed and then rationally organized.
Herrmann, 2007 [39]	These innovations refer to the possibility of blurring the boundaries between physical and virtual entities.
Folgado et al., 2024 [42]	Digitalization forces traditional industrial organizations to rethink and develop existing business models.
Akpan and Offodile, 2024 [43]	Modelling of connections between operations and analysis of real scenarios for a given process, taking into account the internal and external environment of the enterprise, is enabled by computer simulation.
Nota et al., 2020 [45]	Solving various technical problems using simulation modeling is effective, and its effectiveness has been confirmed in scientific studies.
Niekurzak et al., 2023 [46]	A properly designed production process can reduce operating costs in a company by up to 50%.
Agarwal and Ojha, 2024 [47]	The effect of modelling processes, objects and phenomena is to obtain resultsin the form of time dependencies of statistical data.
Zamora-Antuñano et al., 2019 [55]	Modelling of manufacturing processes is possible thanks to computer software for conducting experiments and tests. These tools focus primarily on scene modelling, information visualization and data analysis, thus allowing for the analysis of various production operation scenarios without having to interrupt them in reality.
Krenczyk et al., 2019 [53]	Design is conducted by learning and using knowledge about past events. The analysis of available solutions may come from analogous cases or known external potential improvements.
Gola and Wiechetek, 2017 [56]	Popular tools for managing and controlling production are Production Execution Systems (MES), which are used in the operational area.
Veisi et al., 2018 [58]	Thanks to simulation modelling, manufacturing processes can be analyzed in terms of resource allocation, collecting data necessary for decision-making and the full use of the production capacity.

3. Material and Methods

3.1. Model Assumptions

The following research program was planned:

- Selection of the production line to be the subject of research,
- Collecting and defining input data necessary to create a simulation model,
- Building a model using the FlexSim software,
- Validation of the created model,
- Conducting a simulation reflecting the implementation of a sample order,
- Analysis of the obtained simulation results,
- Isolation of process bottlenecks,
- Presentation of potential process improvements, specification of implementation cost solutions,
- Conducting simulations using selected improvements,
- Summary of the results obtained,
- Choosing the best solution.

A simulation tool called FlexSim was used for this research. FlexSim is 3D simulation modelling software that transforms existing data into accurate predictions. Customers who use the software are engineers and decision-makers who see potential in processes that can be improved. The FlexSim testing program is available at the University's laboratory station [61]. It has a perpetual educational license under Academic Classroom LAN. The purchase of a license, depending on the add-ons and library database, costs around EUR 12,000–30,000. Most modern Windows computers meet the minimum FlexSim requirements, i.e., A 64-bit edition of MS Windows 10 2022 version 22H2 under current Microsoft 365 extended support, 8 GB RAM or more, a GPU supporting OpenGL 3.1 or higher, 3 GB free and the latest .NET Framework [62].

The above-mentioned research scheme is intended to allow for the verification of the thesis indicated in the introduction of this work.

3.2. Preparation of the Simulation Model for Research

As part of the research problem, the production process of lighting poles was analyzed in terms of the selected parameter of the unit of time that should be spent on performing individual activities. The first step in building the model was to define the necessary elements and parameters of the analyzed production process. The analyzed production line consisted of the following elements:

- HD-F "Durma" laser cutter,
- Sheet loading, unloading and transport system (part of the cutting machine),
- Pressing the brake of the "Durma" AD-R 30135,
- Integrated system for laser welding of lighting poles (welding cell),
- Cavity cutting station (Fanuc industrial robot),
- Manual (MIG—Metal Inert Gas, MAG—Metal Active Gas) welding machine.

In addition to the above-mentioned machines, the use of human resources was essential. For the efficient operation of the line, the support of at least 5 employees was required to operate the cutting laser, hydraulic press, welding cell and pole foot welding.

A round lighting pole with a length of 4 m was selected as the production object for the relevant analyses. The results of the measurements in the production environment are presented in Table 2.

When a laser is used, 4 pieces of sheet metal with the given dimensions are obtained from one sheet. Therefore, for analysis, it was decided to divide the working time on this machine by 4 to obtain the processing time of one piece of product. After determining the implementation times of the individual process stages, it was decided to introduce information about all the events causing the operation of the individual machines to stop in the model. For this purpose, an analysis was performed using prediction tools.

Table 2. Average processing time at each stage of the process.

Position	Operation Name	Operating Time [s]
Laser	Table exchange	30
	Sheet metal search	60
	Sheet metal cutting	480
Transfer	Waste collection	130
	Loading + unloading	660
Press brake	Configure	30
	Sheet metal bending	90
	Unlock	30
Welding cell	Snorting	130
	Positioning, fixing	100
	Welding	81.4
	Unlock	30
Transfer	Transport to the next station	55
Cutting station exit from the alcove	Preparation for cutting	90
	Cutting a cavity	230
	Lowering after cutting	120
Pole foot welding station	Preparation for welding	30
	Welding of the column foot	60
	Pulling the bar down	30

Source: own study.

Among the many possible disruptions to the production process, an undoubtedly harmful phenomenon is the failure of the machines carrying out their production processes. Each defect causes the device to be excluded for a certain period. In the case of not very complicated damage, this time may range from a few to several minutes (repair by employees of the maintenance department). However, in the event of a serious fault, the machine's downtime may reach one or even several days (need to call the service center or order a damaged component or part). The fault causes an inability to perform a technological operation on a given machine, as well as a loss of smoothness of subsequent operations resulting from the technological route. This situation also affects the remaining tasks in the schedule. To conduct the presented research, it was necessary to collect historical data on failures and defects. The source of this data was service books for technological machines constituting part of the company's machinery. The books contained information on all activities carried out on 12 machines by employees of the maintenance department. However, data on failures were selected for research, and inspections (annual and periodic) were carried out (Table 3).

Because most failure cases were recorded in 2017, data from this period were used for the analysis. The acquisition and appropriate processing of historical data allowed for their implementation in the STATISTICA 13.1 StatSoft Polska system. The collected data were saved in the form of a table using the appropriate variables:

- Failure_1 and Failure_2—variables informing about a failure (value "1") or its absence (value "0"),
- Time_1 and Time_2—variables defining the number of days from the beginning of the observation to the occurrence of the event,
- Time_12—variable defining the time between the first and second failure,
- Complete_1 and Complete_2—variables informing about the observation status (complete—"1", abscissa—"0"),
- Previous—a variable defining whether the failure was related to the previous failure ("1"—yes, "0"—no),
- Time_review < 90—variable indicating whether the fault occurred within 90 days of an inspection ("1"—yes, "0"—no),

– Next—a variable informing about subsequent failures ("1"—further failures occurred, "0"—no further failures).

Table 3. Summary of data on the failure rate of technological machines in 2017.

1	2	3	4	5	6	7	8	9	10	11
Machine	**Failure _1**	**Time_1**	**Failure _2**	**Time_2**	**Time_12**	**Complete_1**	**Complete_2**	**Previous**	**Time_Review < 90**	**Next**
Machine 1	0	365	0	365	0	0	0	0	0	0
Machine 2	1	10	1	74	64	1	1	0	0	0
Machine 3	1	199	0	365	166	1	0	0	1	1
Machine 4	0	365	0	365	0	0	0	0	0	0
Machine 5	1	77	1	246	169	1	1	0	0	0
Machine 6	1	170	1	213	42	1	1	0	0	0
Machine 7	1	247	1	266	19	1	1	0	1	1
Machine 8	1	73	1	231	158	1	1	0	0	0
Machine 9	0	365	0	365	0	0	0	0	0	0
Machine 10	1	220	0	365	145	1	0	0	0	0
Machine 11	1	67	1	260	193	1	1	0	0	0
Machine 12	1	3	1	45	42	1	1	0	0	0

Source: own study.

Presenting the data in this form made it possible to perform research using the tools available in the Survival Analysis module of the STATISTICA package (version 12). The basic tool for duration analysis is mortality tables. This technique belongs to the group of the oldest methods of analyzing duration data. For each interval, the number of complete and censored observations is determined. Based on these data, the number of cases at risk, the proportion of cases failing or the proportion of cases surviving were calculated (Table 4).

Table 4. Survival tables for cases from 2017.

Compartment	Initial Range	Midpoint	Compartment Widths	Number Incoming Observations	Number Truncated Observations	Number of Threats	Number of Deaths	Proportionality Deaths	Proportionality Experiences	Cumulative Probability of Survival	Probability Density	Hazard Rate
L.Initial_1	0.000	13.035	26.071	14	0	14.000	3	0.214	0.785	1.00	0.008	0.009
L.Initial_2	26.071	39.107	26.071	11	0	11.000	0	0.045	0.954	0.758	0.001	0.001
L.Initial_3	52.142	55.178	26.071	11	0	11.000	3	0.274	0.724	0.750	0.007	0.012
L.Initial_4	78.214	91.250	26.071	8	0	8.000	0	0.062	0.935	0.545	0.001	0.002
L.Initial_5	104.285	117.321	26.071	8	0	8.000	0	0.062	0.935	0.511	0.002	0.002
L.Initial_6	130.357	143.392	26.071	8	0	8.000	0	0.062	0.935	0.479	0.004	0.005
L.Initial_7	156.428	169.464	26.071	8	0	8.000	1	0.125	0.875	0.441	0.002	0.012
L.Initial_8	182.500	195.535	26.071	7	0	7.000	2	0.284	0.714	0.391	0.002	0.008
L.Initial_9	208.574	221.607	26.071	5	0	5.000	1	0.200	0.800	0.280	0.001	0.008
L.Initial_10	234.642	247.678	26.071	4	0	4.000	1	0.250	0.750	0.224	0.000	0.010
L.Initial_11	260.714	273.750	26.071	3	0	3.000	0	0.167	0.833	0.165	0.000	0.006
L.Initial_12	286.785	299.821	26.071	3	0	3.000	0	0.167	0.833	0.140	0.000	0.006

Source: own study.

The summary of data in the above form allows for the analysis of failure rates in specific time intervals. Another parameter worth noting is the hazard rate. It is defined as the probability per unit of time that a case that has survived to the beginning of a given interval will fail in that interval. This indicator increases its value in intervals where the number of failures increases, so observing its variability may be a good element in predicting future events. Compiling data in the form of survival tables is the basic element

of creating probability charts. This process involves fitting the theoretical distribution of failures over time to empirical data. In this way, life expectancy function graphs (Figure 2) and probability density distributions (Figure 3) can be obtained. Because the Weibull distribution is a distribution quite often used in analyses related to failure rates, such a distribution was chosen when generating the charts. The estimation is based on three different estimation procedures—least squares and two weighted least squares methods.

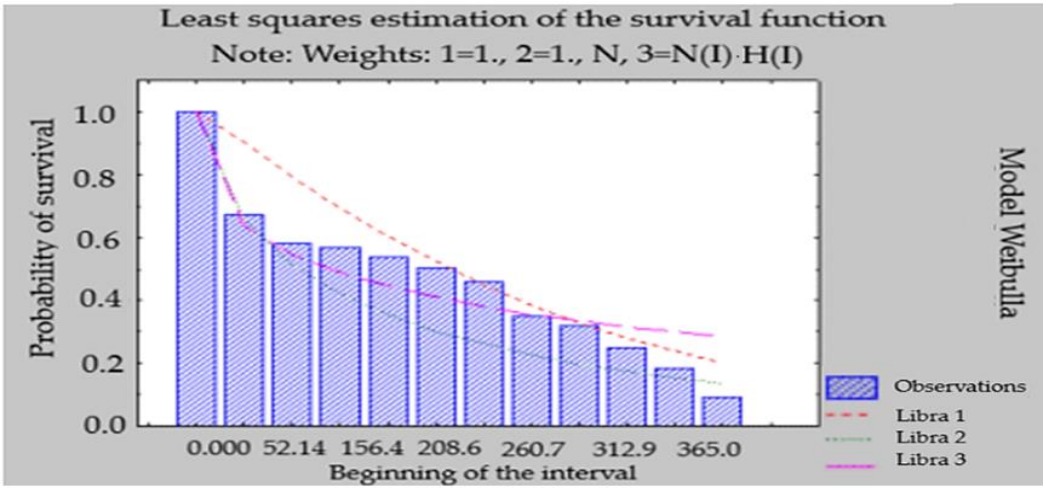


Figure 2. Matching the distribution of failures over time to the empirical distribution for 2017. Source: own study.

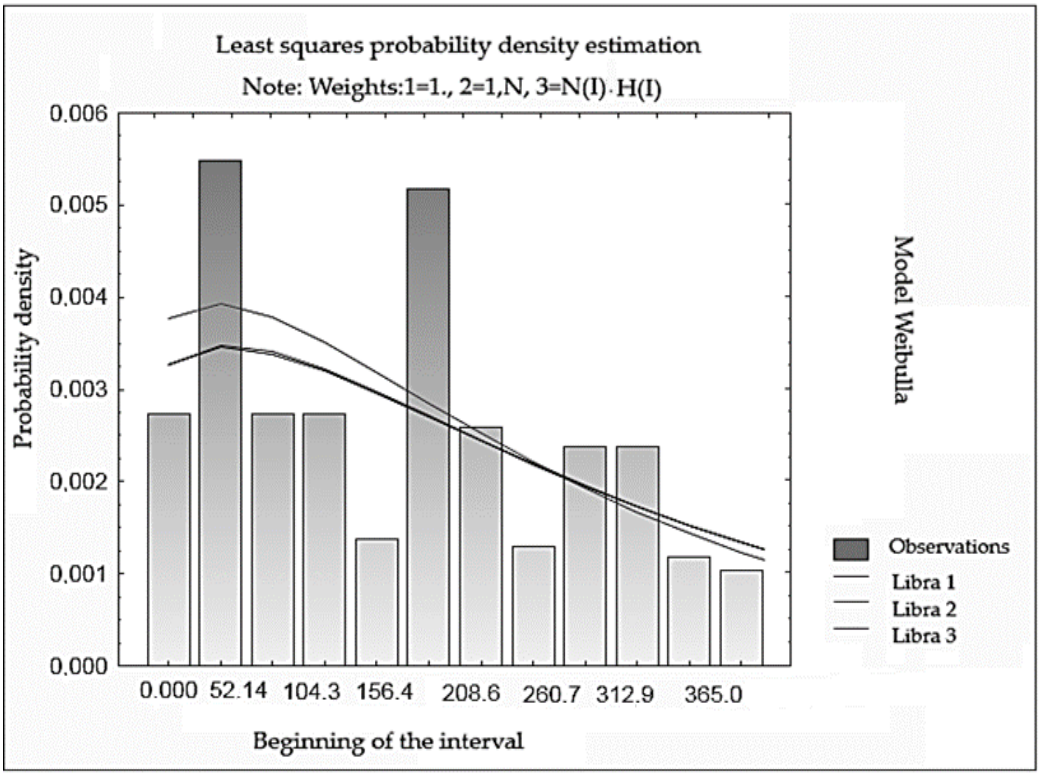


Figure 3. Probability density charts for 2017. Source: own study.

Analyzing the obtained estimation results, it should be stated that in the case of the observations from 2017, the approximate distribution charts overlapped to a significant extent with the empirical distributions. In the case of the estimated distributions for the observations from 2017, it should be stated that the procedures used gave very similar

results. Nevertheless, the distribution matching tool allows us to illustrate, to some extent, the nature of the analyzed failures.

In the case of estimating the probability density curves, we also observed that for the data from 2017, various approximation methods took a similar shape. It should be noted that the obtained graph has a typical shape of the probability density function of the Weibull distribution. Based on the results obtained, it was possible to determine the shape parameter, which is one of the parameters of the fitted distribution.

In the subsequent part of this study, the survival function was estimated directly from continuous survival times using the Kaplan–Meier method. Each interval then contained exactly one event. The survival function is defined as the product of successive probabilities from the intervals [62]:

$$S_n(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{r_j}\right) \quad (1)$$

where:

$S_n(t)$—estimated survival function,

$\prod_{t_j \leq t}$—product symbol,

d_j—number of events in the period t_j,

r_j—number exposed per period t_j.

The results of the estimation of the survival function for the examined empirical data are presented in Figure 4.

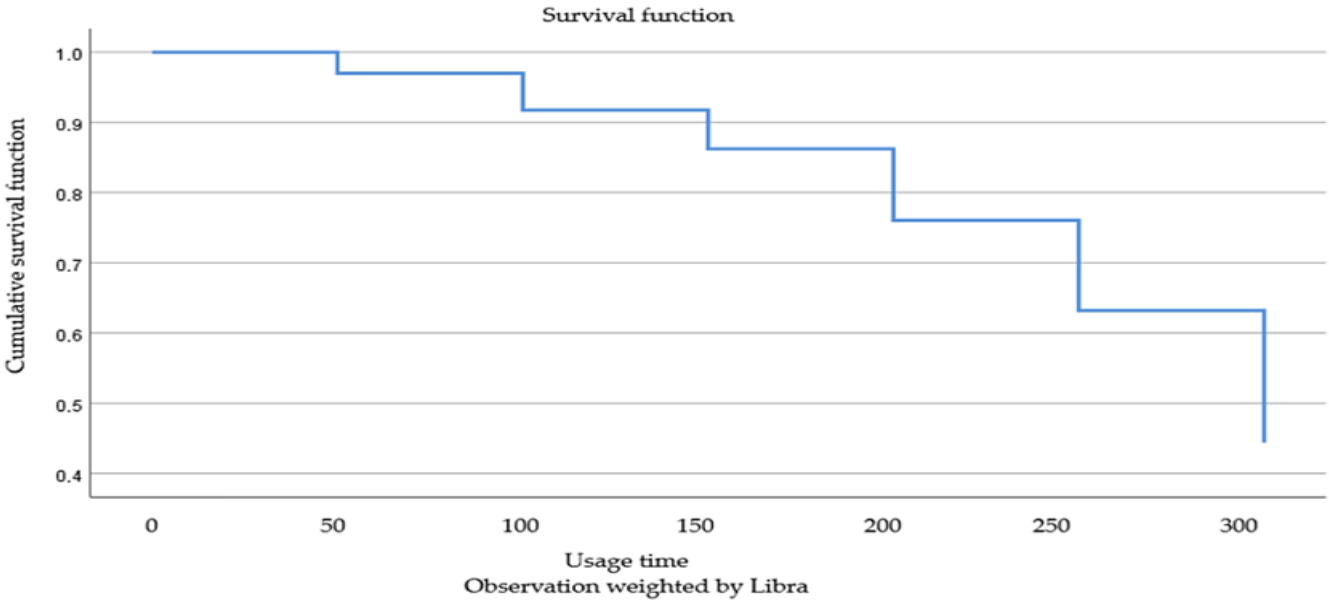


Figure 4. Survival function charts for 2017. Source: own study.

By analyzing the above data, we can determine the value of the survival function in a specific time interval. For example, around 140–170 days of work, the survival rate was 90%. Analyzing the graphs of the tested samples, it should be concluded that the occurrence of the second and subsequent failures significantly affected the probability of survival. Its value dropped below 0.5 already within 100 days, while for machines experiencing only 1 failure, it remained at the level of 0.7 even up to the 220th day. This is valuable information that allows for the identification of machines that potentially pose a threat to the smooth operation of the production process. Selected statistical tests were used to examine the statistical significance of the analyzed data. The results of the significance tests are presented in Table 5.

Table 5. Results of significance tests—2017.

Comparing Trials: Occurrence or absence of failure #2	
Test name	Level of significance p
F Coxa	0.001
T. Coxa–Mantela	0.004
log-rank	0.004
Comparing trials: Occurrence of subsequent failures.	
Test name	Level of significance p
F Coxa	0.04
T. Coxa–Mantela	0.07
log-rank	0.1

Source: own study.

Analyzing the above results of the significance tests, it should be stated that in the case of comparing the tests concerning failure no. 2, all the tests showed that the results obtained were statistically significant ($p < 0.05$). Unfortunately, when comparing the samples in terms of subsequent failures, the significance level in the case of the two tests was above $p = 0.05$, which means that it cannot be concluded that the presented differences were statistically significant. The results of comparing the tests allowed us to conclude that both the occurrence of a second failure and the number of days that pass from the inspection to the occurrence of the failure have a significant impact on the course of the survival function. Machines that have only had one failure are more likely to survive than those that have failed again. Similarly, machines whose fault occurrence time is longer than 90 days from the last inspection are characterized by higher survival function values than those where the fault occurred within 90 days of inspection.

Moreover, before building the simulation model in FlexSim software, it was decided to make the following assumptions:

- The line is open on weekdays from 6:00 a.m. to 4:00 p.m. and closed at other times and on Saturdays and Sundays,
- Sheet metal and finished column feet are available in sufficient quantities to process the order,
- The finished poles are loaded for transport one at a time,
- Each machine has an assigned operator who is responsible for operating it,
- Each operator is assigned to one machine,
- Orders are placed for one type of product, so there is no need to change the equipment,
- Orders are fixed and accepted every Monday (150 pcs.) and Wednesday (200 pcs.).

Based on the above data and the assumptions made, a simulation model was built in the FlexSim program. The following data and objects were used for this purpose:

- Source, i.e., the order generator; the arrival times of individual orders were introduced by importing an Excel spreadsheet. This element was also selected to reflect the composition of the column feet, which are delivered from the workshop once a week in the amount of 500 pieces. Figure 5 shows the program window with the characteristics of the order generator.
- A queue, reflecting the composition of the sheet metal. As a result of accepting the order, sheets appear in this place, which is an input element for processing at subsequent stations. It is assumed that the stack can hold a maximum of 500 sheets.
- A processor equivalent to a laser cutting machine, a press brake, a welding cell and a cavity cutting station. For each station, the processing time, the time between failures and the repair time of the device are specified. Figure 6 shows windows from the FlexSim program with information about one of the machines.
- A connector, corresponding to the welding station of the column foot. As with the previous machines, the processing time of the product was determined, and the "join" operation was fixed so that the final product was a pole with a foot.

- A sink, which is a place to store finished products.
- Operators, assigned one to each machine,
- A transporter, transporting finished products to the field for shipment.

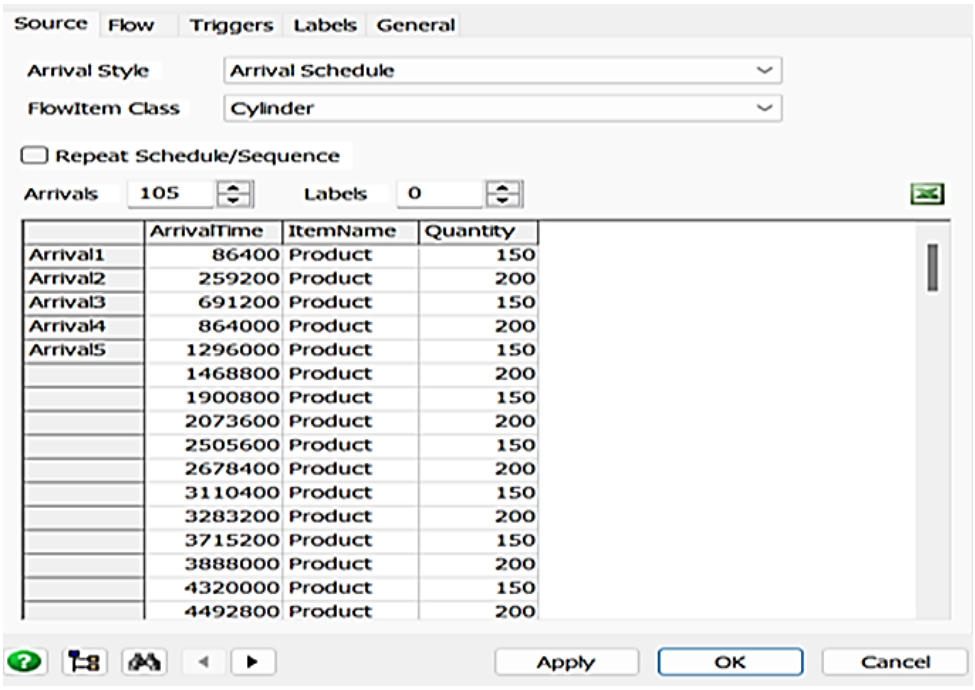
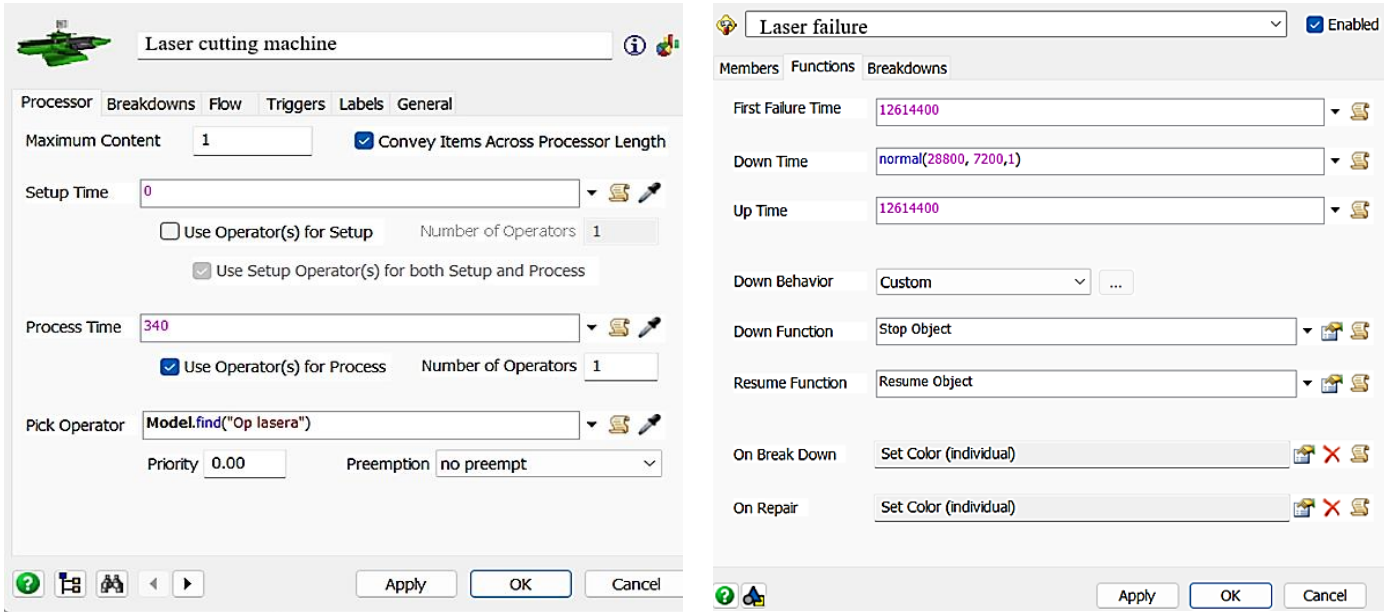


Figure 5. Characteristics of the order generator. Source: own study.



(**a**) (**b**)

Figure 6. Machine characteristics: (**a**) for laser cutting, (**b**) for laser failure. Source: own study.

Figure 7 shows the diagram of the built model.

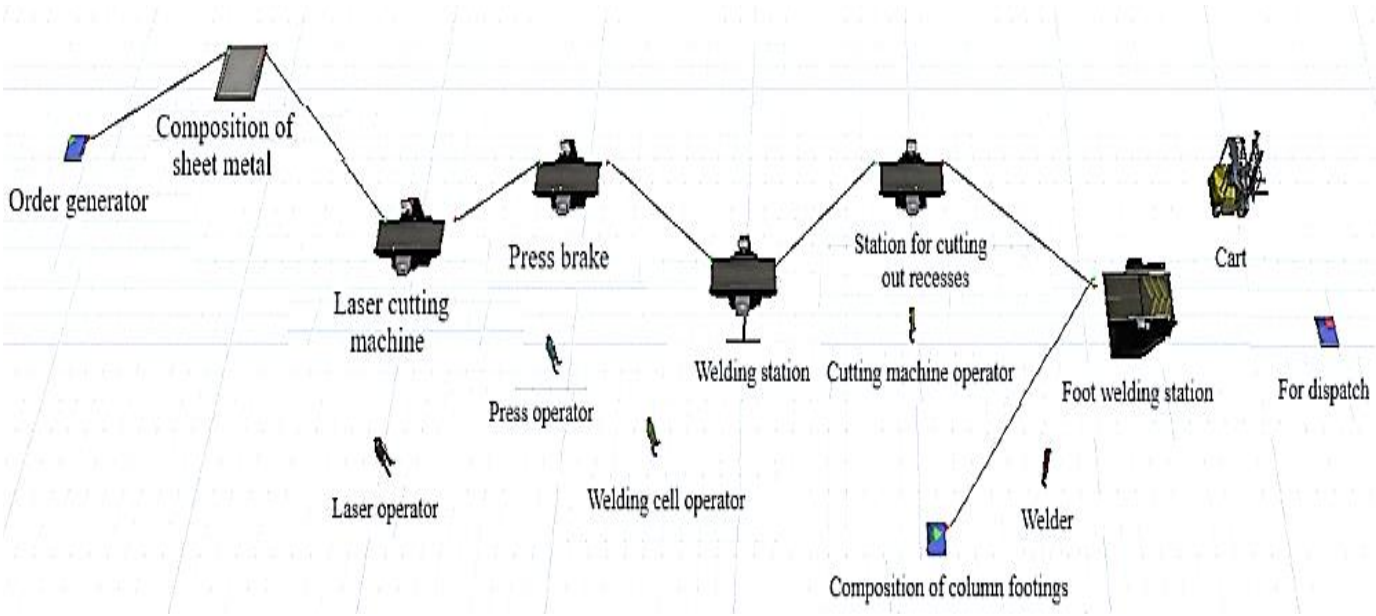


Figure 7. Simulation model of the production line in FlexSim. Source: own study.

The finished products are transported for shipment by a forklift. The logic of its operation is based on the time of completion of the last operation of the production process. Figure 8 shows the process of transporting finished products.

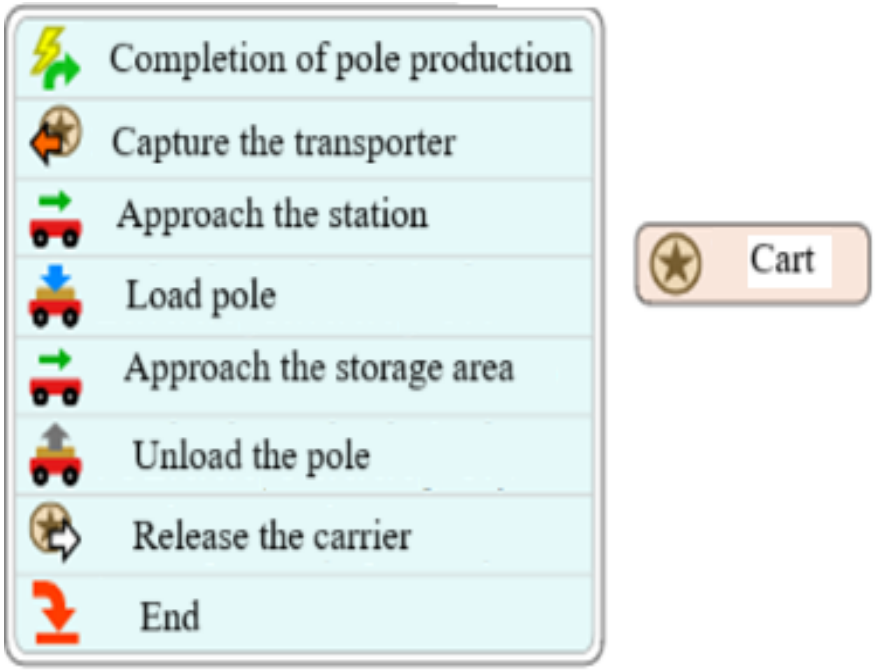


Figure 8. The logic of the transporter's operation. Source: own study.

To maintain the logic and consistency of the analyses, the line operating time was also determined. For this purpose, information about the hours when the system will perform scheduled operations on particular days of the week was entered into the model. Figure 9 shows the program window reflecting this aspect of the model.

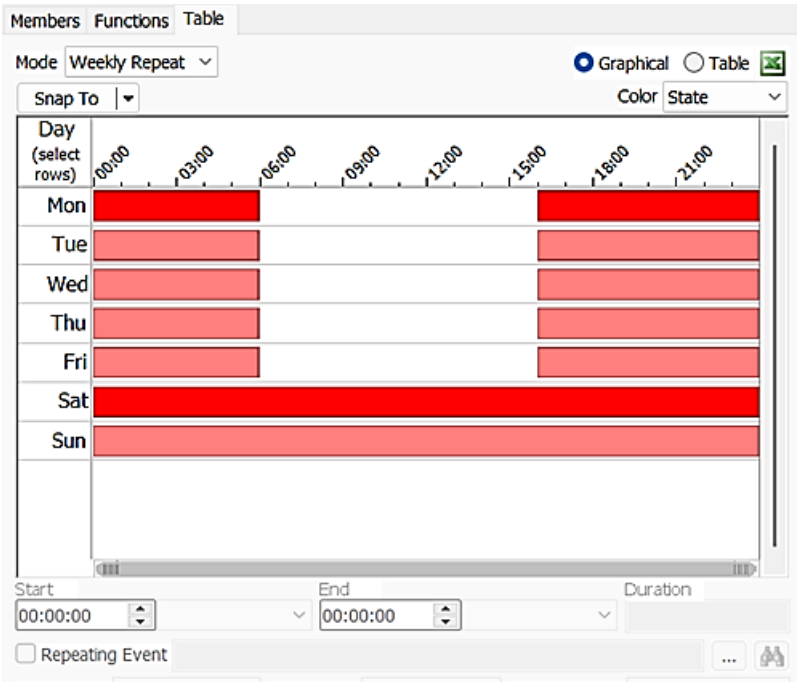


Figure 9. A calendar showing the operating time of the system. Source: own study.

Once the simulation model was built in FlexSim, the validation process began. For this purpose, a simulation of the implementation of a historical order carried out by a production plant was carried out. Then, the actual lead time of the production order was compared with the result obtained with the specialized FlexSim software. After the model was validated, the production line was analyzed in terms of the assessment of the achieved efficiency of individual production sections. For this purpose, a simulation covering a 365-day system operation cycle was carried out based on the previously adopted assumptions.

4. Results and Discussion

4.1. Simulation Results

After entering the data into the model and taking into account the previously adopted assumptions, the simulation of the production line operation was carried out. Figure 10 presents the data generated with the FlexSim program, which illustrate the individual characteristics of the individual production stages and thus the percentage share of the use of the selected machines in the analyzed production process.

By analyzing the presented data, a so-called bottleneck located on the cavity cutting section was identified. For about 86% of the simulation time, the station performed operations on the product. For the rest of the time, the machine waited for the failure to be repaired at the previous stations to resume production. In addition, it was noted that the column foot welding station was idle most of the time (only about 24% of the simulation time was spent by the machine working on product finishing). The second section stage that attention was paid to was the welding cell. Its operation reduced the efficiency of the press brake, which spent most of its time in the simulation waiting for the possibility of transferring the machined product to the next stage of production.

Another area of analysis was the number of products that could be produced during one shift with the current system functioning. Figure 11 shows an analysis of the daily output of poles over a year.

The data analysis indicated that the production line was capable of producing about 80 pieces of product per day. Due to failures and different repair times of individual machine sections, this value was not fixed for all days. After examining the "for ship-

ment" field, it was determined that in the simulated time (365 days), the line produced 18,200 lighting poles.

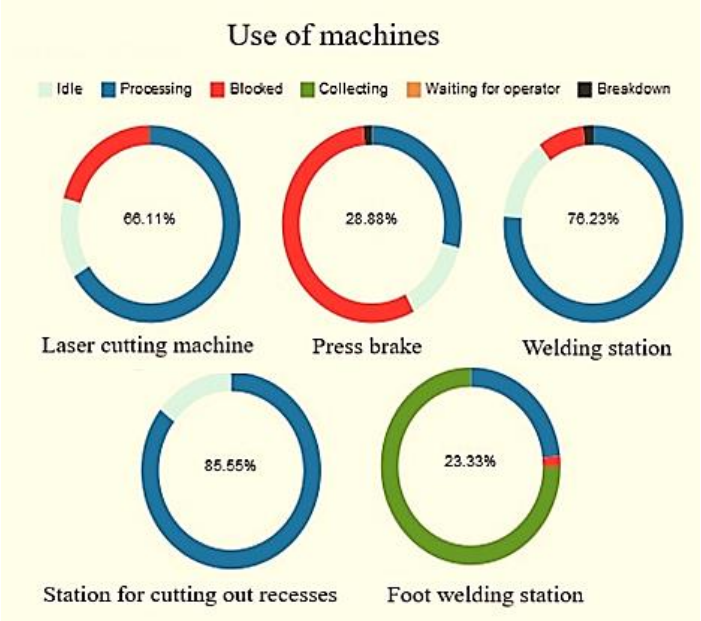


Figure 10. Utilization of individual machines on the production line in the base model percentage share. Source: own study.

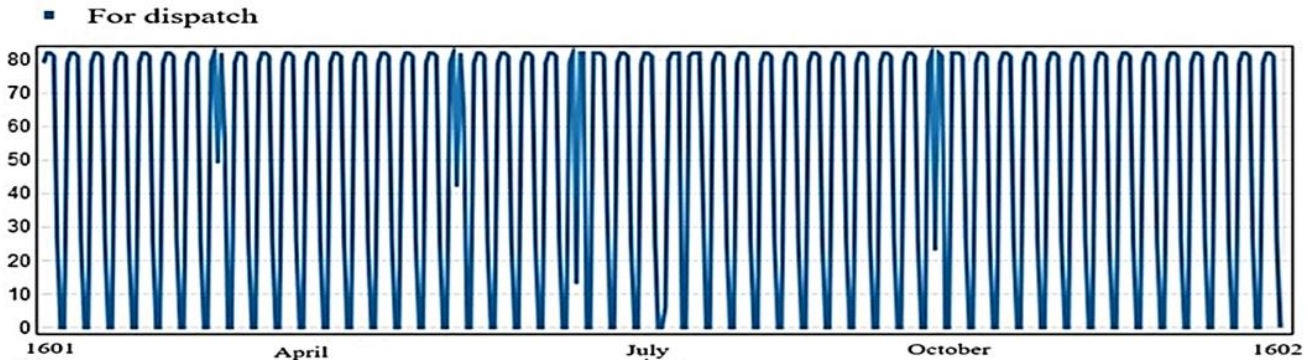


Figure 11. Daily production in pieces of the base model. Source: own study.

In addition, the influence of the sheet metal composition state section was studied. Figure 12 shows a chart reflecting the execution of orders throughout the year.

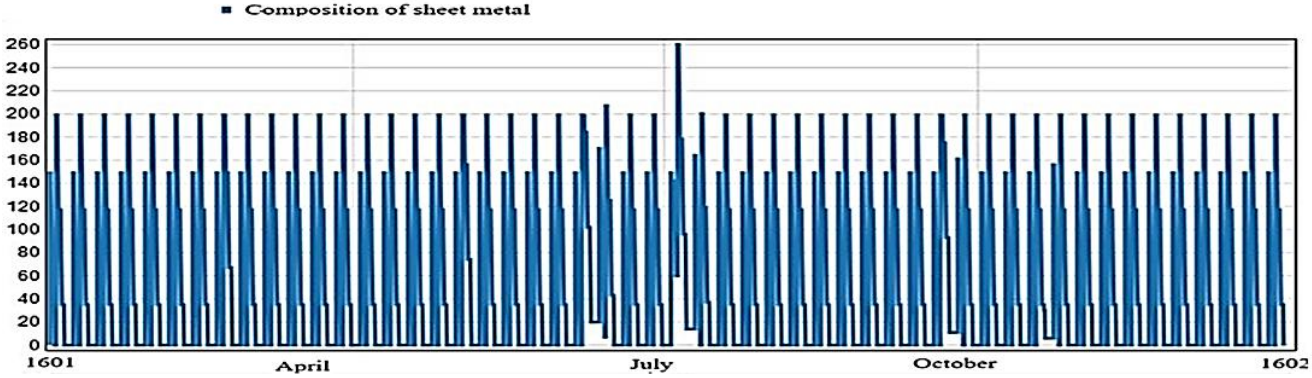


Figure 12. The state of the sheet metal in the base model. Source: own study.

The analysis of the presented data confirmed that the sheet metal storage section could meet the demand for the manufactured product. All of the above values were considered as the initial state to which the results of the next simulation will be compared.

4.2. The Concept of Changes in the Production Process—Analysis and Evaluation

The cavity cutting station of the analyzed semi-finished product was identified as a bottleneck of the analyzed line. The residence time of the semi-finished product in this section slowed down the entire system and reduced the production capacity. Due to the large area of the production hall, it was possible to introduce a second nest-cutting station into the system. The concept of the changes assumed the purchase of a second Fanuc robot and the involvement of an additional employee.

To assess whether the presented solution would improve the process and to what extent the individual parameters of the system would change, it was decided to use the FlexSim software and modify the original model of the system. The following changes were made to the base model:

- A station for cutting out recess No. 2 with the same processing time as station No. 1 was added,
- A second operator of the cutting machine was introduced,
- The logic of selecting the position for the excision of the cavity due to availability was established,
- A queue in front of the foot welding station was added.

After making these changes, the model shown in Figure 13 was obtained.

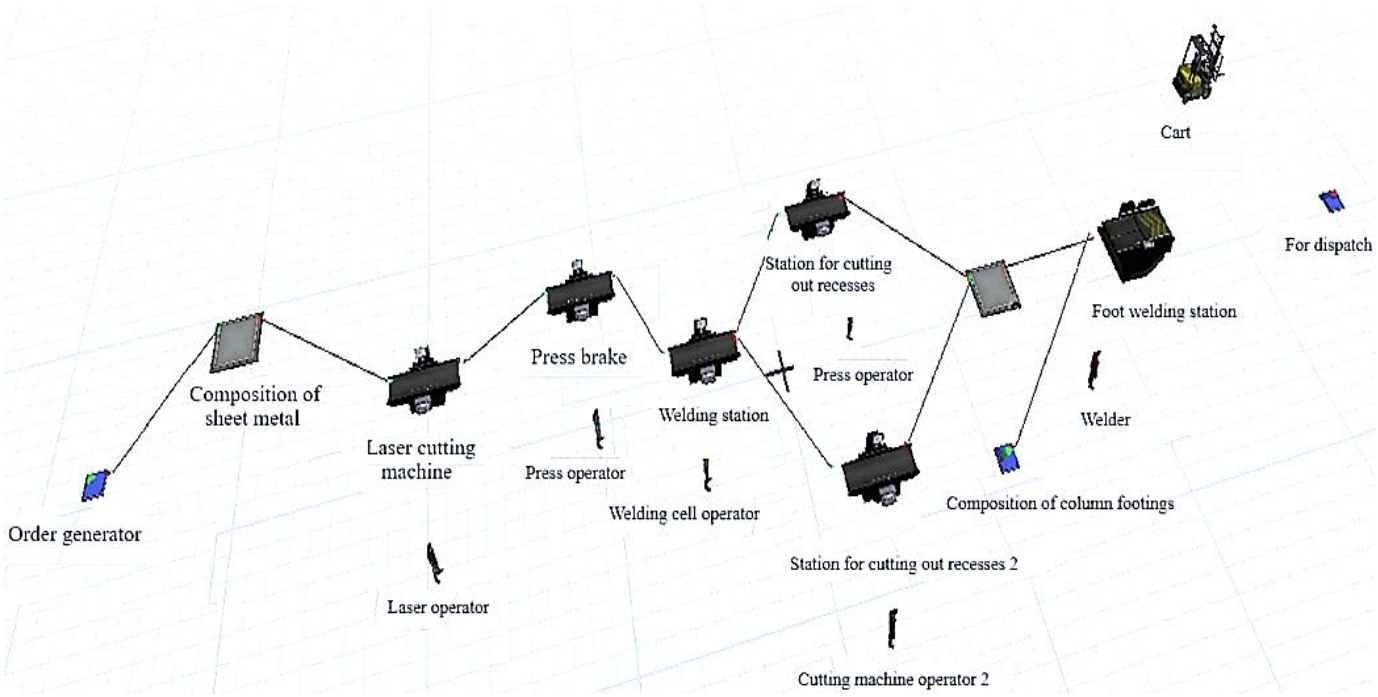


Figure 13. Design of a new model production line with an additional station. Source: own study.

The next step was to simulate the modified system. As with the original model, the duration of the experiment was set at 365 days. Figure 14 shows the daily production of products, Figure 15 shows the graphs showing the use of machines in the analyzed production line and Figure 16 shows the sheet metal composition in the new conceptual model.

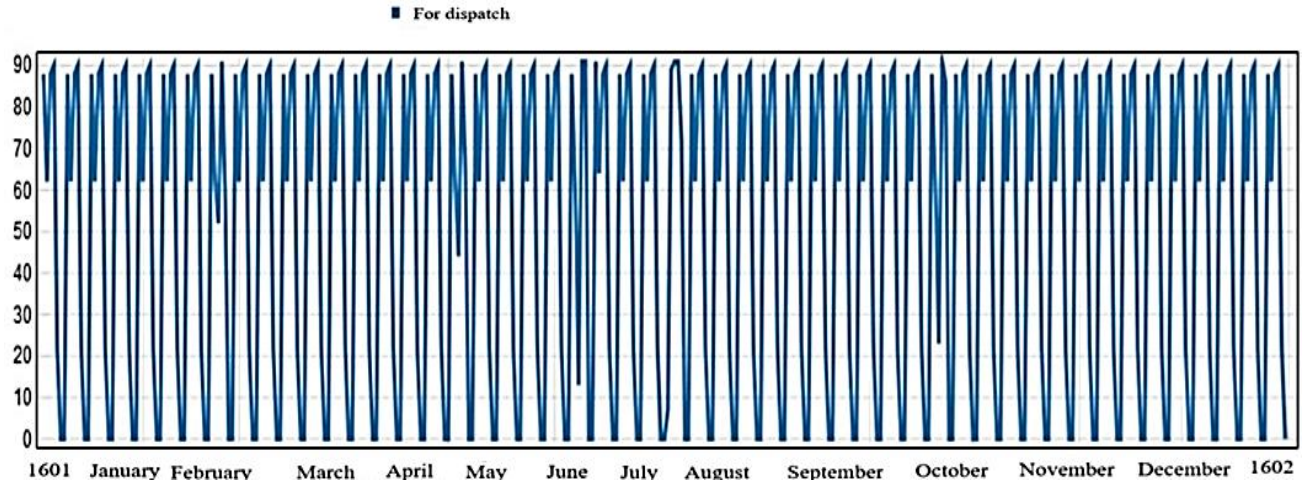


Figure 14. Daily production of products in the concept model. Source: own study.

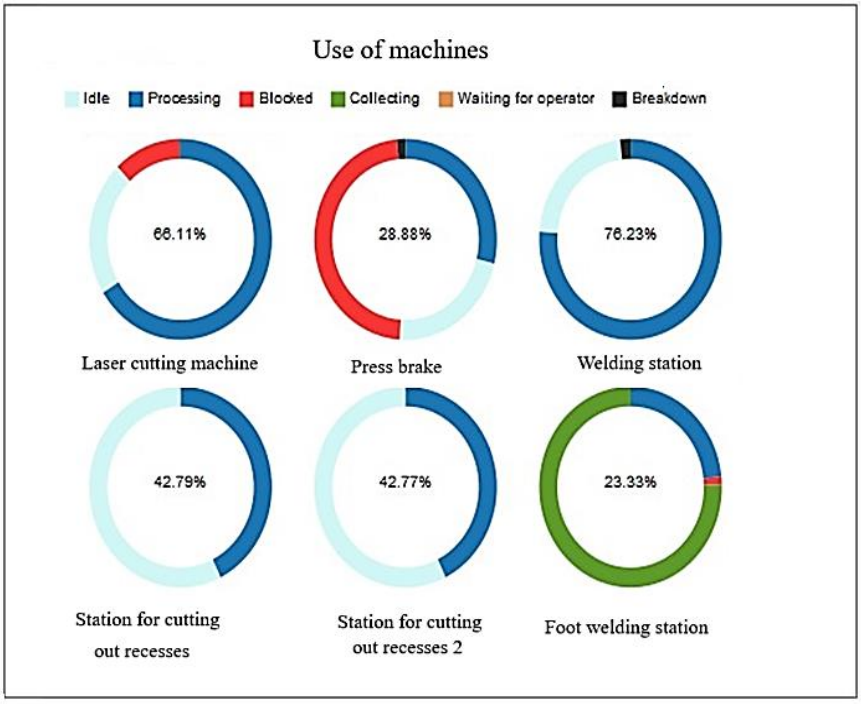


Figure 15. The use of individual sections—machines in the conceptual model—percentage share. Source: own study.

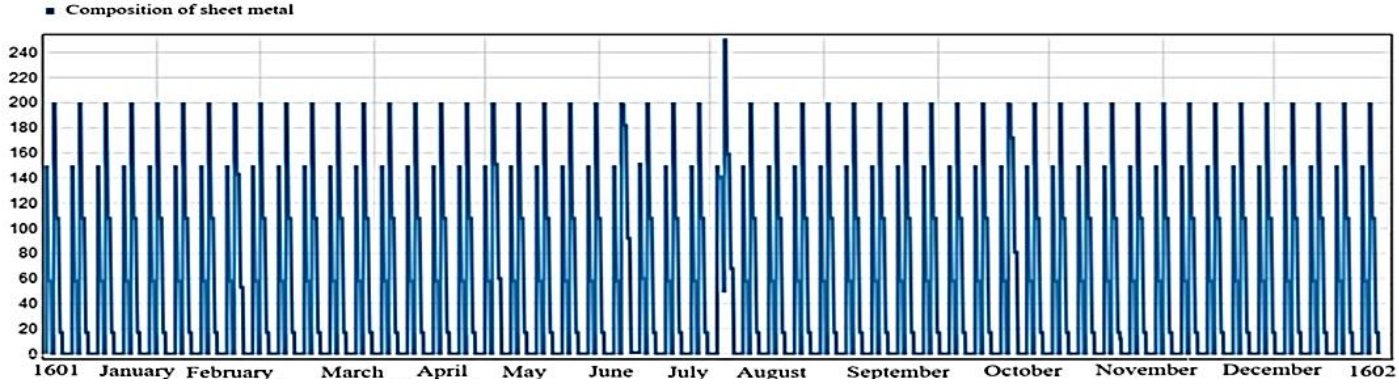


Figure 16. The state of the sheet metal composition in the conceptual model. Source: own study.

After analyzing the data from the conceptual model, it was found that:

- The maximum daily production of the final product increased to about 90 pieces,
- An increase in the phenomenon of inactivity was observed in the selected sections, which resulted from the faster execution of accepted orders,
- The introduction of a second cavity-cutting station eliminated the bottleneck phenomenon at this stage,
- In the remaining sections, it was found that the machines were used to the same extent, but the blocking time of these devices was shortened,
- The stock level of sheet metal was more likely to reach a coefficient of 0.

Based on observations of both the original production line and the conceptual one, it was found that the introduction of an additional welding station resulted in an increase in production efficiency compared to the original model of the production line. A noticeable result of these changes was an increase in production from about 80–90 units. In addition, it was noticed that the condition of the machines preceding the stand had changed. In the case of welding, the percentage of time during which equipment was blocked and waiting for the possibility of transferring the semi-finished product to the next stages of the production process was significantly reduced.

5. Conclusions

Observations of production environments indicate that many companies are currently implementing the idea of Industry 4.0 [63]. This process builds the belief that the implementation of related solutions will allow for significant efficiency improvements, cost reductions, optimization and the harmonization of production processes on the way to sustainable development. One of the key solutions to achieve this success seems to be the use of simulation modelling and the use of specialized tools such as FlexSim to diagnose the entire production system or its selected components [64].

There is no doubt that simulation modelling enables the optimization of product flow and, consequently, the full use of available resources. This effect is possible by conducting simulation experiments on the production line without the need to stop production [54]. The case of using a simulation on the example of the selected good analyzed in this article certainly confirmed how complicated modern production processes can be. In a way, it indicates how many factors affect the difficulties in proper production planning and how important it is to reduce unnecessary activities to optimize production.

Based on the presented research, the following conclusions were formulated:

1. Simulation models can be used to identify adverse phenomena occurring in the area of production processes. They reflect the functioning of systems in a production environment, which are subjected to all kinds of analyses. With their help, it is possible to test the validity of the proposed conceptual and model solutions without the need to make actual changes to the production system and without incurring significant financial outlays.
2. Tools such as FlexSim contain extensive libraries of 3D objects. This makes it possible to faithfully reproduce the analyzed process—the production line. This allows for the construction of detailed models of production systems, the conducting of experiments and the analysis of the achieved results. It is important to emphasize that advanced algorithms together with statistical tools enable the mapping of the natural variability of the process. Thanks to these tools, it is also possible to analyze a large number of alternative scenarios to find a solution, thus confirming the validity of strategic decisions, finding the optimal solution from hundreds of possibilities or even creating the best production plan. In addition, these programs have the function of integration with other tools (e.g., Microsoft 365 Excel 2016 or © 2024 Autodesk Inc. CAD—Computer-Aided Design—software).
3. The use of the FlexSim software made it possible to propose solutions that could bring positive effects to the existing production line without incurring additional costs. Furthermore, the tool used significantly reduced the time spent on research and

development. The simulation of the system allowed for the 365-day life cycle of the modified line to be analyzed in one day, which is an extremely short time compared to the period during which the data were obtained.

4. According to the authors, the FlexSim modelling and simulation tool can also be used to create new production lines. At the same time, it provides a solid foundation supporting the digital transformation of Industry 4.0. This software allows for the analysis of project costs and potential profits from production. This makes it possible to decide whether to start or reject a project of building a production line without incurring additional costs.

Thus, the authors postulate that in the era of growing interest in the idea of Industry 4.0, it is worth considering investing in simulation software, especially in the case of corporations that can use it in many areas and branches of a given group. Furthermore, simulation software can not only improve individual internal processes but also harmonize the coordination of the entire supply chain. The benefits of simulation will then also be noticeable in other links in the production chain. The authors agree with the statement promoted by other researchers that the possibilities offered by simulations are certainly worth considering and will probably become a standard as part of the Industry 4.0 idea [65]. The business sphere should recognize that simulations offer great opportunities to create new processes and systems and even to design entire supply chains and improve their operation, regardless of the complexity of the process [66].

Although the presented considerations do not exhaust the entire topic, they were intended to draw attention to the modelling and optimization of production processes in a real production environment as part of the idea of Industry 4.0 [67]. Undoubtedly, increasing efficiency and reducing production costs are serious challenges that cannot be successfully achieved without breakthrough changes in production paradigms and the underlying technologies and applications. According to the authors, appropriate tools, such as those discussed in this article, have the potential to play a significant role in this process of evolution from Industry 4.0 to Industry 5.0 [68,69] in key areas such as simulation, system integration, autonomous systems, cloud computing, augmented reality, big data and data analysis [70].

Author Contributions: Conceptualization, M.N. and W.L., methodology, M.N.; software, M.N. and W.L.; validation, M.N. and W.L.; formal analysis, M.N. and W.L.; investigations, M.N. and W.L.; resources, M.N.; data curation, M.N. and W.L.; writing—preparation of the original draft, M.N., W.L. and J.W.; writing—reviewing and editing, M.N., W.L. and J.W.; visualization, M.N.; supervision, W.L.; project administration, M.N. and W.L.; obtaining financing, W.L. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: The research was financed as part of a project carried out within the Faculty of Economics of the West Pomeranian University of Technology in Szczecin under the name Green Lab. Research and innovation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

1. Nascimento, D.L.M.; Alencastro, V.; Quelhas, O.L.G.; Caiado, R.G.G.; Garza-Reyes, J.A.; Rocha-Lona, L.; Tortorella, G. Exploring Industry 4.0 technologies to enable circular economy practices in a manufacturing context: A business model proposal. *J. Manuf. Technol. Manag.* **2019**, *30*, 607–627. [CrossRef]
2. Laskurain-Iturbe, I.; Arana-Landín, G.; Landeta-Manzano, B.; Uriarte-Gallastegi, N. Exploring the influence of Industry 4.0 technologies on the circular economy. *J. Clean. Prod.* **2021**, *321*, 128944. [CrossRef]

3. Jamwal, A.; Agrawal, R.; Sharma, M.; Giallanza, A. Industry 4.0 technologies for manufacturing sustainability: A systematic review and future research directions. *Appl. Sci.* **2021**, *11*, 5725. [CrossRef]
4. Tavera Romero, C.A.; Castro, D.F.; Ortiz, J.H.; Khalaf, O.I.; Vargas, M.A. The synergy between circular economy and industry 4.0: A literature review. *Sustainability* **2021**, *13*, 4331. [CrossRef]
5. Breznik, M.; Buchmeister, B.; Vujica Herzog, N. Assembly line optimization using MTM time standard and simulation modelling—A case study. *Appl. Sci.* **2023**, *13*, 6265. [CrossRef]
6. Heshmat, M.; El-Sharief, M.A.; El-Sebaie, M.G. *Simulation Modeling of Production Lines: A Case Study of Cement Production Lines*; Assiut University: Assiut, Egypt, 2013.
7. Zahraee, S.M.; Golroudbary, S.R.; Hashemi, A.; Afshar, J.; Haghighi, M. Simulation of manufacturing production line based on Arena. *Adv. Mater. Res.* **2014**, *933*, 744–748. [CrossRef]
8. Pehrsson, L.; Frantzen, M.; Aslam, T. Aggregated Linear Modeling for Simulation and Optimization of Production Systems. In Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA, 6–9 December 2015.
9. Gajda, J.B. *Forecasting and Simulations in Economics and Management*; C.H. Beck Publishing House: Warsaw, Poland, 2017.
10. Lee, K.; Kang, K.C.; Lee, J. Concepts and guidelines of feature modelling for product line software engineering. In Proceedings of the International Conference on Software Reuse, Austin, TX, USA, 15–19 April 2002; Springer: Berlin/Heidelberg, Germany, 2002; pp. 62–77.
11. Cezarino, L.O.; Liboni, L.B.; Oliveira Stefanelli, N.; Oliveira, B.G.; Stocco, L.C. Diving into emerging economies bottleneck: Industry 4.0 and implications for circular economy. *Manag. Decis.* **2021**, *59*, 1841–1862. [CrossRef]
12. Maciąg, A.; Pietroń, R.; Kukla, S. *Prognozowanie i Symulacja w Przedsiębiorstwie*; PWE: Warszawa, Poland, 2013.
13. Ruwaida Aliyu, A.A.M. Research Advances in the Application of FlexSim: A Perspective on Machine Reliability, Availability, and Maintainability Optimization. *J. Hunan Univ. Nat. Sci.* **2021**, *48*, 518–563.
14. Dubai, K. *Practical Application of the FlexSim Tool in Simulation Modelling of Production Systems*; Scientific Papers of the Upper Silesian Academy, No. 6/2023; Akademii Górnośląskiej: Katowice, Poland, 2023; pp. 23–31. [CrossRef]
15. Stawowy, A. *Simulation Method*; Kraków, Poland, 2020.
16. Toczyńska, J. *Modelling and Simulation of the Management System of the Education Process at the University*; Scientific Papers of the Silesian University of Technology, 1949/2016; Silesian University of Technology: Gliwice, Poland, 2016.
17. Gotowała, K.; Patyk, R. *Application of Modern Modelling and Simulation Methods in the Design of Machines and Equipment*; Autobusy 8/2016; 2016.
18. Fritzkowski, P. *Computer Modeling and Simulations*; Poznan University of Technology: Poznań, Poland, 2016.
19. Łatuszczyńska, M. *Computer Simulation Methods—An Attempt at Logical Classification*; Studies & Proceedings of the Polish Association of Knowledge Management, No. 41; Uniwersytet Szczeciński: Szczecin, Poland, 2011; pp. 163–165.
20. Maciąg, A.; Pietroń, R.; Kukla, S. *Forecasting and Simulation in the Enterprise*; Polskie Wydawnictwo Ekonomiczne: Warsaw, Poland, 2013.
21. Łatuszczyńska, M. Modeling and simulation in production management. *Organ. Rev.* **2015**.
22. Pawlewski, P. *Practical Application of the Multimodal Approach in Simulation Modelling of Production and Assembly Systems*; Springer: Cham, Switzerland, 2022.
23. Gołda, G.; Gwiazda, A.; Kampa, A.; Monica, Z. *The Use of Computer-Aided Systems in Planning Logistics Activities*; University of Dąbrowa Górnicza: Dąbrowa Górnicza, Poland, 2015.
24. Bernat, P. Computer-aided technical preparation of production. In *Applications of Computer Science in Production Engineering*; Lublin, Poland, 2009.
25. Hellmuth, R.; Frohnmayer, J.; Sulzmann, F. Design and application of a digital factory model for factory restructuring. *CIRP Procedia* **2020**, *91*, 158–163. [CrossRef]
26. Stolarska-Szeląg, E. Gemba Walk in Manufacturing Companies—Implementation Process and Benefits. *ZN WSH Manag.* **2022**, *23*, 63–76. [CrossRef]
27. Huikkola, T.; Kohtamäki, M.; Ylimäki, J. Becoming an Intelligent Solution Provider: Reconfiguring the product manufacturer's strategic capabilities and processes to facilitate innovation in the business model. *Technology* **2022**, *118*, 102498.
28. Leminen, S.; Rajahonka, M.; Wendelin, R.; Westerlund, M. Business models of the Industrial Internet of Things in the context of machine-to-machine. *Ind. Mark. Manag.* **2020**, *84*, 298–311. [CrossRef]
29. Kohtamäki, M.; Rabetino, R.; Huikkola, T. Learning in Strategic Alliances: Reviewing literature streams and setting the agenda for future research. *Ind. Mark. Manag.* **2023**, *110*, 68–84. [CrossRef]
30. Kończak, A.; Passławski, J. Decision support in planning the production of precast panels based on simulation and learning from mould examples. *Procedia Eng.* **2015**, *122*, 81–87. [CrossRef]
31. Roháč, A.; Weißenfels, S.; Strassburger, S. Concept of comparing intralogistics projects with real factory layouts using augmented reality, SLAM and tag-based tracking. *CIRP Procedia* **2020**, *93*, 341–346. [CrossRef]
32. Okubo, Y.; Mitsuyuki, T. Ship Production Planning Using Shipbuilding System Modeling and Discrete Time Process Simulation. *J. Mar. Sci. Eng.* **2022**, *10*, 176. [CrossRef]
33. Baroroh, D.K.; Chu, C.H.; Wang, L. A systematic literature review on augmented reality in smart manufacturing: Human-computational intelligence collaboration. *J. Prod. Syst.* **2021**, *61*, 696–711. [CrossRef]

34. Qin, J.; Liu, Y.; Grosvenor, R. A categorical production framework for Industry 4.0 and beyond. *CIRP Procedia* **2016**, *52*, 173–178. [CrossRef]
35. Forcael, E.; Gonzalez, M.; Soto, J.; Ramis, F.; Rodriguez, C. Simplified Building Construction Process Planning Using Discrete Event Simulations. In Proceedings of the LACCEI's 16th International Conference on Engineering, Education and Technology: "Innovation in Education and Inclusion", Lima, Peru, 19–21 July 2018.
36. Li, G.; An, S.-M.; Liao, J.-J. Optimization and simulation of the production line for sprawling cars. *Ind. Eng.* **2009**, *11*, 71–76.
37. Dziadosz, A.; Kończak, A. An overview of selected methods of supporting the decision-making process in the construction industry. *Civ. Eng.* **2016**, *62*, 111–126.
38. Ljubenkov, B.; Dukic, G.; Kuzmanic, M. Simulation methods in the design of shipyard processes. *Stroj. Vestn. J. Mech. Eng.* **2008**, *54*, 131–139.
39. Herrmann, F. *Simulation der Regelung von Unternehmensprozessen—Ein Enterprise Ressource Planning System am Beispiel von SAP R/3*; Nordhäuser Hochschultexte—Schriftenreihe Betriebswirtschaft; FH Nordhausen: Nordhausen, Germany, 2007.
40. Blok, C.; Kreimeier, D.; Kuhlenkötter, B. A holistic approach to teaching IT skills in a production environment. *Prod. Procedia* **2017**, *23*, 57–62. [CrossRef]
41. Ghasemi, A.; Farajzadeh, F.; Heavey, C.; Fowler, J.; Papadopoulos, C.T. Simulation optimization applied to production scheduling in the era of industry 4.0: A review and future roadmap. *J. Ind. Inf. Integr.* **2024**, *39*, 100599. [CrossRef]
42. Folgado, F.J.; Calderón, D.; González, I.; Calderón, A.J. Review of Industry 4.0 from the Perspective of Automation and Supervision Systems: Definitions, Architectures and Recent Trends. *Electronics* **2024**, *13*, 782. [CrossRef]
43. Akpan, I.J.; Offodile, O.F. The Role of Virtual Reality Simulation in Manufacturing in Industry 4.0. *Systems* **2024**, *12*, 26. [CrossRef]
44. Martínez-Olvera, C.; Mora-Vargas, J. A comprehensive framework for the analysis of Industry 4.0 value domains. *Sustainability* **2019**, *11*, 2960. [CrossRef]
45. Nota, G.; Nota, F.D.; Peluso, D.; Toro Lazo, A. Energy efficiency in Industry 4.0: The case of batch production processes. *Sustainability* **2020**, *12*, 6631. [CrossRef]
46. Niekurzak, M.; Brelik, A.; Lewicki, W. The economic potential for the recovery and recycling of silicon photovoltaic cells and non-ferrous metals as part of the transition to a circular economy. *Econ. Environ.* **2023**, *86*, 202–224. [CrossRef]
47. Agarwal, A.; Ojha, R. Prioritizing implications of Industry-4.0 on the sustainable development goals: A perspective from the analytic hierarchy process in manufacturing operations. *J. Clean. Prod.* **2024**, *444*, 141189. [CrossRef]
48. Illgen, B.; Sender, J.; Flügge, W. Simulation-based production support system for steel structures for large offshore structures. *CIRP Procedia* **2019**, *81*, 204–209. [CrossRef]
49. Niekurzak, M.; Mikulik, J. Modelling of Energy Consumption and Reduction of Pollutant Emissions in a Walking Beam Furnace Using the Expert Method—Case Study. *Energies* **2021**, *14*, 8099. [CrossRef]
50. Wróblewski, P.; Niekurzak, M. Assessment of the possibility of using various types of renewable energy sources installations in single-family buildings as part of saving final energy consumption in Polish conditions. *Energies* **2022**, *15*, 1329. [CrossRef]
51. Renna, P.; Mancusi, V. Controllable processing time policy in job shop manufacturing systems: Design and evaluation by simulation modelling. *Int. J. Serv. Oper. Manag.* **2017**, *27*, 366. [CrossRef]
52. Jia, Y.; Tian, H.; Chen, C.; Wang, L. Predicting the availability of production lines by combining simulation and surrogate model. *Adv. Prod. Eng. Manag.* **2017**, *12*, 285–295. [CrossRef]
53. Krenczyk, D.; Davidrajuh, R.; Skolud, B. Comparing Two Methodologies for Modeling and Simulation of Discrete-Event Based Automated Warehouses Systems. In *Advances in Manufacturing II. Lecture Notes in Mechanical Engineering*; Hamrol, A., Kujawińska, A., Barraza, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2019.
54. Mikulik, J.; Niekurzak, M. *Impact of a Photovoltaic Installation on Economic Efficiency on the Example of a Company with High Energy Consumption*; Zeszyty Naukowe, Organizacja i Zarządzanie seria nr 169; Politechnika Śląska: Gliwice, Poland, 2023.
55. Zamora-Antuñano, M.A.; Cruz-Salinas, J.; Rodríguez-Reséndiz, J.; González-Gutiérrez, C.A.; Méndez-Lozano, N.; Paredes-García, W.J.; Altamirano-Corro, J.A.; Gaytán-Díaz, J.A. Statistical Analysis and Data Envelopment Analysis to Improve the E_ciency of Manufacturing Process of Electrical Conductors. *Appl. Sci.* **2019**, *9*, 3965. [CrossRef]
56. Gola, A.; Wiechetek, Ł. Modelling and simulation of production flow in a job-shop production system with enterprise dynamics software. *Appl. Comput. Sci.* **2017**, *13*, 87–97. [CrossRef]
57. Greenwood, A.; Pawlowski, P.; Bocewicz, G. A conceptual design tool to facilitate simulation model development: Object flow diagram. In Proceedings of the 2013 Winter Simulation Conference, Washington, DC, USA, 8–11 December 2013; pp. 1292–1303.
58. Veisi, B.; Farughi, H.; Raissi, S. Two-Machine Robotic Cell Sequencing under Di_erent Uncertainties. *Int. J. Simul. Model.* **2018**, *17*, 284–294. [CrossRef]
59. Tran, N.-H.; Park, H.-S.; Nguyen, Q.-V.; Hoang, T.-D. Development of a Smart Cyber-Physical Manufacturing System in the Industry 4.0 Context. *Appl. Sci.* **2019**, *9*, 3325. [CrossRef]
60. Matheson, E.; Minto, R.; Zampieri, E.G.G.; Faccio, M.; Rosati, G. Human-robot collaboration in manufacturing applications: A review. *Robotics* **2019**, *8*, 100. [CrossRef]
61. Hietanen, A.; Latokartano, J.; Pieters, R.; Lanz, M.; Kamarainen, J.-K. AR-based interaction for human-robot collaborative manufacturing. *Robot. Comput.-Integr. Manuf.* **2020**, *63*, 101891. [CrossRef]
62. Long, P.; Chevallereauo, C.; Chablat, D.; Girin, A. An industrial security system for human-robot coexistence. *Ind. Robot* **2018**, *45*, 220–226. [CrossRef]

63. Satyro, W.C.; Contador, J.C.; Monken, S.F.d.P.; Lima, A.F.d.; Soares Junior, G.G.; Gomes, J.A.; Neves, J.V.S.; do Nascimento, J.R.; de Araújo, J.L.; Correa, E.d.S.; et al. Industry 4.0 Implementation Projects: The Cleaner Production Strategy—A Literature Review. *Sustainability* **2023**, *15*, 2161. [CrossRef]
64. Nigischer, C.; Reitererer, F.; Bougain, S.; Grafinger, M. Finding the proper level of detail to achieve sufficient model fidelity using FlexSim: An industrial use case. *Procedia CIRP* **2023**, *119*, 1240–1245. [CrossRef]
65. Sikora, A. An example of the use of the FlexSim simulation program in improving the functioning of a warehouse. In Proceedings of the 21st International Conference on Urban Transport and the Environment, València, Spain, 2–4 June 2015.
66. Caiado, R.G.G.; Scavarda, L.F.; Azevedo, B.D.; de Mattos Nascimento, D.L.; Quelhas, O.L.G. The challenges and benefits of sustainable Industry 4.0 for supply chain operations and management—A framework to achieve the 2030 Agenda. *Sustainability* **2022**, *14*, 830. [CrossRef]
67. Rosário, A.T.; Dias, J.C. How Industry 4.0 and Sensors Can Benefit Product Design: Opportunities and Challenges. *Sensors* **2023**, *23*, 1165. [CrossRef]
68. Wolniak, R.; Saniuk, S.; Grabowska, S.; Gajdzik, B. Identification of energy efficiency trends in the context of the development of Industry 4.0 on the example of the Polish steel sector. *Energies* **2020**, *13*, 2867. [CrossRef]
69. Martín-Gómez, A.M.; Agote-Garrido, A.; Lama-Ruiz, J.R. A Framework for Sustainable Manufacturing: Integrating Industry 4.0 Technologies with Industry 5.0 Values. *Sustainability* **2024**, *16*, 1364. [CrossRef]
70. Hozdić, E.; Jurković, Z. Cognitive Cyber-Physical Production Systems: A New Concept of Manufacturing Systems on the Route to Industry 5.0. In Proceedings of the International Conference "New Technologies, Development and Applications", Sarajevo, Bosnia and Herzegovina, 22–24 June 2023; Springer Nature: Cham, Switzerland, 2023; pp. 201–212.

applied sciences

MDPI

Article

Dynamic Job and Conveyor-Based Transport Joint Scheduling in Flexible Manufacturing Systems

Sebastiano Gaiardelli [1,†], Damiano Carra [1,*,†], Stefano Spellini [1] and Franco Fummi [2]

[1] Department of Computer Science, University of Verona, 37134 Verona, Italy
[2] Department of Engineering for Innovation Medicine, University of Verona, 37134 Verona, Italy
* Correspondence: damiano.carra@univr.it
† These authors contributed equally to this work.

Abstract: Efficiently managing resource utilization is critical in manufacturing systems to optimize production efficiency, especially in dynamic environments where jobs continually enter the system and machine breakdowns are potential occurrences. In fully automated environments, co-ordinating the transport system with other resources is paramount for smooth operations. Despite extensive research exploring the impact of job characteristics, such as fixed or variable task-processing times and job arrival rates, the role of the transport system has been relatively underexplored. This paper specifically addresses the utilization of a conveyor belt as the primary mode of transportation among a set of production machines. In this configuration, no input or output buffers exist at the machines, and the transport times are contingent on machine availability. In order to tackle this challenge, we introduce a randomized heuristic approach designed to swiftly identify a near-optimal joint schedule for job processing and transfer. Our solution has undergone testing on both state-of-the-art benchmarks and real-world instances, showcasing its ability to accurately predict the overall processing time of a production line. With respect to our previous work, we specifically consider the case of the arrival of a dynamic job, which requires a different design approach since there is a need to keep track of partially processed jobs, jobs that are waiting, and newly arrived jobs. We adopt a total rescheduling strategy and, in order to show its performance, we consider a clairvoyant scheduling approach, in which job arrivals are known in advance. We show that the total rescheduling strategy yields a scheduling solution that is close to optimal.

Keywords: scheduling; heuristic; makespan minimization

1. Introduction

Industry 4.0 has significantly reshaped manufacturing system paradigms, emphasizing the move toward the complete automation of the production process [1,2]. In this transformative context, where the goal is to fully exploit the capabilities of machines, the precise planning of all production *tasks* becomes imperative, aiming to minimize machine idle time. Consequently, the scheduling problem has been the subject of extensive study in recent decades [3–5]. In its simplest form, this scheduling challenge is commonly referred to as the job shop scheduling (JSS) problem [6]. Given a set of *machines* and a set of *jobs*, where each job comprises a set of *tasks* to be processed in a specific order by different machines, the objective is to find a task assignment that minimizes an objective function, such as the total completion time or *makespan*.

The above problem formulation assumes that the time is divided into slots (e.g., each day is a time slot): the job requests are collected during a time slot and scheduled for the next one. By relaxing this assumption, we can insert the jobs as they arrive into the schedule. This scenario is usually referred to as *dynamic JSS* (DJSS). With the evolution of manufacturing technology, in which a single machine may perform multiple task types, we can further extend JSS to include this flexibility; in such a case, we have the flexible

DJSS problem (FDJSS or simply FJSS). The FJSS problem has been particularly interesting in recent years in the context of *Industry 4.0* [7], in which the production lines include advanced features that allow the machines to communicate and be reorganized and reconfigured to meet the increasingly challenging production constraints.

A crucial aspect that is not fully explored in the existing literature is the automation of *job transfers* between machines. While task execution on machines can be characterized by processing times, the transfer process involves various settings, including (i) the means of transportation, such as automatic guided vehicles (AGVs) or conveyor belts, and (ii) input and output buffers at the machines, determining their ability to store jobs for processing (e.g., while finishing another task). The store processed jobs (while awaiting the availability of the transport system) (Our definition of *transfer* pertains to the time between the end of one task and the start of the next for a given job).

Certain combinations of transport systems and buffers may be modeled using constant time [8], as is the case when the transport facility is always available, and the machines have sufficiently large input and output buffers. In such scenarios, incorporating transfer time into the processing time is straightforward, and numerous solutions proposed in the literature [4] can be applied. In other cases, transfer time depends on the availability of the transport system and the destination machine [1,9,10]. Overall, no single model can cover all alternatives, necessitating ad-hoc modeling for specific combinations.

In this study, we consider a scenario inspired by a *fully automated production line*, incorporating a conveyor belt as a means of transportation and no buffer at the machines. In this setup, once a task is completed, the job is immediately placed on the belt to free the machine. If the next machine is occupied, the job remains on the belt until the destination machine becomes available. The JSS problem is well-known to be NP-hard [11], and the variant considered here is at least as challenging [12]. In order to address this complexity, we propose a heuristic named SCHED-T, which falls under the *stochastic local search* (SLS) approach [13]. SLS encompasses well-known algorithms such as simulated annealing and tabu search.

A primary challenge is evaluating potential moves when exploring the solution space. A slight change in the scheduling sequence has a cascading effect on the remaining tasks, as the sequence depends on transfer times, which, in turn, depend on the execution sequence. While existing solutions rely on an approximate evaluation of each move (e.g., the computation of the critical path), we adopted a randomized approach, accurately assessing a few random neighbors. Given SCHED-T's ability to quickly compute complex schedules, especially in the case of dynamic job arrivals, we employed the "total rescheduling" strategy [14]. This strategy considers newly arrived jobs alongside those not yet scheduled, resulting in a new schedule that integrates with the current one. This approach is versatile, accommodating events beyond job arrivals, such as scheduled machine maintenance [15].

Here, we evaluate SCHED-T on a set of instances publicly available without transfer times [16]. In such cases, our heuristic achieves results comparable to those of the previously proposed versions. Subsequently, we analyze instances generated from a production line where the transfer times are available. Although the problem can be modeled using mixed linear integer programming (MILP) approaches, standard MILP solvers fail to find a solution in a reasonable time. Consequently, SCHED-T emerges as the only viable approach. Our results demonstrate that when applying the schedule found by SCHED-T to a real-world production line, the predicted makespan closely aligns with the actual outcome. Conversely, using a schedule derived without considering the transport system, as seen in the literature, results in a makespan that is up to 30% larger than that achieved using SCHED-T's schedule.

In our approach, we consider dynamic arrivals and compare the total rescheduling strategy using clairvoyant scheduling, i.e., a scheduling type wherby future job arrivals are known in advance, and show that the total rescheduling strategy produces close-to-optimal scheduling.

This article builds upon the findings presented in [17] in several ways. Specifically, we delve into the scenario of dynamic arrivals, where the schedule must be recalculated,

which is in contrast to the static case examined in [17], where all jobs were available at the onset of the scheduling period. This necessitated a redesign of the scheduler to manage partially processed, waiting, and newly arrived jobs. Consequently, this updated definition encompasses the static case as a special instance, thus representing a generalization of the problem. Furthermore, we present a clear system model and problem formulation, along with an expanded comparison of our heuristics. This comparison includes additional instances, additional performance metrics, such as the running time of our scheduler, and diverse layouts for real-world experiments.

In summary, the contributions of our work are the following. We provide a heuristic for the JSS problem that takes into account the transport system based on a conveyor belt and no buffers at the stations. We consider a dynamic scenario where jobs continuously arrive while we managed the partially processed jobs and the newly arrived jobs, making for a total rescheduling strategy. We tested our solution in a real-world production line to show that if transport times are not considered, the scheduler produces results that may contain large errors.

In summary, the contributions of our work are as follows:

- We propose a heuristic for the Job Shop Scheduling (JSS) problem that incorporates a transport system based on a conveyor belt and no buffers at the stations.
- We address a dynamic scenario where jobs continuously arrive while we effectively manage partially processed jobs and newly arrived jobs by adopting a total rescheduling strategy.
- We validated our solution through experimentation on a real-world production line, demonstrating that neglecting transport times can lead to substantial errors in the scheduler's results.

The paper is structured as follows. In Section 2, we present the case study that serves as the motivation for our work. Specifically, we examine a production line where machines are interconnected through a conveyor belt. We then delve into the existing body of related work and highlight their limitations. Section 3 is dedicated to formalizing the problem, introducing a model that defines the key variables. This formalization enables us to cast the problem as a minimization challenge. Given the impracticality of finding an exact solution within a reasonable timeframe, Section 4 outlines our proposed heuristic, SCHED-T. This heuristic belongs to the class of *randomized iterated improvement* algorithms tailored to address the nuances of the specific problem. In Section 5, we conduct a comparative analysis of the solutions obtained with SCHED-T against those generated by state-of-the-art heuristics. We explore scenarios involving the transport system that were previously unsolvable with tools available in the literature. Finally, Section 6 encapsulates the key findings and draws conclusions from our study.

2. Background and Related Work

Production line with a conveyor belt: The manufacturing system, taken as a reference, consists of a set of machines dedicated to specific tasks disposed in a general *layout*. Each machine is connected with the others through a transport system that moves the materials between the machines circularly. Both raw and finished materials are stored in a vertical warehouse. The production line available at our research facility, *the ICE Laboratory* [18], follows the same layout. Figure 1 shows the plant configuration. It is composed of a set of "production cells", specifically tailored to a specific manufacturing process. Starting from the right, the laboratory includes a multi-tool milling machine, a robotic assembly station, a quality control cell, and a vertical warehouse. The conveyor belt moves materials on top of 10 pallets across various *belt segments*. Each pallet is identified by an RFID tag, which is detected by RFID sensors located near the switching mechanisms (blue squares in Figure 1). In particular, the belt segments can be differentiated into the following:

- *The four machine segments*: These are in charge of loading and unloading materials from the machines (i.e., moving pallets toward and away from the processing area).

Since these segments are the only access route to the machines, they can contain only one pallet at a certain time instance;

- *The main belt segments*: These are in charge of moving a pallet near the machine when it is ready to process a new task. The pallets waiting for a machine to become free must loop near the destination machine, implementing a circular buffer, while free pallets loop along the long main segments. The main belt is composed of two long belts running in opposite directions. Along these main segments, it is possible to traverse from one to the other by exploiting the switching mechanisms placed in certain positions.

The two types of belt segments differ in terms of their control policy. Machine segments are activated in the desired direction to perform loading or unloading pallet operations, and the main belt segments are always active, pushing in the same direction. A switching mechanism can move a pallet by choosing one of the following directions: (1) forward, following the direction of the actual belt; (2) backward, changing direction by switching onto the opposite conveyor; and (3) towards a machine working area. If two pallets must move in the same direction at the same time, the pallet exiting from an unloading segment or moving forward along the main belt has precedence.

Literature review: In the last decades, the JSS problem has been extensively analyzed by the research community, along with the different variants, such as DJSS and FJSS. The solutions that have been proposed over the years are summarized in [4,5] and [19]. Our work differs from the standard problem because it specifically addresses the transport time between machines and the impact of this on the schedule. Therefore, we concentrated our analysis of the related works on this specific topic.

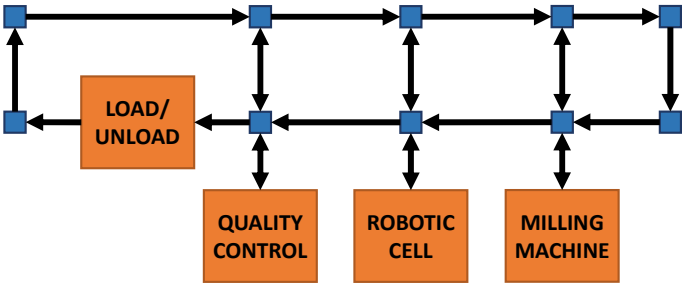


Figure 1. The layout of the production line used as a case study.

The transfer time between two machines is a variable delay that depends on the distance between the two machines and the availability of the destination machine. A solution proposed in [20] handles variable delays as sequence-dependent setup times. However, this solution considers the machines busy during the setup time, whereas, in our case, the destination machine can process other tasks while the materials of the next task are moving toward the machine.

The authors in [10,21–23] proposed a solution for the problem of job scheduling combined with transport time, but they consider only automated guided vehicles (AGVs) as a means of transportation. The limit of considering AGVs is that until a vehicle picks up a task from a machine, the machine itself cannot execute other activities. Similarly, when an AGV is waiting to perform a pick operation from a machine, it wastes transport resources. Moreover, if the machine is completing the previous task when the AGV reaches the destination, it must wait until the machine completes the task, freeing its working area. This is not a valid scenario in our case study, in which the machine immediately unloads the task to the conveyor belt once the task is completed. Another difference is that the jobs waiting to be processed by a machine do not impact on the unloading and loading capabilities of the tasks of other machines.

The aim of the different heuristics proposed in the literature is to find the best schedule that optimizes an objective function [12,20,24]. These state-of-the-art solutions are based on graph-based representations of the execution sequence. By evaluating the graph properties,

such as dependencies and the critical path, it is possible to infer the feasibility of a solution by also approximating the objective function. In our case, these approaches cannot be applied due to the inter-dependency between transfer time and the execution sequence. There are also solutions, such as [16], based on MILP optimization models, which minimize the objective function. A major drawback of these approaches consists of the execution time. In fact, by modeling the constraints of our transport system, solving even small instances cannot be carried out in a reasonable amount of time.

In summary, none of the existing works take into account the influence of a transport system on comprehensive scheduling when machines lack input and output buffers. In this specific setup, the sequence of jobs affects the transport times, and conversely, the transport times impact the job sequence. This intricate circular dependency necessitates specialized tools that are currently absent in the literature.

3. Problem Description

3.1. Model

The flexible job shop problem concerning no buffers and a transportation system can be described as follows. The facility has a set of m machines ($M = \{1, 2, \ldots, m\}$) that are used to process a set of n jobs ($J = \{1, 2, \ldots, n\}$). Since we consider a dynamic environment in which jobs continuously arrive, the set of jobs ($J(t)$) depends on the time (t), i.e., we have $n(t)$. Such a dynamic set is composed of (i) the jobs that have arrived but have not been started yet and (ii) the newly arrived jobs.

Each job (i) comprises h_i operations or *tasks* (τ_{ij}, where i denotes the job, and $j = 1, 2, \ldots, h_i$), and each task can be executed on a subset of machines. The processing time (p_{ijk}) for a given task (j) of a job (i) on a machine (k) is considered known. If a task (τ_{ij}) cannot be performed on machine (k), we set $p_{ijk} = \infty$. The tasks within a job may have *precedence constraints*, meaning that a task cannot start until its preceding tasks (if any) are completed. For each job (i), the precedence constraints are summarized by a square matrix (U_i). An element ($U_{ijj'}$) is set to 1 if task τ_{ij} precedes task $\tau_{ij'}$, and this is 0 otherwise.

Each machine (k) is associated with a set of available times ($A_k(t)$), representing intervals during which tasks can be executed on that machine. Outside these intervals, the machine is unavailable either because it is busy with other tasks (e.g., a previous schedule is still being executed) or because it is undergoing maintenance. Each machine performs, at most, one task of any job at a time. With each new arrival, the availability sets for all machines are recomputed, considering the currently running jobs.

The time required to move a job from machine k to machine l, denoted by t_{kl}, is a key aspect of our work, particularly as we assume that machines lack input or output buffers and a conveyor belt serves as the means of transportation. More specifically,

$$t_{kl} = t^0_{kl} + t^c_{kl} \cdot n^c \tag{1}$$

Here, t^0_{kl} represents the minimum time to travel from machine k to machine l, t^c_{kl} is the cycle time in case machine l is unavailable, and n^c is the number of cycles the job needs to complete until the machine becomes available. These values are contingent on the layout of the production line (machines and conveyor belts) and can be readily measured once such a layout is defined. It is worth noting that by appropriately setting these values, we can encompass scenarios commonly explored in the literature: if $t^c_{kl} = 0 \; \forall k, l$, the transport time remains constant; if $t_{kl} = 0 \; \forall k, l$, the model excludes the transport system. Hence, our model is versatile, covering various layouts and accounting for cases where transport time is not considered.

Table 1 summarizes the notation used for the model and the formulation of the problem.

Table 1. Notation summary.

Inputs	
M	Set of machines, $\{1, 2, \ldots, m\}$
$J(t)$	Set of jobs $\{1, 2, \ldots, n(t)\}$
h_i	Number of tasks of job i
τ_{ij}	jth task of job i, $j \in [1, h_i]$
p_{ijk}	Processing time for τ_{ij} on machine k
U_i	Precedence matrix for job i: 1 if τ_{ij} precedes $\tau_{ij'}$; otherwise, it is 0
$A_k(t)$	Availability intervals for machine k
t_{kl}	Transfer time between machines k and l
Auxiliary Variables	
E_{ijk}	1 if τ_{ij} is executed on machine k; otherwise, it is 0
s_{ijk}	Execution start time of τ_{ij} on machine k
C_{ij}	Completion time of τ_{ij}

3.2. Problem Formulation

We consider a dynamic environment in which jobs continuously arrive. We adopt a total rescheduling strategy, i.e., we recompute the schedule at every job arrival; we discuss such a choice in detail in Section 4.4. The aim of the scheduling is to minimize the *makespan*, i.e., the total execution time required to process all the tasks of the current set of jobs $J(t)$. Formally, we have

$$\underset{i \in [1, n(t)]}{\text{minimize}} \quad \max\ C_{ij} \tag{2}$$

$$\text{s.t.}$$

$$s_{ijk} + p_{ijk} = C_{ij} \tag{3}$$

$$s_{ijk} + p_{ijk} + t_{kl} \leq s_{ij'l} \qquad \forall \tau_{ij}, \tau_{ij'} | U_{ijj'} = 1 \tag{4}$$

$$\sum_k E_{ijk} = 1 \qquad \forall \tau_{ij} \tag{5}$$

$$s_{ijk} + p_{ijk} < s_{i'j'k} \qquad \forall \tau_{ij}, \tau_{i'j'}, k \in [1, m] \tag{6}$$

$$\Big[s_{ijk},\ s_{ijk} + p_{ijk}\Big] \in A_k(t) \qquad \forall \tau_{ij},\ \forall k \tag{7}$$

Equation (3) specifies that the end time of each task must be equal to its start time plus its processing time. Here, we omit the necessary time to store the material produced by the last task of each job. Equation (4) models the dependencies between tasks within the same job. If a task (j) is assigned to a machine (k), it has precedence over another task (j') assigned to the machine (l); the starting time of the task (j') must be greater or equal to the end time of the task (j) plus the transport time between the two machines. Equation (5) imposes the maximum number of machines to which a task can be assigned, i.e., one. Equation (6) limits the number of tasks a machine can process in parallel. Finally, Equation (7) indicates that tasks can be executed on the machines only during their availability intervals.

The provided equations can be utilized to formulate and solve the constraints of a mixed linear integer programming (MILP) model. However, conventional MILP solvers, such as IBM CPLEX, face challenges in finding a solution within a reasonable time frame; more specifically, even with a modest number of jobs, they are unable to converge within a 12 h timeframe on standard hardware. This computational inefficiency arises due to the incorporation of transport time in Equation (4), where t_{kl} is contingent on the values derived from Equation (1), significantly escalating the computational complexity of the model. Consequently, we must turn to alternative methodologies that involve exploring the solution space.

4. Exploring the Solution Space

4.1. Overview

The general framework for any heuristic based on stochastic local search comprises the following steps [13]:

1. Construct a solution and compute the objective function;
2. Explore the neighborhood, assessing the objective function for each neighbor;
3. Choose the neighbor based on a specified criterion;
4. Repeat all steps starting from step 2 until a stop condition is satisfied.

In addition, given a solution, we must define (i) how we can explore its neighborhood (i.e., new solutions); (ii) how the objective function is computed (i.e., what are the parameters that we want to optimize); (iii) the policy used to choose the next solution; and (iv) when the exploration should terminate (i.e., stop conditions). The computational complexity of these approaches depends on the neighborhood size. For instance, given a task sequence (s_i) containing n elements, we may define (as a neighbor) any sequence (s_j) that differs from s_i for the position of two tasks (the minimum possible change). This would imply that the number of elements in the neighborhood is at least equal to $n(n-1)/2$.

Given the high number of neighbors used to explore each solution, a common procedure consists of limiting the exploration to the subset of the *most significant* neighbors. This set is composed of all the neighbors except those that are less likely to improve the current solution, e.g., all the neighbors that swap two tasks belonging to the same job since they have dependencies that limit the time interval in which they can be allocated.

The cost function must be computed for each explored neighbor. Therefore, its complexity heavily impacts the execution time. A cost function should be a trade-off between precision and complexity, exploring a wide range of neighbors in a certain time slot and choosing the most promising one. In our case, the complexity of the cost function is given by two main factors. The first one consists of the inner dependencies between tasks of the same job. This implies that small changes in one task can cause cascading effects on the subsequent portion of the schedule. The second factor is related to the transfer time between two machines. This transfer time is strictly dependent on both the availability of the destination machine (see Equation (1)) and the task execution order. While the first problem can be addressed by exploiting graph-based representations, the circular dependency between execution order and transport time makes it difficult to precisely or approximately compute the objective function.

As such, we are forced to compute the objective function for each neighbor explored from scratch, making the selection of the most promising neighbors to explore a critical step.

4.2. Randomized Approach

Given a set of jobs to schedule, the number of possible solutions increases exponentially as the number of jobs increases. Therefore, an entire exploration of the solution state space could require hours or could be unfeasible in a reasonable amount of time. For this reason, the exploration is always guided by heuristics that reduce the number of evaluated solutions. An idea inspired by the results proposed in [25–27] is to explore a set of neighbors chosen randomly. This is indeed a variant of the stochastic local search approach called *probabilistic random improvement*.

A random neighbor (i.e., solution) is described by the sequence of jobs. For each neighbor, we compute the allocation of tasks on the machines and evaluate the objective function. Given a solution, we consider its set of *randomly selected* neighbors. Then, we choose the best neighbor that improves the current solution. This process is repeated to increase the exploration's precision. Thus, we can also tune the precision of the exploration by increasing or decreasing the total number of random neighbors to explore.

During the exploration, we must avoid being stuck in a local minimum. We can do that by adopting known techniques used in tabu search or simulated annealing, in which a new solution is accepted with a certain probability and without considering the optimality degree of the solution. Instead, we adopt another methodology based on sampling, which

is inspired by the fisheye view [28] and fisheye routing [29]. These techniques sample solutions at different distances to check if there are other *valleys* that exist to explore. In our context, a distant neighbor is obtained with more complex changes in the task execution order. In addition, for each remote neighbor, we perform a limited local exploration to test if such a remote neighbor could indeed improve the current solution. Figure 2 shows an example of the exploration process.

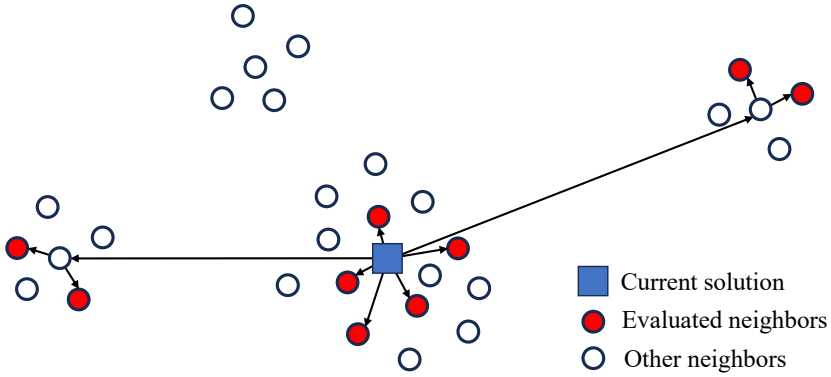


Figure 2. Example of the randomized exploration approach. We start from the current solution and evaluate some random close neighbors, along with remote neighbors (chosen randomly).

4.3. Detailed Solution

Our scheme for the exploration of the neighborhood uses a two-level hierarchical approach. At a higher level, we have jobs, while at the lower level, we have tasks.

Initialization: In the case of a cold start, in which the system is completely unloaded, and a set of jobs has been collected and are ready to be executed (e.g., during the night, the production line does not work and arriving orders are collected), we initialize the order of the jobs according to the *longest processing first* policy since it has been shown to improve the makespan compared other policies [30]. We define the processing time of a job as the sum of the processing time of its tasks. If a task (τ_{ij}) can be processed by more than one machine, then we consider the minimum (p_{ijk}) (see Algorithm 1, procedure INITTASKORDER, and parameter *BySize* set to TRUE). In the case of a running system, the order is given by the output of the previous schedule (considering only the jobs not yet launched), with the newly arrived job placed at the end.

We then establish the order of tasks within each job by grouping tasks with the same precedence and randomly arranging the tasks within each group. The final output is an ordered sequence of tasks, denoted as s_l. The makespan value is computed based on this sequence (refer to Algorithm 1 and procedure EVALSOLUTION). We allocate each task to a machine, considering machine availability and dependencies for previous tasks. If multiple machines can perform a task, we choose the machine where the task terminates earlier, including the transportation time.

Neighborhood: For a given sequence (s_l) of tasks, we define a *close neighbor* as the feasible solution where we work at a low level. In other words, we switch tasks to obtain a new sequence ($s_{l'}$) for evaluation. A random neighbor is obtained by (i) uniformly selecting a task (τ_{ij} in s_l) and (ii) uniformly selecting another task ($\tau_{i'j'}$). The latter must belong to the same job ($i' = i$) or to a job that comes before or after job i in the ordered set J. If the tasks belong to the same job, our choice is limited to tasks with no dependencies on task τ_{ij} (such a subset can be pre-computed upon job arrival).

We define a *remote neighbor* as a new solution in which jobs are swapped in the ordered set

J (i.e., higher-level permutations). Thus, remote neighbors allow for the exploration of new solution areas, and close neighbors can be used to fine-tune the current solution.

Algorithm 1: Initialization and Evaluation

```
1  Procedure InitTaskOrder(J, BySize):
2      if BySize then
3          foreach job i do
4              i(size) ← Σ_j min_k p_ijk ;
5          J ← OrderBySize(J);
6      s_l ← ∅;
7      foreach job i ∈ J do
8          G_i ← GroupByPrecedence(i);
9          foreach group g ∈ G_i do
10             g ← RandomOrder(g);
11             AddTasks(s_l, g));
12     return s_l
13 Procedure EvalSolution(s, {A}):
14     foreach task τ_ij ∈ s do
15         q ← arg min_k(t_lk + p_ijk);
16         Update(A, q, τ_ij);
17     return max C_ij;
```

Exploration: The number of close and remote neighbors (solutions) explored during an iteration is governed by the budget, B. At each iteration, we update the current solution with a new one if it improves the current one (see Algorithm 2). The exploration starts by considering the current solution s_i and αB as close neighbors, chosen uniformly at random among the feasible close neighbors (lines 6–12). The parameter α is given as input, with $0 < \alpha < 1$. The exploration continues by evaluating R remote neighbors (lines 13–14), selected uniformly at random. For each of them, it explores its neighbors locally uniformly at random with a budget $(1 - \alpha)B/R$ (lines 15–21).

The exploration stops when one of the following conditions is reached: (i) the maximum number of iterations $T_{\max}$; (ii) no improvements are found for t_{idle} iterations (not shown in Algorithm 2).

Figure 3 shows a high-level view of the processing flow.

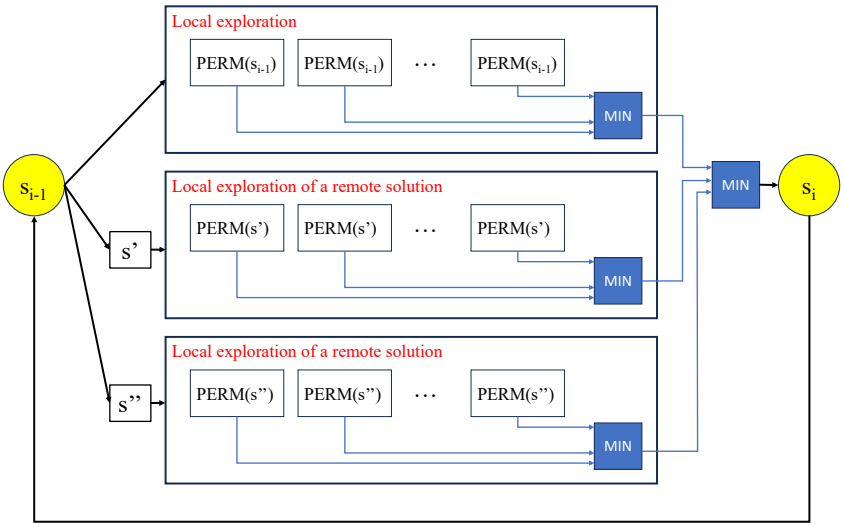


Figure 3. Exploration flow: starting from a solution at time $t - 1$, we explore the local neighbors, along with a set of distant neighbors. If we find a better solution, we update the current solution.

Algorithm 2: SCHED-T

input: J, set of jobs to be scheduled
input: M, set of machines
input: $\{A_0\}$, availability intervals
input: *BySize*, True if jobs need to be ordered by size
input: B, α, R, parameters for exploration

$s_0 \leftarrow$ INITTASKORDER(J, *BySize*);
$\mathcal{O} \leftarrow$ EVALSOLUTION($s_0, \{A_0\}$);
$i = 1$;
while $i \leq T_{max}$ **do**
  $s_i \leftarrow s_{i-1}$;
  for αB *times* **do**
    $s'_i \leftarrow$ RANDLOCALPERM(s_{i-1});
    $\mathcal{O}' \leftarrow$ EVALSOLUTION($s'_i, \{A_{i-1}\}$);
    if $\mathcal{O}' < \mathcal{O}$ **then**
      $s_i \leftarrow s'_i$;
      $\mathcal{O} = \mathcal{O}'$;
      $\{A_i\} \leftarrow$ UPDATEAVAL(s_i);
  for R *times* **do**
    $s'_{i-1} \leftarrow$ RANDREMOTEPERM(s_{i-1});
    for $(1-\alpha)B/R$ *times* **do**
      $s'_i \leftarrow$ RANDLOCALPERM(s'_{i-1});
      $\mathcal{O}' \leftarrow$ EVALSOLUTION($s'_i, \{A_{i-1}\}$);
      if $\mathcal{O}' < \mathcal{O}$ **then**
        $s_i \leftarrow s'_i$;
        $\mathcal{O} = \mathcal{O}'$;
        $\{A_i\} \leftarrow$ UPDATEAVAL(s_i);
  i++;

Complexity: When evaluating a single solution, since we build it from scratch, we run through the ordered list of tasks and assign the task to the available machine. Each task may be executed in more than one machine, but we assume that the number of alternative machines is bounded. The complexity of each evaluation, therefore, is $\mathcal{O}(H)$, with $H = \sum_i h_i$ total number of tasks considering all the jobs to be scheduled. We perform $T_{\max}$ iterations, and we evaluate B possible solutions; therefore, the complexity of SCHED-T is $\mathcal{O}(HBT_{\max})$. In Section 5, we show the running time for different instances.

4.4. Discussion

When considering dynamic job arrivals, there are different approaches for integrating the new job into the existing scheduling. For instance, one can exploit the gaps in the current schedule and run the newly arrived job during such gaps, with little or no impact on the other jobs. When we take into account the transport system, it is not easy to understand if the gaps can be fully exploited. In fact, even a small change in the current scheduling, such as shifting a task to accommodate the new one, results in a cascading effect that disrupts the whole scheduling.

Such interdependence between the job execution and the transport system is the reason behind our randomized approach to the exploration of the solution space. Every time we evaluate a possible solution, we need to compute the schedule from scratch. The solutions in the literature that adapt the current schedule to accommodate the new arrivals can not be easily extended when we consider the transport system, and the only available option is to recompute the whole schedule.

The use of a heuristic allows us to trade accuracy with speed, i.e., we accept sub-optimal solutions (with an error of less than 5%) that are obtained in a few seconds;

in Section 5, we show the running time for different instance sizes. Compared to the job processing time (e.g., tens of minutes), the scheduling processing time is very small, and therefore, the approach based on total rescheduling is justified. In case of high job arrival rates, rather than recomputing the scheduling at every arrival, it is possible to collect some new jobs before running SCHED-T. As a rule of thumb, if the scheduling processing time is P_{proc} seconds, and the average arrival rate (estimated considering the last arrivals) is V_{arr} jobs/s, then the new scheduling can be computed once $\lceil \beta \cdot V_{\text{arr}} \cdot P_{\text{proc}} \rceil$ jobs have arrived, with the parameter $\beta \geq 1$ that controls the trade-off between the delay and processing load.

Alternatively, at every job arrival, it is possible to consider only a subset of jobs waiting to be processed (e.g., the last W jobs in line) as part of the set of jobs that are included in the new schedule computation so that the other subset can be used to feed the production line during the computation.

5. Experimental Results

In this section, we demonstrate the validity of our proposed heuristic SCHED-T. First, we compare our heuristic with state-of-the-art approaches for solving the FJSS problem, using public benchmarks without the transport system. Then, we apply our heuristic to both a real-world scenario and a simulated scenario to evaluate its performance when the transport system is present.

5.1. Experimental Methodology and Settings

In order to compare SCHED-T with other heuristics, we consider the set of publicly available instances described in [12,16]. These instances have been created using the instance generator available at [31]. The generated instances contain information regarding the jobs, assuming that transportation is part of the task-processing time. We use these instances to evaluate the performance of our proposed scheduling heuristic with respect to the state-of-the-art. In all the comparisons, we use a cost function based on the *makespan*, i.e., the latest completion time of the entire set of scheduled tasks. The built-in parameters that guide our heuristics are (i) the number of iterations and (ii) the available budget at each iteration, divided between the budget dedicated to the exploration of local and remote neighbors. We analyzed the impact of these parameters on the makespan with a sensitivity analysis.

As for the transport system, unfortunately, public instances that include this aspect, such as the one used in [32], consider different means of transport (AGVs rather than conveyor belts). Therefore, they cannot be used in our comparison. Therefore, we consider a real-world use case related to our ICE lab, in which we were able to create a set of instances that include a transport system.

Finally, we consider the case of a sequence of arrivals. We assume, as is carried out in the literature [14], that a set of jobs is present at time zero and that another set randomly arrives. We evaluate the makespan in such a dynamic case (updated at every arrival), and we compare it with the makespan computed by an ideal clairvoyant scheduler, i.e., a scheduler that knows all the future arriving jobs and provides a single, optimized schedule.

SCHED-T is implemented in Python, and the experiments are carried out on a 3.3 GHz Intel Core i7 (Intel, Santa Clara, CA, USA) with 16 Gb of RAM (Microchip Technology Inc., Chandler, AZ, USA).

5.2. Instances with No Transport System

We compared SCHED-T with a set of heuristics that were presented in [12], in which authors applied iterated local search, genetic algorithms, differential evolution, and tabu search on a set of 50 *large* instances defined in the same paper, along with another set of 50 *large* instances proposed by the same authors in [16]. Each instance differs in the number of available machines (up to 97) on which tasks can be allocated, the number of jobs (up to 200), and the number of total tasks (up to 2000). None of these heuristics outperformed

the others in all the instances: for some instances, differential evolution was better than the others, and for other heuristics, tabu search provides the best makespan. Rather than listing the values of each heuristic, we consider, for each instance, the best and the worst makespan obtained by these heuristics: we used such values to represent an interval for which the gap is much more narrow than the one found with the CP approach. Within this range lie the results of the four heuristics cited before (iterated local search, genetic algorithm, differential evolution, and tabu search). Therefore, we consider such a gap as a reference with which to compare. Instead of showing the absolute values for each instance, we can normalize the maximum makespan with respect to the minimum makespan so that the gaps of the different instances are comparable and can be put in a single graph. We also normalize the makespan obtained by SCHED-T so that it is simple to understand if it falls in the gap provided by other heuristics.

Makespan: In Figure 4, we compare the makespan obtained on the set of instances by SCHED-T with the one obtained with the other reference heuristics. The results show that SCHED-T is able to find a makespan within the gap defined by the other state-of-the-art heuristics. In some cases, it outperforms the other heuristics by 2%, whereas in two cases, it finds a larger makespan by 3%. We notice that every heuristic proposed in the literature, including our solution, provides an approximate solution, and there are no guarantees on the error bound. Therefore, we can only observe how well these heuristics perform on real datasets. Figure 4 shows that the approximate solutions have a 4–12% gap; therefore, such an error can be considered acceptable. From these results, SCHED-T performs similarly to other state-of-the-art heuristics on instances without considering the transport time.

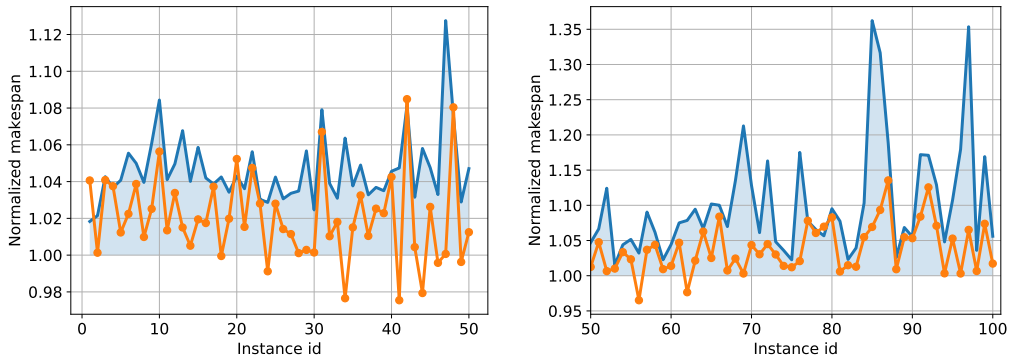


Figure 4. Normalized makespan for the set of the 50 + 50 large instances in [12,16]. The shaded area represents the gap (minimum and maximum makespan) found with different heuristics in [12], while the orange points are the results of SCHED-T.

Running time: Figure 5 shows the running time (in minutes) required to find the solutions of the 50 instances in [16]. The instances are ordered by increasing the number of tasks (shown at the top of the graph), e.g., instance 50 has almost 1000 tasks. Even with a large number of tasks, SCHED-T is able to obtain a solution that is comparable to the results found in the literature in less than 10 min.

Sensitivity analysis: SCHED-T has a set of parameters that tune the depth of the exploration phase. At the task level, we can exchange tasks belonging to the same job or to the next or previous job by considering the current scheduling order. At the job level, we can exchange jobs. During our tests, we noted that switching jobs provides most of the gain, while at the task level, the impact of changing the order on the makespan is less significant.

Therefore, the two main parameters we consider are the number of moves we explore $T_{\max}$ and the number of neighbors we evaluate in each iteration, R. These parameters influence the time required to compute the solution. Figure 6 shows two views for analyzing the impact of these parameters on the normalized makespan, i.e., the makespan found with a combination of $T_{\max}$ and R divided by the makespan found with the highest values of $T_{\max}$ and R used in our test. If we maintain $T_{\max}$ as fixed and we increase R (Figure 6, left), when

R is sufficiently large ($R > 10$), no additional improvement is observed. However, even for smaller values of R, the makespan only slightly increases (2–5%) with respect to the best makespan found. This is also confirmed if we keep R fixed and vary $T_{\max}$ (Figure 6, right); the two figures refer to two different instances, and they are representative of the general behavior that we observed for all the instances.

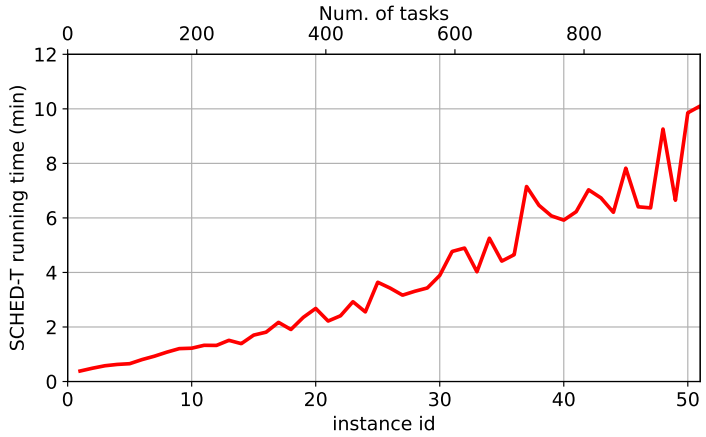


Figure 5. SCHED-T running time for finding the solution is shown in Figure 4. The top axis indicates the number of tasks associated with each instance.

Therefore, the two main parameters we consider are the number of moves we explore, denoted as $T_{\max}$, and the number of neighbors we evaluate in each iteration, denoted as R. These parameters influence the time required to compute the solution. Figure 6 provides two views for analyzing the impact of these parameters on the normalized makespan. In other words, it shows the makespan found with a combination of $T_{\max}$ and R divided by the makespan found with the highest values of $T_{\max}$ and R used in our test. If we keep $T_{\max}$ fixed and increase R (Figure 6, left), when R is sufficiently large ($R > 10$), no additional improvement is observed. Even for smaller values of R, the makespan only slightly increases (2–5%) compared to the best makespan found. This observation holds even if we keep R fixed and vary $T_{\max}$ (Figure 6, right). The two figures represent different instances but are indicative of the general behavior observed for all instances.

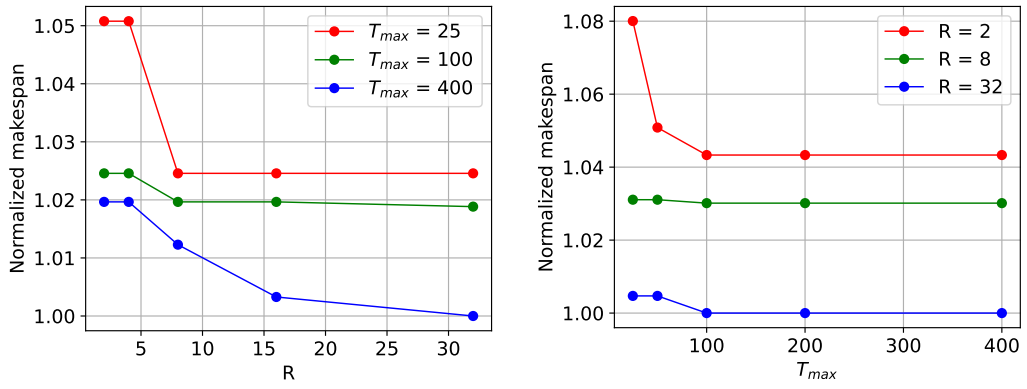


Figure 6. Sensitivity analysis of SCHED-T: impact of the parameters on the makespan (makespan normalized to the best value found). **Left** and **right** figures refer to different instances.

Overall, SCHED-T is not extremely sensitive to parameter settings and provides good results across a wide range of values. In terms of computational time, using small $T_{\max}$ and R allows for obtaining a solution in less than 30 s, while larger $T_{\max}$ and R may take up to 10 min. However, since large $T_{\max}$ and R do not significantly improve the results, the use

of a small $T_{\max}$ and R is appropriate, especially in dynamic contexts where scheduling is recomputed when new jobs arrive.

In any case, the tuning process can be performed in a simulated environment where data collected from the production line is replayed, and different values of $T_{\max}$ and R are tested. If the production load does not vary much from one day to the next, this offline analysis should provide an indication of the best parameter settings for that specific production line.

ICE instances: In order to test SCHED-T in a real test case, we generated six instances that can be executed within our research laboratory (described in Section 2). The number of tasks and jobs for each instance are summarized in Table 2. Without considering the transport time, we were able to solve the problem using a MILP model and compare the results with SCHED-T (Table 2, column **Opt**).

Table 2. ICE instances: characteristics (cols. 2 and 3); makespan with no transport (cols. 4, 5, and 6).

Id	#Jobs	# Tasks	Opt	SCHED-T	Err.
1	5	32	900	900	0%
2	10	106	3580	3650	2.0%
3	15	100	2930	3014	2.9%
4	20	142	3530	3698	4.8%
5	25	174	4935	4998	1.3%
6	30	263	7475	7570	1.3%

By using SCHED-T, we are able to find a solution in less than 30 s. The difference between the local optimal and the global optimal solution is less than 5% (columns **SCHED-T** and **Err.** of Table 2). As mentioned earlier, the literature deems an error gap of up to 10% acceptable, acknowledging the trade-off between accuracy and computational speed provided by the approximate solutions. Regarding the variability in error, we were unable to pinpoint any specific job characteristic that might influence such variability. It can be regarded as an inherent variability intrinsic to the considered problem.

5.3. Instances with the Transport System

Introducing a transport system increases the complexity of the MILP model exponentially. In fact, a MILP formulation given to standard MILP solvers, such as IBM CPLEX, cannot be solved after hours of computation. As such, we do not have any values that can be used as a global optimal comparison reference to our results. The results presented in Table 3, column **SCHED-T** of **Experim.-I**, report the makespan found by SCHED-T. Next, we describe the test we performed on a real-world testbed to validate such results.

Table 3. Reference layout scheduling instances with transport.

Id	Experim.-I			Experim.-II		
	SCHED-T	Actual	Err.	SCHED-T	Actual	Err.
1	1511 s	1556 s	−2.89%	1221 s	1555 s	−21.48%
2	2814 s	2861 s	−1.64%	2243 s	2875 s	−21.98%
3	3557 s	3674 s	−3.18%	3321 s	4248 s	−21.82%
4	4654 s	4892 s	−4.87%	4366 s	5577 s	−21.71%
5	5820 s	5963 s	−2.40%	5473 s	7012 s	−21.95%
6	7025 s	7227 s	−2.80%	6461 s	8377 s	−22.87%

5.4. Real-World Experiments

We consider the set of instances described in Table 2, and we run them on an actual production line built in our lab (described in Section 2). The production line is governed by

service-oriented manufacturing (SOM) software architecture similar to the one presented in [33,34], which automatically manages the production line, interacting with both the manufacturing execution system (MES) and the machines. On top, we have developed a module that implements our proposed scheduling heuristic, and the execution is forwarded to the SOM architecture. For each instance's job, we have fixed the task execution time. This allows for a fair evaluation of the accuracy of SCHED-T, removing processing time variability. The case of a stochastic execution time will be considered in future works.

Given the schedule returned by SCHED-T, we ran the jobs on the production line and recorded the makespan. The results are shown in Table 3, columns **Experim.-I**, and **Actual**. SCHED-T is able to predict the actual makespan obtained from a real-world production line with an error smaller than 5%. Additionally, in this case, there is no specific job characteristic that might influence the small variability for the different instances.

Then, in column **Experim.-II**, we consider what happens if the scheduler does not take into account the transport system, as most of the schedulers in the literature do. In this case, they would produce a schedule that is not optimal. In particular, we run SCHED-T by setting all the transport times to zero and we obtain a schedule that is fed to our production line. We record the makespan obtained in this case and compare the error obtained with the one computed in columns **Experim.-I**. If the scheduler does not include the transport times, the resulting schedule contains an error of up to 23%, which is 18% greater than the one obtained by the schedule considering the transport times. Thus, the makespan obtained using the real system by considering the transport time improves the makespan by up to 14% with respect to the one found without considering the transport time.

Table 4 shows the results obtained by applying SCHED-T to a different plant configuration depicted in Figure 7. This new configuration consists of a modified version of our real-world case study: we added a bay and a machine of the same type on the opposite side of each existing one to implement production redundancy. We built this new configuration by exploiting Tecnomatix Plant Simulation, a commercial state-of-the-practice discrete event simulation tool. The simulation allows for testing our proposed algorithm on a different scenario, enabling the calculation of simulated production times (e.g., makespan) and demonstrating that SCHED-T can also be applied to different plant configurations. We executed the same test using Table 3, generating new instances of the same size. The results show that when considering transport time, SCHED-T allows for reducing the error of the estimated makespan, keeping it within 5% and lowering it by 17% when not considering transport time.

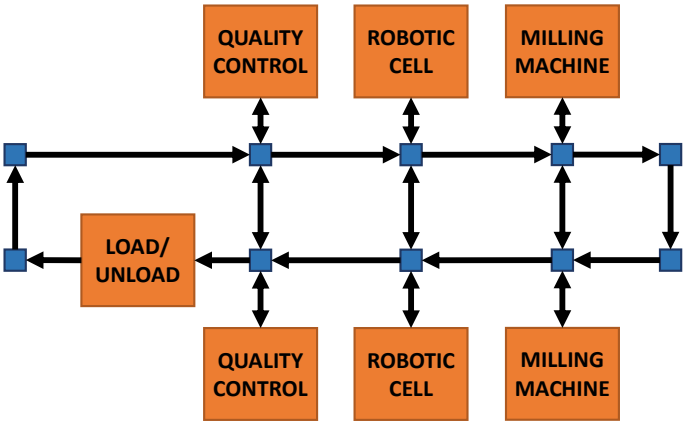


Figure 7. The modified layout of the production line used to test another scenario. We added a bay and a machine of the same type on the opposite side of each existing one.

Table 4. Modified layout scheduling instances with transport.

Id	Experim.-I			Experim.-II		
	SCHED-T	Actual	Err.	SCHED-T	Actual	Err.
1	1362 s	1343 s	+1.41%	1107 s	1412 s	−21.60%
2	2444 s	2436 s	+0.33%	2214 s	2721 s	−18.63%
3	3575 s	3685 s	−2.99%	3231 s	4239 s	−23.78%
4	4796 s	4775 s	+0.44%	4428 s	5557 s	−20.32%
5	6041 s	5929 s	+1.89%	5354 s	6945 s	−22.91%
6	6973 s	7189 s	−3.00%	6488 s	8428 s	−23.02%

To sum up, our proposed solution allows us to estimate the makespan within a reasonable time frame with high precision. By correctly estimating the transfer times time, we are able to build a more precise schedule.

5.5. Multiple Arrivals

The results discussed above focused on the efficiency of the solution at each single job arrival. In this section, we evaluate how the schedule changes as new jobs arrive. We consider the same set of instances without the transport system (Section 5.2) and with the transport system (Section 5.3). We remove p% of the jobs, compute the initial makespan without these jobs, and uniformly spread the arrival of the removed jobs at random between zero and the initial makespan. For each arrival, we update the schedule and the makespan.

In order to evaluate the quality of the final makespan obtained in this way, we consider an *ideal* scheduler that knows all the jobs (the initial ones and the future ones, with their arrival times) in advance and computes a single optimize schedule; we call this scheduler *clairvoyant*.

Figure 8 shows the difference (in percentage) between the final makespan obtained after updating the initial makespan at every job arrival and the makespan obtained with the clairvoyant scheduler for different percentages of dynamic jobs (jobs removed from the instances).

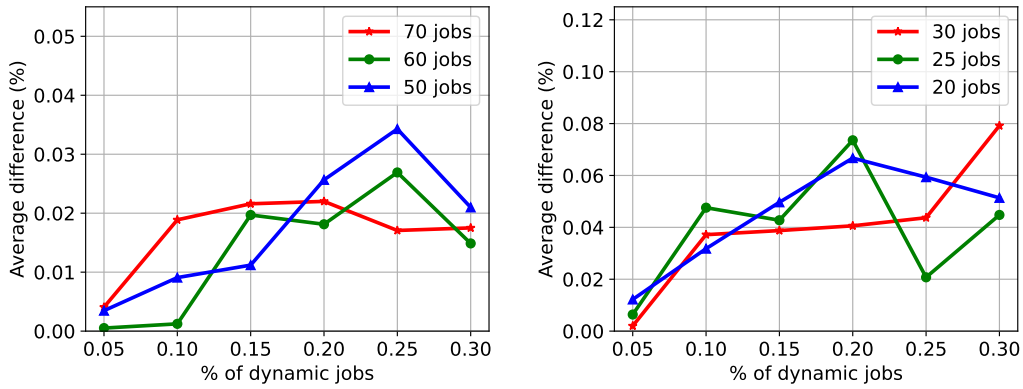


Figure 8. Difference (in percentage) between the makespan updated at every job arrival and the makespan from the ideal clairvoyant scheduler. Left and right figures refer to instances in [16] and the ICE instances.

In both cases (with and without the transport system), updating the schedule at every arrival provides a solution that is comparable to an ideal scheduler that has all the information in advance. The case with the transport system shows a higher deviation since the instances have a lower number of jobs. Overall, the total rescheduling strategy provides near-optimal solutions with a limited use of the processing resources.

6. Conclusions

Estimating the start and finish time of tasks belonging to jobs is fundamental for improving the efficiency of the manufacturing system. This is even more important in a fully automated environment, in which tasks are typically frequent but short. Integrating task duration with transport times enables more robust control of the overall process. To this aim, we designed a scheduler based on stochastic local search, which explores the solution space with a randomized approach.

Our proposed heuristic, SCHED-T, allows for building accurate and near-optimal schedules in a limited amount of time. This allows for efficiently managing the dynamic scenarios in which jobs continuously arrive or when unexpected events, such as machine breakdowns or maintenance activities, may occur.

In the future, we plan to extend our heuristic to include stochastically variable task-processing times and evaluate their impact on overall efficiency.

Author Contributions: Conceptualization, S.G. and D.C.; investigation, S.G. and D.C.; methodology, D.C.; software, S.G. and D.C.; writing—original draft preparation, S.G. and D.C.; writing—review and editing, S.S.; supervision, S.S. and F.F.; project administration, F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially funded by the Italian Ministry for University and Research (MUR) under the project *"Joint Communication and Sensing: CSI-Based Sensing for Future Wireless Networks (CSI-Future)"*, PRIN 2022-NRRP Mission 4, CUP D53D23016040001. The research was partially supported by the Italian National Group for Scientific Computation (GNCS-INDAM). Thisstudy was carried out within the PNRR research activities of the consortium iNEST (Interconnected North-Est Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5–D.D. 1058 23/06/2022, ECS_00000043). This manuscript reflects only the Authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

Terminology

Availability	Intervals of times when machines can be used (i.e., it is not under maintenance or booked for task execution).
Job	A set of operations or tasks that create a new product either directly from raw materials or components.
Machine	Equipment that processes, forms, or shapes raw materials or the output or other machines.
Makespan	Total execution time to process all the jobs currently submitted to the facility.
Precedence	Constraint on a task that can not be started before the completion of one or more tasks.
Processing time	Time required by a machine to perform a task.
Task	A unit of work performed on a single machine.
Transport times	Time required to move the partially processed materials from one machine to the next.

References

1. Wen, X.; Qian, Y.; Lian, X.; Zhang, Y.; Wang, H.; Li, H. Improved genetic algorithm based on multi-layer encoding approach for integrated process planning and scheduling problem. *Robot. Comput. Integr. Manuf.* **2023**, *84*, 102593. [CrossRef]
2. Grau, A.; Indri, M.; Bello, L.L.; Sauter, T. Industrial robotics in factory automation: From the early stage to the Internet of Things. In Proceedings of the IECON 2017—43rd Annual Conference of the IEEE Industrial Electronics Society, Beijing, China, 29 October–1 November 2017; pp. 6159–6164.
3. Luo, Q.; Deng, Q.; Xie, G.; Gong, G. A Pareto-based two-stage evolutionary algorithm for flexible job shop scheduling problem with worker cooperation flexibility. *Robot. Comput. Integr. Manuf.* **2023**, *82*, 102534. [CrossRef]
4. Chaudhry, I.A.; Khan, A.A. A research survey: Review of flexible job shop scheduling techniques. *Int. Trans. Oper. Res.* **2016**, *23*, 551–591. [CrossRef]
5. Zhang, J.; Ding, G.; Zou, Y.; Qin, S.; Fu, J. Review of job shop scheduling research and its new perspectives under Industry 4.0. *J. Intell. Manuf.* **2019**, *30*, 1809–1830. [CrossRef]
6. Pinedo, M.L. *Scheduling*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 29.
7. Kourtis, G.; Kavakli, E.; Sakellariou, R. A rule-based approach founded on description logics for Industry 4.0 smart factories. *IEEE Trans. Ind. Informatics* **2019**, *15*, 4888–4899. [CrossRef]
8. Zhang, G.; Sun, J.; Liu, X.; Wang, G.; Yang, Y. Solving flexible job shop scheduling problems with transportation time based on improved genetic algorithm. *Math. Biosci. Eng.* **2019**, *16*, 1334–1347. [CrossRef]
9. Yao, Y.J.; Liu, Q.H.; Li, X.Y.; Gao, L. A novel MILP model for job shop scheduling problem with mobile robots. *Robot. Comput. Integr. Manuf.* **2023**, *81*, 102506. [CrossRef]
10. Li, Y.; Gu, W.; Yuan, M.; Tang, Y. Real-time data-driven dynamic scheduling for flexible job shop with insufficient transportation resources using hybrid deep Q network. *Robot. Comput. Integr. Manuf.* **2022**, *74*, 102283. [CrossRef]
11. Pavlov, A.A.; Misura, E.B.; Melnikov, O.V.; Mukha, I.P. NP-Hard Scheduling Problems in Planning Process Automation in Discrete Systems of Certain Classes. In *Advances in Computer Science for Engineering and Education*; Hu, Z., Petoukhov, S., Dychka, I., He, M., Eds.; Springer: Cham, Switzerland, 2019; pp. 429–436.
12. Lunardi, W.T.; Birgin, E.G.; Ronconi, D.P.; Voos, H. Metaheuristics for the online printing shop scheduling problem. *Eur. J. Oper. Res.* **2021**, *293*, 419–441. [CrossRef]
13. Hoos, H.H.; Stützle, T. *Stochastic Local Search: Foundations and Applications*; Elsevier: Amsterdam, The Netherlands, 2004.
14. Wang, Z.; Zhang, J.; Yang, S. An improved particle swarm optimization algorithm for dynamic job shop scheduling problems with random job arrivals. *Swarm Evol. Comput.* **2019**, *51*, 100594. [CrossRef]
15. Baykasoğlu, A.; Madenoğlu, F.S.; Hamzadayı, A. Greedy randomized adaptive search for dynamic flexible job-shop scheduling. *J. Manuf. Syst.* **2020**, *56*, 425–451. [CrossRef]
16. Lunardi, W.T.; Birgin, E.G.; Laborie, P.; Ronconi, D.P.; Voos, H. Mixed Integer linear programming and constraint programming models for the online printing shop scheduling problem. *Comput. Oper. Res.* **2020**, *123*, 105020. [CrossRef]
17. Gaiardelli, S.; Carra, D.; Spellini, S.; Fummi, F. On the Impact of Transport Times in Flexible Job Shop Scheduling Problems. In Proceedings of the 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation, Stuttgart, Germany, 6–9 September 2022; pp. 1–8.
18. Industrial Computer Engineering (ICE) Lab. Available online: https://www.icelab.di.univr.it/ (accessed on 1 June 2023).
19. Li, X.; Guo, X.; Tang, H.; Wu, R.; Wang, L.; Pang, S.; Liu, Z.; Xu, W.; Li, X. Survey of integrated flexible job shop scheduling problems. *Comput. Ind. Eng.* **2022**, *174*, 108786. [CrossRef]
20. Naderi, B.; Zandieh, M.; Balagh, A.K.G.; Roshanaei, V. An improved simulated annealing for hybrid flowshops with sequence-dependent setup and transportation times to minimize total completion time and total tardiness. *Expert Syst. Appl.* **2009**, *36*, 9625–9633. [CrossRef]
21. Sun, Y.; Chung, S.H.; Wen, X.; Ma, H.L. Novel robotic job-shop scheduling models with deadlock and robot movement considerations. *Transp. Res. Part E Logist. Transp. Rev.* **2021**, *149*, 102273. [CrossRef]
22. Zhang, X.-j.; Sang, H.-y.; Li, J.-q; Han, Y.-y; Duan, P. An effective multi-AGVs dispatching method applied to matrix manufacturing workshop. *Comput. Ind. Eng.* **2022**, *163*, 107791. [CrossRef]
23. Lu, J.; Ren, C.; Shao, Y.; Zhu, J.; Lu, X. An automated guided vehicle conflict-free scheduling approach considering assignment rules in a robotic mobile fulfillment system. *Comput. Ind. Eng.* **2023**, *176*, 108932. [CrossRef]
24. Li, X.; Xing, K. Iterative Widen Heuristic Beam Search Algorithm for Scheduling Problem of Flexible Assembly Systems. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7348–7358. [CrossRef]
25. Kleywegt, A.J.; Shapiro, A.; Homem-de Mello, T. The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* **2002**, *12*, 479–502. [CrossRef]
26. Mitzenmacher, M. The power of two choices in randomized load balancing. *IEEE Trans. Parallel Distrib. Syst.* **2001**, *12*, 1094–1104. [CrossRef]
27. Yang, J.; Wang, Y.; Wang, Z. Efficient Modeling of Random Sampling-Based LRU. In Proceedings of the 50th International Conference on Parallel Processing, Lemont, IL, USA, 9–12 August 2021; pp. 1–11.
28. Furnas, G.W. Generalized fisheye views. *ACM Sigchi Bull.* **1986**, *17*, 16–23. [CrossRef]

29. Pei, G.; Gerla, M.; Chen, T.W. Fisheye state routing: A routing scheme for ad hoc wireless networks. In Proceedings of the 2000 IEEE International Conference on Communications. ICC 2000. Global Convergence Through Communications, Conference Record, New Orleans, LA, USA, 18–22 June 2000; Volume 1, pp. 70–74.
30. Chen, J.; Chen, K.; Wu, J.; Chen, C. A study of the flexible job shop scheduling problem with parallel machines and reentrant process. *Int. J. Adv. Manuf. Technol.* **2008**, *39*, 344–354. [CrossRef]
31. FJS Instance Generator. Available online: https://github.com/willtl/online-printing-shop (accessed on 1 June 2023).
32. Zeng, C.; Tang, J.; Yan, C. Scheduling of no buffer job shop cells with blocking constraints and automated guided vehicles. *Appl. Soft Comput.* **2014**, *24*, 1033–1046. [CrossRef]
33. Gaiardelli, S.; Spellini, S.; Panato, M.; Lora, M.; Fummi, F. A Software Architecture to Control Service-Oriented Manufacturing Systems. In Proceedings of the 2022 Design, Automation & Test in Europe Conference & Exhibition (DATE), Antwerp, Belgium, 14–23 March 2022; pp. 1–4.
34. Beregi, R.; Pedone, G.; Háy, B.; Váncza, J. Manufacturing Execution System Integration through the Standardization of a Common Service Model for Cyber-Physical Production Systems. *Appl. Sci.* **2021**, *11*, 7581. [CrossRef]

applied sciences

MDPI

Article

SysML4GDPSim: A SysML Profile for Modeling Geometric Deviation Propagation in Multistage Manufacturing Systems Simulation

Sergio Benavent-Nácher *, Pedro Rosado Castellano and Fernando Romero Subirón

Department of Industrial Systems Engineering and Design, Universitat Jaume I. Av. Vicent Sos Baynat, s/n, 12006 Castellón de la Plana, Spain; rosado@uji.es (P.R.C.); fromero@uji.es (F.R.S.)
* Correspondence: benavens@uji.es

Abstract: In recent years, paradigms like production quality or zero-defect manufacturing have emerged, highlighting the need to improve quality and reduce waste in manufacturing systems. Although quality can be analyzed from various points of view during different stages of a manufacturing system's lifecycle, this research focuses on a multidomain simulation model definition oriented toward the analysis of productivity and geometric quality during early design stages. To avoid inconsistencies, the authors explored the definition of descriptive models using system modeling language (SysML) profiles that capture domain-specific semantics defining object constraint language (OCL) rules, facilitating the assurance of model completeness and consistency regarding this specific knowledge. This paper presents a SysML profile for the simulation of geometric deviation propagation in multistage manufacturing systems (SysML4GDPSim), containing the concepts for the analysis of two data flows: (a) coupled discrete behavior simulation characteristic of manufacturing systems defined using discrete events simulation (DEVS) formalism; and (b) geometric deviation propagation through the system based on the geometrical modeling of artifacts using concepts from the topologically and technologically related surfaces (TTRS) theory. Consistency checking for this type of multidomain simulation model and the adoption of TTRS for the mathematical analysis of geometric deviations are the main contributions of this work, oriented towards facilitating the collaboration between design and analysis experts in the manufacturing domain. Finally, a case study shows the application of the proposed profile for the simulation model of an assembling line, including the model's transformation to Modelica and some experimental results of this type of analysis.

Keywords: MBSE; SysML; model consistency; manufacturing system simulation; geometric deviation analysis; TTRS

1. Introduction

In recent years, manufacturing systems design, in alignment with initiatives like Industry 4.0, has promoted different operational paradigms like production quality [1] or zero-defect manufacturing [2] to address quality improvement and waste reduction. The production quality bases propose an integration of quality, productivity, and maintenance evaluation, an orientation adopted in this paper to analyze, jointly, production and quality. Among product quality characteristics, this work focuses on the geometric quality of both manufacturing resources and processed products, combined with productivity analysis using multidomain simulation model during early design stages.

During the design and development of a system, and especially during the verification and validation stages of the adopted solution principles, the development of mathematical models enables the analysis of the referent system's behavior and performance. In the case of simple systems, an analytical solution can usually be obtained, but complex systems analysis, such as of manufacturing systems, usually requires the development

of simulation software systems to support their high uncertainty and/or nonlinearity, combining quantitative/qualitative features and continuous/discrete behaviors [3]. Although specific tools can be used for the simulation of certain domains (including the digital manufacturing tools focused on the manufacturing systems domain), the adoption of generic and/or standardized tools and languages (e.g., Modelica) facilitates the modeling of multi-domain simulation systems during initial design stages, an orientation aligned with model-based system engineering principles that promote model integration and consistency assurance [4]. Moreover, simulation models have been a widely explored solution, as evidenced in the increase in the research works in this field [5].

In this context, the capability of generic simulation languages (e.g., Modelica) to verify and validate simulation models is limited to assure the syntactic correctness of the models; for example, checking that the set of variables and mathematical relationships is sufficient to obtain a solution. An interesting alternative to overcome these limitations is the adoption of domain-specific modeling languages (DSML) [6], which include the specific semantics of a studied domain, which would enable the model consistency validation based on these specific semantics. Generic descriptive modeling languages such as SysML [7] support the development of DSML, defining specific profiles that extend the language by adding stereotypes and detailing the semantics through OCL expressions [8]. OCL defines a set of functions to define queries on the model without side effects. The results of these queries are processed by logical functions to check the accomplishment of specific conditions (invariants). Their implementation as part of stereotypes definition at the profile level enables the evaluation of stereotyped model constructions and the detection of inconsistencies. In addition to the creation of profiles, the use of SysML and the adoption of a systems modeling approach also allows for the linking of the simulation models with other models of the referent system (e.g., specification or design models), checking consistency between models to support the collaboration between design and analysis experts. The design and validation of the simulation model using SysML have been studied in previous works, such as [9]. This work also addresses the subsequent model transformation to executable models using languages like Modelica by applying the SysML4Modelica profile [10]. A similar transformation mechanism is proposed in [11], also based on the application of an SysML profile, to enable an automated transformation. Other authors have explored the use of SysML activity diagrams to describe workflows for connecting simulation and design models and facilitating collaboration between system architects and analysis experts [12,13].

In the manufacturing domain, there is a lack of developments adopting this approach, and it is hard to find works focused on geometric deviations and their propagation in multi-stage systems. In [14], SysML is used for the preliminary design of a multidomain simulation system that integrates the analysis of productivity and geometric quality to study how different control strategies influence performance measures. However, all the previously mentioned works use SysML as a modeling language only to define the descriptive models facilitating communication between engineers.

The present paper explores the SysML's capability to define DSMLs as SysML profiles, including domain-specific semantics to develop simulation software systems. Specifically, this research proposes a language with which to define simulation models for manufacturing systems, focused on a joint analysis of productivity and geometric quality performance. The proposed SysML profile (geometric deviation propagation Simulation—SysML4GDPSim) integrates the discrete behavioral aspects of a manufacturing system [15] and topologically and technologically related surfaces (TTRS) [16] concepts to model the geometrical quality characteristics of the products.

The content and structure of this document is summarized as follows. Section 2 briefly discusses the complexity and main characteristics of the supersystem to be analyzed (including the interactions between the product, the process, and the resources), the simulation systems for their analysis, and the deviated geometry modeling to support quality analysis. Section 3 is dedicated to the proposed SysML profile, presenting the metamodel

description, the implementation of the profile, and its application during library model definition. Section 4 presents a case study to exemplify the application of the proposal and validate its suitability in the context of manufacturing simulation. Finally, Section 5 presents a final discussion and Section 6 summarizes some conclusions and future works.

2. Modeling and Simulation of Product–Process–Resource Systems

A simulation system model is the model of (executable) software that allows for the performance analysis of a referent system in response to different conditions derived from the properties of the referent system and its environment. Simulation models are usually developed for complex referent systems. In the case of this research, the referent system is a system of systems, that is, a system that is made up of systems, according to the definition of [17]. This concept is also named a "supersystem" in other references, like [18]. Specifically, the term supersystem is adopted in [19] to express a temporal construction to describe the production lifecycle phase of a product, supporting links with other elements like manufacturing resources. In this paper, this orientation is adopted to define a system composed of resources (the manufacturing assets), the processes executed by these resources, and the manufactured products. All these elements together compound the PPR system that has been studied in previous works on manufacturing domain modeling, like [20]. In order to simulate PPR systems, it is necessary to acquire deep knowledge of the referent system and its representation in the simulation system, including all the aspects that have influence in the proposed analysis.

2.1. PPR Systems Basis

This research is focused on the study of discrete manufacturing systems that support multistage processes, with reconfigurable and automated resources. Therefore, the manufacturing system is composed of configurable resources supporting various manufacturing processes with small changes. In this context, a resource is any mechatronic system that plays a role in the production of goods (products).

Resources in a manufacturing system are structured in a multilevel hierarchy in which simple resources are part of complex ones. In fact, the manufacturing system as a whole is a complex resource. The modeling of the manufacturing system requires the establishment of an adequate decomposition level to support the analysis goals. Considering the interests of this research, this atomic level, according to the classification proposed in [21], is the workstation level. A workstation is a mechatronic system with a behavior that directly influences the final geometry of the product. In a broader sense, the term resource is also applied to simpler elements such as fixtures, tools, etc., but in this work, these elements are treated as physical artifacts without their own behavior, and they will be considered for the geometrical modeling of process assemblies without granting them the category of resource.

Moreover, resources in a manufacturing system can be classified into processing and control resources. The processing resources change the properties of the processed batches, while control resources merely monitor the batches and make decisions to improve their properties. Processing resources can be classified into: (a) logistical resources that modify the batch properties as a unit; and (b) transformer resources that modify the properties (especially geometric features) of individual products of the batch.

The transformer resources (mainly assembling and machining workstations) can support several configurations to execute alternative manufacturing processes. In this case, it is necessary to consider a setup stage (before the processing) to introduce the necessary changes (e.g., changes in fixtures or tools, loading an NC code, etc.).

Another basic element of the PPR system is the product to be manufactured. The product type can have representations from different viewpoints and degrees of complexity. In this research, products are represented as entities with the necessary attributes to describe the geometric deviations from the product specification. These products are grouped in

batches which have their own properties, such as size, start time of manufacturing, waiting or storage times, processing duration, etc.

The third element of the PPR system is the manufacturing process. Each product type has a generic process plan, defining the necessary manufacturing processes, and a native process plan which details the specific resource assigned for each process stage. Although the process plan can be established with different levels of detail, in this research, the selected atomic level is the subphase, which aggregates all the operations executed in a machine using the same clamping (part-fixture assembly).

Each subphase of the native process plan establishes an interaction between the product (in a certain state) and specific resources of the manufacturing system. This research is focused on the physical interactions between the product and the manufacturing system that produce the new geometric features of the product, as studied in [22]. Specifically, fixture–workpiece–machine interactions are studied in [23] in order to characterize the geometric deviations on these types of assembly. This type of construction is also promoted in [24], which is focused on inspection planning. In this paper, this assembly model is named "process assembly", and it is defined to support the analysis of the geometric deviations resulting from the execution of a certain manufacturing subphase. The propagation of these deviations to subsequent stages must be simulated in parallel to the material flow. In this way, the geometric quality of the final product and the indicators related to PPR system productivity can be jointly assessed, as promoted by the production quality paradigm.

2.2. Simulation System Modeling

As mentioned above, a simulation system is a soft system that emulates the behavior of a referent systems to analyze its response to certain conditions. These conditions are generally defined by a set of parameters whose values are assigned for each experiment in order to quantify the influence of conditions on the performance metrics. Simulation models are built from a set of software objects that share data through connecting ports and perform calculations. These objects are permanent entities, called "resources" in simulation [25], that have existence throughout simulated time (i.e., their life is not finite). This characteristic is one of the main differences compared to flow units (also called "entities" in [25]), which are elements that flow through the system and have a finite life (they are created and destroyed at certain moments). In PPR system simulation, permanent entities primarily emulate manufacturing system resources, while flow units represent individual products, product batches, manufacturing orders, etc.

Moreover, permanent entities can include the definition of a certain behavior, expressed in the form of algorithms and mathematical equations. In the scope of this research, different behaviors are identified corresponding to the typical functional units that participate in manufacturing the system simulation, supporting both the simulation of the referent system and its environment and other functionalities specific to the proposed analysis. These main functional units address issues such as the following:

- Generation of flow units based on a defined schedule or behavior: Objects with this functionality represent the beginning of the flow, and they have ports with which to send the data of the generated flow units.
- Processing flow units: Objects with this functionality have at least two ports with which to send and receive flow units. Moreover, a certain behavior can introduce changes in the data of the flow units. In the simulation of PPR systems, this function is mainly linked to the emulation of transforming resources (e.g., assembly or machining workstations) and logistic resources (stores or transportation resources).
- Destruction of flow units: Objects with this functionality represents the end point of the flow, where flow units are destroyed. These objects, generally named sinks, must have at least one port to receive flow units.
- Monitoring data and decisions making: Objects with this functionality have ports with which to receive data about key performance information. If they process data to make control decisions, they have ports with which to send data about these

decisions. In a PPR systems simulation, these objects are linked to the control resources representation.

All these types of functional units are characterized by a discrete behavior described by the principles of the discrete event system (DEVS) formalism [26]. According to DEVS, the behavior of a complex system (coupled behavior) emerges from the behavior of its simplest components and their interactions. In the same way, the complex behavior simulation of the PPR system emerges from the simpler behavior simulation of its components. Although there are various ways to represent this type of discrete behavior, one of the most common alternatives is the use of state machines, where a set of system states are defined and linked by transitions that are triggered by certain events. In addition, detailed behaviors can be defined when a component enters, exits, or stays in each state.

2.3. Representation and Calculation of Geometric Deviations

All the components of the process assembly (products, fixtures, tools, etc.) are imperfect realities whose geometric elements have deviated from the nominal ones specified in their designs. The analysis of these geometric deviations is mainly supported by the study of the interactions between product and resource devices during each single process. The representation of geometric deviations has been addressed in different specific and standardized languages (ISO, Ansi, etc.) to represent the maximum allowable deviations, specified as tolerance zones. Beyond tolerance representation, different mathematical models have been also proposed to deal with tolerance analysis, such as the Jacobian–Torsor model [27], the small displacement torsor [28], T-map [29], or matrix transformation [30], among others summarized in [31].

SysML models have also been used for tolerance representation, although they are not as widespread. In [32], TTRS concepts are adopted to describe the tolerance specifications with a SysML model, but it does not include the mathematical constructions to support a quantitative analysis. To overcome this limitation, [33] presents the SysML for tolerance analysis (SysML4TA) profile, which defines a DSML to support the TTRS-based mathematical characterization of geometric surfaces and their intra- and inter-part relationships. The work in question also details the mathematical operations necessary to compute deviations and verify the specification fulfilment. In [34], this orientation is adopted for developing a Modelica library for the geometric analysis of mechanical assemblies or parts. These works also highlight the compatibility of TTRS with geometric dimensioning and tolerancing (GD&T) standards [35], allowing the specification of parts to be addressed and unambiguously transferring the dimensioning scheme to a mathematical formulation suitable for simulation.

This paper adopts the principles and concepts formulated in a SysML4TA profile [33] to support the geometric modeling of the process assemblies necessary to solve each single-stage problem (subphase analysis). Additionally, the simulation system must have sufficient elements to transmit the resulting product geometric data to subsequent stages, as commented on in Section 3.

3. Proposed SysML4GDPSim Profile

This section presents the SysML4GDPSim profile, first describing the concepts and relationships considered in the metamodel of the developed language, depicted using some conceptual diagrams with UML notation. The Section 3.2 introduces some brief notes on the profile implementation, whose stereotypes are described in detail in Appendix A. Finally, the Section 3.3 briefly describes the SysML libraries developed to facilitate the modeling of the simulation systems under study and their transformation to Modelica executable models.

3.1. Metamodel Description

Assuming many of the concepts and bases discussed in Section 2, this subsection presents the metamodel description, commenting on the concepts related to the basic

structure of the simulation system, the representation of the manufacturing system and the products, and finally, the artifacts modeling for the analysis of geometric quality.

3.1.1. Simulation System Basic Structure

According to the bases commented on in Section 2.2, Figure 1 represents the basic components of the simulation system (*SimulationSystem4MS*), which simulates the behavior of a manufacturing system (*MS_sim*) in a specific environment (defined by the rest of the elements of the simulation system). Before the description of the *MS_sim*, which constitutes the central component of the simulation system, the different concepts that define the simulation environment supporting specific functionalities are presented below:

- *Configurator* identifies the component that groups all the characteristic parameters and offers them to the other components (*MS_sim*, *InputGenerator*, *OutputCollector*, etc.) through the reference relationships shown in Figure 1.
- *InputGenerator* represents components that create the *FlowUnits*, representing material supply (inputs for the *MS_sim*). An *InputGenerator* is the initial point of the data flow that represents the logistics flow.
- *OutputCollector* represents components that destroy *FlowUnits*; so, the *OutputCollector* is a sink or end point of the data flow that represents the materials flow.
- *ResultsManager* represents components that collect and process the information of simulation executions to compute the desired performance metrics.

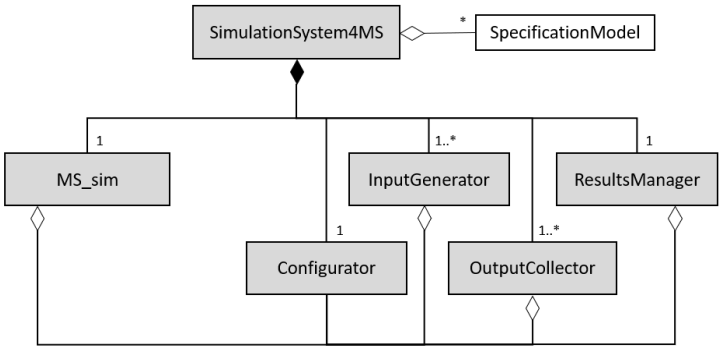


Figure 1. Abstract syntax of simulation systems. The notations "1" and "1..*" are multiplicities ("only one" and "one or more" respectively).

As represented in Figure 1, a *SimulationSystem4MS* is composed of at least one *MS_sim*, a *Configurator*, a *ResultsManager*, and at least one *InputGenerator* and *OutputCollector*. Moreover, the *SimulationSystem4MS* can have an aggregation relationship with the referent system specification to represent the necessary collaboration between the design and analysis tasks and facilitate the assurance of consistency between both models.

3.1.2. The Manufacturing System in the Simulation System

As mentioned, the *MS_sim* identifies the central component of the simulation system that emulates the behavior of the manufacturing system, transforming the input materials to obtain the output products with different geometrical properties. As represented in Figure 2, the *MS_sim* is a complex transformer resource (*TransformResource_sim*) because material units (a) flow through the resource (*ProcessingResource_sim* specialization) and (b) are transformed (i.e., there are changes in their physical characteristics).

A manufacturing system (and its simulation system) involves different types of resources; so, the metamodel includes different concepts as represented in Figure 2 and briefly described below:

- *ManufResource_sim* (abstract) represents any resource of the simulated manufacturing system, i.e., both processor and control resources, *ProcessingResource_sim* and *ControlResource_sim*, respectively.

- *ProcessingResource_sim* is a specialization of the *ManufResource_sim* to emulate a resource through which batches of products flow, executing processes that modify some properties of the batches (e.g., location, flow times, etc.) or the contained products (e.g., its state or its geometric characteristics). A *ProcessingResource_sim* can be composed of other *ManufResource_sim* (i.e., both processing and control resources) and must have at least two *FU_Ports* to send and receive *FlowUnits*.
- *TransformResource_sim* is a specialization of the *ProcessingResource_sim* that emulates a transformer resource that modifies the properties of the *FlowUnits*; so, incoming *FlowUnit* type (input product) is different from the outgoing *FlowUnit* type (processed product).
- *LogisticResource_sim* is a specialization of the *ProcessingResource_sim* that emulates a logistical resource, participating in the materials flow without modifying *FlowUnit* properties so the type of incoming and outgoing *FlowUnits* is the same.
- *ControlResource_sim* is a specialization of the *ManufResource_sim* that emulates a control resource. *ControlResource_sim* does not participate in the logistics flow but exchanges data to support monitoring and decision making.
- *FlowUnit* represents a flowing unit, that is, a batch of products. Therefore, it will be composed of one or more *Product_sim*, which represents each material or product unit.
- *DataPort* represents a generic port as an interaction point to exchange data through connections that specify relationships between components of the simulation system.
- *FU_Port* is a specialization of the *DataPort* that supports the transfer of data on the simulated *FlowUnits* (product batches), supporting the logistics flow. As represented via a dependency relationship in Figure 2, this type of port must be typed by a *FlowUnit*.
- *C_Port* is a specialization of the *DataPort* that supports data transfer in the form of communication between resources to synchronize tasks and behaviors.
- *Product_sim* represents a simulated product unit characterized by properties that include the deviations for its key geometric characteristics. As a part of the *FlowUnit*, it supports geometric deviation propagation through the simulated manufacturing stages.

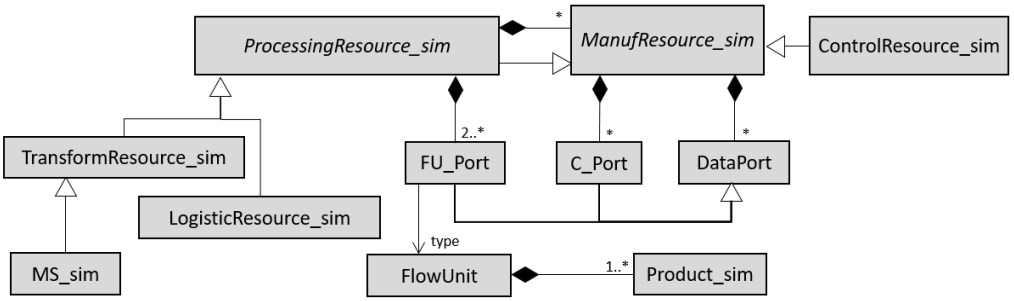


Figure 2. Abstract syntax of simulated manufacturing systems.

3.1.3. Behavior Modeling in the Simulation System

In addition to identifying previously commented concepts, it is important to clarify the modeling of the behavior of some of them (*MS_sim*, *InputGenerator*, *OutputCollector*, etc.). Although various approaches to modeling discrete behaviors could be valid, with each one focused on certain aspects (events, activities, processes, etc.), in this proposal, the DEVS formalism is adopted and described using state machines. According to DEVS formalism, a behavior can be defined by coupling simpler behaviors; so, the behavior of composite elements emerges from their behavioral components and their interactions. In this metamodel, the *BehavioralElement_sim* concept represents the generalization of any component with behavior, differentiating between atomic and coupled behavior with the *isAtomic* Boolean property (Figure 3). It is established that atomic elements (when atomic = true) must have an explicitly defined behavior and cannot be composed by parts of *BehavioralElement_sim* type (i.e., parts with their own behavior). This atomic behavior must be defined as a

state machine, represented by *STM_MainBehavior*. In contrast, composite elements (when atomic = false) have no explicitly defined behavior, but they must be composed of other *BehavioralElement_sim* (atomic or not). Both cases are depicted as composite relationships in Figure 3, buy they are mutually exclusive.

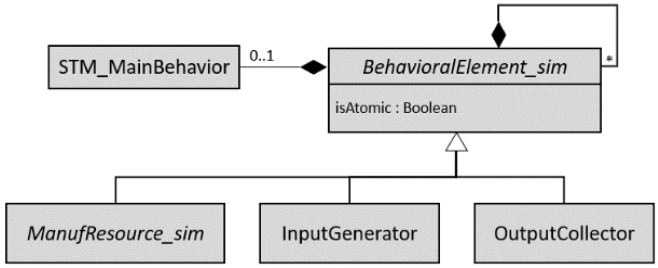


Figure 3. Abstract syntax of elements with behavior.

3.1.4. The Product in the Simulation System

Another fundamental element in the simulation of PPR systems is the processed product (*Product_sim*). As shown in Figure 4, the *Product_sim* has a reference relationship with product specification (*ProductSpecif_data*), which includes at least two descriptions of its possible transformation states (*StateSpecif*) corresponding to the initial and final product states.

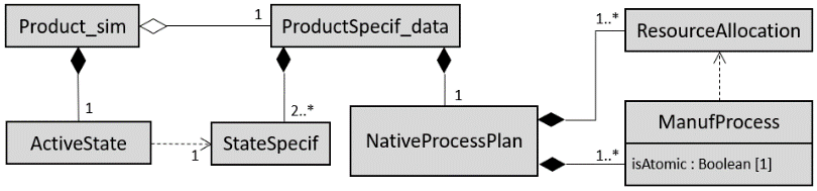


Figure 4. Abstract syntax of emulated product and its specification.

Moreover, *ProductSpecif_data* is composed by a native process plan (*NativeProcessPlan*) described in an activity diagram. Each *NativeProcessPlan* is composed of one or more actions (*ManufProcess*) that represent manufacturing processes. Atomic *ManufProcess* (identified by the Boolean attribute *isAtomic*) correspond to subphases. To support the resource assignment, the *NativeProcessPlan* has *ResourceAllocations* to define a specific resource for each atomic *ManufProcess*, as shown by the dependence relationship.

On the other hand, *Product_sim*, which represents each product unit, is characterized with some basic properties, including the identification of the current state (*ActiveState*), which points to one of the *StateSpecif*, as shown by the dependence relationship.

3.1.5. Product–Resource Geometric Interaction in the Simulation System

One of the fundamental aspects addressed in the proposed simulations is the analysis of the geometric quality of the emulated products, an issue that is based on the geometrical modeling of products and resources and the physical interactions between them.

On the one hand, the geometric modeling of the products is addressed in Figure 5. The product specification (*ProductSpecif_data*) must include at least one *ProductArtifact* (specialization of the *Artifact* concept from SysML4TA) to define the nominal geometries and tolerances according to the SysML4TA profile constructions. *ProductDeviations* is defined as a dataset (vector or matrix) that stores the deviations of each of the key geometric features identified in the specification of a simulated product unit (*Product_sim*). *ProductDeviations* specializes the *DeviationVector*, a concept of the SysML4TA profile.

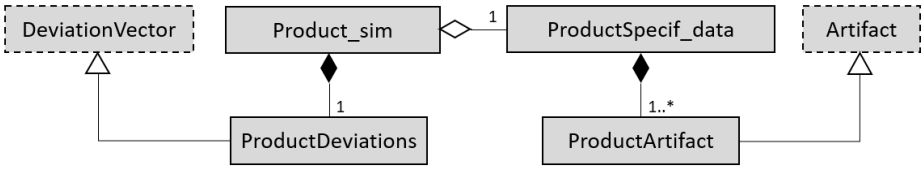


Figure 5. Abstract syntax of emulated products incorporating geometric deviations.

On the other hand, in a similar way, the geometric modeling of the transformer resources is addressed in Figure 6. A resource specification (*ManufResourceSpecif_data*) is composed by at least one *ConfiguredMachine*. Each *ConfiguredMachine*, which is a specialization of the *Artifact* concept from SysML4TA, describes the resource configuration for a subphase, including the geometric specification of the machine, fixtures and/or tools, and the assembly relationships between them. The particular deviations of each *TransformResourceSim* are defined by a *ResourceDeviations* (specialization of *DeviationVector* from the SysML4TA profile).

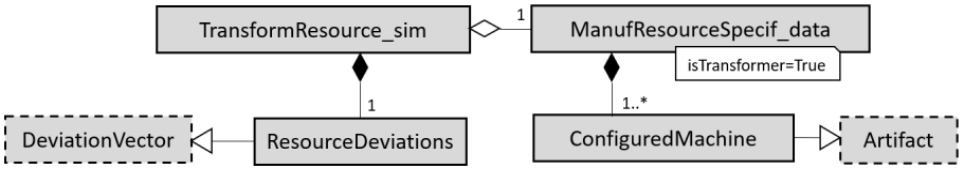


Figure 6. Abstract syntax of simulating resources with geometric deviations.

As shown in Figure 7, the *ProcessAssembly* is an artefact owned by an atomic *TransformResource_sim* which has references to at least one *ProductArtifact* and a *ConfiguredMachine* and includes the assembly relationships between these artifacts. Therefore, the *ProcessAssembly* involves the mathematical expressions necessary to obtain the quality characteristics of the resulting product in a single-stage problem. These results obtained in a workstation are transmitted to subsequent workstations to support the propagation of the in-process product geometric deviations, influencing other *ProcessAssembly* definitions until the final product quality is obtained.

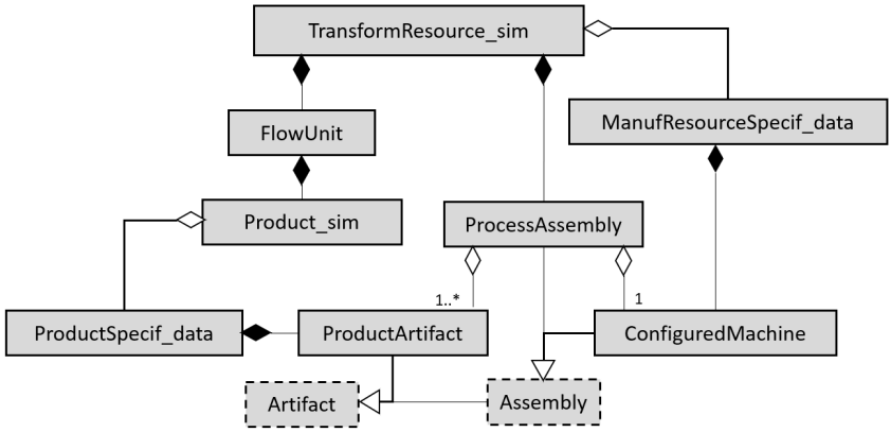


Figure 7. Abstract syntax of emulated process assembly.

3.2. SysML4GDPSim Profile

From the described metamodel (i.e., concepts and their relations and restrictions), the SysML4GDPSim profile has been developed. In order to define the profile, each concept of the metamodel has been defined as a stereotype, transferring part of the abstract syntax through extension relationships with UML metaclasses or specializations of the SysML profile, as shown in Figure 8. The rest of the relationships and semantic considerations have been included in the profile through the implementation of OCL rules for each

SysML4GDPSim stereotype. These OCL rules are identified in the description of the stereotypes included in Appendix A and exemplified in Appendix B.

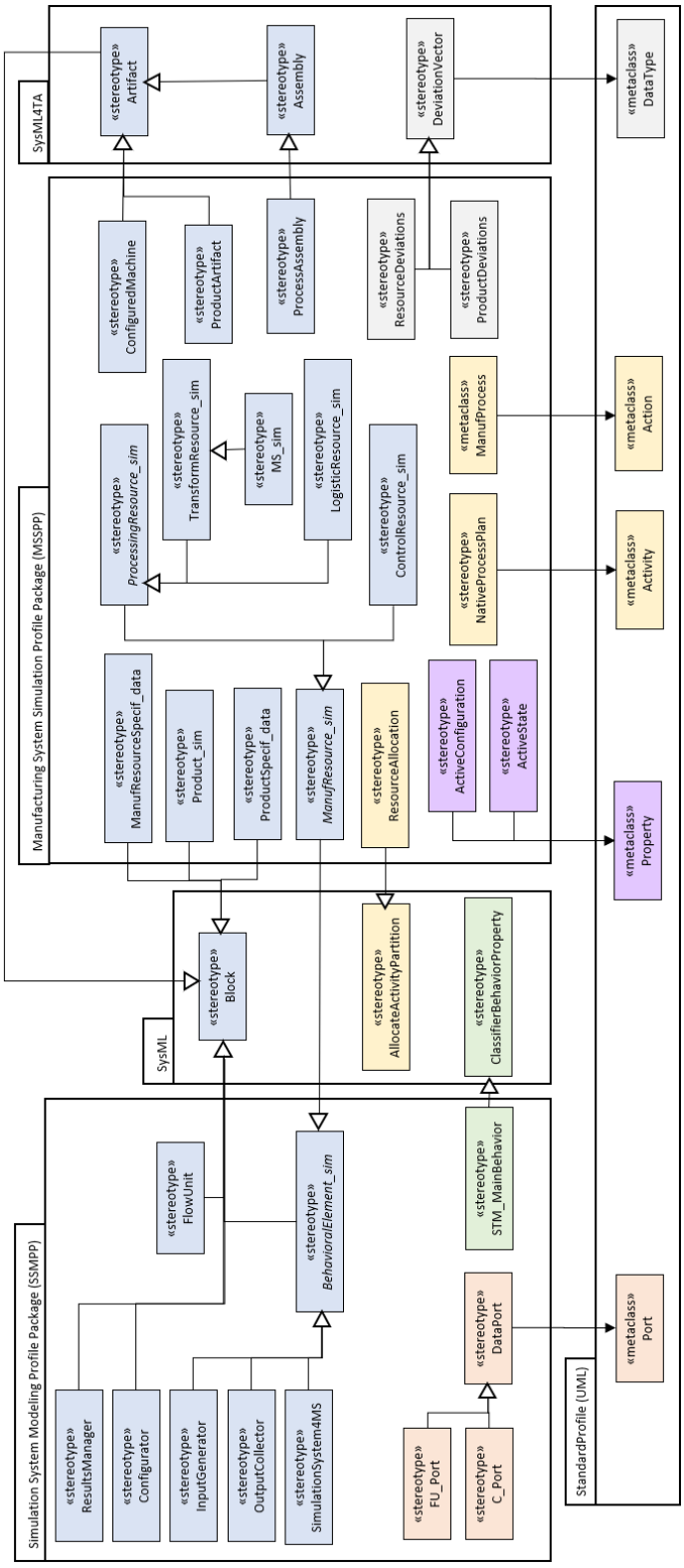


Figure 8. SysML4GDPSim profile package diagram. The color code groups stereotypes that extend the same UML metaclass, or specialize the same SysML stereotype.

It should be noted that in order to structure the content of this profile, two main packages have been proposed (Figure 8). The simulation system modeling profile package (SSMPP) includes the most general concepts of the simulation system; so, in future work, this package and its stereotypes can be used to develop other alternative simulation systems. The manufacturing system simulation profile package (MSSPP) includes specific stereotypes for the simulation of manufacturing systems, and in particular, for the analysis of productivity and geometric quality. As mentioned in the metamodel description and shown in Figure 8, some stereotypes are imported from the SysML4TA profile, which enables the nominal geometric specification of artifacts (i.e., product, fixtures, . . .) and assembly relationships in order to compute the geometric deviations on the process assemblies.

3.3. Developed Libraries

Following the SysML4GDPSim profile definition, some modeling libraries have been developed to facilitate user modeling tasks. Both the SysML4GDPSim profile edition and the libraries definition have been carried out in the Papyrus environment, one of the most widespread free environments for modeling with UML and SysML in academia. Figure 9 shows a screenshot of this modeling environment, showing an error identified in the defined model. During profile development, some inconsistencies are forcibly introduced to verify the correct functioning of the OCL rules. In this case, rule "C8" is defined to assure that a «SimulationSystem4MS» block has a part typed by a «MS_sim» Block. This rule is violated in the model (i.e., multiplicity must be "1", and it is defined as "1..*").

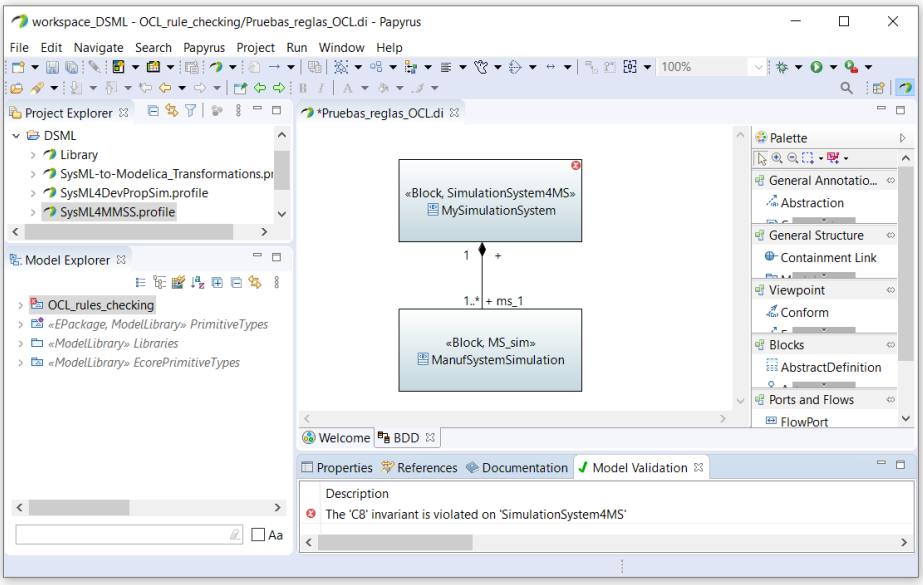


Figure 9. Papyrus environment and errors identified during model validation.

Additionally, equivalent libraries have been developed using Modelica. These Modelica libraries have the same package structure and elements as the SysML libraries, and, as explored in [11], they can be obtained via manual or automatic model transformation, which is beyond the scope of this paper. In this way, when a consistent simulation model is defined using the SysML libraries, it can be transformed into a Modelica model ready to be executed in an easy and quick way.

The package structure of these libraries, presented in detail in [11], is shown in Figure 10. As can be observed, there are three main packages with elements for (a) the analysis of material flow in multistage manufacturing systems; (b) the mathematical modeling of artifacts passed into TTRS concepts; and (c) the analysis of material flow and the propagation of geometric deviations in manufacturing systems.

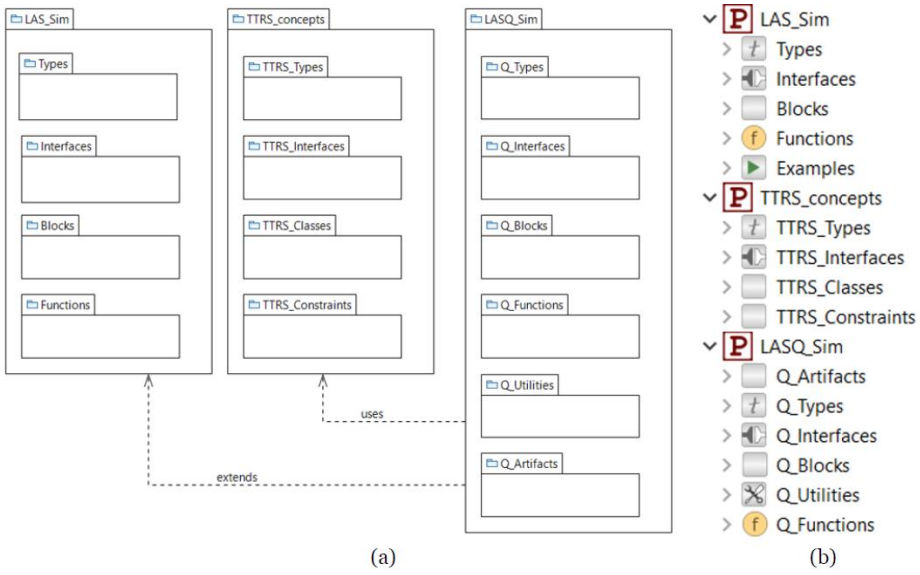


Figure 10. Libraries packages developed in SysML (**a**) and Modelica (**b**).

4. Case Study

This section presents a case study to exemplify the SysML modeling of a simulation system applying the SysML4GDPSim profile, as well as the description of some results obtained after its transformation and execution. It has been decided to limit the case to a 2D analysis of a multistage assembly process using a simple isostatic localization pattern for each assembly stage. Moreover, the methodology proposed in [11] is applied, whose procedure encompasses the tasks of modeling the referent system and modeling the simulation system (including the geometric modeling of artifacts based on TTRS), the definition of experiments, and their to-be-executed transformation to Modelica.

4.1. Referent System Description

This case is focused on the analysis of the manufacturing of the product depicted in Figure 11, a frame manufactured from the union of four metal parts. Each part has one or more holes where the axles of a series of gears are inserted. The manufacturing system (MyLAS) designed is made of three welding workstations, each one composed of a welding station and two stores for the incoming parts, as shown in the block definition diagram (BDD) presented in Figure 12. The resource connections, shown in the internal block diagram (IBD) of Figure 13, supports a linear flow of the product batches.

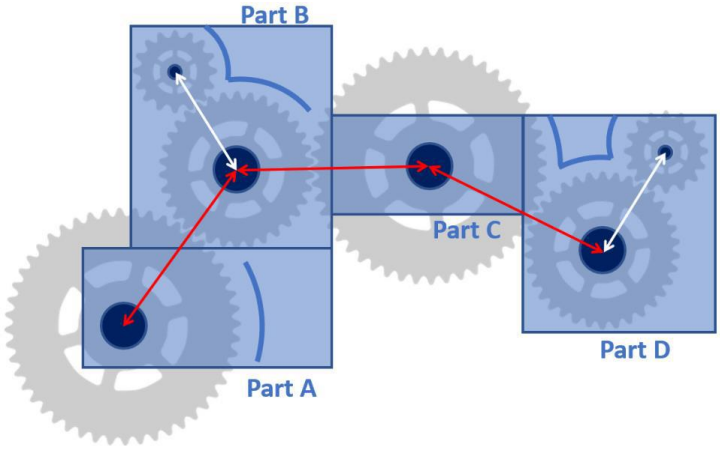


Figure 11. Graphic representation of the analyzed product.

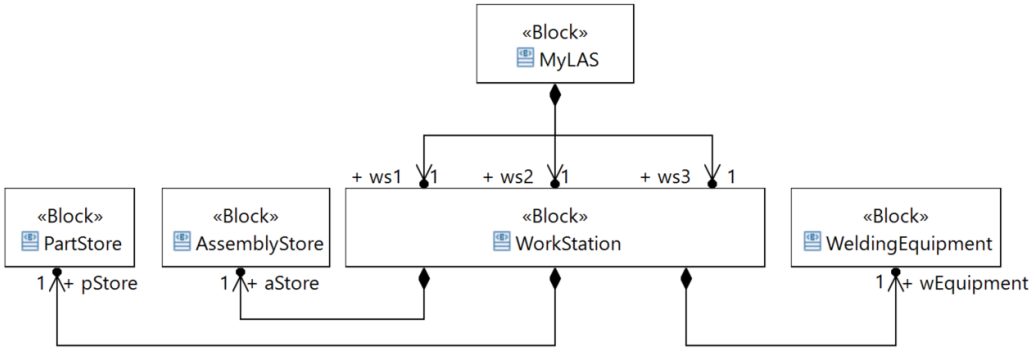


Figure 12. BDD to represent the structure of the linear assembly system.

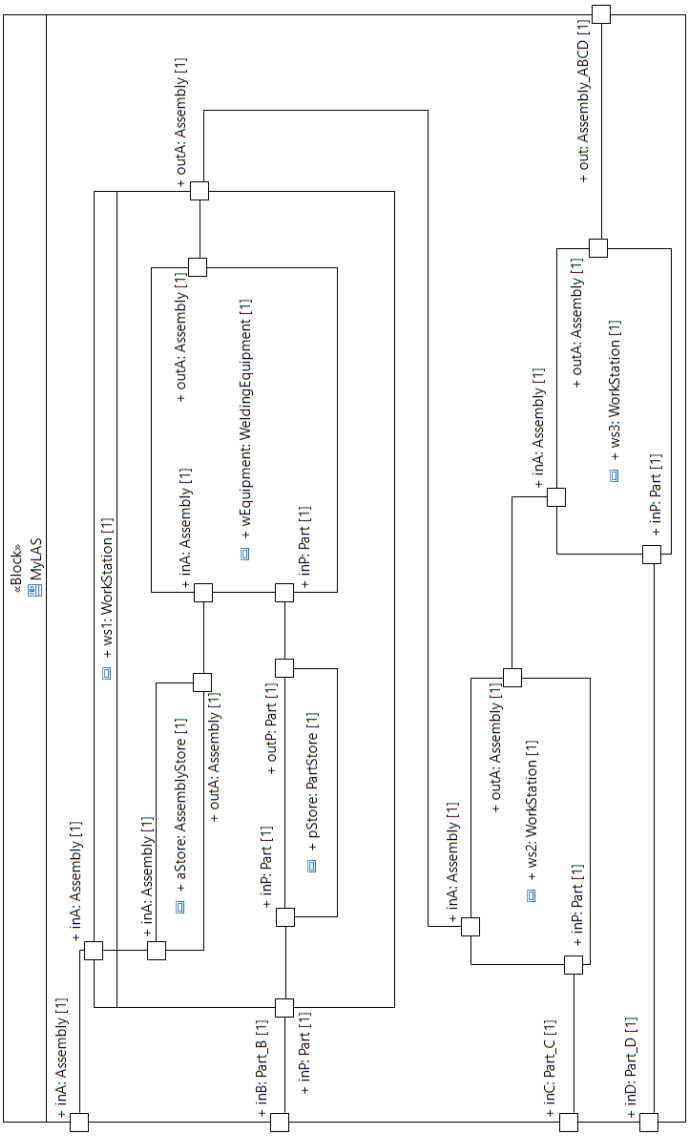


Figure 13. Internal structure of the specified assembly line (MyLAS block).

The main goal of the proposed study is the modeling of a simulation system to analyze its productivity and product quality performance. The native process plan for the assembly process is represented in Figure 14 by means of an activity diagram where the three assembling stages and intermediate storage stages are shown, each one assigned to a specific workstation of the manufacturing system.

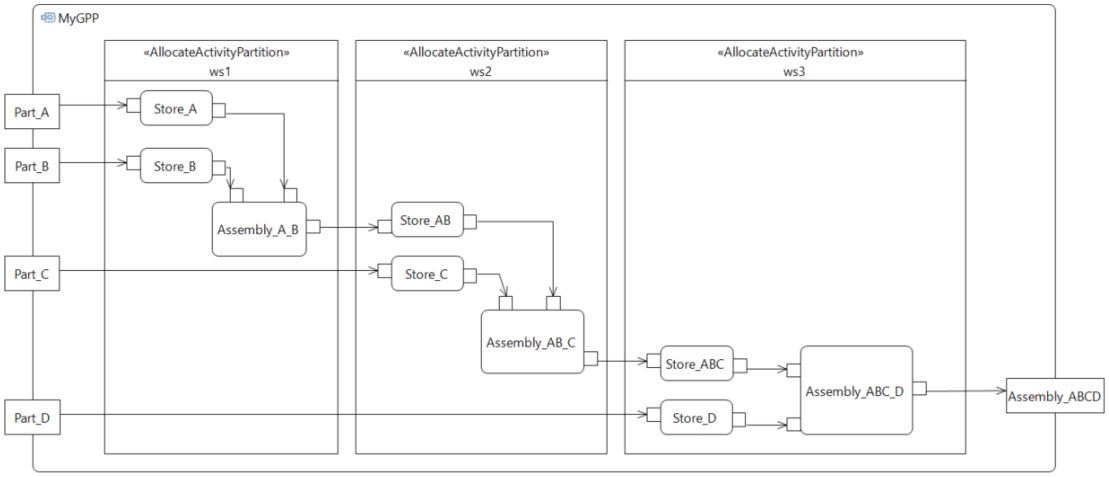


Figure 14. Activity diagram of the generic process plan specification (MyGPP).

In Figure 15, the process assemblies for each process stage are shown. As can be observed, for each process stage, the incoming parts or subassembly are located on the fixture using the same isostatic localization pattern: a four-way locator (p1, p3) and a two-way-locator (p2, p4).

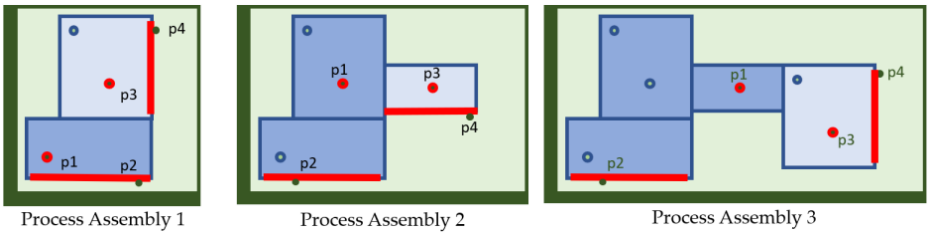


Figure 15. Process assembly corresponding to each of the stages considered.

4.2. Simulation System Modeling

Figure 16 shows a BDD of the simulation system (MySimulationModel) proposed for analyzing the performance of the manufacturing system described.

MySimulationModel includes the following: (a) a block to hold the parameter values («Configurator»); (b) the input materials generators («InputGenerator»); (c) the simulated manufacturing system itself («MS_sim»); and (d) the end of the material flow («OutputCollection»). These elements, as well as the connections that support the data flow between them, are also represented in the IBD depicted in Figure 17.

4.3. Geometric Artifact Modeling in the Simulation System

To support the analysis of geometric deviations in the assembled products, each of the blocks that emulates an assembly stage has a process assembly model defining the geometrical representation the artifacts (i.e., fixtures and parts) involved, as well as the assembly relations between them. This artifact model is defined based on TTRS concepts using the SysML4TA profile [33]. Figure 18 shows an IBD of ProcessAssembly1, defined for the first assembly stage where two parts (saA and pB) are positioned on the fixture (f1) using the pattern previously described, i.e., a four-way location (assembly relationships c9_a and c9_b) and a two-way location (assembly relations c11_a y c11_b).

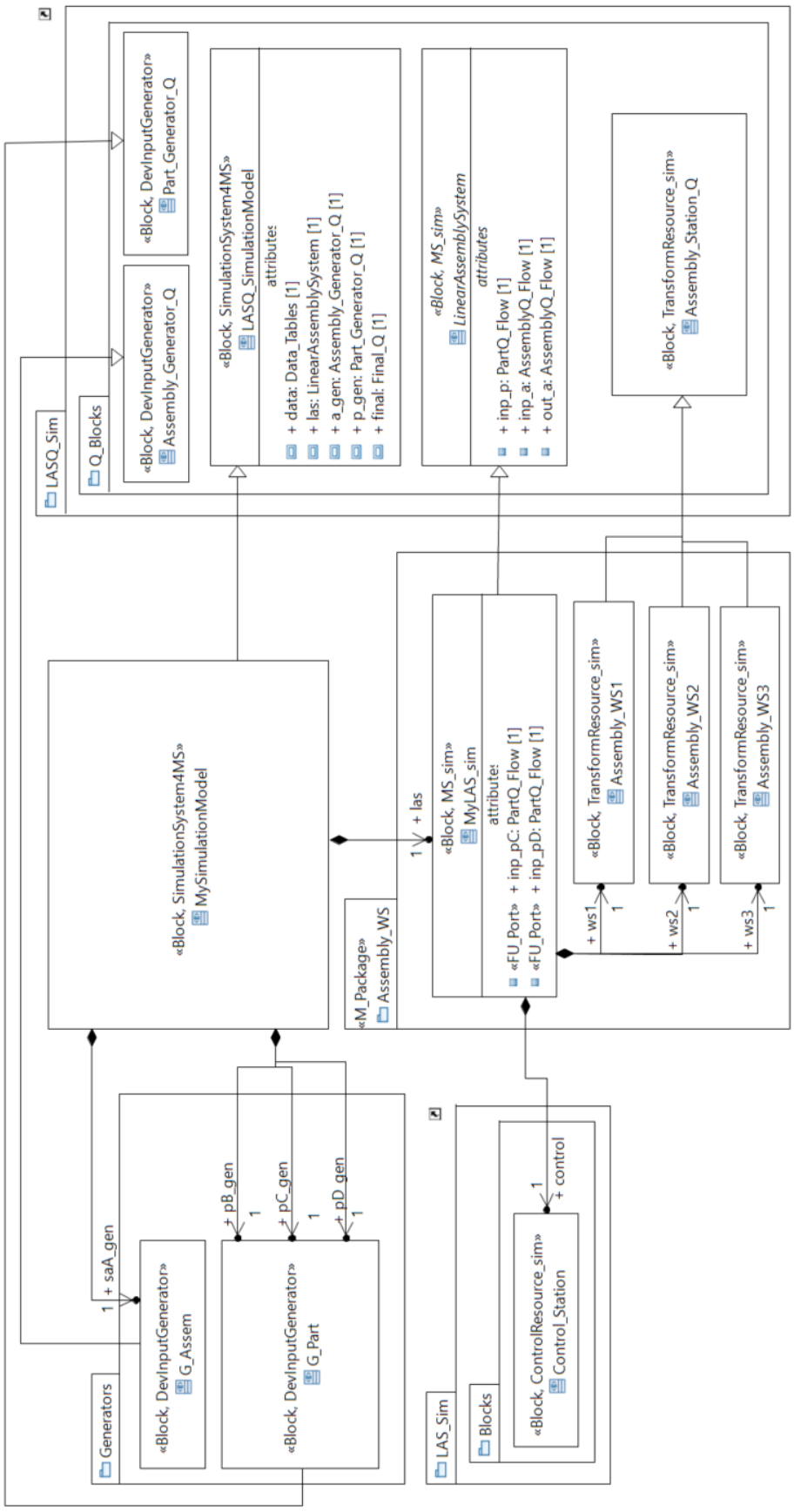


Figure 16. Block definition diagram of the simulation system (MySimulationModel).

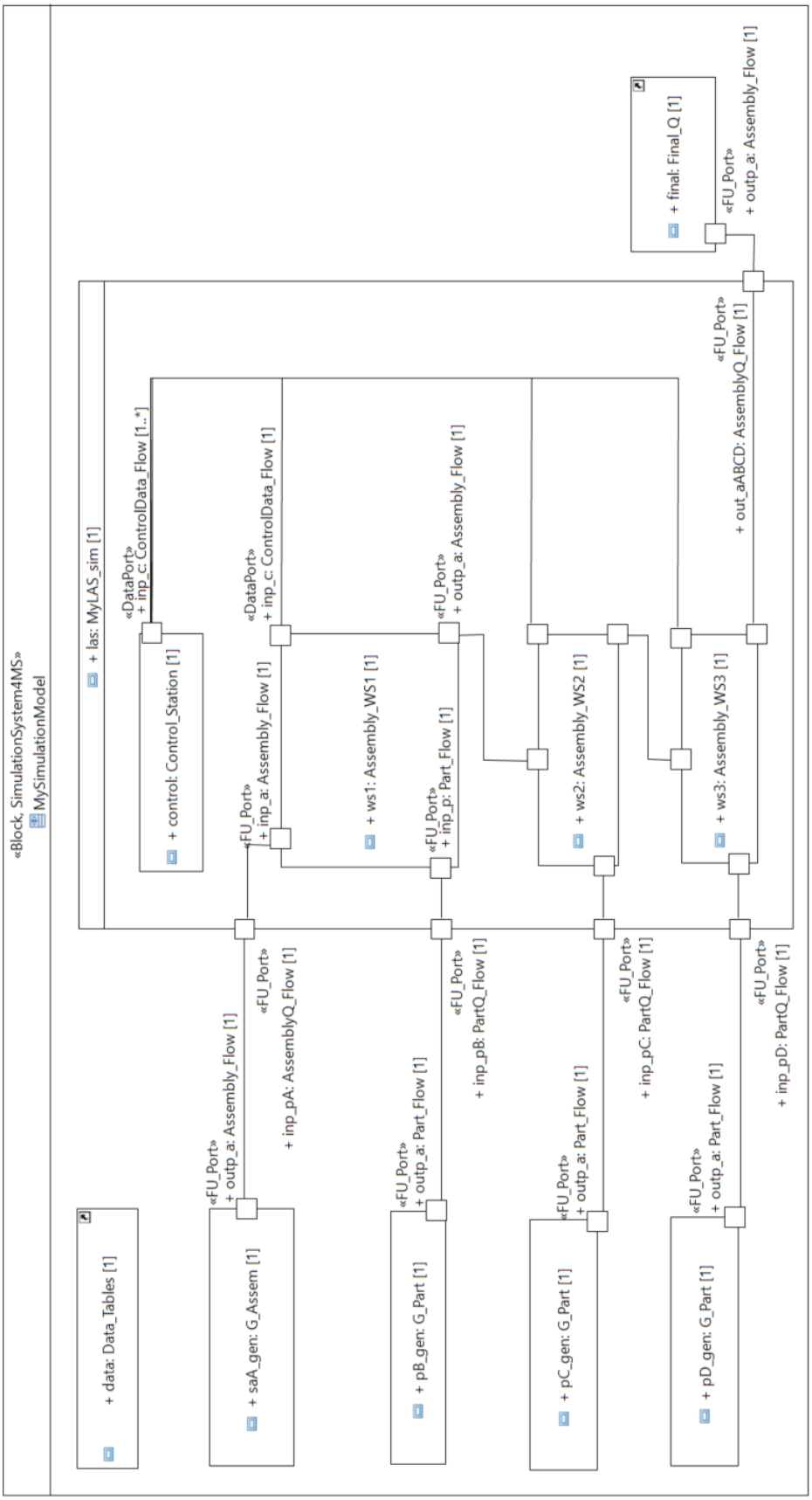


Figure 17. Internal block diagram of the simulation system (MySimulationModel).

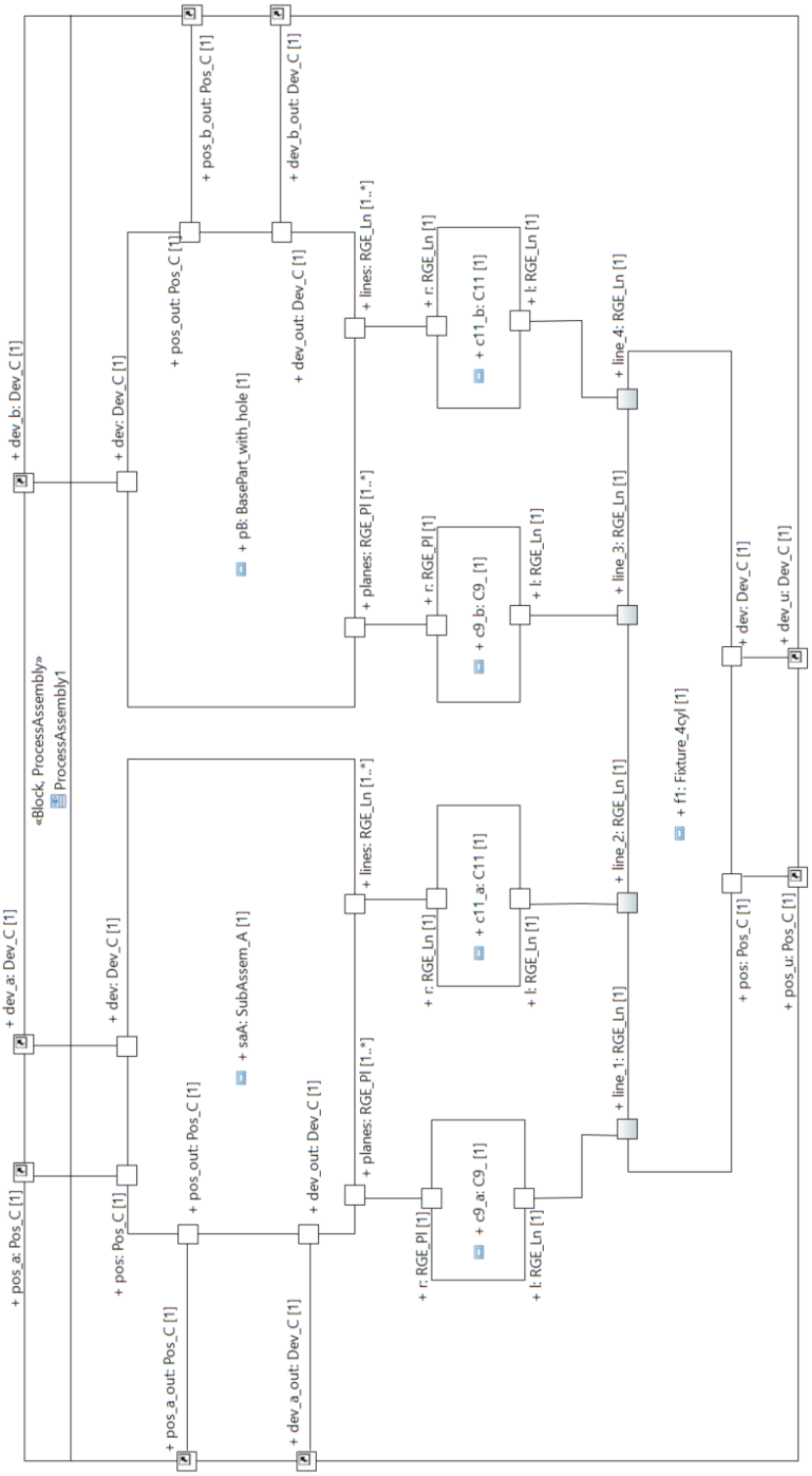


Figure 18. Internal block diagram of ProcessAssembly1.

4.4. Experiment Definition and Model Transformation

Once the structural modeling of the simulation system has been completed, and considering that the atomic behaviors are previously defined in the employed library elements, it is time to model the experiment to be executed; so, specific values are defined to each simulation parameter. In SysML, the particular values of the instantiated elements are defined by means of instance specifications. Figure 19 shows an example of a SysML block (Data_Tables) instantiation.

data_inst: Data_Tables

slots

max_parts : Integer = 4
part_models : Integer = 4
total_stages : Integer = 4
assembly_stages : Integer = 3
part_by_assembly : Integer = 4
max_steps : Integer = 4
Plan_data : Integer = {{2, 1, 2, 1, 200, 20, 10}, {2, 3, 4, 2, 190, 19, 9}, {2, 5, 6, 3, 190, 19, 9}, {2, 7, 7, 4, 60, 6, 3}}
Assemblies_data : Integer = {{1, 0, 0, 0}, {2, 0, 0, 0}, {1, 2, 0, 0}, {3, 0, 0, 0}, {1, 2, 3, 0}, {4, 0, 0, 0}, {1, 2, 3, 4}}
part_arrival_data : Integer = {1 / 20.25, 1 / 15.67, 1 / 35.67, 1 / 50}
part_arrival_data : Integer = {1 / 20.25, 1 / 15.67, 1 / 35.67, 1 / 50}

Figure 19. Instance specification for the Data_Tables block.

Finally, the model validation is executed, checking that all the OCL rules defined in the profile are met (see Figure 9). After that, the validated SysML model is transformed to Modelica model using model-to-model and model-to-text transformations (MMT and MTT). MMT is mainly supported by the application of another SysML profile (such as SysML4Modelica [10]), while automated MTT requires the execution of an algorithm. This content is out of the scope of this paper and can be consulted in [11]. Figure 20 shows a screenshot of the resulting textual model in the OpenModelica simulation environment.

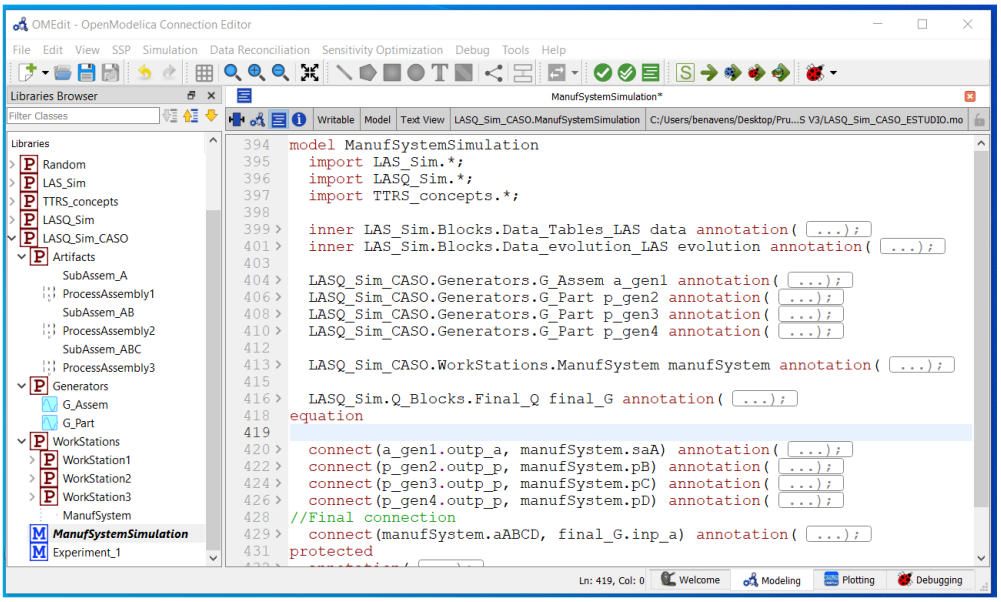


Figure 20. Modelica executable model.

4.5. Experiment Results

In order to exemplify some of the many results offered by the proposed analysis, Figure 21 shows a comparison of the orientation deviation obtained for the last assembled part (part_D), calculated with two alternative experiments in which the position of the fixture locators has been modified. In Case 1, fixture locators are defined in the positions shown in Figure 15. In Case 2, two-way locators (p2 and p4) are positioned closer to the four-way locators (p1 and p3), reducing the distance between them. Data shown in Figure 21 correspond to the orientation deviation in the third assembly station; so, the deviation has no value until the first finished product is processed. After this point, each value change corresponds to a different processed product. As expected, the obtained results show that in Case 1, the deviations of the incoming product parts results in a greater orientation deviation of the resultant assembly (higher mean value) compared to Case 2. However, a closer locator position (Case 2) subtly increases the variability (standard deviation) of this orientation deviation in the simulated products.

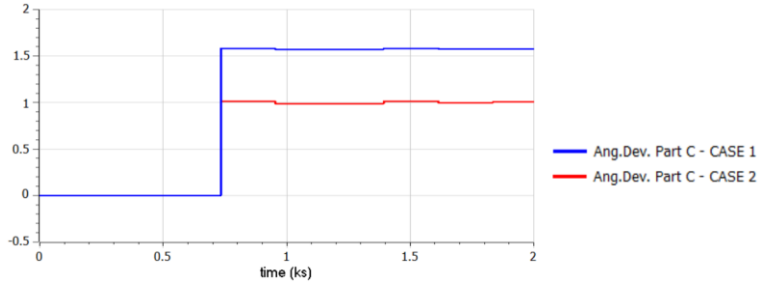


Figure 21. Orientation deviation of part D in the final product. Comparison between cases.

The simulations also provide data related to the productivity of the analyzed system, such as the throughput, blocked-station times, simulating stops for maintenance or repair of breakdowns, etc. Furthermore, these results can be crossed with aspects related to quality, for example, establishing limits on deviations for product acceptance and calculating the productivity of parts that meet the specifications. Table 1 summarizes some results of the Case 1 experiment execution with a simulation time of 5000 s (14 h approximately). During this time, 203 products were finished, which represent a throughput of 14.6 products per hour. Considering the geometric specifications and the calculated deviations for each simulated product, only 164 units meet these specifications, obtaining a real throughput of 11.8 products per hour.

Table 1. Main productivity results of the Case 1 experiment execution.

Variable	Results
Finished products	203
Throughput (Prod./h)	14.61
Finished products meeting specifications	164
Throughput with products meeting specifications (Prod./h)	11.81

5. Discussion

The research presented in this paper is focused on the multidomain simulation of multistage manufacturing systems, integrating the analysis of productivity and geometric quality. These types of analysis have been widely studied separately and are reported in multiple works. On the one hand, [36] presents the state of the art on the use of simulation models for manufacturing systems analysis during their design, and especially on process planning and material flow analysis. On the other hand, [37] is a review of tolerance-related works, ranging from tolerance specification to their mathematical analysis. Although paradigms like production quality promote the joint analysis of these aspects, in the revised literature, few works have proposed a detailed solution adopting this approach.

This orientation has already been explored by our research group in previous works. In [14], the simulation of the material flow is enriched with geometric data to support the productivity and geometric quality analysis. These multidomain simulation models integrate different control logics and/or strategies for a more realistic analysis of the system, quantifying the quality improvement and the influence of measurement processes and control decisions on productivity indicators. Specifically, the quality analysis proposed in [14] adopts of the stream of variation (SoV) technique [38,39], a mathematical model based on the state space formalism [40], to simulate the propagation of geometric deviations in multistage systems. However, the adoption of SoV has certain limitations, highlighted in [11]. To overcome these limitations, the SysML4TA profile proposed in [33], based on TTRS concepts, is adopted in this paper. The concepts of SysML4TA enable the definition of geometric artifact models and the mathematical expressions necessary to simulate geometric deviations without assumptions or simplifications.

Another substantial improvement over [14] is the use of SysML for the definition of a profile that considers the specific semantics of modeling simulation systems and supports

consistency checking based on these semantics. Some previous works have addressed the development of SysML profiles for geometric modeling or tolerance analysis, such as [41–43]. These DSML are basically a set of concepts (tags), but they do not take advantage of SysML's capabilities to formalize the semantics of the proposed concepts.

During the definition of the proposed SysML4GDPSim profile, certain difficulties or limitations have been detected in SysML, alongside its capabilities with respect to supporting the modeling of certain aspects of the system. For example, since SysML is fundamentally a descriptive language, it does not have sufficient elements to address the modeling of detailed behaviors, usually defined as opaque expressions implemented with other modeling languages, or even describing the behaviors in a natural language. The content of these opaque expressions is not based on any metamodel, so they are defined as annotations that can only be interpreted by users but without the necessary formalism to be interpreted by computers. At the current stage of this research, user models are created from library components, assuming that their behavior is well-defined, using opaque expressions written with other languages (like Modelica). However, other formal ways of defining detailed behaviors should be explored to support the consistent validation and subsequent transformation of user models into executable models. Despite these limitations, the SysML4GDPSim profile has proven to be a valid language for simulation systems design.

6. Conclusions and Future Work

The developed SysML4GDPSim profile supports the necessary concepts for modeling simulation systems to analyze manufacturing systems, especially those focused on the analysis of productivity and geometric quality. The formalization of these specific domain semantics supports the consistency assurance of models developed to analyze products and manufacturing systems during their design. Although the current state of the proposal addresses mainly intra-model consistency, this approach can also be applied to manage inter-model consistency between simulation models and specification models of the referent system. During this experience, both SysML and OCL have proven to be valid languages for creating DSML and supporting domain-specific semantics, implemented through rules defined in each concept.

The proposed profile has been successfully applied in the development of libraries of reusable elements, which have been used in the development of a case study, highlighting the assurance of consistency supported by the proposed profile. Furthermore, since the proposed profile is a well-defined DSML with its own metamodel, it is possible to establish relationships with other languages and perform automatic transformations. Although this issue is out of the scope of this paper, transformation mechanisms presented in referenced works have been applied during the case study to obtain the executable simulation models.

Based on the presented proposal, some future work lines are proposed to continue this research line. For example, the behavioral modeling in SysML must be deeply explored, reaching the formal and detailed level necessary to support the complete definition of simulation systems and to enable automatic transformation to executable models. Moreover, the proposed DSML can be extended to support alternative analysis, facilitating the definition of multidomain simulations integrating productivity analysis and other key aspects of manufacturing systems design. In a similar way, libraries can also be extended to support the analysis of other PPR systems, including other manufacturing processes (such as machining) or non-linear systems, in which more complex flows are considered. Finally, another future work line is the definition of the metamodel concepts with an ontological language such as ontology web language (OWL). This ontological approach would allow for consistency validation at both model and instance level (individuals in an ontological model) and the use of reasoners.

Author Contributions: Conceptualization, S.B.-N., P.R.C. and F.R.S.; methodology, S.B.-N. and F.R.S.; software, S.B.-N. and P.R.C.; validation, S.B.-N. and P.R.C.; formal analysis, S.B.-N. and F.R.S.; investigation, S.B.-N. and F.R.S.; resources, S.B.-N.; data curation, S.B.-N.; writing—original draft preparation, S.B.-N., P.R.C. and F.R.S.; writing—review and editing, P.R.C.; visualization, S.B.-N.; supervision, P.R.C. and F.R.S.; project administration, P.R.C.; funding acquisition, P.R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by "Pla de Promoció de la Investigació a l'UJI" of UNIVERSITAT JAUME I. Project title: "Metodología para el modelado e implementación del gemelo digital de un sistema de fabricación multietapa orientado a la optimización del funcionamiento". Reference: UJI-B2022-49.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Description of main SSMPP stereotypes.

Estereotype	Description
«*BehavioralElement_sim*»	Abstract stereotype that specializes the «Block» to represent any behavioral part of the simulation system (or the whole simulation system). A Boolean attribute (isAtomic) differentiates atomic and composite behavioral elements. Atomic behavioral elements have a «STM_MainBehavior» behavior (C1) and they cannot own behavioral parts (C2). Composite behavioral elements cannot have an explicitly defined behavior (C3) but they must own at least one behavioral part (C4).
«STM_MainBehavior»	«ClassifierBehaviorProperty» specialization to represent the behavior of an atomic «BehavioralElement_sim» Block, defined as a «StateMachine» (C5).
«SimulationSystem4MS»	«Block» representing the whole simulation system. A «SimulationSystem4MS» Block own a part typed by a «Configurator» block (C6), a part typed by a «ResultsManager» block (C7) a part typed by a «MS_sim» block (C8), at least one part typed by a «InputGenerator» block (C9), and at least one part typed by a «OutputCollector» block (C10).
«Configurator»	«Block» for data structuration and parameter definition in a simulation system.
«InputGenerator»	«*BehavioralElement_sim*» specialization to identify a block in which a flow unit starts, generating the unit flows with a periodicity defined in its behavior. An «InputGenerator» block must own at least one «FU_Port» port (C11).
«OutputCollector»	«*BehavioralElement_sim*» specialization to identify a block in which a Flow unit finishes. An «OutputCollector» block must own at least one «FU_Port» port (C12).
«ResultsManager»	«Block» defined to compute the performance measures from simulation data. A «ResultsManager» block must own at least one «DataPort» port (C13) to receive data from other simulation system parts.
«DataPort»	«Port» defined for the data exchange with other simulation system parts.
«FU_Port»	«DataPort» specialization to identify ports exchanging flow units. A «FU_Port» ports must be typed by a «FlowUnit» block (C14).
«C_Port»	«DataPort» specialization to identify ports defined for exchanging data related with the communication and processes synchronization.
«FlowUnit»	«Block» defined to represent a flow unit, representing a product batch. A «FlowUnit» block is composed by at least one part typed by «Product_sim» Block (C15).

Table A2. Description of main MSSPP stereotypes.

Stereotype	Description
«*ManufResource_sim* »	Abstract stereotype that specializes the «*BehavioralElement_sim*» to identify any manufacturing resource. A block stereotyped by a «*ManufResource_sim*» specialization must have an aggregation relationship (reference) with a «ManufResourceSpecif_data» block (C16).
«ManufResourceSpecif_data»	«Block» defined to support specification data about a manufacturing resource type. A boolean property (isTransformer) identifies the specifications about transformer resources. A transformer resource specification (isTransformer = True) must have at least one part typed by a «ConfiguredMachine» block (C17).
«ConfiguredMachine»	«Artifact» specialization to define the TTRS_based representation of a specific configuration for a transformer resource.
«*ProcessingResource_sim*»	Abstract stereotype that specializes the «*ManufResource_sim*» to identify a processing resource definition, that is, a resource through which material units flow. A block stereotyped by a «ProcessingResource_sim» specialization must have at least two «FU_Port» ports (C18).
«TransformResource_sim»	«*ProcessingResource_sim*» specialization to represent transformer resources where product characteristics are modified. A «TransformResource_sim» block must own an «ActiveConfiguration» property (C19) and at least one part typed by «ResourceDeviations» data type (C20), and its two «FU_Port» ports must be typed by different blocks (C21).
«MS_sim»	«TransformResource_sim» specialization to identify the block that emulates the whole manufacturing system.
«ActiveConfiguration»	«Property» owned by a «TransformResource_sim» block (C22) identifying the current configuration of a transformer resource.
«ResourceDeviations»	«DeviationVector» specialization to define deviation values for the key geometric characteristics in a resource artefactual representation.
«LogisticResource_sim»	«*ProcessingResource_sim*» specialization to represent logistic. A «LogisticResource_sim» block cannot own any party typed by a «TransformResource_sim» block (C23), and its two «FU_Port» ports must be typed by the same block (C24).
«ControlResource_sim»	«*ManufResource_sim*» specialization to represent a control resource, that is, a resource that supports the monitoring, control and decision-making functionality. A «ControlResource_sim» block cannot have any «FU_Port» port (C25), but it must have at least one «C_Port» or «DataPort» port (C26).
«Product_sim»	«Block» defined to support data about product units. A «Product_sim» block have an aggregation relationship (reference) with a «ProductSpecif_data» block (C27) and an «ActiveState» property (C28).
«ProductDeviations»	«DeviationVector» specialization to define deviation values for the key geometric characteristics in a product artefactual representation.
«ActiveState»	«Property» of a «Product_sim» block (C29) used to define the current product state.
«ProductSpecif_data»	«Block» defined to support specification data about a product type considered in the simulation system. A «ProductSpecif_data» block must include a behavior defined by a «NativeProcessPlan» activity (C30) and at least two parts typed by different «ProductArtifact» blocks (C31) to support the artefactual representations of the product at different manufacturing states.
«ProductArtifact»	«Artifact» specialization to define the TTRS_based representation of a specific state for a product type.
«NativeProcessPlan»	«Activity» defined to stablish the manufacturing stages of a product and the resources where they are executed. All the activities included in a «NativeProcessPlan» activity must be stereotyped as «ManufProcess» (C32) and they must be contained in a «ResourceAllocation» allocate activity partition (C33).
«ManufProcess»	«Action» owned by a «NativeProcessPlan» activity (C34) representing a manufacturing stage. A Boolean attribute (isAtomic) identifies the atomic processes, in this case, the subphases.

Table A2. *Cont.*

Stereotype	Description
«ResourceAllocation»	Specialization of the «AllocateActivityPartition» to assign particular resources to each «ManufProcess». It must be defined in a «NativeProcessPlan» Activity (C34). Every contained action must be stereotyped as «ManufProcess».
«ProcessAssembly»	«Assembly» specialization to define the TTRS_based representation of a process assembly, so a «ProcessAssembly» block has at least one reference to a «ProductArtifact» block and another reference to a «ConfiguredMachine» block.

Appendix B

Table A3. OCL expressions corresponding to some described rules.

Rule	OCL Expression
C2	if self.isAtomic=true then self.base_Class.allAttributes()->select(a | a.type.oclIsKindOf(UML::Class)). type.oclAsType(UML::Class).getAppliedStereotypes().allParents()->select(b | b.name = 'BehavioralElement_sim')->isEmpty() endif
C5	self.oclIsKindOf(UML::StateMachine)
C11	self.base_Class.allAttributes()->select(a | a.type.oclIsKindOf(UML::Port)). getAppliedStereotypes()-> select(b | b.name = 'FU_Port').size() = 1

References

1. Colledani, M.; Tolio, T.; Fischer, A.; Iung, B.; Lanza, G.; Schmitt, R.; Váncza, J. Design and management of manufacturing systems for production quality. *CIRP Ann.* **2014**, *63*, 773–796. [CrossRef]
2. Psarommatis, F.; May, G.; Dreyfus, P.A.; Kiritsis, D. Zero-defect manufacturing: State-of-the-art review, shortcomings and future directions in research. *Int. J. Prod. Res.* **2020**, *58*, 1–18. [CrossRef]
3. Zhang, L.; Zhou, L.; Ren, L.; Laili, Y. Modeling and simulation in intelligent manufacturing. *Comput. Ind.* **2019**, *112*, 103123. [CrossRef]
4. Henderson, K.; Salado, A. Value and benefits of model-based systems engineering (MBSE): Evidence from the literature. *Syst. Eng.* **2021**, *24*, 51–66. [CrossRef]
5. Ferreira, W.D.; Armellini, F.; Santa-Eulalia, L.A. Simulation in industry 4.0: A state-of-the-art review. *Comput. Ind. Eng.* **2020**, *149*, 106868. [CrossRef]
6. Vještica, M.; Dimitrieski, V.; Pisarić, M.; Kordić, S.; Ristić, S.; Luković, I. An application of a DSML in Industry 4.0 production processes. In Proceedings of the IFIP International Conference on Advances in Production Management Systems (APMS), Novi Sad, Serbia, 30 August 2020.
7. OMG Systems Modeling Language (SysML), v. 1.6. Available online: https://sysml.org/.res/docs/specs/OMGSysML-v1.6-19-11-01.pdf (accessed on 17 January 2024).
8. OMG Object Constraint Language (OCL), v. 2.4. Available online: https://www.omg.org/spec/OCL/ (accessed on 13 January 2023).
9. Gauthier, J.-M.; Bouquet, F.; Hammad, A.; Peureux, F. Tooled process for early validation of SysML models using Modelica simulation. In Proceedings of the Fundamentals of Software Engineering (FSEN 2015), Tehran, Iran, 22–24 April 2015.
10. OMG, SysML-Modelica Transformation, Version 1.0, Object Management Group. Available online: https://www.omg.org/spec/SyM/1.0/PDF (accessed on 6 February 2024).
11. Benavent-Nácher, S. Modelado y Simulación Híbrida de Sistemas de Fabricación Multietapa Orientado a la Evaluación de la Calidad Geométrica y la Productividad. Ph.D. Thesis, Universitat Jaume I, Castelló de la Plana, Spain, 2024.
12. Höpfner, G.; Jacobs, G.; Zerwas, T.; Drave, I.; Berroth, J.; Guist, C.; Rumpe, B.; Kohl, J. Model-based design workflows for cyber-physical systems applied to an electric-mechanical coolant pump. In Proceedings of the 19th Drive Train Technology Conference (ATK 2021), Aachen, Germany, 9–11 March 2021.
13. Wagner, H.; Zuccaro, C. Collaboration between system architect and simulation expert. In Proceedings of the IEEE International Symposium on Systems Engineering, Vienna, Austria, 24–26 October 2022.
14. Benavent-Nácher, S.; Rosado, P.; Romero, F. Multidomain simulation model for analysis of geometric variation and productivity in multi-stage assembly systems. *Appl. Sci.* **2020**, *10*, 6606. [CrossRef]

15. Lefeber, E.; Rooda, J.E. Modeling and analysis of manufacturing systems. In *Handbook of Dynamic System Modeling*, 1st ed.; Chapman and Hall/CRC: New York, NY, USA, 2007.
16. Clément, A. The TTRSs: 13 Constraints for dimensioning and tolerancing. In *Geometric Design Tolerancing: Theories, Standards and Applications*; Springer: Boston, MA, USA, 1998; pp. 122–131.
17. INCOSE. *Systems Engineering Handbook: A Guide for Systems Life Cycle Processes and Activities*, 4th ed.; John Wiley & Sons: Hoboken, NY, USA, 2015.
18. Dictionary by Merriam Webster. Available online: https://www.merriam-webster.com/dictionary/supersystem (accessed on 6 February 2024).
19. Gedell, S.; Claesson, A.; Johannesson, H. Integrated product and production model—Issues on completeness, consistency and compatibility. In Proceedings of the 18th International Conference on Engineering Design (ICED 11), Lyngby/Copenhagen, Denmark, 15–19 August 2011.
20. Kathrein, L.; Meixner, K.; Winkler, D.; Lüder, A.; Biffl, S. A meta-Model for representing consistency as extension to the formal process description. In Proceedings of the 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 10–13 September 2019.
21. Bifll, S.; Lüder, A.; Gerhard, D. *Multidisciplinary Enginering for Ciber-Physical Production System*; Biffl, S., Lüder, A., Gerhard, D., Eds.; Springer: Cham, Switzerland, 2017.
22. Michaelis, M.T. Function and process modeling for integrated product and manufacturing system platforms. *J. Manuf. Syst.* **2015**, *36*, 203–215. [CrossRef]
23. Souilah, M.; Tahan, A.; Abacha, N. A small displacement torsor model to evaluate machining accuracy in the presence of locating and machine geometric errors. In Proceedings of the Canadian Society for Mechanical Engineering International Congress 2021, Charlottetown, PE, Canada, 27–30 June 2021.
24. Bruscas-Bellido, G. Modelo Basado en Elementos Característicos Para la Planificación Supervisora de la Inspección. Ph.D. Thesis, Universitat Jaume I, Castelló de la Plana, Spain, 2015.
25. Law, A.M. *Simulation Modeling and Analysis*, 5th ed.; Mc Graw Hill: Tucson, AZ, USA, 2015.
26. Kim, T.G.; Zeigler, B.P. The DEVS formalism: Hierarchical, modular systems specification in an object-oriented framework. In Proceedings of the 1987 Winter Simulation Conference, Atlanta, GA, USA, 14–16 December 1987.
27. Tian, A.; Liu, S.; Chen, K.; Mo, W.; Jin, S. Spatial expression of assembly geometric errors for multi-axis machine tool based on kinematic Jacobian-torsor model. *Chin. J. Mech. Eng.* **2023**, *36*, 44. [CrossRef]
28. Mu, X.; Yuan, B.; Wang, Y.; Sun, W.; Liu, C.; Sun, Q. Novel application of mapping method from small displacement torsor to tolerance: Error optimization design of assembly parts. *J. Eng. Manuf.* **2022**, *236*, 955–967. [CrossRef]
29. Wang, H.; Lin, Y.; Yan, C. T-Maps-based tolerance analysis of composites assembly involving compensation strategies. *ASME J. Comput. Inf. Sci. Eng.* **2022**, *22*, 041007. [CrossRef]
30. Jiang, Q.; Ou, Y.; Zou, Y.; Zhou, C.-G.; Huang, S.; Qian, C.-Q. Analysis and optimization of tolerance design for an internal thread grinder. *Int. J. Adv. Manuf. Technol.* **2023**, *125*, 5369–5383. [CrossRef]
31. Umaras, E. A New Method of Rigid Assemblies Stochastic 3D Tolerance Analysis Including Thermal Performance. Ph.D. Thesis, Universidade de São Paulo, Sao Paulo, Brazil, 2022.
32. Monica, F.D.; Patalano, S.; Choley, J.; Mhenni, F.; Gerbino, S. A hierarchical set of SysML model-based objects for tolerance specification. In Proceedings of the IEEE International Symposium on Systems Engineering, Edinburgh, UK, 3–5 October 2016.
33. Benavent-Nácher, S.; Rosado Castellano, P.; Romero Subirón, F.; Abellán-Nebot, J.V. SYSML4TA: A SysML profile for consistent tolerance analysis in a manufacturing system case application. *Appl. Sci.* **2023**, *13*, 3794. [CrossRef]
34. Aguilella-Antolí, D.; Rosado-Castellano, P.; Benavent-Nácher, S. A Modelica library to simulate geometrical and dimensional deviations in process assemblies. In Proceedings of the 10th Manufacturing Engineering Society International Conference, Sevilla, Spain, 28–30 June 2023.
35. ASME. *Dimensioning and Tolerancing. Y14.5*; American Society of Mechanical Engineers: New York, NY, USA, 2019.
36. Mourtzis, D. Simulation in the design and operation of manufacturing systems: State of the art and new trends. *Int. J. Prod. Res.* **2020**, *58*, 1927–1949. [CrossRef]
37. Hallmann, M.; Schleich, B.; Wartzack, S. From tolerance allocation to tolerance-cost optimization: A comprehensive literature review. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 4859–4912. [CrossRef]
38. Zhou, S.; Huang, Q.; Shi, J. State space modelling of dimensional variation propagation in multistage machining process using differential motion vectors. *Trans. Robot. Autom.* **2003**, *19*, 296–309. [CrossRef]
39. Abellan-Nebot, J.V.; Liu, J.; Romero, F. Design of multi-station manufacturing processes by integrating the stream-of-variation model and shop-floor data. *J. Manuf. Syst.* **2011**, *30*, 70–82. [CrossRef]
40. Delchamps, D.F. *State Space and Input-Output Linear Systems*; Springer: New York, NY, USA, 2011.
41. Barbedienne, R.; Penas, O.; Choley, J.Y.; Rivière, A.; Warniez, A.; Della Monica, F. Introduction of geometrical constraints modeling in SysML for mechatronic design. In Proceedings of the 10th Europe-Asia Congress on Mechatronics, Tokyo, Japan, 27–29 November 2014.

42. Kernschmidt, K. Interdisciplinary Structural Modeling of Mechatronic Production Systems Using SysML4Mechatronics. Ph.D. Thesis, Technische Universitat Munchen, Munchen, Germany, 2019.
43. Nachmann, I.; Rumpe, B.; Wortmann, A.; Berroth, J.; Hoepfner, G.; Jacobs, G.; Spuetz, K.; Zerwas, T.; Guist, C.; Kohl, J. Modeling mechanical functional architectures in SysML. In Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, Virtual Event, Canada, 16–23 October 2020.

applied sciences

MDPI

Article

Risk Management in Good Manufacturing Practice (GMP) Radiopharmaceutical Preparations

Michela Poli, Mauro Quaglierini, Alessandro Zega, Silvia Pardini, Mauro Telleschi, Giorgio Iervasi and Letizia Guiducci *

Officina Farmaceutica, Institute of Clinical Physiology, National Research Council (CNR), 56124 Pisa, Italy; michela.poli@cnr.it (M.P.); mauro.quaglierini@cnr.it (M.Q.); alessandro.zega@cnr.it (A.Z.); silvia.pardini@cnr.it (S.P.); mauro.telleschi@cnr.it (M.T.); giorgio.iervasi@cnr.it (G.I.)
* Correspondence: letizia.guiducci@cnr.it

Abstract: Risk assessment and management during the entire production process of a radiopharmaceutical are pivotal factors in ensuring drug safety and quality. A methodology of quality risk assessment has been performed by integrating the advice reported in Eudralex, ICHQ, and ISO 9001, and its validity has been evaluated by applying it to real data collected in 21 months of activities of ^{18}F-FDG production at Officina Farmaceutica, CNR-Pisa (Italy) to confirm whether the critical aspects that previously have been identified in the quality risk assessment were effective. The analysis of the results of the real data matched the hypotheses obtained from the model, and in particular, the most critical aspects were those related to human resources and staff organization with regard to management risk. Regarding the production process, the model of operational risk had predicted, as later confirmed by real data, that the most critical phase could be the synthesis and dispensing of the radiopharmaceuticals. So, the proposed method could be used by other similar radiopharmaceutical production sites to identify the critical phases of the production process and to act to improve performance and prevent failure in the entire cycle of radiopharmaceutical products.

Keywords: quality risk management; ^{18}F-FDG PET production; good manufacturing practice

1. Introduction

Positron Emission Tomography (PET) is a nuclear medicine technique used mostly for the diagnosis, staging, and follow-up of cancers [1]. ^{18}F-FDG is the most widely used radiopharmaceutical in PET clinical investigations; it is used in over 95% of the PET examinations performed [2].

Over the years, alternative radiopharmaceuticals to ^{18}F-FDG have been studied, but research in this environment is complex due to the high costs for development (from 20 to 60 million dollars in 2013), the long period for development (minimum 7–9 years) [3], and high risk (one new radiopharmaceutical for clinical use out of ten thousand molecules tested at the beginning) [4].

^{18}F-FDG will still be the most used radiopharmaceutical in nuclear medicine for years to come, although there are malignant diseases with poor uptake of F-FDG and other benign diseases that could cause false positives [5].

The radiopharmaceutical industrial production requires compliance with Good Manufacturing Practice (GMP): the guidelines for GMP implementation are enclosed in Volume 4 of the Eudralex "The rules governing medicinal products in the European Union" [6].

The GMP rules should be applied to all phases of a drug's life cycle, starting from clinical trials through technology transfer and production until the final product's retirement [6].

The Eudralex also suggests the use of other guidelines, among which ICHQ 10 guidelines (ICHQ:International Conference on Harmonisation of Technical Requirements for

Registration of Pharmaceuticals for Human Use) were finalized at setup for the pharmaceutical quality system (PQS) to be applied throughout the product life cycle.

ICHQ10 guidelines recommend the integration of the GMP regulations with the ISO 9001:2015 quality concepts related to the whole process [7], with the aim of harmonizing the entire production cycle of the drug.

A key aspect of the PQS is represented by Quality Risk Management (QRM), as described in another ICHQ guideline (ICHQ 9) [8].

In fact, to date, QRM has assumed a pivotal role in the (radio)pharmaceutical industries for the assessment, control, and communication of the risks associated with product safety and quality.

The literature reports several cases of QRM applied to radiopharmaceutical production, which are mostly applied to segments of the production process but not to the whole process [9–11] or are intended to prevent injury to the operators or patients [12] but without quantification of the risk throughout the product life cycle.

The analyzed literature also shows that the integrated application of some quality standards, such as GMP, ISO 9001, ICHQ, and EFQM (European Foundation for Quality Management), could contribute to obtaining the quality of a radiopharmaceutical [13] through quality risk assessment and guaranteeing performance and efficacy in the entire process.

Therefore, the application of different quality standards, such as GMP, ISO 9001, and ICHQ, guarantees the quality and safety of a radiopharmaceutical [13] and contributes to optimizing the performance and efficacy of the entire production process.

This paper describes the conceptualization and the setting up of a quality risk assessment methodology to be applied to the production of sterile PET radiopharmaceuticals under the GMP regulations.

This methodology has been developed in our public research institution, taking into consideration all phases of the product life cycle, starting from development, passing through technology transfer, and finally, commercial delivery.

This methodology of quality risk management has been performed by integrating the recommendations reported in Eudralex, ICHQ, and ISO 9001, and its validity has been evaluated by applying it to real data collected in 21 months of activities of 18F-FDG production at Officina Farmaceutica, CNR-Pisa (Italy), to confirm whether the critical aspects, which have previously been identified in the quality risk assessment, were effective.

2. Materials and Methods

The quality risk assessment has been carried out, highlighting criticalities at both management and operational levels. The flowchart in Figure 1 shows the rationale applied to develop the methodology.

Figure 1. Flowchart of quality risk assessment methodology to be applied to production of sterile PET radiopharmaceuticals under GMP regulations.

Our institution is authorized to manufacture radiopharmaceuticals for diagnostic use (2-^{18}F fluoro-2-deoxy-D-glucose and ^{18}F-Fluoromethylcholine) with marketing authorization, and it is also authorized to produce fluorinated radiopharmaceuticals intended for clinical trials.

Synthesis and dispensing of ^{18}F-FDG are carried out in a classified room in accordance with annex 1 of the EU; the raw materials and the finished product enter and leave the clean room bypass through ventilated boxes.

Inside the clean room, there are five shielded cells and a Class A class isolator equipped with a Class B transfer chamber. The shielded cells contain three automatic synthesis modules. Briefly, the production of ^{18}F-FDG is carried out with the fully automated IBA Synthera® multipurpose synthesizer, and the radiopharmaceutical is then dispensed into the single vials through the semi-automatic fractionation system, which is equipped with a dose calibrator. The product is dispensed in a grade-A environment and filtered through a 0.22-micron Millipore filter. After dispensing, the filter integrity is verified using the Bubble Point test system. The quality control of the finished product is made in an unclassified environment equipped with the instrumentation for chemical-physical and microbiological (content of bacterial endotoxins) tests.

2.1. Organizational and Management Risk Assessment

For the assessment of organizational and management risks, a matrix has been developed to identify the factors that influence the production of radiopharmaceuticals. These factors have been identified considering the indications contained in the ISO 9001: 2015 standard, which is considered a golden standard for the management of an organization. For each requirement of the ISO 9001 standard, we asked ourselves a number of questions, and based on the answers, we identified the relative risks. The aspects taken into consideration were the context of Organization, Leadership, Planning, Support, Operating Activities, Performance evaluation, and Improvement.

Table 1 reports management risks individuated based on the ISO 9001:2015 standard.

Table 1. Risks identification of the management process.

Requirement 9001:2015	What Did We Ask Ourselves	Main Risks Identified
	Context of the Organization	
Understand the organization and its context	What are the factors of the context that influence the ability to achieve the expected results?	Risks related to the external context: legal, technological, competitive, market, social, cultural, national, and international. Risks related to the internal context: values, organizational capacity, and culture.
Understand the needs and expectations of stakeholders	Who are the stakeholders, and what are their needs?	Risks associated with failing to meet the needs of patients and the MA holder
Quality management system and related processes	Is there a QMS that meets the requirements of the standard?	Quality management system that is effectively an aid to the production process
	Leadership	
Leadership	Does leadership show commitment to quality improvement by taking responsibility? Have goals and policies been defined?	Uncommitted leadership poses a great risk for a manufacturing site like ours, as the business is frowned upon by the scientific community.
Roles, responsibilities, and authorities	Has management assigned responsibilities? Are there personal assignments with duties and references to the QMS?	Responsibilities must be clearly defined to avoid gray areas.

Table 1. *Cont.*

Requirement 9001:2015	What Did We Ask Ourselves	Main Risks Identified
	Planning	
Actions to address risks and opportunities	We have determined what the risks and opportunities are Have actions been planned to achieve the objectives?	Without the planning of actions to achieve the objectives, there is a risk of uncoordinated actions.
	Support	
People	Have we determined and made the necessary people available? Have we determined the skills needed? Have we promoted staff awareness of the objectives to be achieved?	People are the wealth of the organization, and the risk is high if there are few people who are not adequately trained and not aware of the role they play in achieving the objectives.
Infrastructure	Is the infrastructure properly maintained? Have the critical parts been identified and periodically checked?	The risk is linked to inefficient infrastructure and is kept under control through the use of preventive maintenance and periodic calibration with primary standards.
Documented information	Do we have procedures and records?	The risk is mainly linked to redundant document systems that generate bureaucracy.
	Operating activities (Treated separately with the FMAE method)	
Operational planning and control	We planned, implemented, and monitored the processes	(see FMEA Analysis)
Requirements for products and services	Have we determined and reviewed the requirements of the products?	The risk related to requirements not specifically defined
Design and development of products and services	During the design and development planning, have I taken into account all those factors that can guarantee the success or failure of the project?	The risk in the design phase is that of not properly evaluating the resources available to complete the project. Among the factors to be taken into consideration are the complexity of the activities and the available resources
External suppliers	Do the externally supplied processes comply with the specified requirements? How do we evaluate external suppliers?	The risk is linked to the possibility of entrusting parts of the process to unreliable suppliers.
Properties that belong to customers or external suppliers	Are we able to look after the property of customers or external suppliers?	The risk is linked to the possibility that the MA holder's information may be lost or disclosed.
Change control	In the event that a change is necessary, can we demonstrate that we have reviewed and controlled the process changes do not affect the quality of the product or process?	The risk is that changes may be made that affect the quality of the product.
Product release	Are there documents reporting the compliance of the products with the acceptance criteria? Are there documents proving traceability to the person(s) authorized to issue products and services?	The risk is the uncontrolled release of the product.
Control of non-compliant outputs	If there are outputs that do not comply with the requirements, are we able to identify them and keep them under control to prevent their inadvertent use or delivery?	The risk is the release of non-compliant products.

Table 1. *Cont.*

Requirement 9001:2015	What Did We Ask Ourselves	Main Risks Identified
	Performance evaluation	
Management review	Have we determined what and when needs to be monitored and measured? Is the system reviewed to ensure alignment with strategic guidelines?	The risk is linked to non-adherence to strategic objectives.
	Improvement	
Non-compliance and corrective and improvement actions	Do I keep track of non-conformances and take actions to keep them under control and correct them, or avoid their re-occurrence?	The risk is linked to the fact that I do not solve the problems.

Each identified risk was analyzed and defined as irrelevant, tolerable, moderate, effective, or intolerable by examining the impact of the risk on the production process and the probability of occurrence. Table 2 shows the risk quantification matrix.

Table 2. Matrix for risk quantification, risk index.

Impact/Probability	Not Very Likely	Likely	Very Likely
Low	Irrelevant	Tolerable	Moderate
Medium	Tolerable	Moderate	Effective
High	Moderate	Effective	Intolerable

Based on the obtained score, corrective and preventive actions have been identified to mitigate the risk. The actions were identified according to the criterion shown in Table 3.

Table 3. Corrective and preventive actions based on risk index.

Risk Type	Action Required
Irrelevant	No Action required
Tolerable	No further control actions are required. If necessary, improvement actions can be identified.
Moderate	Risk mitigation actions are required.
Effective	Resources must be assigned in order to reduce the risk.
Intolerable	Activities should not be carried out until the risk is reduced. If it is not possible to reduce the risk even with the use of adequate resources, the activities cannot continue.

2.2. Operational Risks

To assess the operational risk, a similar approach has been used, but rather than ISO9001, the Failure Mode and Effect Analysis (FMEA) methodology was used according to ICH Q9 because FMEA is able to identify and help us understand the risk sources, their causes, and their effects on the system by assigning priorities and implementing corrective actions to address the most serious risks.

At the beginning, the production process was divided into the main activities and then each identified activity was analyzed for the identification of the risks. Figure 2 shows the main process subdivided into the principal activities.

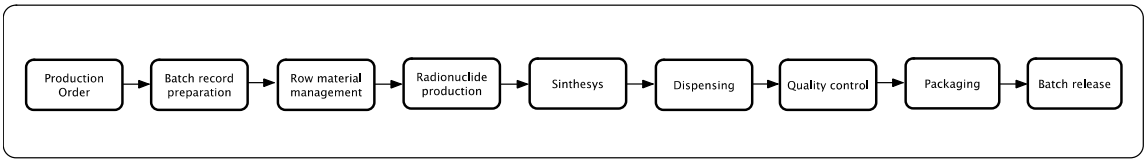


Figure 2. Principal activities of the main process.

For each step of the process, the most critical operations were identified, and the risks were quantified by evaluating the severity, detectability, and probability of occurrence according to the criteria reported in the following tables, attributing 3 points to a severe risk, 2 points to a medium risk and 1 point to a not-relevant risk (see Tables 4–6).

Table 4. Severity.

Score	Severity	Product	Economic Sustainability	Safety
3	Severe	GMP critical deviation	High economic damage (>10 k€)	Serious radiological accident (contamination beyond permitted limits or external impact on the production site);
		Product defect with potentially serious health impact or lethal risk	Severe image loss (public corrective actions);	Harm to operators (risk of death, permanent impairment, or long duration ≥3 months)
			Loss of the customer	
2	Medium	Major GMP deviation	Substantial economic damage (€ 1000–9999);	Modest radiological incident (contamination within limits, event contained within the site)
		Product defect with potential protracted harm to the patient's health (e.g., risk of hospitalization or resolvable impairment with short-term disability	Moderate damage to image (action limited to individual customers)	Medium-sized injury (disability/disability ≥7 days but <3 months)
		Process defect with impact on the pharmaceutical quality of the finished product	Repetition of a run	
1	Not relevant	Minor GMP deviation	Minor economic damage (<1000 €); higher consumption of reagents or raw materials in one run)	Minor radiological accident (contamination that can be removed by decontamination or removal of PPE)
		Non-compliance with GMP without risk to the patient's health	Deviations and non-conformities that do not lead to cost increases	Minor injury manageable with the help of the ward first aid kit)
		Deviations without impact on product safety	No economic damage	NEAR-MISS and deviations or non-compliance with GMPs that do not impact the health of operators and safety in the workplace.

Table 5. Probability of occurrence.

Score		Probability Scale	
1	Almost certainly	General	Almost inevitable or inevitable
		Systems and technical–instrumental management aspects	Aspects possibility of frequent repetition (once a week)
		Batch production	Frequency of occurrence: risk event at least once every 10 batches produced

Table 5. *Cont.*

Score	Probability Scale		
2	Possible	General	Reducible or deferrable
		Systems and technical–instrumental management aspects	Moderate repetition (once a month)
		Batch production	Frequency of occurrence: the risk event at least once every 10–30 batches produced
3	Rare	General	Almost completely avoidable
		Systems and technical–instrumental management aspects	Low repeatability (less than once a month)
		Batch production	Frequency of occurrence: risk event once over 30 batches produced

Table 6. Detectability.

Score	Detectability Scale		
3	Hardly detectable	General	Post-facto evidence
		Systems and technical–instrumental management aspects	Possibility of maintenance over time without identification
		Batch production	Evidence downstream of use
2	Detectable	General	Detection before the end of the expected function
		Systems and technical–instrumental management aspects	Detection before the start of a production process
		Batch production	Detection before use of the product
1	Immediately detectable	General	Certain detections before starting the performance of the expected function
		Systems and technical–instrumental management aspects	Detection even in the absence of a process to start
		Batch production	Detection during the process

By multiplying the severity of the damage (S) with the probability of occurrence (O) and the detection index (D), the risk index was obtained according to the formula RI = SxOxD.

Based on the risk index (RI) calculated, 3 different risk categories are individuated:

Risks with indexes 27, 18, and 12. They are high risk (NOT acceptable) and require immediate mitigation action.

Risk with index 9 or 6 medium risk: requires corrective actions to be implemented as soon as possible (days).

Risk with index from 4 to 1: low risk; they may require improvement actions to be implemented in the medium–long term (annual planning).

3. Results

Tables 7 and 8 show the results of risk assessment defined as the "Risk Index" performed based on theoretical analysis of the entire production cycle divided into management (Table 7) and operational process risks (Table 8).

Table 7. Risk assessment of the management's risks.

9001:2015 Requirement	Key factor	Risk Identification	Impact	Probability	Risk Index
Contest	Safety and radiation protection regulatory	Radiation protection, chemical risk, load handling	High	Rare	Moderate
	Organizational Skills, culture	Industrial activity carried out in a public organization	High	Possible	Effective
Leadership	Commitment	Uncommitted leadership represents a great risk in a production site like ours as the activity is frowned upon by scientific community	High	Rare	Moderate
	Policy and Strategy	Decisions not based on facts, uncoordinated decisions	Medium	Possible	Moderate
	Objectives definition	Failure to achieve objectives	Medium	Rare	Tolerable
Planning	Actions to address risks and opportunities	Failure to achieve objectives, uncoordinated decisions	High	Rare	Moderate
Support	Human Resources/ Staff Skills	Not enough staff Not competent staff Not motivated staff	High	Possible	Effective
	Management of Infrastructures	Incorrect management of infrastructures	Medium	Possible	Tolerable
	Documented Information	Loss of traceability	Medium	Rare	Tolerable
Operational activities	Planning and development	Bad evaluation of the complexity of the activities and of the available resources	High	Rare	Moderate
	Management of suppliers	Not compliant supplies	Medium	Possible	Moderate
Performance evaluation	Management review	Uncoordinated decision not based on facts	Medium	Rare	Tolerable
Continuous improvement	Not-compliance/ corrective and improvement actions	Not resolution of issues	Medium	Rare	Tolerable

Table 8. Risk assessment of the operational process. The green scale represents the severity of risk.

Failure Mode and Effect Analysis (FMEA)					
Main Process	**Activity**	**Ipoteticals Risks**	**Effects**	**Type of Damage**	**SxOxD**
Production order	Preparation of the production plan (PP)	PP not present Delayed PP Wrong PP Insufficient radioactivity	Customer dissatisfaction, loss of production	Sustainability	2
Batch record	BR Preparation	Incomplete or wrong registration forms	Operator error Loss of traceability	Sustainability	1

Table 8. *Cont.*

Failure Mode and Effect Analysis (FMEA)					
Main Process	**Activity**	**Ipoteticals Risks**	**Effects**	**Type of Damage**	**SxOxD**
Row material management	Entry into the warehouse	Difference between ordered material and arrived material	Insufficient material and loss of production	Sustainability	4
		Missed or incorrect registrationincluding labeling	Loss of traceability	Product safety	1
			Exchange of material, approval of non-conforming material		2
		Sampling	Lack of Retention Sample Loss of traceability	Product safety	1
	Storage	Expired material	Loss of production Non-conformities of the finished product	Product safety	2
		Insufficient inventory		Sustainability	2
		Incorrect storage conditions		Product safety	2
	Exit from the warehouse	Exchange of material	Loss of production Non-compliant finished product	Product safety	2
Radionuclide production	target irradiation	Boot failure	Delayed product shipment or lack of synthesis	Sustainability	2
		Human error in target selection			
		Failed to load the target	Delayed product shipment or lack of synthesis	Sustainability	4
		Failure irradiation	Delayed product shipment or lack of synthesis	Sustainability	4
	Delivery of radionuclides	Failure to transfer	Lack of synthesis	Sustainability	6
		Incomplete transfer	Loss of production, reduction in the transferred activity	Sustainability	6
		Transfer to the wrong cell/module	Delayed product shipment or lack of synthesis	Sustainability	2
Product preparation	Clean room management	Parameters out of specification	Production delay or loss of production	Sustainability	1
			OOS final product	Product safety	1
	Shielded cell management	Inefficient cell	Production delay or loss of production	Sustainability	1
		Loss of radioactivity	Operator contamination	Operator safety	1
	Isolator management	Fault DTC-SAS-LAF	Production delay or loss of production	Sustainability	6
			OOS final product	Product safety	6
Radiopharmaceutical synthesis	Synthesis	Pretest times too long	Production delay	Sustainability	4
		Preliminary tests not passed	Loss of production	Sustainability	4
		Low yield	Customer dissatisfaction	Sustainability	6
		Synthesis failure	Loss of production	Sustainability	6

Table 8. *Cont.*

Failure Mode and Effect Analysis (FMEA)					
Main Process	**Activity**	**Ipoteticals Risks**	**Effects**	**Type of Damage**	**SxOxD**
Dispensing	SAS-LAF material entrance	Blocking moving parts or operating software	Production delay or loss of production	Sustainability	1
		Sanitization of raw materials	DTC bacterial contamination	Product safety	1
	DTC material entrance	Sanitization of raw materials	bacterial contamination of the finished product	Product safety	1
	CRP6 Set-up	Error in assembling the kit	Production delay or loss of production	Sustainability	6
		CRP6 software communication problems	Production delay or loss of production	Sustainability	6
		Incorrect labeling of the vial	OOS final product	Product safety	6
		Vial packaging	Production delay/customer dissatisfaction	Sustainability	6
	Bulk product transfer to DTC	Incomplete transfer to the DTC	Customer dissatisfaction	Sustainability	6
		Failure to transfer to the DTC	Loss of production	Sustainability	6
	DTC-SASLAF operation	Delay in dispensing	Customer dissatisfaction	Sustainability	2
		Inability to continue with dispensing	Loss of production	Sustainability	4
		Loss of the product into the DTC	Personal and environmental contamination	Operator safety	6
		Mixup of the vial and shielded container	Customer dissatisfaction, complaint	Product safety	3
		Breakage of the vial in the delivery output	Personal and environmental contamination	Operator safety	6
			Customer dissatisfaction	Sustainability	6
	Delivery	Vial blocked in the delivery output	Production delay or loss of production	Sustainability	6
		Fall of the container	Break of the vial	Sustainability	4
			Personal and environmental contamination	Operator safety	4
	Bubble point test	Test failed	Loss of production	Sustainability	3
CQ	PH measurement	Value out of range in calibration	Production delay or loss of production	Sustainability	1
		Failure of the instrument	Production delay	Sustainability	1
		Value out of range in sample measurement	OOS final product	Product safety	2
	Half-life analysis (dose calibrator)	Device faulty or not calibrated	Production delay or loss of production	Sustainability	1
		Value out of range in sample measurement	OOS final product	Product safety	2

Table 8. *Cont.*

Failure Mode and Effect Analysis (FMEA)					
Main Process	**Activity**	**Ipoteticals Risks**	**Effects**	**Type of Damage**	**SxOxD**
CQ	Analysis LAL test	Device faulty or not calibrated	Production delay or loss of production	Sustainability	1
		Value out of range in sample measurement	OOS final product	Product safety	4
	GC residual solvent analysis	Error in the standards preparation	Production delay	Sustainability	2
		Device faulty or not calibrated	Production delay or loss of production	Sustainability	1
		Value out of range in sample measurement	OOS final product	product safety	4
	Chemical and radiochemical purity analysis by HPLC	Error in the standards preparation	Production delay	Sustainability	4
		Device faulty or not calibrated	Production delay or loss of production	Sustainability	2
		Value out of range in sample measurement	OOS final product	Product safety	4
	Chemical and radiochemical purity: TLC analysis	Error in the standards preparation	Production delay	Sustainability	1
		Device faulty or not calibrated	Production delay or loss of production	Sustainability	2
		Value out of range in sample measurement	OOS final product	Product safety	2
	Radionuclide purity: gamma spectrometry analysis	Device faulty or not calibrated	Production delay or loss of production	Sustainability	4
		Value out of range in sample measurement	OOS final product	Product safety	2
	Chemical purity: determination of kriptofix	Value out of range in sample measurement	OOS final product	Product safety	3
	Sterility control	Loss of samples, shipment not carried out	GMP deviation	Product safety	2
		Value out of range in sample measurement	GMP deviation	Product safety	3
	Gamma spectrometry analysis at 72 h	Value out of range in sample measurement	GMP deviation	Product safety	1
	Microbiological control	Value out of range	OOS final product	Product safety	2
		Device faulty or not calibrated (oven and SAS)	GMP deviation	Product safety	1
Packaging of the final product	Preparation for shipment	Suitcases not available	Inability to ship	Sustainability	8
		Shipping documents not available	Shipping delay	Sustainability	1
		Mixup of the shielded container and Suitcases	Customer dissatisfaction	Product safety	2
	Delivery to the carrier	Courier not available	Inability to ship	Sustainability	2
Batch Release	BR verification	Human error in data verification	release of non-compliant product	Product safety	3

The most critical activities (productivity, reliability, non-conformities) were analyzed as suggested in Chapter 1 of the EUGMP guidelines Volume 4 for a period of 21 months of activity of the production site after restart.

The results of these analyses are summarized in the following figures.

The first analysis concerns the productivity of the site evaluated through the number of productions/months. Figure 3 shows that over those 21 months, the productivity went from 2 productions per month to 27 productions per month.

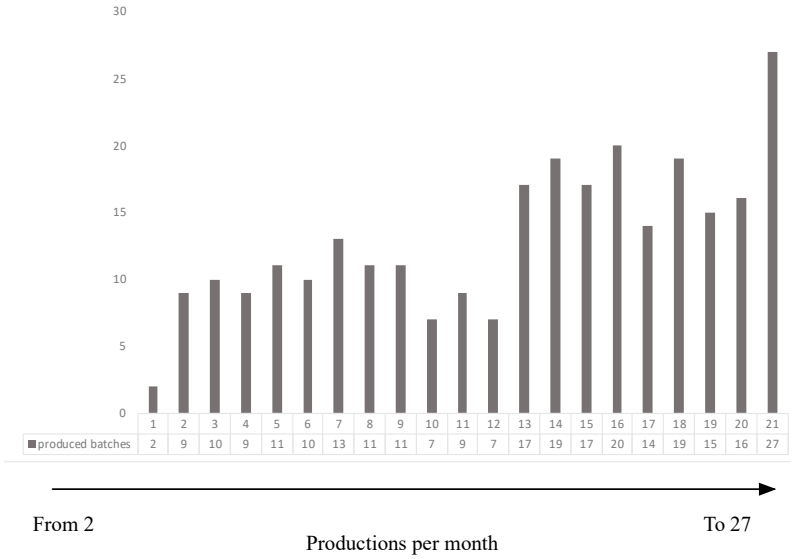


Figure 3. Increase in productivity.

Regarding reliability evaluation, the collected data were grouped into 17 slots of 20 productions, and for each slot, the reliability percentage value was calculated using the following formula: Reliability = scheduled lots/released lots × 100.

The average reliability is 84%, and the trend is positive, as shown in Figure 4.

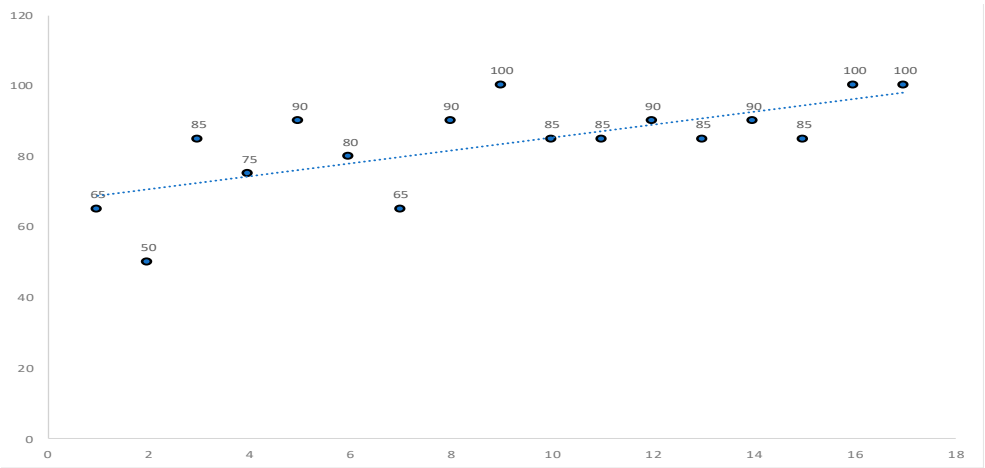


Figure 4. Increase in reliability over time (each circle represents the % scheduled lots/released lots).

All non-conformities (n = 143) were recorded in the 21-month period and were analyzed in accordance with the internal procedures that contemplate a classification related to the items reported below:

- Production process deviations;
- Quality control deviations;
- Raw material deviations;

- General site deviations;
- Documentation deviations.

Figure 5 shows how non-conformities are distributed within these five areas.

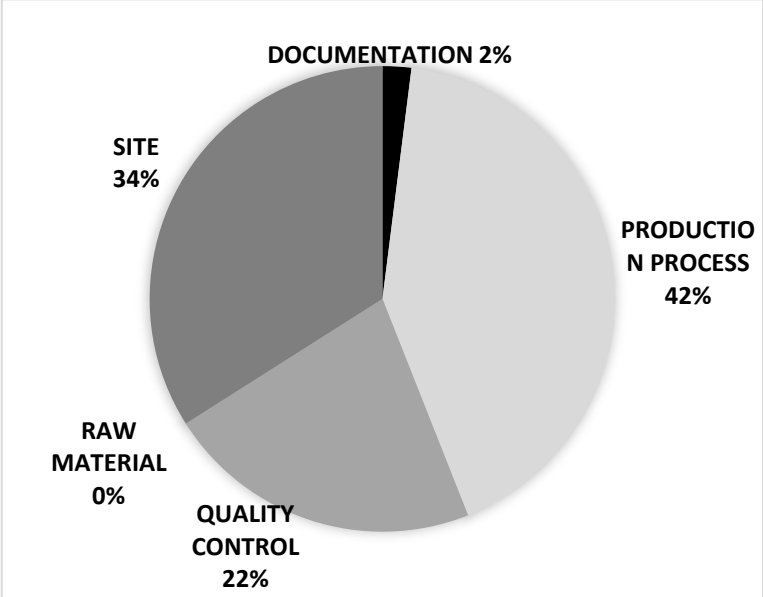


Figure 5. Non-conformities distributed within these 5 areas.

To verify if the risk assessment process had correctly identified the most critical segments of the process, we further subdivided the registered non-conformities along the entire production process as scheduled in Figure 6 (main steps process), as shown in Figure 6.

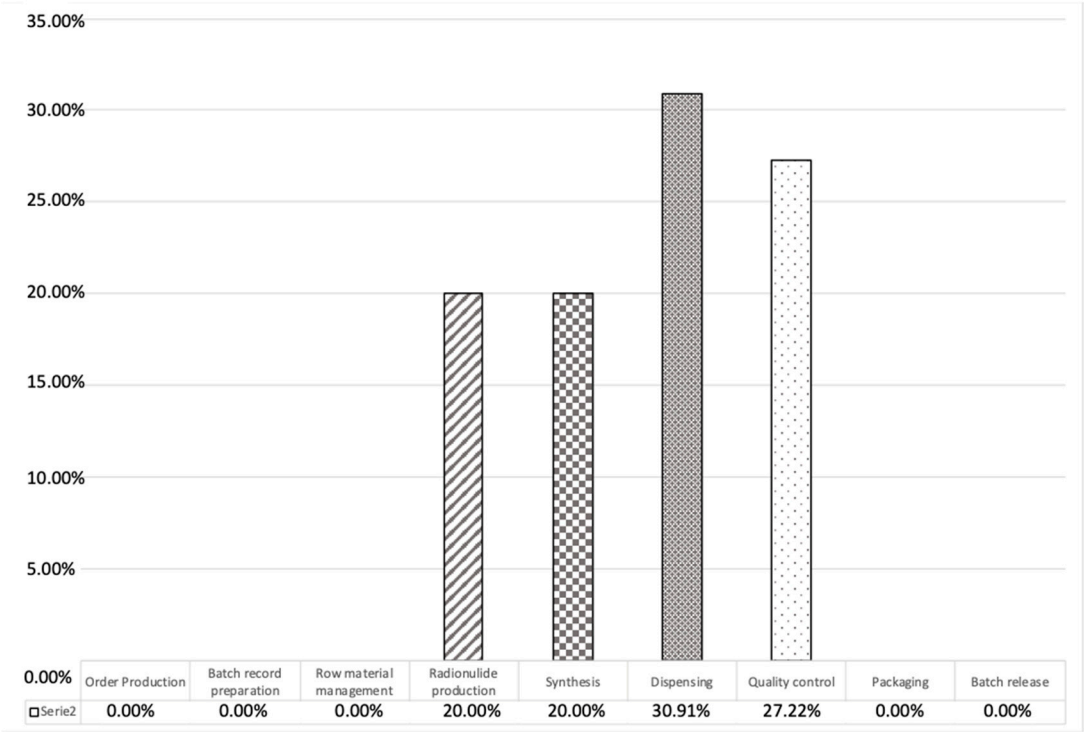


Figure 6. Percentage distribution of registered non-conformities along the entire production process steps.

We have also quantified the impact of the non-conformities on the three types of failures identified during the risk assessment: product quality, personnel safety, and sustainability.

In general, the risk to the quality of the product never occurred because the system was able to detect potential risks to product quality, including with sterility control. To guarantee the sterility of the finished product, specific actions are adopted to ensure that the entire production chain takes place in quality assurance conditions, despite the test result being available to the patient at least 15 days after the injection.

For the operator risk, the data analysis shows that the most critical process is the dispensing in 8.6% of non-conformities; radioactive material has leaked out, especially inside the shielded isolator, due to incorrect assembly of the sterilization filter. This aspect was carefully monitored, and corrective actions were taken, which led to risk mitigation. Another risk for operator safety is the breaking of the vial in the delivery channel (1.2% of non-conformities), but also, in this case, the radioactivity monitoring systems have effectively detected the contamination, and the automatic locking systems for the opening of the cells have prevented the contamination of the operator.

As regards the failure to sustainability, the failure to release the lot generally had an important impact on the site's sustainability.

Figure 7 shows the main causes of the rejection of a lot of radiopharmaceuticals.

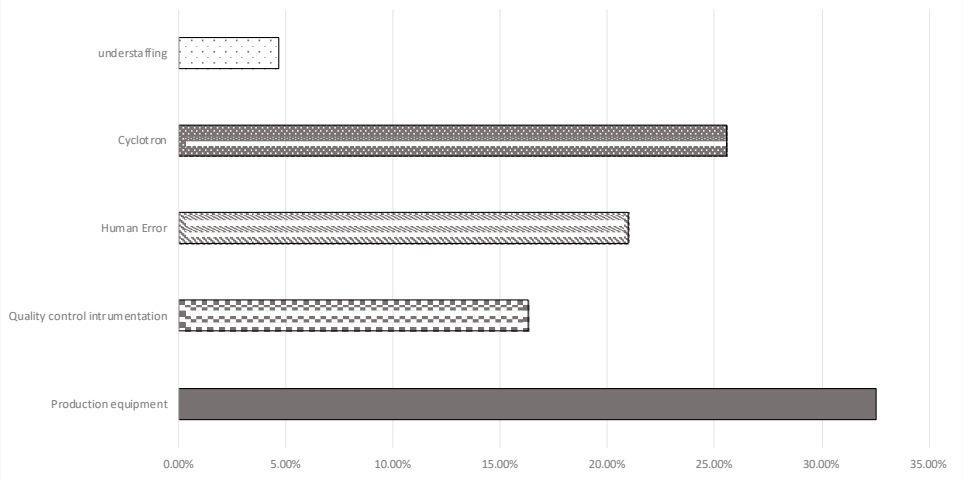


Figure 7. Main causes of lot rejection.

4. Discussion

In general, the radiopharmaceutical manufacturing process starts with the production of the radionuclide and continues with the synthesis of the radiopharmaceutical according to the formulation described in the Marketing Authorization (MA) dossier or Pharmacopoeia Monograph [14].

The production of PET radiopharmaceuticals is extremely complex due to a series of factors [15]: the delivery of the final product that takes place with the sterility control still in progress; the short half-life of the tracer; the very high number of batches produced compared to conventional drugs; the activity carried out mainly during the night; and the handling of radioactive substances. The complexity of this system implies that risk assessment is even more necessary than in other production systems.

Fortunately, the GMP standard helps us keep the critical aspects of the production process under control by indicating the tools to be used "to manage" the risk [16,17].

Our GMP experience has allowed us to develop a risk assessment, which is conducted by separately considering organizational and management risks and operational risks.

This report describes the experience of a public radiopharmaceutical site that has adopted the GMP standards and which, therefore, produces radiopharmaceuticals following the indications of the EU-GMP for industrial production [18]. Since the start of GMP

radiopharmaceutical production, the site has used an approach founded on risk-based thinking and followed regulatory updates over the years.

In 2021, the production site radically changed the production process, moving from production with final sterilization in an autoclave to production using aseptic techniques. Before starting with the new production process, a risk assessment was carried out, which identified the most critical segments of the process. Twenty-one months after the restart of the production, which was intended for distribution on the national territory, the data that emerged were collected and analyzed to verify if the risk assessment process carried out ex ante had been able to identify the main risks.

Compared to the initial risk assessment, the evidence collected during these twenty-one months allows us to state that the method, adopted both for the assessment of management risks and operational risks, is a suitable tool for identifying the critical phases of the process.

At the managerial level, the greatest criticality was mainly linked to personnel management.

The results of the analysis also show that production is the most critical phase of the process (asepsis and radioactivity). In particular, the dispensing phase has proven to be the most critical, implying a delay or even a loss of production along with customer dissatisfaction. In addition, there are problems during the dispensing phase; as an example, the breakage of a product vial could result in personal and environmental contamination and put an operator's safety in danger (Table 2).

In general, the preventive actions to be undertaken to mitigate the risk can be grouped into two main types: the continuous training of staff, with attention to motivation and awareness of the complexity of the system, and the increasingly detailed processes so that the risks can be reduced as much as possible and, therefore, achieve an overall risk reduction.

In conclusion, having a risk management process in place is critical for GMP production facilities, and the identification of risks in the matrix is a useful tool. Our implemented model has identified and quantified the risks, and, given the overlap between the risks identified and the analysis of real data, the proposed method could be used by other similar radiopharmaceutical production sites to identify the critical phases of the production process, improve the performance, and prevent damages to the entire cycle of the radiopharmaceutical product.

Author Contributions: Conceptualization, M.P., G.I. and L.G.; methodology, A.Z., S.P. and M.T.; software, M.Q.; validation, M.T. and A.Z.; formal analysis, M.P. and M.Q.; investigation, G.I. and L.G.; data curation S.P. and M.P.; writing—original draft preparation, M.P.; writing—review and editing, G.I. and L.G.; visualization, L.G.; supervision L.G. and G.I.; project administration, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Boellaard, R.; Delgado-Bolton, R.; Oyen, W.J.; Giammarile, F.; Tatsch, K.; Eschner, W.; Verzijlbergen, F.J.; Barrington, S.F.; Pike, L.C.; Weber, W.A.; et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: Version 2.0. *Eur. J. Nucl. Med. Mol. Imaging* **2015**, *42*, 328–354. [CrossRef]
2. Chierichetti, F.; Pizzolato, G. 18F-FDG-PET/CT. *Q. J. Nucl. Med. Mol. Imaging* **2012**, *56*, 138–150. [PubMed]
3. Shaw, R.C.; Tamagnan, G.D.; Tavares, A.A.S. Rapidly (and Successfully) Translating Novel Brain Radiotracers From Animal Research Into Clinical Use. *Front. Neurosci.* **2020**, *14*, 871. [CrossRef] [PubMed]
4. Nerella, S.G.; Singh, P.; Sanam, T.; Digwal, C.S. PET Molecular Imaging in Drug Development: The Imaging and Chemistry Perspective. *Front. Med.* **2022**, *9*, 812270. [CrossRef] [PubMed]

5. Wright, C.L.; Binzel, K.; Zhang, J.; Knopp, M.V. Advanced Functional Tumor Imaging and Precision Nuclear Medicine Enabled by Digital PET Technologies. *Contrast Media Mol. Imaging* **2017**, *2017*, 5260305. [CrossRef] [PubMed]
6. Vol 4 EU GMP (Guidelines for Good Manufacturing Practices for Medicinal Products for Human and Veterinary Use—Pharmaceutical Quality System). Available online: https://health.ec.europa.eu/medicinal-products/eudralex/eudralex-volume-4_en (accessed on 8 October 2003).
7. ICH Q10 Pharmaceutical Quality System—Scientific Guideline. 2015. Available online: https://www.ema.europa.eu/en/ich-q10-pharmaceutical-quality-system-scientific-guideline (accessed on 28 May 2014).
8. ICH Q9 Guideline of the International Conference on Harmonization (ICH Q9). 2015. Available online: https://www.ema.europa.eu/en/ich-q9-quality-risk-management-scientific-guideline (accessed on 26 July 2023).
9. Lange, R.; ter Heine, R.; Decristoforo, C.; Peñuelas, I.; Elsinga, P.H.; van der Westerlaken, M.M.; Hendrikse, N.H. Untangling the web of European regulations for the preparation of unlicensed radiopharmaceuticals: A concise overview and practical guidance for a risk-based approach. *Nucl. Med. Commun.* **2015**, *36*, 414–422. [CrossRef] [PubMed]
10. Elsinga, P.; Todde, S.; Penuelas, I.; Meyer, G.; Farstad, B.; Faivre-Chauvet, A.; Mikolajczak, R.; Westera, G.; Gmeiner-Stopar, T.; Decristoforo, C.; et al. Guidance on current good radiopharmacy practice (cGRPP) for the small-scale preparation of radiopharmaceuticals. *Eur. J. Nucl. Med. Mol. Imaging* **2010**, *37*, 1049–1062. [CrossRef] [PubMed]
11. Chitto, G.; Di Domenico, E.; Gandolfo, P.; Ria, F.; Tafuri, C.; Papa, S. Risk assessment and economic impact analysis of the implementation of new European legislation on radiopharmaceuticals in Italy: The case of the new monograph chapter Compounding of Radiopharmaceuticals. *Curr. Radiopharm.* **2013**, *6*, 208–214. [CrossRef] [PubMed]
12. Dondi, M.; Torres, L.; Marengo, M.; Massardo, T.; Mishani, E.; Ellmann AV, Z.; Solanki, K.; Bischof Delaloye, A.; Estrada Lobato, E.; Nunez Miller, R.; et al. Comprehensive Auditing in Nuclear Medicine through the International Atomic Energy Agency Quality Management Audits in Nuclear Medicine Program. Part 2: Analysis of Results. *Semin. Nucl. Med.* **2017**, *47*, 687–693. [CrossRef] [PubMed]
13. VanDuyse, S.A.; Fulford, M.J.; Bartlett, M.G. ICH Q10 Pharmaceutical Quality System Guidance: Understanding Its Impact on Pharmaceutical Quality. *AAPS J.* **2021**, *23*, 117. [CrossRef] [PubMed]
14. *Monograph for (^{18}F) Fludeoxyglucose Injection*; Working Document QAS/17.733; WHO: Geneva, Switzerland, 2017.
15. Hansen, S.B.; Bender, D. Advancement in Production of Radiotracers. *Semin. Nucl. Med.* **2022**, *53*, 266–275. [CrossRef] [PubMed]
16. Todde, S.; Kolenc Peitl, P. Guidance on validation and qualification of processes and operations involving radiopharmaceuticals. *EJNMMI Radiopharm. Chem.* **2017**, *2*, 8. [CrossRef] [PubMed]
17. Gillings, N.; Hjelstuen, O.; Behe, M.; Decristoforo, C.; Elsinga, P.H.; Ferrari, V.; Kiss, O.C.; Kolenc, P.; Koziorowski, J.; Laverman, P.; et al. EANM guideline on quality risk management for radiopharmaceuticals. *Eur. J. Nucl. Med. Mol. Imaging* **2022**, *49*, 3353–3364. [CrossRef] [PubMed]
18. Gillings, N.; Hjelstuen, O.; Ballinger, J.; Behe, M.; Decristoforo, C.; Elsinga, P.; Ferrari, V.; Kolenc Peitl, P.; Koziorowski, J.; Laverman, P.; et al. Guideline on current good radiopharmacy practice (cGRPP) for the small-scale preparation of radiopharmaceuticals. *EJNMMI Radiopharm. Chem.* **2021**, *6*, 8. [CrossRef] [PubMed]

applied sciences

MDPI

Article

Multi-AGV Scheduling under Limited Buffer Capacity and Battery Charging Using Simulation Techniques

Jin-Sung Park and Jun-Woo Kim *

Department of Industrial and Management Systems Engineering, Dong-A University, Busan 49315, Republic of Korea; pjs0958@dau.ac.kr
* Correspondence: kjunwoo@dau.ac.kr; Tel.: +82-51-200-7687; Fax: +82-51-200-7697

Abstract: In recent years, automated guided vehicles (AGVs) have been widely adopted to automate material handling procedures in manufacturing shopfloors and distribution centers. AGV scheduling is the procedure of allocating a transportation task to an AGV, which has large impacts on the efficiency of an AGV system with multiple AGVs. In order to optimize the operations of multi-AGV systems, AGV scheduling decisions should be made with consideration of practical issues such as buffer space limitations and battery charging. However, previous studies have often overlooked those issues. To fill this gap, this paper proposes a simulation-based multi-AGV scheduling procedure for practical shopfloors with limited buffer capacity and battery charging. Furthermore, we propose three kinds of rules: job selection rules, AGV selection rules, and charging station selection rules, for AGV scheduling in practical shopfloors. The performance of the rules is evaluated through multi-scenario simulation experiments. The FlexSim software v.2022 is used to develop a simulation model for the experiments, and the experimental findings indicate that the job selection rules have larger impacts on the average waiting time than the other kinds of rules.

Keywords: material handling automation; automated guided vehicles (AGVs); AGV scheduling; simulation; FlexSim software

1. Introduction

Material handling procedures have significant impacts on the cost and delivery of materials or goods; however, traditional material handling methods that rely on human labor or manned vehicles have limitations in terms of speed, accuracy, and flexibility. Thus, automated material handling systems have emerged as an essential tool for optimizing production, reducing costs, and ensuring the timely delivery of products. In particular, an automated guided vehicle (AGV) that transports specific objects within a workspace is one of the most promising technologies for automation of material handling procedures [1]. With the recent proliferation of AGVs in the manufacturing and logistics industries, much attention is paid to the operations management of large AGV systems with multiple AGVs [2–5].

AGVs are used to transport materials from one place to another within manufacturing shopfloors and distribution centers, demonstrating a high degree of autonomy and flexibility [6–8]. To operate an AGV system efficiently, AGV scheduling, a decision-making process that allocates appropriate AGV vehicles and tasks to multiple AGVs, plays an important role. AGV scheduling, which focuses on efficient utilization of AGVs, involves several sub-decisions such as job and AGV selection and routing [9–11]. However, given the complexity of modern manufacturing environments, characterized by complex shopfloor conditions and diverse operational requirements, establishing an efficient AGV schedule is not an easy task and requires significant effort [12,13].

This paper proposes a simulation-based approach for establishing an AGV scheduling policy for practical shopfloors. The AGV scheduling policy in this paper consists of three

kinds of rules: the job selection rule, the AGV selection rule, and the charging station selection rule. Conventional AGV scheduling methods prioritize the first two types of rules; however, it is crucial to acknowledge the significant impact of battery charging on the efficiency of AGV systems.

Another practical issue considered in this paper is the limited buffer capacity, where a buffer denotes work-in-process at a specific place. Many manufacturing plants impose limitations on the buffer capacity and maximum buffer size within their production process due to congestion control and safety concerns. In such shopfloors, production is halted when the buffer size reaches the limitation, and it is resumed after the buffer size decreases. The buffer capacity limitation can also affect the performance of AGV scheduling.

This paper applies the discrete event simulation (DES) technique to multi-AGV scheduling under limited buffer capacity and battery charging. A simulation model for shopfloors with a multi-AGV system is developed using the commercial DES software, FlexSim. Note that the structure of the simulation model is motivated by the practical shopfloor of an automotive part manufacturer in Korea. Experimental results demonstrate that optimal AGV scheduling can be obtained through multi-scenario simulation experiments.

The remainder of this paper is organized as follows: Section 2 reviews existing studies addressing the AGV scheduling problem. Section 3 discusses the main issues related to buffer capacity limitations and battery charging in a practical shopfloor. We also propose some scheduling rules related to work and AGV selection, as well as charging station selection. Section 4 presents a simulation model for multi-AGV scheduling in a practical shopfloor and demonstrates that the optimal AGV scheduling policy can be derived through multi-scenario experiments. Finally, Section 5 summarizes the conclusions of this study, presents future research directions, and concludes.

2. Literature Review

Numerous researchers have employed different approaches to address AGV scheduling issues. Typically, AGV scheduling entails determining job selection rules for assigning specific transport goods to AGVs and AGV selection rules for choosing suitable AGVs from the available pool. These two key decisions, job selection and AGV selection, are fundamental to AGV scheduling and have been the subject of numerous prior studies [14–18]. However, early research in this field mainly focused on small-scale AGV operations in limited-sized workshops, and practical operational issues such as AGV battery charging and buffer capacity limitations were not considered.

Meanwhile, in recent years, researchers have dedicated significant attention to the development of practical AGV scheduling methods applicable in real work environments. These studies propose scheduling models and optimization algorithms that consider real-world operational issues such as buffer capacity limitations and AGV battery charging.

In real manufacturing environments, buffer capacity limitations play an important role in the efficiency and flexibility of the production process. Some physical manufacturing sites have limited buffer capacity for a variety of reasons, including the need to effectively manage work-in-process (WIP) and meet short lead times. Therefore, it is important to consider factors such as buffer capacity limitations in order to establish scheduling suitable for real environments. There is existing research that addresses the impact of buffer capacity limitations on job scheduling.

Wang et al. [19] discussed the issue of scheduling jobs for multiple multiload AGVs. The study considered the transportation time between these AGVs and the processing time between machines, addressing a situation where there is a limitation on the input buffer capacity of each machine. In a similar study to the previous work, the literature by Wang et al. [20] also addressed the multi-load AGV scheduling problem due to limited buffer capacity. A heuristic algorithm based on simulated annealing was proposed to minimize the maximum completion time and utilize it in a real environment. Naeem et al. [21] proposed a mathematical model that integrates the scheduling of yard cranes and AGVs at

a container terminal and solved the scheduling problem by reflecting new constraints in the model to consider limited buffer capacity and storage allocation of containers.

Meanwhile, studies on the topic of AGV battery charging have also looked into issues like power management and energy efficiency. Li et al. [22] developed a multi-objective nonlinear programming model based on a genetic algorithm to minimize the maximum completion time in a flexible manufacturing cell, considering the battery charging and waiting time of AGVs. Additionally, the influence of various loads on AGV power consumption was discussed. Yang et al. [23] focused on optimizing AGV scheduling in automated container terminals (ACTs) and specifically conducted research on scheduling strategies considering battery replacement and speed control. Additionally, they proposed a mixed integer programming (MIP) model designed to minimize carbon dioxide emissions and delays caused by AGV operation and developed a solution based on genetic algorithms. Sun et al. [24] addressed the AGV scheduling problem considering the battery status of the AGV to solve the problem of AGV utilization rate and high workload imbalance in ACTs and proposed a collaborative scheduling model based on simulated annealing (SA) and genetic algorithm (GA). Singh et al. [25] introduced a mixed integer linear programming (MILP) model and a heuristic algorithm as a solution to the AGV scheduling problem and proved that efficient scheduling considering AGV battery constraints is possible through experiments using data from the actual shopfloor. Meanwhile, Abderrahim et al. [26] addressed job scheduling considering battery constraints in AGV-based manufacturing operations and showed that productivity and energy consumption in manufacturing facilities can be optimized through AGV scheduling and battery charging.

The common feature of the above studies is that they focus on improving work optimization and energy efficiency in manufacturing environments by approaching the scheduling problem by considering buffer capacity limitations and AGV battery charging. Each study uses different optimization techniques and models to increase the utilization of AGVs and proposes effective scheduling solutions considering real-world environmental constraints, such as limited buffer capacity, battery charging, etc. However, although these studies present scheduling models applying metaheuristics and mathematical methodologies, it can be seen that studies using discrete event-based simulations for AGV scheduling are not clearly addressed. These simulations help model AGV behavior in realistic operating environments and proactively identify potential problems and optimization opportunities. Therefore, the scheduling study with discrete event-based simulation of AGVs is an important study that can show better performance in the real operational environment. Therefore, this paper aims to develop a simulation model for a shopfloor with limited buffer capacity and battery charging and apply it to find the optimal AGV scheduling rules.

3. Multi-AGV Scheduling Rules for Practical Shopfloors

3.1. Practical Manufacturing Shopfloor with Limited Buffer Capacity and Charging Policy

Figure 1 depicts the transportation procedure of AGVs considered in this paper, where AGVs are used to transport a batch of finished goods from the production line to the delivery point. The finished goods are produced in the production line and forwarded to the pickup location. The transportation batches waiting at the pickup location constitute the output buffer of the production line, and they have to wait at the pickup location until an AGV comes and picks them up.

Initially, AGVs stay in charging stations, and one of them is requested to perform the transportation procedure whenever a transportation batch appears at the pickup location. Then, the AGV visits the pickup location to load the batch to transport, travels to the drop-off location, and unloads the batch. After the transportation batch is unloaded, we have to check the start-charge condition, and the AGV returns to a charging station if and only if this condition is met. Otherwise, the AGV continues to perform the transportation procedure for the next transportation batches. In this paper, the start-charge condition is defined by two sub-conditions.

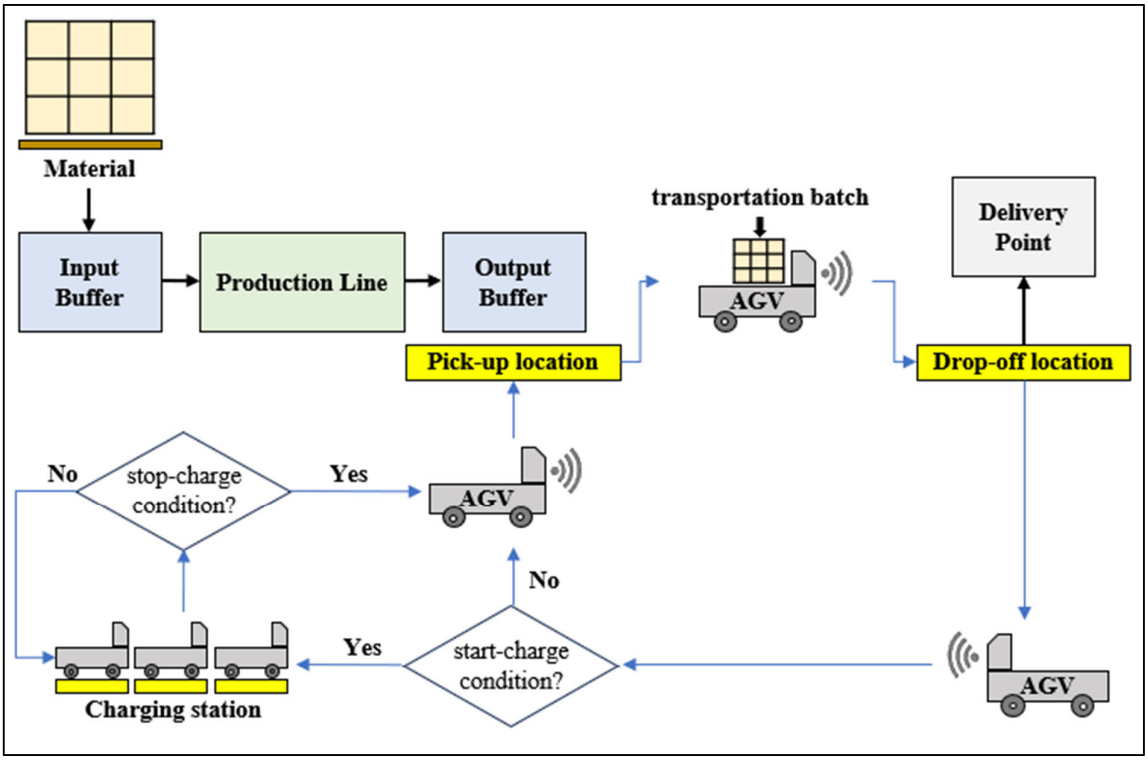


Figure 1. Transportation procedure of AGVs.

Sub-condition 1 for start-charge: every pickup location in the entire shopfloor is empty. Sub-condition 2 for start-charge: the battery level of the AGV is lower than the pre-defined threshold, $T_{charging}$. The start-charge condition is met if one of those two sub-conditions is satisfied. Typically, multi-AGV systems contain multiple charging stations. In this paper, a charging station for an AGV satisfying the start-charge condition is chosen by applying the charging station selection rules listed in Table 1.

Table 1. Charging station selection rules.

Rule	Charging Station to Choose
Dedicated charging station (DCS)	Charging station where the AGV stays at time $t = 0$
Random charging station (RCS)	A random empty charging station
Furthest charging station (FCS)	The furthest empty charging station
Nearest charging station (NCS)	The nearest empty charging station

For the transportation procedure in Figure 1, this paper considers two practical issues: limited buffer capacity and battery charging. First, limited buffer capacity indicates that the capacity of the pickup location is limited. Let L_i denote the ith production line in a given shopfloor, $W(L_i)$ the number of transportation batches in the pickup location for L_i, and $W_{max}(L_i)$ the buffer capacity of L_i. L_i can forward finished goods to the pickup location only if $W(L_i) < W_{max}(L_i)$. In contrast, the manufacturing process of L_i is temporarily halted when $W(L_i) = W_{max}(L_i)$, and L_i can resume its process when $W(L_i) < W_{max}(L_i)$. This constraint reflects the limited space for storing products or items, and it can help mitigate the excessive work-in-process (WIP).

Second, battery charging is also an important issue in practical multi-AGV systems, and the performance of any AGV scheduling approach will be overestimated if battery charging is ignored. The transportation procedure in Figure 1 contains two conditions related to battery charging: start-charge and stop-charge. While the former is explained

above, the latter is used to determine if an AGV at a charging station can receive the transportation request. In other words, an AGV can start the transportation procedure for a transportation batch if and only if it satisfies the stop-charge condition, which is characterized by two sub-conditions as follows:

Sub-condition 1 for stop-charge: there are one or more production lines that have a non-empty output buffer.

Sub-condition 2 for stop-charge: the battery level of the AGV is higher than the pre-defined threshold, $T_{stop-charging}$, and an AGV at the charging station can start the transportation procedure if and only if both of the two sub-conditions are satisfied.

3.2. AGV Scheduling Rule

AGV scheduling involves a comprehensive set of decision-making mechanisms that determine the assignment of jobs to AGVs, the selection of AGVs for specific transport tasks, and the utilization of charging stations. Figure 2 illustrates the overall procedure of multi-AGV scheduling in this paper, which utilizes two kinds of decision rules: the job selection rule and the AGV selection rule. Let job_k denote a transportation job associated with a single transportation batch to be moved from one of the production lines to an arbitrary drop-off location. The job selection rules are used to choose one of the existing transportation jobs, considering factors such as job urgency, priority, pickup location, and drop-off location. The chosen transportation job is assigned to an AGV chosen by applying AGV selection rules. In other words, AGV selection rules are used to choose one of the available AGVs, and factors such as battery level and the current location of AGVs can be considered in identifying the best AGV for a specific job. Note that an AGV is available for a new transportation job if and only if it does not satisfy any of the following sub-conditions:

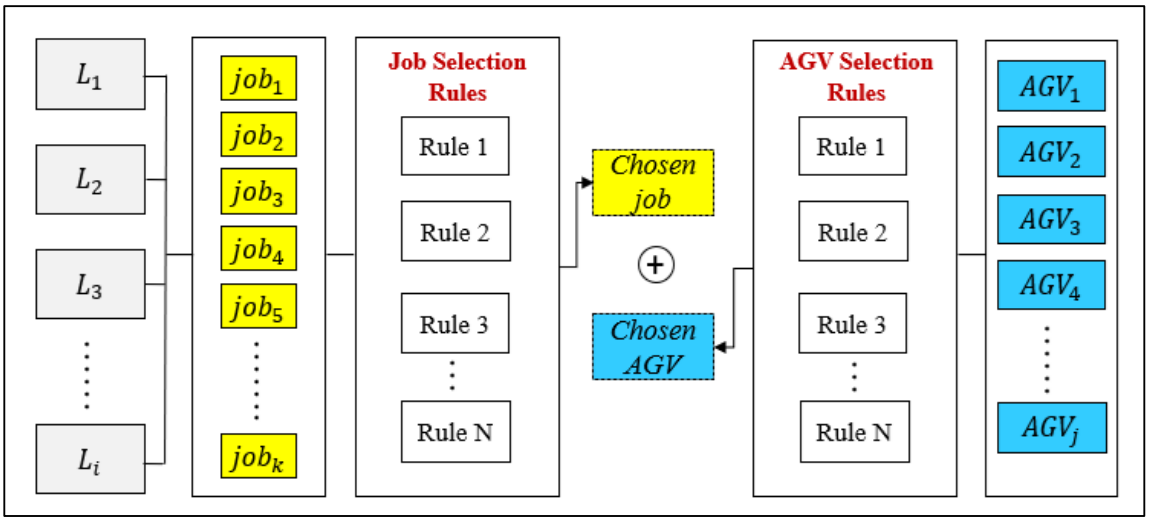


Figure 2. Multi-AGV scheduling procedure.

Sub-condition 1 for an unavailable AGV: The AGV is carrying a previous transportation batch.

Sub-condition 2 for an unavailable AGV: The AGV is empty, and it is heading to a charging station since it satisfies the start-charge condition after unloading the previous transportation batch.

Sub-condition 3 for an unavailable AGV: The AGV is empty and staying at a charging station; however, it does not satisfy the stop-charge condition yet.

Sometimes, a transportation job must wait for an available AGV when all AGVs are unavailable, which can cause a long waiting time for the job and poor service in the multi-AGV system. Thus, job selection and AGV selection rules should be carefully

designed. Examples of the job selection rules and the AGV selection rules are listed in Tables 2 and 3, respectively.

Table 2. Job selection rules.

Rule	Job to Choose
First in first out (FIFO)	A job with the longest waiting time
Longest distance to destination (LDD)	A job with the longest distance between pickup location and drop-off location
Shortest distance to destination (SDD)	A job with the shortest distance between pickup location and drop-off location
Longest queue in buffer (LQB)	A job at pickup location with the highest $W(L_i)$
Shortest queue in buffer (SQB)	A job at pickup location with the lowest $W(L_i)$
Random job (RJ)	A job randomly chosen from $\{job_k\}$

Table 3. AGV selection rules.

Rule	AGV to Choose
Longest idle time (LIT)	An available AGV with the longest idle time after last transportation
Highest battery level (HBL)	An available AGV with the highest battery level
Lowest battery level (LBL)	An available AGV with the lowest battery level
Longest distance to pickup location (LDP)	An available AGV with the longest distance to pickup location
Shortest distance to pickup location (SDP)	An available AGV with the shortest distance to pickup location
Random AGV (RA)	An available AGV randomly chosen from $\left\{AGV_j\right\}$

4. Multi-AGV Scheduling Rules

4.1. Simulation Model for a Multi-AGV System

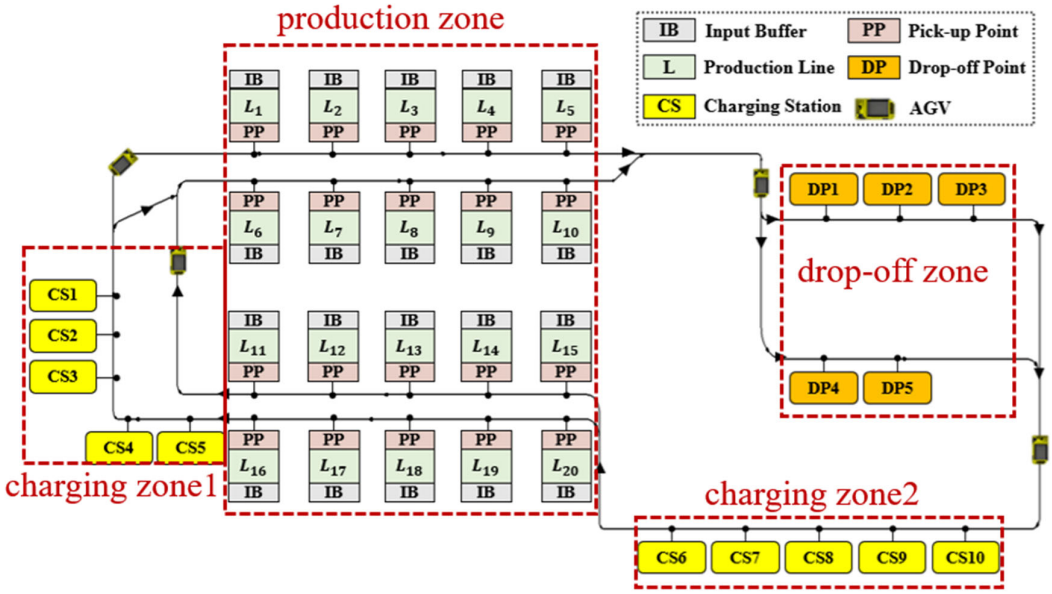
Simulation modeling for a multi-AGV system is a non-trivial task due to the intricate interactions and dependencies among the entities, including production lines and AGVs. Moreover, in order to apply domain-specific issues, such as limited buffer capacity and battery charging, to a simulation model, complex custom logics need to be developed. In this paper, the simulation model for a multi-AGV system is developed using the FlexSim software, which provides rich 3D visualization for manufacturing shopfloors, a built-in library for modeling AGV systems, and a flexible logic builder for developing custom logics. These advantages of the FlexSim software help to model and simulate multi-AGV systems in an efficient manner.

Figure 3a depicts a schematic representation of the shopfloor to be modeled, which can be divided into four zones: production zone, drop-off zone, charging zone 1, and charging zone 2. This structure for a shopfloor is motivated by an automobile part manufacturer in South Korea.

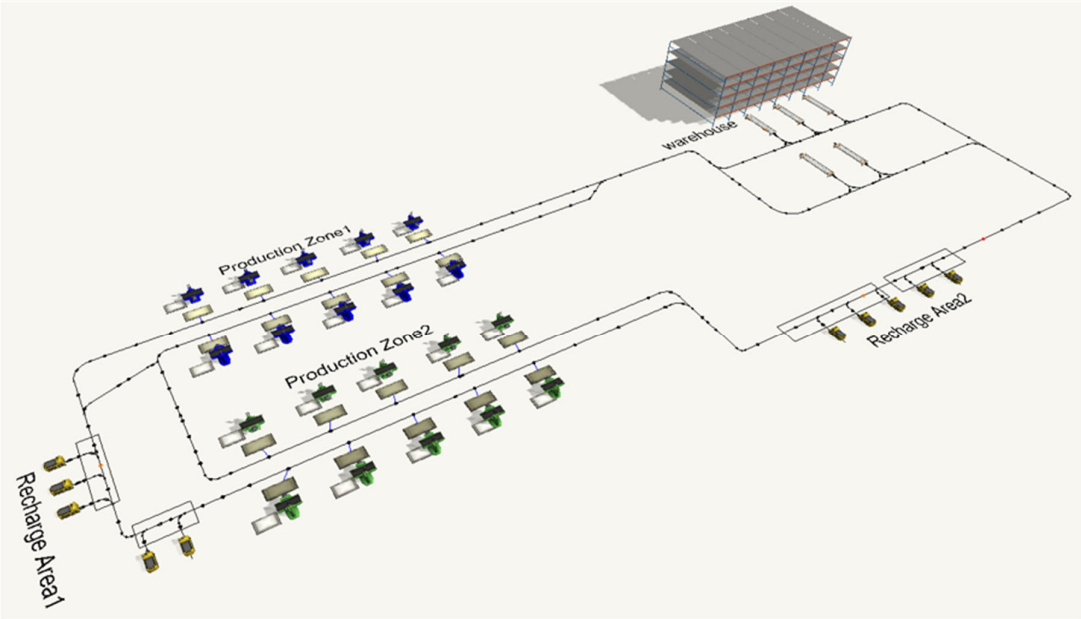
The production zone contains 20 production lines, and each production line has its own pickup point. The transportation batches at pickup points need to be transported to one of five drop-off points within the drop-off zone by AGVs. We have 10 AGVs and 10 charging stations, and every AGV stays at a charging station at time $t = 0$. The charging stations are grouped into two charging zones, charging zone 1 and charging zone 2, where each charging zone has five charging stations. Note that the path lines are directed, and this can affect the proximity between a production line and a charging zone. For instance, an AGV at charging zone 2 can move to L_{11} more quickly than other AGVs at charging zone 1. The drop-off zone contains five drop-off points. Moreover, Figure 3b shows the layout of the simulation model developed using the FlexSim software.

Let S_{job} denote the set of transportation jobs to be completed, S_{AGV} the set of available AGVs, and S_{CS} the set of empty charging stations. During the simulation experiment, new elements can be added to these sets. For instance, a new transportation job will be added

to S_{job} whenever a production line produces a transportation batch. Sometimes, an existing element can be deleted from the sets. Assume that a transportation job is assigned to an AGV. Then, the AGV must be deleted from S_{AGV}.



(a) Schematic representation of shopfloor



(b) FlexSim model layout

Figure 3. Layout of shopfloor.

This paper proposes a multi-AGV scheduling procedure summarized in Figure 4. The following three kinds of rules must be determined by the analyzer before the simulation experiment: the job selection rule r_{job}, the AGV selection rule r_{AGV}, and the charging station selection rule r_{CS}. At first, a job is chosen from S_{job} by applying r_{job}. If $n(S_{job}) = 0$, we have to wait until a new transportation batch is generated. After a job is chosen, an AGV is chosen from S_{AGV} by applying r_{AGV}. If there is no available AGV, we have to wait for an available one. Once an AGV is chosen, the chosen job is assigned to the AGV.

When the AGV unloads the associated transportation batch at the drop-off point, we have to check the start-charge condition before handling the next transportation job. If the condition is met, a charging station is chosen from S_{CS} by applying r_{CS}, and the

AGV travels to the chosen station. An AGV staying at a charging station can perform a transportation task if and only if it satisfies the stop-charge condition. Note that the performance of the proposed multi-AGV scheduling procedure is affected by the following three kinds of rules: r_{job}, r_{AGV}, and r_{CS}.

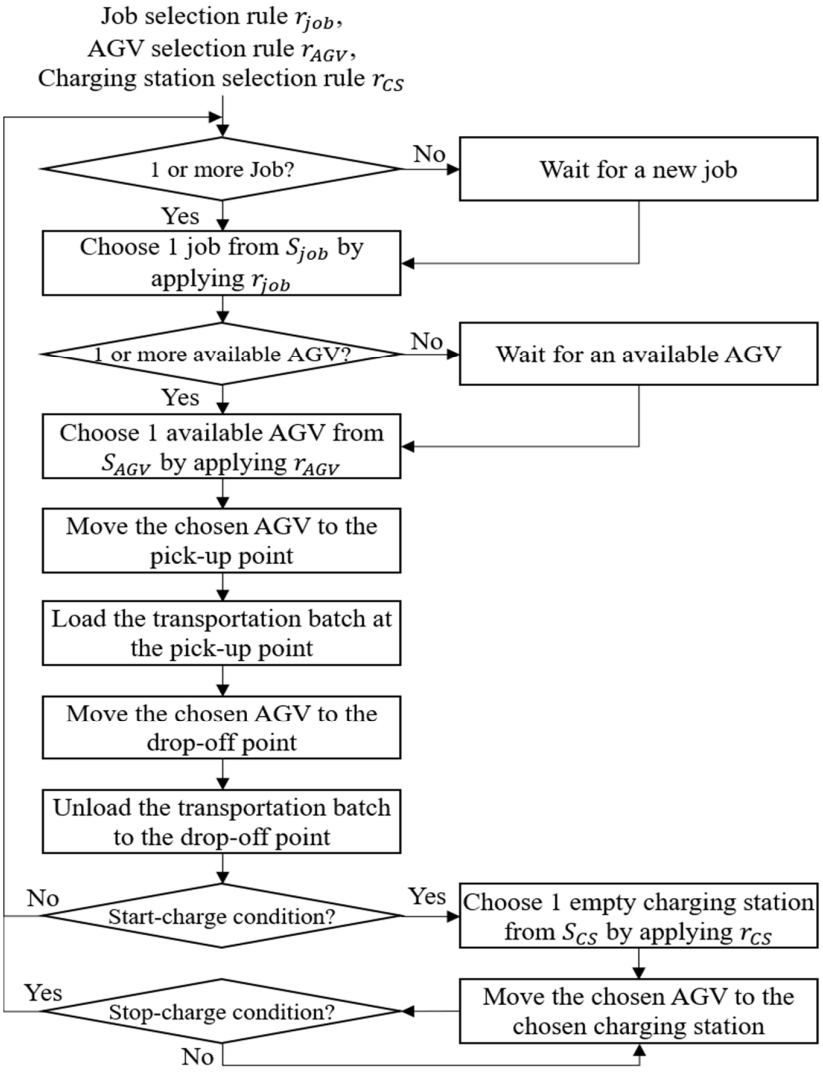


Figure 4. Overall multi-AGV scheduling procedure.

In addition, the following model parameters and constraints are applied to the simulation model:

1. Each production line produces five types of products, and the processing time of each type can be found in Table 4. Each type accounts for 20% of the entire product.
2. The raw materials arrive at the input buffer of each production line. The capacity of the input buffer is unlimited, and the inter-arrival times of raw materials are uniformly distributed between 100 and 200 (s).
3. The finished products are forwarded to the output buffer with limited capacity $W_{max}(L_i)$. L_i stops operation when $W(L_i) = W_{max}(L_i)$.
4. An AGV can transport only one product at a time. That is, a product is a transportation batch in the simulation model.
5. The detailed parameters of AGV are as follows:
    - Speed on the straight section of a path line: 1.8 m/s;
    - Speed on the curved section of a path line: 1.5 m/s;
    - Acceleration/deceleration: 1 m/s^2;

- Battery capacity: 100 (Ah);
- Battery C-rate when busy (discharging): 0.25 C;
- Battery C-rate when idle (discharging): 0.005 C;
- Battery C-rate when charging: 0.5 C;
- Pickup and drop-off time: 5 (s);

6. Two or more AGVs cannot occupy a single charging station at a time.
7. Overtaking of the AGV is not allowed.
8. The shopfloor is operating 24 h a day with regular break periods (12:00–13:00, 18:00–19:00). During the break period, the production lines do not operate, but AGVs can perform transportation tasks if $n(S_{job}) > 0$. Consequently, $W(L_i)$s tends to be small at the end of a break period.
9. The simulation end time for a single experiment is 604,800 s (7 days).

Table 4. Processing time by item type.

Type	Processing Time (s)
1	100
2	120
3	140
4	160
5	180

4.2. Experiment Results

The performance of the proposed multi-AGV scheduling approach is evaluated through a multi-scenario simulation experiment, as depicted in Figure 5. One simulation scenario is defined by combining a job selection rule, an AGV selection rule, and a charging station selection rule. Thus, we have 144 scenarios, where 10 replications of simulation experiments are conducted for each scenario. The performance measure is the average waiting time (AWT) of transportation batches, and the objective of the experiment is to find the best scenario that minimizes the AWT.

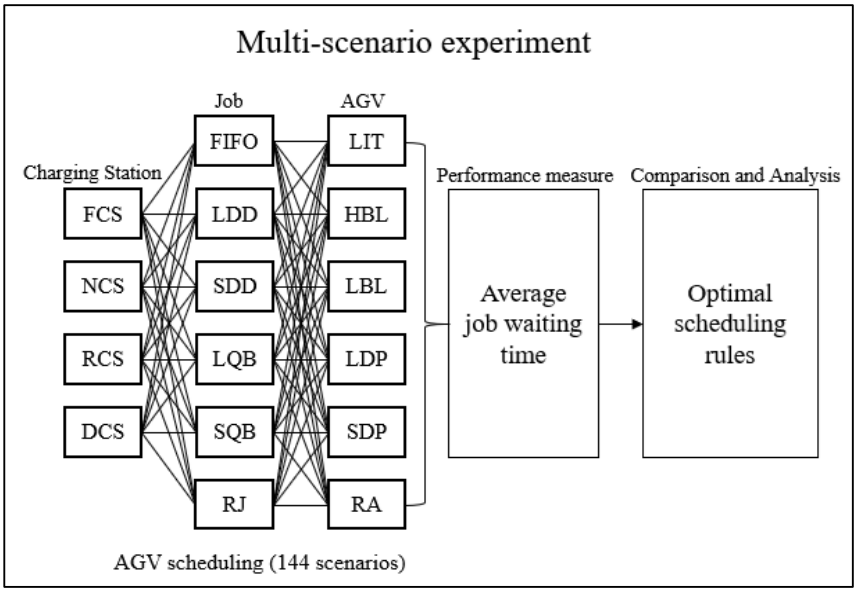


Figure 5. Multi-scenario experiment procedure.

Figure 6 presents a boxplot graph, where one boxplot indicates the distribution of AWTs for a single scenario. The 144 boxplots are divided into 24 groups, where each group

has six scenarios. Note that the six scenarios in a single group use identical charging station selection rules and job selection rules, while their AGV selection rules are different from each other.

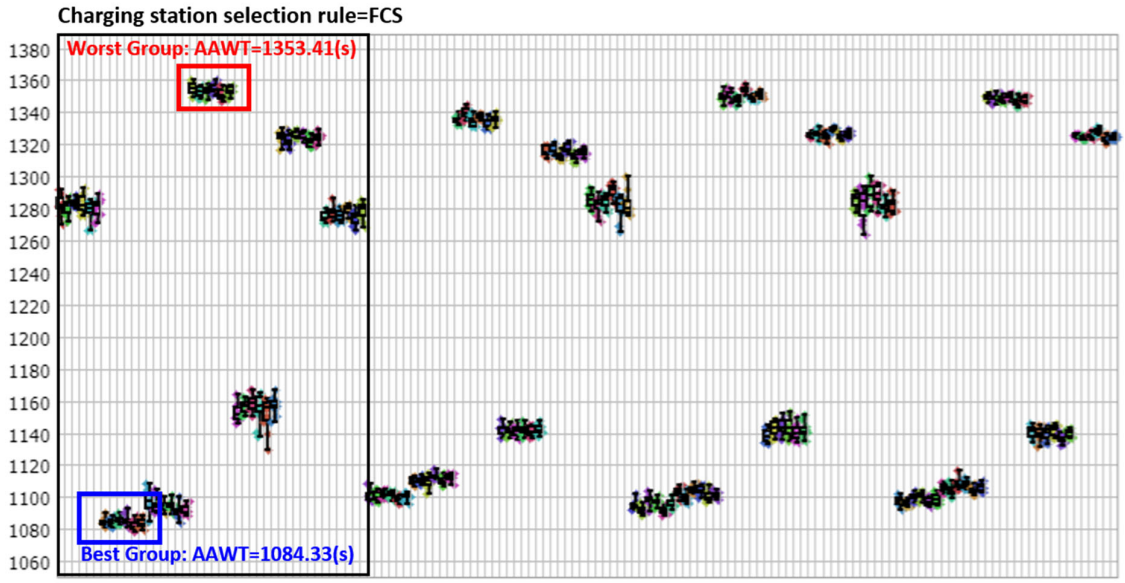


Figure 6. Distribution of AWTs of entire scenarios.

From Figure 6, we can make several observations: First, there are significant variations in the AWTs of the given scenarios, which means that the scheduling rules for a specific multi-AGV system in a practical shopfloor must be carefully chosen. For example, the second-left group containing scenario 7~scenario 12, which uses FCS and LDD, showed the best performance in terms of the AWT. The average AWT (AAWT) of those six scenarios is 1084.33 (s). In contrast, the worst group is the second-fourth one, which uses FCS and LQB, and the AAWT of this group is 1353.41 (s).

Second, the variations in the AAWT of the six scenarios in a single group are not large in that the boxplots for them are at similar heights, as can be seen in Figure 6. Which means that the six types of AGV selection rules considered in this paper have relatively small impacts on the AWT.

Third, the charging station selection rule has relatively small impacts on the AWT, too. In Figure 6, every six consecutive groups use the identical charging selection rule. For example, the first six groups choose the charging station by using the FCS rule. Every six consecutive groups shows similar patterns, where the second and fourth groups are the worst and the best groups, respectively. This means that the AWT is not much affected by the charging station selection rule.

Fourth, the job selection rule has significant impacts on the AWT. The six scenarios of a group use identical job selection rules, and Figure 6 shows that the AWT of a group is highly dependent on the type of job selection rule.

The details of the AAWTs of the entire scenarios are summarized in Appendix A, where Tables A1–A4 summarizes the performances of the six consecutive groups that use the identical charging stations rule. Note that the first columns of those tables specify the type of charging station selection rule with the AAWT of the cluster of six consecutive groups. Similarly, the second column specifies the type of job selection rule with the AAWT of the associated group in the cluster. Finally, the type of AGV selection rule and the AAWT of the associated scenario can be found in the third and fourth columns, respectively.

In order to further investigate the performances of the three kinds of AGV scheduling rules, the performances of alternatives for a specific kind of AGV scheduling rule are compared in Figures 7–9. For instance, 24 out of 144 scenarios use the LDD rule for job selection, and Figure 8 shows that the AAWT of them is 1094.67 (s).

Figure 7 indicates that the AAWT value varies slightly depending on the type of charging station selection rules. Among the four alternatives, NCS (AAWT = 1213.55 (s)) and DCS (AAWT = 1217.17 (s)) are the best and worst ones; however, the difference between their performances was not large, which suggests that the charging station selection rule has limited impacts on the AWT.

In contrast, the impact of the job selection rule is evident, as shown in Figure 8. For example, LDD (AAWT = 1094.67 (s)) and SDD (AAWT = 1103.41 (s)) rules produce a low AWT, indicating their usefulness for reducing the AWT. On the other hand, RJ (AAWT = 1322.74 (s)) and LQB (AAWT = 1347.19 (s)) rules show poor performances. In Figure 9, we can see that SDP (AAWT = 1213.47 (s)) and LDP (AAWT = 1217.80 (s)) are the best and worst rules for AGV selection, respectively. However, the deviation in their performances is not significant.

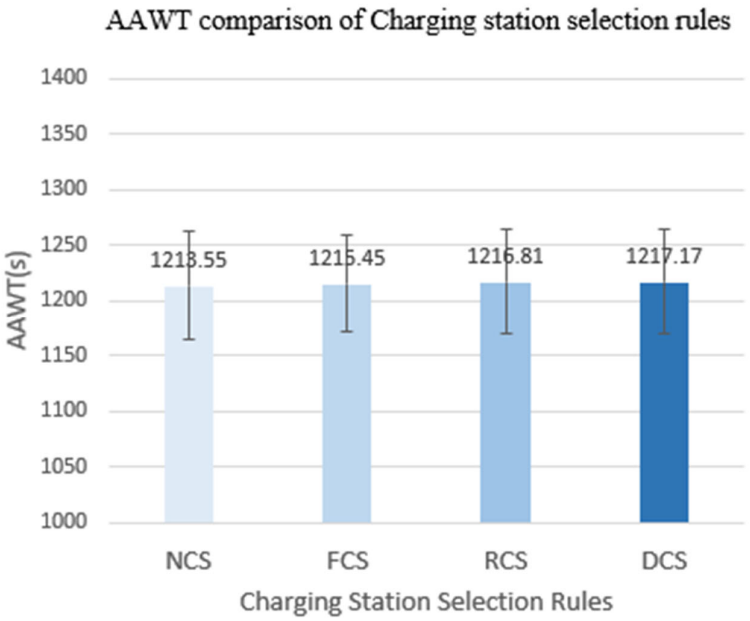


Figure 7. AAWT by type of charging station selection rules.

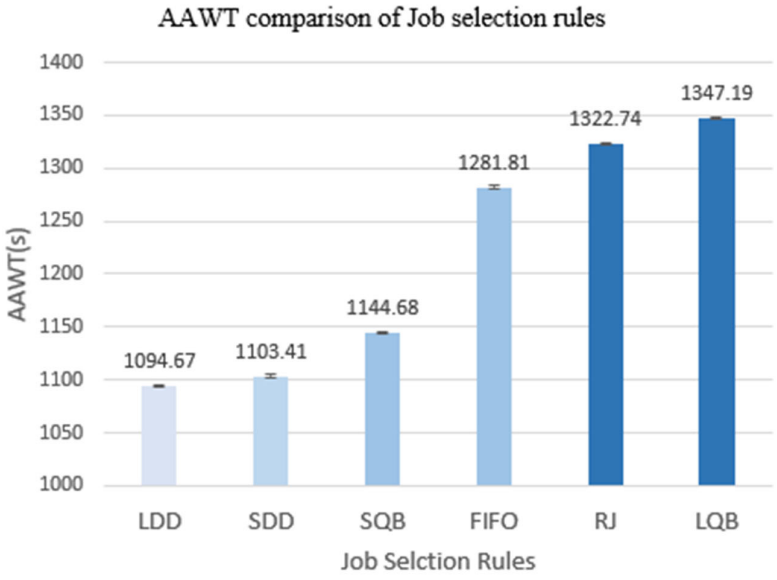


Figure 8. AAWT by type of job selection rules.

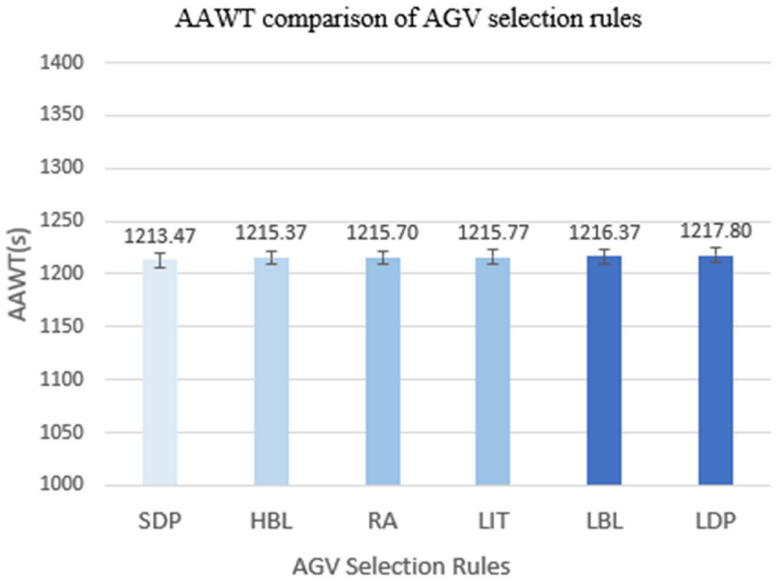


Figure 9. AAWT by type of AGV selection rules.

Note that the AAWT was minimized under charging station selection rule = FCS, job selection rule = LDD, and AGV selection rule = SDP, as shown in Table A1. LDD is the best rule for job selection, as shown in Figure 8. Similarly, Figure 9 shows that SDP is the best AGV selection rule. On the contrary, FCS is not the best charging station rule, as shown in Figure 7. This means that AGV scheduling rules must be determined by conducting a multi-scenario simulation experiment. In other words, a combination of the best rules in Figures 7–9 does not guarantee optimal AGV scheduling.

5. Conclusions and Further Remarks

In this paper, we propose a simulation-based AGV scheduling approach for multi-AGV systems within practical shopfloors with limited buffer capacity and battery charging. The proposed approach considers the following three kinds of scheduling rules: job selection, AGV selection, and charging station selection rules; the optimal combination of those rules can be determined through multi-scenario simulation experiments. In this paper, the scheduling approach was applied to minimize the AWT of the transportation batches. The experimental results led to several important conclusions.

First, notable variations in the AWT were observed based on the applied scheduling rule. Experimental results confirmed that applying the optimal rule set improved performance by more than 25% compared to applying the worst rule set. In other words, this means that in a multi-AGV system environment, scheduling rules must be carefully selected considering the shopfloor operating situation and requirements.

Second, it is worth noting that the choice of charging station rules and AGV selection rules had relatively small effects on the AWT compared to the impact of job selection rules. This is probably because job selection rules have a greater impact on the overall efficiency of the AGV system, whereas charging station selection rules and AGV selection rules have a relatively limited impact. This is partly inconsistent with a previous study by Le-Ahn and van der Meer (2004), which suggests that the SDP rule for AGV selection is quite helpful in reducing the AWT. Note that only the AGV selection rule was considered in the paper, while this paper combines three kinds of rules for multi-AGV scheduling. Thus, the scheduling approach proposed in this paper can make a more flexible decision about AGV scheduling. On the contrary, the experimental result of this paper is partly consistent with the previous study in that SDP is the best AGV selection rule, as shown in Figure 9; however, its impact on the AWT is limited.

Additionally, because the charging station selection rules and AGV selection rules are mainly related to the battery status of AGVs, their impact on the AWT may be relatively

small compared to the overall work efficiency of AGVs. On the other hand, job selection rules determine the job order and priority of AGVs and are therefore directly linked to the overall work efficiency, which has a greater impact on the AWT. However, this does not mean that charging station selection rules and AGV selection rules are not important. These two rules remain important factors related to battery management in AGVs and are therefore essential to maintaining the efficiency of the overall system. In the context of energy efficiency or battery-related performance metrics, the criteria for selecting AGVs or charging stations may have greater significance than the rules governing job selection.

Third, it is important to note that achieving the optimal combination of charging stations, job selection, and rules for AGV selection does not automatically ensure the most efficient AGV schedule. The average AWT of the entire scenario was recorded at 1084.33 (s), with the best outcome observed when the charging station selection rule was set to FCS, the job selection rule to LDD, and the AGV selection rule to SDP. However, Figures 7–9 indicate that LDD and SDP are preferable for job selection and AGV selection rules, respectively, whereas FCS is not the optimal selection. In other words, the combination of rules that seems optimal under specific conditions may not necessarily generalize well to different operational scenarios. These implications underscore the importance of an integrated approach that considers various variables and situations. In conclusion, we believe that it is desirable to evaluate the performance of AGV scheduling rules through multiple scenario experiments.

This paper will provide practical insights into AGV scheduling. However, the proposed scheduling approach has some limitations. First, the types of scheduling rules are limited in this paper. For instance, only four types of charging station rules are used to define an experiment scenario. More diverse types of job selection, AGV selection, and charging station selection rules might be useful to optimize the scheduling policy. Second, the proposed scheduling approach is static in that it does not change the types of scheduling rules during the operation of the shopfloor. The performance measure might be further improved if the types of scheduling rules can be changed dynamically.

Third, only one performance measure, the AWT, is considered in this paper, while various performance measures can be used to evaluate the performance of the AGV system. For instance, job tardiness, machine utilization, or battery charging-related measures can be more useful in some shopfloors. Fourth, this paper assumes the number of charging stations equals the number of AGVs, which might lead to little impact of charging station selection rules on the performance of the AGV system. The number of charging stations and their locations are another future research topic related to multi-AGV systems. Therefore, we plan to study a revised version of the proposed AGV scheduling approach to overcome the above limitations.

Author Contributions: Conceptualization, J.-W.K.; methodology, J.-S.P. and J.-W.K.; software, J.-S.P.; validation, J.-S.P. and J.-W.K.; formal analysis, J.-S.P. and J.-W.K.; investigation, J.-S.P.; data curation, J.-S.P. and J.-W.K.; writing—original draft preparation, J.-S.P. and J.-W.K.; writing—review and editing, J.-S.P. and J.-W.K.; visualization, J.-S.P.; supervision, J.-W.K.; project administration, J.-W.K.; funding acquisition, J.-W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A03093301).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. AAWT under charging station selection rule = FCS (Scenario 1 to 36).

Charging Station Selection Rule	Job Selection Rule	AGV Selection Rule	AAWT (s)
FCS (AAWT = 1215.45)	FIFO (AAWT = 1281.35)	LIT	1282.26
		HBL	1280.98
		LBL	1283.28
		LDP	1282.85
		SDP	1278.29
		RA	1280.46
	LDD (AAWT = 1084.33)	LIT	1084.54
		HBL	1084.08
		LBL	1085.67
		LDP	1085.59
		SDP	**1082.25 ****
		RA	1083.89
	SDD (AAWT = 1094.14)	LIT	1097.5
		HBL	1094.79
		LBL	1094.14
		LDP	1095.02
		SDP	1090.61
		RA	1092.78
	LQB (AAWT = 1353.41)	LIT	1354.97
		HBL	1352.72
		LBL	1353.2
		LDP	1354.65
		SDP	1351.23
		RA	1353.68
	SQB (AAWT = 1155.29)	LIT	1152.7
		HBL	1156.4
		LBL	1158.9
		LDP	1155.03
		SDP	1151.67
		RA	1157.07
	RJ (AAWT = 1324.19)	LIT	1324.79
		HBL	1322.71
		LBL	1326.1
		LDP	1325.61
		SDP	1321.67
		RA	1324.27

****: Global optimal AAWT for overall charging station selection rules.**

Table A2. AAWT under charging station selection rule = NCS (Scenario 37 to 72).

Charging Station Selection Rule	Job Selection Rule	AGV Selection Rule	AAWT (s)
NCS (AAWT = 1213.55)	FIFO (AAWT = 1276.63)	LIT	1274.83
		HBL	1277.45
		LBL	1275.66
		LDP	1279.02
		SDP	1274.36
		RA	1278.47
	LDD (AAWT = 1100.77)	LIT	1101.01
		HBL	1100.45
		LBL	1102.16
		LDP	1101.85
		SDP	**1098.83 ***
		RA	1100.35
	SDD (AAWT = 1111.02)	LIT	1110.09
		HBL	1110.5
		LBL	1109.72
		LDP	1113.11
		SDP	1110.25
		RA	1112.42
	LQB (AAWT = 1336.05)	LIT	1335.23
		HBL	1339.64
		LBL	1334.49
		LDP	1337.42
		SDP	1333.85
		RA	1335.68
	SQB (AAWT = 1141.82)	LIT	1142.25
		HBL	1141.83
		LBL	1141.41
		LDP	1142
		SDP	1141.04
		RA	1142.39
	RJ (AAWT = 1315.03)	LIT	1315.86
		HBL	1316.71
		LBL	1312.79
		LDP	1317.7
		SDP	1312.15
		RA	1315

***: Local optimal AAWT for a single charging station selection rule.**

Table A3. AAWT under charging station selection rule = RCS (Scenario 73 to 108).

Charging Station Selection Rule	Job Selection Rule	AGV Selection Rule	AAWT (s)
RCS (AAWT = 1216.81)	FIFO (AAWT = 1284.52)	LIT	1285.9
		HBL	1281.97
		LBL	1284.54
		LDP	1291.06
		SDP	1280.81
		RA	1282.85
	LDD (AAWT = 1095.07)	LIT	1095.14
		HBL	1091.98
		LBL	1098.28
		LDP	1097.06
		SDP	**1091.88 ***
		RA	1096.08
	SDD (AAWT = 1102.32)	LIT	1103.01
		HBL	1099.95
		LBL	1104.61
		LDP	1104.61
		SDP	1100.54
		RA	1101.21
	LQB (AAWT = 1350.57)	LIT	1350.38
		HBL	1350.55
		LBL	1348.05
		LDP	1354.05
		SDP	1349.19
		RA	1351.21
	SQB (AAWT = 1141.9)	LIT	1138.23
		HBL	1142.93
		LBL	1144.14
		LDP	1143.89
		SDP	1140.33
		RA	1141.87
	RJ (AAWT = 1326.48)	LIT	1325.91
		HBL	1326.82
		LBL	1325.59
		LDP	1329.13
		SDP	1324.97
		RA	1326.44

***: Local optimal AAWT for a single charging station selection rule.**

Table A4. AAWT under charging station selection rule = DCS (Scenario 109 to 144).

Charging Station Selection Rule	Job Selection Rule	AGV Selection Rule	AAWT (s)
DCS (AAWT = 1217.17)	FIFO (AAWT = 1284.72)	FCFS	1285.22
		HBL	1282.27
		LBL	1288.19
		LDP	1288.44
		SDP	1281.38
		RA	1282.79
	LDD (AAWT = 1098.51)	FCFS	1097.55
		HBL	**1096.26** *
		LBL	1100.23
		LDP	1101.09
		SDP	1097.22
		RA	1098.69
	SDD (AAWT = 1106.14)	LIT	1105.79
		HBL	1104.38
		LBL	1108.5
		LDP	1107.99
		SDP	1103.88
		RA	1106.3
	LQB (AAWT = 1348.71)	LIT	1349.45
		HRB	1349.37
		LRB	1348.22
		FPL	1350.18
		SPL	1346.47
		RAS	1348.56
	SQB (AAWT = 1139.71)	LIT	1140.2
		HRB	1139.73
		LRB	1139.68
		FPL	1141.71
		SPL	1136.98
		RAS	1139.92
	RJ (AAWT = 1325.27)	LIT	1325.66
		HRB	1324.47
		LRB	1325.31
		FPL	1328.2
		SPL	1323.51
		RAS	1324.46

***: Local optimal AAWT for a single charging station selection rule.**

References

1. Mehrabian, A.; Tavakkoli-Moghaddam, R.; Khalili-Damaghani, K. Multi-objective routing and scheduling in flexible manufacturing systems under uncertainty. *Iran. J. Fuzzy Syst.* **2017**, *14*, 45–77.
2. Witczak, M.; Majdzik, P.; Stetter, R.; Lipiec, B. Multiple AGV fault-tolerant within an agile manufacturing warehouse. *IFAC-PapersOnLine* **2019**, *52*, 1914–1919. [CrossRef]

3. Zhang, X.J.; Sang, H.Y.; Li, J.Q.; Han, Y.Y.; Duan, P. An effective multi-AGVs dispatching method applied to matrix manufacturing workshop. *Comput. Ind. Eng.* **2022**, *163*, 107791. [CrossRef]
4. Okumuş, F.; Dönmez, E.; Kocamaz, A.F. A cloudware architecture for collaboration of multiple agvs in indoor logistics: Case study in fabric manufacturing enterprises. *Electronics* **2020**, *9*, 2023. [CrossRef]
5. Zając, J.; Małopolski, W. Structural on-line control policy for collision and deadlock resolution in multi-AGV systems. *J. Manuf. Syst.* **2021**, *60*, 80–92. [CrossRef]
6. Draganjac, I.; Miklić, D.; Kovačić, Z.; Vasiljević, G.; Bogdan, S. Decentralized control of multi-AGV systems in autonomous warehousing applications. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 1433–1447. [CrossRef]
7. Zacharia, P.T.; Xidias, E.K. AGV routing and motion planning in a flexible manufacturing system using a fuzzy-based genetic algorithm. *Int. J. Adv. Manuf. Technol.* **2020**, *109*, 1801–1813. [CrossRef]
8. Vlachos, I.; Pascazzi, R.M.; Ntotis, M.; Spanaki, K.; Despoudi, S.; Repoussis, P. Smart and flexible manufacturing systems using Autonomous Guided Vehicles (AGVs) and the Internet of Things (IoT). *Int. J. Prod. Res.* **2022**, *2022*, 2136282. [CrossRef]
9. Jin, J.; Zhang, X.H. Multi agv scheduling problem in automated container terminal. *J. Mar. Sci. Technol.* **2016**, *24*, 5.
10. Fazlollahtabar, H.; Hassanli, S. Hybrid cost and time path planning for multiple autonomous guided vehicles. *Appl. Intell.* **2018**, *48*, 482–498. [CrossRef]
11. Małopolski, W. A sustainable and conflict-free operation of AGVs in a square topology. *Comput. Ind. Eng.* **2018**, *126*, 472–481. [CrossRef]
12. Zou, W.Q.; Pan, Q.K.; Meng, T.; Gao, L.; Wang, Y.L. An effective discrete artificial bee colony algorithm for multi-AGVs dispatching problem in a matrix manufacturing workshop. *Expert Syst. Appl.* **2020**, *161*, 113675. [CrossRef]
13. Dang, Q.V.; Singh, N.; Adan, I.; Martagan, T.; van de Sande, D. Scheduling heterogeneous multi-load AGVs with battery constraints. *Comput. Oper. Res.* **2021**, *136*, 105517. [CrossRef]
14. Liu, C.I.; Ioannou, P.A. A comparison of different AGV dispatching rules in an automated container terminal. In Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems, Singapore, 6 September 2002; pp. 880–885.
15. Koo, P.H.; Jang, J. Vehicle travel time models for AGV systems under various dispatching rules. *Int. J. Flex. Manuf. Syst.* **2002**, *14*, 249–261. [CrossRef]
16. Ho, Y.C.; Liu, H.C. A simulation study on the performance of pickup-dispatching rules for multiple-load AGVs. *Comput. Ind. Eng.* **2006**, *51*, 445–463. [CrossRef]
17. de Koster, R.M.B.; Le-Anh, T.; van der Meer, J.R. Testing and classifying vehicle dispatching rules in three real-world settings. *J. Oper. Manag.* **2004**, *22*, 369–386. [CrossRef]
18. Le-Anh, T.; De Koster, M.B.M. A review of design and control of automated guided vehicle systems. *Eur. J. Oper. Res.* **2006**, *171*, 1–23. [CrossRef]
19. Wang, Z.; Wu, Y. An Ant Colony Optimization-Simulated Annealing Algorithm for Solving a Multiload AGVs Workshop Scheduling Problem with Limited Buffer Capacity. *Processes* **2023**, *11*, 861. [CrossRef]
20. Wang, Z.; Hu, J.; Wu, Y.; Wang, Y. Improved Simulated Annealing Algorithm for Scheduling of Multi-load AGV Workshop with Limited Buffer Capacity. In Proceedings of the 2022 4th International Academic Exchange Conference on Science and Technology Innovation (IAECST), Guangzhou, China, 9–11 December 2022; pp. 1509–1515.
21. Naeem, D.; Eltawil, A.; Iijima, J.; Gheith, M. Integrated Scheduling of Automated Yard Cranes and Automated Guided Vehicles with Limited Buffer Capacity of Dual-Trolley Quay Cranes in Automated Container Terminals. *Logistics* **2022**, *6*, 82. [CrossRef]
22. Li, J.; Cheng, W.; Lai, K.K.; Ram, B. Multi-AGV Flexible Manufacturing Cell Scheduling Considering Charging. *Mathematics* **2022**, *10*, 3417. [CrossRef]
23. Yang, X.; Hu, H.; Cheng, C.; Wang, Y. Automated Guided Vehicle (AGV) Scheduling in Automated Container Terminals (ACTs) Focusing on Battery Swapping and Speed Control. *J. Mar. Sci. Eng.* **2023**, *11*, 1852. [CrossRef]
24. Sun, B.; Zhai, G.; Li, S.; Pei, B. Multi-resource collaborative scheduling problem of automated terminal considering the AGV charging effect under COVID-19. *Ocean Coast. Manag.* **2023**, *232*, 106422. [CrossRef] [PubMed]
25. Singh, N.; Dang, Q.V.; Akcay, A.; Adan, I.; Martagan, T. A matheuristic for AGV scheduling with battery constraints. *Eur. J. Oper. Res.* **2022**, *298*, 855–873. [CrossRef]
26. Abderrahim, M.; Bekrar, A.; Trentesaux, D.; Aissani, N.; Bouamrane, K. Manufacturing 4.0 operations scheduling with AGV battery management constraints. *Energies* **2020**, *13*, 4948.

applied sciences

MDPI

Article

Machine Learning Applied to Logistics Decision Making: Improvements to the Soybean Seed Classification Process

Djonathan Luiz de Oliveira Quadras [1,*], Ian Cavalcante [2], Mirko Kück [3], Lúcio Galvão Mendes [1] and Enzo Morosini Frazzon [1]

[1] Graduate Program in Production Engineering, Federal University of Santa Catarina, Florianopolis 88040-970, SC, Brazil; enzo.frazzon@ufsc.br (E.M.F.)
[2] Neosilos, Curitiba 81280-340, PR, Brazil
[3] Faculty of Production Engineering, University of Bremen, 28359 Bremen, Germany
* Correspondence: djonquadras@gmail.com; Tel.: +55-479-9626-1051

Abstract: Soybean seed classification is a relevant and time-consuming process for Brazilian agribusiness cooperatives. This activity can generate queues and waiting times that directly affect logistics costs. This is the reason why it is so important to properly allocate resources, considering the most relevant factors that can influence their performance. This paper aims to present an approach to predicting the average lead time and waiting queue time for the soybean seed classification process, which supports the decision regarding the number of workers and machines to be deployed in the process. The originality of the paper relies on the applied approach, which combines discrete event simulation with machine learning algorithms in a real-world applied case. The approach comprises three steps: data collection to structure the simulation scenarios; simulation runs to generate artificial historical data; and machine learning applications to predict lead and queuing times. As a result, various scenarios using the data generated by machine learning were simulated, making it possible to choose the one that generated the best trade-off between performance, investments, and operational costs. The approach can be adapted to support the solution of different logistic-related decision-making problems that combine human and equipment resources.

Keywords: machine learning; discrete event simulation; agribusiness 4.0

1. Introduction

The fourth industrial revolution, so-called Industry 4.0, has changed several sectors of the economy. The data generated in every process, which used to be a differentiator, is now necessary to maintain competitiveness in the market [1]. Besides the industrial scenario, agribusiness can also take advantage of the new technologies. Genetic engineering of plants, nanotechnology, biometric sensing, electrical agricultural machinery, computer vision integrating robotics with artificial intelligence (AI), and blockchain are examples of emerging technologies that have the potential to address challenges in all field states of the agricultural chain, pre-field, in-field, and post-field [2]. Agriculture 4.0 is based on principles such as data-driven decision-making, connectivity of devices, and growth of productivity, with added goals of adaptation to climate change and reduction of food waste [3–5]. Digital transformation can help farmers forecast the weather, control plagues, control temperature, and control moisture as well as uncountable other applications [6,7].

Capacity planning was approached by the examination of capacity investment decisions employing a multiperiod model to study the optimal one-time processing and storage capacity investment decisions [8] or for lot-sizing and pricing when supply and price-sensitive demand are uncertain [9]. Moreover, ref. [10] developed a capacity planning model for transport to estimate the number of locomotives and shifts, the number of bins, and the delays to harvesting operations resulting from harvesters waiting for bin deliveries. Nevertheless, machine learning algorithms could excel when facing decision-making

problems based on historical data [11]. According to [12], "machine learning describes the capacity of systems to learn from problem-specific training data to automate the process of analytical model building and solve associated tasks". It can forecast, cluster, and classify data. Thus, it can generate knowledge that helps decision-makers with problems such as how many machines to buy, how many employers to hire, or how many products to manufacture in the next month [13].

Machine learning methods have been successfully applied to solve several problems in production and logistics, such as forecasting customer demands [14], predicting energy consumption [15,16], or making travel time predictions [17]. In agribusiness, machine learning technique applications include crop yield production [18], predicting soil properties [19], irrigation management [20], weather prediction [21], crop quality [22], harvesting [23], demand prediction [24], detecting vegetable diseases [25], and determining crop production [26]. For transportation, machine learning algorithms were applied for the vehicle routing problem [27–29] and to model the potential distribution of different plant species [30].

Nevertheless, simulation was applied to agribusiness to understand factors influencing the long-term viability of an intermediated regional food supply network [31]; to analyze firms' choices of spatial pricing policy [32]; to explore a hypothesis regarding the adoption of capabilities for entrepreneurship [33]; to model human behavior in food supply chains with asymmetric information about food quality and food safety [34]; to understand how technology, market dynamics, environmental change, and policy intervention affect a heterogeneous population of farm households and resources [35]; to analyze how adaptation affects the distribution of household food security and poverty under current climate and price variability [36]; to understand how government payments to enhance public values in social-ecological systems can contribute to the resilience of the system [37]; to explore nutrient mitigation potentials of different policy instruments [38]. Additionally, some studies use simulation to outline different what-if scenarios, e.g., for farmer decision-making on crop choice, fertilizer, and pesticide usage [39] or to identify global change impacts on farmland abandonment and test policy and management options [40].

In 2020, Brazil's agribusiness sector played a significant role, accounting for a substantial 26.7% of the country's Gross Domestic Product (GDP). Within this sector, the farming segment alone contributed 7% to the nation's overall economic output. Notably, soybean production has taken center stage in the modernization of Brazilian agriculture over recent decades [41]. Back in 2007, the Brazilian Ministry of Agriculture, Livestock, and Supply (MAPA) established the standard governing the classification of soybean seeds [42]. This classification process precision is of the utmost importance, as it serves a dual purpose. Firstly, it ensures that soybean products maintain the highest quality standards. Secondly, it plays a crucial role in meeting food safety regulations [43]. These regulations are imperative for upholding product integrity and safeguarding consumer health. As a result, this task necessitates a comprehensive and dedicated approach to guarantee the accurate pricing of the highest-quality soybean seeds.

While a cooperative operates as a non-profit organization, its foremost objective remains the equitable distribution of benefits among its members, thus underscoring the importance of achieving operational excellence [44]. Agro-industrial cooperatives confront notable challenges in managing their operations. Firstly, resource constraints often curtail their capacity to invest in advanced inventory management systems or expand their workforce, leading to member dissatisfaction stemming from issues like stockouts, overstocking, and delayed service. These issues directly impact the cooperative's competitiveness, resulting in financial inefficiencies and a diminished ability to attract new members. Secondly, the unpredictable nature of agribusiness, influenced by factors such as seasonal demand, climatic variability, or inefficient crop management practices, creates demand spikes that strain operations, leading to prolonged lead times and customer waiting queues. This, in turn, curtails the cooperative's bargaining power. Consequently, it becomes imperative to

embrace new technologies that streamline manual processes, thus reducing cycle times and enabling the cooperative to effectively address these challenges.

Despite the wide application opportunities for simulation and machine learning in agriculture, no papers were found regarding the usage of simulation to generate data regarding process performance to input a machine learning technique [45]. In addition, no paper considered using machine learning outputs for investment in process improvement. The present paper aims to present a novel approach combining computational simulation and machine learning techniques to enhance decision-making for agribusiness logistics. To forecast the average lead time and waiting queue time to help decision-makers determine the best number of workers to be hired and machines to be installed for the soybean seed classification process. To the best of the authors' knowledge, this study represents the first endeavor to employ such an approach for assisting in decision-making. In particular, the case displays a method to determine the best configuration for the process, i.e., how many machines to implement and how many workers to hire. The paper is structured as follows: Section 2 presents the soybean injury classification process and describes the applied methodology. Section 3 shows the results of the conducted case study. Section 4 concludes the paper.

2. Materials and Methods

2.1. Scenario Description

The soybean seed classification process addressed in this study is based on the real-life case of a Brazilian agribusiness cooperative located in the state of Paraná. This process is performed twice: at the arrival and departure of grains.

The current process, to ensure transparency among all stakeholders, including farms and cooperatives, stores physically classified samples for up to three days before being discarded. However, there are still weaknesses in the process, such as the manual input of classification results into the system. This raises concerns among managers, who report attempted fraud and bribery every year.

As shown in Figure 1, the system comprises the following stages: Trucks arrive with a soybean grain load and wait in a queue until they are served. Once served, the worker performs the "Liming" step, during which samples are taken from different parts of the truck. The samples proceed to the "Quartering" stage, where the sample is homogenized (i.e., grains collected from different parts of the truck are mixed to form a single sample). Subsequently, the samples are separated into two parts for two parallel processes: the first involves the "Impurities Classification", while the second involves both "Moisture Measurement" and "Injury Classification". The "Impurities Classification" stage analyzes the presence of contaminants in the sample, such as small stones, branches, leaves, and dust. It is performed using a machine. In the subsequent process, "Moisture Measurement" assesses the humidity level of the sample, also carried out by a machine. "Injury Classification" is the stage where soybean grains are manually inspected by workers to identify the quantity of damaged grains, rotten grains, and unripe grains. Finally, a report is generated summarizing all the analyses conducted, and the truck exits the system.

Human vision cannot reliably detect tenuous injuries, and the "Injury Classification" step necessitates cutting the seed and observing its health, leading to extended processing times (averaging 10 min) and lengthy truck queue times. Currently, workers perform a superficial visual analysis, which is highly inefficient and can lead to substantial fines for misclassification (averaging 3 min). However, on days with a high influx of trucks, the workers do not perform the "Injury Classification" step, claiming that this activity requires a significant amount of time, increasing the total lead time, and consequently resulting in long queues and waiting times.

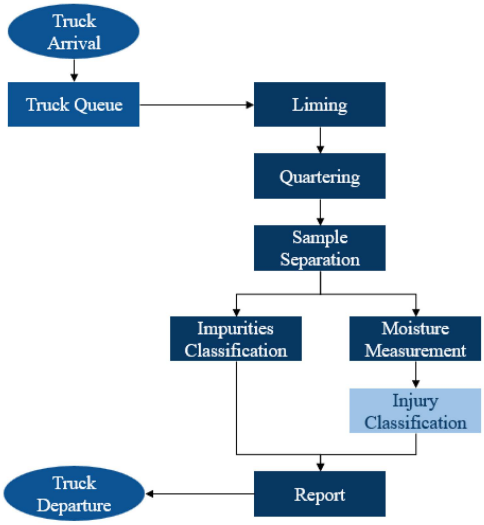


Figure 1. The soybean seed classification process.

Thus, the company chose to invest in intelligent methods for carrying out the task. It was considered to use a machine capable of performing the "Injury Classification" step efficiently using computer vision, a technique that has been used by researchers in order to identify varieties or defects in different types of grains such as soybeans [43,46,47], wheat [48,49], corn [50,51], rice [52], and many others. The machine's current lead time proves to be better than the ideal analysis time when performed by a worker, and the machine supplier anticipates that with new software updates, this lead time will be reduced to half.

Then, considering the different lead times for the same step, it is necessary to understand how the whole system will be affected when including the machine and changing parameters. Therefore, an approach combining computational simulation and machine learning techniques is applied to forecast the average lead time and waiting queue time to help decision-makers.

2.2. Methodological Procedure

The methodological procedure was split into three steps, as presented in Figure 2.

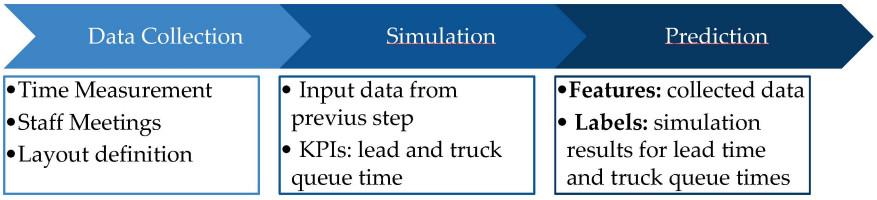


Figure 2. Methodological process.

The study's initial phase was conducted through on-site visits to the cooperative, focusing on data collection, process comprehension, and time measurement through discussions with employees. Then, historical data was gathered on truck arrivals, waiting times, and lead times. To mitigate potential biases, special care was given to consulting with the staff and ensuring cooperative representation. Staff consultations were conducted separately from data collection to prevent external influences on individual employee responses. Additionally, data was collected from various sites within the cooperative to provide a more comprehensive and unbiased assessment of its performance, reducing the risk of concluding based on a single site's characteristics.

Following the data collection phase, the study embarked on the development of multiple scenarios to support the simulation process. Given the innovative nature of the machine intended for injury classification, the absence of prior research or studies

elucidating its logistical impact on system performance necessitated the generation of artificial data through simulation. Within this context, the cooperative outlined three options for staffing levels and four for potential machine quantities, encompassing 1, 3, or 5 dedicated workers and 0, 1, 3, or 5 machines available for acquisition. Subsequently, the central objective revolved around determining the workforce composition (1, 3, or 5 employees to be hired) and machine deployment (0, 1, 3, or 5 machines to be purchased). This consideration involved exploring various permutations of the number of machines (m) and workers (w) while ensuring that $m \leq w$. In situations where there were more workers than machines ($m < w$), it was assumed that workers could simultaneously participate in the analysis process alongside the machines, eliminating the necessity for idle waiting periods. Moreover, it was firmly established that each worker could manage a single truck at a time, with the stipulation that once a worker initiated service for a truck, the process would proceed without interruption.

In the second step, a computer simulation model was designed to evaluate all the described scenarios. A Discrete Event Simulation using AnyLogic University v.8.5.2 was developed. As presented in Figure 3. The truck arrivals were input according to the historical data since 2017 provided by the cooperative. The simulation can explore different scenarios mixing real and virtual data, generating useful results for decision-making [53].

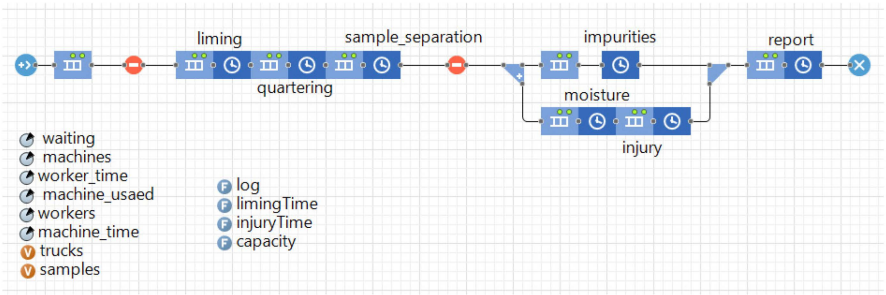


Figure 3. Process simulation in AnyLogic.

The simulation inputs are presented in Table 1. "Workers" and "Machines" determine the quantity w of workers and m machines available for the process. The values for both can be 1, 3, or 5 available workers/machines for future scenarios (as suggested by the decision-makers). "Worker Wait" is a Boolean parameter. When $m < w$, "Worker Wait" takes "True" if the worker waits for a machine to finish the task to use it or "False" if the worker executes the task in parallel. "Human Time" and "Machine Time" are the lead times for a human or a machine, respectively, to execute the task. "Truck Arrivals" are all the truck arrivals in the system.

Table 1. Simulation inputs for different scenarios.

Parameter	Values
Workers	[1, 3, 5]
Machines	[0, 1, 3, 5]
Worker Wait	[False, True]
Human Time	[195, 615] (s)
Machine Time	[60, 120] (s)
Truck Arrivals	All from 2017 until 2022

The case of simulating different scenarios by varying the number of workers without including the machine was also considered. In this case, 0 machines and 1, 3, or 5 workers were considered. It is important to note that this scenario does not directly reflect the real-world situation. In the simulated scenario, it is established that the workers must perform the classification manually with a lead time of 195 or 615 s, which is not currently the case in the real world where the "Injury Classification" is not even performed.

Additionally, all the machines are shared by all workers, as the worker needs to move between each of the machines or workstations to carry out the activities and complete the process. Two KPIs were collected: (i) truck queue time and (ii) lead time. The simulation results for each scenario constituted the database utilized in the subsequent stage, involving the application of machine learning. Thus, simulation was employed to generate historical data based on what-if scenarios, producing a robust historical foundation that was used as labels for the machine learning algorithms.

Finally, in the third step, a machine learning method was applied. This step was considered to understand how the different scenarios would impact the system in the following periods. The forecasting step was divided into two components: (1) predicting truck queue times and (2) estimating lead times. Both processes followed the procedural steps outlined in Figure 4.

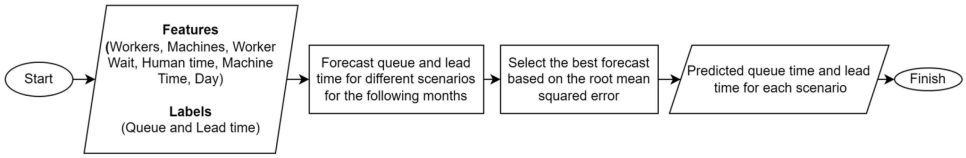


Figure 4. Forecasting process.

Table 2 presents the input variables, including "Workers", "Machines", "Worker Wait", "Human Time", and "Machine Time", which align with the input variables utilized in the simulation phase. Additionally, we introduce "Day", "Month", and "Day of Week" as time-based parameters linked to truck arrivals dating back to 2017. The incorporation of these temporal variables allows the model to capture potential seasonal patterns and trends that have evolved, enhancing its ability to learn and adapt to time-related dynamics. The Labels are the simulated truck queue time and lead time resulting from the simulation, considering all scenarios described in Table 1.

Table 2. Machine learning features.

Parameter	Values
Workers	[1, 3, 5]
Machines	[0, 1, 3, 5]
Worker Wait	[False, True]
Human Time	[195, 615]
Machine Time	[60, 120]
Day	All from 2017 until 2022
Month	All from 2017 until 2022
Day of the Week	From Monday

Algorithm 1 presents the pseudocode for the machine learning step. Initially, data scaling was performed to ensure that features with differing scales or units exerted an equitable influence during model training. Next, the data was partitioned into training and testing sets. In the following phase, machine learning models were trained on the training sets and subsequently tested. The algorithms considered included Linear Regression [54], Random Forest [55], Gradient Boosting [56], and Decision Tree [30] as they are well applied in the literature [57,58]. Model selection was based on the root mean squared error (RMSE), a metric that offers a comprehensive assessment of prediction accuracy, assigning equal importance to both small and large errors. This attribute makes RMSE sensitive to both overestimation and underestimation [59]. Finally, the model with the smallest RMSE was chosen for forecasting future scenarios.

Algorithm 1. Machine learning pseudocode.

```
Inputs:
- features←Dataframe
- scenarios←Dataframe (dates just for 4 first months of 2023)
- labels <- "queue_time" (double) or "lead_time" (double)
- training_models <- ["Linear Regression", "Random Forest",
"Gradient Boosting", "Decision Tree"]

Output:
- forecasted_data <- "queue_time" (double) or "lead_time" (double)

# Scale features
scaler <- initialize scaler
scaler(features)
scaler(labels)

# Split into test and training sets
feature_train, feature_test, label_train, label_test <- split(features, labels)

# Train Models
for model in training_models:
        train(feature_train,label_train)
        test(feature_test_, label_test)
        new_rmse <- test_rmse
        if new_rmse < old_rmse
                best_model <- model
                old_rmse <- new_rmse

# Forecast
forecasted_data <- best_model(scenarios)
return forecasted_data
```

Following model training and selection, the forecasting model was applied to all days during the first four months of 2023. Subsequently, for queue time, the total waiting time was aggregated, while for lead time, the average was computed for each month. Finally, for both targets, the monthly averages were calculated.

3. Results and Discussion

Table 3 presents the results for all the scenarios presented in Table 2. All the times are in minutes, and the table is sorted by queue time, ascending.

Table 3. Forecasts (time in minutes).

Scenario	Workers	Machines	Worker Wait	Machine Time	Worker Time	Queue Time	Lead Time
1	5	5	No	1	3.25	0.62	6.12
2	5	3	No	1	3.25	0.64	6.12
3	5	5	Yes	1	10.25	0.64	6.11
4	5	3	Yes	1	10.25	0.65	6.12
5	5	3	No	1	10.25	0.66	6.11
6	5	3	Yes	1	3.25	0.66	6.12
7	5	5	Yes	1	3.25	0.67	6.10
8	5	5	No	1	10.25	0.67	6.11
9	5	5	Yes	2	3.25	1.27	7.11
10	5	5	No	2	10.25	1.28	7.11
11	5	5	No	2	3.25	1.28	7.13

Table 3. *Cont.*

Scenario	Workers	Machines	Worker Wait	Machine Time	Worker Time	Queue Time	Lead Time
12	5	3	No	2	3.25	1.30	7.12
13	5	5	Yes	2	10.25	1.31	7.12
14	5	3	Yes	2	3.25	1.31	7.12
15	5	3	Yes	2	10.25	1.33	7.12
16	5	3	No	2	10.25	1.36	7.14
17	5	1	Yes	1	10.25	1.49	6.35
18	5	1	Yes	1	3.25	1.58	6.35
19	5	1	No	1	10.25	1.59	6.32
20	5	1	No	1	3.25	1.60	6.35
21	5	0	-	-	3.25	3.20	8.37
22	3	3	No	1	3.25	17.54	6.12
23	3	3	Yes	1	3.25	17.59	6.12
24	3	3	Yes	1	10.25	17.59	6.11
25	3	3	No	1	10.25	17.85	6.11
26	3	1	Yes	1	3.25	25.71	6.25
27	3	1	No	1	3.25	25.78	6.26
28	3	1	Yes	1	10.25	25.90	6.24
29	3	1	No	1	10.25	26.09	6.26
30	3	3	No	2	3.25	54.14	7.12
31	3	3	Yes	2	10.25	55.14	7.13
32	3	3	No	2	10.25	55.39	7.13
33	3	3	Yes	2	3.25	55.47	7.11
34	5	1	Yes	2	10.25	67.81	7.98
35	5	1	No	2	10.25	68.47	7.98
36	5	1	No	2	3.25	68.64	7.96
37	5	1	Yes	2	3.25	68.89	7.95
38	3	1	Yes	2	3.25	87.93	7.43
39	3	1	Yes	2	10.25	88.47	7.43
40	3	1	No	2	10.25	89.30	7.44
41	3	1	No	2	3.25	89.85	7.43
42	3	0	-	-	3.25	140.45	8.36
43	5	0	-	-	10.25	218.99	15.37
44	3	0	-	-	10.25	1255.33	15.37
45	1	1	Yes	1	3.25	9249.24	6.13
46	1	1	No	1	3.25	9250.02	6.12
47	1	1	Yes	1	10.25	9254.78	6.13
48	1	1	No	1	10.25	9262.24	6.12
49	1	1	Yes	2	3.25	17806.45	7.11

The results show that the variation in the lead time is considerably lower compared to the queuing time. Its values predominantly varied from 6 to 8 min, with a few exceptions close to 15 min. The queue time, on the other hand, presents a high variation, returning values lower than 1 min or greater than 270 days.

Table 4 presents the computed statistics for the predicted queue times, including averages, maximum values, and standard deviations, for each of the scenarios over the projected period. An observed trend reveals that a reduction in the number of machines and workers correlates with decreased system responsiveness during historically high-demand periods. This trend is corroborated by the observed escalation in standard deviation values. Indeed, as the average queue time increases, both maximum values and deviations follow suit. These findings underscore the nuanced relationship between resource allocation and system responsiveness in the context of fluctuating demand scenarios.

Table 4. Statistics for queue time forecasting (time in minutes).

Scenarios	Average	Max	Standard Deviation
1, 3, 7, 8	0.1	27.0	0.9
2, 4, 5, 6	0.1	27.7	0.9
9, 10, 11, 13	0.4	59.4	3.0
12, 14, 15, 16	0.4	63.1	3.2
17, 18, 19, 20	0.6	101.6	5.0
21	1.6	160.4	9.6
22, 23, 24, 25	4.4	312.3	20.4
26, 27, 28, 29	5.6	356.8	24.1
30, 31, 32, 33	10.5	474.4	36.4
34, 35, 36, 37	13.5	557.4	45.9
38, 39, 40, 41	15.5	574.5	48.5
42	25.7	677.9	68.4
43	37.4	829.1	92.2
44	957.4	7266.5	1918.4
45, 46, 47, 48	2051.3	14307.7	3781.4
49	3428.9	23684.1	6000.7

In contrast, when analyzing the statistics related to lead time in Table 5, a minimal variation can be observed across all values. However, particular attention is drawn to scenarios 43 and 44. In these scenarios, where classifications are conducted without the aid of machines and workers are required to manually split grains in half, a significant increase in total processing time is evident, nearly doubling the results obtained in previous scenarios.

Table 5. Statistics for lead time forecasting (time in minutes).

Scenarios	Average	Max	Standard Deviation
1, 3, 7, 8, 22, 23, 24, 25, 45, 46, 47, 48	6.2	7.0	0.6
2, 4, 5, 6	6.2	8.1	0.6
26, 27, 28, 29	6.4	9.7	0.7
17, 18, 19, 20	6.6	12.4	0.9
9, 10, 11, 13, 30, 31, 33, 34, 49	7.2	8.0	0.6
12, 14, 15, 16	7.2	10.0	0.6
38, 39, 40, 41	7.6	12.7	0.8
21, 42	8.4	9.2	0.7
34, 35, 36, 37	8.9	17.4	2.2
43, 44	15.4	16.2	0.7

Tables 4 and 5 present the results, while Table 6 offers a correlation map between input variables, averages, maximum values, and standard deviations. Concerning lead time, there is a direct proportionality with machine processing time, indicated by a factor of 0.82. Additionally, fewer machines result in longer average and maximum times, highlighting the significant impact of machine quantity on service speed. Other factors did not yield significant results.

Regarding queue time, the number of workers has a more substantial impact than machine quantity. Increasing the number of workers leads to shorter waiting times for trucks. Surprisingly, both machine and worker processing times have limited impacts on queue time, contradicting expectations of a more significant influence. These findings indicate that service capacity is a more critical factor in queue time than the speed at which services are performed.

Table 6. Correlation map.

Feature	Lead Time			Queue Time			
	Avg	Max	Std	Avg	Max	Std	Scale
Nº Workers	−0.05	0.25	0.29	−0.45	−0.59	−0.56	−1.00
Nº Machines	−0.43	−0.54	−0.26	−0.29	−0.35	−0.33	−0.50
Worker Wait	−0.24	−0.10	0.04	−0.14	−0.13	−0.14	0.00
Worker Time	0.19	0.12	0.00	0.11	0.09	0.10	0.50
Machine Time	0.82	0.35	0.28	0.11	0.12	0.10	1.00

Besides the discussion above, a final aspect needs to be explored. The machine currently has a lead time of 2 min. Nevertheless, according to the supplier, the new version of the software can carry out the task in half of it (1 min). Then, considering that the workers are not analyzing in parallel and the possible improvements in the machine, Table 7 shows the best alternatives for both scenarios.

Table 7. Best alternatives.

Machine Time	Workers	Machines	Avg. Queue Time	Avg. Lead Time
2	5	5	1.27	7.11
	5	3	1.31	7.12
	3	3	55.47	7.11
1	5	3	0.66	6.12
	5	5	0.67	6.10
	5	1	1.58	6.35
	3	3	17.59	6.12

Based on Table 7, the best decision is to use 3 machines, as it requires less investment and has the same efficiency as 5 machines. Additionally, considering that trucks can wait in the queue and that the machine can be improved, the best number of workers is 3 as well. Then, it can be concluded that in the best scenario, when the machine is improved, the best arrangement is 3 workers to 3 machines. Nevertheless, if the machine is not improved, then the best arrangement is 5 workers to 3 machines.

4. Conclusions

This paper presented a novel approach combining Discrete Event Simulation and machine learning to enhance decision-making in logistics. A systematic three-step methodology was employed. The initial phase involved conducting on-site visits to the cooperative to gather relevant data. Subsequently, a simulation model was developed to generate data encompassing various scenarios. Finally, in the third step, a machine learning approach was applied to predict both the queue time of trucks and the lead time. This comprehensive methodology enabled a structured and analytical approach to tackle the problem at hand. This study aimed to understand the logistical impact of selecting different scenarios to support decision-makers in determining the best configuration for the process. Therefore, in the real case of a Brazilian cooperative, the best number of workers to be hired and machines to be allocated to the Soybean Seeds Classification Process was calculated. The results confirm the expected strong connection between the queue time and the lead time, i.e., the higher the lead time, the higher the queue time. In addition, the queue time is affected mainly by the system capacity (number of workers), while the lead time is affected by machine capacity.

With the integration of the new machine into the system, the process gains accuracy, ensuring that the results of "Injury Classification" are truthful, mitigating the possibility

of fraud and bribery, thereby generating greater reliability for both the company and its clients. Additionally, with this suggested configuration, the machine's implementation should not result in significant queue time, which would render its utilization impractical. As a result, the process becomes more efficient and effective.

The utilization of machine learning presents significant potential by facilitating decision-making. This data-driven approach empowers decision-makers to provide informed recommendations for achieving the optimal configuration of the process, thereby enhancing efficiency and productivity. Moreover, the combination of discrete event simulation with machine learning highlights the capacity to leverage diverse techniques in addressing real-world challenges. For theoretical implications, this is the first article, to the best of the authors' knowledge, to employ a combined approach of simulation and machine learning to generate data that aids decision-making regarding investments in process improvements.

Regarding the contributions inherent in this work, the first relates directly to the field of agriculture. Against the backdrop of the growing importance of the efficiency of agricultural systems, a method is proposed that aims to better plan the allocation of critical resources—in this case, directly in the classification process—that can be used by other similar operations, considering other types of grains or products, reducing logistical costs, waiting times, and lead times.

In terms of practical implications, this work has contributed to the capacity planning process. In the approach proposed in this research, the company's managers obtained improved forecast data for future scenarios, supporting better decision-making in line with the company's operational objectives. Although the results cannot be generalized to other companies as this is a specific case, companies can adapt this method to make accurate decisions based on hypothetical scenarios, benefiting from a simple and low-cost implementation process. Additionally, the model could encompass more forecasting models or include different data for training.

As a limitation of this research, the costs of hiring workers and purchasing machinery were not analyzed. Furthermore, the proposed approach was not compared to traditional methods such as Little's law, the funnel model, or factory physics. These, then, are some recommendations for future directions of this research. Furthermore, the development of a framework that enhances the applied method to achieve generalization for other cases is recommended.

Author Contributions: Conceptualization, D.L.d.O.Q., I.C. and E.M.F.; methodology, D.L.d.O.Q., M.K. and L.G.M.; software, D.L.d.O.Q., M.K. and E.M.F.; validation, D.L.d.O.Q., M.K., I.C. and L.G.M.; formal analysis, D.L.d.O.Q., M.K., I.C., E.M.F. and L.G.M.; investigation, D.L.d.O.Q. and I.C.; resources, D.L.d.O.Q.; writing—original draft preparation, D.L.d.O.Q., I.C. and L.G.M.; writing—review and editing, D.L.d.O.Q., M.K., I.C., E.M.F. and L.G.M.; visualization, D.L.d.O.Q. and M.K.; supervision, M.K. and E.M.F.; project administration, I.C. and E.M.F.; funding acquisition, I.C. and E.M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Research Foundation (DFG), grant number FR 3658/4-1, and by the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES), grant number 88881.364431/2019-01, in the scope of the Collaborative Research Initiative on Smart Connected Manufacturing program and by the National Council for Scientific and Technological Development (CNPq), grant number 424195/2021-6.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bernardo, S.M.; Rampasso, I.S.; Quelhas, O.L.G.; Filho, W.L.; Anholon, R. Method to Integrate Management Tools Aiming Organizational Excellence. *Production* **2022**, *32*, 1–14. [CrossRef]
2. da Silveira, F.; Lermen, F.H.; Amaral, F.G. An Overview of Agriculture 4.0 Development: Systematic Review of Descriptions, Technologies, Barriers, Advantages, and Disadvantages. *Comput. Electron. Agric.* **2021**, *189*, 106405. [CrossRef]
3. Liu, Y.; Ma, X.; Shu, L.; Hancke, G.P.; Abu-Mahfouz, A.M. From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges. *IEEE Trans. Industr. Inform* **2021**, *17*, 4322–4334. [CrossRef]
4. Braun, A.-T.; Colangelo, E.; Steckel, T. Farming in the Era of Industrie 4.0. *Procedia CIRP* **2018**, *72*, 979–984. [CrossRef]
5. Belaud, J.-P.; Prioux, N.; Vialle, C.; Sablayrolles, C. Big Data for Agri-Food 4.0: Application to Sustainability Management for by-Products Supply Chain. *Comput. Ind.* **2019**, *111*, 41–50. [CrossRef]
6. Mendes, J.A.J.; Carvalho, N.G.P.; Mourarias, M.N.; Careta, C.B.; Zuin, V.G.; Gerolamo, M.C. Dimensions of Digital Transformation in the Context of Modern Agriculture. *Sustain. Prod. Consum.* **2022**, *34*, 613–637. [CrossRef]
7. Baierle, I.C.; da Silva, F.T.; de Faria Correa, R.G.; Schaefer, J.L.; Da Costa, M.B.; Benitez, G.B.; Benitez Nara, E.O. Competitiveness of Food Industry in the Era of Digital Transformation towards Agriculture 4.0. *Sustainability* **2022**, *14*, 11779. [CrossRef]
8. Boyabatl, O.; Nguyen, J.; Wang, T. Capacity Management in Agricultural Commodity Processing and Application in the Palm Industry. *Manuf. Serv. Oper. Manag.* **2017**, *19*, 551–567. [CrossRef]
9. Golmohammadi, A.; Hassini, E. Capacity, Pricing and Production under Supply and Demand Uncertainties with an Application in Agriculture. *Eur. J. Oper. Res.* **2019**, *275*, 1037–1049. [CrossRef]
10. Higgins, A.; Davies, I. A Simulation Model for Capacity Planning in Sugarcane Transport. *Comput. Electron. Agric.* **2005**, *47*, 85–102. [CrossRef]
11. Sendhil Kumar, K.S.; Anbarasi, M.; Shanmugam, G.S.; Shankar, A. Efficient Predictive Model for Utilization of Computing Resources Using Machine Learning Techniques. In Proceedings of the 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 29–31 January 2020; IEEE: New York, NY, USA, 2020; pp. 351–357.
12. Janiesch, C.; Zschech, P.; Heinrich, K. Machine Learning and Deep Learning. *Electron. Mark.* **2021**, *31*, 685–695. [CrossRef]
13. Khoa, B.T.; Huynh, T.T. Is It Possible to Earn Abnormal Return in an Inefficient Market? An Approach Based on Machine Learning in Stock Trading. *Comput. Intell. Neurosci.* **2021**, *2021*, 2917577. [CrossRef] [PubMed]
14. Kück, M.; Freitag, M. Forecasting of Customer Demands for Production Planning by Local Nearest Neighbor Models. *Int. J. Prod. Econ.* **2021**, *231*, 107837. [CrossRef]
15. Amasyali, K.; El-Gohary, N.M. A Review of Data-Driven Building Energy Consumption Prediction Studies. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1192–1205. [CrossRef]
16. Gumz, J.; Fettermann, D.C.; Frazzon, E.M.; Kück, M. Using Industry 4.0′s Big Data and IoT to Perform Feature-Based and Past Data-Based Energy Consumption Predictions. *Sustainability* **2022**, *14*, 13642. [CrossRef]
17. Zhang, Y.; Haghani, A. A Gradient Boosting Method to Improve Travel Time Prediction. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 308–324. [CrossRef]
18. Pantazi, X.E.; Moshou, D.; Alexandridis, T.; Whetton, R.L.; Mouazen, A.M. Wheat Yield Prediction Using Machine Learning and Advanced Sensing Techniques. *Comput. Electron. Agric.* **2016**, *121*, 57–65. [CrossRef]
19. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Soil Moisture Forecasting by a Hybrid Machine Learning Technique: ELM Integrated with Ensemble Empirical Mode Decomposition. *Geoderma* **2018**, *330*, 136–161. [CrossRef]
20. Goap, A.; Sharma, D.; Shukla, A.K.; Rama Krishna, C. An IoT Based Smart Irrigation Management System Using Machine Learning and Open Source Technologies. *Comput. Electron. Agric.* **2018**, *155*, 41–49. [CrossRef]
21. McNider, R.T.; Handyside, C.; Doty, K.; Ellenburg, W.L.; Cruise, J.F.; Christy, J.R.; Moss, D.; Sharda, V.; Hoogenboom, G.; Caldwell, P. An Integrated Crop and Hydrologic Modeling System to Estimate Hydrologic Impacts of Crop Irrigation Demands. *Environ. Model. Softw.* **2015**, *72*, 341–355. [CrossRef]
22. Folberth, C.; Baklanov, A.; Balkovič, J.; Skalský, R.; Khabarov, N.; Obersteiner, M. Spatio-Temporal Downscaling of Gridded Crop Model Yield Estimates Based on Machine Learning. *Agric. For. Meteorol.* **2019**, *264*, 1–15. [CrossRef]
23. Haghverdi, A.; Washington-Allen, R.A.; Leib, B.G. Prediction of Cotton Lint Yield from Phenology of Crop Indices Using Artificial Neural Networks. *Comput. Electron. Agric.* **2018**, *152*, 186–197. [CrossRef]
24. Hofmann, E.; Rutschmann, E. Big Data Analytics and Demand Forecasting in Supply Chains: A Conceptual Analysis. *Int. J. Logist. Manag.* **2018**, *29*, 739–766. [CrossRef]
25. Rahamathunnisa, U.; Nallakaruppan, M.K.; Anith, A.; Kumar, K.S.S. Vegetable Disease Detection Using K-Means Clustering and Svm. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; IEEE: New York, NY, USA, 2020; pp. 1308–1311.
26. Ramachandran, A.; Sendhil Kumar, K.S. Tiny Criss-Cross Network for Segmenting Paddy Panicles Using Aerial Images. *Comput. Electr. Eng.* **2023**, *108*, 108728. [CrossRef]
27. Qiang, L.; Jiuping, X. A Study on Vehicle Routing Problem in the Delivery of Fresh Agricultural Products under Random Fuzzy Environment. *Int. J. Inf. Manag. Sci.* **2008**, *19*, 673–690.
28. Padilla, M.P.B.; Canabal, P.A.N.; Pereira, J.M.L.; Riaño, H.E.H. Vehicle Routing Problem for the Minimization of Perishable Food Damage Considering Road Conditions. *Logist. Res.* **2018**, *11*, 1–18.

29. Rabbani, M.; Farshbaf-Geranmayeh, A.; Haghjoo, N. Vehicle Routing Problem with Considering Multi-Middle Depots for Perishable Food Delivery. *Uncertain Supply Chain. Manag.* **2016**, *4*, 171–182. [CrossRef]
30. Lorena, A.C.; Jacintho, L.F.O.; Siqueira, M.F.; De Giovanni, R.; Lohmann, L.G.; De Carvalho, A.C.P.L.F.; Yamamoto, M. Comparing Machine Learning Classifiers in Potential Distribution Modelling. *Expert. Syst. Appl.* **2011**, *38*, 5268–5275. [CrossRef]
31. Krejci, C.C.; Stone, R.T.; Dorneich, M.C.; Gilbert, S.B. Analysis of Food Hub Commerce and Participation Using Agent-Based Modeling: Integrating Financial and Social Drivers. *Hum. Factors* **2016**, *58*, 58–79. [CrossRef]
32. Graubner, M.; Balmann, A.; Sexton, R.J. Spatial Price Discrimination in Agricultural Product Procurement Markets: A Computational Economics Approach. *Am. J. Agric. Econ.* **2011**, *93*, 949–967. [CrossRef]
33. Ross, R.B.; Westgren, R.E. An Agent-Based Model of Entrepreneurial Behavior in Agri-Food Markets. *Can. J. Agric. Econ.* **2009**, *57*, 459–480. [CrossRef]
34. Tykhonov, D.; Jonker, C.; Meijer, S.; Verw, T.; Verwaart, T. Agent-Based Simulation of the Trust and Tracing Game for Supply Chains and Networks. *J. Artif. Soc. Soc. Simul.* **2008**, *11*, 1–30.
35. Schreinemachers, P.; Berger, T. An Agent-Based Simulation Model of Human-Environment Interactions in Agricultural Systems. *Environ. Model. Softw.* **2011**, *26*, 845–859. [CrossRef]
36. Wossen, T.; Berger, T. Climate Variability, Food Security and Poverty: Agent-Based Assessment of Policy Options for Farm Households in Northern Ghana. *Environ. Sci. Policy* **2015**, *47*, 95–107. [CrossRef]
37. Schouten, M.; Opdam, P.; Polman, N.; Westerhof, E. Resilience-Based Governance in Rural Landscapes: Experiments with Agri-Environment Schemes Using a Spatially Explicit Agent-Based Model. *Land Use Policy* **2013**, *30*, 934–943. [CrossRef]
38. Zheng, C.; Liu, Y.; Bluemling, B.; Mol, A.P.J.; Chen, J. Environmental Potentials of Policy Instruments to Mitigate Nutrient Emissions in Chinese Livestock Production. *Sci. Total Environ.* **2015**, *502*, 149–156. [CrossRef] [PubMed]
39. Malawska, A.; Topping, C.J. Evaluating the Role of Behavioral Factors and Practical Constraints in the Performance of an Agent-Based Model of Farmer Decision Making. *Agric. Syst.* **2016**, *143*, 136–146. [CrossRef]
40. Brändle, J.M.; Langendijk, G.; Peter, S.; Brunner, S.H.; Huber, R. Sensitivity Analysis of a Land-Use Change Model with and without Agents to Assess Land Abandonment and Long-Term Re-Forestation in a Swiss Mountain Region. *Land* **2015**, *4*, 475–512. [CrossRef]
41. Medina, G.d.S. The Economics of Agribusiness in Developing Countries: Areas of Opportunities for a New Development Paradigm in the Soybean Supply Chain in Brazil. *Front. Sustain. Food Syst.* **2022**, *6*, 842338. [CrossRef]
42. Ministry of Agriculture, Livestock, and Food Supply. *Brazil Normative Instruction No. 11 of May 15, 2007*; Ministry of Agriculture, Livestock, and Food Supply: Brasilia, Brazil, 2009; pp. 1–9.
43. Huang, Z.; Wang, R.; Cao, Y.; Zheng, S.; Teng, Y.; Wang, F.; Wang, L.; Du, J. Deep Learning Based Soybean Seed Classification. *Comput. Electron. Agric.* **2022**, *202*, 107393. [CrossRef]
44. Anheier, H.K.; Toepler, S. *International Encyclopedia of Civil Society*; Springer Science & Business Media: Berlin, Germany, 2009.
45. Utomo, D.S.; Onggo, B.S.; Eldridge, S. Applications of Agent-Based Modelling and Simulation in the Agri-Food Supply Chains. *Eur. J. Oper. Res.* **2018**, *269*, 794–805. [CrossRef]
46. Zhao, G.; Quan, L.; Li, H.; Feng, H.; Li, S.; Zhang, S.; Liu, R. Real-Time Recognition System of Soybean Seed Full-Surface Defects Based on Deep Learning. *Comput. Electron. Agric.* **2021**, *187*, 106230. [CrossRef]
47. Lin, P.; Xiaoli, L.; Li, D.; Jiang, S.; Zou, Z.; Lu, Q.; Chen, Y. Rapidly and Exactly Determining Postharvest Dry Soybean Seed Quality Based on Machine Vision Technology. *Sci. Rep.* **2019**, *9*, 17143. [CrossRef] [PubMed]
48. Laabassi, K.; Belarbi, M.A.; Mahmoudi, S.; Mahmoudi, S.A.; Ferhat, K. Wheat Varieties Identification Based on a Deep Learning Approach. *J. Saudi Soc. Agric. Sci.* **2021**, *20*, 281–289. [CrossRef]
49. Zhao, X.; Que, H.; Sun, X.; Zhu, Q.; Huang, M. Hybrid Convolutional Network Based on Hyperspectral Imaging for Wheat Seed Varieties Classification. *Infrared Phys. Technol.* **2022**, *125*, 104270. [CrossRef]
50. Javanmardi, S.; Miraei Ashtiani, S.-H.; Verbeek, F.J.; Martynenko, A. Computer-Vision Classification of Corn Seed Varieties Using Deep Convolutional Neural Network. *J. Stored Prod. Res.* **2021**, *92*, 101800. [CrossRef]
51. Xu, P.; Tan, Q.; Zhang, Y.; Zha, X.; Yang, S.; Yang, R. Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning. *Agriculture* **2022**, *12*, 232. [CrossRef]
52. Koklu, M.; Cinar, I.; Taspinar, Y.S. Classification of Rice Varieties with Deep Learning Methods. *Comput. Electron. Agric.* **2021**, *187*, 106285. [CrossRef]
53. Quadras, D.; Frazzon, E.M.; Mendes, L.G.; Pires, M.C.; Rodriguez, C.M.T. Adaptive Simulation-Based Optimization for Production Scheduling: A Comparative Study. *IFAC PapersOnLine* **2022**, *55*, 424–429. [CrossRef]
54. de Oliveira, R.C.; Mendes-Moreira, J.; Ferreira, C.A. Agribusiness Intelligence: Grape Production Forecast Using Data Mining Techniques. In Proceedings of the Advances in Intelligent Systems and Computing; Springer: Berlin/Heidelberg, Germany, 2018; Volume 747, pp. 3–8.
55. Rusli, N.I.A.; Zulkifle, F.A.; Ramli, I.S. A Comparative Study of Machine Learning Classification Models on Customer Behavior Data. In Proceedings of the 7th International Conference on Soft Computing in Data Science 2023, Virtual Event, 24–25 January 2023; Springer: Singapore, 2023; pp. 222–231.
56. Gao, B.; Zhang, L.; Ou, D.; Dong, D. A Novel Deep Learning Model for Short-Term Train Delay Prediction. *Inf. Sci.* **2023**, *645*, 119270. [CrossRef]

57. Xu, Z.; Kurek, A.; Cannon, S.B.; Beavis, W.D. Predictions from Algorithmic Modeling Result in Better Decisions than from Data Modeling for Soybean Iron Deficiency Chlorosis. *PLoS ONE* **2021**, *16*, e0240948. [CrossRef]
58. Li, H.; Yao, B.; Yan, X. Data-Driven Public R&d Project Performance Evaluation: Results from China. *Sustainability* **2021**, *13*, 7147. [CrossRef]
59. Soares, L.D.; Franco, E.M.C. BiGRU-CNN Neural Network Applied to Short-Term Electric Load Forecasting. *Production* **2022**, *32*, e20210087. [CrossRef]

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Applied Sciences Editorial Office
E-mail: applsci@mdpi.com
www.mdpi.com/journal/applsci

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-4406-7