Universidad de Valladolid

ESCUELA DE INGENIERÍAS INDUSTRIALES

UNIVERSIDAD DE VALLADOLID

ESCUELA DE INGENIERIAS INDUSTRIALES

Grado en Ingeniería de Organización Industrial

Influence of AI-Generated Data in the simulation of an industrial process

Autor:

Redondo García, Elena

Ángel Manuel Gento Municio

Hochschule Albstadt-Sigmaringen

Valladolid, Septiembre 2024.

TFG REALIZADO EN PROGRAMA DE INTERCAMBIO

TÍTULO:	Influence of AI-Generated Data in the simulation of an industrial process
ALUMNO:	Elena Redondo García
FECHA:	24.08.2024
CENTRO:	Faculty of Engineering
UNIVERSIDAD:	Hochschule Albstadt-Sigmaringen
TUTOR:	Prof. Dr. Lutz Sommer

Resumen

Este trabajo se centra en analizar los efectos del uso de herramientas de machine learning en un problema concreto de logística. Para ello se ha creado un modelo logístico en el que existen cinco tipos de productos que pasan por ciertas fases antes de ser distribuidos a sus respectivos centros específicos. Las fases en las que consiste el problema son: recepción, almacenamiento, preparado de pedidos y distribución. Se ha utilizado el software Witness para realizar una simulación del sistema. Al principio se han establecido parámetros aleatorios para recopilar datos que posteriormente nos sirvan para alimentar la memoria de la inteligencia artificial utilizada para el análisis. Para finalizar, se han integrado todos los datos en el programa Orange Data Mining. Gracias al diagrama de flujo establecido, se ha hecho que los datos sean evaluados y mediante un código se busque la combinación óptima para la gestión de la simulación.

Palabras clave: Inteligencia Artificial, Logística, Optimización, Simulación de procesos, Software Witness.

Abstract

This work focuses on analyzing the effects of using machine learning tools in a specific logistics problem. To achieve this, a logistics model has been created in which there are five types of products that go through certain phases before being distributed to their respective specific centers. The phases involved in the problem are reception, storage, order preparation and distribution. The Witness software was used to simulate the system. Initially, random parameters were set to collect data that would later be used to feed the memory of the artificial intelligence used for the analysis. Finally, all the data was integrated into the Orange Data Mining program. Thanks to the established flow diagram, the data was evaluated and through a code, the optimal combination for managing the simulation was sought.

Key words: Artificial Intelligence, Logistics, Optimization, Process Simulation, Witness Software.

Table of Contents

List of Figures

List of Abbreviations

AI…………………………………………………...…………………Artificial Intelligence

Academic Honesty Declaration

I declare that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or whole, for an assignment at this or any other University.

Formulations and ideas taken from other sources are cited as such. This work has not been published.

Albstadt, August 24th, 2024

E. Redondo García

…………………………………

Abstract

In this thesis we investigate the impact of using artificial intelligence within the simulation of an industrial process. The study focuses on simulating a logistics system using the Witness software. Once this system is completed, it integrates input data generated in the Orange tool through machine learning processes. The experiment aims to analyze how AI-generated data influences process efficiency and to provide insights for decision-making. The simulation consists of a reception area, storage, order preparation, and final distribution of products to their respective points of sale. It includes four types of products with random arrivals destined for specific demand centers. After running several simulations with the generated random data, the model is examined for inefficiencies or potential bottlenecks. Once this is done, the data generated by Orange is integrated to optimize the flow of products. Upon completion of the study, the results demonstrate how the incorporation of AI improves process performance. Some of the advantages shown include significant reductions in delivery times, leaving room for potential future research on other aspects, such as storage costs. This research provides an overview of the basic concepts involved in the study, as well as concise results that demonstrate how AI has a beneficial impact on industrial logistics management.

Keywords: Process Simulation, Artificial Intelligence, Logistics, Optimization, Witness Software.

1 Introduction

1.1 Problem Definition

In this thesis, we will describe the development of the simulation of a logistics process and the implementation of data generation to observe the different results that can be obtained, trying to find a continuous improvement of the process.

The structure of the thesis is as follows:

- A comprehensive review of fundamental concepts, including the principles of simulation, definition of industrial processes and the basics of AI. In addition to the specifications for the development of the model and the systems used.
- The methodology section that describes the design process and details each step involved in the development of the model and the implementation of the data generated by AI, specifying each step followed in the software used.
- The results section which presents the outputs obtained in the simulation.
- Lastly, the conclusion that summarizes the findings of this work and provides recommendations for future implementations in this area.

In the last decade, the concept of artificial intelligence has been progressively integrated into all types of sectors and industries. Many companies have been able to optimize their processes, improve efficiency, and reduce costs thanks to the capabilities it offers. Industrial process simulation has benefited greatly from advances in AI, allowing researchers and professionals to model and predict the behavior of complex systems with greater accuracy (OpenAI, 2024).

Despite significant advances, there is a continued need to improve the accuracy and efficiency of industrial simulations. A crucial question is how AI-generated data can influence the simulation of logistics processes, specifically the management of product receipt, storage and distribution. This study seeks to answer this question.

1.2 Literature Review

In this section, we are going to discuss about the fundamental theoretical concepts that form the background of this thesis. These concepts are crucial for understanding the interplay between artificial intelligence, data generation, and industrial logistics processes. Some of the topics that are developed are the definition of AI, industrial process and simulation. Additionally, there is a section that explain the relationship between these concepts to establish a coherent framework that supports the research objectives.

1.2.1 Basic Concepts of AI and Data Generation.

Artificial intelligence is a branch of computer science dedicated to developing systems capable of performing tasks that usually necessitate human intelligence, including learning, reasoning, and perception. These systems can observe their surroundings, reason based on knowledge, process information from data, and make decisions to accomplish specific objectives (Gobierno de España, 2023).

It uses algorithms and mathematical models to analyze large datasets and make decisions based on patterns and rules established through machine learning. This is the capability of a machine to learn independently from data without explicit programming. This enables AI to enhance its accuracy and efficiency over time (Gobierno de España, 2023).

There are various classifications, but we will focus on the following two types:

- Software: virtual assistants, image analysis software, search engines or voice and face recognition systems.
- Integrated artificial intelligence: robots, drones, autonomous vehicles or the Internet of Things.

(Gobierno de España, 2023)

Data mining involves analyzing a dataset, associating it with patterns, and extracting information deemed useful from these patterns. Some of the techniques associated with this concept include machine learning, statistics, and database methods, which identify potential hidden relationships and trends. Data classification, regression, clustering, and association rules are some of the methods used.

1.2.2 Industrial Processes.

We can define an industrial process as a series of operations and procedures designed to convert raw materials into finished or intermediate products using machinery, tools, and technology. These processes are crucial for producing goods and services across various industries and can encompass activities such as manufacturing, assembly, chemical processing, and materials handling (Elecproy, 2023).

We can differentiate four types of industrial processes:

- Make-to-order production: there is a prior demand.
- Batch production: products are generated in specific quantities and at a particular time.
- Mass production: used to produce items in large quantities.
- Continuous production: involves the uninterrupted production of goods that maintain the same exact characteristics.
  (Elecproy, 2023)

There are many industrial processes; in this work, it has been decided to focus on a typical logistics process. Therefore, it is necessary to introduce this type of process.

A logistics process encompasses all the activities a product goes through from its manufacturing to its delivery to the final customer, including transportation, storage, and subsequent distribution. The main objective of the process is to deliver the requested quantity of materials at the right time and place. To maintain an efficient supply chain, it is necessary to properly manage all integrated activities.

Some notable examples of logistics processes include the procurement of materials or goods, the storage of any item, inventory management, order preparation and dispatch, and the transportation and delivery of orders.

1.2.3 Principles of Simulation.

The concept of simulation involves constructing a model that represents a real system, aiming to replicate the behavior of that system in a controlled environment. The goal is to extract the necessary information to understand the system's behavior and evaluate potential scenarios to identify failures, improvement opportunities, and even predict the impact of certain changes on the system. In summary, its purpose is to analyze the model's behavior and assess possible strategies to follow.

Below is a list of the basic elements that every simulation must contain.

- Entities: These are the elements that move, change state, and interact with the system's resources.
- Resources: These are permanent elements that are always present in the system and perform actions on the entities.
- Activities: These are the functions that the resources perform on the entities.
- Attributes: These are the characteristics used to describe entities and resources.
- Parameters: These are special attributes of the resources whose modification leads to new what-if scenarios.
- Variables: These are elements subject to some type of change. There are input, output, and internal variables.
- State: This is the situation in which an entity or resource is at a given moment.
- Events: These are occurrences at a specific point in time that cause changes in the state of an entity or resource.
- Buffers: These are sets of entities waiting for their turn because the resources they wish to use are not available.
- Functional relationships: These are the connections that exist between the system's resources, which serve to define the system's processes.

(Sanz)

To develop a simulation, we must follow the following steps:

- Definition of the objectives, scope and level of detail of the system we want to study.
- Collection of data necessary for the construction of the model.
- Structure the model and estimate its parameters to finally build the computer model with all its elements.
- Run the model and evaluate the outputs obtained.
- Ensure that the model is correct and conforms to the specifications of the real system. Therefore, the system must be verified and validated.
- Conduct experiments to analyze different scenarios.
- Evaluate the results and make decisions.

(Sanz)

1.2.4 Summary of the correlation of terms.

Once we introduced each concept, this section is going to show the correlation that exists for the execution of this work.

The application of artificial intelligence in a logistics process can be integrated at various stages. Algorithms can be used to analyze historical data and trends to predict future product demand at each point of sale. It can also be employed to optimize the amount of stock each storage should keep on hand to minimize costs and avoid stockouts. Another potential optimization can be achieved through machine learning to calculate the most efficient transportation routes, thereby reducing time and distribution costs. Additionally, the presence of integrated artificial intelligence allows for the prediction of equipment failures, enabling proactive maintenance (OpenAI, 2024).

We must point out that associated with artificial intelligence is the concept of big data. This concept refers to the huge amount of data that can be generated, processed, and utilized through digital tools, and it directly supports AI by providing all the necessary information for its operation.

Thanks to data generation and subsequent analysis, we will be able to create a simulation that fits the established process and can present optimizations using the available resources.

1.3 Research Question

Lately, industrial logistics processes are becoming more complex due to the need to increase profitability through efficient resource allocation, product delivery optimization, and the management of all involved operations. The traditional methods of managing these processes often fail to adapt to the changing demands of today. Artificial intelligence offers an opportunity to enhance these processes by providing data-driven insights and predictive capabilities.

Given the challenges faced by logistics managers, this thesis seeks to answer the following research question:

"How can data generated by artificial intelligence improve the efficiency and profitability of logistics and distribution processes in industrial environments through simulation and data analysis?"

1.4 Hypothesis

As mentioned in the previous section, this thesis is trying to analyze the impact of generated data on a logistics process. Therefore, the main hypothesis is that data generated by artificial intelligence can significantly improve the efficiency and profitability of industrial logistics processes through optimized simulation models. This study posits that AI-driven data analysis not only enhances decision-making but also leads to better resource allocation and process optimization in logistics operations. In addition to the main idea, there are certain hypotheses directly related to the topic of study that can be examined. These are referred to as sub-hypotheses, and this work does not focus on proving all of them.

Main Hypothesis:

H1: *Data generated by artificial intelligence improves the efficiency of logistics processes by optimizing resource allocation through simulation models.*

Sub-hypotheses:

H1a: *The use of AI in data generation can reduce delivery times in logistics processes by predicting demand more accurately.*

H1b: *AI-driven simulation models can minimize storage costs by optimizing inventory levels and reducing holding times.*

H1c: *The integration of AI enhances decision-making in logistics by providing predictive insights and actionable data.*

2 Method

2.1 Subject / Object

The principal aim of this project is to evaluate how the data generated by a software can influence to the simulation of an industrial process, specifically focusing on a logistics process

Furthermore, the work has some specific objectives.

- Develop a simulation of a logistics process using a specialized software.
- Generate random data that simulate the product arrivals and subsequently refine data generation using machine learning techniques.
- Analyze the simulation results and apply AI-driven improvements to optimize the logistics process.
- Compare and evaluate the outcomes before and after optimization.

2.2 Procedure and Methods

This section outlines the procedural framework, and methodologies employed in this research. The approach consists of several phases, each designed to address the specific objectives of the project. Before detailing the steps, it is crucial to classify the nature of the process being developed. Therefore, we must say that we have a process:

- Dynamic, since the evolution of the system over time is considered.
- Stochastic, since the input components are random variables.
- At first, we will consider it descriptive, since we will observe its behavior through the relationships that exist within. And later it will become prescriptive since we will seek to optimize the system by trying to identify the best action to take.
- Discrete time, since the state of the system changes at specific moments in time.

Once we have established the characteristics of the process, we proceed to divide the work into four distinct phases. These are explained below.

Phase 1: Process Simulation Development

In this initial stage, we decided the appropriate software for creating the correct model. Some of the possible software options considered were 'Process Simulator' and 'Altair,' but the decision was ultimately made to use 'Witness.' This choice was based on both the ease of downloading it from the server and the advantages offered by the program itself.

Witness is an advanced tool that allows for the modeling and optimization of industrial processes through simulation. It offers a wide range of options to ensure the model can be tailored to the previously established specifications, providing great flexibility in the simulation. Additionally, it features an easy and educational learning process.

Once the simulator has been selected then we start with the develop of the simulation model. Key components such as product arrivals, storage, order preparation, and subsequent distribution have been incorporated into the process. It is also necessary to set the appropriate process parameters, processing times, and transportation durations to ensure that a realistic logistics operation is reflected.

The entire process of creating a simulation model is detailed in Appendix A.

Phase 2: Data Generation and Integration

In this section, we need to choose again an appropriate software to generate the necessary data. The choice was straightforward as we opted directly for the ‘Orange’ software. So, we will generate the random data that represent the arrivals of each type of product to the model. These data will be integrated into the simulation to observe real-world variability.

Once the simulation is executed with complete randomness, we start applying machine learning techniques to the generated data for the input data, with the goal of making more accurate predictions for product arrivals.

Phase 3: Process Optimization

The next step is to collect the simulation results and try to identify any bottlenecks, inefficiencies, or areas for improvement. Then we applied the optimization algorithms to enhance the overall efficiency of the process. The new data obtained will be reintroduced to evaluate the impact of the improvements on the process.

The process carried out with the Orange tool is explained in Appendix B.

Phase 4: Results Comparison and Evaluation

Once the entire process is completed, we proceed to compare all the data obtained to derive appropriate results from the process. These will be presented in the following section.

Finally, the results will be used to validate or refute the initial hypotheses, contributing to the overall conclusions of the study.

This has been an explanation of the steps to be followed in the study referred to in the thesis. As mentioned, the tools used will be Witness and Orange. The following sections present the results obtained and the final reflection of the study.

3 Results

This section will present the results obtained from the programs used.

First, we'll start with the results from Witness. The simulation was initiated and stopped after 24 hours of operation. This allowed us to observe how the product arrivals were generated, and the model progressed according to its specifications. In the process diagram, we can see the workflow within the model. There is a queue before the storage area, which is due to the wide range of storage times, which could potentially cause issues in the long run. This could be avoided by adjusting the machine's specifications. On the conveyors, we can see that each product takes its corresponding time to reach the end, as shown in the final counter.

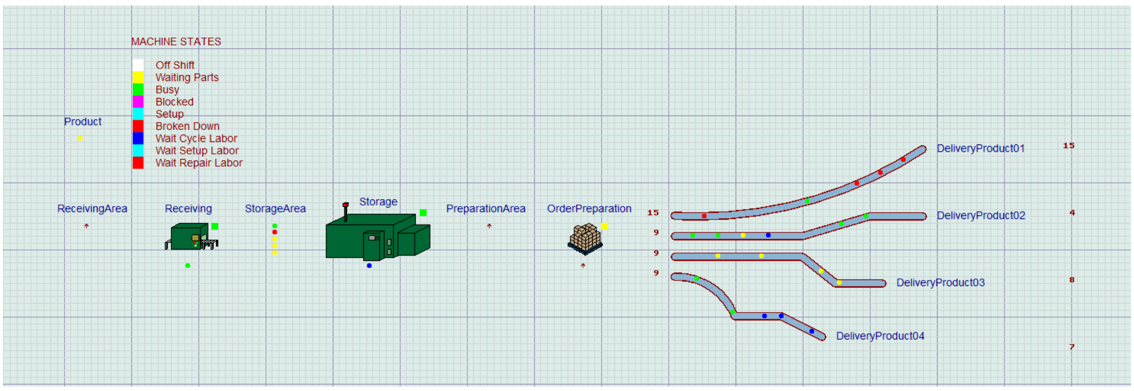

Figure 1: Run model in Witness.

In addition to the model simulation, certain statistics were extracted from each area. This allows us to observe the number of products that entered the model and completed their cycle, as well as those still in process.

Part Statistics Report by On Shift Time

Name	Product
No. Entered	49
No. Shipped	22
No. Scrapped	0
No. Assemble	0
No. Rejected	0
W.I.P.	27
Avg W.I.P.	17.96
Avg Time	527.67
Sigma Rating	6.00

Figure 2: Report of Part Statistics.

Furthermore, all the statistics for the queues, machines, and conveyors are shown in the following figures.

Buffer Statistics Report by On Shift Time

Name	ReceivingAr	StorageArea	PreparationA
Total In	49	48	42
Total Out	49	43	42
Now In	0	5	0
Max	1	6	2
Min	0	0	0
Avg Size	0.00	2.88	0.07
Avg Time	0.00	86.48	2.44
Avg Delay Co			
Avg Delay Tim			
Min Time	0.00	0.00	0.00
Max Time	0.00	158.05	19.72

Figure 3: Report of Buffer Statistics.

Machine Statistics Report by On Shift Time

Name	Receiving	Storage	OrderPrepar
% Idle	66.67	0.69	50.66
% Busy	33.33	99.31	42.44
% Filling	0.00	0.00	0.00
% Emptying	0.00	0.00	0.00
% Blocked	0.00	0.00	6.90
% Cycle Wait	0.00	0.00	0.00
% Setup	0.00	0.00	0.00
% Setup Wait	0.00	0.00	0.00
% Broken Do	0.00	0.00	0.00
% Repair Wait	0.00	0.00	0.00
No. Of Operati	48	42	42

Figure 4: Report of Machine Statistics.

Conveyor Statistics Report by On Shift Time

Name	DeliveryProd	DeliveryProd	DeliveryProd	DeliveryProd
% Empty	8.91	43.50	5.46	15.26
% Move	91.09	56.50	94.54	84.74
% Blocked	0.00	0.00	0.00	0.00
% Queue	0.00	0.00	0.00	0.00
% Broken Do	0.00	0.00	0.00	0.00
Now On	5	6	4	5
Total On	18	6	8	10
Avg Size	4.89	1.62	3.50	3.17
Avg Time	391.33	387.76	629.96	457.15

Figure 5: Report of Conveyor Statistics.

Once the data was input into the Orange model, the following distribution was obtained, showing how the preparation time of products affects the arrival frequency of each product type. We see that the model programmed more preparations between 11 and 16 minutes, with a peak in the distribution of product B.

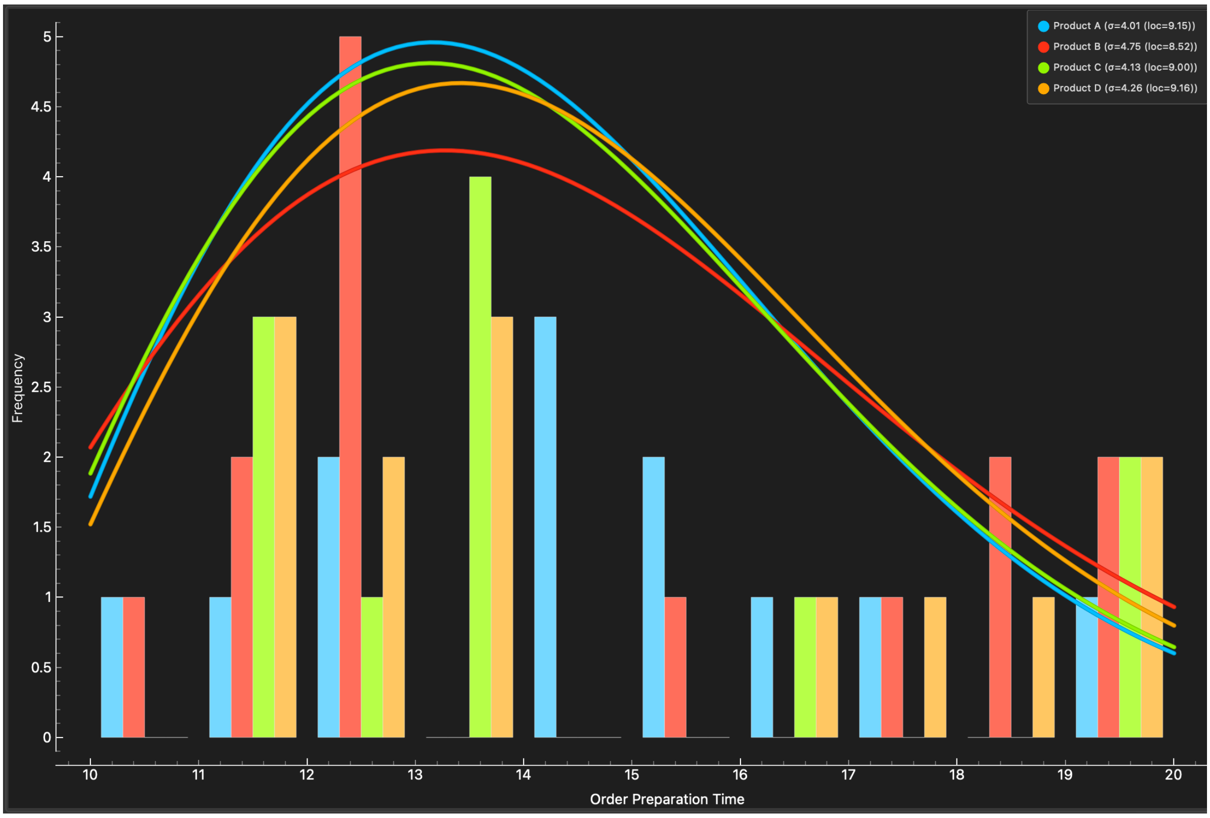

Figure 6: Distribution of Product Type vs Orden Preparation Time.

Finally, Figure 7 shows a list generated by the program. This is an example generated through machine learning that offers an optimized version of how product arrivals can be organized to maximize the use of resources within the model.

	Product	ID	Arrival Time	Storage Time	der Preparation Ti	Delivery Time	End Time
1	Product A	P001	30	51	12	45	148
2	Product D	P002	60	0	12	60	142
3	Product A	P003	90	38	11	45	194
4	Product C	P004	120	5	13	85	233
5	Product C	P005	150	19	11	85	275
6	Product B	P006	180	2	18	105	315
7	Product C	P007	210	15	20	85	340
8	Product A	P008	240	3	10	45	308
9	Product B	P009	270	15	20	105	420
10	Product B	P010	300	10	17	105	442
11	Product C	P011	330	58	13	85	496
12	Product B	P012	360	16	10	105	501
13	Product B	P013	390	26	20	105	551
14	Product D	P014	420	47	11	60	548
15	Product B	P015	450	10	12	105	587
16	Product A	P016	480	15	12	45	562
17	Product B	P017	510	55	12	105	692
18	Product C	P018	540	27	20	85	682
19	Product A	P019	570	7	15	45	647
20	Product D	P020	600	16	13	60	699

Figure 7: Results of the machine learning.

4 Discussion

4.1 Hypothesis Testing

Based on the result of chapter 3 we can affirm that the hypothesis presented in this work is correct.

H1: *Data generated by artificial intelligence improves the efficiency of logistics processes by optimizing resource allocation through simulation models.*

After conducting the study, the results have shown that the use of AI in process simulation is a highly effective tool for achieving optimal outcomes in models. Through its integration, we can analyse a wide range of scenarios and identify the most suitable one according to our needs. The sub-hypotheses have also been corroborated, as an optimized scenario leads to reduced process times due to the absence of idle or overexploited resources, with each resource correctly allocated. This results in minimal process costs, as every resource is utilized to its fullest potential. The ability to demonstrate all these advantages through the integration of algorithms simplifies decision-making for those involved in the process. For all these reasons, we confirm that the integration of AI into industrial process simulation is a competitive advantage in every aspect.

4.2 Comparison with literature

There are certain observations made during the study that can be compared with the theoretical part presented. One example is the advantage provided by AI's ability to generate and store data, meaning it can handle a wide range of data without losing any detail. While this isn't very noticeable in such a simple model, on a larger scale, having such a powerful tool at everyone's disposal is highly beneficial.

Another concept that proved very useful is the integration of the general concept of simulation. Thanks to this technique, an unlimited number of tests can be conducted for each problem posed without the need to spend resources or invest excessive time.

4.3 Limitation

It is true that the model used in this work is quite simple and did not present any difficulties in terms of adjustment, meaning there were no bottlenecks or areas where an excessive number of products accumulated, causing issues in the model. This is also due to the assumptions made when creating the model, such as products being created gradually, resource capacities not being limited, or the times being very similar.

Regarding the software, there were indeed some limitations in establishing a connection between the two. The data generated in CSV format by the Witness program was not entirely accurate, requiring some modifications. Additionally, it was not possible to directly import the data generated with Orange into Witness. Therefore, the model had to be recreated in Orange, and the simulation was carried out directly in that program.

Some recommendations for future studies could include using more complex programs to generate more elaborate results. Adding cost specifications within the model itself could also add value, as it would be an additional factor to aid in decision-making. Adjusting time specifications to the characteristics of the products instead of assuming them to be random is another way to conduct experiments that could be of interest for future studies.

In this case, we used Random Forest predictions. There are a variety of algorithms that offer optimal solutions using different methods, such as genetic algorithms, simulated annealing, etc. For future studies, their use could be considered to observe the differences in the results obtained.

4.4 Implications

As we have just mentioned, this system for generating simulation models is generally very helpful for a wide range of companies, as it allows for the easy analysis of all possible outcomes without investing a significant number of resources.

Regarding the integration of AI in this field, as we also mentioned, it is a tool that helps address future uncertainty by providing strong demand prediction capabilities and helping models adjust to market needs. Moreover, it supports decision-making for any problem that may arise. These solutions are extremely helpful for companies looking to improve their performance by analysing their activities through these means.

4.5 Conclusion

In summary, this study has provided a clear and concise view of how the use of artificial intelligence offers advantages in the logistics sector. It is important to integrate new technologies into all sectors where they can bring innovation and improve performance. We have focused on the logistics sector, but this model is easy to apply to any industrial sector. AI has brought about many advancements, and in this work, we have seen a small part of its data storage and generation capabilities. We have also closely examined the advantages of using simulation models for solving and analysing complex models. Therefore, we can conclude that the thesis provides a simple and adequate approach to the topic, focusing on the key points and clearly outlining current needs and potential future implementations.

References

Gobierno de España (19 April 2023). Qué es la Inteligencia artificial. https://planderecuperacion.gob.es/noticias/que-es-inteligencia-artificial-ia-prtr

Banks, J. C. (2009). *Discrete-Event System Simulation.* Prentice Hall.

Chopra, S. &. (2016). *Supply Chain Management: Strategy, Planning, and Operation* . Pearson Education.

Elecproy. (30 January 2023). ¿Qué es un proceso industrial? https://elecproy.com/es/blog/que-es-un-proceso-industrial/

Lanner Group (2021). *WITNESS Horizon User Guide.* Lanner Group.

Kleijnen, J. P. (2015). Design and analysis of simulation experiments for input/output modeling. *European Journal of Operational Research, 266(3)* , 877-888.

Law, A. M. (2014). *Simulation Modeling and Analysis.* McGraw-Hill Education.

Russel, S. &. (2021). *Artificial Intelligence: A Modern Approach.* Pearson.

Sanz , P. (s.f.). Sistemas de Producción y Fabricación. *Tema 5: Simulación*. Valladolid.

Simchi-Levi, D. K.-L. (2007). *Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies.* McGraw-Hill Education.

Witten, I. H. (2016). *Data Mining: Practical Machine Learning Tools and Techniques.* Morgan Kaufmann.

OpenAI. (2024). ChatGPT https://chat.openai.com/chat

Appendix

Appendix A: Work in Witness

In this section, the entire process of developing the simulation using the Witness software will be explained in detail. To begin, we will describe the complete logistics process, and then focus on the characteristics of each part.

The created simulation represents a logistics process that involves four different types of products. These products are introduced into the system randomly and arrive at a receiving area where they must wait for a specified amount of time before moving to a storage area. There, the products wait for a random period before advancing to the preparation area, which is responsible for preparing each product within a random time frame, and finally distributing each type of product to a final recipient. The image shown below presents a schematic of the complete simulation model as implemented in Witness.

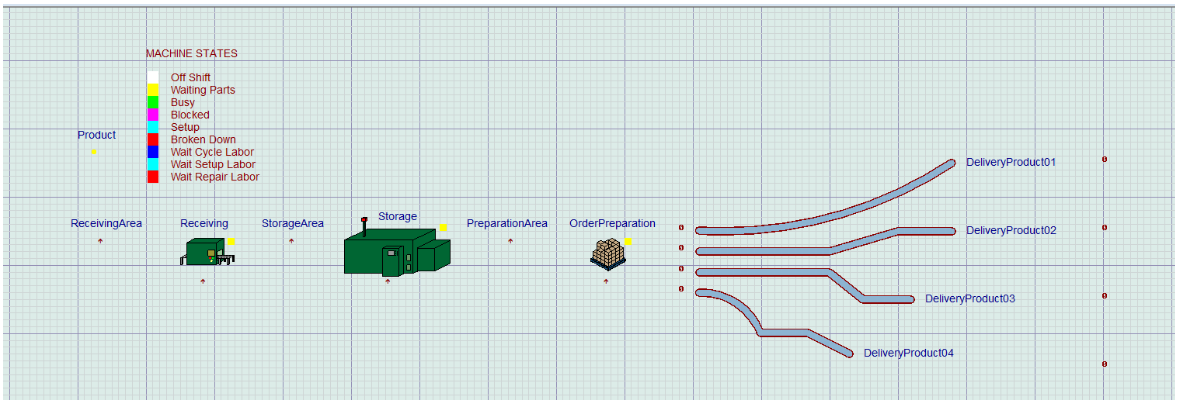

Figure 8: Complete Simulation Model.

The simulation consists of a total of fifteen elements that enable its proper development. Each of these elements will be explained in detail, relating them to the concepts presented in Principles of Simulation, along with the specifications established at each stage.

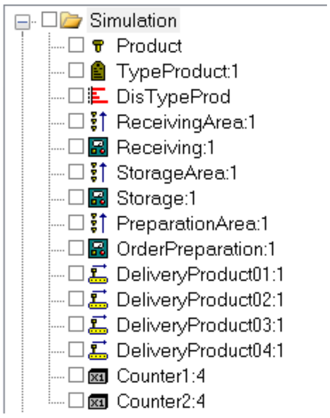

Figure 9: Elements used in the simulation.

To correctly create the concept of the product we intend to use, three distinct elements have been defined. The first element is of type *Part*, which is the only entity in the process and refers to the products in general. In the element details, we set an arrival period of 30 minutes[1] and an unlimited maximum arrival capacity. Additionally, to specify that there are four different types of products, we have defined the *Attribute* element, which, as explained earlier, indicates the characteristics of the product. Specifically, this attribute establishes that there are four different types of products. This attribute is linked to the last element in this section, which is of type *Distribution*. This variable has allowed us to set an equal probability of product arrivals to the model, ensuring randomness. To establish the relationship between the three elements, it is necessary to program a section called "Actions on Create" within the product

[1] It is important to note that the Witness software sets the minute as the default unit of time measurement. This should be considered to ensure no issues arise in the workflow. In our case, all the times used have been directly input in minutes.

details. The code used is shown in Figure 12. Each product has been associated with a colour to make it easier to differentiate them in the model.

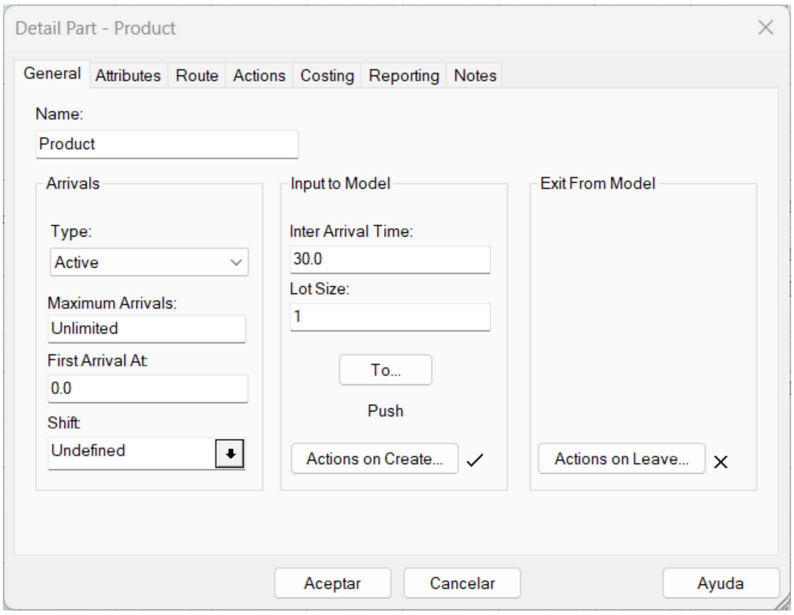

Figure 10: Details of products.

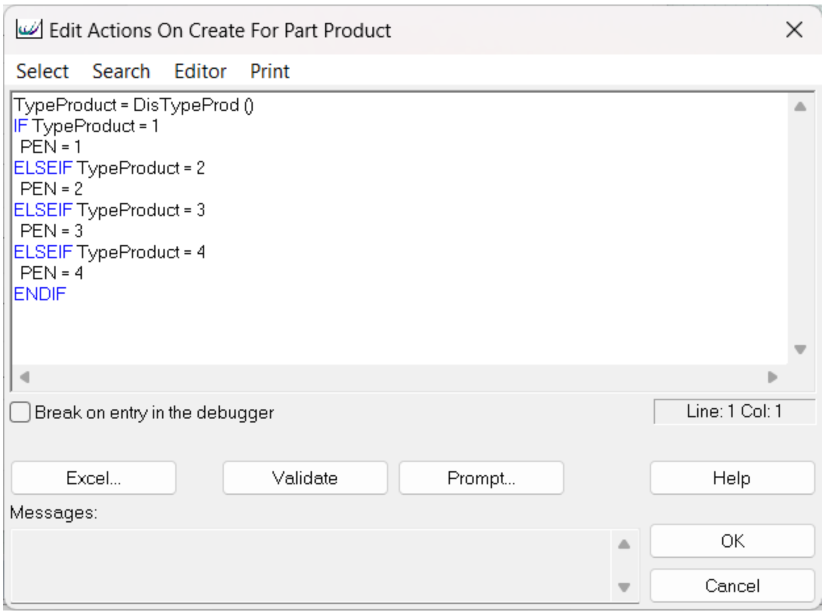

Figure 11: Actions to create products.

Continuing with the model, we have another type of element called a *Buffer*. As the name suggests, this is the area where entities wait before entering the next section. There are three of these elements in the model, referred to as "Area," located in front of each system resource. The only notable characteristic is a maximum capacity of 10 products, which will help us observe issues such as bottlenecks in the model.

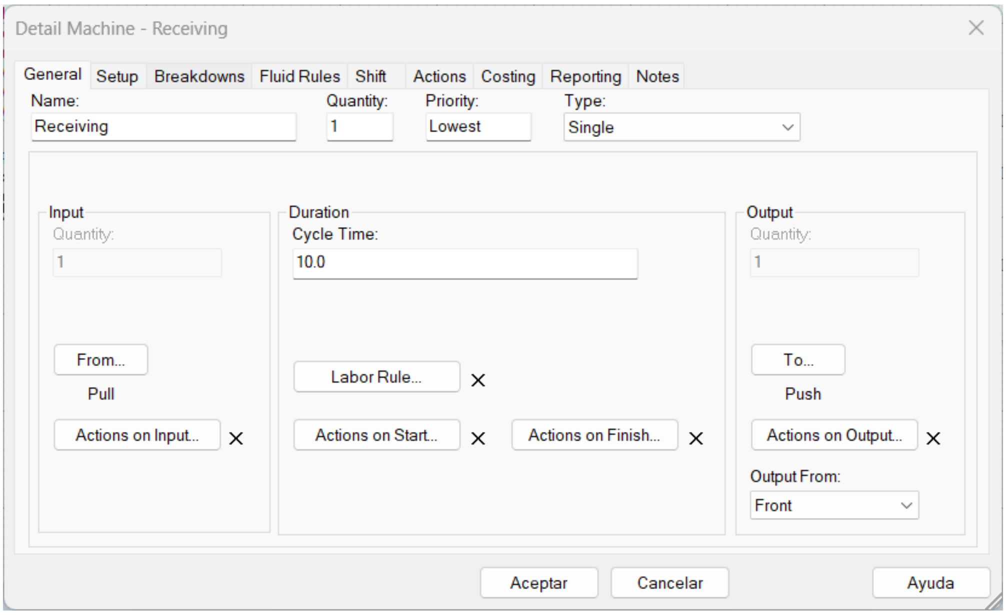

Figure 12: Details of Machine 1, Receiving.

The next elements to discuss are of type *Machine*; these are the resources present in the system, and there are three of these elements, one for each area where activities are performed in the model. We have the receiving machine, the storage machine, and the machine responsible for preparing the orders. As mentioned in the previous paragraph, just before each of these elements is a buffer that serves as a link between them.

The receiving machine has been assigned a processing time of 10 minutes per product; we can see it in Figure 12.

A cycle time has been set on the storage machine using a uniform distribution so that the products are randomly stored for a maximum of 1 hour before moving to the next stage.

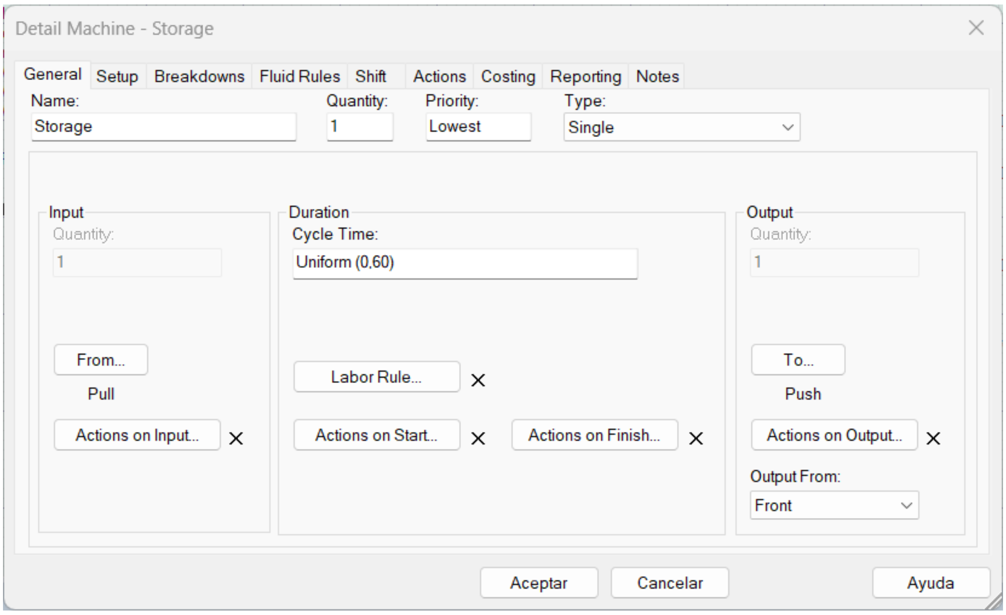

Figure 13: Details of Machine 2, Storage.

The last machine is where the products are prepared to be sorted by type and then distributed. Here, a uniform distribution has also been used, establishing random times between 10 and 20 minutes for each product to be distributed.

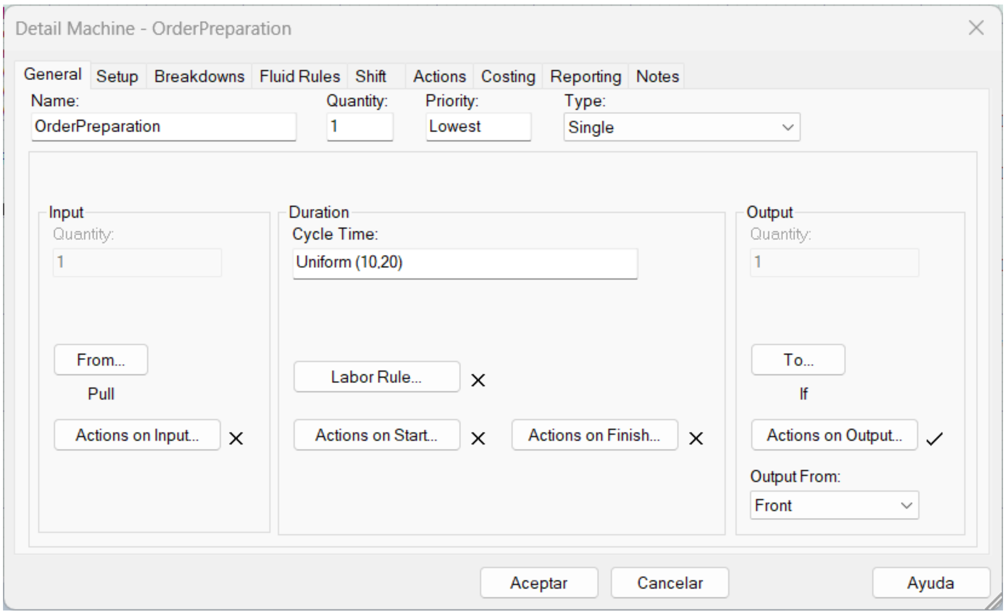

Figure 14: Details of Machine 3, Order Preparation.

Before continuing with the characteristics of this last machine, it is necessary to introduce a new element. This is of type Counter, and there are two in the model. These have been used to count the products passing through the model. One of them is located at the output of the product preparation machine to keep track of the products of each type that are ready for distribution, and the other is located at the end of the distribution lines to confirm the products that have reached the end of the model. To make use of these elements, lines of code have been added both in the output specifications of the preparation machine and in the distribution lines to enable counting. This code is shown in Figures 15 and 21.

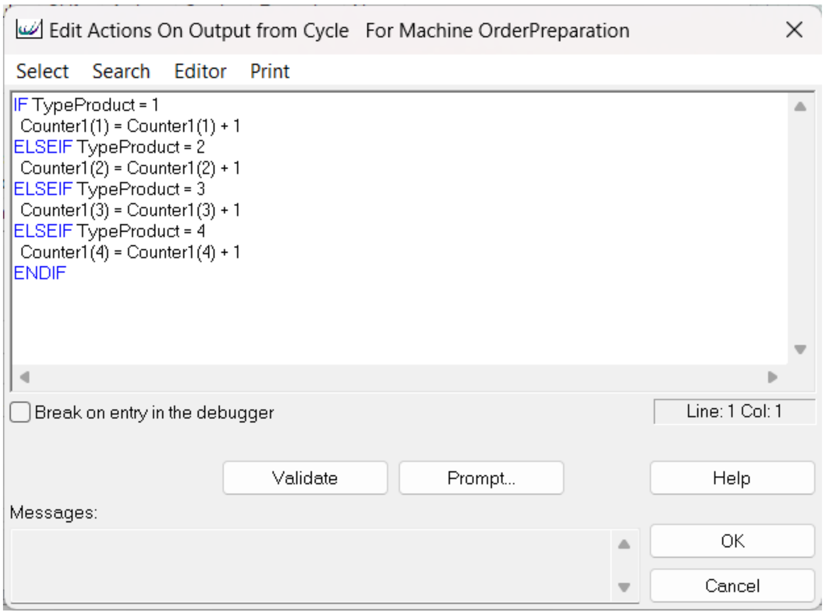

Figure 15: Actions on output for Order Preparation.

In addition to the previous code, an action needs to be added to the preparation machine that separates each type of product into different distribution lines, as shown below.

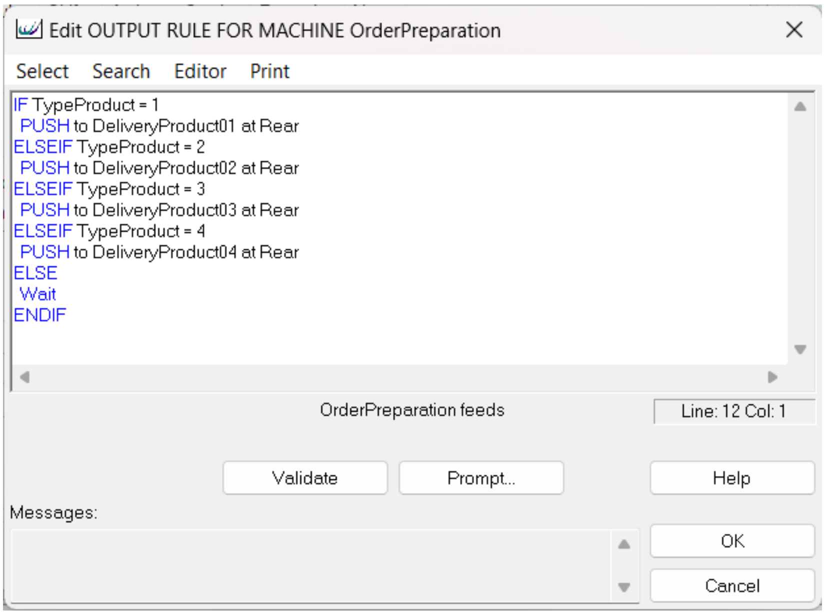

Figure 16: Output rules for Order Preparation.

Finally, the last type of element is a Conveyor. There are four of them in the model, one for each type of product. These elements are responsible for distributing the products to their corresponding destinations.

The characteristics of these elements include setting an equal length for all, a different maximum capacity for each one to match the required demand for each product, and finally, each distribution line has a different time for the product to be distributed. The next four figures show all the specifications assigned to each production line. It is important to note that we added the counter code mentioned earlier in the details of all the lines.

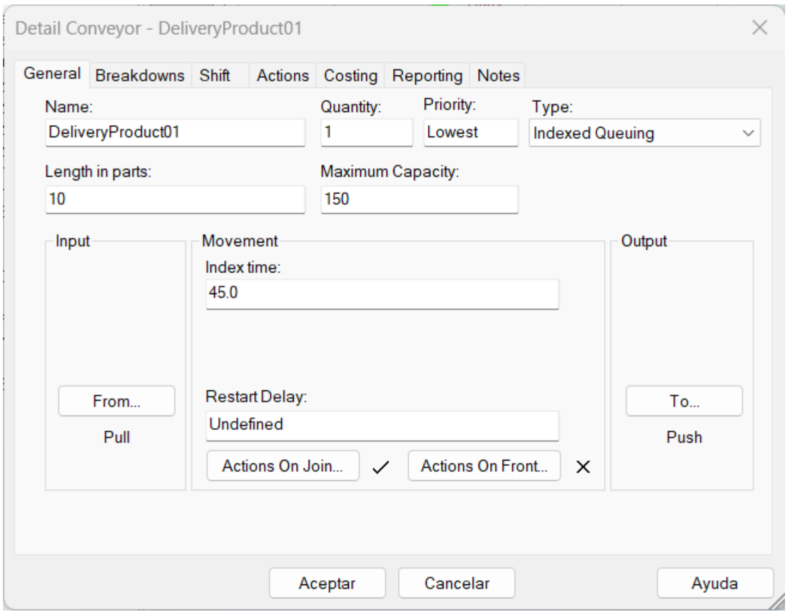

Figure 17: Details of Conveyor for deliver product 1.

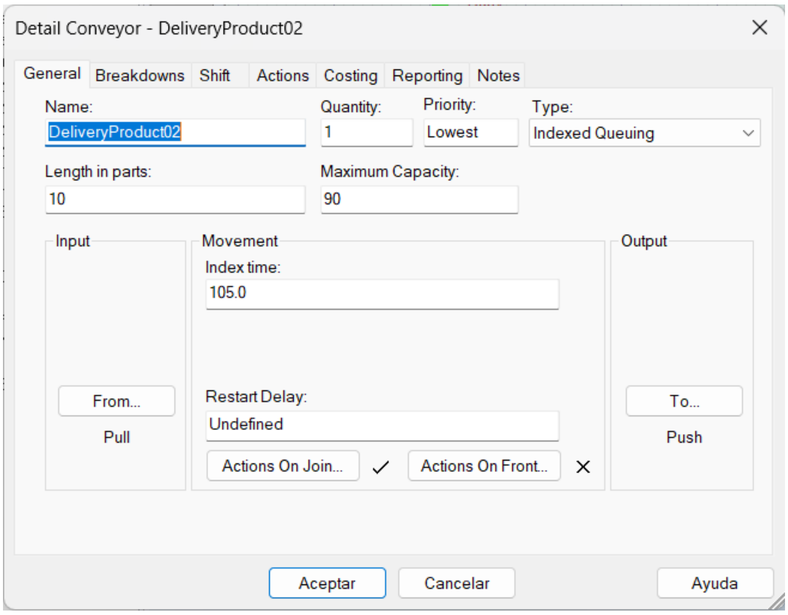

Figure 18: Details of Conveyor for deliver product 2.

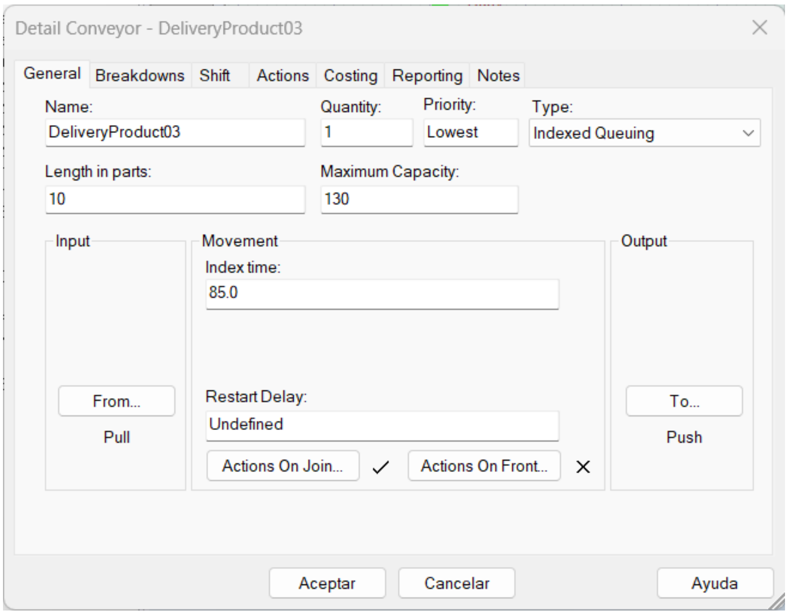

Figure 19: Details of Conveyor for deliver product 3.

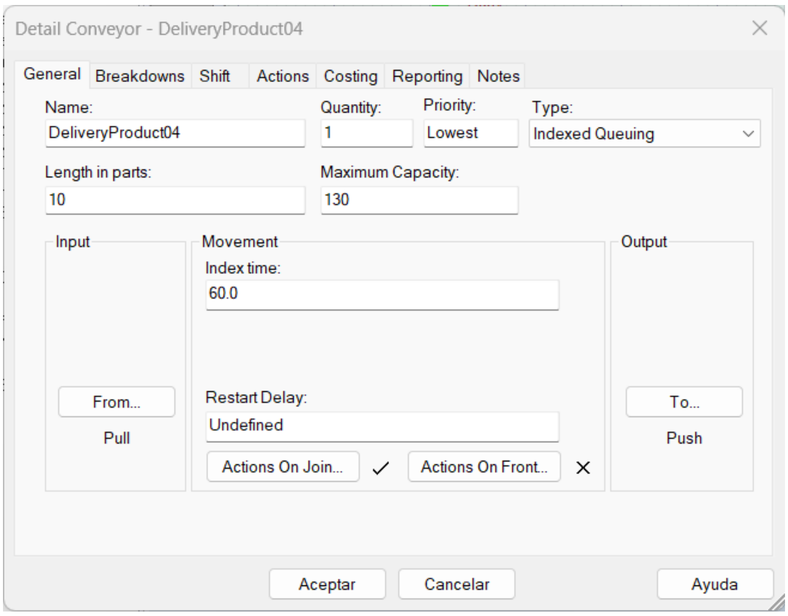

Figure 20: Details of Conveyor for deliver product 4.

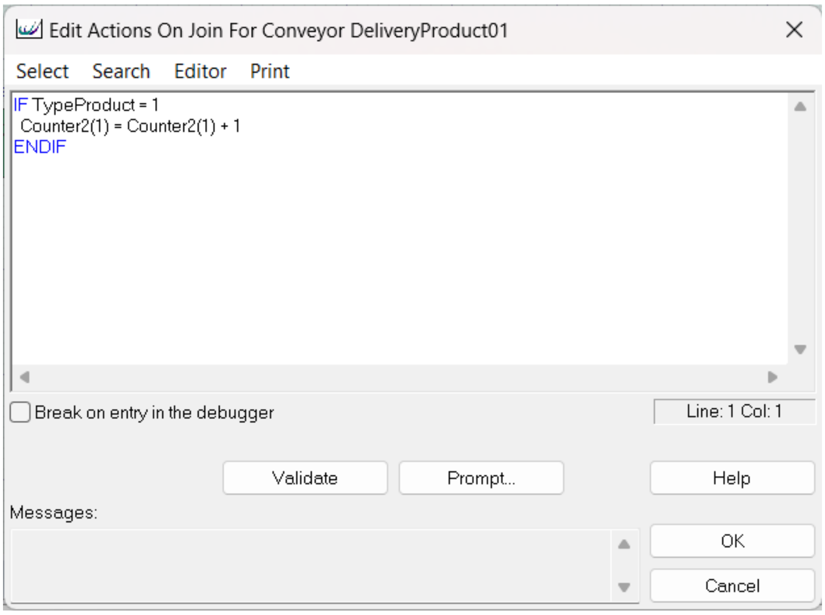

Figure 21: Example action done in every conveyor.

This has been the complete description of the model creation for the simulation in the Witness program.

Appendix B: Work in Orange

To ensure our model functions correctly in the Orange software, one of the first steps was obtaining data from the simulation created with Witness. The compatible file type for this program is CSV. This type of file is used to store tabular data in a simple format, characterized by the arrangement of data in rows representing records and columns where each one corresponds to a field of that record. These files are easy to read, exchange, and overwrite, which makes them very useful. The data exported from Witness had to be adjusted in Excel because, although they were exported correctly, they were not organized properly for the program to recognize them. The different fields in our study include the product ID, product type, arrival time, storage time, preparation time, shipping time to the destination, and the total operation time.

Below is the conceptual map used in the process.

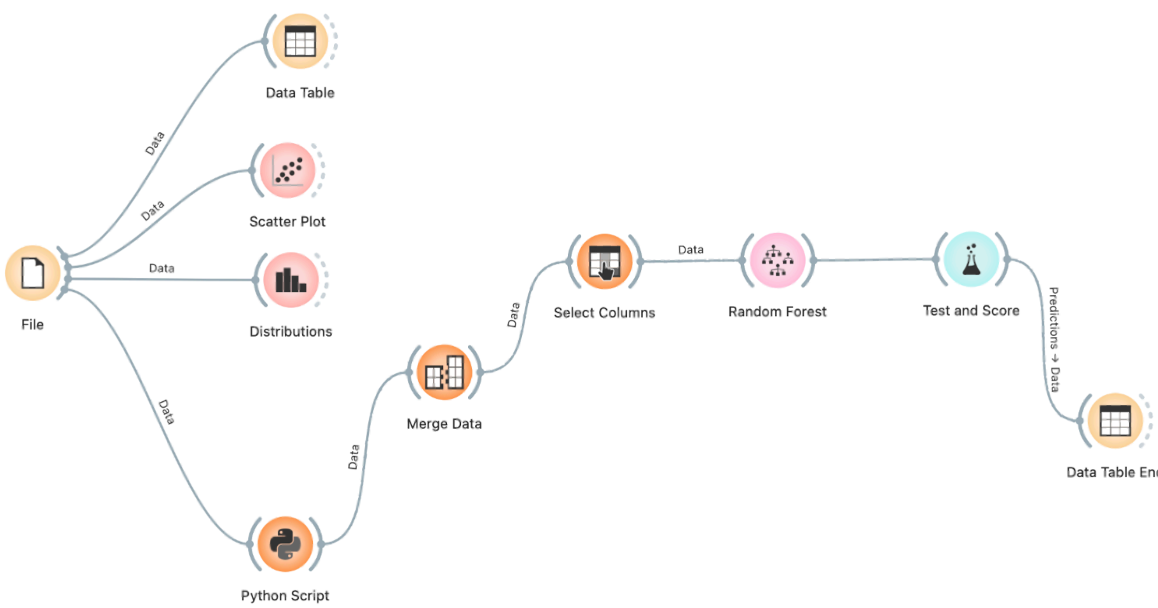

Figure 22: Concept map in Orange.

The CSV file generated was uploaded into the *File* widget. Once the file is uploaded, the program automatically recognizes the fields and assigns them a classification within the model. There are four roles: the first is "feature," which represents a general characteristic of the model, and here all time-related fields are classified, as they are randomly generated in this case study. Another role is "skip," used for fields that can be ignored within the model. For numerical fields, there is the "meta" role, which in this case is only used to identify each record. Finally, the most important role is "target," which identifies the field that is the focus of the study; in our case, since we aim to establish a proper order of product arrivals, the appropriate target field is the product type.

	Name	Type	Role	Values
1	Type	N numeric	skip	
2	Product	C categori...	target	Product A, Product B, Product C, Product D
3	Arrival Time	N numeric	feature	
4	Storage Time	N numeric	feature	
5	Order ...	N numeric	feature	
6	Delivery Time	N numeric	feature	
7	End Time	N numeric	feature	
8	ID	S text	meta	

Figure 23: File features.

Next, I will briefly explain the widgets used in the model. The *Data Table* widget is used to display all the generated data in an organized manner—the first one is used to check that the integrated data is correct, and the second to visualize the output generated by the program. The *Scatter Plot* and *Distributions* widgets are used to visualize initial results and relationships in the generated data file. In the *Python Script* widget, a script was written to generate new data records and optimize the product order. The *Merge Data* widget is used to combine all data into a single document. *Select Columns* is used to specify which data should be used.

For the learning part of the model, the *Random Forest* widget is used. This is a machine learning algorithm that makes predictions in classification and regression problems. It works by creating multiple decision models and combining the results to improve accuracy and prevent overfitting. Finally, the *Test and Score* widget is responsible for evaluating and comparing the performance of the various machine learning models generated. It calculates metrics such as accuracy, error, and sensitivity to determine which model is the most suitable for our dataset.