

Improving Warehouse Layout Design: a Probabilistic Simulation Approach

Diogo Filipe Guedes Bento

Master's Dissertation

Supervisor: Maria Henriqueta Dourado Eusébio Sampaio da Nóvoa



Master's in Industrial Engineering and Management

2023-06-26

Abstract

Layout optimization is far from having a straightforward approach. In fact, general procedures for warehouse design are still lacking. Hence, the project in which this work is inserted aims to fill in this void both in the literature and in the market. The ultimate goal is to develop an application that optimizes any user-given layout. However, there are several - and interrelated - decision variables. So, when considering all the degrees of freedom, the problem of warehouse design becomes increasingly complex, if not impossible to solve.

Therefore, this problem requires a framework that is able to: (a) decompose a given warehouse on its building blocks and generate multiple scenarios for each block, (b) optimize the internal layout of the most promising scenarios, and (c) consider the whole warehouse operation and the impact of each optimized block on the overall solution. In order to tackle this, a hybrid approach combining simulation and optimization was proposed. The hierarchical structure of the models used is sequential: Simulation-Optimization-Simulation, hence the name SOS. Due to the shortened time window of this work, when compared to the overall project duration, only the Picking by Line (PBL) block within the first 'S' will be highlighted.

Nonetheless, when it comes to warehouse design, order-picking emerges as the most daunting hurdle to surmount. In fact, all picking areas are unique, so each case is a case. However, they all can be modelled as a grid. Grid modelling is an innovative approach that allows the construction of accurate replicas of any given picking area based on graph theory.

After providing a means to create multiple scenarios for the PBL block, two models were derived to estimate their performances. The first one divides the picking area into geometric zones and treats them independently. The second one, more complex, considers the picking area as a whole. By leveraging a probabilistic simulation approach, both models estimate the performance of any possible solution by predicting the expected travel distances, inside the picking area, for increasing task sizes. Therefore, quick comparisons of several layouts can be made with residual forecast errors and without having to make any physical modifications.

In order to assess the quality of the proposed models, a real-life scenario was needed. Based on the assessment of the "AS-IS" configuration of a large Portuguese retailer's warehouse, the following conclusions could be drawn: (1) the zone model demonstrates better suitability for tasks that only address one zone and (2) for tasks that involve both zones, the mixed model is the preferred choice. It is also worth mentioning the use of oversampling techniques to improve the estimation of zones with limited data availability.

All in all, these findings resulted in a final Mean Absolute Percentage Error (MAPE) of only 14.20%. This indicates the effectiveness of the derived models to provide accurate insights into the order-picking operations. Hence, new scenarios might be considered to further optimize the most promising ones. Moreover, all the methodologies presented throughout this work can be readily extendible to also encompass the Picking by Store (PBS) block, paving the way for future work in this area.

Resumo

A otimização de layouts está longe de ter uma abordagem simples e direta. De facto, ainda não existem procedimentos gerais para o desenho de armazéns. Logo, o projeto no qual este trabalho está inserido visa preencher este vazio, tanto na literatura como no mercado. O objetivo final passa por desenvolver uma aplicação que optimize qualquer layout fornecido pelo utilizador. No entanto, as variáveis de decisão são diversas e, acima de tudo, inter-relacionadas. Assim, ao considerar todos os graus de liberdade, este problema torna-se cada vez mais complexo, se não impossível de se resolver.

Posto isto, este problema requer uma estrutura que seja capaz de: (a) decompor um determinado armazém nos seus blocos fundamentais e gerar vários cenários para cada bloco, (b) otimizar a disposição interna dos cenários mais promissores, e (c) considerar toda a operação do armazém e o impacto de cada bloco otimizado na solução global. Para tal, foi proposta uma abordagem híbrida que combina simulação e otimização. A estrutura hierárquica dos modelos utilizados é sequencial: Simulação-Otimização-Simulação, daí o nome SOS. Devido ao curto período de tempo deste trabalho, quando comparado com a duração total do projeto, apenas o bloco PBL (*Picking by Line*) dentro do primeiro 'S' será destacado.

Apesar disso, neste tipo de problemas, o *order-picking* surge como o obstáculo mais difícil de se ultrapassar. De facto, todas as áreas de *picking* são únicas, pelo que cada caso é um caso. No entanto, todas elas podem ser modeladas como uma grelha. A modelação em grelha é uma abordagem inovadora que permite a construção de réplicas de qualquer situação tendo por base teoria de grafos.

Depois de fornecer um meio capaz de criar vários cenários para o bloco PBL, foram derivados dois modelos para estimar a performance dos mesmos. O primeiro divide a área de *picking* em zonas geométricas e trata-as de forma independente. O segundo, mais complexo, considera a área de *picking* como um todo. Utilizando uma abordagem de simulação probabilística, ambos os modelos estimam o desempenho de qualquer solução com base nas distâncias percorridas dentro da área de *picking* e para tamanhos de tarefas crescentes. Por conseguinte, é possível efetuar rápidas comparações de várias disposições com erros de previsão residuais e sem ter necessidade de efetuar quaisquer modificações físicas.

A fim de avaliar a qualidade dos modelos propostos, foi necessário analisar um cenário real. Com base na avaliação da configuração "AS-IS", de um armazém de um grande retalhista português, foi possível retirar-se as seguintes ilações: (1) o modelo por zonas demonstra melhor adequação para tarefas que apenas envolvam uma zona e (2) para tarefas que envolvam ambas as zonas, o modelo misto é a escolha preferida. É também de salientar a utilização de técnicas de sobreamostragem para melhorar a estimativa de zonas com disponibilidade limitada de dados.

Em suma, estas conclusões resultaram num MAPE (*Mean Absolute Percentage Error*) final de apenas 14,20%. Este facto realça a eficácia dos modelos desenvolvidos para fornecer informações precisas sobre as operações de *picking*. Assim, novos cenários podem ser considerados para futuramente otimizar os mais promissores. Além disso, todas as metodologias apresentadas ao longo

deste trabalho podem ser facilmente alargadas para englobar também o bloco PBS (*Picking by Store*), abrindo caminho para trabalhos futuros nesta área.

Acknowledgments

All good things must come to an end. Memories, however, are timeless treasures of the heart. Looking back at it, many were the ones who supported me throughout this journey as a student. To all of them, this is my 'thank you' note.

To start with, I would like to express my gratitude to Professor Henriqueta for being such a careful and interested supervisor. You gave me strong insights and I learned a lot with your guidance.

To my Professors at FEUP, thank you for all the never-ending challenges you gave me. They certainly shaped me into the person I am today. A special note to Professors Pedro Amorim, Luís Guimarães, and Bernardo Almada-Lobo for giving me the opportunity to be part of such an extraordinary team. You taught me to always think ahead and to aim for nothing less than the stars.

I also would like to thank João Alves for his endless knowledge of warehouse design. You helped me a lot and you always pushed me forward to seek improvements in my research.

Earning a college degree is not an individual achievement. For that, I would like to express my heartfelt gratitude to a few special persons who have accompanied me throughout these years.

To Ricardo, for being the best workmate I could ever have. It has been quite a ride and I am sure it is not the end yet.

To my closest colleagues - João, Inês, Mariana, Helena, and Catarina - who taught me that any problem can be solved with time and calmness. College friends are for life and I want to thank you all for being part of it. This was not the same without you.

Life gave me a lot since the day I joined my first football team. Almost one decade later, friends who played together still stand together. This is for my favourite A-Team: to Pedro and Rui that sometimes are far far way, yet always close; and to Miguel and Tomás that always backed me up and helped me a lot during this last dance.

Some friendships are so long that they eventually evolve into family. I would like to thank Hugo and Afonso for being always at my side since day one of kindergarten and also to Diogo, who joined our circle later, but feels as if he has been part of it from the beginning.

Family means the world to me and I have such an amazing and numerous one. If I were to thank them all, I would make a thesis about it. Still, I am deeply grateful to my cousins that are like brothers to me.

To Dinis and Tiago, thank you for always believing in me (except in football, but that is debatable). You are the next of the Bento's to conclude your degrees and I wish you the best of luck.

To Paulinha, who gave me infinite strength in the times when I needed the most. Thank you for all the trips, songs, movies, and memories that we had together. When you told me that college was the best time of my life, I did not believe in you but it turned out to be true. However, the same can not be told about Gaia being the better city in the world.

To Francisca, for the silent support and for never expecting less than the best for me. Thank you for being the best sister in the world. Your big brother will always be there for you and when you fall, I will catch you and lift you back up.

To Mum, thank you for your patience in dealing with my occasional bad mood. Thank you for projecting your dreams and expectations onto me. Thank you for always having my back and ensuring that I never lacked support or protection.

Some people do not believe in heroes, but they have not met my dad. To Dad, my "Super Mário", one lifetime is not enough to repay your effort, dedication, and affection. Thank you for being the most brilliant person I know. I would not be where I am if it was not for you.

Finally, I would like to pay tribute to my beloved cat and to my late grandfather. Oliver, Vu, this one is for you.

Diogo Bento

*"Foco, determinação e postura positiva
Esse é o meu lema"*

Chico da Tina

Contents

1	Thesis Overview	1
1.1	The Underlying Project	1
1.1.1	Project Landscape	2
1.1.2	Problem Statement	3
1.1.3	Mission and Objectives	3
1.1.4	Proposed Approach	4
1.2	From Project to Thesis	5
1.2.1	Timeline and Scope	5
1.2.2	Proposed Approach	6
1.3	Thesis Structure	6
2	Literature Review	7
2.1	Theoretical Background	7
2.1.1	Retail Market	7
2.1.2	Supply Chain	8
2.2	Warehouse Science	10
2.2.1	The Purpose of Inventory and Warehouses	10
2.2.2	Warehouse Operations	11
2.2.3	Design and Control of Order-Picking Processes	12
2.3	Wrap-Up	19
3	Methodology	21
3.1	The 'SOS' Framework	21
3.2	The PBL Block	22
3.2.1	Grid Modelling	23
3.2.2	Path Cells and Non-Path Cells	24
3.2.3	Aisle-Based Cell Grouping	25
3.2.4	Understanding Layout Zones	27
3.3	Gap to Literature	28
4	First 'S': Simulating Layout Zones	29
4.1	Notation	29
4.2	Assumptions	30
4.3	Mathematical Model	31
4.4	Synthesis	39

5	First 'S': Simulating The PBL Block	41
5.1	Motivation	41
5.2	Notation and Assumptions	42
5.3	Graph Theory	44
5.4	Mathematical Model	45
5.4.1	Breadth First Search Heuristic	47
5.4.2	Handling Multiple Shortest Paths	49
5.4.3	Relaxation	50
5.5	Synthesis	51
6	Case Study And Results	53
6.1	Study Subject	53
6.1.1	The Company	53
6.1.2	The Warehouse	54
6.1.3	Warehouse Data	54
6.2	The PBL Block	56
6.2.1	Hard Constraints	56
6.2.2	Layout Variables	56
6.2.3	Zoning: Variable or Constraint?	58
6.3	Results	58
6.3.1	Zone Model	58
6.3.2	Mixed Model	60
6.3.3	Monte Carlo Simulation	61
6.4	Validation of the Analytical Models	64
6.5	Deriving the Optimal Model	67
7	Conclusions and Future Work	69
7.1	Key Findings and Insights	69
7.2	Scientific Contributions	70
7.3	Main Limitations and Recommendations	70
7.4	Future Research Directions	72
7.4.1	Short Term	72
7.4.2	Long Term	72
7.5	Final Remarks	73

Acronyms and Symbols

BFS	Breadth First Search
COI	Cube-per-Order Index
DES	Discrete Event Simulation
KPI	Key Performance Indicator
MAPE	Mean Absolute Percentage Error
MPE	Mean Percentage Error
PBL	Pick by Line
PBS	Pick by Stock
SKU	Stock Keeping Unit
WMS	Warehouse Management System

List of Figures

1.1	Typical warehouse zones and main activities	2
1.2	Product flow type in PBL vs PBS	3
1.3	An overall perspective of the SOS methodology and its building blocks	5
1.4	Project's timeline	5
1.5	Schema for estimating the performance of any user-defined layout	6
2.1	Supply chain network structure	8
2.2	Integrating and managing business processes across the Supply Chain	9
2.3	Inventory is held at points along the chain as well as in transit	10
2.4	Complexity of order-picking systems (design and control)	12
2.5	Typical distribution of an order picker's time	13
2.6	Typical internal layout decisions in order-picking design	14
2.7	Class-based storage assignment policies	17
2.8	Routing methods for a single-block picking area	18
3.1	An overall perspective of the SOS methodology and its building blocks	23
3.2	Cell as a ground floor location of the PBL block	24
3.3	PBL block modelled as a grid	24
3.4	Flow for a given store with both active and buffer cells	25
3.5	Aisle-based cell grouping	26
3.6	Modelling retrieval aisles as forbidden path cells	26
4.1	Possible picking area with vertical aisles only (Cell granularity level)	29
4.2	Possible picking area with vertical aisles only (Aisle granularity level)	32
4.3	Illustration of all the instances when an aisle $a \in D$ needs to be visited	35
4.4	Illustration of all the instances when an aisle $a \in I$ needs to be visited	37
4.5	Rotations and flips allow the modelling of any zone with only one orientation	39
5.1	Differences between a picking area with one and both orientations	41
5.2	Order of all the possible picking positions for a picking area with mixed zones	42
5.3	Grid convergence problem if $l_c \neq w_c$	44
5.4	Modelling path cells as an oriented graph	44
5.5	All possible movements to fully traverse a path cell	45
5.6	Illustration of the breadth first search algorithm	48
5.7	Construction of all the shortest paths to traverse the maze	49
6.1	Division of the warehouse layout on its building blocks	54
6.2	Daily distribution of the PBL block's picking tasks	55
6.3	Average daily boxes per operation and per work hour	55

6.4	2D top view of the PBL block's current layout	57
6.5	Heat map of the number of picks per picking position	57
6.6	Expected distance for a picking task of size N inside each zone	58
6.7	Modelling the PBL block's current layout as a grid	60
6.8	Expected distance for a picking task of size N inside the mixed zone	60
6.9	Distance distribution of randomly generated tasks of size N	61
6.10	Gap between the mixed model and the Monte Carlo simulation	62
6.11	Distribution of the number of tasks for increasing task sizes	62
6.12	Distribution of the number of tasks for increasing task sizes for each task type . .	63
6.13	3D View of the PBL block's current layout using DES	64
6.14	Daily distribution of the average daily task size	64
6.15	Evolution of each model's errors for an increasing task size of vertical tasks . . .	67
6.16	Evolution of each model's errors for an increasing task size of mixed tasks	68
7.1	Difference between von Neumann's and Moore's neighbourhoods	71
7.2	Grid convergence problems hard to solve for uncommon layouts	71
7.3	Two different store allocation policies for the same layout solution	72

List of Tables

2.1	Picking task examples for the PBL and PBS flows	15
2.2	A summary of the related research	19
3.1	Division of the grid in cell clusters	27
4.1	Cases to address for calculating $p(V_a)$	34
6.1	Expected daily picking tasks	55
6.2	Probability of a pick addressing an aisle	59
6.3	Picking task data deconstructed by task type	65
6.4	Accuracy metrics for the derived models	66
6.5	Accuracy metrics deconstructed by task type	66

Chapter 1

Thesis Overview

Uncertainty became even more of a buzzword since the first COVID-19 outbreak. Shopping in a disrupted world, by the pandemics, opened the path for e-commerce to grow and led customers' expectations to change. Nevertheless, one thing remained certain: retailers, more than ever, need agile supply chains to face this new paradigm.

An efficient supply chain must be based on efficient warehouse operations. However, warehouses are frequently associated with large costs. In fact, a survey in Europe, conducted at Kearney (2022), identified the cost of inventory as being 13% of the total logistics costs, whilst warehousing accounted for a further 24%. This data establishes important motivations to improve warehouse productivity.

Besides, among the various warehouse activities, order picking is the most labour-intensive one, accounting for 60–70% of the warehousing costs. It is, therefore, a first candidate for improvement. By leveraging a probabilistic simulation approach, the main goal of this work is to develop a means to quickly assess, and compare, the expected performance of any possible picking layout.

Bearing this in mind, this chapter aims to first introduce the project, in which this thesis is inserted. Then, the problem at hand will be discussed and the approach used to overcome it is presented afterwards. Finally, the thesis' scope will be defined within the overall project goal and, to conclude, a concise overview of the structure used, to best convey this work, will be provided.

1.1 The Underlying Project

Each warehouse is unique since they are built to best fulfil their companies' needs. Hence, there is not a one-size-fits-all approach regarding warehouse design. The main concern of the project, in which this study is inserted, is to overcome this pain point by developing a product that could easily be adjusted into all these tailor-made layouts. The goal is to, given a certain warehouse, dynamically change its layout in order to maximize efficiency.

1.1.1 Project Landscape

Despite there are not two warehouses exactly equal, every warehouse still comprises three major zones – inbound, storage/picking, and outbound. Their configuration and disposal along the shop floor are what differ. Figure 1.1 highlights those distinct areas and their main activities.

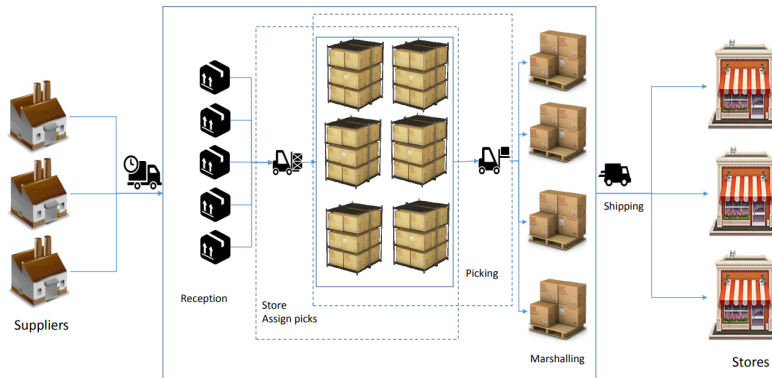


Figure 1.1: Typical warehouse zones and main activities

The inbound area is where suppliers deliver the ordered products according to a pre-defined schedule. Then, operators perform quality and quantity control checks to validate the inbound orders before buffering the arrived items into pallets. Operators also register these movements into the warehouse information system to keep track of all stocks. The pallets will then follow different paths according to the picking policy in use. This project encompasses both the Picking by Line (PBL) and Picking by Store (PBS) policies.

In PBL, the picking task starts with a complete pallet and finishes with it empty. Pickers drag the prepared pallets (picking pallets) throughout their respective circuits inside the picking area, with each circuit being a sequence of ground-floor locations to visit. Moreover, each location represents one store and has store pallets (one active and others in buffer). Hence, pickers deliver the ordered products, by each store, transporting those items in the picking pallet to their respective store pallets. When a store pallet reaches its full capacity, the picker closes it, sends it to a buffer position, and opens a new store pallet on the spot. Besides, pickers also perform retrieval tasks and move the store pallets (in buffer positions) onto the marshalling area. In fact, those tasks can be done simultaneously with the picking tasks, but by a different picker, since each picker can only perform one task at a time.

In PBS, the prepared pallets are stored instead of being dragged along the picking area. Therefore, locations represent SKUs, instead of stores, since each pallet will mostly have one SKU. Thus, picking tasks are the opposite of those at PBL as they start with an empty pallet and finish with the pallet complete. Pickers now drag the store pallet alongside the picking area in order to collect the ordered products from each store. When a store pallet reaches its maximum, it is moved onto the marshalling area. Figure 1.2 illustrates the main differences between PBL and PBS.

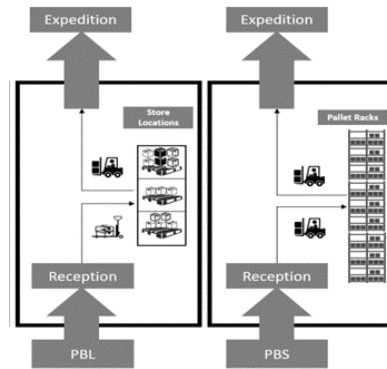


Figure 1.2: Product flow type in PBL vs PBS

Finally, the marshalling zone has staging lanes where store pallets are waiting to be shipped. Similarly to the inbound area, trucks arrive according to a planned schedule. They are then loaded with store pallets and dispatched onto the respective stores. In fact, the reception and expedition often occur in different time windows. Consequently, many warehouses opt to use the inbound and marshalling zones as the same one. However, in this project, we consider them separately.

1.1.2 Problem Statement

The focal question lies in the dynamic arrangement of the picking and marshalling zones. Considering the PBL flow type, a picking task has stores to visit. Therefore, the picking area can be viewed as a series of locations that can be rearranged. This flow has the particularity of having products received and dispatched on the same day. Thus, it is preferable to have the stores with a high number of orders, for that day, closer to the receiving zone, in order to minimize the distance travelled.

Since demand has trends and seasonality, stores can easily vary the number of orders throughout time. So, their positions inside the picking area should vary too. A similar reasoning could be found for PBS, but instead of stores, one shall optimize the SKU assignment rules.

In addition, store pallets are moved onto the marshalling area, if their limit has been reached or the store order has no more items to be picked. They are then allocated to their position on a staging lane. The same rationale can be applied here: If a store has many orders for the following period, the respective store pallets should be buffered closer to the truck they will be loaded into, instead of having a fixed queue location.

1.1.3 Mission and Objectives

As mentioned before, the main goal of this project is to develop a product that optimizes the layout of any given warehouse. For that matter, the problem at hand can be viewed, in a nutshell, as a dynamic allocation problem to warehouse locations. To be more precise, this product intends to improve picking and marshalling productivity since these operations pose the most arduous challenges to overcome.

By optimizing what should be assigned to each location, for each period, it is possible to minimize the total travelled distances and, consequently, the time spent performing those activities.

However, each warehouse is unique and solutions derived from one layout might only be optimal for that particular case. Hence, this product can be viewed as an application that requires a layout, as an input, to then output the best allocation rules for each user-defined configuration. This enables users to quickly test multiple layouts and analyze key performance metrics to compare between different scenarios.

Moreover, there are several interrelated decision variables. So, every choice that is made, at one level, narrows down the solution space possibilities from that point onwards. For example, in PBL, the way picking tasks are done will depend on the way picking pallets are prepared. However, the way picking pallets are prepared must also consider the orders to fulfil.

This adds an additional layer of complexity to the problem at hand, thus requiring a framework that is able to: (a) decompose a given warehouse on its building blocks, (b) optimize each block's internal layout independently, and (c) consider the whole warehouse operation and the impact of each optimized block in the overall solution.

1.1.4 Proposed Approach

In order to tackle those questions, a hybrid approach combining simulation and optimization was developed. The hierarchical structure of the models used is sequential: Simulation-Optimization-Simulation, hence the name SOS. The first simulation has the purpose of being an analytical model enhancer since its results are used to improve the optimization model parameters. The second simulation is used to complement and generate the solution provided by the optimizer. According to Figueira and Almada-Lobo (2014), approaches that combine both simulation and optimization are used to better answer to uncertainty.

Warehouses are then split into functional blocks according to the nature of their operations. As previously discussed, every warehouse comprises three major zones – inbound, storage/picking, and marshalling. Besides, picking can also be decomposed into PBL or PBS whether storage hardly plays a role or not. After defining these boundaries, the SOS framework can be used:

1. Having the processes mapped in-depth, a first simulation model can be developed to deploy scenario testing within each block and without interactions between different blocks. The main goal of this primary phase is to enhance an optimization model by feeding it with the most promising inputs.
2. Afterwards, the optimization model is used to further refine the chosen scenarios. For the PBL block, the goal is to dynamically change the storage assignment policy to minimize the total picking travel time. This need for scenarios is due to the fact that different factor combinations lead to different results. In fact, the optimal storage assignment policy heavily depends on the picking area design variables: internal corridor disposal, routing constraints, whether there is zoning or not, etc. So, this step improves the best combinations to choose the best among the best.

- However, any combination of optimized blocks can be sub-optimal since interactions between blocks were not considered prior to this point. Hence, a validation phase is followed to test the robustness of the generated solutions. This can be made with a second simulation model. By submitting the blocks to stress conditions, it is possible to uncover bottlenecks. This is done to substitute any underperforming blocks with the second-best ones attained by the first simulation model, thus being an iterative approach.

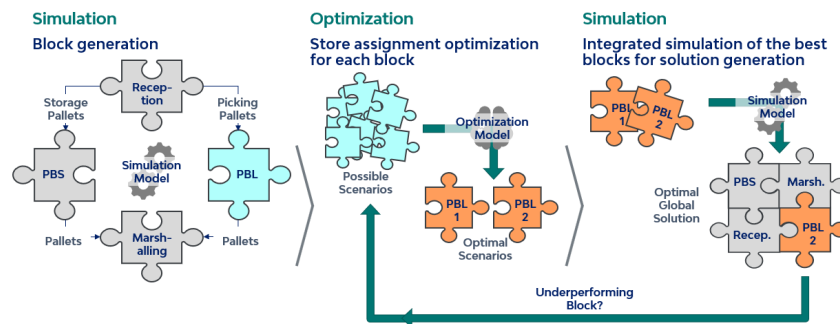


Figure 1.3: An overall perspective of the SOS methodology and its building blocks

1.2 From Project to Thesis

Figure 1.3 presents a schematic perspective of the SOS methodology. Only the PBL block, within the first 'S' is highlighted, since it is the focal point of this dissertation. This is due to the shortened time window of this work when compared to the overall project duration.

1.2.1 Timeline and Scope

Given the absence of such a product in the market, and due to its complexity, this project is a joint effort of several companies. Besides, it is expected to last three more years after this thesis completion. Because of that, the present dissertation has its focal point "only" towards the modelling and performance estimation of picking areas under a PBL flow type. In Figure 1.4, the several steps are described alongside their respective milestones.

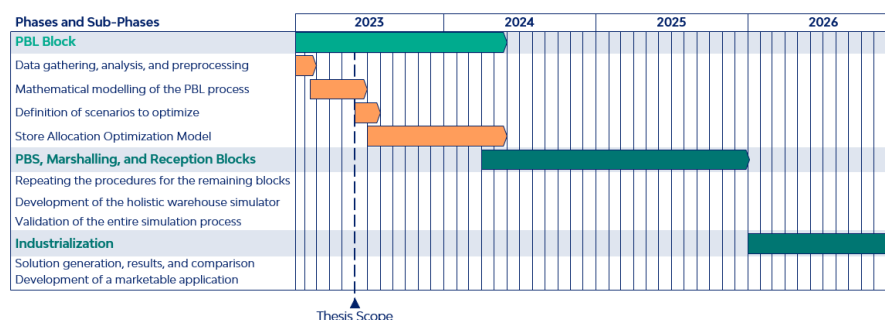


Figure 1.4: Project's timeline

1.2.2 Proposed Approach

Since the pallets have standardized dimensions, the ground floor locations assigned to the stores should also be normalized. So, in theory, The PBL block can be conceptualized as a grid, where each cell represents a possible location that can be designated for a store.

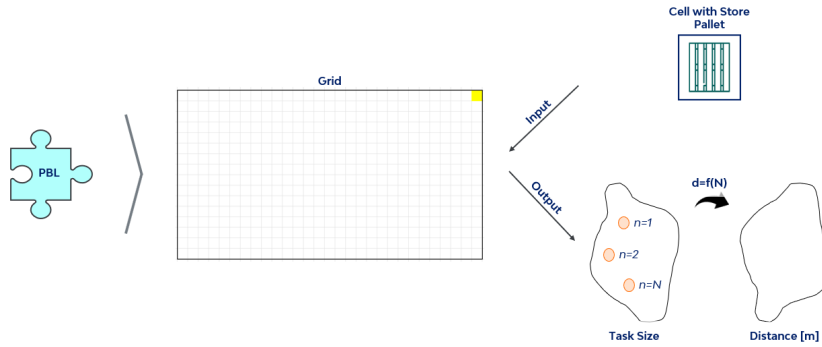


Figure 1.5: Schema for estimating the performance of any user-defined layout

Bearing this in mind, the following key questions must be addressed to challenge the current *status quo* of a given warehouse's picking area: (1) How does the current layout perform and what KPIs must be used to evaluate this? (2) What scenarios shall be considered to improve picking performance? (3) For a given layout, what combinations of factors pose the most promising scenarios? Section 2.2.3 explores a myriad of different factors that affect order-picking performance: layout design, storage assignment policies, routing methods, zoning, etc. The goal is to, given any picking area, estimate its performance based on the distance needed to perform a picking task.

1.3 Thesis Structure

This dissertation is structured into six different chapters organized as follows: Chapter 2 is provided with: (i) a theoretical background on why warehouses are needed, as well as their functions in modern-day retail-market supply chains; (ii) a revision on the current state-of-the-art literature, on order-picking, to better understand the decisions faced when designing and controlling this operation, and (iii) the techniques that exist to diagnose and improve order-picking performance in order to identify the gaps relative to the problem under study.

Chapter 3 extends on the problem scope and on the grid methodology used to tackle it. In order to estimate the expected performance of a given picking area, two models were derived with an increasing degree of detail. A comprehensive analysis and explanation of these models are conducted in Chapters 4 and 5, respectively. Nevertheless, they are both based on the expected travel distance that is going to be needed to perform a picking task.

Since this thesis aims to develop a general model, there is a need to validate the used methodology in a particular case study. Hence, Chapter 6 will be used to: (i) present the study subject, and (ii) analyse the results obtained. To finalize, Chapter 7 withdraws the main conclusions and points out some suggestions and improvements for future work.

Chapter 2

Literature Review

Warehouses are, and always have been, central points of the retail market supply chains. Therefore, this chapter will first present a brief overview of why warehouses are needed and what are their main functions in modern-day supply chains. Afterwards, state-of-the-art literature on warehouse science and order picking is reviewed.

Most research only focuses on the design phase of the warehouse while the operational side is often disregarded. So, the final part of this chapter is dedicated to the research more intricately related to the work of this thesis.

2.1 Theoretical Background

2.1.1 Retail Market

The retail market is, in a nutshell, the sale of goods (or services) to the last elements of the supply chain. Usually, the retailer acquires large volumes from the suppliers and sells smaller quantities to the end consumers to maximize profit. The modern retail environment is now global, dynamic and ever-growing but changed a lot in the past few years.

The year 2020 might have been a year of online shopping, but as the restrictions faded away, people eagerly returned to stores. In fact, brick-and-mortar stores will always have a place in the retail landscape and the reason is simple – people enjoy shopping and spending time with each other. It is an experience that cannot be replicated on a screen, despite some attempts have been made to emulate that social element, such as live stream selling or squad shopping.

However, the following year was deeply marked by uncertainty. Successive openings and closures forged a new generation of consumers that crave for both online and offline purchases. Therefore, brands succeeded due to their digital transformation. By making use of modern technology, retailers were able to deliver personalized customer experiences across all channels. Thus, speed and flexibility are two key elements to business survival and COVID has only reinforced their importance. Until now, it was all about how fast you can provide information, process transactions, and deliver products. However, companies learned, the hard way, it is also about how well you can adapt to new trends and pivot operations to ensure customer (and employee) satisfaction.

Nowadays, in a world disrupted by consecutive lockdowns and re-openings, commerce is still evolving – and fast. In fact, the modern-day retailer must cope with a multitude of channels in order to thrive. History has proven that those who relied on technology were able to quickly adapt and evolve their businesses once stores were forced to a closure. Hence, new ways of shopping have emerged because of the pandemic, making it crucial to mix physical and digital touchpoints when reimagining retail in a new normal world.

Customer expectations have thoroughly changed since the first COVID-19 outbreak, but one thing remained certain: retailers need agile supply chains to be ready for whatever the future may reserve.

2.1.2 Supply Chain

According to Christopher (2017), one of the most significant paradigm shifts of the millennium is that companies no longer compete by themselves. Instead of “brand versus brand”, it is now “supply chain versus supply chain”. Therefore, the success of retailers relies on their abilities to integrate and manage their network of business relationships. Figure 2.1 provides a simplified version of a company’s network structure.

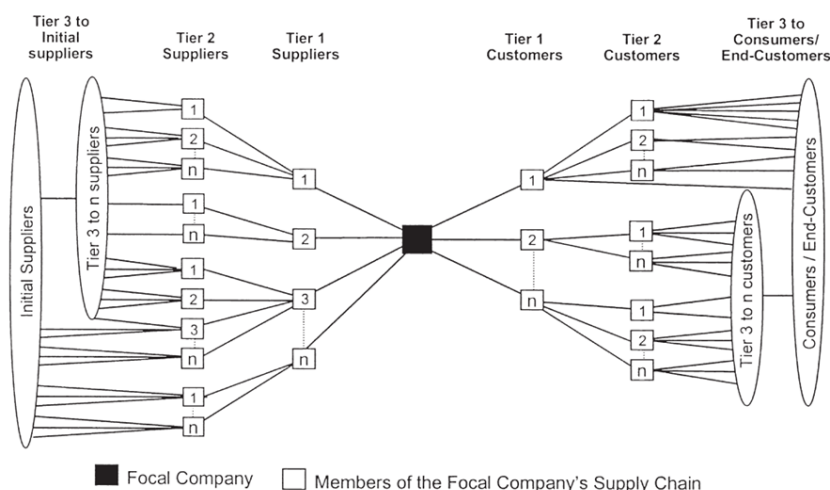


Figure 2.1: Supply chain network structure

The same supply chain might look different from different perspectives – managers see their firms as the focal company, hence viewing different memberships and network structures. In fact, a supplier of a given company can be a customer of another company. Besides, it would be rare for a firm to participate in only one supply chain, making it harder to identify its role.

Countless activities are performed and coordinated within a company. Furthermore, every company is inherently involved with other companies; for example, retailers affect the operations of consumers and vice-versa (Stigler, 1951). In fact, when two companies establish a relationship, certain activities will be linked and managed between both entities (Coase, 1937). Nevertheless, some links are more critical than others (Snehota and Håkansson, 1995) and it is up to management to choose the ones that best fit the circumstances (Lambert et al., 1996).

Lambert and Cooper (2000) proposed a framework to tackle these pain points that stem from supply chain management. They devised a methodology to support tactical decisions based on three closely interrelated elements: network structure, business processes, and management components. The first one aims to uncover key members in the supply chain to link up processes with. The second one identifies which processes should be aligned with each of those key members. The third one measures the strength and level of integration to deploy at each process link.

It is unfeasible to monitor and manage everything – from all the suppliers, back to the point of origin, to all the products and services, out to the point of consumption – given the degree of complexity and resources that would be needed. For that reason, it is crucial to differentiate members similarly to how Porter (1985) distinguishes activities on his “value chain” framework.

As a result, primary members carry out value-adding activities whilst supporting members provide the means to do them. Moreover, processes are a structured set of activities (Davenport, 1993), and according to Snehota and Håkansson (1995), the structure of activities within and between companies is a critical cornerstone of internetwork competitiveness and, consequently, supply chain performance.

Thus, ‘supply chain management’ quickly became a buzzword in the business world. By definition, it is the integration of key business processes, from point of origin to point of consumption, in order to provide products and services that add value for customers and other stakeholders. Hence, operating a holistic supply chain requires continuous information flows that enable timely responses to uncertainty in demand. Figure 2.2 depicts this cross-functional integration across the silos within a company, and across the companies within the chain.

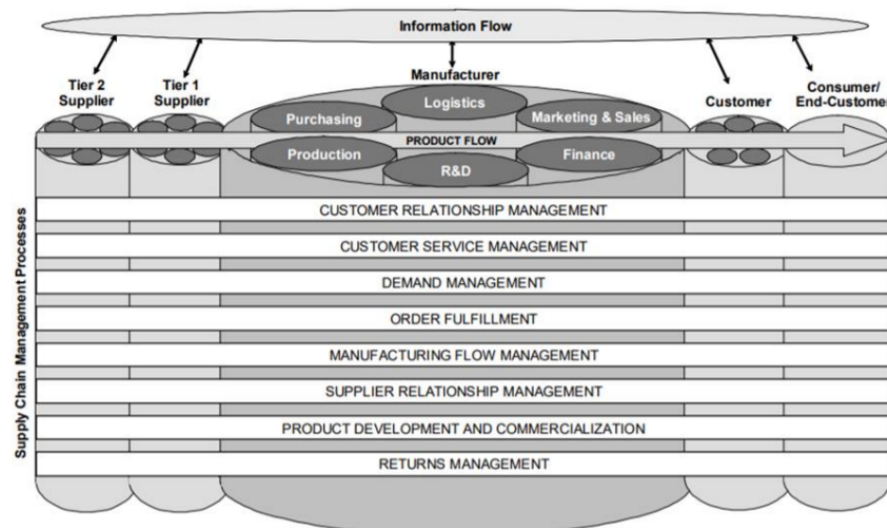


Figure 2.2: Integrating and managing business processes across the Supply Chain

Warehouses and distribution centres are vital points of any supply chain as they often make the bridge between suppliers, manufacturers, and customers. Hence, an efficient supply chain must be based on efficient warehouse operations.

2.2 Warehouse Science

2.2.1 The Purpose of Inventory and Warehouses

Inventory holding and warehousing continue to play an important role in modern-day supply chains since product availability is a key metric in terms of customer service (Frazelle, 2002). However, there are widely different points of view regarding this topic. Factors such as speed to market for new products or responsiveness to new technologies are listed as major drivers for the obsolescence of existing stocks. Therefore, inventory can either be viewed as a source of risk or a means for risk mitigation.

Literature has been mostly focused on the disadvantages of holding inventory by devising strategies for inventory reduction, such as the centralization of inventory in a single decoupling point rather than having several buffers. In fact, van Hoek (1998) suggests the use of production postponement to store similar components instead of having a myriad of finished goods. In addition, Christopher (2011) defends the visibility of consumers' demand across all supply chain members to reduce the "bullwhip effect". Figure 2.3 illustrates a retail supply chain; the triangular shapes represent the components in which inventory is held.



Figure 2.3: Inventory is held at points along the chain as well as in transit

On the other end, there has been little research concerning the need for inventory. Besides, stocks can only be reduced to a certain extent in order to avoid "corporate anorexia" (Radnor and Boaden, 2004). Thus, the question is not whether inventory is "good" or "bad" but how to determine its optimal level. Baker (2007) proposes a framework to better understand the role of inventory, and warehousing, by unifying these two separate but interrelated areas of investigation – inventory management and inventory reduction.

First of all, companies must address whether inventory is required at all. The need for inventory depends on the business model. Some firms have to quickly identify new customer trends and push the product directly into the market, hence keeping stock to a bare minimum. This is especially true in fashion and technology retail where the product life cycle tends to be short and supply chains need to be agile. Christopher et al. (2006) classify the goods that a retailer can sell in two different taxonomies regarding products' life cycle – hard goods and soft goods – with the latter being associated with those fast-moving ones.

Therefore, in agile supply chains, companies make-to-order and inventory is held only at some echelons – if held at all (van Hoek et al., 2001). This is in accordance with the lean school of thought which paints inventory as one of the seven "wastes", thus being considered something to

be reduced as much as possible (Womack and Jones, 1996). Nonetheless, stocking at strategic points can act as a buffer, separating lean production activities from an agile response to volatile marketplaces (Christopher and Towill, 2001).

The need for inventory buffers is demonstrated by the wide difference between the supplier lead time and the customer lead time, known as the lead-time gap (van Hoek et al., 2001). Moreover, globalization created wider supply chains which led to longer supply lead times and the need for more inventory to provide the same service levels (Waters, 2003).

That is why, inventory reduction strategies must be measured against supply chain risks. In Oke and Gopalakrishnan (2009), the different potential risks in a retail supply chain and their main mitigation strategies are addressed. Besides, Lowson (2002) demonstrates that bulk purchases offer higher discounts, larger batch sizes imply lower production costs, and full container loads allow economies of scope. Hence, companies must understand these supply chain trade-offs to define optimal inventory policies and provide the desired service level to customers.

Finally, this discussion about the purpose of inventory can naturally be extended to the purpose of warehouses. In fact, inventory holding and order fulfilment from that inventory are key warehouse functions (Rushton et al., 2014). The term ‘warehouse’ is used if the main function is storage or buffering. If, additionally, transportation is core, the term ‘distribution centre’ is preferred whereas ‘transshipment’ or ‘cross-dock’ centre are often used if storage hardly plays a role.

According to Lambert et al. (1998), warehouses contribute to several of the company’s critical missions such as:

- Overcoming the time and space differences that may exist between suppliers and customers;
- Achieving production economies (e.g., make-to-stock);
- Taking advantage of bulk purchase discounts and forward buys;
- Achieving transportation economies (e.g., combined shipment and full-container load);
- Providing customers with a mix of products on each order (e.g., consolidation);
- Providing temporary storage of material to be disposed or recycled (e.g., reverse logistics);
- Minimizing logistics costs while maintaining a desired level of customer service.

2.2.2 Warehouse Operations

According to Tompkins et al. (2010), warehouses are commonly split into functional areas. Each area is held accountable for a set of operations. In their work, de Koster et al. (2007) describe the four main operations that occur inside a warehouse: receiving, put-away, order-picking, and shipping.

Everything that comes in will, eventually, come out. Hence, receiving is the first activity to take place. The ordered products are delivered to the reception area at a predefined schedule and

workers are assigned to unload those items from the transport carrier. Then, inspections are made to ensure order quality and completeness before the update of inventory information systems.

After that, the newly arrived products need to be stored. Thus, put-away is in charge of transferring the incoming products from the reception area to the storage area. The flow of products between functional areas is usually made with pallets to minimize handling costs.

Having finalized their inbound process, products will wait to be picked. Order-picking is, by definition, the retrieval of stored products in response to specific customer requests. Besides, it is the most labour-intensive and costly activity for almost every warehouse, thus being the highest-priority area for improvements (Rushton et al., 2014). Chen et al. (2015) estimate that order picking accounts for 60–70% of warehousing costs.

Finally, orders often have to be packed and stacked on the right unit load (like a case or a pallet) in order to be shipped. Shipping is the inverse process of receiving. Therefore, transport carriers are only available at specific points in time given the tight customer delivery windows. Hence, orders require a dedicated area for buffering prior to shipment – the marshalling area.

2.2.3 Design and Control of Order-Picking Processes

Companies make decisions on the design and control of order-picking systems at tactical and operational levels respectively (Rouwenhorst et al., 2000). Decisions made at both levels are strongly interdependent. However, including all the decisions, with different time horizons, in only one model is certainly impossible.

Figure 2.4 shows the level of complexity of order-picking systems according to Goetschalckx and Ashayeri (1989) - the farther a system is located from the origin, the harder the system is to design and/or control.

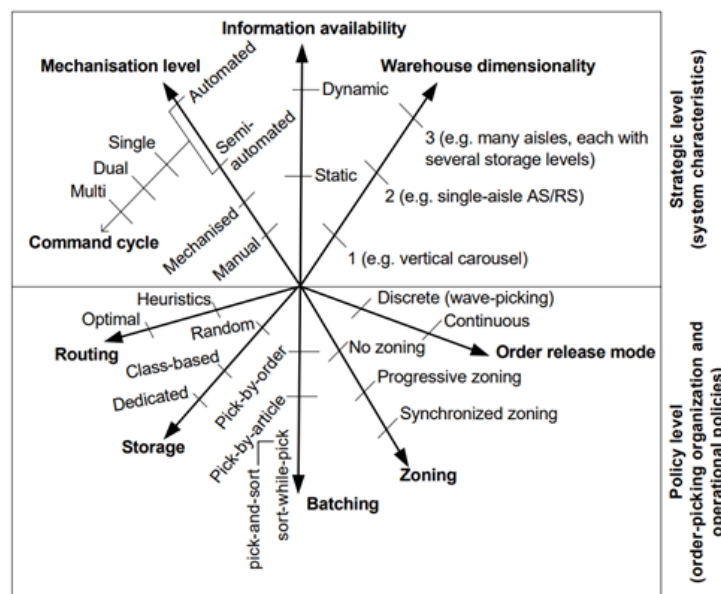


Figure 2.4: Complexity of order-picking systems (design and control)

When all the degrees of freedom for arranging a warehouse are added up, the problem becomes increasingly complex. This is especially true given the onset of e-commerce since more SKUs need to be stored, in fewer quantities, to tackle volatile demand patterns and primarily single-line customer orders (de Koster et al., 2017). Therefore, it is even more relevant to address warehouse problems in an integrated way (van Gils et al., 2018).

Few authors have addressed combinations of decision problems. Hence, there is a need to overcome this gap in order to understand new interaction effects on the order-picking objectives. The most common objective is to maximize the service level. Besides, the faster an order can be retrieved, the sooner it is available for shipping. Therefore, minimizing the picking time is of utmost importance for any order-picking system. Figure 2.5 dissects the order-picking time components in a typical picker-to-parts warehouse according to Tompkins et al. (2010).

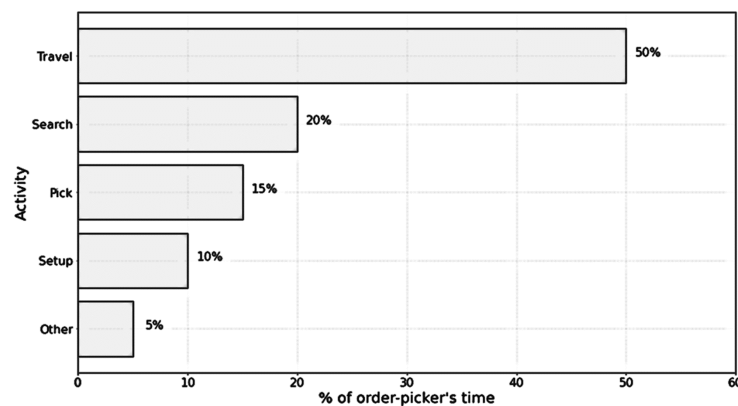


Figure 2.5: Typical distribution of an order picker's time

Travel is often the dominant activity. According to Bartholdi and Hackman (2019), travelling is a waste of time since it costs labour hours and does not add any value. Therefore, it is a natural first candidate for improvement. Moreover, travel time is an increasing function of travel distance. Thus, minimizing the travel distance is also considered a key objective in order-picking optimization (Grosse et al., 2017).

In order to improve picking efficiency, warehouse managers typically face four tactical and operational decision-related problems: (a) layout design, (b) picking policies, (c) storage assignment rules, and (d) routing methods (Chan and Chan, 2011). On top of that, authors like de Koster et al. (2007) also consider zoning as a variable.

2.2.3.1 Layout Design

The layout design concerns two sub-problems: the layout of the facility containing the order-picking system and the layout within the order-picking system. The first one is usually known as the facility layout problem. It defines the boundaries and allocation of the different functional areas on the warehouse floor. In their work, Meiler and Gau (1996) provide an extensive overview of this subject. Nonetheless, the main objective is to minimize the handling cost, which is often represented by a linear function of the travelled distance between departments.

The second one focuses on the design of the picking area. Moreover, the picking area can be deconstructed as a set of blocks such as the ones depicted in Figure 2.6. Consequently, the internal layout design concerns the determination of the number of blocks and the number, length, width, and orientation of the aisles inside each block. Again, the most common objective function is to minimize the travel distance (Sprock et al., 2017).

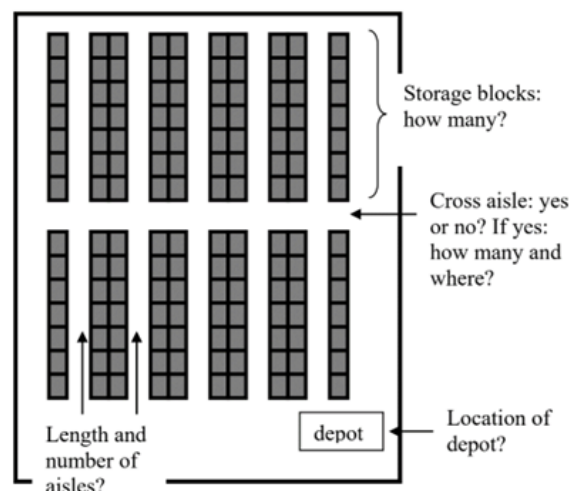


Figure 2.6: Typical internal layout decisions in order-picking design

Much of the existing knowledge is used to optimize warehouse layouts for various storage and routing strategies. However, changes to the facility layout are inside the scope of this thesis. Nevertheless, the following sections briefly review the literature concerned with those topics.

2.2.3.2 Order-Picking Policies

Products flow through the operations described in section 2.2.2. However, there are some nuances concerning the order-picking process. Figure 1.2 illustrates the two main order-picking policies found in warehouses – Picking by Line (PBL) and Picking by Stock (PBS).

Picking by Stock concerns the traditional warehousing methodology. As the name implies, products are stored to fulfil future demand. Each product is then assigned to a storage location, usually a pallet in a pallet rack, giving birth to the term Stock Keeping Unit (SKU).

An SKU is the smallest unit that is tracked within a warehouse – and across the supply chain. SKUs work like barcodes as they identify product types rather than products themselves. Thus, each product is an SKU, but identical products are, to the warehouse, a certain amount of the same SKU. Besides, each pallet in a pallet rack usually stores products of the same type or, in other words, usually stores the same SKU.

Hence, a picking task in a PBS flow type can be seen as a supermarket shopping list – a picker will go through the aisles to collect a list of SKUs, and their respective quantities, in order to fulfil a customer request.

In contrast, the PBL flow type adopts a no-stock policy. Products from suppliers are received and shipped on the same day. Instead of storing them, products are directly distributed through customers' orders. Hence, order picking is done after the reception process is completed. In PBL, each customer order is assigned to a location, whereas in PBS, products are the ones to have locations assigned. Thus, pickers will go through the aisles to drop off (viewed as a pick) the newly received products in customers' locations. These are usually pallets on the shop floor.

Table 2.1 depicts two picking task examples for the PBL and PBS flows. The focus of this work only comprises the Picking by Line flow type.

Table 2.1: Picking task examples for the PBL and PBS flows

PBL Task ID	SKU	Qty.	Store ID	PBS Task ID	SKU	Qty.	Store ID
1	2	2	10	1	2	5	20
1	2	4	20	1	6	8	20
1	2	3	50	1	4	3	20
2	1	5	20	2	5	2	50
2	1	3	10	2	1	3	50

Apart from that, there are multiple order-picking methods regarding whether or not humans are employed to perform the order-picking tasks (de Koster, 2012). The most common one is the picker-to-parts system in which the picker walks or drives along the aisles to pick items. Surprisingly, these systems have received less research attention when compared to others. The reasons for this may have something to do with the complexity and diversity of picker-to-parts order-picking systems. Nonetheless, the main goal of this study is to develop a product that optimizes order picking in every warehouse with such a system.

Within this picking model, the decision over how to split and combine orders also matters, as it may severely affect the picking performance. For large orders with diverse products, each order can be handled individually, with each task giving origin to one picking tour. Alternatively, when orders are small, travel times can be reduced by picking a set of orders in a single picking tour.

There are two criteria for batching: the proximity of pick locations and customers' time windows. Proximity batching assigns each order to a batch considering the distances of its customer's locations to those of other orders. The major pain point is how to measure the proximity among orders, which implicitly assumes a route to visit a set of customer locations. Under time window batching, the orders are grouped as a batch based on the time interval that they can be shipped to the customers.

2.2.3.3 Storage Assignment Policies

Assigning products to storage locations has been thoroughly discussed over the past years. In their work, de Koster et al. (2007) compile the main storage assignment policies under a PBS flow type. Even so, those policies are extensible to the PBL flow type without loss of generalization. As mentioned before, customers are the ones that have locations in a picking by line area. Hence, the scope of this work concerns assigning customers – not products – to locations.

For random storage, every new customer is randomly assigned, with equal probability, to an empty location in the picking area (Petersen, 1997). In the case of a retailer's warehouse, customers correspond to the stores being served. Also, one store can have multiple locations whether they place many orders or not.

This strategy results in higher space utilization at the expense of increased travel distances. Otherwise, stores may have a fixed location defined *a priori*. A disadvantage of dedicated storage is that a position is reserved even if the respective store does not pose any orders. However, order-pickers may become more familiar with the picking layout.

Another possibility is to assign stores based on the first empty location that is encountered. This typically leads to a disproportionate picking area with an uneven concentration of stores. As a result, aisles are full around the depot and gradually more empty towards the back. Depots can be seen as the starting point for picking tasks. According to Hausman et al. (1976), closest open location and random assignment policies have similar performances if products are moved by full pallets only. That is the case with PBL.

Finally, customers can also be allocated according to their turnover. By doing so, stores with the highest order volumes are placed at the easiest accessible locations. For that, Heskett (1963) developed the cube-per-order (COI) index. By definition, the COI of an item is the ratio of the item's total required space to the number of trips needed to satisfy its demand per period. Replacing "item" with "customer's orders", it is possible to devise an algorithm to place the stores with the lowest COI closer to the depot.

There is research that proves that the COI policy is optimal in minimizing the material handling cost. The first work on this matter was published by Kallina and Lynn (1976). The main disadvantage of a full turnover policy is that demand rates are uncertain and ever-changing. Each change would require a new store assortment since the orders to fulfil would change as well, resulting in a large reshuffling every other period. This project aims to optimize this trade-off in order to successfully apply dynamic storage policies.

Gu et al. (2007) formulates the storage assignment problem based on product information. Every new product is assigned to a product class based on its characteristics. Naturally, this can also be applied to new customers. Nevertheless, classes are devised by some kind of demand frequency metric.

Originally, Pareto's method was adapted from Economy to subject items into groups. It states that 80% of the wealth in the world is held accountable by only 20% of the people. Hence, 80% of the warehouse turnover relies on 20% of the customers. This fast-moving class is usually called A-class. The number of classes is often restricted to three – the ABC method. However, in some cases, more classes can provide additional gains with respect to travel times (Petersen et al., 2004). Nevertheless, the pick volume is often preferred since high-end products may bias turnover.

Each class is then assigned to a dedicated area of the picking zone. There are two major contributions on this subject. Jarvis and McDowell (1991) suggest that each aisle should contain only one class (within-aisle storage) whereas Petersen and Aase (2004) claim the opposite (across-aisle storage). Figure 2.7 compares both policies for a single-block layout.

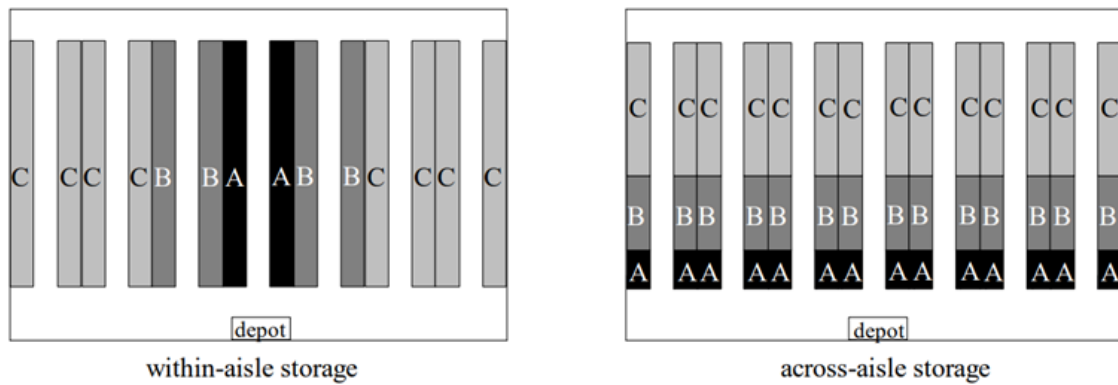


Figure 2.7: Class-based storage assignment policies

Apart from that, storage assignment and routing policies are deeply interconnected. However, this effect seems to be largely neglected in the literature. Le-Duc and de Koster (2005) optimize these storage-class positionings for some routing methods. In their findings, the across-aisle storage is depicted as close to optimal.

The concept of class-based storage may also combine some of the methods mentioned so far. In fact, items inside each class can have their own storage assignment policies. The work of Caron et al. (2000) provides a comparison of the average travel distance of an order-picking task when using a random storage policy and a COI-based storage policy inside each class. The results show that the second one is preferred as the number of picks per task increases.

2.2.3.4 Zoning

Customers may tend to order similar products for the same time periods. In this case, it makes sense to assign their positions relatively close to each other in the picking area. This is called family grouping. To do so, the statistical correlation between customers' orders should be known or, at least, be predictable (Wäscher, 2004). This can be useful to split the picking area into flow types. Just like a supermarket, aisles are divided according to the products' nature. As a result, customers will have more picking locations for the same order – at least one per zone.

de Koster et al. (2007) provide a brief insight on zoning disclaiming that the literature around this topic is very limited. Zoning, as the name suggests, is the partition of the order-picking area into zones. The main leverage of zoning is that each picker is assigned to one zone only, hence traversing a smaller area. Yet, the order needs to be split and consolidated again before shipment to the customer.

There are two approaches to complete an order when using zoning. The first one is called progressive assembly. In this approach, one picker starts the order inside his zone and when he finishes his part a different picker continues to retrieve the rest of the items. For that, this system is also known as pick-and-pass. The second one is called parallel picking. Here, several pickers start working on one order, inside their respective zone, simultaneously or not. Once all the parts of that order are completed, they are merged.

2.2.3.5 Routing methods

Routing order pickers in a warehouse is a special case of the Travelling Salesman Problem, originally thought by Lawler et al. (1985).

As mentioned before, the main goal of order-picking is to minimize the picking time. For that, the travelled distance must be as small as possible. As this is an NP-hard problem, many studies focus on developing heuristic methods for solving it. According to Petersen (1997), the best heuristic solution is, on average, 5% worse when compared to an optimum.

Roodbergen and Koster (2001) distinguish several heuristic methods for routing order pickers in single-block warehouses. Figure 2.8 illustrates some of these methods.

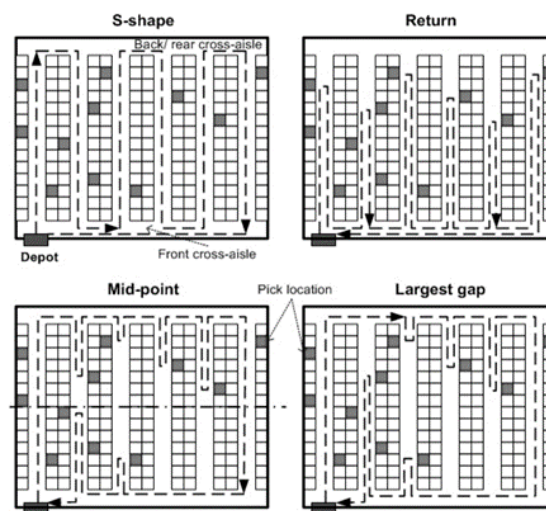


Figure 2.8: Routing methods for a single-block picking area

In the S-shape heuristic, any aisle containing at least one pick is traversed entirely whereas aisles without picks are not entered at all. The return method is similar; however, the picker enters and leaves each visited aisle by the same end. Another possibility is the midpoint method: picks in the front half are accessed from the front cross aisle whilst picks in the back half are accessed from the back cross aisle.

The largest gap strategy extends the midpoint concept. In fact, the gap within an aisle represents either one of these three options: the distance between two adjacent picks, between the current end and the first pick, or between the current end and the last pick. If the largest gap is between two adjacent picks, the order picker will perform a return route from both ends of the aisle. Otherwise, a return route from either the front or the back aisle is used. The largest gap always outperforms the midpoint method (Hall, 1993).

According to Roodbergen and Koster (2001) all these methods can be applied to multiple-block warehouses with some modifications. Even so, problems with routing may arise if, for the same SKU, products are stored throughout multiple locations.

This is especially true for customers in a PBL flow type. In fact, large stores place many orders, thus requiring more than one position in the picking area. If so, a choice has to be made

from which location the pickers should visit (Daniels et al., 1998). Also, congestion may pose a drawback since too many pickers travelling through the same aisle may hamper order retrieval times (Bahrami et al., 2017).

2.3 Wrap-Up

The objective of this section is to provide a concise summary of the fundamental findings that form the basis of this work. To accomplish this, a comprehensive overview of the most significant insights is presented in Table 2.2.

Table 2.2: A summary of the related research

Author(s)	Field	Findings
Tompkins et al. (2010) Rushton et al. (2014) Chen et al. (2015)	Warehouse Operations	Warehouses can be split into functional areas according to their operations Order-Picking is the most labour-intensive and costly operation. Order-Picking is the highest priority area for improving warehouse design
Goetschalckx and Ashayeri (1989) Rouwenhorst et al. (2000) de Koster et al. (2007), Chan and Chan (2011) van Gils et al. (2018)	Order-Picking Design	Complexity levels when designing and controlling order-picking systems Decisions made at tactical and operational levels are interdependent The main decision problems are: (a) internal layout design, (b) picking policy, (c) storage assignment rules, (d) zoning, and (e) routing methods Few authors have considered combinations of decision problems
Tompkins et al. (2010) Grosse et al. (2017) Alves (2015), Amorim-Lopes et al. (2020)	Picking Performance and Optimization	Minimizing the picking time is crucial for any order-picking system Minimizing the travelled distance is also considered a key objective Layout performance can be estimated based on the travelled distance by using a probabilistic simulation model
de Koster et al. (2007) de Koster (2012) Sprock et al. (2017)	Internal Layout Design	Every warehouse comprise three major zones – inbound, storage/picking, and outbound There are multiple order-picking methods. The most common one is the picker-to-parts system The picking area can be deconstructed as a set of blocks
Petersen (1997) Hausman et al. (1976) Heskett (1963), Kallina and Lynn (1976) Gu et al. (2007) Jarvis and McDowell (1991), Petersen et al. (2004) Le-Duc and de Koster (2005) Caron et al. (2000)	Store Allocation Rules	Comparison between random and dedicated strategies Closest open location and random assignment are similar for PBL COI policy allocates items according to their turnover. Proves that this policy is optimal when minimizing the material handling cost Class-based product assignment Each aisle can either contain one class (within-aisle storage) or more (across-aisle storage) The across-aisle storage is depicted as close to optimal Estimate layout performance for various storage allocation rules. The best result obtained was by using a COI-based policy inside each class
Wäscher (2004) de Koster et al. (2007)	Zoning	The statistical correlation between customer's orders should be known Orders are split and consolidated before shipment either in a progressive or parallel way
Petersen (1997) Roodbergen and Koster (2001) Bahrami et al. (2017)	Routing Methods	Best heuristic is 5% worse when compared to an optimum Distinguish several heuristic methods for routing order-pickers Provides solutions while dealing with congestion issues

Each warehouse can be divided into functional areas based on its operations. This helps explain the methodology of the project in which this thesis is inserted. In fact, the first 'S' intends to simulate each block independently in order to generate different scenarios for each area.

Moreover, order-picking is identified as the most labour-intensive and costly operation. So, due to the limited time frame of this work, the focus will be on how to generate multiple scenarios for the PBL block. Hence, the primary objective of this study is to develop a means to estimate the performance of any given picking area.

This is done based on the distance that a picker needs to travel for increasing task sizes. However, modelling a picking area involves numerous - and interrelated - variables. These include the internal layout design, the picking policy in use, the storage assignment rules, the existence of zoning or not, and the routing methods. Thus, each case is a case.

Nevertheless, by leveraging a graph theory approach, all picking areas can be well-modelled under a grid. Hence, it is possible to test any solution composed of any combination of the above-mentioned factors. This study specifically focuses on the PBL picking policy but can easily be extended to comprise the PBS flow type. Additionally, it only considers the picker-to-parts system, which is, by far, the most commonly used approach.

To estimate the distances, two probabilistic simulation models were derived. The first one, more simple, deconstructs the picking area as a set of blocks, as in the research of Caron et al. (2000), Alves (2015), and Amorim-Lopes et al. (2020). The second one, more detailed, is adaptable and applicable to any possible picking area without having to make such simplifications.

Chapter 3

Methodology

Layout optimization is far from having a general and straightforward approach. When considering all the degrees of freedom, in warehouse design, the problem becomes increasingly complex if not impossible to solve. There are several interrelated decision variables at both tactical and operational levels, hence every choice that is made narrows down the solution space possibilities. Because of that, the existing research focuses towards optimizing one decision problem *ceterus paribus*, that is, maintaining all the remaining ones constant.

In fact, most of the knowledge is used to optimize a fixed layout for various storage strategies. Besides, the best storage strategy is the one that minimizes the total travelled distance, which logically depends on the best routing method that can be deployed. Moreover, the routing constraints are imposed by the internal layout design. So, what happens if the internal layout, that was initially fixed, changes? This scenario-based rationale is the basis behind the derived methodology. Furthermore, those optimizations are mainly applied to order-picking systems only. However, the same line of thought can easily be transposed to the reception and marshalling areas. By doing that, how will the optimized areas perform altogether?

This adds an additional layer of complexity to the problem at hand, thus requiring a framework that is able to: (a) decompose a given warehouse on its building blocks, (b) optimize each block's internal layout independently, and (c) consider the whole warehouse operation and the impact of each optimized block in the overall solution. This chapter aims to explain this framework and discuss the scope of this thesis within that framework. Specifically, this study will primarily focus on how to model any given PBL block to estimate its performance. To achieve this, a methodology to represent any picking area, using a grid, is presented afterwards.

3.1 The 'SOS' Framework

In order to tackle those questions, a hybrid approach combining simulation and optimization was developed. The hierarchical structure of the models used is sequential: Simulation-Optimization-Simulation, hence the name SOS. The first simulation has the purpose of being an analytical model enhancer since its results are used to improve the optimization model parameters. The second

simulation is used to complement and generate the solution provided by the optimizer. According to Figueira and Almada-Lobo (2014), approaches that combine both simulation and optimization are used to better answer to uncertainty.

First of all, warehouses must be split into functional blocks according to the nature of their operations. As previously discussed, every warehouse comprises three major zones – inbound, storage/picking, and marshalling. Besides, picking can also be decomposed into PBL or PBS. After defining all these boundaries, the SOS framework can be applied:

1. First 'S' - Probabilistic Simulation: Having the processes mapped in-depth, a first simulation model can be developed to deploy scenario testing within each block and without interactions between different blocks. The main goal is to enhance an optimization model by feeding it with the most promising outputs. Making use of probabilities, it is possible to quickly simulate multiple configurations of each block regarding factors like: internal layout design, picking flow, or storage assortment strategies. Based on this, it is possible to estimate, for an increasing number of order lines, the distance (and time) required to fulfil orders considering different setups. The most interesting scenarios are the ones that minimize the travelled distances.
2. The 'O' - Block Optimization: Afterwards, the optimization model is used to further refine the chosen scenarios. For the PBL block, the goal is to dynamically change the storage assignment policy to minimize the total picking travel time. This need for scenarios is due to the fact that different factor combinations lead to different results. In fact, the optimal storage assignment policy heavily depends on the picking area design variables: internal corridor disposal, routing constraints, whether there is zoning or not, etc. So, this step improves the best combinations to choose the best among the best.
3. Second 'S' - Discrete Event Simulation: However, any combination of optimized blocks can be sub-optimal since interactions between blocks were not considered prior to this point. Hence, a validation phase is followed to test the robustness of the generated solutions. Discrete event simulation allows submitting the blocks to stress conditions and, therefore, uncover bottlenecks such as corridor congestion that were not considered until now. This is done to substitute any underperforming blocks with the second-best ones attained by the first simulation model, thus being an iterative search for the best overall solution.

Figure 3.1 presents a schematic perspective of the SOS methodology. The PBL block is highlighted since it is the focal point of this dissertation. Once more due to its complexity and shortened time window, when compared to the project's expected duration, this work only covers the first S.

3.2 The PBL Block

Considering the PBL flow type, a picking task begins with a pallet full of SKUs that must be distributed to the stores that ordered those items. Hence, a picking task can be viewed as a series of

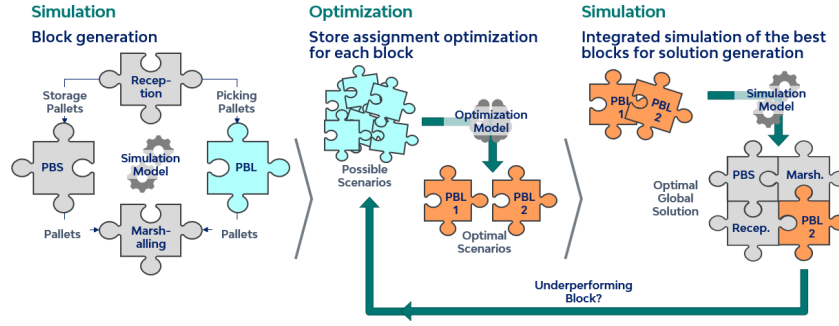


Figure 3.1: An overall perspective of the SOS methodology and its building blocks

stores to visit. Moreover, stores are assigned to one or more ground floor locations whether they pose many orders or not. Each location with a store assigned has a store pallet for the picker performing a picking task to drop off the items respective to that store. Since the pallets have normalized dimensions, the ground floor locations assigned to a store will all have the same dimensions. In theory, the PBL block can be modelled as a grid with each cell representing a potential ground floor location that can be assigned to a store. Figure 3.2 illustrates one of those cells.

3.2.1 Grid Modelling

As mentioned before, each cell can have a store pallet with length, l_p , and width, w_p . Besides, it is a good practice to leave a small gap between two consecutive pallets. This gap is usually positive in the direction parallel to the pallet length and it is denoted by g_l . However, the gap in the direction parallel to the pallet width, g_w , is often null. Nevertheless, equations 3.1 and 3.2 are used to define the cell length, l_c , and cell width, w_c , respectively.

$$l_c = l_p + g_l \quad (3.1)$$

$$w_c = w_p + g_w \quad (3.2)$$

Once the cell dimensions are fully defined, it is possible to parameterize the picking area accordingly. In fact, the PBL block can be modelled as a grid, G , as depicted in Figure 3.3, with N_{row} rows and N_{col} columns. Thus, the number of rows and columns will depend on the dimensions of the picking area. Equation 3.3 estimates the number of rows as being the maximum integer of cells that can be stacked on top of each other within the grid width, w_g . Similarly, equation 3.4 derive the number of columns alongside the grid length, l_g

$$N_{row} = \frac{w_g}{w_c} \quad (3.3)$$

$$N_{col} = \frac{l_g}{l_c} \quad (3.4)$$

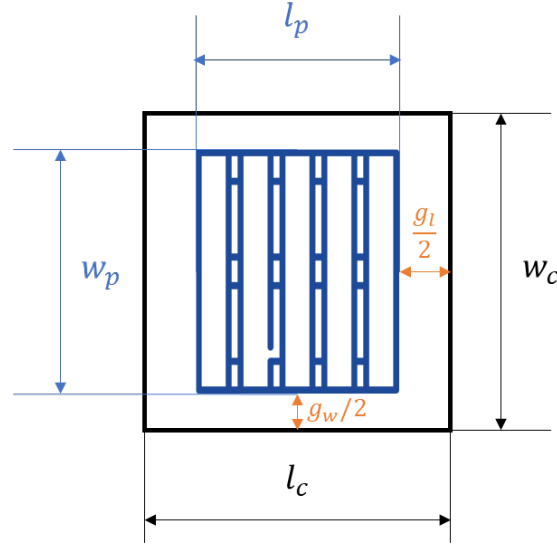


Figure 3.2: Cell as a ground floor location of the PBL block

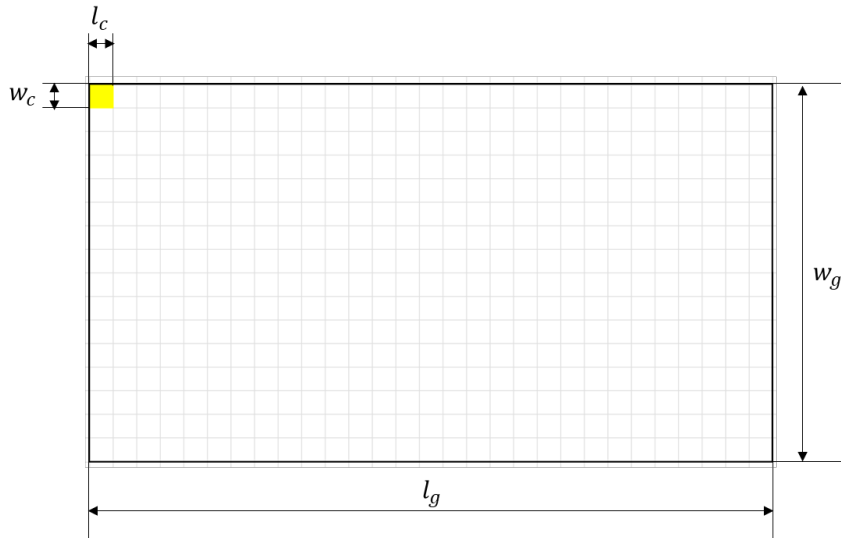


Figure 3.3: PBL block modelled as a grid

3.2.2 Path Cells and Non-Path Cells

Theoretically, each cell can have a store assigned. However, the routing methods presented in section 2.2.3.5 show that, in practice, there are path cells and non-path cells. Logically, path cells are needed to reach the non-path cells. Moreover, these non-path cells can either be active store cells or buffer store cells.

By definition, buffer store cells have a store assigned but they are no longer needed to be reached by the picker performing a picking task. In fact, when a store pallet (in an active store cell) reaches its full capacity, the picker closes it, sends it to a buffer position (a buffer store cell) and opens a new store pallet on the spot. Also, these so-called buffer cells are just next to the respective active cells. Hence, the picker only needs to push the new store pallet towards the old

one in order to send the latter to a vacant buffer position. That is why g_w is often null since pickers only access active store pallets alongside the pallet's length.

Besides, pickers perform retrieval tasks to move the store pallets (in buffer positions) onto the marshalling area. In fact, those tasks can be done simultaneously with picking tasks, but by a different picker, since each picker can only perform one task at a time. Hence, path cells can be further classified into picking cells, retrieving cells or routing cells. Routing cells have no tasks and are solely used to traverse the warehouse.

Figure 3.4 better illustrates this flow for a given store with both active cells (in blue) and buffer cells (in orange). Picking path cells and retrieving path cells are also represented in yellow and grey, respectively. Despite not being mentioned, routing path cells will be left in blank.

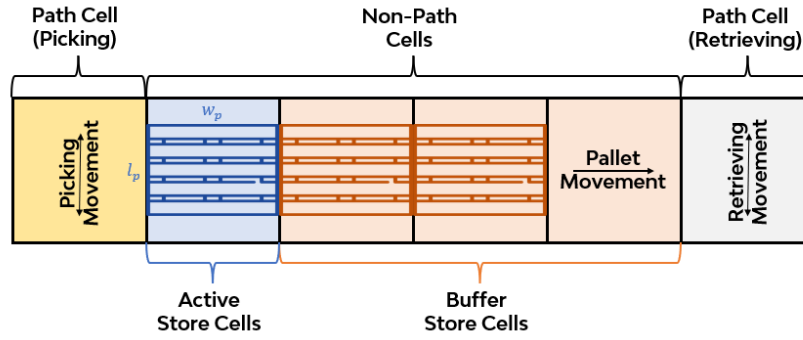


Figure 3.4: Flow for a given store with both active and buffer cells

3.2.3 Aisle-Based Cell Grouping

An aisle is, by definition, a passage between two entities. Bearing this in mind, a picking aisle, a , can be viewed as an oriented set of path cells with active store cells on both its ends. For example, a picking aisle that expands vertically has active store cells both to its right and to its left.

Aisles can either be unidirectional or bidirectional which will affect the routing methods that can be deployed and, consequently, the aisle width, $w(a)$, since there must be at least one path cell per each possible direction. On the other hand, the aisle length, $l(a)$, depends on the number of stores in that aisle. Considering again a picking aisle that expands vertically, there are two sets with $N_{pos}(a)$ active store cells stacked on top of each other at both aisle's ends. In fact, this cell grouping resembles the supermarket aisles. So, it can be said that each aisle has two racks, $r_1(a)$ and $r_2(a)$.

Besides, active store cells can have buffer store cells and, therefore, retrieving path cells associated. Hence, each position of a given rack, r , must comprise one active store cell which can be followed by $N_b(r)$ buffer cells, and by $N_p(r)$ path cells (relative to the retrieving aisle). That is why each position of a rack is associated with only one store, s . Figure 3.5 better depicts this aisle-based cell grouping.

A retrieving aisle, f , can be viewed as an oriented set of path cells with buffer store cells on both its ends. Since this work does not encompass any marshalling activities, the retrieving aisles

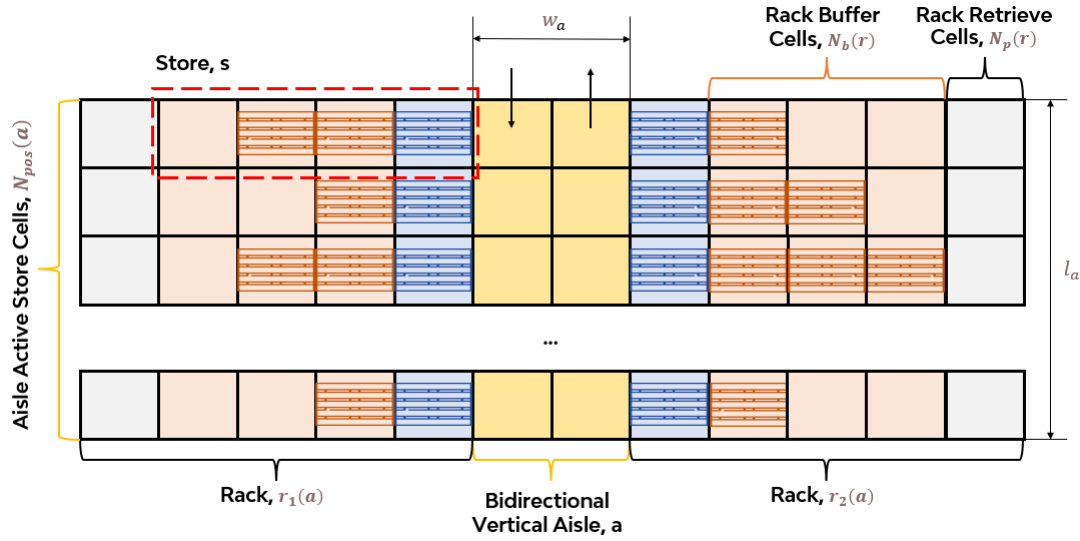


Figure 3.5: Aisle-based cell grouping

will be modelled as forbidden cells, that is, non-path cells belonging to picking aisles' racks. That means that the path cells of the retrieving aisles will be assigned to racks of the picking aisles. In analogy, a retrieving aisle that expands vertically will be split in half so that half of its width could be modelled as forbidden cells of the previous and next picking aisles' racks, respectively, as shown in Figure 3.6.

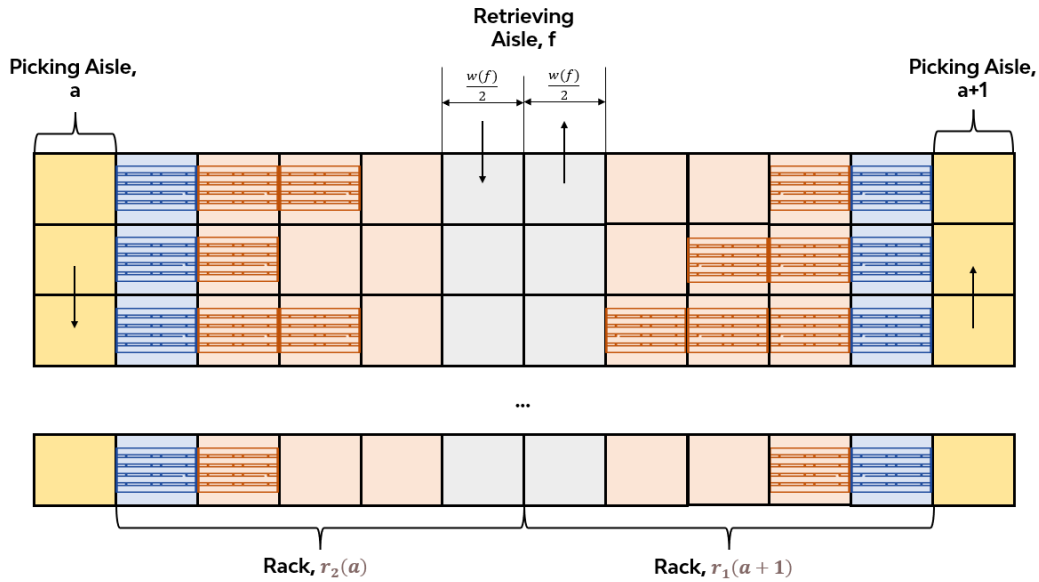


Figure 3.6: Modelling retrieval aisles as forbidden path cells

After this assumption, it is possible to mathematically define disjoint sets of cells regarding their nature. Each set is also assigned with a colour to better visualize their differences throughout this work. Table 3.1 summarizes all the groups that have been made.

Table 3.1: Division of the grid in cell clusters

Set	Subset	Color Code
Path Cells, C_p	Picking Cells, C_{pp}	Yellow
	Routing Cells, C_{pn}	White
	Retrieving Cells, C_{pf}	Gray
Non-Path Cells, C_n	Active Cells, C_{np}	Blue
	Buffer Cells, C_{nn}	Orange
	Impossible Cells, C_{nf}	Black

As mentioned before, retrieving path cells were modelled as forbidden since this work does not encompass any marshalling activities. In practice, they can be seen as non-path cells for the picker performing a picking task, but in theory there are no obstacles blocking their access. In fact, those physical constraints (like structural warehouse pillars) are well represented in the impossible cells subset. This is the main difference between those two groups: retrieving cells shall not be traversed in picking tasks whereas impossible cells can not be traversed at all.

It is also worth mentioning that active cells and buffer cells are allocated to a single store for a given period in time. In case they are not, it is always possible to assign them stores anywhere in the future. Therefore, it could have been made a further division of these cells whether they currently have a store or not. Nevertheless, this distinction is well captured by the path cells. In fact, active store cells are the ones that have picks and can only be reached if the picker is in a picking path cell. Thus, potential active cells don't have picks and, for that reason, they are next to a routing path cell.

3.2.4 Understanding Layout Zones

In practice, it is almost impossible to find two warehouses exactly equal. Nevertheless, the PBL area can be accurately reduced to a grid with aisles and their respective racks. However, the way racks are disposed, largely differs from warehouse to warehouse and, thereby, fill in the grid is a crucial user-defined input. In analogy with the Sea Battle strategy game, there are millions of combinations to fill in a grid, and ships (racks in this case) can be disposed either horizontally, vertically, or a mix of both.

The way racks are connected with each other is also important. In fact, if racks were to be buildings, aisles would be the streets used to reach them. Therefore, it is a good practice to design the picking area in a way that resembles the New York City's topography. By doing that, there may be some zones with horizontal blocks and some zones with vertical blocks. Zoning motivated the division of the first simulation model into two major phases: (1) simulating layout zones and (2) simulating the whole layout (with mixed zones).

Most of the picking areas only have a single horizontal zone or a single vertical zone but there are others that comprise both. Zoning hereby described refers to the geometric decomposition of the picking area based on its racks' (and aisles') orientation. On the other hand, zoning (as defined

in section 2.2.3.4) assumes that a picker is assigned to a single zone in the picking area. This is a managerial decision that usually distinguishes zones based on product familiarity.

When simulating layout zones that picker division assumption has to be made. However, it may not stand for areas with more than one geometric zone. Because of that, the model that emulates layout zones independently will be firstly introduced. Then, the necessary adjustments will be made in order to become general and applicable to any existing picking area.

3.3 Gap to Literature

The SOS framework was initially developed in Alves (2015) and further refined in Amorim-Lopes et al. (2020). However, the analytical model that estimates the expected travel distances can only be applied to picking areas where corridors are disposed either horizontally or vertically, not both.

This thesis made improvements to the existing model in chapter 4, but its major contribution lies in the development, in chapter 5, of a second model that can be applied to any possible picking area. This was only achievable due to the innovative approach of modelling the picking area as a grid.

Besides, the existing methodology was originally thought for PBS only. So, this work aims to introduce all these novelties into the existing literature and extend this rationale to a PBL flow type.

Chapter 4

First 'S': Simulating Layout Zones

In order to select the best possible layout, an analytical model was developed to estimate the expected travel distance inside a given picking area with only one orientation. The model output depends on several factors: task size, store assignment policy, internal zone layout, and routing constraints. This allows for a comparative analysis of multiple user-defined configurations in a very short amount of time. The ultimate goal is to quickly generate and access several layouts before optimizing, in a next phase, the store allocation policy to be deployed.

4.1 Notation

This section has the purpose of clarifying and introducing the remaining notation that will be used throughout this first approach. For that, consider Figure 4.1 that depicts a 2D top view of one possible picking area with vertical racks (and aisles) only and a single inbound/outbound point.

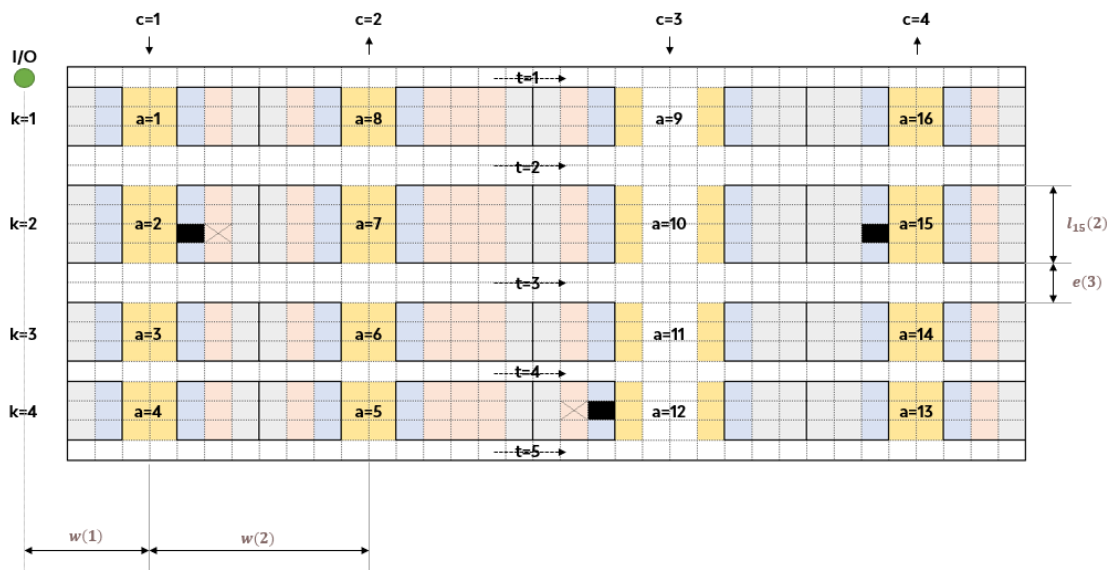


Figure 4.1: Possible picking area with vertical aisles only (Cell granularity level)

The layout is divided vertically into corridors and horizontally into levels. Corridors are indexed from left to right using $c = \{1, \dots, C\}$, whereas levels are coded top to bottom using $k = \{1, \dots, K\}$. In the interception between a corridor and a level, there is an aisle $a = \{1, \dots, A\}$. Besides, each level is located between two cross-corridors, defined top to bottom using $t = \{1, \dots, T\}$.

Each corridor, c , can only be crossed in either one of two directions - top to bottom or bottom to top. Moreover, each corridor is composed by K aisles. Thus, it is possible to define two sets: one comprehending the aisles from corridors that are crossed top to bottom (set D) and another composed by the aisles from corridors that are crossed in the opposite direction (set I).

Aisles are numbered from left to right and alongside the corridor direction. Also, as mentioned before, each aisle can have associated a rack, r , on both its ends. In this case, one to its left, $r_1(a)$, and another one to its right, $r_2(a)$.

Cross-corridors, t , are only traversed from left to right. They are used, as the name implies, to go from one corridor, c to another corridor, c' such that, $c < c'$. This creates shortcuts that pickers may take to avoid following the longest path.

Finally, the measures that fully define a picking area are also represented. Each aisle's length is $l(a)$, the distance between the centre of two consecutive corridors is $w(c)$, and the width of each cross corridor is $e(t)$. Besides, all the aisles in a level have the same length. So, the aisle's length depends on the level in which the aisle is inserted, thereby the new notation $l_a(k)$.

4.2 Assumptions

As previously stated, this first model splits the picking area into horizontal and vertical zones. Moreover, each zone is assumed to be independent with pickers never crossing more than one zone to perform a picking task. Hence, each zone can be analyzed separately.

Also picking tasks are assumed to be independent from each other but, in reality, this assumption may not stand for very large warehouses. To better understand this consider the following: warehouses pose orders to their suppliers in order to meet their store's requirements. Each order has a certain amount of a given SKU and it will originate one or more picking tasks accordingly. So, if an order generated two picking tasks, the route of the second picking task will depend on the stores already fulfilled by the first picking task.

This adds an additional layer of complexity considering that the items of different orders might be combined, in the reception phase, to originate a new picking task. In fact, orders only have one SKU but the same may not be true for the picking tasks. Nevertheless, this work only embraces the PBL block and assumes the preparation of the picking pallets as an input given by the Reception block.

Finally, this first model does not require the cell granularity level that has been introduced until now. This will be most needed for modelling the transitions between different zones. Instead, the degree of detail will decrease to the aisle level. In fact, an aisle is no more than a group of cells as thoroughly seen in section 3.2.3. By doing that, it decreases the computational effort that is necessary to run the zone model.

4.3 Mathematical Model

Most of the picking time is actually spent on travelling between pick locations. Therefore, improving the picking efficiency requires minimizing the total picking time which, in turn, is an increasing function of the travelled distance. For that purpose, not only the zone's internal layout (number, size, direction, and orientation of picking aisles) but also several other factors such as the store allocation policy in use (closest open position, random assignment, COI-based ABC categorization, etc.) may be changed.

All these degrees of freedom will then enable a myriad of scenarios to test. The ultimate goal is to estimate the expected picking time, $E[time]$, that is spent inside a given picking zone, for a given picking task of size N . This can be estimated by equation 4.1, adding up the expected time spent travelling (assuming that pickers travel at an average speed of v) to the expected time spent on actually picking "down" the items, t_{pick} . As mentioned before, deliveries are seen as picks for the PBL flow type.

$$E[time] = \frac{E[L]}{v} + N \cdot t_{pick} \quad (4.1)$$

However, t_{pick} largely depends on the active store pallets that are going to be visited on a given picking task. If they happen to be empty, the picker only has to drop off the ordered items. If they are not, the picker will most certainly have to reorganize the pallet. This is done because, as a rule of thumb, bigger and heavier items must occupy the lower levels of a pallet. In order to estimate t_{pick} , the picking tasks should be known in advance. This knowledge may bias the model since it has to be general enough to estimate times for any picking task that could appear in the warehouse.

Hence, a simplification had to be made and, from now on, the objective of this model is to estimate the expected picking tour length, $E[L]$, for any picking task of size N . In fact, minimizing the travelled distance is also considered a key objective in order-picking optimization. Bearing this in mind, equation 4.2 splits the expected travel distance into two components: (1) the expected travel distance within-aisles, $E[L_W]$, and (2) the expected travel distance crossing-aisles, $E[L_C]$.

$$E[L] = E[L_W] + E[L_C] \quad (4.2)$$

Logically, the expected travel distance within-aisles can easily be obtained by summing up all the expected travel distances within each aisle, $E[L_a]$, as depicted in equation 4.3.

$$E[L_W] = \sum_{a=1}^A E[L_a] \quad (4.3)$$

The expected contribution of an aisle, a , to the total picking tour length also depends on two factors: (1) the distance needed to fully traverse the aisle, d_a , and (2) the probability of traversing it, $p(S_a)$. Therefore, $E[L_a]$ can be expressed as of equation 4.4.

$$E[L_a] = d_a \cdot p(S_a) \quad (4.4)$$

In order to derive d_a , it is assumed that pickers can only traverse the warehouse throughout its corridors' (and cross-corridors') mid-points. In other words, pickers travel alongside the imaginary lines that split these entities in half. This makes sense only when considering the aisle granularity level.

Bearing this in mind, Figure 4.1 can be further simplified onto a scheme like the one depicted in Figure 4.2 and despite the absence of a grid, all the measurements can be parameterized concerning the cell dimensions.

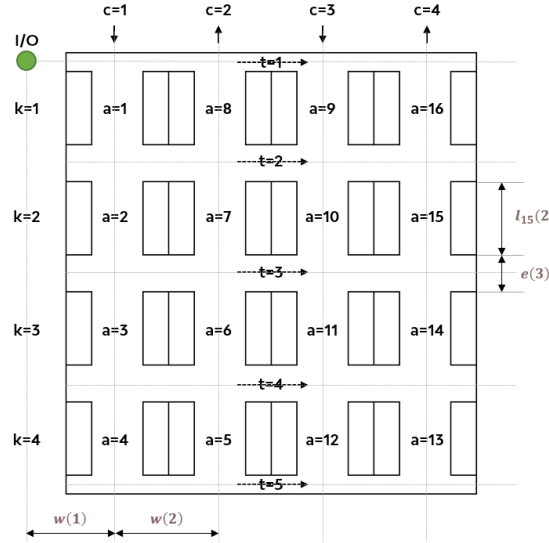


Figure 4.2: Possible picking area with vertical aisles only (Aisle granularity level)

Nevertheless, all the aisles on the same level have the same length. So, the distance needed to fully traverse an aisle is going to be equal to the distance needed to fully traverse the level on which that aisle is inserted. Since each level k is comprised by two consecutive cross-corridors, $t=k$ and $t'=k+1$, that distance can be obtained by the expression in equation 4.5.

$$d_a = \frac{1}{2} \cdot e(k) + l_a(k) + \frac{1}{2} \cdot e(k+1) \quad (4.5)$$

The complicated part is computing $p(S_a)$, or in other words, the probability of traversing an aisle a . There are only two reasons from which an aisle needs to be traversed and they are mutually exclusive: (1) at least one item has to be delivered on that aisle or (2) no items have to be delivered on that aisle despite being part of the shortest path to complete the task. These events can be denoted as E_a and V_a respectively.

To better understand these events, consider Figure 4.2 and suppose that there is a task with only two picks: one located in aisle a_2 and the other one located in aisle a_4 . Thus, only these two aisles will experience the event E_a . Besides, this model assumes that each task always follows the shortest path available. Hence, the event V_a is present for all the aisles without picks involved in the picking tour, which, in this case, is: a_1 through a_4 , plus a_5 through a_8 since the picker has to return to the I/O point.

Therefore, $p(S_a)$ can be expressed, for each aisle, as the arithmetic sum of the individual probabilities $p(E_a)$ and $p(V_a)$, as in equation 4.6.

$$p(S_a) = p(E_a) + p(V_a) \quad (4.6)$$

In order to perform all these calculations, it is crucial to know, *a priori*, the probability of a pick addressing an aisle a . Besides, the probability of a given event is, by definition, the ratio between the number of favourable outcomes and the number of possible outcomes. With that being said, $p(P_a)$ can be represented by equation 4.7.

$$p(P_a) = \frac{\text{Number of Picks in Aisle } a}{\text{Number of Total Picks}} \quad (4.7)$$

The number of total picks can be calculated by estimating each store's requirements for a given period in time. Furthermore, each store is assigned to a position of a rack, belonging to an aisle a , for a given period in time that may or may not be coincident with that used in the demand forecasts. For example, store 1 has projected to order 100 different SKUs for the following month. However, that same store is going to be allocated to different aisles throughout that next month. The total number of picks is 100 but the number of picks in each aisle will depend. Therefore, $p(P_a)$ becomes fully defined when knowing both: (a) the store allocation plan, and (b) the store demand forecasts (or if possible the actual store orders).

It is now clear that $p(P_a)$ is the probability of a pick addressing an aisle a but for a task with a single pick. In fact, if picks were to be marbles numbered with their destination aisles, $p(P_a)$ was the probability of retrieving a single marble, with number a , from a bag with all the picks.

Dissecting equation 4.8, the term $1 - p(P_a)$ can be seen as the probability of a pick not addressing an aisle a , for a task with a single pick. Raising it to the power of N , it computes the probability of a pick never addressing an aisle a , for a task with N picks. Finally, by taking the complementary, it outputs the probability of at least one of the N picks addresses the aisle a , or in other words, $p(E_a)$.

$$p(E_a) = 1 - [1 - p(P_a)]^N \quad (4.8)$$

Having $p(E_a)$ well defined, the following steps will be focusing on how to calculate $p(V_a)$. In order to know the probability of traversing an aisle a , when no items have to be delivered on that aisle, one shall first compute all the routes in which it happens.

As a matter of fact, each picking tour with N tasks can be decomposed in $N+1$ sub-tours. For that, consider m as the last aisle with a picking task, a' as the next aisle with a picking task, and a as the aisle in analysis to calculate the probability $p(V_a)$.

Besides, the picker ends the trip at the same point where it began - the I/O point. Since the I/O point is always part of the route, it can be seen as a virtual aisle with a virtual picking task. Hence, for the first picking task, the last aisle is the Inbound point ($m=0$); for the last picking task, the following aisle is the Outbound point ($a'=A+1$).

Moreover, all the $N+1$ ordered pairs, (m, a') , that compose a picking tour, may or may not have a direct move between the last picking position m and the next picking position a' . In case there is not a direct move, the picker will certainly traverse the set of aisles a that comprise the shortest path between m and a' . These aisles a do not have items to be delivered.

Therefore, the first step to attain $p(V_a)$ is to calculate the probability of going directly from aisle m to aisle a' without having to perform a picking task in between. Equation 4.9 provides a general expression for this probability, denoted from here afterwards by $p(E_m, E_{a'})$.

$$p(E_m, E_{a'}) = \begin{cases} \left[1 - \sum_{j=m+1}^{a'-1} p(P_j) \right]^N, & \text{if } m < a' \\ 0 & \text{otherwise} \end{cases} \quad (4.9)$$

Digging deeper into the above formula, the summation retrieves the probability of having a pick in the path between m and a' (both exclusive), for a task with only one pick. Logically, when taking the complementary and raising it to the power of N , it outputs the probability of a pick never addressing any of the aisles in between, for a task with N picks.

Shortcuts complicate matters further since this model assumes that each task always follows the shortest path available. For example, let $m=1$ and $a'=7$. Considering Figure 4.2, it is clear to see that aisle a_7 can be reached from aisle a_1 by taking a shortcut after aisle a_2 . In this case, aisle a_2 is the only aisle to be traversed without having any pick.

Generalizing, there are three cases in total that have to be addressed for calculating $p(V_a)$ when the "just visiting" aisle is at an odd corridor. This set of aisles is denoted by D as formerly mentioned in Section 4.1. Also, it is possible to compute the probabilities of "just visiting" aisles at even corridors in a similar fashion. This second group of aisles belong to domain I .

Table 4.1: Cases to address for calculating $p(V_a)$

Visiting Aisle, a	Last Pick, m	Next Pick, a'	Probability
Set D	Inbound/Start Point	Any	$p(V_{sd})$
	Set D	Any	$p(V_{dd})$
	Set I	Any	$p(V_{id})$
Set I	Set D	Any	$p(V_{di})$
	Set I	Any	$p(V_{ii})$
	Any	Outbound/End Point	$p(V_{ie})$

Table 4.1 summarizes all these possibilities when a given aisle a needs to be traversed for a route $m \rightarrow a'$. considering that $m < a < a'$. In addition, equation 4.10 describes the mathematical formulation behind this rationale allowing the calculation of $p(V_a)$ for any given aisle a .

$$p(V_a) = \begin{cases} p(V_{sd}) + p(V_{dd}) + p(V_{id}), & \text{if } a \in D \\ p(V_{di}) + p(V_{ii}) + p(V_{ie}), & \text{if } a \in I \end{cases} \quad (4.10)$$

In order to dissect all these cases, this work will firstly deduct the expressions for the visiting aisles a belonging to set D , and afterwards extend them to the visiting aisles of set I . With that being said, consider Figure 4.3 to better understand each one of those three cases depicted in the first branch of equation 4.10.

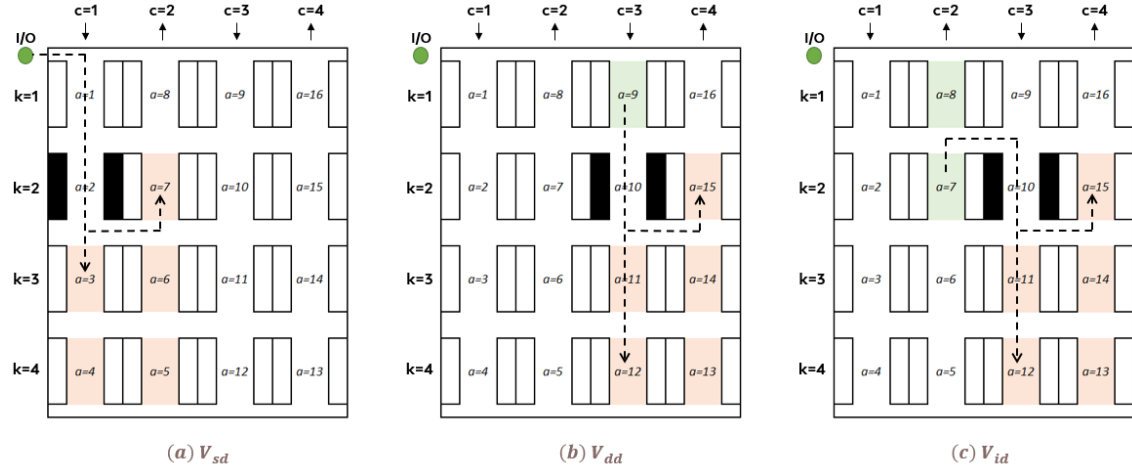


Figure 4.3: Illustration of all the instances when an aisle $a \in D$ needs to be visited

The first case to address is $p(V_{sd})$. This occurs when the picking task has just started and the picker is at the depot. It comprises all the possibilities to traverse an aisle $a \in D$ in order to reach the aisle with the first picking task, a' . In other words, this case encompasses all the possibilities of an aisle a , belonging to an odd corridor, to be in the shortest route between the Inbound Point and the aisle a' .

The probability of doing such a direct path is given by $p(E_0, E_{a'})$. However, the picker only does this path, if there is at least one pick in a' . Hence, the probability of such a route occurring in a picking task with N picks is given by that combination or, mathematically speaking, by $p(E_0, E_{a'}) \cdot p(E_{a'})$.

Bearing this in mind, consider the following colour code: in order to traverse an aisle a (coloured in black), that aisle must be part of the shortest path between two consecutive points of interest in a picking task (coloured in green and orange, respectively). Logically, the points of interest are the I/O point and the aisles with picking locations.

For the example represented in Figure 4.3a, if the picker has to deliver the first item on any of the orange aisles, he must visit the black aisle as well. This occurs when that first item is at an aisle a' that is either: (1) on the same corridor, $c_{a'} = c_a$, but one or more levels above $k_{a'} > k_a$, or (2) on the next corridor $c_{a'} = c_a + 1$ but on the same level or above $k_{a'} \geq k_a$. Generalizing, we obtain equation 4.11:

$$p(V_{sd}) = \sum_{\substack{a' \in D \wedge c'_{a'} = c_a \wedge k'_{a'} > k_a \\ \vee \\ a' \in I \wedge c'_{a'} = c_a + 1 \wedge k'_{a'} \geq k_a}} p(E_0, E_{a'}) \cdot p(E_{a'}) \quad (4.11)$$

One may be wondering that if a picker has to deliver an item in aisle a_{11} , there is a shortest path that traverses a_2 , given by the sequence: $I/O \rightarrow a_1 \rightarrow a_2 \rightarrow a_{11}$. However, the summation's domain does not encompass that situation because this is not a mandatory shortest path. In other words, if the first picking task is on aisle a_{11} , the aisle a_2 can be traversed but it is not mandatory to do that. This model applies the same rationale throughout all the remaining cases.

The second and third cases can occur for the remainder of a picking task. They comprise all the possibilities to necessarily traverse an aisle $a \in D$ for all the shortest paths that connect the last picking location, m , to the next picking location, a' . The key difference between both cases resides on where the last picking position is located.

Nevertheless, there is at least one pick in m . Hence, the probability of doing a route $m \rightarrow a'$ can easily be derived by combining that event's odd. Equation 4.12 provides a general expression to estimate the probability of such a path occurring in a task with N picks. However, if either m or a' happens to be the I/O point, $p(E_m)$ and $p(E_{a'})$ are, respectively, equal to 1. This is true because a picking task always starts and finishes at that point.

$$p(m \rightarrow a') = p(E_m) \cdot p(E_m, E_{a'}) \cdot p(E_{a'}) \quad (4.12)$$

On the one hand, $p(V_{dd})$ can happen when the last pick is at an aisle belonging to set D . Figure 4.3b, better illustrates all the instances in which the picker must visit aisle a on its way to the next picking position. In fact, the black aisle must be traversed for any combination of paths starting at a green aisle and ending at an orange aisle.

This occurs when the aisle with the last picking position m is on the same corridor, $c_m = c_a$, but one or more levels below, $k_m < k_a$. Then, for every aisle m , the next picking position is located at an aisle a' that can either be: (1) on the same corridor, $c_{a'} = c_a$, but one or more levels above, $k_{a'} > k_a$, or (2) on the next corridor, $c_{a'} = c_a + 1$, but on the same level or above, $k_{a'} \geq k_a$. Mathematically, equation 4.13 encompasses all these scenarios.

$$p(V_{dd}) = \sum_{\substack{m \in D: m < a \\ c_m = c_a \\ k_m < k_a}} p(E_m) \cdot \sum_{\substack{a' \in D \wedge c'_a = c_a \wedge k'_a > k_a \\ \vee \\ a' \in I \wedge c'_a = c_a + 1 \wedge k'_a \geq k_a}} p(E_m, E_{a'}) \cdot p(E_{a'}) \quad (4.13)$$

On the other hand, $p(V_{id})$ can only be triggered if the last pick is at an aisle belonging to set I . Considering Figure 4.3c, an aisle a is traversed when the aisle with the last picking position, m , is on the previous corridor, $c_m = c_a - 1$, and at the same level or below, $k_m \leq k_a$. Then, for every aisle m , the next picking positions located at the aisles a' are the same as in the previous cases. This can easily be seen by the pattern of the orange-shaped aisles. Therefore, only the domain of green aisles needs to be changed in order to derive equation 4.14.

$$p(V_{id}) = \sum_{\substack{m \in I: m < a \\ c_m = c_a - 1 \\ k_m \leq k_a}} p(E_m) \cdot \sum_{\substack{a' \in D \wedge c'_a = c_a \wedge k'_a > k_a \\ \vee \\ a' \in I \wedge c'_a = c_a + 1 \wedge k'_a \geq k_a}} p(E_m, E_{a'}) \cdot p(E_{a'}) \quad (4.14)$$

Having tackled all the scenarios in which an aisle a , belonging to set D , must be visited, despite not having any pick, the same has to be done for all the aisles a in I . For that, consider Figure 4.4 to better understand each one of the remaining cases depicted in the second branch of equation 4.10.

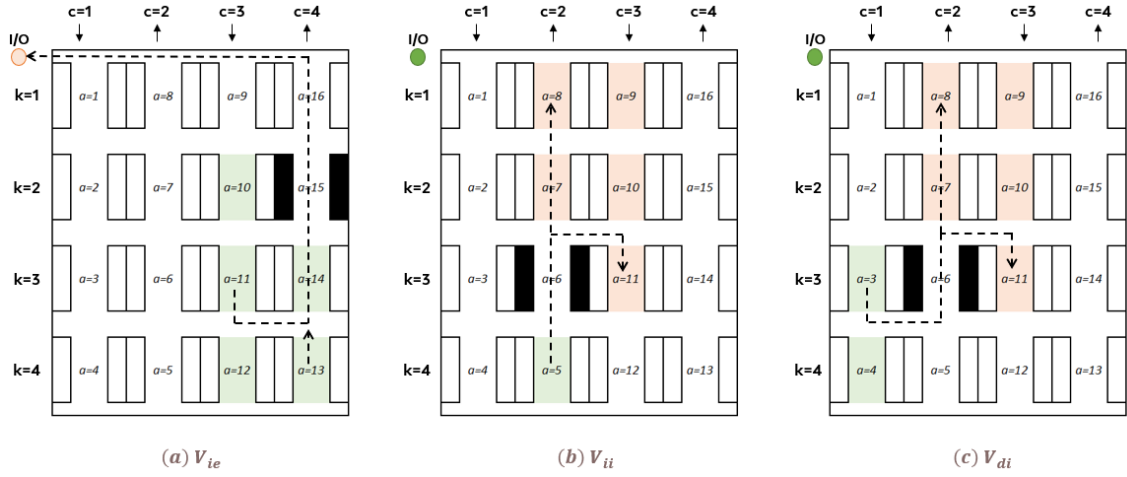


Figure 4.4: Illustration of all the instances when an aisle $a \in I$ needs to be visited

It seems like every case is correlated to one of the cases already seen. As a matter of fact, the first case in Figure 4.4 resembles the first case in Figure 4.3 and so on. Following that line of thought, the first case to address is $p(V_{ie})$. This case covers all the scenarios for when an aisle a needs to be visited and the task is about to finish. In comparison, $p(V_{sd})$ addresses all these possibilities but for when the task has just started. Therefore, equation 4.11 needs to be adapted accordingly.

For the example represented in Figure 4.4a, if the picker has just delivered the last item on any of the green aisles, he must visit the black aisle as well in order to leave the picking area. This occurs when the last pick is at an aisle m that is either: (1) on the previous corridor, $c_m = c_a - 1$, but on the same level or above, $k_m \geq k_a$, or (2) on the same corridor, $c_m = c_a$, but one or more levels above, $k_m > k_a$. Equation 4.15 can be used to capture these changes.

$$p(V_{ie}) = \sum_{\substack{m \in D \wedge c_m = c_a - 1 \wedge k_m \geq k_a \\ \vee \\ m \in I \wedge c_m = c_a \wedge k_m > k_a}} p(E_m) \cdot p(E_m, E_{A+1}) \quad (4.15)$$

This is the exact opposite of $p(V_{sd})$. Instead of being part of shortest paths with the same point of origin, each aisle a is now part of shortest paths with the same point of destination. This point can be seen as a virtual aisle with a virtual pick location. The only nuance is that, for the first case this point serves as an inbound point and for the current one as an outbound point.

The second and third cases, of Figure 4.4, are somehow mirrored images of the ones previously discussed in Figure 4.3. This can be explained because they complement each other in what can happen anywhere in the middle of a picking task. Therefore, $p(V_{ii})$ can easily be computed by

simply changing two things in equation 4.13. These are the inequalities and the sets used in the summations' domains. Set D becomes Set I , whereas a "greater than" gives place to a "smaller than". Nevertheless, one can still deduct the same expression by looking up to the example in Figure 4.4b.

$$p(V_{ii}) = \sum_{\substack{m \in I : m < a \\ c_m = c_a \\ k_m > k_a}} p(E_m) \cdot \sum_{\substack{a' \in D \wedge c'_a = c_a + 1 \wedge k'_a \leq k_a \\ \vee \\ a' \in I \wedge c'_a = c_a \wedge k'_a < k_a}} p(E_m, E_{a'}) \cdot p(E_{a'}) \quad (4.16)$$

Finally, an expression for $p(V_{di})$ is quickly obtained by adapting equation 4.14 in a similar fashion. Again, this can also be done by generalizing the example depicted in Figure 4.4c.

$$p(V_{di}) = \sum_{\substack{m \in D : m < a \\ c_m = c_a - 1 \\ k_m \geq k_a}} p(E_m) \cdot \sum_{\substack{a' \in I \wedge c'_a = c_a \wedge k'_a < k_a \\ \vee \\ a' \in D \wedge c'_a = c_a + 1 \wedge k'_a \leq k_a}} p(E_m, E_{a'}) \cdot p(E_{a'}) \quad (4.17)$$

Tracing back the path to equation 4.3, it is possible to conclude that all of this work cascades upwards to estimate the expected travel distance within-aisles, $E[L_W]$. Therefore, one final step must be addressed to reach the ultimate goal as defined in equation 4.2. For that reason, the expressions that allow the estimation of the expected distance crossing-aisles, $E[L_C]$, will be presented afterwards. Keeping things simple, this can easily be obtained by summing up all the expected travel distances across each corridor.

However, there is a catch. Consider Figure 4.2 and a picking task with only one item to deliver at aisle a_2 : the picker starts at the depot and has to reach corridor c_1 to eventually drop off the item. Then, the picker must leave the picking area, but this can only be accomplished throughout even corridors. Hence, the picker also has to reach corridor c_2 . So, corridors must be addressed in pairs. This is why, $E[L_C]$ must be obtained by adding up all the expected travel distances across each pair of corridors, $E[L_j]$, rather than corridors *per se*. Equation 4.18 describes this step.

$$E[L_C] = \sum_{j=1}^J E[L_j] \quad (4.18)$$

A pair of corridors, j , is composed by two consecutive corridors in a way that each corridor only belongs to one pair. On top of that, pairs are created following an ascending order. Considering Figure 4.2 once again, it can be said that: corridors 1 and 2 morph into the first pair whereas corridors 3 and 4 give origin to the second and last pair. Mathematically, the j -th pair is represented by $j = \{c_{2j-1}, c_{2j}\}$.

In addition, the expected contribution of each pair to the total picking tour length depends on two factors: (1) the distance needed to fully traverse the pair, d_j , and (2) the probability of traversing it, $p(S_j)$. Therefore, $E[L_j]$ can be expressed as of equation 4.19.

$$E[L_j] = d_j \cdot p(S_j) \quad (4.19)$$

It is known that the j -th pair has two corridors of width $w(c_{2j-1})$ and $w(c_{2j})$, respectively. The distance needed to fully traverse the pair will depend on both: (1) the width of that pair's corridors and (2): the width of all the p pairs prior to that one. Moreover, the picker has to return to the I/O point, so this distance needs to be doubled. Hence, d_j can be fully defined by equation 4.20.

$$d_j = 2 \cdot \sum_{p=1}^j [w(c_{2p-1}) + w(c_{2p})] \quad (4.20)$$

Finally, the probability of traversing a pair will depend on the probabilities of traversing the aisles belonging to the corridors of that pair. In theory, $p(S_j)$ would have to comprise all the possibilities to have at least one aisle with a visit. However, this value can be approximated by equation 4.21.

$$p(S_j) = \max_{a \text{ in } c \in j} P(S_a) \quad (4.21)$$

The model is now fully defined and all these expressions shall be used to derive the expected travel distance, $E[L]$.

4.4 Synthesis

This section explains in detail how to estimate the expected picking tour length for a task with N picks. This is done for a picking area with only one orientation. Until now, all the aisles (and their racks) expanded vertically. However, facilities can be rotated and flipped in order to apply these equations to any zone with only one orientation. Rotations transform vertical zones into horizontal zones, whereas flips change the corridors' direction. This can better be seen in Figure 4.5.

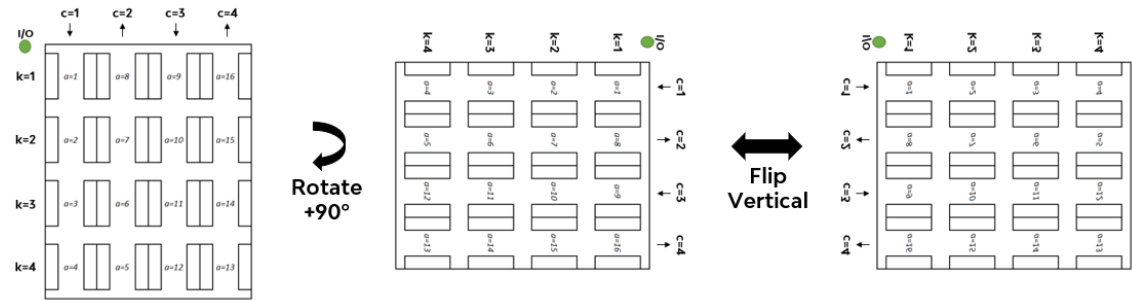


Figure 4.5: Rotations and flips allow the modelling of any zone with only one orientation

All in all, there is only the need for two critical inputs: (a) the store allocations and (b) the store requirements, both for a given time period. This will affect the probabilities, $p(P_a)$, of a pick addressing an aisle a and, inherently, all the other equations. Since aisles are ordered, these probabilities can be defined by a vector $\mathbf{p}$ in a way that for a given zone:

$$E[L] = f(N, \mathbf{p}) \quad (4.22)$$

Chapter 5

First 'S': Simulating The PBL Block

Usually, picking areas only contain aisles (and racks) that are disposed either vertically or horizontally. Few are the warehouses that comprise both situations. Moreover, the ones who do that usually practice zoning, hence dividing the picking area into smaller business units with only one orientation. Therefore, the analytical model developed prior to this point can be applied to the majority of cases.

5.1 Motivation

Nevertheless, the goal of the project, in which this thesis is inserted, aims to develop a product that can be applied to any given scenario. This can only be done by entering at the cell granularity level. To better understand the flaws of the previous model, consider Figure 5.1 which illustrates two possible picking areas - one mixed (cell granularity level) and the other one not (aisle granularity level). It is also worth mentioning that the aisles' racks are condensed for simplicity purposes. Their true aspect is similar to the one depicted back in Figure 4.1.

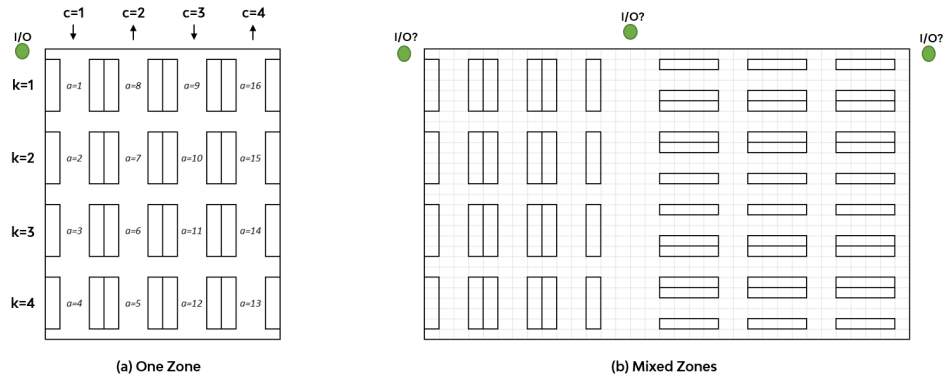


Figure 5.1: Differences between a picking area with one and both orientations

The model assumes that the aisles must be ordered, in such a way, that if a picker had a pick in all possible locations, he would have to follow the shortest path available. Clearly, this depends on where the I/O point is located. This was not a problem, until now, because there is only one

way to traverse all the aisles inside a picking area with a single orientation (as depicted in Figure 5.1a. To be more precise, the depot could be located in any corner of the picking area. However, this can easily be attained by applying any combination of rotations and flips, as seen in Figure 4.5.

Mixed zones complicate matters further. Figure 5.1b illustrates several ways to place the I/O point and each one of them will consequently lead to different aisle ordinations. Moreover, the inbound and outbound points are not necessarily the same. Hence, these two entities must be defined. *a priori*, by the user. On top of that, the aisle granularity level would never work for the example in which the inbound point is at the junction between zones. This is because the same aisle can be visited multiple times on the same picking task. This can better be seen in Figure 5.2.

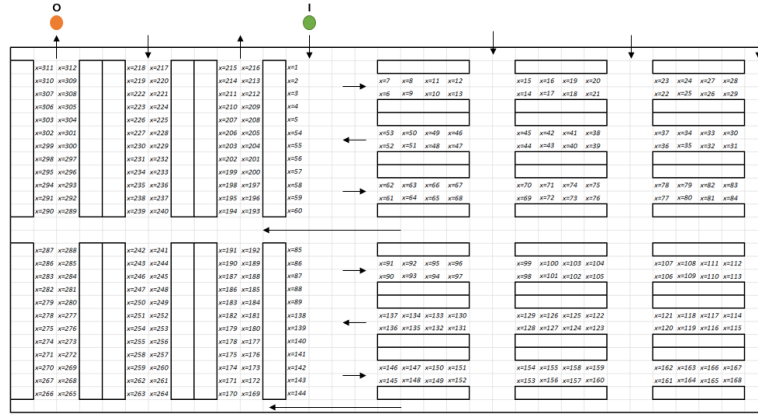


Figure 5.2: Order of all the possible picking positions for a picking area with mixed zones

Applying Dijkstra (1959)'s algorithm to the situation described above, it is possible to know the shortest path to visit all the possible picking locations, starting at the inbound point and finishing at the outbound point. Therefore, it is clear that the picker partially traverses the first vertical aisle (just below the I/O point) before entering the horizontal zone. Then, he must visit all the aisles belonging to the first and second corridors of that zone, before traversing the remainder of the first vertical aisle.

The same aisle was visited two times for the same picking task. However, each cell of that aisle was only visited once. That explains why the cell granularity level will be used from now onwards. Only cells, not aisles, can have a unique order of visit in the warehouse.

5.2 Notation and Assumptions

The model that emulated layout zones independently was firstly introduced. Now, the necessary adjustments will have to be made in order to attain an enhanced version, general enough to be applicable to any picking area.

Most of the notation used, prior to this point, remains the same. Still, the simplification that saw an aisle as the elementary unit must fall. In fact, the derived expressions are only valid if

there is a well-defined visiting order of all the possible picking locations. For picking areas with mixed zones, however, this order may only be achievable at the cell granularity level. So, all the probabilities computed for all the aisles, a , must now be addressed for all the path cells, $x \in C_p$.

For that, it is worth recalling section 3.2. In a nutshell, path cells (Set C_p) are used to reach non-path cells (Set C_n) - in particular active store cells with store pallets (Set C_{np}). Path cells can be further divided into picking cells (Set C_{pp}) or routing cells (Set C_{pn}) whether they are next to an active store cell or not. Logically, routing cells are solely used to traverse the warehouse whilst picking cells are routing cells that can also be used to pick items.

Bearing this in mind, all the picking path cells may have a picking index, i . This is given by the order in which that cell would have been traversed, if there was a task that needed to reach all the possible picking locations or, in other words, all the picking path cells (just like in Figure 5.2).

Since each task always starts and finishes at the same points, the picking index will depend on three things: (a) the inbound location, (b) the outbound location, and (c) the routing constraints. Therefore, these attributes are critical inputs for the model.

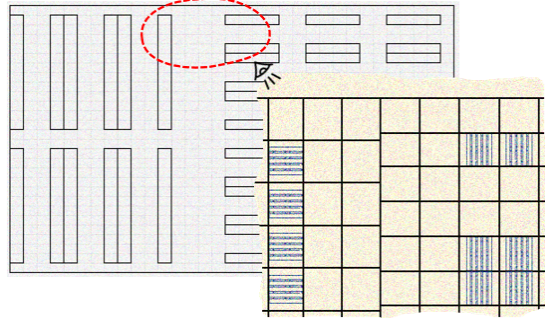
Both the inbound and the outbound points are part of a picking tour, so they can be modelled as path cells. Besides, each path cell is linked to another subset of path cells, either by moving up, down, left, or right. Pairwise combinations - a fancy name for diagonals - were not considered in this model.

Hence, in order to fully define any given picking area, the user must perform all the following steps:

- Input the warehouse dimensions to model the picking area as a grid;
- Create the racks (as in section 3.2.3) and drag them onto the grid like ships on the Battleship game - rack cells are the non-path cells whilst the remaining ones are the path cells;
- Assign which are the impossible cells (e.g. structural pillars) to fully define all the non-path subsets;
- Allocate the inbound and outbound locations to a single path cell. There is only one starting point and one finishing point but they might be non-coincident;
- Define the routing constraints of each path cell to finalize the warehouse design.

Each rack position has one active store cell and, depending on the rack, it can also comprise buffer store cells and at least one retrieving path cell. Even not knowing the store allocation policy in use, it is possible to differentiate between picking path cells and routing path cells. In fact, one can deduct all the potential places from which stores can be reached, or, in other words, all the picking path cells. The model comprehends a function that enables this distinction after reading the user inputs.

Nevertheless, a final assumption has to be made. Considering that racks can assume multiple orientations, the cell length must be equal to the cell width. Otherwise, this would lead to a meshed grid. Figure 5.3 better illustrates this rationale.

Figure 5.3: Grid convergence problem if $l_c \neq w_c$

5.3 Graph Theory

As mentioned before, each path cell, $x \in C_p$, is linked to at least another path cell, either by moving up, down, left or right. However, there might be forbidden movements and, in other cases, one-way links. Just like in traffic, it might be impossible to go backwards. So, once a path cell is reached, the picker's next movement is constrained to that cell's links. This can easily be mapped onto an oriented graph in which path cells, and their links, compose that graph's nodes, and arcs, respectively.

On top of that, there is only one entry path cell - the inbound point - and only one exit path cell - the outbound point. Similarly to what happens in a maze, the picker enters the picking area and must find his way out, after completing a task with N picks. Hence, graphs can be used to uncover the shortest path to perform each task.

Figure 5.4 illustrates the path constraints of a small maze and their representation under a graph's notation. For that, consider the following: (1) each path cell has an array of possible movements in order to represent their links, with letter 'W' allowing the picker to go up, letter 'A' to go left, 'S' to go down, and 'D' to go right; (2) non path-cells are coloured in black and cannot be traversed, thus not being represented in the graph; (3) The inbound point is coloured in green whilst the outbound point is coloured in orange.

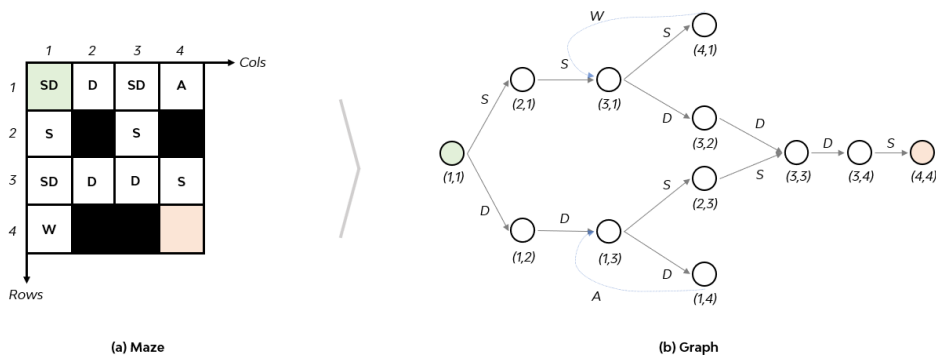


Figure 5.4: Modelling path cells as an oriented graph

In fact, every picking area can be seen as a maze. The only thing that differs is how the non-path cells are disposed alongside the grid. The maze was only used for simplicity purposes.

5.4 Mathematical Model

This section will focus on how to derive a new analytical model that can be applied to any possible picking area. Instead of simulating zones independently, this enhanced version aims to output results without having to make layout partitioning. In order to do this, key adjustments are needed to generalize the expressions attained for the previous model.

First and foremost, picking areas with mixed zones shall be addressed at the cell granularity level. Still, the objective remains the same - how to estimate the expected picking tour length, $E[L]$, for any picking task of size N .

Since aisles are no longer the elementary unit, equation 4.2 is now meaningless. Even replacing the words, it makes no sense to estimate an expected travel distance crossing-cells. This is true because a picker is always at a path cell, $x \in C_p$. Therefore, the expected travel distance, $E[L]$, can simply be obtained by summing up all the expected travel distances within each cell, $E[L_x]$. Equation 5.1 is used to describe this step.

$$E[L] = \sum_{x \in C_p} E[L_x] \quad (5.1)$$

In comparison, the expected contribution of a path cell, to the total picking tour length, can easily be obtained by modifying equation 4.4. As a matter of fact, $E[L_x]$ will depend on the same two factors. These are: (1) the distance needed to fully traverse the path cell, d_x , and (2) the probability of traversing it, $p(S_x)$. Hence, equation 5.2 can be defined as follows:

$$E[L_x] = d_x \cdot p(S_x), \quad \forall x \in C_p \quad (5.2)$$

There are several ways to fully traverse a path cell. This can be done by going either: (a) vertically (from top to bottom or otherwise), (b) horizontally (from left to right or otherwise), or (c) any combination of the previous two that possibly describes a 90° angle. Figure 5.5 helps understand all these cases.

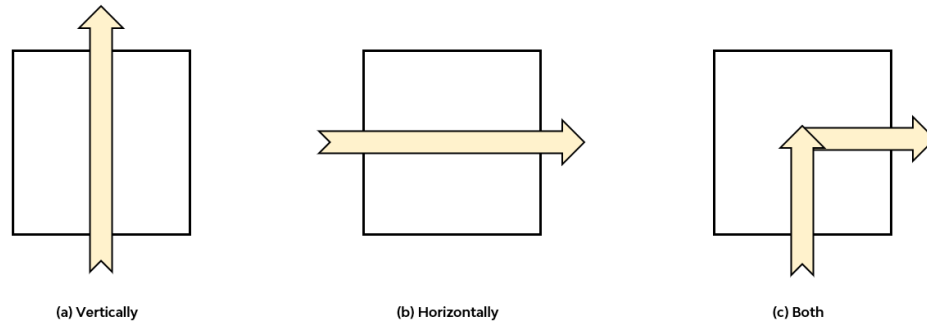


Figure 5.5: All possible movements to fully traverse a path cell

In fact, it was assumed that $l_c = w_c$ in order to avoid grid convergence problems. So, the distance needed to fully traverse a path cell is the same in all possible scenarios. Equation 5.3 provides a general expression for parameter d_x :

$$d_x = l_c, \quad \forall x \in C_p \quad (5.3)$$

Similarly to what happens for the zone modelling, there are only two reasons from which a path cell needs to be reached: (1) at least one item has to be delivered on that cell or (2) no items have to be delivered on that cell despite being part of the shortest path to complete the task. These events are mutually exclusive and can be denoted, in analogy, as E_x and V_x respectively.

Therefore, $p(S_x)$ can be expressed, for each path cell, as the arithmetic sum of the individual probabilities $p(E_x)$ and $p(V_x)$. However, only picking path cells, $x \in C_{pp}$, have items to be delivered on. So, this first component is going to be null for all the routing path cells, $x \in C_{pn}$. In fact, routing cells are solely used to traverse the warehouse whilst picking cells are routing cells that can also be used to pick items. Equation 5.4 analytically depicts both cases.

$$p(S_x) = \begin{cases} p(E_x) + p(V_x), & \text{if } x \in C_{pp} \\ p(V_x), & \text{if } x \in C_{pn} \end{cases} \quad (5.4)$$

As a matter of fact, this expression could be written in a single branch just like in equation 4.6, since $C_p = C_{pp} \cup C_{pn}$.

In order to continue, it is now crucial to know, *a priori*, the probability of a pick addressing a path cell, not an aisle. Hence, $p(P_x)$ can easily be derived by adapting equation 4.7. Moreover, picking path cells are the only ones to have picks. So, applying the same rationale this value will be null for every routing path cell. Nevertheless, equation 5.5 provides a general expression that works for both cases.

$$p(P_x) = \frac{\text{Number of Picks in Path Cell } x}{\text{Number of Total Picks}}, \quad \forall x \in C_p \quad (5.5)$$

Next in line is the probability of at least one of the N picks addressing the path cell $x \in C_p$. As mentioned before, this calculus only makes sense for the picking path cells. Still, it is possible to attain an expression general enough to cover both sets of path cells. This can be done by simply changing the notation used back in equation 4.8. Therefore, $p(E_x)$ can be defined as follows:

$$p(E_x) = 1 - [1 - p(P_x)]^N, \quad \forall x \in C_p \quad (5.6)$$

Things start to become messier from this point onwards. This is because the total number of cases that had to be addressed for calculating $p(V_x)$ would explode in quantity. In fact, the number of path cells is way bigger than the number of aisles and, on top of that, the mixed model does not consider either corridors or levels, since it encompasses more than one zone. Therefore, the expressions derived for each parcel in equation 4.10 are no longer valid.

Still, some of the notation used back in section 4.3 is still valid, with due modifications. Having said that, assume m as the last path cell with a picking task, x' as the next path cell with a picking task, and x as the path cell in analysis to calculate the probability $p(V_x)$. In case there is not a direct move, the picker will certainly traverse all the path cells, x , that comprise the shortest path between m and x' .

So, the first step to attain $p(V_x)$ remains the same. Equation 5.7 is used to calculate that probability, $p(E_m, E_{x'})$, of going directly from picking path cell, m , to picking path cell, x' , without having to perform a picking task in between. Hence, this only makes sense to compute, for every ordered pair (m, x') of picking path cells.

$$p(E_m, E_{x'}) = \begin{cases} \left[1 - \sum_{j=m+1}^{x'-1} p(P_j) \right]^N, & \text{if } m_i < x'_i, \quad \forall m, x' \in C_{pp} \\ 0 & \text{otherwise} \end{cases} \quad (5.7)$$

That is why order matters. Therefore, every picking path cell must have a picking index, i . Like it was mentioned before, this can be obtained by applying Dijkstra (1959)'s algorithm to all points of interest. In fact, this will output the shortest path to visit all the picking path cells, starting at the inbound point and finishing at the outbound point. Hence, these points shall also be numbered with the indexes $i = 0$ and $i = |C_{pp}| + 1$, respectively.

These indexes are then used to assess the validity of a given route, $m \rightarrow x'$. If, $m_i < x'_i$, the probability of a picker doing such route, in a task with N picks, exists and can be obtained by simply adjusting equation 4.12.

$$p(m \rightarrow x') = p(E_m) \cdot p(E_m, E_{x'}) \cdot p(E_{x'}), \quad \forall m, x' \in C_{pp} \quad (5.8)$$

However, this probability encompasses all the possible routes to arrive in x' when departing from m . This was not the case for the previous model so, in order to generalize, one must consider only the shortest routes that connect two picking locations.

For the model by zones, this could be achieved by clustering routes into some few scenarios as the likings of Table 4.1. For the mixed model, however, the same does not stand. So, each ordered pair (m, x') must be analysed individually.

5.4.1 Breadth First Search Heuristic

Bearing this in mind, a function was created to output both the distance and the shortest path of a given route $m \rightarrow x'$. This was done by applying graph theory knowledge under a breadth first search greedy heuristic. To better understand this, consider Figure 5.6.

The picker always starts his tasks at the inbound node. In this case, the inbound node is represented by the path cell in position $(1, 1)$. Then, the picker can only go to an adjacent path cell, if there is a link that enables such movement. For the case being, this can either be done by going downwards, to position $(2, 1)$, or by going rightwards, to position $(1, 2)$. Hence, these two nodes represent a layer with all the possible destinations, after one move inside the picking area.

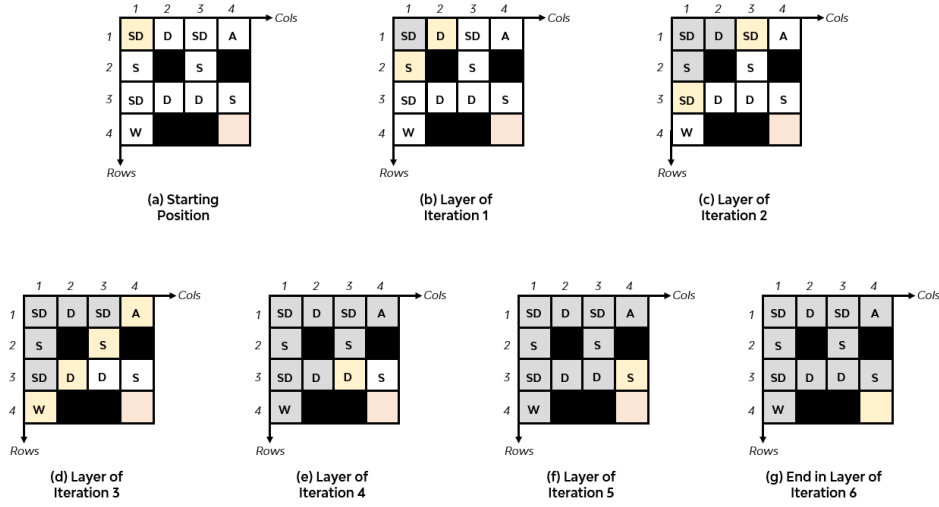


Figure 5.6: Illustration of the breadth first search algorithm

The goal of this algorithm is to uncover the l -th layer that stores the targeted node. Following the same rationale, the picker will have to perform l moves, inside the picking area, in order to reach his destination. This is the least amount of moves between an inbound point and an outbound point. Nevertheless, these points can further be changed to any ordered pair, (m, x') , of picking path cells. Since each move can be seen as a cell to traverse, the minimum distance to perform a given route can be obtained, in meters, by equation 5.9.

$$dist(m \rightarrow x') = d_x \cdot min_{layer}, \quad \forall m, x' \in C_{pp} \quad (5.9)$$

The greedy part is that nodes are addressed by their order of creation. In fact, layer $l+1$ is created after addressing all the elements in layer l . For the example of the first layer, the first node to be created is the path cell in position (2,1). This is because the direction possibilities of the origin nodes are always read in the same way: up-left-down-right.

Moreover, each node has to be assigned to only one layer. So, when analysing a node in layer l , one must flag all its possible destinations as nodes already assigned to layer $l+1$. This can better be seen in iteration 3. Both the nodes in positions (2,3) and (3,2) can reach the path cell in position (3,3). However, only the first node addresses this possibility to avoid the creation of the same node for the fourth layer.

Under this approach, each node has only one way in. This is useful for back-tracing the shortest path between the inbound and the outbound nodes. Nevertheless, there can be multiple shortest paths connecting the same two nodes. The only thing that is certain is the distance needed to perform them. Hence, a distance matrix, $\mathbf{D}$, can be derived as follows:

$$\mathbf{D}_{mx'} = \begin{cases} dist(m \rightarrow x'), & \text{if } m_i < x'_i \\ 0, & \text{otherwise} \end{cases}, \quad \forall m, x' \in C_{pp} \quad (5.10)$$

5.4.2 Handling Multiple Shortest Paths

It is possible to have multiple paths for a given route, $m \rightarrow x'$. However, it was assumed that the picker always choose the shortest path available. In case there are multiple shortest paths, one is randomly chosen.

As previously mentioned, each shortest path between a picking path cell, m , and a picking path cell, x' , will comprise a unique set of path cells, x . Besides, each path cell, x , is not exclusive of a single shortest path. To better understand this, consider once again the maze example.

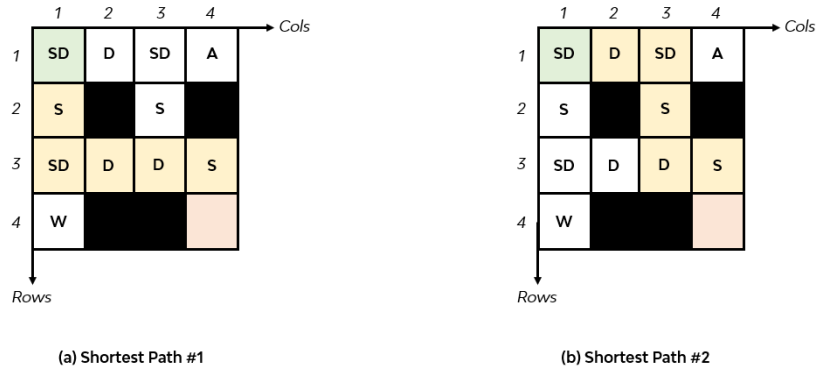


Figure 5.7: Construction of all the shortest paths to traverse the maze

Mazes are used to illustrate the rationale behind routing. They can represent the totality (or even portions) of a picking area. Having said that, let m be the path cell in position (1,1), and x' be the path cell in position (4,4), in such a way that $m_i < x'_i$. Bearing this in mind, Figure 5.7 depicts all the possibilities to reach x' , when departing from m , in the least amount of moves. All the other paths to perform this route are considered to be unfeasible.

For the case being, there are two paths to perform the route $m \rightarrow x'$, but only one can be taken. Assuming that this selection is random, the picker will opt for the first path half of the times. So, the probability of choosing one of the two paths is, logically, half of the probability of performing this route.

Moreover, each path is mutually exclusive. Hence, every path cell, x , can only be traversed once in a given route. However, the same path cell can be part of multiple paths for the same route. For example, position (3,3) is always traversed whatever the path chosen. So, the probability of traversing a path cell, in a route, is given by the probability of choosing a path, of that route, with that cell.

Generalizing, it is possible to obtain $p(V_x)$, for a given route $m \rightarrow x'$, in which B out of its P shortest paths contain the path cell, x .

$$p(V_x) \Big|_{m \rightarrow x'} = \frac{B}{P} \cdot p(m \rightarrow x'), \quad \forall x \in C_p \quad (5.11)$$

Finally, in order to fully define the model, equation 5.11 must be extended to all possible routes. In this way, $p(V_x)$ is accounted for every route in which a path cell, x , could possibly partake. This can be mathematically written as of equation 5.12.

$$p(V_x) = \sum_{m \in C_{pp}} \sum_{\substack{x' \in C_{pp}: \\ m_i < x'_i}} p(V_x) \Big|_{m \rightarrow x'}, \quad \forall x \in C_p \quad (5.12)$$

5.4.3 Relaxation

With great detail, comes greater complexity. In fact, there is always a trade-off when increasing a model's granularity level. On the one hand, the results are more efficient since processes are closer to reality, but on the other hand, the computational cost associated with its resolution augments drastically.

As mentioned before, the first model takes an aisle as the elementary unit. So, all the expressions will have a computational time that depends on the number of aisles, A . In comparison, for the mixed model, this value will depend on the number of path cells, $x \in C_p$. However, the number of path cells is way bigger than the number of aisles.

Figure 5.2 helps understanding this statement. Each picking position in an aisle must now be reached by a picking path cell and, for the case being, 26 aisles were quickly morphed into nearly 400 picking path cells. Moreover, this set C_p also comprises several other hundreds of path cells used solely for routing purposes. Besides, routing path cells are usually more numerous than picking path cells, so the same route will certainly have multiple ways of being completed.

However, it is impossible to attain - in feasible times - all the shortest paths between two picking path cells. This is because of the way paths are created. For instance, imagine a route that can be done in 30 moves. If the first cell is linked to at least more than one cell there can be at least two possible paths after one move. Considering again Figure 5.2, each path cell has, indeed, more than one link. So, after 30 iterations there would be more than a billion paths to analyse.

All these paths must be saved in memory to see how many of them reached the targeted node after 30 iterations. Furthermore, routes done in 30 moves are considered to be small for the majority of picking areas. Hence, a relaxation of the mixed model was needed.

There can be multiple shortest paths for the same route. However, the distance needed to perform them is always the same. These distances can easily be obtained by computing the distance matrix, $\mathbf{D}$. Moreover, the probabilities of doing such routes are known and can simply be calculated as of equation 5.8. Therefore, the expected travel distance, $E[L]$, can be reduced to the summation of the expected travel distances inside each route. All in all, equation 5.13 depicts another way to reach the same goal.

$$E[L] = \sum_{m \in C_{pp}} \sum_{\substack{x' \in C_{pp}: \\ m_i < x'_i}} \mathbf{D}_{mx'} \cdot p(m \rightarrow x') \quad (5.13)$$

The only catch, however, is that $p(V_x)$ and, consequently, $p(S_x)$ do not exist anymore. This was interesting to quickly depict heat maps of cell utilization. Nevertheless, the Second 'S' - Discrete Event Simulation - will also allow that type of insights.

5.5 Synthesis

Despite being applicable to the majority of cases, the first model had its flaws. So, a mixed model was derived in order to overcome those issues. Having said that, this chapter firstly introduced all the key adjustments that were needed, in theory, to generalize the model by zones.

However, some simplifications had to be done to obtain results in a feasible time. In fact, the main goal of this work is to quickly simulate and compare multiple layout scenarios for, in the future, optimize the most promising ones. This can now be done for every possible picking area.

Regardless of the internal layout design, there is still a need for some other critical inputs. Bearing this in mind, the expected picking tour length will depend on: (a) the task size, N , and (b) the probabilities, $\mathbf{p}_x$, of a pick addressing a path cell x . This is similar, in all aspects, to what happened for the first model.

Since order matters, the derived expressions will also depend on the inbound and outbound point locations. However, they can be modelled as path cells and, therefore, part of the internal layout design component. All in all, for a given picking area:

$$E[L] = g(N, \mathbf{p}_x) \quad (5.14)$$

Chapter 6

Case Study And Results

The primary goal of this work is to create a model that is both adaptable and applicable to any possible picking area. By leveraging a probabilistic simulation approach, this model provides a means to estimate the performance of a given picking area, facilitating quick comparisons of multiple layout solutions.

In order to assess the quality of the proposed models, it is essential to conduct a thorough examination of a real-life scenario. Bearing this in mind, this chapter is structured to present a comprehensive case study that demonstrates the usability and effectiveness of the developed models. This was done in three stages.

Part A introduces the selected study subject, which serves as a practical test bed on that matter. Part B delves into providing a detailed description of the warehouse itself, shedding light on its unique characteristics and operational constraints. Finally, Part C presents the outcomes obtained through the statistical approach employed.

6.1 Study Subject

6.1.1 The Company

SONAE is a renowned Portuguese conglomerate with a diversified portfolio of businesses. Founded in 1959, SONAE has evolved into a multinational corporation, operating across various industries, namely retail, real estate, telecommunications, and investment management.

In the context of the retail market, this company has quickly become a powerhouse. With a commitment to quality, affordability, and an expansive network of stores, SONAE has successfully captured a significant market share in the Portuguese retail landscape. This was mainly done through its flagship brand, Continente, by offering a comprehensive range of products all under the same roof.

Despite the vast array of goods under SONAE's umbrella, this work is centred around a specific warehouse dedicated to handling household products. To be more precise, this work only addresses the PBL Flow Type, so storage hardly plays a role. Nevertheless, the following section will deep dive into the characteristics and operational considerations of the warehouse at hand.

6.1.2 The Warehouse

To better understand the dynamics and intricacies of this warehouse, it is essential to dissect its layout into distinct building blocks. As previously mentioned, each building block represents a specific warehouse operation. Figure 6.1 helps visualizing this split, emphasizing the boundaries and connections between those key areas: the reception area, the marshalling area, and the picking area. The latter is further differentiated regarding the picking policy in use (PBS or PBL).

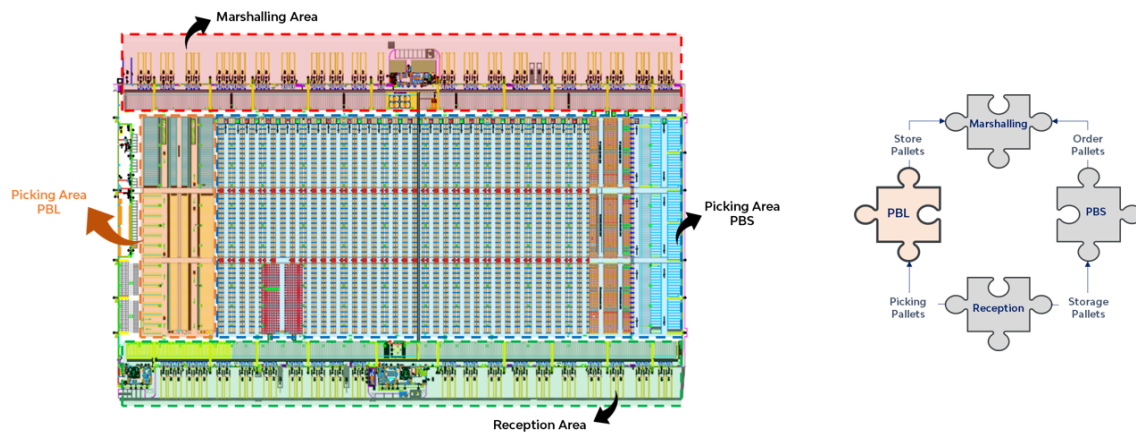


Figure 6.1: Division of the warehouse layout on its building blocks

The scope of this case study specifically focuses on the PBL block and on the picking tasks that occur within this block. In fact, each picking task corresponds to a picking pallet and the preparation of those entities is assumed to be given and based on historical data.

6.1.3 Warehouse Data

In order to develop accurate models, it is crucial to have access to reliable and relevant data. Fortunately, the company has provided a valuable dataset encompassing various aspects of the warehouse's operations.

Besides, the process of cleaning the retrieved data is an essential step to ensure the reliability of future analysis. While the data provided offers a comprehensive snapshot of the day-to-day activities, it is imperative to address potential inconsistencies, outliers, missing values, and other data quality issues that may affect the validity of the results.

The ultimate goal of data cleaning is to obtain representative picking tasks. However, each observation in this dataset corresponds to a box. In fact, picking tasks are no more than a set of boxes and, just like Figure 1.1, it is possible to trace the flow of boxes through the reception, picking, and marshalling operations. Therefore, a box can be viewed as a pick.

The cleaned data revealed a total of 18.346 picking tasks (the equivalent to 518.278 boxes) distributed over a span of nearly three months. These findings highlight the scale and duration of the study, emphasizing the extensive dataset used to assess the warehouse's picking area. Figure 6.2 highlights these results.

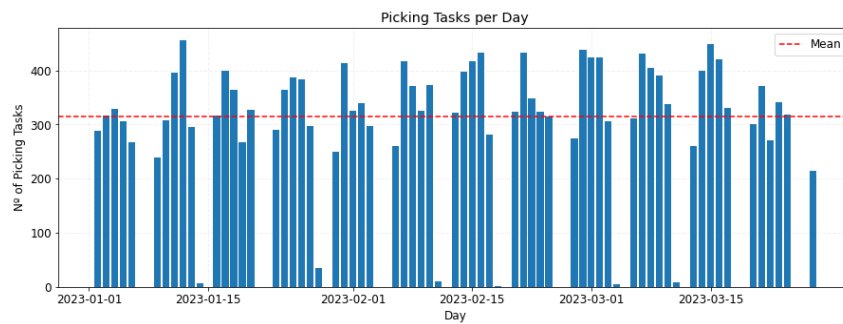


Figure 6.2: Daily distribution of the PBL block's picking tasks

On average there are 314 picking tasks per workday. However, this value is certainly biased by the existence of extra labour during Saturdays. In fact, this anomaly can be attributed to the end-shift hour of the pickers as deductible by Figure 6.3.

That graph also grasps the sequential flow of boxes throughout the warehouse, but this study only comprises the preparation phase - or in other words: the picking tasks.

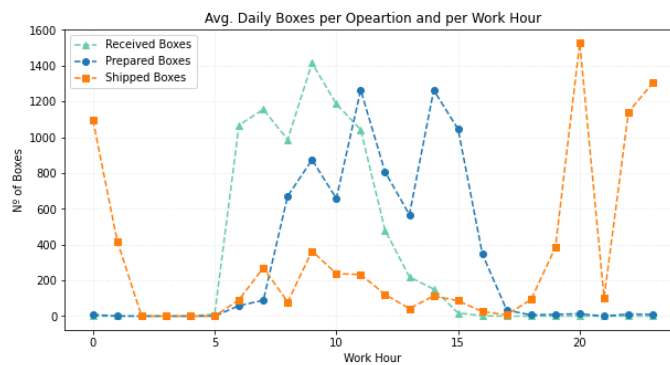


Figure 6.3: Average daily boxes per operation and per work hour

Notably, these tasks can be completed shortly after midnight, theoretically falling under the previous day's jurisdiction. However, in practice, the Warehouse Management System automatically records the precise hour, minute, and second of each pick within a picking task.

Since the warehouse remains closed on weekends, all Saturday's chores are, in fact, part of Friday's late shifts. Table 6.1 shows that data and makes this adjustment.

Table 6.1: Expected daily picking tasks

Weekday	Avg. Picking Tasks	Avg. Boxes	Avg. Ratio
Mon	280.77	7257.15	25.85
Tue	390.83	10665.83	27.29
Wed	374.08	9036.08	24.16
Thu	367.67	8539.83	23.23
Fri	312.33	7078.92	22.66
Sat	5.42	7.25	1.34
All Days	314.24	7735.49	24.62
Real Days	345.15	8496.36	24.62

6.2 The PBL Block

The derived models estimate the performance of a given picking area in distance terms. In fact, travelling is often the dominant activity of a picker's time. So, minimizing the travelled distance is a crucial objective in order-picking.

Warehouse managers commonly encounter several decision problems when trying to enhance picking efficiency. These problems include: (a) the layout design, (b) the picking policy in use, (c) the storage assignment rules, (d) the existence of zoning, and (e) the routing methods.

All those tactical and operational variables are interconnected. So, in this section, the spotlight will be on the following topics: (1) identification of the variables that impose hard constraints on the picking area, (2) examination of the variables that can be modified or changed, and (3) division of the picking layout to implement both the zonal and the mixed models.

6.2.1 Hard Constraints

First and foremost, the primary focus of this work is on the PBL block. So, as the name implies, the scope of this thesis exclusively addresses this specific picking policy. Nonetheless, it is worth noting that the grid methodology presented in this study can be readily extended to also encompass the PBS flow type.

Moreover, the picking area at hand operates on a low-level picker-to-parts system. Picker-to-parts because the picker has to walk or drive along the aisles to pick the items. Low-level because the picking positions are ground-floor marked locations.

A final remark is made to the warehouse pallets since they all share identical dimensions and possess a maximum capacity, whether determined by volume or weight.

6.2.2 Layout Variables

The hard constraints of this problem are the dimensions of the picking area and the nature of the picking tasks. In fact, it is impossible to change the picking policy in use. Hence, changes to the physical disposition of the picking positions were outside the scope of analysis. In other words, the picking area operates always on a low-level picker-to-parts system under a PBL flow type.

However, there are several other aspects that can be considered when estimating layout performances. It is important to note that, for the purpose of this case study, the following presentation will primarily focus on the current "AS-IS" situation. Nonetheless, brief notes will be provided regarding potential modifications that could be made in each one of those areas.

First things first, the PBL Block can be further deconstructed as a set of zones. As a result, the internal layout design concerns: (1) the determination of the number of zones and (2) the number, length, width, and orientation of the aisles inside each zone. Figure 6.4 provides a visual representation of the actual layout, showcasing two zones: one vertical and one horizontal. Bear in mind that this figure is rotated 90° clockwise for simplicity purposes.

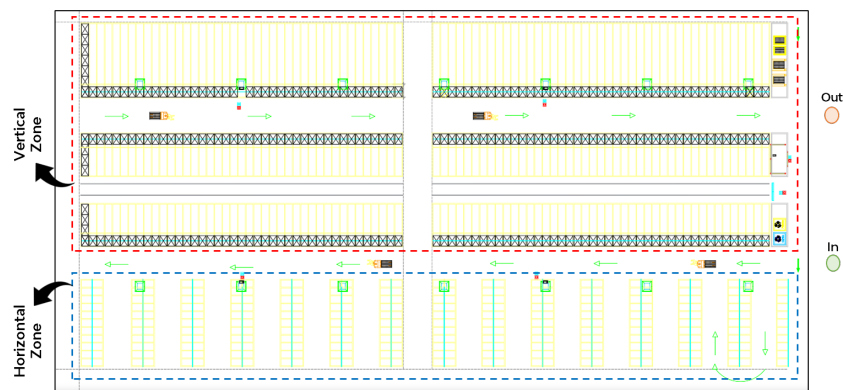


Figure 6.4: 2D top view of the PBL block's current layout

The vertical zone consists of two levels with two aisles each. The first level contains aisles with 45 picking positions, while the second one comprises aisles with 41 picking spots. It is noteworthy that the racks in these aisles also feature several buffer positions.

On the other hand, the horizontal zone is built upon a single level with 14 aisles, each one of them equally equipped with 11 picking locations. In comparison, the racks in these aisles only consist of a single active position, thus indicating a different configuration. The absence of buffer positions can be attributed to the relatively low number of picks addressing this zone. This characteristic is clearly demonstrated in Figure 6.5 by making use of a heat map.

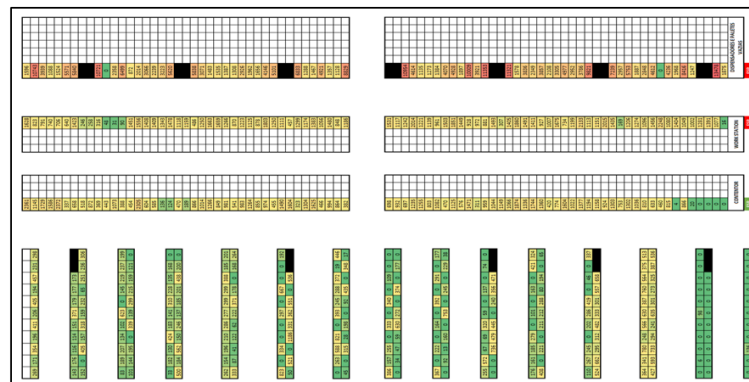


Figure 6.5: Heat map of the number of picks per picking position

The intensity of colour coding is used to depict the frequency of visits. Typically, warmer colours, like red, are used to represent picking positions that have been visited more times, indicating high activity levels. On the other hand, colder colours, like green, are employed to represent picking positions that have been visited less frequently.

However, the heat map is heavily influenced by the store allocation policy in use. As previously mentioned, this involves the assignment of specific stores to warehouse locations. Since the visits to picking locations depend on the store they are associated with, different store allocation policies will result in distinct heat map patterns. By adjusting these rules, warehouse managers can observe changes in the distribution of picking tasks.

Finally, take into consideration that, pickers always follow the shortest path between any two locations comprised in a picking task. However, they may encounter routing constraints such as one-way aisles. This can better be seen in Figure 6.4 as indicated by the orientation of the green arrows. Nevertheless, this can easily be modified to accommodate different routing scenarios.

6.2.3 Zoning: Variable or Constraint?

Moving on to the last problem of layout design, it is important to differentiate between two types of zoning: geometrical zoning and managerial zoning. Geometrical zoning refers to the virtual decomposition of the picking area based on its aisles' orientation. On the other hand, managerial zoning is a strategic decision. This is made at the corporate level to actually divide the picking area by assigning each picker to a single zone.

In the case of the warehouse under study, managerial zoning (as defined in the literature) is not practised. Therefore, the picking area must be considered as a whole. However, most of the picking areas only comprise a single zone. This motivated the division of the analysis into two major phases: (1) simulating layout zones and (2) simulating the entire layout with a single mixed zone.

6.3 Results

In this section, the focus shifts towards presenting the results obtained from the developed models. These two distinct approaches allow for a comprehensive evaluation of the impact of zoning on the overall performance of the current picking area.

However, it is important to emphasize that the methodology presented in this thesis extends beyond the scope of this particular case study. It provides a framework that can be applied to any possible picking area, accommodating different combinations of factors involved in layout design.

6.3.1 Zone Model

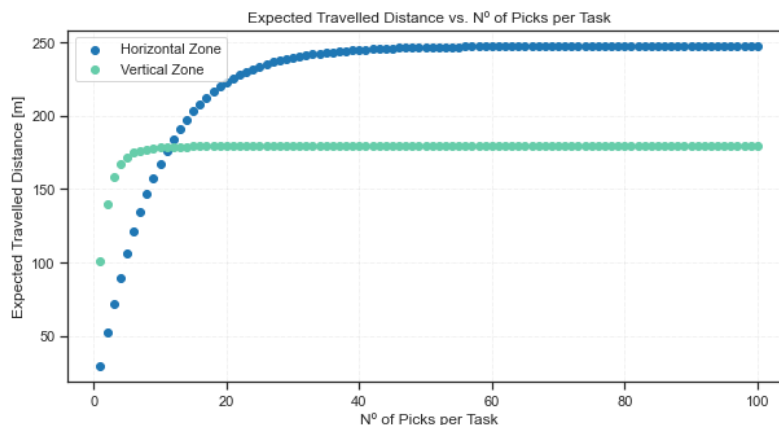


Figure 6.6: Expected distance for a picking task of size N inside each zone

The first model evaluates the distance that a picker would need to travel, within a specific zone, in order to complete a picking task of a given size, inside that zone. This is done without prior knowledge of the task details (e.g.: locations to visit and picks per location).

Since it addresses one zone at a time, this model has no grid-related issues to consider. As a result, the actual cell dimensions can be used ($l_c = 0.9\text{ m}$; $w_c = 1.2\text{ m}$). However, there is not the need to deep dive into the cell detail. It has been established, in chapter 4, that single zones can be accurately modelled using the aisle granularity level. Hence, each pick addresses an aisle rather than addressing the exact picking position inside that aisle.

Table 6.1 shows that each picking task is expected to have around 25 picks. So, it can be concluded, from Figure 6.6, that on average, the picker needs to traverse 179.50 m for a picking task within the vertical zone and 233.50 m for a picking task within the horizontal zone. In fact, for a task of this magnitude, the picker has to traverse almost the entire vertical area.

The convergence of distances within this area occurs at a faster rate. This can be attributed to the lower probability of a picking task involving picks from the horizontal aisles. Moreover, the presence of fewer vertical aisles contributes even more to this phenomenon. Such observations are more clear in Table 6.2.

Table 6.2: Probability of a pick addressing an aisle

Horizontal Zone							
Aisle	Visits	Prob. PBL	Prob. Zone	Aisle	Visits	Prob. PBL	Prob. Zone
H1	136	0.03%	0.22%	V1	37662	7.27%	8.25%
H2	4849	0.94%	7.86%	V2	37818	7.30%	8.28%
H3	9652	1.86%	15.65%	V3	165278	31.89%	36.20%
H4	3591	0.69%	5.82%	V4	215858	41.65%	47.27%
H5	4537	0.88%	7.36%	Total	456616	88.10%	100.00%
H6	3083	0.59%	5.00%				
H7	3206	0.62%	5.20%				
H8	3058	0.59%	4.96%				
H9	6972	1.35%	11.31%				
H10	4161	0.80%	6.75%				
H11	5560	1.07%	9.02%				
H12	3646	0.70%	5.91%				
H13	4419	0.85%	7.17%				
H14	4792	0.92%	7.77%				
Total	61662	11.90%	100.00%				

In order to determine the probability of visiting an aisle, it is necessary to sum the probabilities of visiting the stores assigned to that particular aisle. This information relies on the store orders and the store allocation policy in use. Notably, while the orders remain constant, the allocation policy may be subject to changes.

These probabilities play a crucial role in the construction of both models as they form the basis for all the calculations and expressions involved. In this first model, which operates at the aisle granularity level, the focus is on the probabilities of visiting an aisle. Conversely, for the second model, the emphasis is on the individual store probabilities as this model operates at the cell granularity level.

6.3.2 Mixed Model

Unlike the first model, which assesses zones independently, the mixed model considers the entirety of the picking area. In addition, the elementary unit is now the cell. So, each picking position in an aisle will require a corresponding picking path cell.

As a result, the computational time for the expressions in the mixed model is going to depend on the number of path cells. However, this value is significantly larger than the number of aisles. To be more precise, 18 aisles (14 horizontal + 4 vertical) gave birth to 1544 path cells, from which only 558 are used to perform picking tasks.



Figure 6.7: Modelling the PBL block's current layout as a grid

Figure 6.7 serves as an illustrative example, depicting the different types of cells involved. The non-path cells, that make up the racks of the former aisles, are distinguished by their blue colour. As for the path cells, there are three distinct sets. The picking ones, which are represented by a heatmap; The routing ones, that were left in blank; and the forbidden ones, highlighted in grey.

Despite this level of detail, approximations had to be made. As explained in section 5.2, the cell length was set equal to the cell width.

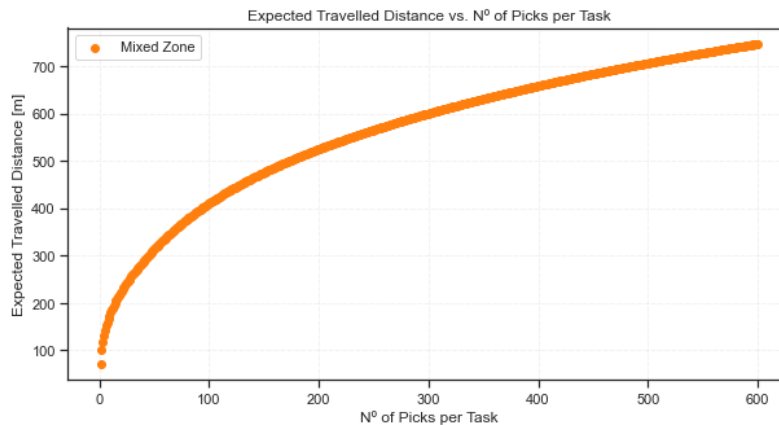


Figure 6.8: Expected distance for a picking task of size N inside the mixed zone

Considering that $l_c = w_c = 0.9 \text{ m}$, the expected distance, for a picking task of size N , was estimated. The results are presented in Figure 6.8.

After examining the graph, it is evident that this model does not converge as expected. This seems unreasonable since, for a significant number of picks, it is highly likely that the picking task would need to traverse the entire picking area.

However, it is important to notice that there are a total of 558 picking positions, and each position can have multiple picks associated with it. Therefore, even for a large value of N , such as $N = 600$, there is no guarantee that all these positions will be visited.

Drawing an analogy to the zone model, a similar pattern can be observed when the number of picks nearly surpasses the number of aisles. In fact, the entire zone is expected to be totally covered only when $N > 10 \cdot A$.

This is due to the presence of cells (and aisles) with low probabilities of being visited. Nevertheless, a Monte Carlo simulation was conducted to further investigate any anomaly.

6.3.3 Monte Carlo Simulation

Monte Carlo simulation is a computational technique that involves running numerous iterations of a process using random sampling. This was employed to assess the expected behaviour of the mixed model. By generating multiple random picking tasks, the simulation provided insights into the variability and distribution of the distances travelled by the picker. This was done 200 times for each N , which allowed to capture a comprehensive range of possible outcomes.

Each picking task has N picks and S stores to visit. However, each store can have more than one pick. So, the way stores were assigned was random, but weighted by the probabilities of visiting that store. Hence, a store with many visits is more likely to appear (and more than once) when generating a picking task. Figure 6.9 presents these results.

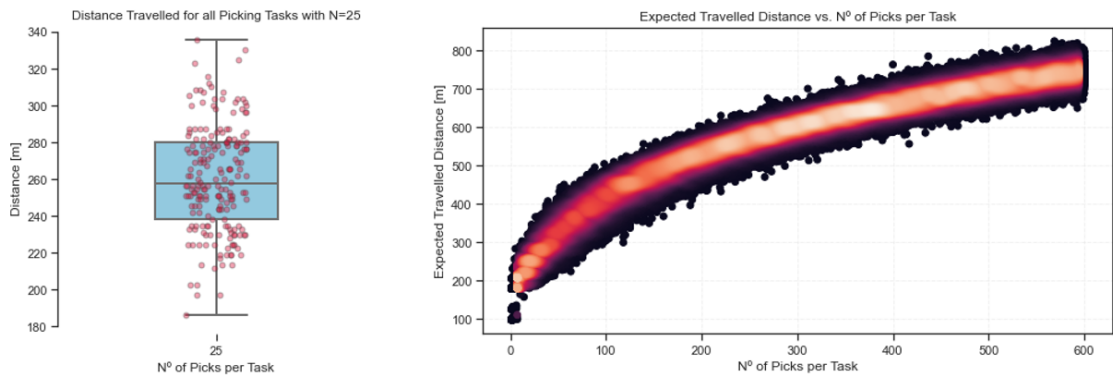


Figure 6.9: Distance distribution of randomly generated tasks of size N

This simulation allowed the evaluation of statistical properties such as mean, standard deviation, and distribution of the distances travelled. By comparing these results to those obtained from the mixed model, it is possible to assess similar patterns, regarding the expected distance for a picking task of size N . Hence, the mixed model appears to be well-constructed.

According to the mixed model, the picker needs to traverse 242.74 m in order to complete an average picking task. However, this value is slightly bigger when considering the 200 iterations of the Monte Carlo simulation. In this case, the average distance travelled is 260.24 m, as shown by the boxplot.

This results in a difference of only -6.73%, indicating that the mixed model marginally underestimates the distance for tasks with 25 picks. To further explore these disparities for different task sizes, Figure 6.10 presents a graph illustrating the gaps between the two models.

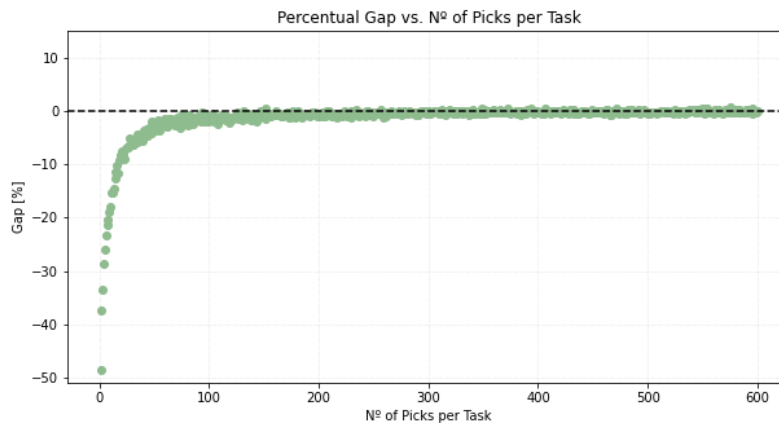


Figure 6.10: Gap between the mixed model and the Monte Carlo simulation

The mixed model exhibits significant deviations primarily for tasks with a low number of picks, but even in these cases, they remain relatively small. For example, for tasks with 10 picks, the mixed model underestimates the distance only by, approximately, 15%. However, tasks with smaller sizes are not uncommon in the PBL Block as they account for roughly 26% of the data.

These tasks usually involve unfinished work from the previous day or occasional deliveries of late supplies. Hence, these tasks do not reflect the typical workload of a picking area. However, the mixed model can yield inaccurate results for these "tasks". Figure 6.11 offers further insights into this phenomenon.

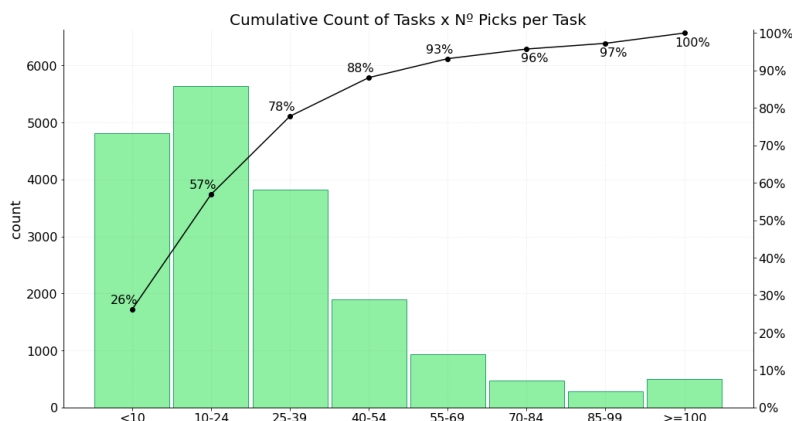


Figure 6.11: Distribution of the number of tasks for increasing task sizes

The pattern observed shows similarities with the distributions presented back in Figures 6.8 and 6.9. This suggests that the mixed model, along with the Monte Carlo simulation, are correlated with the picking tasks. This alignment may indicate that both these approaches serve as a reliable estimator for the distances to travel.

To gain further insights, it would be beneficial to examine this evolution for: (i) tasks limited to the horizontal zone, (ii) tasks confined to the vertical zone, and (iii) tasks that encompass both zones.

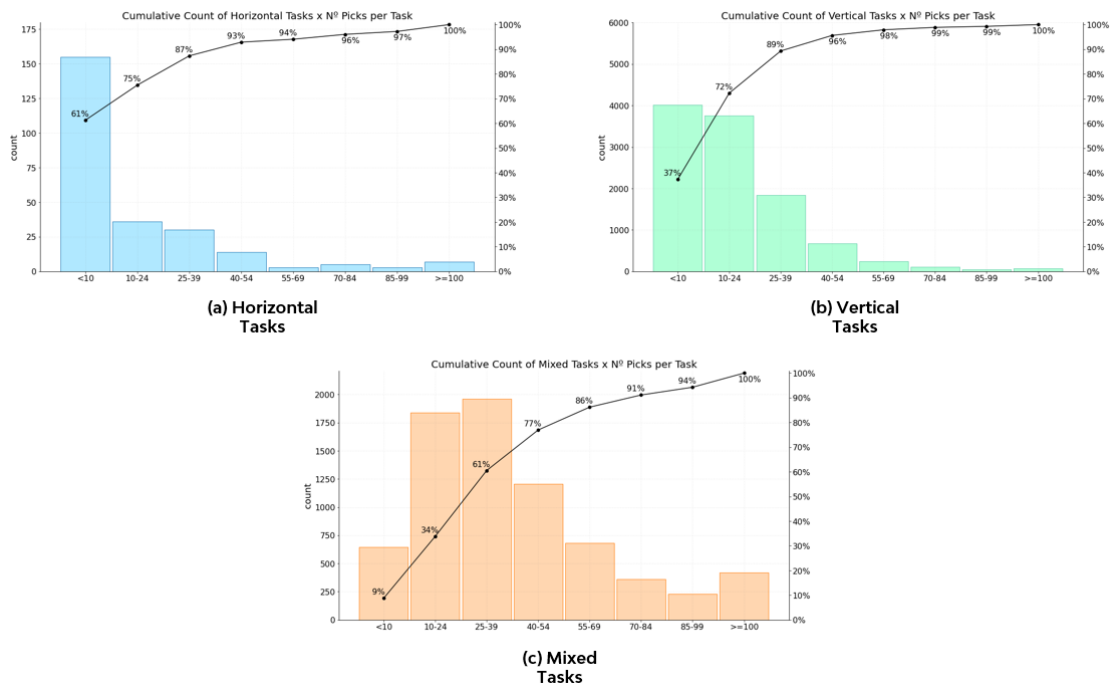


Figure 6.12: Distribution of the number of tasks for increasing task sizes for each task type

The observed patterns for the horizontal and vertical tasks align with the behaviour depicted by the zone model. This can be attributed to the fact that tasks confined to a single zone generally involve a smaller number of picks, as illustrated by Figures 6.12a and 6.12b, respectively.

This trend is particularly pronounced for the horizontal zone, where 61% of the tasks have less than 10 picks. This was expected, given the low probability of entering this zone. In fact, only 253 tasks (1.38% of the total tasks) exclusively address this area.

As for the vertical area, tasks are still small but, in comparison, greater than the horizontal ones. However, there are 10,749 tasks confined solely to this zone, which represents almost 60% of all instances. This can be explained by the placement of high-rotation stores within this area.

The other majority of tasks involve picks in both zones. These tasks, as depicted in Figure 6.12c, tend to be larger in size and exhibit a similar pattern to the one described by the mixed model. Furthermore, only 9% of these tasks have less than 10 picks. So, tasks that accommodate work in progress from the previous day or punctual deliveries of an SKU, are usually confined to a specific area.

In summary, the zone model seems to provide a more accurate response for tasks assigned to a single zone, while the mixed model appears to be a better estimator of mixed tasks. The next section will delve deeper into evaluating the accuracy of this statement, by comparing the estimated distances with the actual distances.

6.4 Validation of the Analytical Models

The calculation of the actual distances was conducted using a discrete-event simulation model. This model emulates, in detail, the current situation of the warehouse at hand. Unlike the analytical models that make certain approximations, it is built upon an AutoCAD file that contains the precise measurements of the picking area. This approach, illustrated in Figure 6.13, ensures a more accurate representation of the PBL block, eliminating the need for grid-based approximations.

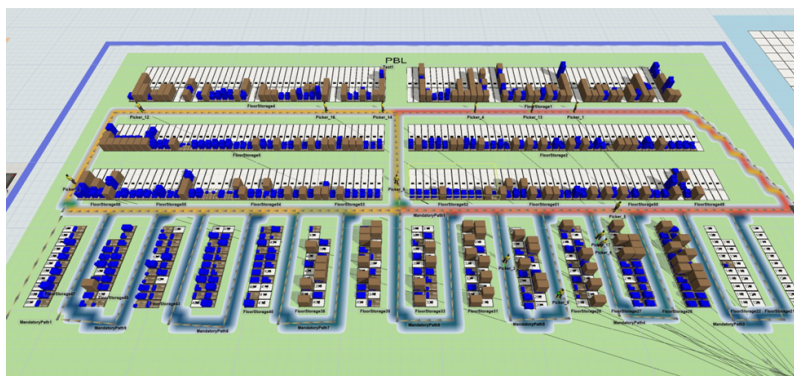


Figure 6.13: 3D View of the PBL block's current layout using DES

The simulation model corresponds to the second 'S' of the SOS methodology. However, it is worth mentioning that this project is still on its embryonic phase. Hence, it cannot simulate the entire time horizon due to its current limitations in terms of running time. This highlights one of the key advantages of the probabilistic models employed - much faster computational performance.

To tackle this problem, only three days were carefully selected. This was done in order to maintain a representative sample of different levels of daily activity: low, medium, and high.

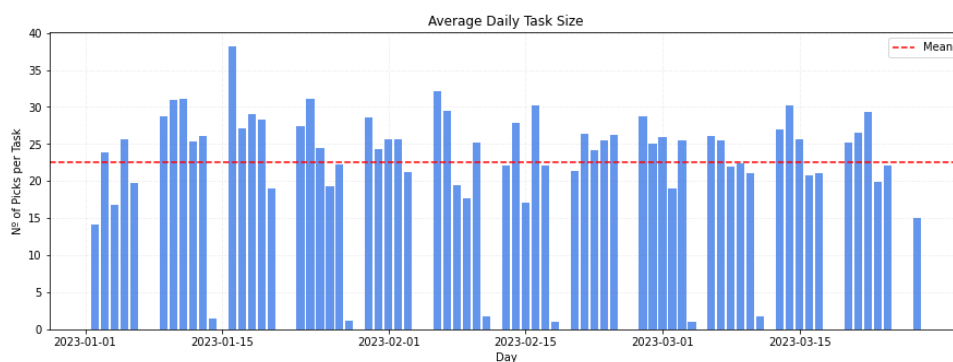


Figure 6.14: Daily distribution of the average daily task size

One shall not select these days based on the number of daily tasks. This metric can be biased by the number of picks per task. Hence, an analysis was made to identify the most suitable days based on the average task size.

By referring to Figure 6.14, it becomes evident that the 2nd of January represents a day with relatively low activity, while the 16th of January corresponds to the day with the highest work level. As for a 'normal' day, the 22nd of February was deemed appropriate for analysis.

Furthermore, Table 6.3 deconstructs the number of tasks and picks for these selected days. It also provides a comparison against all the available data, demonstrating that the sample is well-representative. In fact, the percentages of task types and picks within task types are similar.

Table 6.3: Picking task data deconstructed by task type

Tasks					Picks				
Day	Only H	Only V	Both	Total	Day	Only H	Only V	Both	Total
Low	20	147	57	224	Low	46	1902	1956	3904
Med	11	212	91	314	Med	179	2961	5227	8367
High	1	197	118	316	High	32	4661	7373	12066
Total	32	556	266	854	Total	257	9524	14556	24337
Freq.	3.75%	65.11%	31.15%	100.00%	Freq.	1.06%	39.13%	59.81%	100.00%

Day	Only H	Only V	Both	Total	Day	Only H	Only V	Both	Total
All Days	253	10749	7344	18346	All Days	4776	206051	307451	518278
Freq.	1.38%	58.59%	40.03%	100.00%	Freq.	0.92%	39.76%	59.32%	100.00%

Now, it is time to assess the quality of the derived models based on these data. The Mean Absolute Percentage Error (MAPE) is a commonly used metric for evaluating model accuracy. It measures the average absolute differences between the estimated distance and the actual distance, expressed as a percentage of the actual distance. Mathematically:

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i} \quad (6.1)$$

A separate MAPE is calculated for each model due to their distinct approaches to estimating the distances of a picking task:

- **Zone Model:** The horizontal tasks are estimated using the horizontal model. Likewise, the vertical tasks are estimated by applying the vertical model. Both these models are represented in Figure 6.6. For tasks that encompass both zones, however, a weighted average of these zonal models has to be calculated. The weights are determined by the probabilities of a picker entering inside each zone, as indicated in Table 6.2;
- **Mixed Model:** The estimated distances for the mixed model are obtained directly from the graph in Figure 6.8. This is done regardless of the task type;
- **Monte Carlo:** There is a lot of variability when considering a task of size N. This is well depicted in Figure 6.9. However, to capture the average behaviour, the mean value was computed for each set of randomly generated tasks. This provides a polynomial line that connects the expected values of all the boxplots.

The results obtained are depicted in Table 6.4. In addition, the Mean Percentage Error (MPE) was also analyzed to evaluate whether the model tends to consistently underestimate or overestimate the real distances. Therefore, the MPE does not consider absolute differences.

$$MPE = \frac{100\%}{N} \sum_{i=1}^N \frac{y_i - \hat{y}_i}{y_i} \quad (6.2)$$

Table 6.4: Accuracy metrics for the derived models

	Zone Model	Mixed Model	Monte Carlo
MAPE	31.34%	29.69%	26.38%
MPE	-24.40%	-9.36%	8.43%

Both these metrics return error as a percentage, making it easier to compare values across different models. A good value is dependent upon the use case, but as a rule of thumb, a MAPE lower than 20% is considered good.

MAPE indicates, on average, how distant the model was from the actual values. Bias, on the other hand, refers to persistent forecast error. Therefore, the MPE was used to assess consistent under and/or over-forecasting. However, these metrics can be misinterpreted since they are giving the same importance to all tasks. So, Table 6.5 aims to deconstruct these measures per task type.

Table 6.5: Accuracy metrics deconstructed by task type

Zone Model	Only H	Only V	Both Zones	Mixed Model	Only H	Only V	Both Zones	Monte Carlo	Only H	Only V	Both Zones
MAPE	66.66%	25.30%	39.70%	MAPE	44.15%	32.52%	22.02%	MAPE	20.63%	29.58%	20.39%
MPE	-66.66%	14.89%	-39.21%	MPE	-41.11%	-7.64%	-9.14%	MPE	-6.62%	14.36%	-2.16%

The zone model was specifically designed to emulate independent areas. So, it was expected to yield better results for tasks that only addressed one zone.

This is particularly evident in the case of the vertical tasks, which demonstrate a MAPE of 25.30%. However, it is important to note that this value is slightly higher than expected. One possible explanation for this is the rapid convergence of the vertical model.

In fact, the vertical zone contains stores with a high number of visits and is composed by four aisles only. As a result, even for tasks with a small number of picks, the probability of visiting all four aisles is relatively high. However, each aisle consists of multiple picking positions. Hence, there can be tasks with several picks that do not need to traverse the entire zone. That is why this model overestimates, on average, 14.89% of the travelled distance inside this zone.

As for the other models, they were expected to outperform the zone model for tasks that addressed both zones, This is supported by a MAPE of approximately 20%, indicating very good results for both cases.

In addition, these models were also expected to accurately estimate the distances for vertical tasks, given the majority of picks are in this area. However, they do not differentiate between task types. Unlike the zone model, they solely consider task size. Despite this limitation, they still demonstrate an acceptable MAPE of roughly 30%.

Interestingly, the analytical models struggle when facing horizontal tasks. However, it is worth recalling that these tasks only account for 3.75% of the analyzed instances. Moreover, this value is even smaller when considering the entire dataset.

Despite Monte Carlo's promising results in estimating these distances, it would be premature to deem it the superior model for this specific task type. This is primarily due to the scarcity of data points available for instances that solely address the horizontal zone. In fact, these tasks are so rare that they can be considered negligible when deriving the optimal model.

6.5 Deriving the Optimal Model

In order to further refine this approach, it is crucial to consider the suitability of different models for different task types. The objective is to identify the most effective model for horizontal, vertical, and mixed tasks.

First things first, the Monte Carlo Simulation is going to be used as a proxy for horizontal tasks. This decision stems from the scarcity of data concerning these instances. By employing this approach, multiple scenarios of hypothetical tasks can be randomly generated, taking into account the probabilities of visiting each picking position. This effectively compensates for the lack of information through oversampling.

For the other task types, however, this was not a problem. Thus, for each task size N , with at least 10 data points, the average task error will be calculated. To better understand this, consider the following example: suppose there are 20 vertical tasks with 10 picks. Each individual task will then have a forecast error. Nevertheless, the forecast error for a vertical task with size $N = 10$ can be better expressed by averaging those individual errors.

By doing this for all task sizes that meet this criteria, one can have a glimpse of the evolution of each model gap, for an increasing task size. This was done to determine the most suitable model for vertical and mixed tasks. Figure 6.15 provides an analysis for the vertical ones.

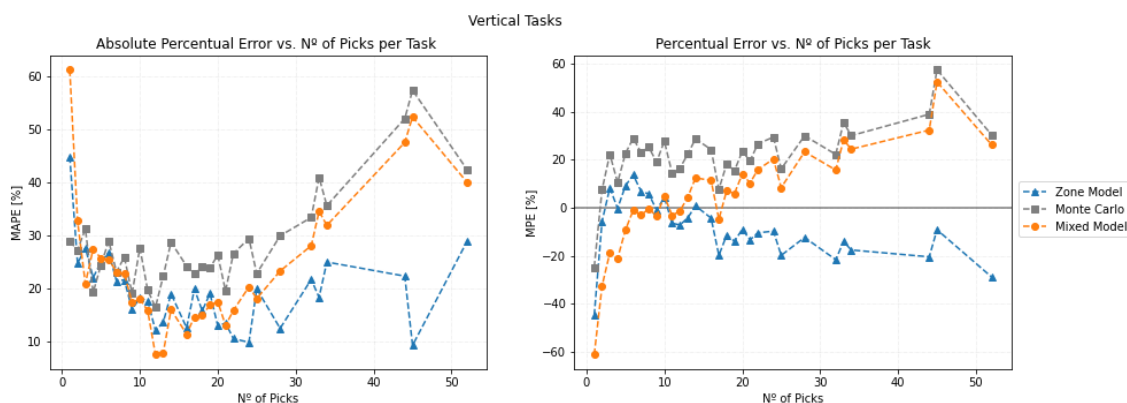


Figure 6.15: Evolution of each model's errors for an increasing task size of vertical tasks

By looking at Figure 6.15a, all models are pretty similar when estimating distances for vertical tasks with less than 10 picks. However, as the task size increases, the zone model demonstrates

a lower absolute percentage error. Interestingly, Figure 6.15b illustrates that the zone model also exhibits a slight tendency to underestimate the actual distances for these tasks.

This reinforces the idea that the zone model is more suitable when addressing vertical tasks. Nevertheless, the other models can still provide reliable distance estimates. However, they tend to overestimate these figures.

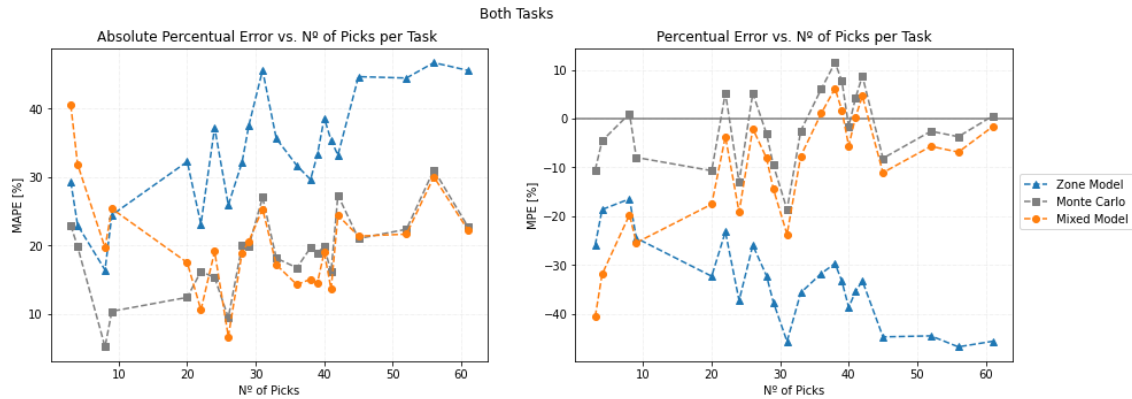


Figure 6.16: Evolution of each model's errors for an increasing task size of mixed tasks

A similar analysis was conducted for tasks that encompass both zones. As anticipated, Figure 6.16a reveals that both the mixed model and the Monte Carlo simulation outperform the zone model for all the spectrum of task sizes. These models exhibit almost the same trend and, therefore, could be used interchangeably.

Additionally, according to Figure 6.16b, they tend to slightly underestimate the distances, being almost on point for bigger tasks. In contrast, the zone model shall not be used since it is worst in all aspects.

Based on these findings, the following conclusions can be drawn: the Monte Carlo simulation is recommended for estimating distances in horizontal tasks, the zone model is more suitable for vertical tasks, and the mixed model is the preferred choice for tasks that address both zones. All of this led to a final MAPE of only 14.20%.

Chapter 7

Conclusions and Future Work

This chapter encompasses the key findings and insights derived from the conducted research. It not only highlights the contributions of the study to the broader field of knowledge but also provides a critical assessment of the methodology employed. Therefore, limitations and constraints are acknowledged, paving the way for suggestions and recommendations for future work.

In essence, this chapter serves as a bridge between the present research and the potential future directions that can be pursued, providing a comprehensive overview of the work done and laying the foundations for further advancements in this field.

7.1 Key Findings and Insights

Literature highlighted that order-picking is the most labour-intensive operation in a warehouse. In addition, travelling is often the dominant activity of an order-picker's time. Hence, minimizing the travel distance is considered a key objective in order-picking - and warehouse - optimization.

All picking areas are unique but they all can be modelled as a grid. However, there are layouts more complex than others. That is why, two different models were derived: the first one divides the picking area into geometric zones and treats them independently; The second one considers the whole picking area without having to make any assumptions.

Both models utilize a probabilistic simulation approach to estimate the performance of a given picking area. This is done by predicting the expected travel distances for an increasing task size, allowing quick comparisons of several layout solutions with residual forecast errors.

Grid Modelling is the framework that allows to replicate, and assess, any possible solution that anyone could have thought of. In fact, in order to improve picking efficiency, warehouse managers typically face various decision problems: internal layout design, picking policy, storage assignment rules, routing methods, zoning, etc. So, each possible combination of those factors will lead to a different solution - and there are millions of them.

Nevertheless, the research conducted in this study aims to overcome this pain point. In reality, this methodology was thought to be inserted in a product that will enable its users to construct

accurate replicas of a given picking area. This will allow to test, assess, and compare several layouts at a distance of a click and without having to make any physical modifications.

In order to assess the quality of the proposed approach, a real-life scenario was needed. Therefore, a warehouse of a large Portuguese retailer was used to validate the derived models. The primary focus was on the PBL Block which mainly handled household products.

Based on the assessment of the "AS-IS" scenario, the following conclusions could be drawn: The zone model demonstrates better suitability for tasks that only address one zone; For tasks that involve both zones, however, the mixed model is the preferred choice.

It is also worth mentioning that the use of oversampling techniques is recommended to improve the estimation of zones with limited data availability. In particular, the application of Monte Carlo simulations yielded more precise results when estimating distances for horizontal tasks.

All in all, these findings resulted in a final MAPE of only 14.20%. This indicates the effectiveness of the proposed models to provide accurate insights into the order-picking operations. Hence, new scenarios might be considered to further optimize the most promising ones.

7.2 Scientific Contributions

This framework was initially developed in Alves (2015) and further refined in Amorim-Lopes et al. (2020). However, the analytical model that estimates the expected travel distances can only be applied to picking areas where corridors are disposed either horizontally or vertically, not both.

This thesis made improvements to the existing model in chapter 4, but its major contribution lies in the development of a second model that can be applied to any possible picking area. This was only achievable due to the innovative approach of modelling the picking area as a grid.

Besides, the existing methodology was originally thought for PBS only. So, this work aims to introduce all these novelties into the existing literature and extend this rationale to a PBL flow type.

Moreover, most of the existing research focuses on a specific order-picking area. Therefore, it is not straightforward to transpose methods developed for a given situation into another situation. General procedures for order picking and warehouse design are still lacking. However, the project in which this thesis is inserted aims to fill in this void in the literature (and in the market).

To conclude, these models can have broader applications beyond the field of study. In fact, it is possible to transpose the derived models to more than just warehouse layout design. For example, the tasks performed by store repositories are no different from what a picker has to do. Operators have a cart with N items and traverse the store aisles to drop off the SKUs as needed. Hence, this could easily be applied to store layout design.

7.3 Main Limitations and Recommendations

The applicability of the proposed models may be limited for very large-sized picking areas. This is especially true for those areas that comprise corridors with different orientations. As previously

mentioned, the modelling of such areas are in need of a more detailed approach. Hence, when the cell is the elementary unit, the computational time required to run the model explodes.

Nevertheless, a relaxation was proposed in order to tackle this pain point. However, obtaining the probabilities of visiting each cell is no longer feasible. This proved to be beneficial in order to assess potential congestion issues.

The next aspect to consider is the way distances are estimated. Currently, both models utilize the von Neumann's neighbourhood to represent the picker's movements. However, this approximation may not be ideal as it leads to robot-like motions similar to those seen in the Snake Game.

In fact, when moving towards the next cell, the picker only considers the four cardinal directions (up, down, left, and right). To address this limitation, it is recommended to consider the Moore's neighbourhood, which allows for diagonal movements as well. Figure 7.1 provides a visual representation of this concept.



Figure 7.1: Difference between von Neumann's and Moore's neighbourhoods

Furthermore, it is important to address the issue of grid convergence problems. This challenge can be mitigated by assigning different dimensions to each cell. By doing so, each cell can still be treated as an individual node but without requiring any adjustments.

Moreover, certain esoteric layouts may pose challenges in terms of modelling due to the rectangular nature of the cells. Figure 7.2 provides a visual representation that helps illustrate this situation. As a result, layout types such as the fishbone layout, the leaf layout, or others that comprise rotated aisles, may not be accurately approximated.

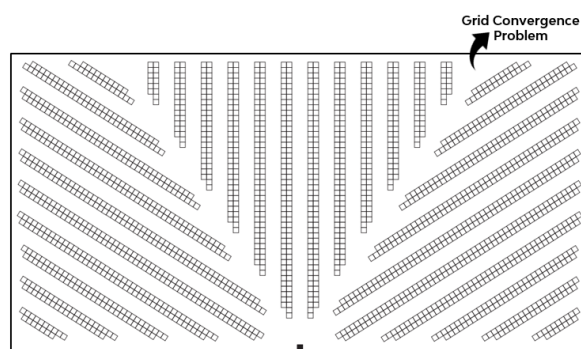


Figure 7.2: Grid convergence problems hard to solve for uncommon layouts

7.4 Future Research Directions

The focus now shifts towards the future work of the project in which this thesis is inserted. All of these further developments are built upon the methodology hereby presented. Hence, this section aims to provide a roadmap for concluding this research.

7.4.1 Short Term

This first part ought to conclude the first 'S' of the SOS Methodology. This entails exploring several different combinations of layout design variables. One approach to achieve this is by exploring different store allocation policies, such as the ones depicted in Figure 7.3.

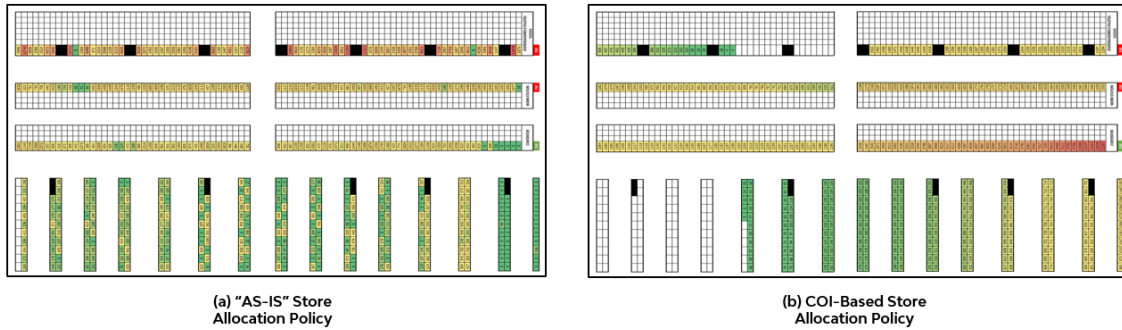


Figure 7.3: Two different store allocation policies for the same layout solution

The first policy represents the current approach, while the second one is based on the COI Index described back in section 2.2.3.3. In fact, the latter decreases the expected travel distances by assigning high-rotation stores closer to the inbound/outbound points. However, this policy will be certainly prone to congestion issues.

Another aspect that can be modified is the internal layout design. This encompasses factors such as the number, orientation, position, and dimensions of the aisles and their racks. Additionally, new routing constraints - like bidirectional corridors - can be considered. Furthermore, the concept of managerial zoning can be explored as well.

In conclusion, it is essential to consider different scenarios in order to optimize the most promising ones. This lays the foundation for the 'O' in the SOS Methodology.

7.4.2 Long Term

In the long term, this research aims to transition from simply estimating layout performances to actually optimize them. However, most of the optimization models output a static store allocation policy for a given layout. Therefore, the main goal will be to do this dynamically and according to customers' orders. Hence, stores with a high number of orders, for a given period, shall be assigned closer to the receiving zone in order to minimize the distance travelled.

To effectively incorporate trends and seasonality, a longer time window is required. For instance, the warehouse in Chapter 6 experiences peaks of activity during events like the start of the school year in September or the Christmas season in December.

The next step involves extending this procedure to all blocks within the warehouse. So, various scenarios for the Reception, PBS, and Marshalling areas shall be simulated and optimized. However, any combination of optimized blocks can be sub-optimal since interactions between blocks were not considered prior to this point. Thus, a validation phase is followed to test the robustness of the generated solution.

This is achieved through the second 'S' in the SOS Methodology. Hence, any underperforming block is replaced with the second-best one obtained in the previous steps. Therefore, this iterative search aims to find the best overall solution for warehouse design.

7.5 Final Remarks

Finally, this framework needs to be tested in other warehouses to further validate its effectiveness. All in all, the ultimate goal is to refine and industrialize this methodology, making it accessible to warehouse managers all over the world. By doing so, this research has the potential to enhance warehouse operations while providing efficient and integrated solutions. This is crucial, particularly at a time when general procedures for warehouse design are still lacking.

Bibliography

- Alves, J. F. L. (2015). Combining simulation and optimization to improve picking performance on a specialized retailer warehouse. Master's thesis, Faculdade de Engenharia da Universidade do Porto.
- Amorim-Lopes, M., Guimarães, L., Alves, J., and Almada-Lobo, B. (2020). Improving picking performance at a large retailer warehouse by combining probabilistic simulation, optimization, and discrete-event simulation. *International Transactions in Operational Research*, 28(2).
- Bahrami, B., Aghezzaf, E. H., and Limere, V. (2017). Using simulation to analyze picker blocking in manual order picking systems. *Procedia Manufacturing*, 11:1798–1808.
- Baker, P. (2007). An exploratory framework of the role of inventory and warehousing in international supply chains. *The International Journal of Logistics Management*, 18(1):64–80.
- Bartholdi, J. J. and Hackman, S. T. (2019). *Warehouse & Distribution Science*. The Supply Chain and Logistics Institute.
- Caron, F., Marchet, G., and Perego, A. (2000). Optimal layout in low-level picker-to-part systems. *International Journal of Production Research*, 38(1):101–117.
- Chan, F. T. S. and Chan, H. K. (2011). Improving the productivity of order picking of a manual-pick and multi-level rack distribution warehouse through the implementation of class-based storage. *Expert Systems with Applications*, 38(3):2686–2700.
- Chen, T. L., Cheng, C. Y., Chen, Y. Y., and Chan, L. K. (2015). An efficient hybrid algorithm for integrated order batching, sequencing and routing problem. *International Journal of Production Economics*, 159:158–167.
- Christopher, M. (2011). *Logistics Supply Chain Management, 4th Edition*. Prentice-Hall, Harlow.
- Christopher, M. (2017). *Relationships and Alliances: Embracing the era of network competition*, pages 272–284. Routledge.
- Christopher, M., Peck, H., and Towill, D. (2006). A taxonomy for selecting global supply chain strategies. *The International Journal of Logistics Management*, 17(2):277–287.
- Christopher, M. and Towill, D. (2001). An integrated model for the design of agile supply chains. *International Journal of Physical Distribution Logistics Management*, 31(4):235–246.
- Coase, R. (1937). The nature of the firm. *Economica*, 4(16):386–405.
- Daniels, R. L., Rummel, J. L., and Schantz, R. (1998). A model for warehouse order picking. *European Journal of Operational Research*, 105(1):1–17.

- Davenport, T. H. (1993). *Reengineering Work through Information Technology*. Harvard Business School Press.
- de Koster, R. (2012). *Warehouse Assessment in a Single Tour*, pages 457–473. Springer-Verlag, London.
- de Koster, R., Le-Duc, T., and Roodbergen, K. J. (2007). Design and control of warehouse order picking: a literature review. *European Journal of Operational Research*, 182(2):481–501.
- de Koster, R. B. M., Johnson, A. L., and Roy, D. (2017). Warehouse design and management. *International Journal of Production Research*, 55(21):6327–6330.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische matematik*, 1(1):269–271.
- Figueira, G. and Almada-Lobo, B. (2014). Hybrid simulation–optimization methods: A taxonomy and discussion. *Simulation Modelling Practice and Theory*, 46:118–134.
- Frazelle, E. (2002). *SUPPLY CHAIN STRATEGY The Logistics of Supply Chain Management*. McGraw-Hill, New York, NY.
- Goetschalckx, M. and Ashayeri, J. (1989). Classification and design of order picking systems. *Logistics World*, 2(2):99–106.
- Grosse, E. H., Glock, C. H., and Neumann, W. P. (2017). Human factors in order picking: a content analysis of the literature. *International Journal of Production Research*, 55(5):1260–1276.
- Gu, J., Goetschalckx, M., and McGinnis, L. F. (2007). Research on warehouse operation: A comprehensive review. *European Journal of Operational Research*, 177(1):1–21.
- Hall, R. W. (1993). Distance approximations for routing manual pickers in a warehouse. *IIE Transactions (Institute of Industrial Engineers)*, 25(4):76–87.
- Hausman, W. H., Schwarz, L. B., and Graves, S. C. (1976). Optimal storage assignment in automatic warehousing systems. *Management Science*, 22(6):629–638.
- Heskett, J. (1963). Cube-per-order index - a key to warehouse stock location. *Transport and Distribution Management*, 3:27–31.
- Jarvis, J. M. and McDowell, E. D. (1991). Optimal product layout in an order picking warehouse. *IIE Transactions (Institute of Industrial Engineers)*, 23(1):93–102.
- Kallina, C. and Lynn, J. (1976). Application of the cube-per-order index rule for stock location in a distribution warehouse. *Interfaces*, 7(1):37–46.
- Kearney (2022). 33rd annual state of logistics report: Out of sync. Technical report, Council of Supply Chain Management Professionals.
- Lambert, D. M. and Cooper, M. C. (2000). Issues in supply chain management. *Industrial Marketing Management*, 29(1):65–83.
- Lambert, D. M., Emmelhainz, M. A., and Gardner, J. T. (1996). Developing and implementing supply chain partnerships. *The International Journal of Logistics Management*, 7(2):1–18.

- Lambert, D. M., Stock, J. R., and Ellram, L. M. (1998). *Fundamentals of Logistics Management*. McGraw-Hill.
- Lawler, E. L., Lenstra, J. K., Kan, A. H. G. R., and Shmoys, D. B. (1985). *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. Wiley, Chichester.
- Le-Duc, T. and de Koster, R. B. M. (2005). Travel distance estimation and storage zone optimization in a 2-block class-based storage strategy warehouse. *International Journal of Production Research*, 43(17):3561–3581.
- Lowson, R. H. (2002). Assessing the operational cost of offshore sourcing strategies. *The International Journal of Logistics Management*, 13(2):79–90.
- Meiler, R. D. and Gau, K. Y. (1996). The facility layout problem: Recent and emerging trends and perspectives. *Journal of Manufacturing Systems*, 15(5):351–366.
- Oke, A. and Gopalakrishnan, M. (2009). Managing disruptions in supply chains: A case study of a retail supply chain. *International Journal of Production Economics*, 118(1):168–174.
- Petersen, C. G. (1997). An evaluation of order picking routeing policies. *International Journal of Operations Production Management*, 17(11):1098–1111.
- Petersen, C. G. and Aase, G. (2004). A comparison of picking, storage, and routing policies in manual order picking. *International Journal of Production Economics*, 92(1):11–19.
- Petersen, C. G., Aase, G. R., and Heiser, D. R. (2004). Improving order-picking performance through the implementation of class-based storage. *International Journal of Physical Distribution & Logistics Management*, 34(7):534–544.
- Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance - With a new Introduction*. The Free Press, New York, NY.
- Radnor, Z. J. and Boaden, R. (2004). Developing an understanding of corporate anorexia. *International Journal of Operations and Production Management*, 24(3-4):424–440.
- Roodbergen, K. J. and Koster, R. D. (2001). Routing methods for warehouses with multiple cross aisles. *International Journal of Production Research*, 39(9):1865–1883.
- Rouwenhorst, B., Reuter, B., Stockrahm, V., van Houtum, G. J., Mantel, R. J., and Zijm, W. H. M. (2000). Warehouse design and control: Framework and literature review. *European Journal of Operational Research*, 122(3):515–533.
- Rushton, A., Croucher, P., and Baker, P. (2014). *The Handbook of Logistics and Distribution Management: Understanding the Supply Chain, 5th Edition*. Konan Page, London.
- Snehota, I. and Håkansson, H. (1995). *Developing Relationships in Business Networks*. Routledge.
- Sprock, T., Murrenhoff, A., and McGinnis, L. F. (2017). A hierarchical approach to warehouse design. *International Journal of Production Research*, 55(21):6331–6343.
- Stigler, G. J. (1951). The division of labor is limited by the extent of the market. *Journal of Political Economy*, 59(3):185–193.
- Tompkins, J. A., White, J. A., Bozer, Y. A., and Tanchoco, J. M. A. (2010). *Facilities Planning, 4th Edition*. Wiley, Hoboken, NJ.

- van Gils, T., Ramaekers, K., Braekers, K., Depaire, B., and Caris, A. (2018). Increasing order picking efficiency by integrating storage, batching, zone picking, and routing policy decisions. *International Journal of Production Economics*, 197:243–261.
- van Hoek, R. I. (1998). Reconfiguring the supply chain to implement postponed manufacturing. *The International Journal of Logistics Management*, 9(1):95–110.
- van Hoek, R. I., Harrison, A., and Christopher, M. (2001). Measuring agile capabilities in the supply chain. *International Journal of Operations and Production Management*, 21(1-2):126–148.
- Waters, D. (2003). *Inventory Control and Management*. Wiley, Chichester.
- Womack, J. P. and Jones, D. T. (1996). *Lean thinkingbanish waste and create wealth in your corporation*. Simon & Schuster, London.
- Wäscher, G. (2004). *Order Picking: A Survey of Planning Problems and Methods*, pages 323–347.