Designing a virtual warehouse operator integrating fatigue, recovery and learning using agent-based modeling

Worten Case Study

Régis Jean François Malpas

Department of Engineering and Management, Instituto Superior Técnico

Abstract

The retail market in Portugal has undergone significant changes due to the rise of online retail. Worten, a market leader in consumer electronics and home appliances, has successfully adapted to this transformation by expanding its online offerings and adopting an omnichannel distribution strategy. However, the company's supply chain, particularly the store and online items preparations area responsible for sorting and packing items, is unable to meet the growing demand. To address this issue, Worten plans to automate operations and implement an automated item sorter in its Azambuja warehouse. To finalize this project, it is important to understand its impact on the workforce. This study consisted in creating a virtual agent using agent-based modeling to simulate operators' behavior and sensitivity to their environment. The goal was to develop and calibrate a human-like agent that can be used for future work to define the key characteristics of the equipment and required workforce to operate the automated sorter using agent-based simulation. By developing a human-like agent, productivity and accuracy were analyzed considering various environment control variables. The methodology involved selecting relevant factors, researching mathematical models, collecting data, calibrating and validating the models, and developing a simulation model. The results of the simulation model analyzed the impact of changes in schedules, experience level, task intensity, rest quality, and layout on operator productivity and accuracy. The findings revealed that fatigue and experience significantly influence operators, with factors such as task intensity, rest quality, layout, and schedules affecting their performance. Overall, the developed virtual agent provides valuable insights for defining the equipment characteristics and workforce requirements to operate the automated sorter effectively.

Keywords: human factors, agent-based modeling, fatigue, recovery, learning, human error, warehouse

1. Introduction

The retail market in Portugal has gone through significant changes in recent years, particularly with the rise of online retail. Customers have become more knowledgeable and demanding. Worten, a consumer electronics, home appliances, and entertainment market leader in Portugal, has adapted to the changing retail landscape by expanding their online offerings and adopting an omnichannel distribution strategy. This omnichannel distribution system merges the physical world with stores providing specialized services and digital channels offering a wide variety of products.

However, Worten's supply chain cannot currently meet the predicted demand. More specifically, the store and online items preparations area which sorts and packs items for shipping is impacted. There is a need for the automation of operations to increase capacity, throughput and reduce errors. As a result, the company plans to automate some of its processes and redesign its operations by implementing an automated item sorter for the online and retail area in its Azambuja warehouse. Worten's engineering and logistics teams have studied the design and implementation of an automated sorter. There is now a need to study the impact on operations, more specifically on the workforce before the final implementation of the project.

This study focuses on creating a virtual agent imitating operators' behavior and sensitivity to their environment in the context of warehousing using agent-based modeling. This virtual operator allows the measure of productivity and accuracy, influenced by various environment control variables. The environment's control variables include the layout and scheduling of the operations, the fatigue (physical intensity of the tasks and recovery rate of rest periods) and the experience (the learning rate and the previously acquired experience) of the operator. This agent is developed for the

Worten case study based on an existing sorting operation, the put to light sorting activity of the preparation area dedicated to retail, which uses human resources in a similar way as the automated sorter. The goal is to develop and calibrate a human-like agent that can be used for future work to define the key characteristics of the equipment and required workforce to operate the sorter using agent-based simulation.

2. Case study

2.1 Worten

Worten is a leading electronics, consumers, and entertainment products retailer in Portugal. It was founded in 1996 and has become one of the largest electronics retail chains in the Iberian Peninsula.

2.2 Worten warehouse

At first, Worten had a linear supply chain where suppliers delivered items to the central warehouse in Azambuja which prepared orders to be shipped to retail stores in Portugal. Over the years, this supply chain underwent a deep reconfiguration, focusing on the omnichannel strategy and the consequent introduction of different flows of products and information. Worten moved on to a networked supply chain which resulted in the creation of different flows such as logistic hubs, physical supply stores, and online stores, as well as the integration of drop shipping and cross docking concepts.

Worten's warehouse located in Azambuja covers an area of 50,000 square meters, employs around 300 workers and stocks 70,000 stock keeping units (SKU). The warehouse faces a significant challenge due to the diversity of products offered. The products are split in two main groups, small (708) and big-sized appliances (701).

2.3 Warehouse flows

The flows are divided into two categories, inbound and outbound flows. These flows can be divided again according to their recipient. If the recipient is a business, it corresponds to a Business-to-Business (B2B) flow, if the recipient is a final consumer, it is called a Business-to-Consumer (B2C) flow.

The inbound flows consist of the path taken by merchandise received by the warehouse before it is used to prepare orders. This flow is only B2B because merchandise is only provided by other businesses. This flow is then divided into three paths depending on the characteristics of the products and of its supplier: pick by line (PBL), pick by store (PBS) and long tail (LT). PBL flow is essentially crossdocking. In PBS flow, items are stored before shipping and not assigned to orders yet. LT flow corresponds to customers' online orders of SKU that are only available on Worten's marketplace.

The outbound flows consist of the path taken by merchandise once it is used to prepare orders. The B2B outbound flows are split into three categories, which are PBL, PBS and online. The B2C outbound flows are split into two categories, which are Online and complementary service home delivery. The PBL and PBS flows correspond to orders from the retail channel and are shipped to stores. PBL corresponds to the crossdocking component, and PBS to items available in storage. Online B2B corresponds to store delivery of online orders. Online B2C corresponds to online orders from customers.

2.4 Warehouse operations

Inbound operations consist of all the activities regarding merchandise reception and processing before it is used to prepare orders. Outbound operations consist of all the activities taking place when an order is received. The focus of this study are the Online and PBL flows and operations which will be affected by the sorter. The manual sorting activity of these flows will be automated by the sorter.

The online area must provide short delivery times, following Worten's commitment to deliver items within 24 hours. Customers can either receive their order at home (HD, home delivery) or pick it up in store (PIS, pick up in store). The online preparation area receives products from the reception area (LT products) and from the 708 storage areas, therefore performing crossdocking and traditional picking. There are different ways of processing orders, depending on their type, size and destination. Most online orders are picked, packed and labelled. Then, they are dropped on a conveyor belt and travel to an accumulation table. At the accumulation table, items are sorted in different totes depending on their end destination. Through the put to light processing, the items are dropped into different store boxes, then consolidated in the PTS.

PBL operations are composed of three different processes, which are the sorting activities PTZ 1 and PTL, and the assembly and consolidation activity PTS. Items are sorted in the PTZ 1 area according to the different PTL sector. The PTL activity is divided in three parts depending on the size of the item processed: GRA, MED and PEQ, corresponding to big, medium and small items. In the PTL area, each box represents a Worten store. Operators scan each product from the tote. PTL technology lights the corresponding box for that SKU along with the quantities required. The operator drops the items in the box and pushes a button to confirm the drop off. These store boxes are then taken to the PTS area to store pallets.

2.5 Problem definition

Worten has witnessed a consistent rise in sales, both through their online platform and physical stores. As a result of this growth, the store (PBL flow) and online small items (Online flow) preparation segment has been affected. This involves the packing and sorting of items into boxes for shipping. The introduction of automated solutions is imperative to expand capacity, increase speed, and eliminate errors. However, such a machine represents a substantial investment, and it is crucial to evaluate the impact on operations, especially regarding workforce requirements and planning. To do so, a simulation model was developed using AnyLogic and used to dimension the sorter, its layout and various characteristics. To properly evaluate the impact on workforce in terms of system throughput, FTE productivity and interactions with workers, there is a need to accurately represent the human factor in the simulation. To this end, this research focuses on building a virtual operator agent sensitive to its environment that behaves like a human worker.

3. Literature review

The literature review first focuses on defining the origin of the problem. Then, the importance of operators is explained. Thirdly, agent-based models are introduced. Finally, the factors affecting operators' productivity and accuracy are analyzed.

3.1 Introduction

The onset of Industry 4.0 in the warehousing and logistics industry opened the door to new innovations and redefined the role of the operator (Gorecky, et al. 2014). Despite the advent of automation in recent years, the logistics industry is heavily dependent on human labor, which drives most of the operational cost of warehouses. Cimini et al. note the importance of the human factor in logistics: *"During a transition to Logistics 4.0, it is necessary to consider all the issues that affect logistics, among which human factors are considered to be the most important."* (Cimini, et al. 2021). The human factors (or ergonomics) are defined as: *"[…] the scientific discipline concerned with the understanding of interactions among humans and other elements of a system."* (IEA 2016). Decision support models therefore need to integrate human factors to provide realistic and reliable insights for decision makers. They are widely used in production and logistics, as they are valuable tools for decision making. However, they often overlook the human factor, leading to unrealistic planning outcomes and harmful work schedules (Grosse, et al. 2017). Grosse et al. insist on the importance of considering human factors to produce valid decision support models. Such incorporation of human factors into simulation models is an emerging field of research, for which the necessity was outlined 20 years ago (Boudreau, et al. 2003). Grosse et al. note that this call was only answered recently, with a constantly growing number of research on incorporating human factors in decision support models for production and logistics (Grosse, et al. 2015). However, Grosse et al. also note the remaining gap in literature concerning the interaction between the worker and its environment, stating that research carried out so far only had a minor focus on warehousing and focused on production and assembly, which is peculiar due to the labor-intensive nature of warehousing (Grosse, et al. 2017). Following an extensive review of the state of the literature, they concluded that decision support models for logistics have mostly neglected human factors so far. There seem to be a gap in literature concerning the design of a realistic logistics human operator (Liang, et al. 2022). This work is a first step to fill this gap.

3.2 Workforce

The operational cost of warehouses is mostly driven by labor cost (Herrmann, et al. 2019). Cost efficiency is vital for survival in manufacturing and warehousing but is challenging due to the demand for shorter delivery times and higher service levels. Businesses must maximize the use of physical and human capital. This means planning the workforce to match the workload using workforce planning, by determining the number of employees required for each planning horizon. In order to properly assess the work capacity, it is crucial to include uncertainties due to the heterogeneity of employees, leading to differences in productivity and accuracy (De Bruecker, et al. 2015).

3.3 Agent-based models

Decision support models can be implemented using different programming paradigms. In agent-based models (ABMs), there is the possibility to simulate every individual of a population to almost any level of detail (DeAngelis and Diaz 2019). Each individual agent can integrate state variables which represent its internal states and have a unique record of interactions with its environment and other agents.

3.4 Human factors affecting performance

The most important consequences of human factors concerning performance are task duration and error probability (Petronijevic, Etienne and Dantan 2019). Using fundamental concepts of interactions between humans and systems, the relevant human factors for logistics and production systems are divided into four categories (Grosse, et al. 2015): perceptual, cognitive, motor and psychosocial. Due to the labor-intensive nature of logistics and production, which heavily relies on physical work, the focus is on the motor category. More specifically, the focus is on fatigue. The cognitive abilities through learning are also explored. This is particularly relevant in an industry that relies heavily on temporary workers to answer peaks in demand, with many activities showing clear learning curves (Glock, et al. 2019).

4. Methodology

In order to design a human-like operator agent to implement in the automated sorter project, the following methodology was developed. Firstly, the relevant factors conditioning the operator behavior are selected. Secondly, an existing activity in Worten warehouse having a similar use of human resources as the automated sorter project is selected. Thirdly, mathematical models of the relevant factors are researched and adapted to the case study. Fourthly, the data needed to calibrate the mathematical models is defined and collected. Fifthly, the data is used to calibrate the agent parameters. Sixthly, the mathematical models of the relevant factors are validated by comparison with the data. Lastly, the simulation model is developed. The selected activity is modelled with its specific characteristics matching reality. The human operator agent is designed using the calibrated parameters and the mathematical models of the factors influencing its behavior.

4.1 Relevant factors

After careful consideration of all possible factors in the literature review, two factors influencing workers' productivity are considered for this case study: fatigue and experience. The impact of layout, scheduling, task intensity and rest quality will be assessed through these two factors. The layout will determine the distances travelled, therefore influencing fatigue. Scheduling will determine the length of work and rest periods, therefore influencing fatigue and experience. Task intensity and rest quality determine the rate at which fatigue increases during work and decreases during rest periods. These factors were selected using the review of similar research in logistics and by performing informal interviews with Worten's warehouse operator and managers.

The first model depicts fatigue. The changes in fatigue will be translated into variable task duration, variable travel speed of the operator, and variable probability of errors. The second model depicts experience. Operators acquire experience through their cumulated work time. The increase in experience will be translated into reduced task length and lower probability of errors.

4.2 Relevant activity

After analyzing the implementation of the sorter project extensively, an existing activity taking place in the warehouse using human resources in a similar manner can be selected. Once the automated sorter is implemented, the items will be picked manually, placed on the conveyor belt, and automatically transported and dropped in the right box according to their destination. Then, operators will have to arrange packages into the boxes to maximize the utilization of space, close and push them onto a conveyor belt when full. This is similar to the put to light sorting activity of items in the preparation area of Worten's warehouse. More specifically, this study is carried out in the PTL GRA sorting area dedicated to retail stores. This activity concerns the PBL flow of items below 6 liters of volume ordered by Worten's stores to replenish their stock.

4.3 Fatigue and experience model

The mathematical modeling of fatigue can take different directions. However, most researchers in logistics and warehousing opt for exponential models for fatigue (Givi, Jaber and Neumann 2015). Fatigue is denominated by $I \in [0,1[$, $d \in [0,1]$ is a coefficient indicating the physical intensity of the task, and w is the cumulated work time in hours, as depicted in Equation 1 (Ferjani, et al. 2015).

$$I(w) = 1 - e^{-dw} \tag{1}$$

Similarly, rest is denominated by $R \in [0,-1[$ and $r \in [0,1]$ is a coefficient indicating the quality of the rest, and x is the cumulated resting time in hours, as depicted in Equation 2 (Ferjani, et al. 2015).

$$R(x) = e^{-rx} - 1 \tag{2}$$

The current state of fatigue $F(w,x)$ is therefore given by the sum of the accumulated fatigue and the accumulated rest.

Many models regarding learning have been developed over the years (Glock, et al. 2019). They evolved to become the S-curve, adapted for warehouse operations (Carlson 1973).

$$y_w = y_1\,(M + (1-M)(w+B)^{-b}) \tag{3}$$

$$b = -\log(LR) / \log(2) \quad (4)$$

In these equations, y_w is the time taken to produce a unit after w hours of work, y_1 the time to produce the first unit. M represents the automation level of the process. $M = 0$ if the process is fully manual and $M = 1$ if the process is fully automated and not subject to learning. Prior experience is denoted as B. The learning exponent is b, and LR designates the learning rate.

The experience factor $E(w, B)$ corresponding to a change in task length after w hours of acquired work experience, including B hours of previous experience is obtained in Equation 5.

$$E(w, B) = y_w / y_1 = M + (1 - M)(w + B)^{-b} \quad (5)$$

This is particularly relevant for the case study because it integrates previously acquired experience.

4.4 Impact on productivity and accuracy

An adjusted task duration D_a can be calculated, considering the state of fatigue $F(w, x)$, the state of experience $E(w, B)$ and the original planned task duration D_0, using a similar logic as Petronijevic et al. (Petronijevic, Etienne and Dantan 2019).

$$D_a = \left(1 + \alpha\, F(w, x)\right) * E(w, B) * D_o \quad (6)$$

The coefficient α is used to bind the effect of fatigue on task duration. Similarly, the agent travel speed is affected by fatigue. Digiesi et al. demonstrated that the movement speed of workers slows down as fatigue increases (Digiesi, et al. 2009). The agent travel speed is denoted S_f and S_b is the base travel speed without fatigue.

$$S_f = S_b * (1 - \alpha F(w, x)/2) \quad (7)$$

Givi et al. developed a mathematical model estimating the human error probability (HEP) when performing an assembly job under the influence of learning-forgetting and fatigue-recovery (Givi, Jaber and Neumann 2015). In this case, the model is adapted to the context of warehousing not to consider the forgetting aspect. The HEP is given by Equation 8.

$$HEP = 10^{\log\left(0.5 * \left(\frac{D_a}{D_0} + \frac{F(w,x)}{0.99}\right)\right)} \quad (8)$$

It takes into account the adjusted task duration D_a, and the original planned task duration D_0. Because D_a is a function of experience and fatigue, the HEP takes into account fatigue and experience.

4.5 Data collection

The goal of the data collection is to calibrate the virtual operator agent to match the characteristics of a real human operator operating in the PTL GRA activity of Worten's warehouse. To do so, data showing the impact of fatigue and experience is collected. By fitting the fatigue and experience mathematical functions on the data, the parameters of the curves are calibrated. These parameters are then implemented in the virtual agent, and simulation is used to assess the impact of layout, scheduling, task intensity and rest quality on productivity and accuracy.

The task studied must be precisely defined. The time to process a full tote containing a batch of items is measured, with the number of different SKUs in the batch. The full batch time is divided by the number of SKUs, giving an average time to process one SKU (composed of one or multiple identical items).

To measure the impact of the fatigue level on task duration, the experience factor $E(w, B)$ must be kept constant. To do so, data is collected on a highly experienced operator. By measuring the difference of processing time of the first and last units of an experienced operator, the parameter α is set.

To measure the impact of experience on task duration, the fatigue factor $F(w, x)$ must be kept constant. To do so, the task duration of operators with different experience levels is measured at the beginning of their shift, so that fatigue has no influence. By comparing the task durations of these operators with different experience levels, their experience factor and learning rate are assessed.

4.6 Fatigue and experience model calibration

Using data collected on a novice and an experienced operator, the parameters of the agent can be calibrated to fit the operations of the PTL activity in Worten's warehouse. To fit the parameters to the data, trendlines are used. The function type of the trendline is based on the factor that is modelled, exponential function for fatigue, and power function for experience. A trendline displays the global pattern of the data, which is necessary because the PTL sorting process is highly variable regarding processing time of a single SKU. Due to the high variability, the expected processing time provided by the fatigue and experience model is quite unrepresentative of the data collected when the data points are represented individually (for one SKU), instead of averaged over time periods, which is more representative of the average processing time. This is why, to assess the quality of the fit, the r-squared indicator is used on the average processing time over periods of 30 minutes.

The novice operator Régis had no previous experience. The data collected on this operator is used to calculate the learning rate LR, and the base duration of the PTL sorting activity, D_0. The effects of fatigue on task length are not considered in this case because the effects of learning are of greater scale than the effects of fatigue during the first hours of learning a new process. First, the base duration of the PTL sorting activity, D_0 is calculated as the average time to process a SKU during the first hour. The value obtained is $D_0 = 70{,}6\ sec$. Then, by averaging the values of the estimated learning rate for each of the one hour long periods, the learning rate obtained is $LR = 0{,}863$. Using this parameter, the theoretical learning curve can be plotted and compared to the collected data over the workday.

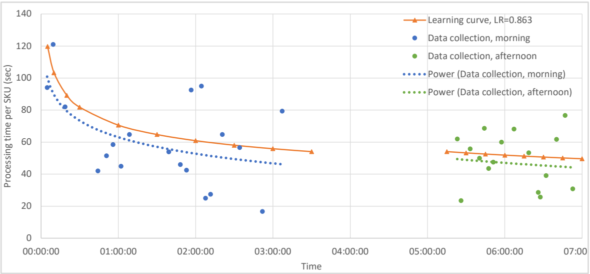

Figure 1 Processing time per SKU, novice operator

The dots represent the data collected. The orange line is the plot using the learning curve function with the learning rate set to $LR = 0{,}863$. Note that the origin of the time axis is set at the beginning of the shift. The blank zone represents lunch break. The blue dotted line is a power trendline of the collected data in the morning ($r^2 = 0{,}30$) and the green in the afternoon ($r^2 = 0{,}54$). The plot of the learning curve displays a similar behavior as the trendlines. Overall, the plot of the learning curve is consistent with the data for a learning rate of $LR = 0{,}863$. The learning curve becomes a great predictor just after a few hours of accumulated experience.

The experienced operator Flavia has three years of experience. Her cumulated experience time is estimated at $B = 3\,000\ hours$ (250 days a year, 4 hours a day of performing the sorting activity). A full workday is 8 hours, but also consists of other tasks. The time to process a SKU D_a with null fatigue is calculated as the average time to process a SKU over the first hour of work, so that fatigue has no impact. The result is $D_a = 20{,}2\ sec$. Knowing her cumulated experience and the base duration $D_0 = 70{,}6\ sec$, the learning rate is $LR = 0{,}897$ and the experience factor $E = 0{,}286$. With 3 years of experience, it takes 71,4% less time to process a SKU compared to someone without any experience. This result is consistent with the LR calculated for the novice operator.

To calibrate the fatigue factor, the fatigue binding factor α, the rest quality factor r and the task intensity factor d must be estimated. The data collected is graphed along with exponential trendlines. The approximated fatigue curve is graphed, and its parameters are adjusted to fit the trendline.

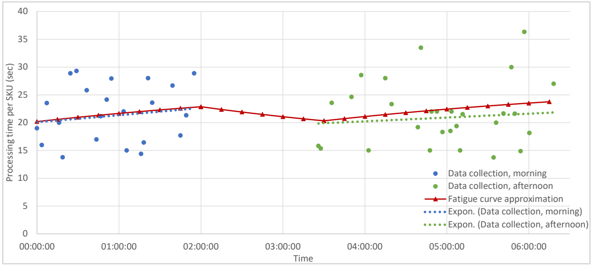

Figure 2 Processing time per SKU, expert operator

The red curve represents the fatigue curve approximation once the parameters are calibrated to fit the trendlines. The blue dotted line is an exponential trendline of the data in the morning ($r^2 = 0{,}38$) and the green in the afternoon ($r^2 = 0{,}30$). The fatigue curve approximation and the trendline match well. However, in this model the task intensity factor is constant, and the fit is not as accurate for the afternoon period. The fatigue curve is slightly above the trendline, meaning that fatigue did not affect the operator as much as expected. Now, the fatigue binding factor α can be estimated. The average time to process a SKU during the last hour of the day is $23{,}6\ sec$. This is an increase of 16,8% in processing time compared to the first hour. The operator estimated that their fatigue level was around 40% at the end of the day. Therefore, a fatigue level of 100% would induce an increase in processing time of 42%. The fatigue binding factor is set to $\alpha = 0{,}4$ so that the maximum effect of fatigue is an increase of 40% of task length. Using the data collected and the fatigue curve approximation, the fatigue and rest factors were estimated by fitting the fatigue curve as accurately as possible on the trendlines. The results obtained are $r = 0{,}25$ and $d = 0{,}20$.

4.7 Fatigue and experience model validation

The parameters of the model regarding the fatigue and experience factors are now calibrated. The model is now tested by collecting data on a third operator, Carla, who has almost 2 years of experience. The test consists in comparing the data collected on Carla to the fatigue and experience curves plotted using the parameters calibrated on Régis and Flavia's data. This will test the quality of the model and its adaptability to other operators with different levels of experience and on different schedules.

Using the parameters set by Régis and Flavia, Carla's experience factor E and expected average time per SKU D_a can be estimated. The results obtained for an experience level $B = 1\,800\ hours$ are $E = 0{,}310$ and $D_a = 21{,}88\ sec$. The expected task duration taking into account fatigue and experience according to the work schedule is graphed.

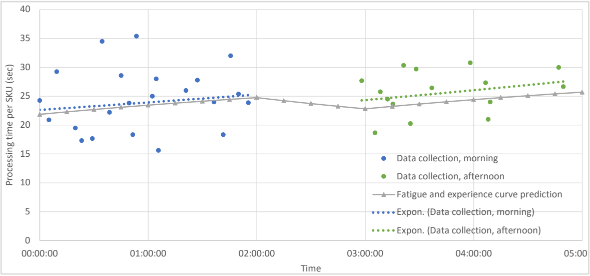

Figure 3 Processing time per SKU, Flavia

The blue dotted line represents the exponential trendline of the data in the morning ($r^2 = 0{,}33$), and the green in the afternoon ($r^2 = 0{,}64$). The grey line is the task length prediction taking into account Carla's experience and based on the parameters calibrated using Régis and Flavia's data. Overall, the prediction is a good approximation of the trendlines, except that the rest period has not been as effective as expected. Indeed, task length after lunch was expected to be lower.

4.8 Simulation model development

The model is developed using AnyLogic, a multimethod simulation software supporting different modeling paradigms. It provides a powerful platform for creating simulations and conducting complex analysis of systems. This model is composed of the main environment, and two agent populations operators and packages. The main environment sets up the process through which agents evolve. Each agent category has specific parameters, variables and functions. The focus of this study is the operator agent, which is developed to contain internal variables setting tasks duration and mistake probability and can therefore be adapted for other applications. However, due to its complexity, it won't be explored in detail.

Main environment

The main environment models the PTL activity. The package agent uses the operator agent as a resource to complete a task. The main environment is composed of the process, the space markup, and the agent components. The PTL activity process is represented as follows.

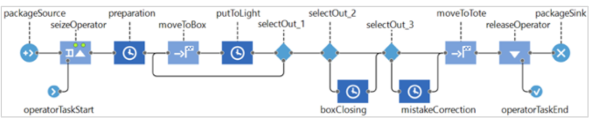

Figure 4 PTL activity simulation process

First, the *packageSource* generates totes containing a customizable number of packages at customizable arrival times. If an operator is idle, he is seized and starts the process. He picks up and scans the item, modeled by the *preparation* delay block. Then, he moves to one of the 24 boxes of the PTL activity. The box is chosen randomly, with each box having an equal probability of being selected. In Worten's PTL activity, the most visited boxes are placed closer to the operator to minimize distances travelled. Once the operator arrives at the box, he performs the activity, by putting the corresponding number of items in the box and pushing the button to confirm the drop off. This is represented by the delay block *putToLight*. From this point, the operator goes through three *selectOut* blocks, defining 4 scenarios: there is another box to visit for this SKU, there is a full box to close, the operator notices a mistake, or there is no other box to visit for this SKU. The probability of being redirected by the first block is 65%, corresponding to visiting on average 2,8 boxes for each SKU processed. The probability of being redirected by the second block to close a box is 10,5%. These numbers are based on the data collected on the operator Flavia. The probability of being redirected by the third *selectOut* block is set by the human error probability function. Depending on the operator's fatigue and experience levels, this number varies between 0,3 and 0,8%. However, this number does not represent the errors measured by the retailers receiving the boxes of packages from the PTL activity. Instead, it represents the number of mistakes noticed during the process and corrected right away. Worten's business intelligence team estimates the percentage of error in the PTL retail sector to be around 0,11% in 2022. These errors are only noticed once the retailers receive and check the contents of the box. It is therefore assumed that around one out of five mistakes are not noticed and cannot be corrected right away.

The time spent in travel blocks depends on travel speed (influenced by fatigue level) and on the distance to travel (determined by the layout). Using the space markup category of AnyLogic, the physical layout of the environment is modelled.

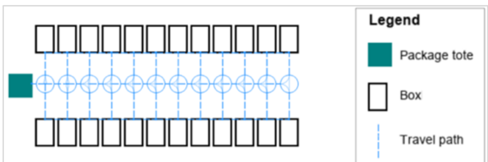

Figure 5 Space markup PTL

The PTL boxes are represented by black rectangular nodes arranged in two rows facing each other. 24 boxes of dimension 50x60cm are represented. The travel path is represented by the blue dashed lines. The green tote contains the packages.

Agent operator

The operator agent represents a human operator in the PTL activity. The agent has been designed in such a way that it calculates task duration and error probability according to its own internal state of fatigue and experience. That way, this agent can be easily adapted for other logistics applications.

4.9 Conclusions

Using the data collected on two operators Régis and Flavia with different levels of experience performing the same PTL sorting task, the parameters of the fatigue and experience factors were calibrated and modelled. Then, the expected performance of a third operator, Carla, was estimated using the previously calibrated model. By collecting data on Carla, the model estimates were confirmed to accurately represent the performance of the operator through the day. They provided accurate task duration estimates, based on the operator's experience and fatigue level. The methodology to calibrate and validate the model has been demonstrated. It can now be applied to any similar activity, such as the automated sorter project. The parameters values estimated for the PTL activity are expected to be similar to the ones representing the sorter project. The two activities use human resources in a similar way, with operators performing comparable tasks.

5. Results analysis

This chapter consists of the presentation and analysis of the results obtained from the simulation model. Using the calibrated virtual operator in a simulation model of Worten's PTL GRA retail sorting activity, the impact of the environment on the operator's behavior is studied. More specifically, the impact of changes in schedules, experience level, task intensity, rest quality and layout are studied. Based on this study's objectives, two metrics are defined. First, operator throughput measuring the average number of SKUs processed per hour. Second, operator accuracy measuring the probability of making a mistake. Scenarios representative of Worten's activity are studied.

5.1 Schedule and operator experience

This scenario is used to measure the impact of scheduling and experience on productivity and accuracy. It simulates operators with different levels of experience on various demand levels. Worten is subject to seasonal variations in demand for its products. The year can be divided into four seasons for which four schedules with their corresponding work time are considered: low (4 hours), average (6 hours), high (8 hours) and extreme demand (10 hours).

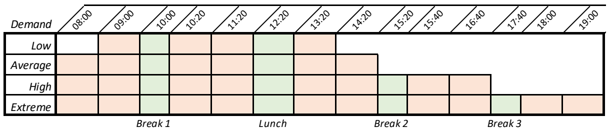

Figure 6 Simulation schedules

The experiences levels are as follow: novice (a week, B=20 hours), proficient (1 year, B=1000h), and expert operator (3 years, B=3000h).

5.2 Task intensity and rest quality

This scenario is used to measure the impact of task intensity and rest quality on productivity and accuracy. Task intensity and rest quality impact the fatigue level, which in turn impacts task duration, travel time and error probability. The operator is an expert so that learning has no impact, and the demand is average to simulate a typical day.

Firstly, task intensity is tested on three different levels: 0,2, 0,4 and 0,6. The first is the value found using the collected data, the second represents a more physically demanding activity and the third is a very intense activity (induces a fatigue of 70% after 2 hours). Secondly, rest quality is tested on 4 levels: 0, 0,25, 0,5 and 0,75. In the first case, resting has no impact on fatigue. The second is the value found using the collected data. The third and last represent a high rest quality.

5.3 Layout

This scenario is used to measure the impact of layout on productivity and accuracy. The layout determines the distance travelled by the operator to reach the boxes. It is changed by assuming different sizes for the boxes. In the warehouse, the boxes are around 50 cm wide, taking into account a small gap between each box. There are 12 boxes next to each other on each side of the alley. The alley is therefore 6 meters long in this base scenario. Two other box widths are studied, 30 and 70 cm, altering the alley's total length. The operator is an expert so that learning has no impact, and the demand is average.

6. Results analysis and discussion

This chapter consists in the analysis of the simulation results.

6.1 Schedule and operator experience

Results on the impact of schedule and experience on throughput are presented in Table 1.

Table 1 Impact of schedule and experience on throughput

THROUGHPUT (SKU/hour)		Operator experience		
		Novice	*Proficient*	*Expert*
Demand	*Low*	101,0	148,3	161,3
	Average	99,2	144,2	155,7
	High	97,9	137,3	150,4
	Extreme	95,7	136,2	150,4

A novice operator's average throughput goes down by 5% changing from low to extreme demand. The change is 7% for an expert. This difference is due to the novice operator benefiting from reduction in processing time due to learning. On an average demand, the proficient operator processes 50% more units than the novice, and the expert 57%. Results on the impact of schedule and experience on accuracy are presented in Table 2.

Table 2 Impact of schedule and experience on accuracy

ACCURACY (%)		Operator experience		
		Novice	*Proficient*	*Expert*
Demand	*Low*	99,48%	99,64%	99,67%
	Average	99,37%	99,53%	99,56%
	High	99,30%	99,46%	99,49%
	Extreme	99,26%	99,42%	99,45%

A novice operator has a 0,52% chance of making a mistake while processing a SKU during a low demand day. This number climbs up to 0,74% during extreme demand. This is because the error probability increases with fatigue. On an average day, a novice has 0,63% of making a mistake, versus 0,44% for an expert, which is a 30% decrease.

6.2 Task intensity and rest quality

Figure 7 depicts the impact of task intensity.

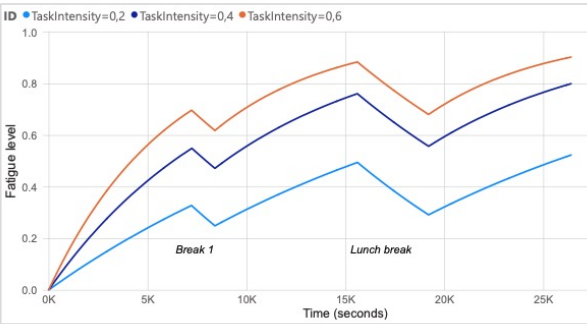

Figure 7 Influence of task intensity on fatigue level

With a task intensity of 0,2 the operator ends the workday with a fatigue level of 0,52. When set to 0,4 this number climbs up to 0,80, and to 0,90 for an intensity of 0,6. This level of fatigue directly impacts the operator throughput. A fatigue level of 1 corresponds to an increase in task duration of 40%. By the end of the day, the production for the 0,4 and 0,6 task intensities are respectively 6 and 10% lower compared to the 0,2 task intensity. Similarly, accuracy is down by 0,16 and 0,21%.

Figure 8 shows rest quality impact on fatigue.

Figure 8 Influence of rest quality on fatigue level

With a rest quality of 0,75, the operator's fatigue level reaches 0 after lunch break. In this case study, the rest quality is estimated at 0,25, giving an estimated fatigue level of around 0,3 after lunch. By the end of the day, the production for the 0 rest quality is 4,7% lower compared to the 0,5 rest quality. Similarly, accuracy is down by 0,16%.

6.3 Layout

The layout impacts the distance travelled, and therefore the processing time. The impact of variations in box sizes is depicted in Table 3.

Table 3 Impact of layout

Average demand, expert		Box size		
Metrics	**Units**	**30cm**	**50cm**	**70cm**
Throughput	*SKU/h*	287,5	233,5	193,3
Accuracy	*%*	99,56%	99,56%	99,56%
Avg process time	*sec/SKU*	18,7	23,0	27,7
SKUs processed	*SKU*	1150	934	773
Distance travelled	*m*	10.357	13.847	15.933
Distance per SKU	*m/SKU*	9,0	14,8	20,6
End fatigue level		0,52	0,52	0,52

Taking the 50cm box as the reference situation, throughput is down 17% for the 70cm box, and up 23% for the 30cm box. The distance travelled is down by 25% for the smaller boxes, and up 15% for the larger ones. This translates into a diminution in processing time per SKU of 18% for the smaller boxes and an increase of 21% for the larger boxes. More interestingly, the operator travels 9, 14,8 and 20,6 meters for each SKU respectively for the small, medium and large box. Note that this simulation does not take into account the fact that smaller boxes have to be closed and replaced more often, and larger boxes less often.

6.4 Conclusions

The operator's productivity and accuracy are influenced by fatigue and experience. Fatigue decreases productivity and accuracy. Experience increases productivity and accuracy. Fatigue is determined by the intensity of the task, the quality of the rest, the layout, and the schedules. The intensity of the task reflects the physical strain or fatigue induced by work. The quality of the rest reflects the impact of resting on fatigue. The layout determines the distances travelled. The schedules define the work and rest periods. Experience is determined by the schedule and the accumulated work hours.

7. Conclusions, limitations, future work

In conclusion, a virtual agent imitating operators' behavior and sensitivity to their environment in the context of warehousing using agent-based modeling, allowing the measure of productivity and accuracy influenced by various environment control variables was created. This agent was developed and calibrated for the Worten case study based on an existing sorting operation, the PTL sorting activity of the preparation area dedicated to retail, which uses human resources in a similar way as the automated sorter project. The agent can be used for future work to define the key characteristics of the equipment and required workforce to operate the automated sorter using agent-based simulation.

It is important to highlight the limitations of this study. The first limitation concerns the data collection and the calibration of the model. Data was collected on two operators for calibration, and one for validation. A larger sample would provide more accurate results. The second limitation resides in the complexity of human behavior and interactions with the environment. Modeling human characteristics is important, as demonstrated, but highly complex. While many factors influence human behavior, only two were selected. The third limitation is that the PTL retail sorting activity uses human resources in a similar but not identical way as the automated sorter.

The agent can now be implemented in the simulation of the automated sorter developed by Worten's engineering and logistics teams. The sorter project will benefit from integrating the virtual operator by providing a more realistic and accurate modeling of human resources.

Bibliography

Boudreau, J., Hopp, W., McClain, J., & Thomas, L. (2003). On the interface between operations and human resources management. *anufacturing & Service Operations Management, 5*(3), 179-202.

Carlson, J. G. (1973). Cubic learning curves-precision tool for labor estimating. *Manufacturing Engineering & Management, 71*(5), 22-25.

Cimini, C., Lagorio, A., Pirola, F., & Pinto, R. (2021). How human factors affect operators' task evolution in Logistics 4.0. *Human Factors and Ergonomics in Manufacturing & Service Industries, 31*(1), 98-117.

De Bruecker, P., Van den Bergh, J., Beliën, J., & Demeulemeester, E. (2015). Workforce planning incorporating skills: State of the art. *European Journal of Operational Research, 243*(1), 1-16.

DeAngelis, D. L., & Diaz, S. G. (2019). Decision-making in agent-based modeling: A current review and future prospectus. *Frontiers in Ecology and Evolution, 6*, 237.

Digiesi, S., Kock, A. A., Mummolo, G., & Rooda, J. E. (2009). The effect of dynamic worker behavior on flow line performance. *International Journal of Production Economics, 120*(2), 368-377.

Ferjani, A., Ammar, A., Pierreval, H., Trabelsi, A., & Aubière, F. (2015, October). A heuristic approach taking operators' fatigue into account for the dynamic assignment of workforce to reduce the mean flowtime. *International Conference on Computers and Industrial Engineering, CIE45, 43*, 65-80.

Givi, Z. S., Jaber, M. Y., & Neumann, W. P. (2015). Modelling worker reliability with learning and fatigue. *Applied Mathematical Modelling, 39*(17), 5186-5199.

Glock, C. H., Grosse, E. H., Jaber, M. Y., & Smunt, T. L. (2019). Applications of learning curves in production and operations management: A systematic literature review. *Computers & Industrial Engineering, 131*, 422-441.

Gorecky, D., Schmitt, M., Loskyll, M., & Zühlke, D. (2014). Human-machine-interaction in the industry 4.0 era. *2014 12th IEEE International Conference on Industrial Informatics (INDIN)* (pp. 289-294). Porto Alegre, Brazil: IEEE.

Grosse, E. H., Calzavara, M., Glock, C. H., & Sgarbossa, F. (2017). Incorporating human factors into decision support models for production and logistics: current state of research. *IFAC-PapersOnLine, 50*(1), 6900

Grosse, Glock, C., Jaber, M., & Neumann, W. (2015). Incorporating human factors in order picking planning models: framework and research opportunities. *International Journal of Production Research, 53*(3), 695-717.

Herrmann, J., Perez, F., Trautwein, V., & Weidmann, M. (2019). *Getting a handle on warehousing costs*. Retrieved from McKinsey & Company: https://www.mckinsey.com

IEA. (2016). *Definition and Domains of Ergonomics*. Retrieved from International Ergonomics Association: http://www.iea.cc/whats/index

Liang, X., Luo, L., Hu, S., & Li, Y. (2022). Mapping the knowledge frontiers and evolution of decision making based on agent-based modeling. *Knowledge-Based Systems, 250*, 108982.

Petronijevic, J., Etienne, A., & Dantan, J. Y. (2019). Human factors under uncertainty: A manufacturing systems design using simulation-optimisation approach. *Computers & Industrial Engineering, 127*, 665-676.