Using Simulation to Improve Checkout Management: a Case Study in a Retail Company

Diogo Emanuel Pereira Miranda

Master's Dissertation

Supervisor: Prof. José Luís Moura Borges

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Mestrado Integrado em Engenharia e Gestão Industrial

2018-06-29

Abstract

Grocery retail is a highly competitive industry, where companies battle for a share of the population's income while dealing with thin profit margins. In this environment, the shopping experience plays a crucial role in the customers' perception of service excellence, being intimately linked with their loyalty and preference for one company over the others. When addressing the in-store customer experience, the process of payment and store-exit taking place in the checkout area is of utmost importance, being acknowledged by the retailers as one of the areas to which greater focus should be devoted. The main challenge lies on the definition of the ideal number of each type of checkout to fulfill customers' demand and expected service level. The dilemma acquires an increasing importance considering the wide range of checkout solutions available nowadays and the high fluctuation of arrival rates across the day, week and year.

The current work was motivated by the case of a Portuguese food retailer, with the objective of providing an analytical support concerning the checkout dimensioning of its stores. The project started by understanding which are the current checkout concepts available in the stores and their specificities. This stage also encompassed the modeling of the key parameters that characterize each checkout typology, the customers' profiles and the global checkout process of each store. Afterwards, it is proposed a two-fold methodology based on queueing theory and simulation to tackle the checkout dimensioning problem. The first step consists on the definition of a preliminary ideal configuration for each store, by using queueing models to represent the expected impact of each checkout configuration. Subsequently, the simulation model tests the configuration's robustness and appropriateness by simulating the store's operation throughout an entire year, suggesting possible required adjustments to the preliminary ideal configuration. The performance assessment includes a sensitivity analysis to evaluate the trade-off between service level and associated costs, as well as robustness tests that evaluate its suitability to changes in the arrival rate. Although in this work the methodology has been applied to the specific case of the studied retailer, the procedures used can easily be extended to other retailers' contexts.

The application of the proposed methodology led to a significant change in the recommended number of checkouts to be available in the stores. The suggested modification of the configurations is expected to lead to a relevant reduction in the costs associated with the checkouts while maintaining a high service level. Additionally, the results obtained also contributed to the identification of improvement opportunities in the in-store checkout process, which could bring benefits for both the retailer and customers.

Resumo

A indústria do retalho alimentar é altamente competitiva, na qual as empresas lutam por uma parcela do rendimento da população operando simultaneamente com margens de lucro reduzidas. Nesse meio, a experiência de compra desempenha um papel crucial na percepção dos clientes quanto à excelência do serviço, estando intimamente ligada à sua lealdade e preferência por uma empresa sobre as demais. Relativamente à experiência do cliente na loja, o processo de pagamento e saída de loja que ocorre na área de *checkout* assume extrema importância, sendo reconhecido pelos retalhistas como uma das principais áreas onde se devem focar. O principal desafio reside na definição do número ideal de cada tipo de checkout para responder à procura dos clientes e ao nível de serviço esperado. O dilema adquire uma importância crescente, considerando a ampla gama de soluções de checkout disponíveis hoje em dia e a alta flutuação das taxas de chegada de clientes ao longo do dia, semana e ano.

Este trabalho foi motivado pelo caso de um retalhista alimentar português, com o objetivo de fornecer um suporte analítico relativamente ao dimensionamento de *checkouts* das suas lojas. O projeto começou pelo levantamento dos conceitos atuais de *checkout* disponíveis nas lojas e das suas especificidades. Esta etapa também englobou a modelação dos parâmetros-chave que caracterizam cada tipologia de *checkout*, os perfis de clientes e o processo global de *checkout* de cada loja. Posteriormente, propõe-se uma metodologia dividida em duas fases para abordar o problema de dimensionamento de *checkouts*, sendo esta baseada em teoria de filas de espera e simulação. A primeira fase consiste na definição de uma configuração ideal preliminar para cada loja, usando modelos de filas de espera para representar o impacto esperado de cada configuração de *checkout*. De seguida, o modelo de simulação testa a robustez e conformidade da configuração simulando a operação da loja durante um ano inteiro, sugerindo possíveis ajustes necessários à configuração ideal preliminar. A avaliação de desempenho inclui uma análise de sensibilidade para avaliar o trade-off entre o nível de serviço e os custos associados, bem como testes de robustez que avaliam a adaptabilidade da configuração às mudanças na taxa de chegada. Embora neste projeto a metodologia tenha sido aplicada ao caso específico do retalhista estudado, os procedimentos usados podem facilmente ser extendidos à realidade de outros retalhistas.

A aplicação desta metodologia conduziu a uma mudança significativa no número recomendado de *checkouts* a estar disponível nas lojas. Espera-se que as alterações sugeridas das configurações leve a uma redução relevante nos custos associados aos *checkouts*, mantendo um nível de serviço elevado. Além disso, os resultados obtidos também contribuíram para a identificação de oportunidades de melhoria no processo de *checkout* na loja, o que poderia trazer benefícios tanto para o retalhista como para os clientes.

Acknowledgements

During my five-year journey as an engineering student, I got along with many special people that contributed to such an incredible experience. Assuming the risk of forgetting someone, some of them deserve an honorable mention and a particular acknowledgement.

First, I would like to thank my supervisor, Professor José Luís Borges, for all the support during the last months. Your insights and guidance were a precious help for the success of this dissertation.

To Professors Bernardo Almada Lobo, Luís Guimarães and Pedro Amorim, I am truly grateful for the opportunity of working in such a challenging company as LTPlabs. The whole team, my fellow LTPeers, welcomed me with open arms and made each day more interesting. To all of you, a sincere "thank you" for teaching me so much and presenting me with such a passionate and amazing working environment.

I would like to give a special praise to Dr. Teresa Bianchi de Aguiar, my manager and supervisor at LTPlabs, for providing me with such a careful and interested guidance. Thank you for always being available to answer my doubts and concerns. I definitely learned a lot from you on how to go around the obstacles that inevitably appear in our day-to-day. Not forgetting the ones who also worked with me in this project, I would like to thank José Queiroga, for his advices and excellence while revising my work, and Ricardo Cunha, to whom I owe great part of what I learned about simulation.

Back at home, there was always the family support that enabled me to push forward during the hardest times. To Dad, thank you for being so proud of me. Even when I doubted myself, your trust in my success never trembled. To Mum, no words are enough to describe my gratitude. You have always been there for me, every step of the way, being my safe-haven. I wish I could give you the world, you deserve that and so much more! While that isn't possible, I can only thank you for making me who I am today. To Sis, thank you for always caring for your little brother as you have been doing since I was born. You are the one to whom I look up to, my strongest role-model. Never leave me, ok?!

Of course, I could not leave behind my friends in this thank you note. You were the ones that uplifted this journey from good to awesome! In these five years, we grew together and shared our lives, dreams, fears and expectations. Thank you for being there celebrating with me in the good moments and cheering me up in the bad ones. To my "bros", hurrah! We did it! And if we stayed together through this hell of a ride, we all know our friendship is pretty much forever!

Diogo Miranda

"Everything negative is all an opportunity for me to rise.
The moment you give up, is the moment you let someone else win."

Kobe Bryant

Contents

Acronyms and Symbols

KPI	Key Performance Indicator
DE	Discrete-Event
AB	Agent-Based
CT	Checkout Typology
SL	Service Level
DSS	Decision Support System

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Competitive pressure in the stationary grocery retail environment is steadily increasing due to customer requirements and ongoing market consolidation (Mentzer and McCarthy-Byrne, 2011). As Schuller (2018) stated, retailers have to deal with aggressive competition in this industry, together with thin profit margins. Contrasting with the past, where supermarkets were simply a physical space to dispose and sell goods essential for living, nowadays they are also a space for technological and innovational experiences, with the excellence of this experience as the key strategy (Terblanche, 2018). This transformation in retail is intimately related with the evolution of the customer itself.

In a hastier day-to-day context, customers are more demanding, with convenience and service quality as two aspects highly appreciated. Moreover, they are significantly more informed and connected to their network, sharing their experiences with their acquaintances and being influenced by theirs. At the same time, in the last few decades retailers have been observing an increasing autonomy of the customer in the store. The traditional format of an assisted service has been progressively evolving to a self-service model in different areas and moments of the buying process, one of them being the checkout area. Beholding these facts, the current challenge for retailers reside in providing a consistent and quality in-store service at a competitive cost, in most cases with demanding circumstances such as high local competition or aggressive promotional activity. Such framework also stimulates the investment in innovative solutions that may enable them to differentiate from the competitors.

Regarding the customer experience, the process of payment and store-exit is one of the most critical for the service's global evaluation, with impact in the probability of repurchase and brand loyalty (Van Riel et al., 2012). In grocery retail, where buying cycles are characterized by a shorter frequency, assuring a consistent and convenient in-store experience is of utmost importance. In that sense, the perception of long queues and long waiting times taunts the customer evaluation and, therefore, its satisfaction and preference over the competition. In order to provide a high-quality service in the checkout process, retailers need to take into account multiple variables, so that they can offer an adequate response to the variable customers' inflow rate that is, in most cases, unpredictable. If on the one hand a wide range and number of checkouts available allows to reduce waiting times and queues, on the other it becomes greatly expensive to occupy physical space with checkouts barely used, due to associated costs such as equipment acquisition and maintenance.

The challenge on today's environment is to be able to have an analytical support behind the process of defining the checkout area configuration, in order to be more confident that the choice made is actually the most suitable for each store's profile.

1.2 Project description

This project was undertaken in a Portuguese food retailer that possesses around 250 stores spread around the country. This retailer has been having troubles defining the number and type of checkouts to be placed in its stores. The current dimensioning process is empirical and there is a lack of an analytical support for decisions related to this issue. Also, the retailer recognizes the need of a way to prioritize the urgency to reconfigure the checkout of current stores so that the budget available is directed to the stores in most critical state.

In an effort to improve checkout management in this context, the main goal of this project is to provide the retailer with an analytical background to determine the number of checkouts of each typology that each store should have, in order to efficiently fulfill the demand while accomplishing a minimum service level. Additionally, the development of a method to assess the checkout performance of current stores and rank the urgency of checkout reconfiguration is considered to be a key element to support the company's decision-making in this business area. It should be noted that, in this work, due to confidentiality agreements, absolute values related to KPIs or other metrics are omitted and will be used relative values instead.

In Figure 1.1, the major steps of this project are outlined with the respective chronogram.

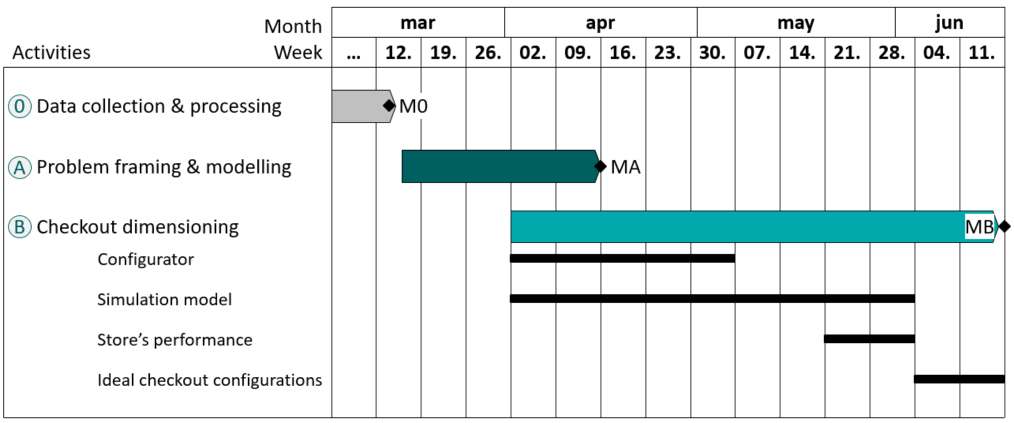


Figure 1.1: Project's timeline

The first stage of the project consisted in collecting and processing the data required to the forthcoming steps. This activity was done in close collaboration with the IT department of the retailer company.

In a second phase the problem was framed, consisting on understanding which are the current checkout concepts and store configurations used by the company. Moreover, the key parameters from the checkout queueing system (e.g. clients' arrival rate) were defined and a study was conducted for the identification of the factors that explain these parameters and how to model them accordingly.

The third stage was the checkout dimensioning, with its complexity requiring the division into four sub-stages. Firstly, the construction of a configurator, which consists of a tool based on queueing theory to identify the configurations to be tested in the simulator. Secondly, the development of the simulation model representative of the actual checkout process. This sub-stage also included the validation of the model's effectiveness by comparing the results with historical records and KPIs such as utilization time. Thirdly, the evaluation of the performance of current stores' configurations and their classification according to their checkout reconfiguration need. Lastly, the identification of the ideal store configuration for under-performing stores by exploiting the features of different checkout concepts to better answer the store's characteristics.

1.3 Dissertation structure

This dissertation is organized in six chapters with the following structure:

Chapter 2 aims to provide the theoretical background of the most relevant topics for this thesis. First of all, the most significant checkout solutions are described. Secondly, a holistic characterization regarding queueing theory is provided. Thirdly, the different simulation types are enumerated and explained as a way to tackle queueing theory limitations. Finally, a literature review on the state of art related with retail checkout area analysis is conducted.

Chapter 3 describes in higher detail the problem being addressed and the retailer's context, as well as a depiction of the current dimensioning process and a preliminary performance assessment. Also, a brief summary of this work's proposed approach to tackle the dimensioning problem is presented.

Chapter 4 presents the detailed methodology developed for the checkout dimensioning problem. Firstly, it identifies how the key parameters that characterize each store's profile were modeled. Afterwards, the construction of the configurator is described. Lastly, the logic that supports the mechanism of the simulation model is specified and its validation is described.

Chapter 5 depicts the results of the implementation of the proposed methodology in the retailer's case study. It begins to show a detailed report for a specific store before describing the global outcomes.

Chapter 6 summarizes the main conclusions drawn, as well as a description of future work and improvement opportunities.

Chapter 2

Theoretical Background

The goal of this chapter is to induce the reader into the theoretical background that guided the development of this project, related not only with the checkout area theme but also the basis of the methodology implemented. In section 2.1, the most relevant checkout solutions currently used in retail stores are enumerated and described as an intent to contextualize the existing offer. This description encompasses not only the diversity of checkout typologies that exist nowadays but also some novel technologies that have been surfacing in recent years. In section 2.2, queueing theory is covered in a holistic perspective to provide a global characterization of the elements and the notation of one of the methodology pillars. In section 2.3, simulation is presented as one way to answer queueing theory limitations, explaining the different methods available when choosing to adopt this approach. Finally, in section 2.4, the state of the art regarding retail checkout area and in-store customer behavior analysis is addressed, as well as possible gaps in the existing research.

2.1 Checkout typologies and technologies

Regular checkout

The most common type of checkout found in stores is the so called "regular checkout", where the checkout procedure is facilitated by a store employee (a cashier) who helps the customer register the products, calculates the total value and accepts the payment (Hartman, 2008). However, the ever-increasing competition in grocery retail puts pressure on retailers to find ways to differentiate themselves from their competitors in many ways, one being the customer service. This has led the retail stores to attempt to improve the checkout processes for their customers in many ways, one of them by offering the customers different alternatives when they arrive to this stage of the shopping experience (Rossetti and Pham, 2015). With the technological developments that have been occurring in the last decades, this range of alternatives is partially based on self-service equipments, which usually imply changes in the checkout process when comparing to the aforementioned procedure.

Self-checkout

In the early 90's an alternative to regular checkouts has emerged: the self-checkout. These systems or electronic point of sale scanning devices, allow consumers to scan their products and pay without, theoretically, any assistance from a register attendant (Davison, 1993). Since their first appearance, self-checkouts have been evolving through the years to better fit the ambitioned requirements, and nowadays they are expected to bring advantages both to the retailers and customers. For the former, reduce labor costs, increase employee availabilities or reclaim valuable floor space for additional sales are aspects mentioned as a reason for investment in this service (Amorim et al., 2016). For the customers, the expectation falls upon a more efficient service with lower waiting times and a higher control and privacy during the checkout process. However, the customer perception and satisfaction towards this technology play a significant role in its level of acceptance, being impacted by demographic factors like age, education or gender (Weijters et al., 2016).

The main downsides of this solution is that, in general, it is restricted to small basket sizes (carts are not allowed) and, even with experienced clients, the scanning operation isn't done as fast as with a trained assistant (Hartman, 2008). Moreover, Demirci Orel and Kara (2014) demonstrated that the service quality of self-checkout positively influences the customer satisfaction although it was not proven that it has direct effect on customer loyalty. Finally, in Opara-Nadi (2005) self- and regular checkouts are compared. Various hypothesis tests were conducted, with the results indicating that, although in terms of time efficiency self-checkouts were better, chances of system's errors increase when comparing with the regular checkouts.

Self-scanning

A technology close to the self-checkout is the self-scanning, in which the customer scans the products while picking them from the shelfs, offering a convenient way of interactive shopping (Pantano, 2014). The payment process is similar to the self-checkout despite the products being already scanned. To ensure that the customer has scanned all of its products, the retailers occasionally select a few customers to be audited and verified (Ekman, 2016). The customer is presented with two options to self-scan: use handheld scanners provided by the retailer or use their own smartphone.

The strengths of self-scanning are essentially related to a shorter waiting time due to a more expedited checkout process, since the scan is already done and there is no need to re-pack the items, and the labor saving to the retailer. Nonetheless, the chance of being audited too often can be perceived in a negative way by customers (Ekman, 2016). Also, like the self-checkout, the occurrence of system-errors leads to customers being annoyed and disturbed when using this service (Zagel, 2016).

Express-checkout

Once there are many different customer's profiles, one way of separating them in the checkout process is by the number of items they carry (Kwak, 2017). Express checkouts appeared to answer this topic and break up waiting times between customers that buy many items and the small-buying ones. This system is usually used as a complement of regular checkouts and is expected to reduce the waiting times of these small-buying customers. However, Schimmel (2013) analysis concluded that, although express checkouts have a positive impact on the expected waiting time of express customers, they strongly harm this metric's value for the regular customers. In addition, retailers quarrel about the number of items that should be defined as the threshold for this division between express and regular checkouts. Nonetheless, minor research has been made on this topic and on the impact express checkouts could have in specific contexts (e.g. store's location or particular customers' profiles).

Hybrid-checkout

As aforementioned, the self-checkouts have the disadvantage of the scanning being slower than in regular checkouts. At peak hours, queues may form in this service and not dissolve as fast as they would at an attended checkout. Therefore, the hybrid-checkout emerged as a solution to tackle this problem, once it is essentially a self-checkout but allows an attendant to take over the scanning from the customers to speed up the process when needed. This solution aims to provide more flexibility to the retailer side, allowing to save floor space while having a good responsiveness to high customers' demand (Hartman, 2008). Nevertheless, it may lead to customers' misperception on why the checkout does not always remain in assisted mode.

Scanning Tunnel

A restricted group of retailers has been testing a novel technological solution related with the checkout area: the scanning tunnel or portal scanner. This device consists on a conveyor where the items are placed and automatically conveyed through a code-reading machine with over 20 cameras that read the items' barcode. The manufacturers of this technology infer that it captures the barcodes with an accuracy of 98% and can speed up scanning time by up to 50% (Vardon, 2015). Besides the efficiency gains and eventual labor cost reduction, it can improve customer satisfaction due to less effort handling products and shorter checkout time.

Multiple-queue vs Single-queue

Besides choosing the checkouts typologies, retailers need to decide on which queue configuration to couple with. To this aspect, there are essentially to main options: multiple-queue and single-queue.

Multiple-queue is the case where there is a separate queue for each service desk or cashier and in this project are acknowledged as the one used for regular checkouts. This system provides flexibility to the customers, since they get to decide on which queue to join according to their criteria (e.g. people in queue). Also, it maintains the illusion that there is more service available (Klukin, 2014). However, multiple-queue presents itself with some drawbacks. Firstly, it can create a sense of unfairness, as a customer that joins a queue later may be served before. Moreover, it stimulates uncertainty and anxiety on customers over which queue to join, diverting their attention to monitoring queues rather than queue merchandise (LineLogic, 2011).

In other perspective, single-queue groups customers in one single line that feeds multiple services or cashiers. Typically, each station is numbered and a digital signage is used to let the first customer on the queue know which cashier he should approach. This system tackles some limitations of multiple-queue, since it mitigates the uncertainty of queue choice and promotes fairness – works in a "first come, first served" basis. Furthermore, the randomness associated with the selection of the cashier reduces the risk of a phenomenon called "sweethearting", that occurs when a cashier neglects to scan some items as a favor to a friend or family member. Nonetheless, single-queue also has disadvantages. On the customer side, a single-queue can appear daunting to whom fail to understand that a longer line doesn't necessarily imply a longer waiting time (Lu et al., 2013; Klukin, 2014). On the retailer side, a single-queue usually requires more in-store space and some studies, like Shunko et al. (2017), indicate that the productivity of the cashiers diminishes under this system due to, for example, not being uniquely responsible for the queue length.

2.2 Queueing Theory

Queueing theory is the study of waiting in a queue in all the various contexts it can occur. It uses queueing models to represent the various types of queueing systems that arise in practice, where formulas for each model indicate how the corresponding system should perform, including the average amount of waiting that will occur, under a variety of circumstances (Hillier and Lieberman, 2001).

These queueing models are very useful for determining the most efficient way of operating a queueing system, balancing the trade-off between excessive cost (too much capacity) and excessive waiting time (not enough capacity).

2.2.1 Basic structure of queueing models

In the basic process usually assumed for the queueing models, customers are generated over time in the *input source*. These customers enter the queueing system by joining a *queue* and then are selected by some rule called queue discipline before being served. This service is next performed by a *service mechanism* and customers exit the system. This macro process is depicted in Figure 2.1.

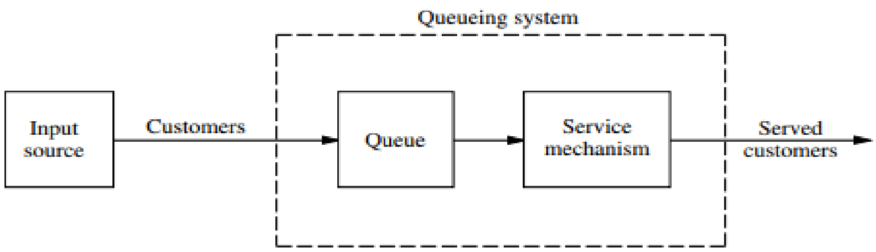


Figure 2.1: Basic queueing elements (Hillier and Lieberman, 2001).

Input source

The input source can be described by a set of characteristics, the first being its size where the population can either be infinite or finite. Due to the calculations' easiness when in an infinite case, this assumption is often also made for very large finite populations.

The arrival's description is also very relevant, being required to depict, for instance, if the customers arrive one by one or in group, if the arrivals are controllable (e.g. pre-scheduled) or uncontrollable (e.g. supermarket or emergencies in a hospital) and its distribution over time. For the last one, the common assumption is that arrivals follow a Poisson distribution and, as a consequence, the time between consecutive arrivals - interarrival time - follows an exponential distribution.

Queue

As Hillier and Lieberman (2001) stated, the queue is where customers wait before being served. This element is characterized by its capacity, which can be finite or infinite, and its number (single or multiple queue). The infinite assumption is very commonly used as a result of analysis' simplification. Also, the order in which customers are selected from the queue to be served - queue discipline - have to be specified, with first-come-first-served as the most typical case.

Service Mechanism

The service mechanism consists of one or more service facilities, each one containing one or more parallel service channels, called servers (Hillier and Lieberman, 2001). In case there's more than one service facility, the customer is served by a sequence of these.

The time a customer remains in a server, known as service time, should be specified in terms of a probability distribution for each server, although most commonly it is assumed to be the exponential distribution and the same for every server. In some cases the service time can be a constant value - known as degenerate distribution - or follow other distribution form (e.g. gamma).

2.2.2 Notation

Unless stated otherwise, the following notation - according to Hillier and Lieberman (2001) - will be the one used from now on:

Table 2.1: Notation of the general queueing theory parameters

$N(t)$	=	number of customers in queueing system at time t ($t \geq 0$)
P_n	=	probability of exactly n customers in queueing system
s	=	number of servers (parallel service channels) in queueing system
λ_n	=	mean arrival rate (expected number of arrivals per unit time) of new customers when n customers are in system
λ	=	mean arrival rate constant for all n
μ_n	=	mean service rate for overall system (expected number of customers completing service per unit time) when n customers are in the system.
μ	=	mean service rate for overall system constant for all n
ρ	=	$\frac{\lambda}{\mu}$ = utilization factor (expected fraction of time the individual servers are busy)

As Hillier and Lieberman (2001) mentioned, queueing theory tends to focus on the steady-state condition, where the state of the system is independent of the initial system and the analytical analysis becomes simpler. The coming notation assumes the steady-state condition:

Table 2.2: Notation of the four fundamental quantities of queueing theory

L	=	expected number of customers in queueing system = $\sum_{n=0}^{\infty} nP_n$
L_q	=	expected queue length = $\sum_{n=s}^{\infty} (n-s)P_n$
W	=	expected waiting time in system (includes service time) for each individual customer
W_q	=	expected waiting time in queue (excludes service time) for each individual customer

Little's formulas

As Little (1961) first proved, there are relationships between the four fundamental quantities - *L, L_q, W and W_q* - enabling an easier calculation once one of them is determined. In case λ_n is constant λ for all n, two relationships are verified:

$$L = \lambda W \tag{2.1}$$

$$Lq = \lambda W_q \tag{2.2}$$

If λ_n is not equal for all n, λ should be replaced by the average arrival rate, $\overline{\lambda} = \sum_{n=0}^{\infty} \lambda_n P_n$.

Moreover, assuming the the service time is also a constant for all n, the following relationships are verified:

$$W = W_q + \frac{1}{\mu} \tag{2.3}$$

$$L = L_q + \frac{\lambda}{\mu} \tag{2.4}$$

Kendall-Lee Notation

For an easier identification of each type of queueing system, Kendall (1951) proposed a notation that consists on 5 parameters *A / B / C : (D / E)*, where:

A: specifies the statistical model of the interarrival time;
B: specifies the statistical model of the service process;
C: is the number of servers;
D: denotes the queue discipline (e.g. FCFS - First Come First Served);
E: denotes the maximum number of customers in the system.

According to Benvenuto and Zorzi (2011), the last two letters are optional and, in case of omission, it is assumed $D =$ FCFS and $E = \infty$. Also, A and B typically assume one of the following distributions:

M: *Memoryless* or *Markovian*, represents the exponential distribution;
D: *Deterministic*, indicates a constant value;
E_r: $Erlang_r$, denotes an Erlang Distribution with index r;
G: *Generic*, indicates a generic distribution.

Additionally, as a rule of thumb (see Benvenuto and Zorzi (2011)), M/M/S systems are much easier to study, especially the M/M/1. Nonetheless, M/G/S have a wider scope of application, since they can assume any service distribution, although this coming at the cost of a more complex analysis. Finally, the G/G/1 system, which can undertake any arrival and service distributions, is barely applied due to some unknown expressions for these situations (for instance, the mean waiting time).

2.2.3 The Birth-and-Death process

The majority of queueing models presuppose that the system's inputs and outputs follow a *birth-and-death* process. In this context, a birth refers to a customer's arrival and a death to a departure of a served customer. The birth-and-death process characterize probabilistically how $N(t)$ changes over time and is based upon 3 assumptions (Hillier and Lieberman, 2001):

1. Given $N(t) = n$, the current probability distribution of the remaining time until the next birth is exponential with parameter λ_n;
2. Given $N(t) = n$, the current probability distribution of the remaining time until the next death is exponential with parameter μ_n;

3. The random variable of assumption 1 and the random variable of assumption 2 are mutually independent, meaning the next transition in the state of the process is either a single birth or a single death.

These assumptions are summarized in Figure 2.2. The arrows show the possible transitions between states and each one have associated its rate. It is important to note that the birth-and-death process assume the system is in a steady-state (equilibrium), which only can occur if $\rho/S < 1$. If this does not happen, the queues would grow infinitely and the system would be unsustainable in the long run.

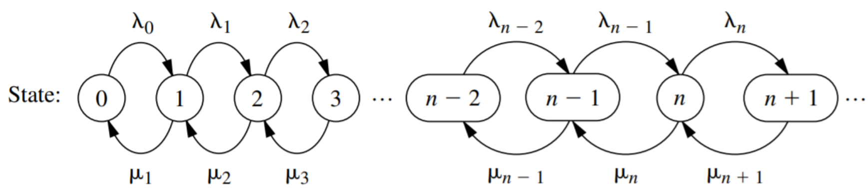


Figure 2.2: Diagram for the birth-and-death process (Hillier and Lieberman, 2001)

The birth-and-death process lead to the **Rate in = Rate out principle**: For any state of the system n, mean entering rate = mean leaving rate. This principle is the base for the obtainment of the balance equations for every system's state (Figure 2.3) and the calculation of all the unknown probabilities P_n (plus an equation specifying that the sum of probabilities must be 1).

State	**Rate In = Rate Out**
0	$\mu_1 P_1 = \lambda_0 P_0$
1	$\lambda_0 P_0 + \mu_2 P_2 = (\lambda_1 + \mu_1) P_1$
2	$\lambda_1 P_1 + \mu_3 P_3 = (\lambda_2 + \mu_2) P_2$
$\vdots$	$\vdots$
$n-1$	$\lambda_{n-2} P_{n-2} + \mu_n P_n = (\lambda_{n-1} + \mu_{n-1}) P_{n-1}$
n	$\lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} = (\lambda_n + \mu_n) P_n$
$\vdots$	$\vdots$

Figure 2.3: Balance equations for the birth-and-death process (Hillier and Lieberman, 2001).

The resolution of the balance equations and the calculation of the four fundamental quantities aforementioned involve summations with an infinite number of terms. Opportunely, these summations have been a subject of study for many years and there are analytic solutions for some specific cases, like for instance the M/M/S system, contributing to its characteristics being more simplistically determined and summarized.

Queueing models with non-exponential distributions

By using the birth-and-death process as basis for queueing models, both interarrival and service times are required to follow an exponential distribution (Hillier and Lieberman, 2001). Although this distribution applies to most of the cases, in some situations, for instance repetitive services where the service times are quite similar, the approximation is not so accurate. Thus, it is important to have queueing models that use alternative distributions. Yet, the mathematical analysis of these models is considerably more complex due to not being able to resort to balance equations. Fortunately, for some models such as M/G/1 or M/D/S, it is possible to use simplifications for the four fundamental quantities (see Pollaczek-Khintchine formula in Ya. Khinchin (1967)).

2.2.4 Special cases

The models previously discussed have all been obtained via analytical resolution, with resort to mathematical manipulations to achieve the expressions to calculate the fundamental parameters. When presented with real-life problems, some hypotheses assumed for the analytical models do not stand or need some adaptation. Here are some examples:

1. Non-stationary arrivals: Traffic queues are typical cases of non-stationary arrivals. The arrival rate greatly increases during peak hours, thinning out in the remaining day periods. This generates continuously growing queues during peak periods, until the arrival rate decreases to values quite below the system's capacity so that queues slowly start to fade out. In this regard, Nelson and Gerhardt (2011) proposed a method to model and simulate non-stationary arrivals in a more intuitive way.
2. Queueing networks: Some queueing systems don't work isolated, being more a sequence of service facilities where customers must receive service at some or all of them (e.g. OR service). Thus, the mathematical expressions do not work in a straightforward way in these situations, leading to a quite extensive and dedicated research on this issue, even though its inherent complexity (see for example Kameda (1984); Mani et al. (1990); Balsamo (1993); Bar-Lev et al. (2009)).
3. Priority queue discipline: For contexts where the queue discipline differs from FCFS, the potential of the queueing models decreases significantly, especially for problems with large dimension (Hillier and Lieberman, 2001).
4. State-dependent arrival rate: In the presence of situations where the arrival rate isn't constant and depends on the current state of the system, simplified expressions for the resolution of the balanced equations no longer apply. In Hillier and Lieberman (2001) is presented a mathematical manipulation of the steady-state results based on a pressure coefficient that enables to incorporate this condition. Yet, approximate solutions[1] are only tabulated for a limited range of coefficients and number of servers, inhibiting a complete analysis on

[1] Available in W. Conway and L. Maxwell (1962)

the impact of this limitation. Furthermore, other authors, like Abouee-Mehrizi and Baron (2016), discussed adjustments of M/G/1 models to represent a state-dependent context, although they lack an explicit explanation on how to generalize these adjustments for models with multiple servers.

With regard to circumstances like these, simulation is largely used to reproduce the behavior of the systems over time. Despite the inconvenience of higher costs and longer developing period, simulation possesses unquestionable benefits at the level of modeling flexibility.

2.3 Simulation

Simulation is defined as the process of designing a model of a real system and conducting experiments with such model, with the intent of either understanding the behavior of the system or evaluating various strategies for the operation of the system (Shannon, 1975). This technique has surfaced to tackle real-world problems where it is not affordable to experiment with real objects in order to find the right solutions due to, for example, being too expensive or just impossible to make physical changes. It also grants an analysis of systems where methods such as analytic calculation or linear programming usually fail (Grigoryev, 2015).

Supported by high computational capacity, simulation allows to test alternative scenarios and designs through a time horizon, recording and comparing the performance of each one of them before choosing the best. Notwithstanding, it requires the gathering of accurate data and a solid validation of all the assumptions made in the model to assess its representativity of the real system. Furthermore, the development of the simulation model can be quite costly and time-consuming, indicating it should not be used when there is a less expensive procedure available that can provide the same information (Hillier and Lieberman, 2001).

Using Grigoryev (2015) definition, in simulation a method is a framework used to map a real world system to its model. In Figure 2.4 is represented the difference between the three methods in terms of abstraction level. The increasing use of simulation led to its evolution and diversification, where, to address a wide range of applications, three different methods were developed: system dynamics (SD), agent-based modeling (AB) and discrete-event modeling (DE).

System dynamics is generally used in long-term strategic decisions, assuming a higher level of abstraction that provides a more holistic view of the system. It ignores the fine details of the system, such as particular events or elements' characteristics, and represents a general picture of a complex system. It is described as a set of stocks (characterizing the system state and which can be, for example, people or materials) and flows (rates at which the states changes). This approach is commonly used for situations where there are properties of the whole that can't be found by connecting properties of the individual elements (Grigoryev, 2015). Aside from the importance of this approach, DE and AB are more focused on medium and medium-low abstraction levels (tactical and operational) and thus more aligned with the purpose of this study. There are differences and similarities between these two approaches that are justifiable to analyze.

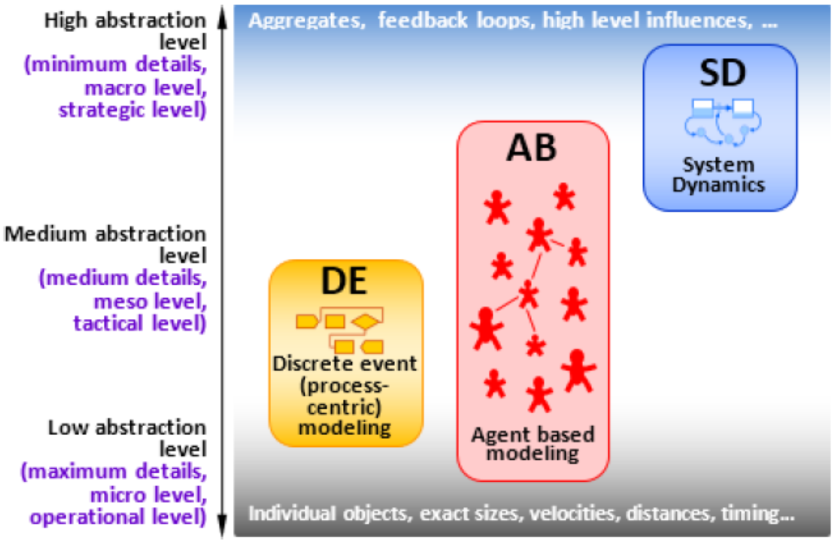


Figure 2.4: Types of Simulation (Grigoryev, 2015).

AB is a relatively new method compared to SD and DE. Its appearance came to offer the modeler another way to look at simulation: with the increasing complexity of systems it may be hard to know how the system behaves or to identify the process flows, but the modeler may have insights of the individual behavior of each entity. Afterwards, these entities are connected and put in an environment where they can interact with each other. The system's global behavior results from these interactions of many concurrent individual behaviors (Grigoryev, 2015). Hence, this method follows a bottom-up approach, where a system is modeled as a collection of autonomous decision-making entities called agents who continuously interact, providing a strong insight about the real-system internal dynamics (Bonabeau, 2002). One of the main advantages of this method is the flexibility it offers to the modeler, ranging from the easy creation of new agents to the different levels of aggregation and complexity between them (Baldwin et al., 2015).

DE requires the modeler to think about the system as a sequence of operations performed by agents. The model is represented as a process flowchart where blocks reproduce operations that have a service time and an agent arrival time that are usually stochastic. Thus, the model itself will be stochastic, meaning multiple replications should be run to produce meaningful outputs (Grigoryev, 2015). Opposite to AB, DE follows a top-down approach, representing a collection of events that impact the system instead of the individual agents in it. Therefore, this method is a good choice when it is known how the system reacts to a given situation (Baldwin et al., 2015). Other advantage of DE is its simplicity in creating the model, since it is based on inputs, states and outputs and the modeler is not required to know the internal workings of the system. As such, the modeler codes the model to react according to actual observations (Belli et al., 2006).

There is also a specific type of DE, the Pedestrian Modeling (PM). This type intends to represent crowds, people and traffic dynamics. According to Grigoryev (2015), pedestrians follow basic rules that have been determined by theoretical studies, based on the aforementioned events that are characteristic of DE. PM can be especially important when designing a supermarket or railway station, where travel time and pedestrians' congestion are aspects to be considered.

With all this multitude of different methods available, simulation has gained its space as a paramount tool to bolster complex decision-making. As Zeigler et al. (2000) pointed out, modeling and simulation's area of influence is as multifarious as the enormity of real-world problem possibilities, enhancing the relevance of simulation as a problem-solving methodology in this context. To carry out a methodology of this kind, one of the most trendy modeling software is Anylogic®, allowing the implementation of all the simulation techniques previously introduced. According to (Hao and Shen, 2008), it is the only hybrid simulator in the market that delivers complex modeling capacities in an all-in-one solution. Anylogic® combines a DE environment with integrated active agents, ensuring a higher controllability over the simulation components.

One of the main assets of Anylogic® is related to being an object-oriented tool, going along with the strong evolution of programming languages towards an Object Oriented Paradigm (OOAD). Furthermore, this software is 100% coded in JAVA, which is acknowledged as a easy-to-learn language and independent of the platform where it is used. Being backed up by JAVA opens the possibility of customizing any entity's properties and functions, providing flexibility and friendliness to the users (Hao and Shen, 2008). Once Anylogic® presents itself with a number of leverage points over the alternatives available, the simulation project was exclusively constructed in this software.

2.4 Simulation in retail stores and checkout area

In the process of dimensioning a checkout area, not only it is important to choose the types of checkout that are going to be used but also to select the best combination between them in order to optimize the checkout process efficiency. Aligned with this perspective, some researchers have been applying quantitative approaches, mainly based on simulation, to study the best layout of the checkout area considering multiple factors as, for instance, the customer behavior in the store.

J Williams et al. (2003) attempted to use simulation to define the staffing policies at a retail store with the objective of minimizing the waiting time of a customer in the checkout. Alvarado and Pulido (2008) studied the optimal combination between cashiers and baggers that minimizes the waiting time in the checkout, gathering data from specific supermarkets in Colombia. The results indicated that when 40% of the cashiers are open, having dedicated baggers have a significant impact in decreasing the customers' expected time in the system. However, this study may vary between locations, meaning that the same method applied to other instance might lead to different outcomes. Yamane et al. (2012) conducted a simulation study to examine the best layout of the checkout area in a specific store in Japan. Nonetheless, the main focus of this study was to mitigate congestion issues instead of minimizing the waiting time. Moreover, Rossetti and Pham (2015) examined the impacts of two factors in customers' total time in the system: on the one hand the customer's criteria when picking a checkout lane and, on the other hand, the payment process being separated from the checkout station (where the scanning and bagging occur). The results showed that the lane choice criteria did not have a significant impact in the checkout time. By contrast, having the payment separated from the checkout station caused a significant drop in the average waiting time.

To understand customer behavior in stores, which may impact the checkout area configuration, Chan et al. (2010) built an "hybrid discrete-continuous simulation model with proactive, autonomous, and intelligent entities" that is based in AB models described in section 2.3. This approach intends to improve the customers' shopping experience as well as to optimize the many processes that occur in a store, checkout included. Additionally, Terano et al. (2009) used the AB method to study the customers' movement in a store along with their shopping behavior. Data from real customers was analyzed to serve as input to the simulator, allowing to investigate the effect of promotions in customer in-store flow. The results led to a set of recommendations on where to locate promotions in the store. Schwenke et al. (2010) tried to model the customers in a different perspective, assigning them 3 main actions: think, move and act. This technique provided a way to study checkout configurations and introduce simulation as a more realistic manner to support sales forecasting. Finally, and purely focused on the customer perspective, Buell (2017) investigates whether people exhibit last place aversion in grocery stores' queues and its implications. The study revealed that waiting in last place diminishes wait satisfaction and increases the odds of switching or abandoning queues, being a topic worth of concern for the retailers.

Broad-spectrum, there have been manifold investigations attempting to use simulation to address checkout area related issues. However, each focuses on different elements of the business scope, without apparent interconnection between them to enable result's benchmark. Additionally, each research is highly tailored to the context where it was developed, containing specificities that limit its generalization to other retailer's background. Moreover, it is believed that the absence of articles describing integrated methodologies used by big world retailers is connected to non-disclosure agreements and corporate confidentiality with the intent of having an advantage over the competitors. With this in mind, this work presents a two-fold methodology based on queueing theory and simulation to address the integrated problem of defining the configuration of the checkout of a store. The results presented are obtained by applying this methodology in the context of a food retailer.

Chapter 3

Problem Description

Planning the configuration of a checkout area poses itself as a difficult problem, mainly since increasing the number of checkouts available to reduce queue sizes is not advisable in a straightforward manner due to the costs involved (Serasinghe and Vasanthapriyan, 2017). If from the customer standpoint, long waiting times in the checkout downgrade the perceived quality of the shopping experience, from the retailer's perspective having underutilized checkouts represents a large investment in the acquisition and maintenance of equipment that is not being efficiently used. Moreover, the specific characteristics of each store and its customers' profile inhibits the use of a "one fits all" kind of solution. The main challenge is to have placed in the stores the minimum number of checkouts of each typology that satisfy the demand and expected service level.

The problem acquires a higher complexity when the wide range of checkout solutions is incorporated. If defining the total number of checkouts to be placed in a store is not a trivial matter, specifying how this number is divided among the various checkout typologies available is an even more complex decision. Furthermore, the considerable fluctuations of demand and arrival rates along a day or week (causing the differentiation between peak and off-peak periods) increases the difficulty in finding the ideal configuration for each store.

This chapter intends to assess the magnitude of the problem at hands. In the first section, the stores' scope is addressed, where a description is given of the different stores' formats the case study retailer has, as well as the identification of the checkout typologies currently being used. In section 3.2 it is described the checkout dimensioning process currently in use in the studied company, pointing out its most noticeable limitations. Next, in section 3.3, a preliminary quantitative assessment of the opportunity for improvement is conducted purely based on the data originally available. Finally, in section 3.4, a holistic description of the proposed approach to improve checkout management is provided.

3.1 Stores' scope

This project covers the full scope of the retailer's stores, which totals approximately 250 establishments. These stores are acknowledged to follow the retailer's store structure, which divides them into three formats depending on their characteristics. With regards to the formats, they can be described as following:

Format A Stores of large dimension and with a wide range of products, usually located in major cities and areas of dense population;

Format B Medium-sized stores, generally placed in the surroundings of major cities and with a medium range of products;

Format C Smaller stores, with a narrower range of products and located to serve specific population hubs.

The format of a store has a significant influence in the checkout solutions available in each store. Stores of format A, due to its bigger size, are able to offer a wider diversity of checkout alternatives, while stores of format C, in general, possess only one checkout typology (CT).

In regard to the CTs described in chapter 2, the case study retailer has in its stores the following alternatives:

1. Regular checkout in multiple queue (one queue per checkout);
2. Regular checkout in single queue (one queue for a group of checkouts);
3. Self-checkouts in single queue;
4. Express-checkouts in multiple queue;
5. Self-scanning with payment in single queue.

The first three alternatives presented are the core typologies of the retailer and will be the main focus of this study. An important remark is that the alternatives 4 and 5 are only present in a very small number of stores. Also, the data available from the operation of these two alternatives is not standardized and clearly identified, making it infeasible to develop an accurate analysis.

3.2 Current dimensioning process

The case study retailer uses a very simplistic process for managing the decisions regarding the checkout dimensioning. This process can be divided into two subcategories: (i) Define the checkout configuration for a new store; (ii) Reconfigure a current store. After the description of these two subprocesses, the identified limitations are enumerated.

3.2.1 Configuration for a new store

The current process for defining the configuration of a new store is empirical. It starts by selecting another store that has a profile similar to the one that is expected for the new store. This selection is based on aspects such as geographical location, store's size and format, and customers' characteristics. However, it encompasses little or no analytical reasoning.

After having selected the "mirror" store, its annual sales volume is considered to be a good approximation for the new store (with possible fluctuations due to factors that are, once more, identified from past experience). From the expected sales volume, the number of checkouts is defined to be proportional to that value to each type of checkout. For instance, a regular checkout is equivalent to a percentage of total sales while a self-checkout corresponds to half of that value (as a result of an inferior service rate). The mix of typologies to be used is then determined by resorting to the experience and knowledge of the decision-makers, in a way that takes into account the limited space available in the store.

3.2.2 Reconfiguration of a current store

The reconfiguration of a store can be triggered by two different aspects. On the one hand and more commonly, a need for remodeling due to the equipment wear-out period, which depends on the CTs. In this case, the worn out equipment can be either replaced by one of the same kind or by one from a different typology. On the other hand, each store's manager may request an alteration of the configuration to the operations department, due to a perception of a change in the customers' requirements or a shift in the demand for a specific type of checkout.

3.2.3 Limitations

An analysis of the processes previously described decoded some inherent shortcomings:

No analytical support Currently both processes are of an empiric nature, lacking an analytical support of the decisions made regarding the checkout dimensioning. These decisions are based on the know-how of the stakeholders and on benchmarking with similar stores, with the latter also being a result of knowledge gathered from experience.

Lack of performance evaluation The fact that no monitoring is conducted on the stores' checkout performance leads to an absence of a prioritization on the urgency of reconfiguration between the stores. In case there is a large number of managers' requests for an adjustment of the checkout layout, the company has no way of prioritizing the investment in order to allocate the available budget to the stores with most pressing issues. Also, this fact contributes to the nonexistence of a trigger for checkout reconfiguration besides the wear-out period. The checkout performance could be evaluated by, for instance, the utilization ratio of the checkouts, which is obtained by determining the percentage of the checkouts available in the store that are actually used during store operation. The performance analysis is further detailed in section 3.3.

No risk assessment A change in the configuration of the checkout area implies a large investment for the retailer, both in the acquisition of new equipment and in operational and logistic expenses for removing a current checkout from the store and replace it by a new one. Thus, there is a high risk associated with this type of investment and, currently, the retailer does not possess a way of anticipating and simulating the impact of a change in the configuration. An alteration that was originally expected to have a great impact on the performance can easily disappoint if it ends up not being particularly suited to fit the store's needs. Moreover, the retailer has no way of testing the impact emerging checkout solutions could have if implemented in the stores.

Incorporation of peak hours The methodology currently in place fails to elucidate on how the checkout dimensioning should incorporate the peak hours characteristics of each store. To which period should the dimensioning be made: the day of the year with higher customer inflow or an average scenario over the year? The first possibility may efficiently respond to the most crowded days, but during a large percentage of the year it may be greatly over-dimensioned, implying larger maintenance costs and occupying space that could be used to exhibit products. On the contrary, dimensioning to an average situation may lead to a more frequent occurrence of longer queues, affecting the customer evaluation of the store. More than defining this reference period, the retailer is lacking a tool to assess the performance of the checkout area during the whole year.

3.3 Preliminary performance assessment

As mentioned in the beginning of this chapter, a quantitative assessment of the opportunity for improvement was performed. Since the data available concerning the checkout process refers only to tasks performed in the checkouts' operating system, there is no basis for an analysis of indicators characterizing the customer experience, such as waiting times in queue or queue sizes. That said, the data available allowed an analysis on the usage of the checkouts placed in each store through the calculation of KPIs such as the utilization and occupation ratios, which are now described.

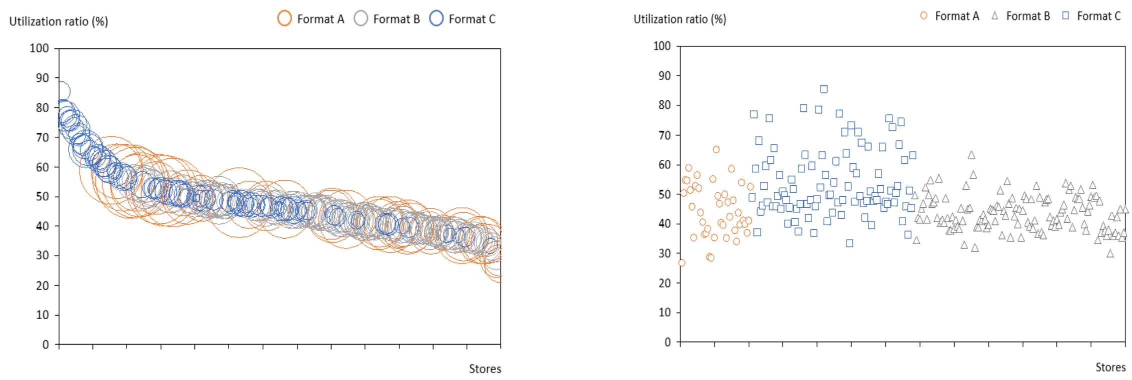
The utilization ratio reflects the percentage of the checkouts available in each store that were actually used during store operation (see Equation 3.1). This indicator can be calculated at the hour-level of granularity, allowing to capture variations even along a single day. Its output provides guidance on the dimensioning effectiveness, potentially indicating if the checkout area is over or under-dimensioned.

$$Utilization\,ratio = \frac{Number\,of\,checkouts\,used}{Number\,of\,checkouts\,available} \tag{3.1}$$

In addition, the occupation ratio represents the percentage of the time each checkout was open that were indeed serving customers, as represented in Equation 3.2. This KPI can also be computed at the hour-level, shedding light on a more operational degree of performance related with the checkout opening criteria.

$$Occupation\,ratio = \frac{Time\,serving\,customers}{Total\,operating\,time} \tag{3.2}$$

Starting with the former, Figures 3.1a and 3.1b depict the average of the utilization ratio per hour of each store in 2017. Each circle represents a single store and its color is representative of its format.



(a) Ranked average utilization ratio

(b) Dispersion of utilization ratio values by format

Figure 3.1: Average utilization ratio for each store in 2017

In Figure 3.1a, where the size of each circle is representative of the sales volume of the corresponding store, the utilization ratio is illustrated in descendant order of performance between the stores. The variance of values is quite significant, ranging from 30% to 80%, indicating that some stores have a relatively good performance but there is still margin for improvement in multiple others. The top performance values appear to mainly belong to format C, although Figure 3.1b, which depicts the dispersion of utilization ratios values inside the same format, indicates the performances have large variability even between stores with similar characteristics. Also, stores of format A, which represent the highest sales volumes, cover the full range of values, not suggesting the existence of a clear pattern for the causes of inefficiencies. Finally, format B is associated with lower variability between values and there are very few examples of a utilization ratio over 50%.

Bottom line, a noteworthy number of stores has an average utilization ratio below 40% and most likely the stores possess an excessive number of checkouts. However, an important remark is that these are mean values for the year, which might mask situations where the utilization ratio was significantly higher or lower than the mean. To try to identify if its a recurrent situation, the worst performing stores of each format were selected to carry out a more detailed analysis on fluctuations of this KPI throughout the year. Results are shown in Figure 3.2.

As it is possible to observe, there is a variation in the weekly average of the utilization ratio per hour during the year, which is common to all formats. Using the store of format C as an example for the following remarks, there is a gap of 15 percentage points between the week with the lowest utilization and the one with the highest, being a good indicator of the variation aforementioned. Also, is it perceived that many weeks have a value of this KPI that is above its year's mean for this store. As a result, special attention should be given to not use the mean as a single indicator of the current performance of the store.

In other perspective, it is common to all the 3 formats that the weeks with highest utilization

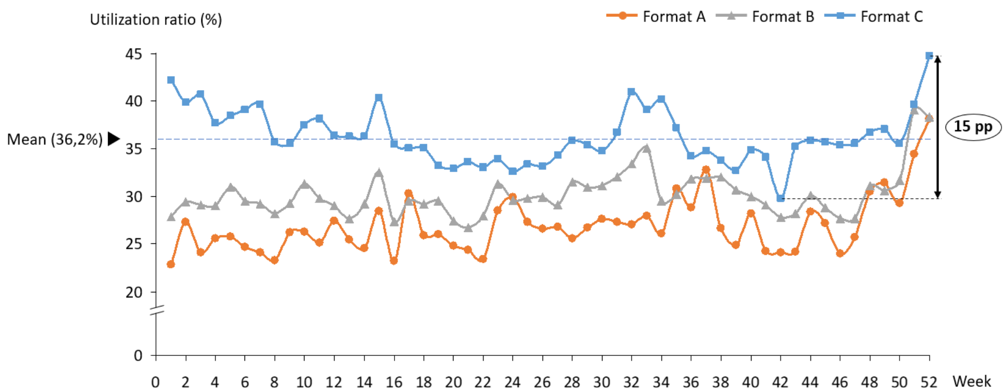


Figure 3.2: Detailed analysis of the weekly average of the utilization ratio per hour throughout the year

ratio occur at the end of year, matching the Christmas and New Year's Eve season which usually yield a higher customer inflow to the stores. However, the peak weeks for each store are subordinated to its features, such as the region where it is located. For instance, a store placed in a region well-known for its affluence during summer vacations will most likely have its peak weeks in the course of this season. Hence, the key note to be withdrawn is that each store's unique characteristics have to be taken into account while addressing its dimensioning instead of assuming the same reference for the whole stores' universe.

An aspect to also take into consideration is the different CTs available. In some stores, especially the ones of format A, part of the regular checkouts are used in a multiple-queue system and others in a single-queue one, with differences in terms of utilization between them that may be hidden when only contemplating each store's average value. In this regard, Figure 3.3 shows the results from an analysis conducted on the stores of this format with higher volume of sales, so that an eventual discrepancy in performance between the two CTs would be outlined. An important remark is that self-checkouts, since they do not require a cashier to operate, are always "open" and, consequently, the utilization ratio equals 100%. Thus, this CT will not be included in this preliminary analysis.

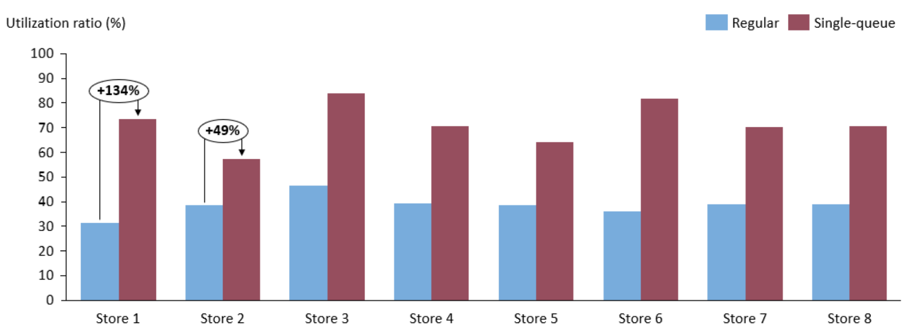


Figure 3.3: Differences in utilization ratio between regular and single-queue checkouts

A quick examination of the graph suggests that single-queue checkouts systematically outperform multiple-queue ones in the matter of utilization. In fact, in some stores (e.g. store 1) the utilization ratio of the former more than doubles the one of the latter. This can be a potential indicator that stores preferably open the single-queue checkouts rather than opening multiple-queue ones.

Regarding the occupation ratio, Figure 3.4 shows the differences in this indicator between the three store formats when intersected with the utilization ratio.

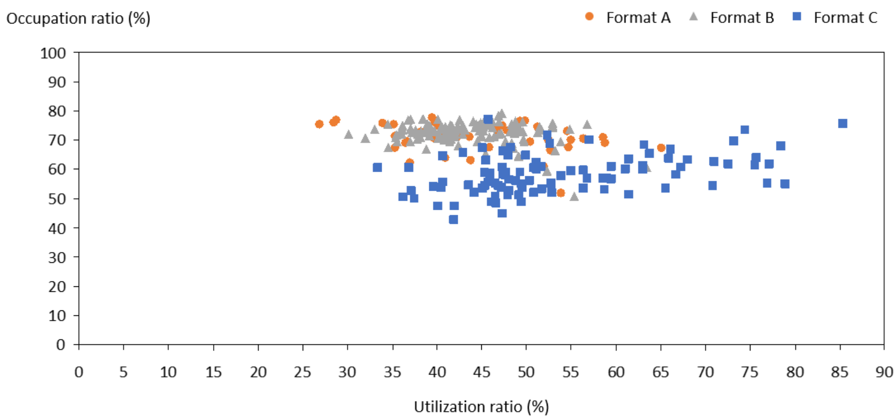


Figure 3.4: Occupation ratio vs utilization ratio for the three formats

While in formats A and B the occupation values are somewhat stable around 70% for any utilization ratio, in format C this indicator has a wider variability and lower values, pointing out a possible disparity and inefficiency in the checkout opening criteria. These lower values can also be occurring due to lower customers' inflow rates in stores of format C, which may lead to the existence of checkout idle time even with only minimal resources being used (one checkout open in the entire store). However, the operational component of the checkout area, particularly the checkout opening criteria, will not be the main focus of this project.

3.4 The proposed approach

The proposed approach to address the checkout dimensioning problem can be described by the flowchart presented in Figure 3.5. The first step encompasses the data retrieval and processing in order to be used in the remaining methodology. Besides gathering the current configuration of each store and analyze its actual utilization, a crucial phase is the input parameters' modeling. This phase will consist on modeling the key parameters that represent the store's and each CT most differentiating characteristics, such as the customers' basket size or the average scanning time. These parameters will then be used in the simulation model, allowing an accurate representation of the stores' features while conferring variability and reduced computational effort to the simulation. For the case of dimensioning new stores, a "mirror" store should be used to model each of the parameters, including the expected demand.

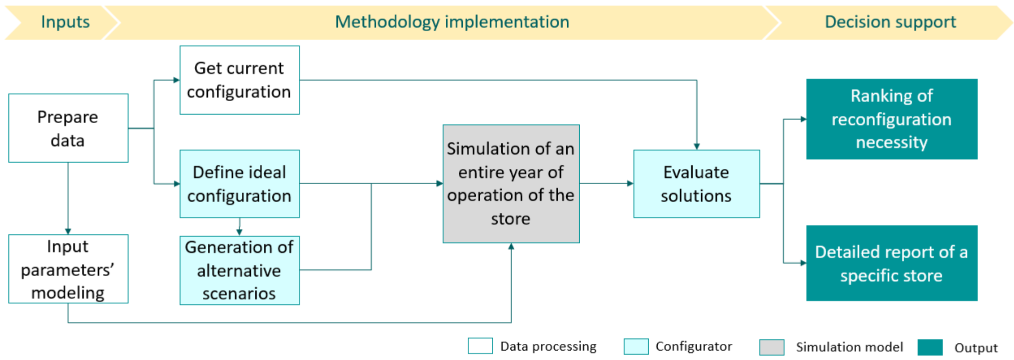


Figure 3.5: Proposed approach to address the checkout dimensioning process

The second major step is the development of a configurator. This methodology is based on queueing theory and will be responsible for guiding the simulation model into which may be the ideal configuration for each store according to the customers' demand. This dimensioning will not be made for the period with highest demand, since it would lead to low utilization ratios during the rest of the year. Instead, the indicated configuration will be the one that allows the attainment of an annual service level of 99% in terms of working hours fulfilled while respecting a maximum queue size. With the preliminary ideal configuration defined, two additional tasks are assigned to the configurator: the evaluation of the solution's cost/benefit by comparing it with the current configuration, as well as the definition of alternative scenarios (e.g. demand variation) to be later tested using the simulation model.

The third and final stage is the development of a simulation model representative of the checkout process. This model will allow to test the robustness of the proposed ideal configuration for an entire year of store's operation without the inherent simplifications of queueing theory, assessing its impact in the determined KPIs, namely the average queue size, the utilization ratios and the annual service level. The final recommendation regarding the proposed configuration for a specific store is only conducted after the analysis through the simulation model. Also, the simulation will enable to evaluate the effects of the alternative scenarios defined by the configurator.

There will be two outputs to support the retailer's decision-making regarding the checkout dimensioning: a ranking of reconfiguration necessity of all the stores and detailed reports of the dimensioning of specific stores. Regarding the first, a preliminary ranking based on the gap between the current configuration and the one suggested as a result of the proposed methodology is provided. The difference between the solutions' fixed costs (equipment and maintenance) will also be presented. The main goal of this ranking is to shed light on the stores that need a more urgent and detailed analysis. Secondly, through the interaction between the configurator and the simulation model, detailed reports of the dimensioning of specific stores will be provided, enabling a more complete assessment of the solution (for example, the attainment of the solution's total costs, including personnel). Also, the simulation of alternative scenarios (for instance, an increase in the store's demand) will allow to test the robustness of the proposed configuration.

Chapter 4

Methodology

In this chapter it is described the methodology developed to overcome the problems previously identified and to improve the checkout management process. After the summary presentation of the proposed approach in section 3.4, each stage of the methodology is now depicted in detail in sections 4.1 to 4.3, encompassing the input parameters' modeling, the configurator and the simulation model, respectively. Once again, the data available and used in this work refers to the checkout transactions of all stores that occurred during the year of 2017. To refer that, in Appendix A, it is provided a preliminary presentation of the decision support system that is going to be delivered to the retailer as an interface to the implementation of the proposed methodology. Notwithstanding, this system is still in an early stage of development.

4.1 Input parameters' modeling

The first step of the methodology chosen to tackle the identified problems is to model the input parameters. The goal is to find a way to accurately represent the parameters that will be used further on to characterize each store's profile. Each parameter has its unique characteristics and explanatory variables, requiring the individual description provided below.

4.1.1 Basket size

The basket size is a fundamental input to the simulation model, influencing other parameters that will be later described. It consists on determining the number of items a client carries when arriving to the checkout area, a number that will impact the choice of the CT, the processing time, among others.

The modeling of this parameter should capture the characteristics of each store with regard to its customers' profile, at the same time it confers variability to the simulation, allowing to test the robustness of the proposed solutions. That said, the observation of historical data induced an adjustment to a log-normal profile, which was confirmed as able to seize the aforementioned characteristics. An example of the data's fitness to this distribution form is shown in Figure 4.1. The bars represent the number of transactions with each number of items in a specific day while the line represents a log-normal distribution that fits these records. The objective is to have a probability of occurrence of each basket size.

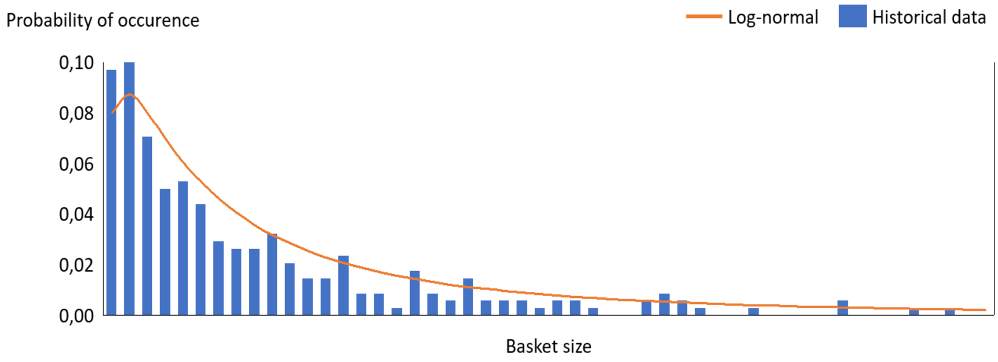
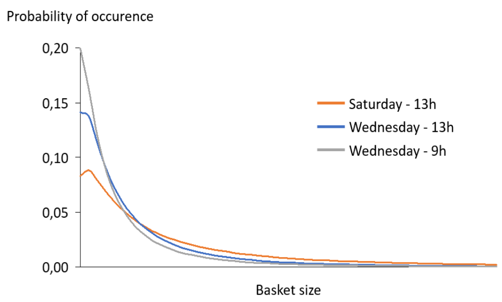


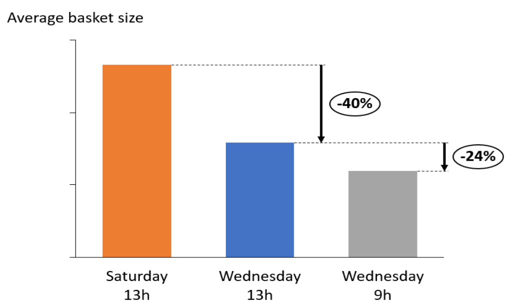
Figure 4.1: Adjustment of log-normal distribution to the historical records of basket sizes

It is possible to observe that the probability is much higher for smaller basket sizes, capturing well the higher number of occurrences in this range. Also, for very large basket sizes the probability is close to zero, being a good representation of the real situation.

Besides the characteristics of each store, there are other factors that originate variations in the basket size profile inside the same store. Through the analysis of the data, the day of the week and the hour of the day were acknowledged to have a significant impact in the values of the distribution that models this parameter. These differences are evidenced in Figures 4.2a and 4.2b, leading to the decision of using a distribution for each combination of store-weekday-hour.



(a) Variations in log-normal distribution profile



(b) Variations in the average basket size

Figure 4.2: Variations of the basket size with day of the week and hour of the day

4.1.2 Choice between checkout typologies

Upon a customer arrival at the checkout area, a decision is made regarding which checkout he'll choose to perform this process. This choice is based on the current state of the system (queue sizes) but also taking into account the customer's preferences over the CTs available in the store.

For instance, a hypothetical scenario can occur where there is a smaller queue in the self-checkout, yet the customer chooses the regular checkout due to his personal predilection for this typology.

In an endeavor to accurately represent the customers' real conduct, probabilities of choosing between assisted or self-service were estimated from historical records and split into two groups according to the basket size: 1) Up to 15 items; 2) More than 15 items. This threshold is the division point conventionalized by the retailer between customers using a basket or a cart to carry the items, an aspect that conditions the typology choice. The estimation of each probability is summarized in Equations 4.1 and 4.2.

$$P_{self} = \frac{\text{Number of transactions in self-service}}{\text{Total number of transactions}} \tag{4.1}$$

$$P_{assisted} = 1 - P_{self} \tag{4.2}$$

Figure 4.3 graphically describes the methodology proposed for this parameter. Between the typologies available in assisted service, the choice will be made according to the state of the system.

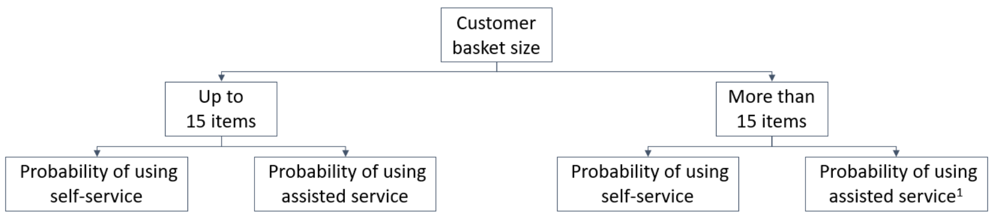


Figure 4.3: Checkout typology (CT) choice criteria

4.1.3 Arrivals to checkout queues

As already referred, the data available is relative to checkout transactions, which begin with the scanning process. Hence, the number of transactions per hour in a store is biased by the number of checkouts open at that moment. In order to make an accurate analysis of the checkout management, urges the need to know the moment the customer arrived at the queue, instead of the moment he initiated the checkout process.

The approach followed attempts to have an estimate of the average in-store time of each customer in function of the number of items in two shopping contexts: off-peak hours and peak hours. A peak hour is one where there were no records of time intervals between two transactions larger than 2 minutes, indicating that the checkouts were under pressure. This approach for the identification of peak hours is also used for the modeling of the between-ticket time explained in subsection 4.1.6.

[1] Although a customer carrying a cart is not allowed to use the self-checkout, it is possible to use this CT while carrying two baskets and, thus, more than 15 items. This probability intends to represent these situations.

Then, for a customer with a certain basket size, the difference between the in-store time in peak hours (1) and off-peak hours (2) outputs an estimate of the time spent by the customer in the queue (3), under the assumption that during off-peak hours customers don't have to wait in a queue to get served. Figure 4.4 describes the approach in a graphical way. To attain the estimates of the average in-store time in both periods, records from customers who printed their discount coupons when entering the store and used them in the checkout process were utilized and fit into one regression model for each period.

Finally, Figure 4.5 demonstrates the difference in the number of customers arriving at the checkout along one day when considering two different approaches: the moment of arrival at the queue (by applying the aforementioned methodology) and the beginning of transaction. Although being slight, these differences may have a significant impact in the dimensioning due to a higher concentration of customers in a certain period of the hour.

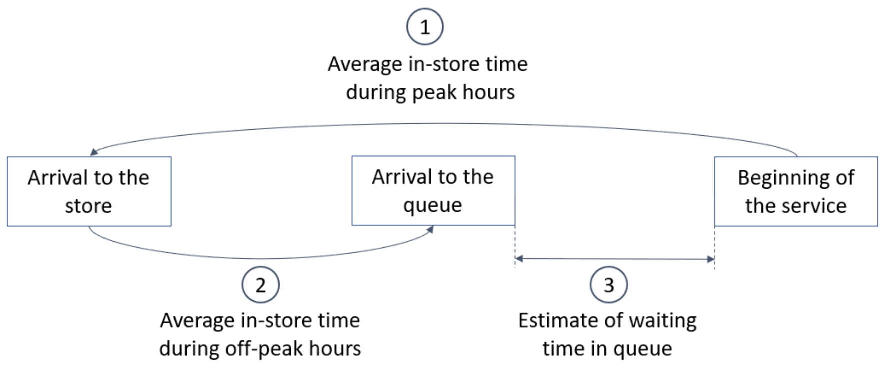


Figure 4.4: Approach to estimate waiting time in queue

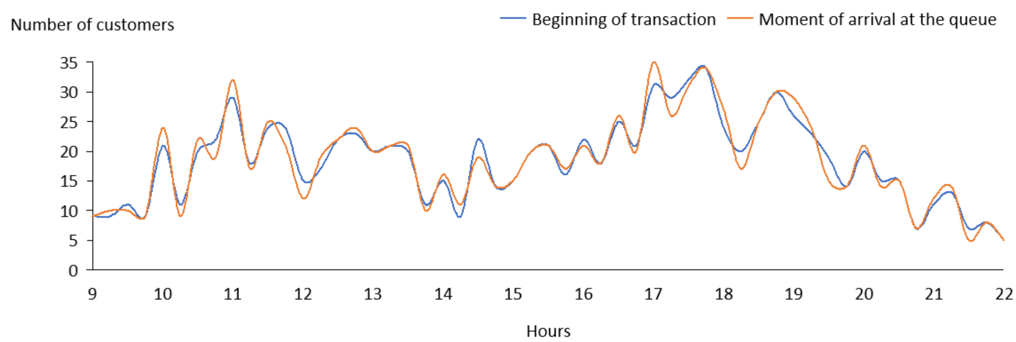


Figure 4.5: Differences in the number of customers between the moment of arrival at the queue and the beginning of transaction

4.1.4 Scanning time per item

A crucial variable to determine the checkout process duration is the scanning time per item, which regards the time it takes to scan each item the customer is buying. Analyzing the data available, the conclusion is this parameter depends on the cashiers' productivity, item's type (size and weight) and CT. Since it is utopian to have a standard scanning time for each different item, a distribution was obtained to approximately represent the data at store-level of granularity. This way, the effect of both productivity and general diversity of items usually bought in each store is accounted for, while some variability is also conferred to this parameter. Once again, the log-normal distribution was chosen as the most suited for the modeling. Figure 4.6 illustrates the fitness of the log-normal distribution to the historical data of scanning time per item in one store for regular checkouts.

Nonetheless, there is a particular aspect that characterizes one CT present in the stores: in the self-checkout, the scanning operation is done by the customer himself, leading to a lower productivity in this process. Therefore, for each store were assigned two different log-normal distributions for the scanning time: one for the self-checkout and other for the remaining CTs.

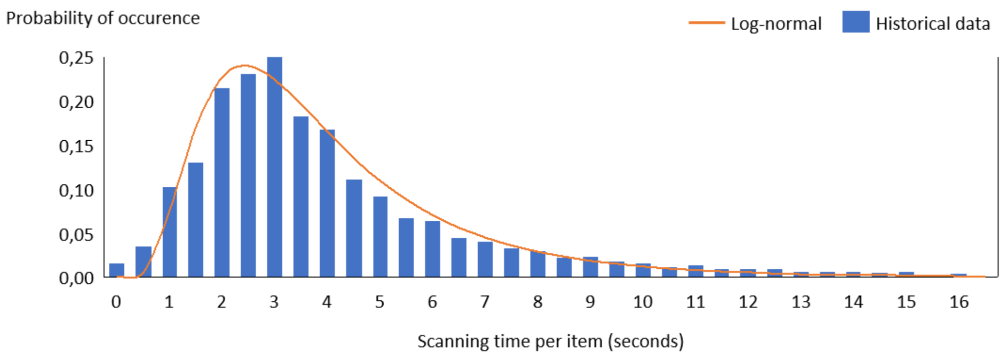


Figure 4.6: Adjustment of log-normal distribution to the historical records of scanning time per item

4.1.5 Post-scanning time

After the scanning operation is complete, the post-scanning process initializes, where the customer not only pays for his items but also may occur other payment related issues, such as the use of loyalty card and discount coupons. As a consequence, the modeling of the post-scanning time involves two steps. Firstly, the definition of the customers' post-scanning profile, consisting on the probability of choosing each payment type and using or not the loyalty card and coupons. Secondly, the attainment of the impact each element of the profile has on the post-scanning time.

Post-scanning profile

Through the analysis of historical records, no significant differences were identified in the post-scanning profile between the various CTs or the basket size threshold previously mentioned. However, relevant discrepancies were encountered between stores, indicating a store-level of granularity for this parameter is pertinent.

The profile division follows a tree-structure, with conditional probabilities calculated for each branch. With regards to the payment type, three were considered: cash, card and others, with the latter including all the remaining possibilities. After the definition of the payment method, the customer may use or not the loyalty card and, in the positive case, may also use discount coupons.

Multiple regression

In order to have a prediction of the post-scanning time of each customer, the choice fell back on a stepwise regression approach. The idea is to quantify the impact of the variables that most affect this parameter and to capture the correlations between them.

The intercept of the regression (γ_{0_s}) was obtained considering a base scenario where the payment is done in cash and without loyalty card (and, consequently, without the use of coupons). The post-scanning time is then modeled using the following formula (Equation 4.3):

$$T_{is} = \gamma_{0_s} + \gamma_{1_s} \cdot C_i + \gamma_{2_s} \cdot L_i + \gamma_{3_s} \cdot D_i + \gamma_{4_s} \cdot S_i \tag{4.3}$$

where:

T_{is}	=	Post-scanning time of customer *i* in store *s*
C_i	=	Dummy variable which takes value 1 when customer *i* pays with credit card
L_i	=	Dummy variable which takes value 1 when customer *i* uses loyalty card
D_i	=	Dummy variable which takes value 1 when customer *i* uses discount coupons
S_i	=	Basket size of the customer *i*
γ_s	=	Regression coefficients of store *s*

An important remark is the addition of the S_i component, in an attempt to capture the effect of the items' bagging done by the customer, an aspect that impacts his speed during the post-scanning.

The accuracy of the regression is considered as not much satisfactory, with an average R-squared of 40%. These results can be mainly explained by three aspects. Firstly, there is an extensive variability in the post-scanning time due to the unpredictability associated with customers' behavior. Additionally, variables that could capture that behavior in some extent, like the customers' age, are protected by privacy policies and thus cannot be accounted for in the regression. Finally, the variables used in this regression may not impact the post-scanning time in a completely linear way. Nonetheless, the results obtained were discussed with the retailer's team and considered to be plausible, based on their know-how of the day-to-day operations.

4.1.6 Between-ticket time

The between-ticket time will determine the intrinsic time between the conclusion of a customer service and the beginning of the scanning process of the subsequent one in the queue. Its main goal is to seize the effects of aspects that delay the initiation of the next customer service and are not accounted for in the data, such as remaining repercussions of a slow bagging process or the time of printing the ticket and hand it to the customer. This parameter has a notorious underlying randomness and varies significantly between the self-checkout and regular ones, since the former has a predefined machine setup time before starting other service. Also, the basket size has a slight impact on this parameter, hence a differentiation is also done in this regard.

That said, the approach chosen encompassed two steps. In a first stage, only peak hours were considered for the modeling of this parameter. This way, atypical behaviors and idle checkout time are disregarded. Secondly, a distribution of the post-scanning times was obtained for each bucket of 10 items in each CT, capturing the effects of the explanatory variables previously mentioned. Since there is an absence of data to make a comparison between the values obtained and the reality, the validation of the modeling results was made together with the retailer's team.

4.2 Configurator

The main goal of the configurator is to guide the simulation model into which checkout configuration is theoretically ideal for each store, taking into account the customers' demand. Additionally, the configurator provides a preliminary assessment of the stores that might be in higher need of a reconfiguration. To define this preliminary ideal configuration, the first step encompassed the development of an M/M/S model for each hour of the year and each CT available in the stores.

The use of an hour approach enables to somehow work around the arrivals' stationarity inherent to queueing theory, whereas by developing a model for each CT, the differences between them in the arrival rate λ and service rate μ are seized. The λ is obtained based on the number of customers that used each CT in each hour (already "anticipated" recurring to the methodology described in 4.1.3), while the μ is based on the transaction time per server in that hour in each CT.

The M/M/S models have the following characteristics:

- The condition $\frac{\rho}{S} < 1$ has to be verified so that the model is solvable.
- The arrival rate is independent of the system's state ($\lambda_n = \lambda, \forall n$);
- The service rate μ is the same in all servers of the same CT and $\mu_n = \begin{cases} n \cdot \mu & \text{if } n < S \\ S \cdot \mu & \text{if } n \geq S \end{cases}$

A key note is that these models correspond to the general case of a single-queue, with S servers in parallel. That said, it is important to point out that the distribution of customers between the regular checkouts available (that in reality have one queue for each checkout) is in fact optimized. Moreover, M/M/S models assume that the interarrival and service times follow exponential distributions. To verify these hypotheses, the K-S Lilliefors test (Lilliefors, 1969) was used to evaluate the adjustment of the data points to an exponential distribution with unknown mean. Regarding the interarrival time, the hypothesis stood for the majority of the stores, whereas concerning the service time it was only validated in a small sample of stores, enhancing the importance of the simulation model.

To implement the M/M/S models, the following simplified expressions for the probability of each system's state were used (Equations 4.4 and 4.5):

$$P_0 = \frac{1}{\left(\sum_{n=0}^{S-1} \frac{\rho^n}{n!}\right) + \frac{\rho^S}{S!} \cdot \frac{1}{1-\rho/S}} \tag{4.4}$$

$$P_n = \begin{cases} \frac{\rho^n}{n!} \cdot P_0 & \text{if } n \leq S \\ \left(\frac{\rho}{S}\right)^n \cdot \frac{S^S}{S!} \cdot P_0 & \text{if } n > S \end{cases} \tag{4.5}$$

When solving queueing problems to determine the optimal number of servers (S), the usual assumption is to attribute a cost to both an additional server and customers' waiting time, with the objective being minimizing the total cost of the system. Notwithstanding, the attribution of a cost to waiting time is not a simple procedure, with the usual criteria being too subjective (Kleinrock, 1967). Due to the lack of ease for the retailer to define this cost, and having into account the importance given by the customers to the size of the queues, this cost was replaced by a restriction related with the number of customers in the system. Thus, the imposed restriction is that the probability of having more than 3 customers per server (1 being served plus 2 in the queue) is never greater than 1% (Equation 4.7). The value of this threshold was strategically defined along with the retailer, although it is customizable and adaptable to other requirements. The cost of the solution will then be evaluated *a posteriori*.

In short, for each M/M/S model at the hour-CT level of granularity, the mathematical formulation of the problem can be summarized by the Equations 4.6 to 4.9.

$$\text{Minimize} \quad S \tag{4.6}$$

$$\text{subject to:} \quad \sum_{n=3\cdot S+1}^{\infty} P_n < 1\% \tag{4.7}$$

$$\frac{\rho}{S} < 1 \tag{4.8}$$

$$S > 0 \tag{4.9}$$

Then, for the synthesis of the preliminary ideal configuration for each store, it is chosen the number of S of each CT obtained in the M/M/S models that allows to achieve an annual service level of 99%. In other words, this annual service level means that the proposed number of checkouts is able to fulfill the demand of 99% of the hours of the year[1], while complying with the maximum of 2 customers per queue. This decision was made along with the retailer's team, in an attempt to have a dimensioning that is not overfitted to the hour of the year with the highest requirement of checkouts. For a better understanding of the described approach, Figure 4.7 depicts the suggested number of regular checkouts based on the M/M/S models for each hour of the year of one store. As it is possible to observe, by pointing to an annual service level of 99% the ideal configuration would allow to fulfill the majority of the hours while meeting the maximum queue

[1] In an entire year there are approximately 5000 hours of store's operation.

size threshold. However, there are some hours of the year that would require a greater number of checkouts to meet this criterion. In these hours, to comply with the maximum queue size the retailer acknowledges that he has to trigger a backup response mechanism, essentially based on mobile checkouts that are only placed on the store in these occasions. A deeper analysis on the fulfillment of the service level with the proposed configuration will be conducted later using to the simulation model.

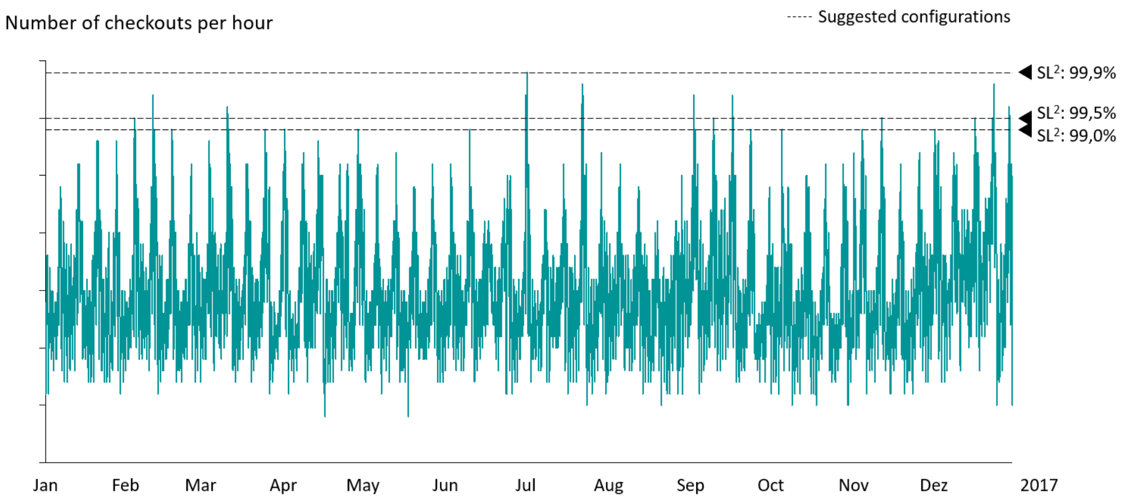


Figure 4.7: Example of the number of regular checkouts suggested by the M/M/S models in each hour of the year

4.3 Simulation model

This section presents the developed simulation model representative of the checkout process of a store. The main components and the connection between them are described, as well as a validation of the model obtained by comparing it with historical records and data gathered while visiting one store. The objective of the simulation model is to test the robustness of the solution proposed by the configurator when implemented in a store throughout an entire year. Queueing theory has a set of inherent limitations and approximations that might hide situations where the proposed configuration could not perform as expected. Thus, the simulation model provides a thinner level of detail, allowing the seizure of these situations in case they exist and enabling the stakeholders to adjust the configuration accordingly. Furthermore, it allows the gathering of a set of KPIs, such as average queue length or utilization ratios, with higher precision and smaller granularity level. Finally, it grants the possibility of developing multiple sensitivity analysis on the trade-offs between service level and associated costs.

The simulation model was entirely developed in Anylogic®. This software, as mentioned in chapter 2, supports an object-oriented modeling through its intrinsic Java environment. All the

[2]Annual service level.

functions implemented were coded in Java language, allowing not only a high customization of the model but also the access to an extensive number of external libraries.

4.3.1 Selection of the simulation method

Before start developing the model, it was necessary to choose the simulation method that better suited the problem, between the ones presented in chapter 2. Since the checkout process is of close operational character, Agent-Based (AB) and Discrete-Event (DE) methods emerged as the most suited for the purpose. Among these two, since the checkout process can be well represented as a flowchart with a sequence of known operations, DE is acknowledged to be the best choice. Inside DE simulation, PM could be considered as a valid alternative, given that the goal is the replication of pedestrians' flow in a store. However, the higher computational effort required by this type of simulation (Grigoryev, 2015), along with the intent to simulate an entire year for every store, jeopardized its selection. For the reasons aforementioned, the final decision fell back on DE as the best method to accurately represent the checkout process.

4.3.2 Model components

The structure that supports the simulation model is composed of five agents[3] and their interaction. For the development of this simulation model, not only was important to guarantee that the complexity of the checkout process was well represented but also to ensure the model was modular and flexible enough to enable its extension or modification according to new CTs available or changes in the checkout process. Thus, each CT is not represented as a separate agent, instead, it is composed of a sequence of smaller agents. For instance, a checkout may be composed of a queue, a scanning, and a post-scanning agents. By adopting this approach, the simulation model acquires a generic nature that can then be tailored for many retailers' contexts, as it happens when applied to the case study of the presented retailer.

In this section, the features and parameters that characterize each agent are described, being the basis for the simulation of the real scenario.

Customer agent

The customer agent represents each customer that arrives at the checkout area of a store. At the moment of agent's creation, a set of parameters must be defined:

Id: Unique identifier of the customer.

Number of items: Represents the basket size of the customer.

Payment method: Identifies if a customer will pay with credit card, cash or other methods.

Discount usage: Identifies if a customer will use the loyalty card and/or discount coupons.

[3]Note that the term "agent" does not refer to Agent-Based Modeling, being merely the way Anylogic® allows to organize the model in different building blocks.

Checkout sequence: Details the sequence of checkout related agents (queue, scanning, and post-scanning) that the customer will follow to reproduce the checkout process. This sequence is limited to the checkout alternatives available in the store.

These elements are all generated based on the parameters' modeling computed for the store being simulated, according to the methods described in section 4.1. Finally, auxiliary variables such as $checkoutTime$, $sequenceIndex$ or $checkoutIndex$ keep a record of information related with the time each customer spends in the system, as well as the current and next checkout stage of the customer.

Queue agent

The queue agent represents each queue that composes a CT available in the store. As already mentioned, the queue can be of two types: single-queue and multiple-queue. However, the logical representation between the two is the same, with the only aspects that change being how many checkouts the queue is going to serve and how the customers are "pulled out" of the queue. Each queue agent possesses a set of parameters:

Id: Unique identifier of the queue.

Type: Identifies if the queue is working in a single or multiple system.

State: Identifies if the queue is currently open or closed.

Size: Identifies how many customers are in the queue at each moment of time.

All the queues work in FCFS as the queueing priority to accurately represent the real situation. Also, when a customer enters a queue, he is subject to one of two situations: the checkout connected to this queue is empty and the customer can immediately exit the queue ("push" scenario); the checkout is occupied and the customer has to wait in the queue, only being called when the checkout becomes free ("pull" scenario).

Scanning agent

The scanning agent represents the process of scanning all the items each customer is going to buy. It also includes the act of placing the items on the checkout belt before initiating the scanning operation. In fact, the scanning agent allows the presence of two customers at the same time: one placing the items on the belt and one being assisted by the cashier. The exception to this situation is the self-checkout, where only one customer is allowed at the same time and the scanning is completely done by the customer himself. Each scanning agent has the following parameters:

Id: Unique identifier of the scanning station.

Type: Identifies if it is a regular scanning or a self-checkout.

State: Identifies if the scanning is currently open or closed. This state is directly associated with the state of the queue attached to this agent.

Scanning time: Identifies the time it takes to scan each item in this checkout type. This time is based on the distribution that characterizes the store, as explained in subsection 4.1.4

The scanning agent has an identical behavior with the queue in regard to the "push" and "pull" scenarios: if a customer arrives at the belt and the scanning station is free, the scanning process immediately initializes (the act of placing the items on the belt and the scanning process occur simultaneously). Otherwise, if the scanning station is occupied, the customer on the belt has to wait until it becomes free. This agent also has a set of auxiliary variables, which allows the correct implementation of the its logic and the storage of data essential for the KPIs attainment, like $timeServingCustomers$ or $timeOpen$.

Post-scanning agent

The post-scanning agent represents the operation related with payment issues that follow the scanning and end the customer's checkout process. In a standard checkout flow, this process immediately succeeds the scanning. However, the modular structure of the simulation model also allows to simulate situations where there is an intermediate queue between the scanning and the post-scanning process, something that has already been experimented in some European retailers (Rossetti and Pham, 2015). As a result, this agent has the following parameters:

Id: Unique identifier of the post-scanning station.

Connection: Identifies the agent's id to which the post-scanning is connected (queue or scanning)

Post-scanning time: Identifies the duration of post-scanning operation for each customer. This time is based on the distribution that characterizes the store, as explained in subsection 4.1.5.

Since this agent is the last stage of the customer's checkout process, it is also where the between-ticket time is accounted for. This parameter is incorporated by implementing a delay in the customer's exit, with its duration being obtained through the use of the modeled distribution described in section 4.1.6.

Main agent

The main agent represents the environment where all the agents described interact. Also, it is in this agent that the general settings are defined, namely the movement's speed or the unit of time to be used in the simulation, as well as it is where all the input data is received from external sources.

4.3.3 Functional interactions

After the description of the model components individually, it is important to understand how they interact and the sequence of events that produce the dynamics of the simulation.

The simulation starts by importing all the data regarding the store's characteristics and the configuration to be tested. Concerning the former, all the distributions that represent the input parameters' are recorded in collections, so that they can be continuously used during the simulation while the customers enter the checkout area of the store. With regards to the latter, the number of checkouts that allows to reproduce the configuration to be tested is generated on simulation's startup.

The creation of customers is based on a two-fold approach: firstly, it is considered the weekly demand of the store for every week of the year 2017; secondly, that weekly demand is divided into buckets of 15 minutes for every day of the week, according to the average profile of the store and using the parameter described in subsection 4.1.3. Therefore, an event occurs every 15 minutes, where customers are randomly inserted into the system during that time bucket. Each customer will then have a sequence, considering the probabilities of choice between CTs (see section 4.1.2), that determines the CT he will use to perform the checkout process.

With the CT defined, the customer still has to make a decision of which queue to go to. Thus, a function was developed to replicate this decision process, evaluating the current state of the system and choosing the queue with fewer customers, out of the checkouts currently open of that CT. The number of open checkouts at each moment of time is the result of an event taking place every 5 minutes (customizable). In this event, the size of the queues is evaluated and, in the case of surpassing a certain threshold (also customizable and defined as 2 in the case of this retailer), a new checkout of that CT opens and is ready to receive customers. To mimic the real circumstances where, in the case of the checkouts in a multiple-queue system, when a new checkout opens, the customers in the adjacent queues are allowed to switch, an algorithm was developed to reproduce this behavior. On the other hand, every 15 minutes, if the size of the queues is below a certain value, one checkout closes and no more customers are allowed to enter that queue until the event re-occurs. A key note is that, in accordance with the retailer policy, at least one checkout of each CT has to be always operating. Once again, self-checkouts are an exception to this opening/closing criterion, since they do not need a cashier to operate and are permanently open.

Summing up, in Figure 4.8 it is possible to observe the customer flow in the simulation model. In each "goTo" element, the current state of the customer's sequence is evaluated and he is directed to the next stage of the checkout process, in accordance with the path already performed.

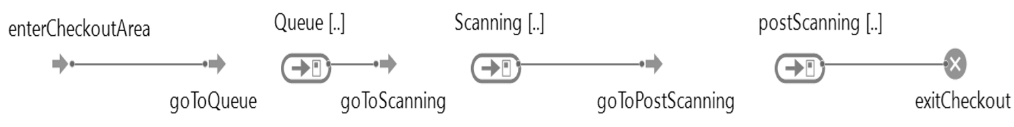


Figure 4.8: Customers' flow in the simulation model

To assess the performance of the tested configuration, a set of KPIs are recorded during the execution of the simulation. First of all, to contemplate the customer perspective, every 2 minutes the queue sizes of each CT are measured and recorded. Secondly, regarding the checkouts' performance, the operating time (time open) and the time serving customers are recorded, allowing the calculation of the occupation and utilization ratios of each checkout available in the store. Finally, the service level of each CT in every hour is attained through the event that evaluates the checkout opening/closing criterion. Hence, every time a new checkout should be open, due to the queue sizes surpassing the predefined threshold, and there is no more checkouts available of that CT (which is limited by the configuration being tested), an auxiliary counter is incremented. The hourly service-level is calculated by dividing this counter by the total number of evaluating periods in an hour – which, in the case of the event occurring every 5 minutes, totals 12 periods. This KPI is of crucial importance, since it allows to understand the number of times the tested configuration would not be enough to maintain the queue sizes below the established cap.

4.3.4 Validation of the simulation model

This section intends to expose the analysis conducted to validate the simulation model and assure its accurate representation of the reality. Two aspects were essential to be verified: the correct representation of the queue sizes and the checkout opening/closing criterion, as well as the fitness of the parameters' modeling. Hence, two tests were conducted. Firstly, a one-day visit to a store allowed the gathering of the number of open checkouts and queue sizes every 2 minutes. This data was then compared with the results outputted by the simulation model while simulating this specific day (see Figure 4.9). In this test, the modeled distributions were not used. Instead, it was used the customers' actual transaction data (basket size, scanning time, among others), in an attempt to validate the logic behind the simulation.

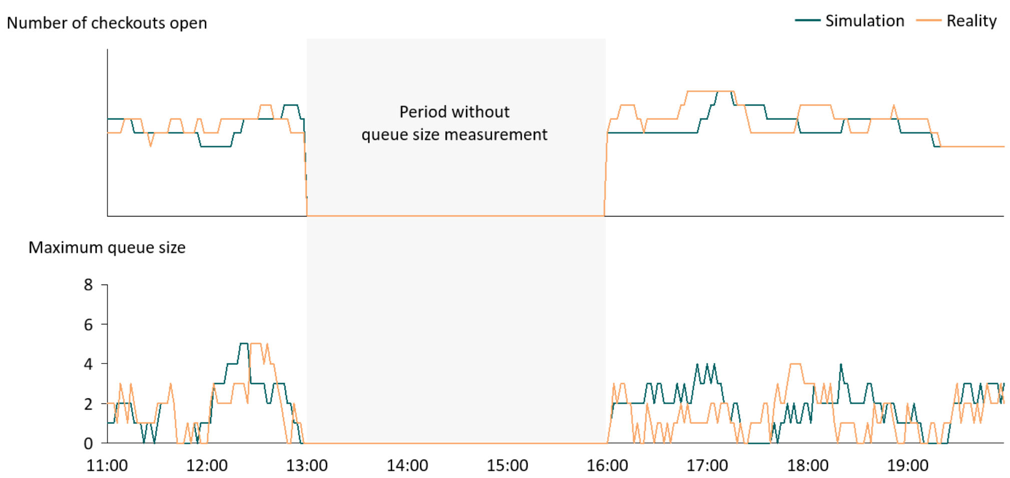


Figure 4.9: Comparison of the number of checkouts open and maximum queue size between the simulation model and reality for a specific day

Through the analysis of the upper chart of Figure 4.9, it is possible to observe that the checkout opening/closing criterion is quite similar between the simulation model and the reality. However, it seems to exist a slight offset when there is a need to open a new checkout, with the simulator being a bit slower to react to an increase in the queue size. This aspect is believed to be explained by the visibility the store's manager has regarding the customers present in the store, enabling him to, in some way, anticipate a higher affluence to the checkouts. The simulation model is blind to this factor, leading to a slower reaction. Nonetheless, even with this delay, the simulator responds well and opens more checkouts when necessary, smoothing the queue sizes.

On the other hand, to validate the accuracy of the distributions, two simulations were executed of an entire month of one store, one using real data and the other the modeled distributions (Figure 4.10).

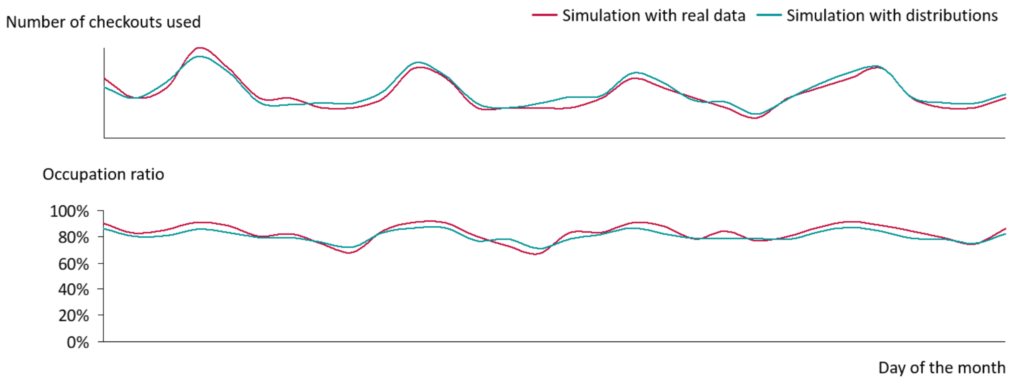


Figure 4.10: Comparison of the number of checkouts open and occupation ratio between the simulation model with real data and with distributions

As it is possible to observe, the results are also very close between the two scenarios tested. An important remark to be made is that, for the simulation with distributions, due to the stochastic nature of the simulation model, 10 replications were used for each run in order to have more solid results. Thus, in Figure 4.10 the results shown are the averages of each run's replications. These results suggest that the simulation with distributions seems to capture the behavior of the real data, including the variations between the days of the week.

In conclusion, the developed simulation model appears to accurately reproduce the checkout reality, with the residual differences having a plausible explanation. Therefore, the simulation is considered to be validated.

Chapter 5

Case Study Results

The current chapter presents the results of the developed methodology in the case of the studied retailer. The presentation is conducted in two separate steps. In section 5.1, a detailed report of the dimensioning of one specific store is provided, demonstrating the multitude of analyses that the combination configurator-simulation model allows to compute while simulating an entire year of store's operation. Afterwards, in section 5.2, the global results contemplating all of the retailer's stores are shown, as well as a ranking of the ones in greatest need of a reconfiguration.

In the context of this retailer, it was acknowledged that the demand for regular checkouts with multiple and single-queue should not be divided when addressing the dimensioning of the stores. The checkout equipment is exactly the same and the customers' choice is only based on the queue sizes at the moment they arrive at the checkout area. Thus, the recommended number of regular checkouts for each store will be the total amount, with the division between multiple and single-queue being carried out *a posteriori* by the retailer. Also, the space restriction is only analyzed by the retailer after already having the recommended configuration for each store.

5.1 Detailed report of one store

The results presented in this subsection are with respect to a store of format A (see section 3.1) chosen as the reference for this demonstration. After a presentation and brief examination of the contrast between the current and the proposed configuration, an analysis concerning a few alternative scenarios is also provided and evaluated.

5.1.1 Proposed configuration

Figure 5.1a illustrates the difference between the configuration currently in place in the store and the one that results from the implementation of the proposed methodology. As it is possible to observe, to satisfy the demand for each CT and have an annual service level (SL) of at least 99%, the proposed configuration indicates a reduction of 38,9% of regular checkouts and 16,7%

in self-checkouts. Figure 5.1b, which depicts the fixed costs of both the current and proposed configurations, shows that this reduction leads to savings in the fixed costs of 16,0%. The highest share comes from the equipment, which is valued at the acquisition cost, with the remaining savings resulting from a reduction in the maintenance cost related to the checkouts wear. The operational cost (cost with dedicated cashiers) is not included in this comparison, since this cost with the current configuration is more dependent of the checkout opening/closing policy than of the dimensioning itself.

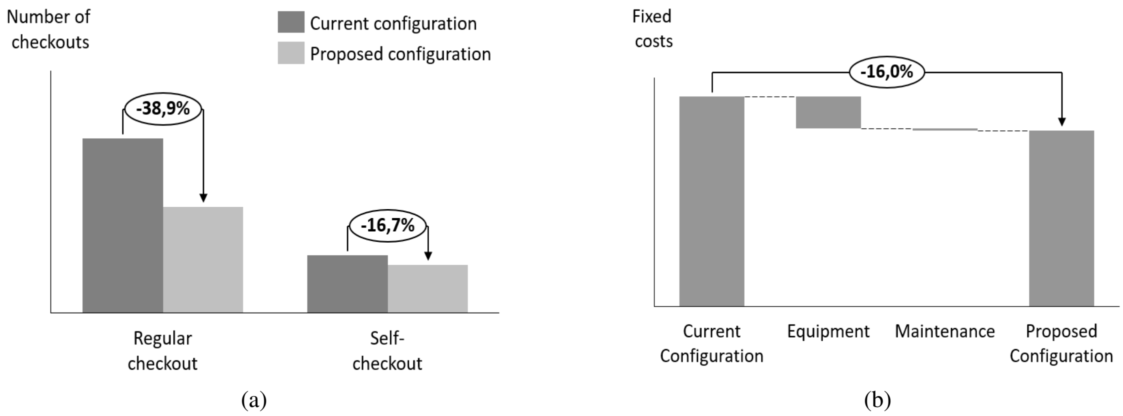


Figure 5.1: Configurations and respective equipment and maintenance costs for one specific store

To better understand the differences in the number of checkouts used between the reality – with the current configuration – and the simulation with the proposed configuration, Figure 5.2 depicts the fluctuation of this number for the regular checkouts in the presented weeks (the 10th week noted in the graph represents the 10th week with the highest checkout demand in 2017). It is also indicated the current and proposed amount of regular checkouts of this store. The number of checkouts used is consistently lower in the simulation than in reality, leading to a better occupation of the checkouts available. Moreover, the gap between the number of checkouts placed in the store and the ones actually used along the week is considerably narrower with the proposed configuration.

Regarding the average queue size, in the 10th and 5th this indicator is always below the defined threshold of 2 customers. However, in the 1st week, particularly on Saturday, this limit is surpassed in a set of hours, meaning that these hours are part of the group that the proposed configuration does not guarantee to satisfy without the queue size going over the predefined threshold. As a result, the annual SL for this store with the proposed configuration, obtained through the simulation model, is 99,39%, corresponding to approximately 30 hours with a response capacity not fully ensured and accomplishing the threshold of 99% used in the dimensioning.

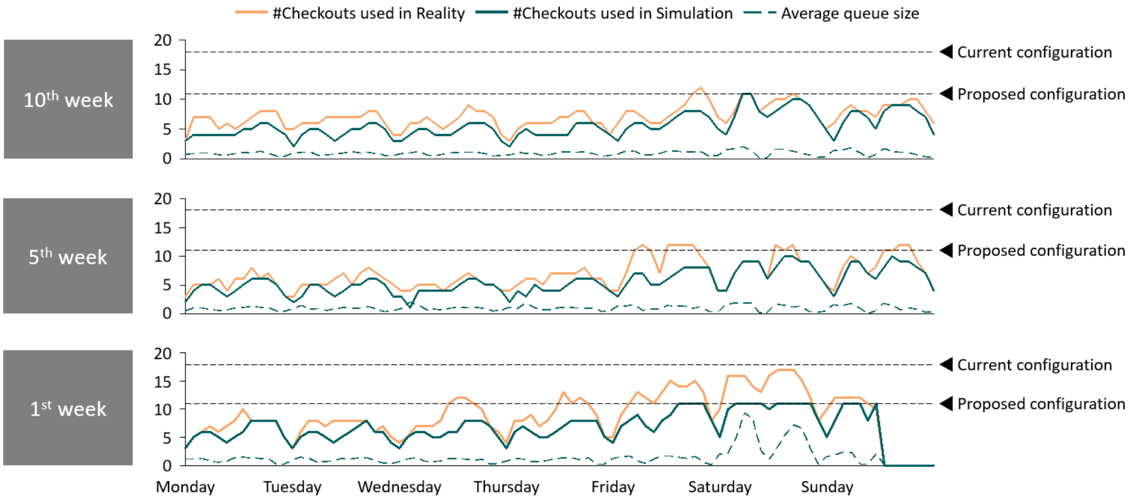


Figure 5.2: Number of checkouts used in reality vs in simulation and simulated average queue size. The average queue size could only be computed for the simulation scenario, since for the reality there is no record of the queue sizes' behavior in the represented weeks.

5.1.2 Demand transfer between checkout typologies

The first alternative scenario to be tested was the impact a demand transfer between CTs would have in the annual SL and in the operational cost[1] without changing the proposed configuration. Since the occupation ratio of self-checkouts is typically lower than the one of regular checkouts, transferring part of the demand to self-checkouts could lead to a better distribution of the customers between the available typologies. Taking into account the fact that it is not allowed to use self-checkouts while carrying a cart, the analysis was conducted only regarding transactions with until 15 items (threshold between basket and cart). Figures 5.3a and 5.3b depict how the annual SL and operational cost, respectively, are affected by a variation in the percentage of transactions performed in self-checkouts.

By transferring a part of the demand from basket sizes until 15 items to self-checkouts (from the current 32% to 50%), the annual SL increases 0,5 percentual points to a value of 99,92%, meaning that the amount of hours where the maximum queue size criterion is not fully fulfilled drops from 30 to only 4 hours during a year (Figure 5.3a). In fact, the demand transfer leverages the slack in the self-checkouts with the proposed configuration, which, by pointing the dimensioning to a SL of 99%, is actually able to meet the queue size criterion in the entire year. Nonetheless, it is important to notice that there is a practical limit to the percentage of the demand that can be transferred to self-checkouts, imposed not only by the impossibility to use this CT while carrying a cart but also by the customers' acceptance. Thus, the value of 50% is obtained by benchmark with the store with the highest percentage of transactions[2] carried out in self-checkouts, being considered as the upper bound of the demand transfer. In terms of cost, since the self-checkouts do not

[1] Cost related with the cashiers dedicated to the checkout operation.
[2] Transactions with until 15 items.

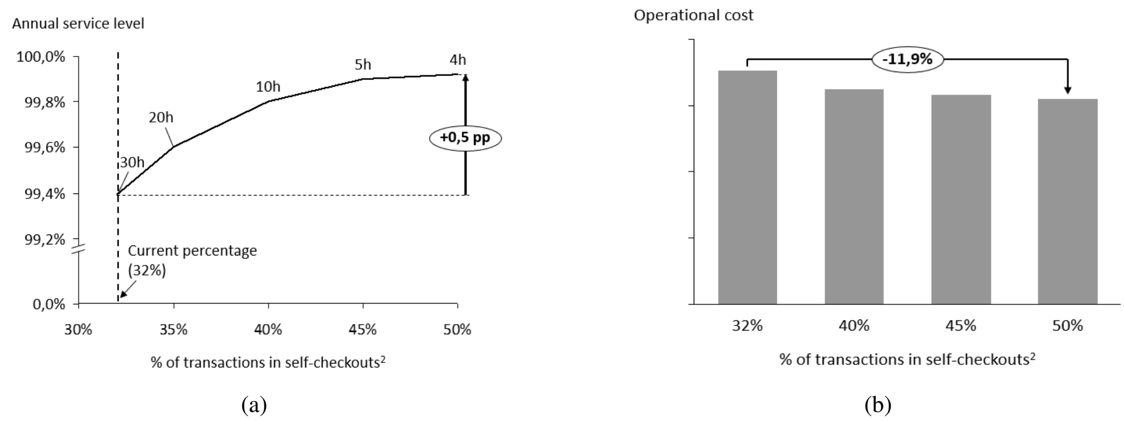


Figure 5.3: Annual service level (SL) and operational cost in the scenarios of demand transfer from regular to self-checkouts

require a cashier to operate, the operational cost is constant, independently of its demand. Hence, the demand transfer leads to a reduction of 11,9% in the operational cost of regular checkouts and, consequently, in the global operational cost.

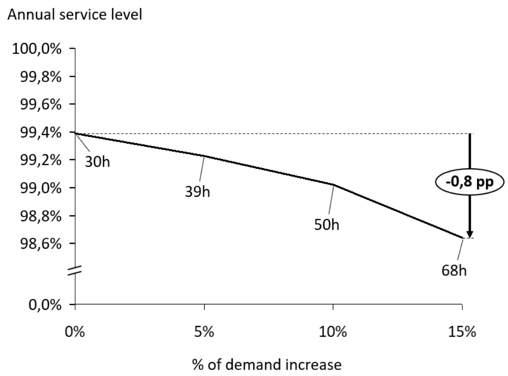
The results of this scenario brought up an improvement opportunity regarding the in-store checkout management. In case the retailer can implement a system that encourages an increase of the use of self-checkouts by the customers with small basket sizes, there are clear benefits for both parts, with the retailer reducing its operational costs and the customers facing shorter queues.

5.1.3 Demand increase

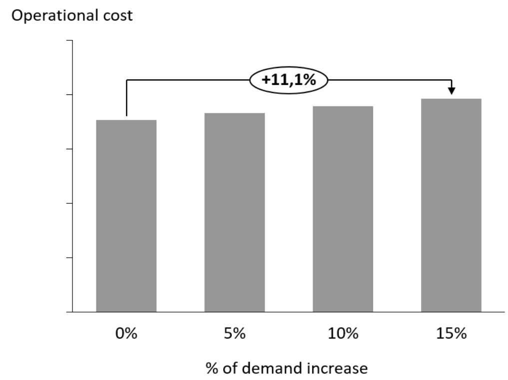
Other alternative scenario tested was to assess the robustness of the proposed configuration to potential increases in the annual demand, resulting in more transactions per day. The idea was to evaluate the magnitude of the impact this increase would have in the annual SL while maintaining the proposed configuration. Also, it was important to provide information about how the operational costs would escalate in the eventuality of this demand growth. This scenario's results are outlined in Figure 5.4a and 5.4b, detailing the impact of the demand increase in the annual SL and operational cost, respectively.

The key insight to retrieve from Figure 5.4a is that, even with a 10% increase in the demand, the annual SL remained within the threshold used in the dimensioning methodology (99%). This indicates that the proposed configuration is not overfitting the current demand, conferring enough slack to accommodate an increase in the number of customers shopping in the store. Only above a 10% increase, the amount of hours where the queue size criterion is not fully met surpasses 1%, corresponding to approximately 68 hours. Regarding the operational costs (Figure 5.4b), a 15% increase in the demand would bring an increment of 11,1% in this indicator, related with higher checkouts' operating time to counterbalance the queue sizes.

[2] Transactions with until 15 items.



(a)



(b)

Figure 5.4: Annual service level (SL) and operational cost in the scenarios of increased demand

To complete the analysis of this scenario, a validation was also conducted with respect to the queue sizes. Therefore, Table 5.1 shows the variations in the average queue size, in the three weeks already used as examples, resulting from the demand increase. From this table, one realizes that, in the 10^{th} most demanding week, there is enough slack to take in the demand increase without escalating the queue sizes. However, in the cases of more onerous weeks, with special emphasis to the 1^{st} one, the average queue size grows almost proportionally to the demand increase, suggesting that a greater attention should be given to assuring an adequate backup response mechanism in these periods.

Table 5.1: Average queue size in each week for the various scenarios of demand increase

Week	Demand increase						
	0%	5%	Δ_{rel}	10%	Δ_{rel}	15%	Δ_{rel}
10^{th}	0,901	0,903	0,2%	0,910	1,0%	0,937	1,04%
5^{th}	0,939	0,964	2,6%	0,988	5,1%	1,013	7,8%
1^{st}	1,533	1,626	5,3%	1,693	10,4%	1,767	15,1%

5.2 Global results

After the presentation of the detailed report for one specific store, it was also relevant to provide a holistic view of the results of the methodology when applied to the whole universe of the retailer's stores. Therefore, Figure 5.5 depicts a comparison between the number of checkouts of each CT currently placed in the stores and the one suggested by applying the proposed methodology. Additionally, other pertinent outcomes are illustrated: firstly, it is shown the gap between the current configuration and the maximum number of checkouts used in 2017; secondly, it is also presented the total number of checkouts of each CT that would be required to achieve an annual

SL of 99,99%, 99,50% and 99% in each of the stores. The percentages inside each bar are relative to the current configuration.

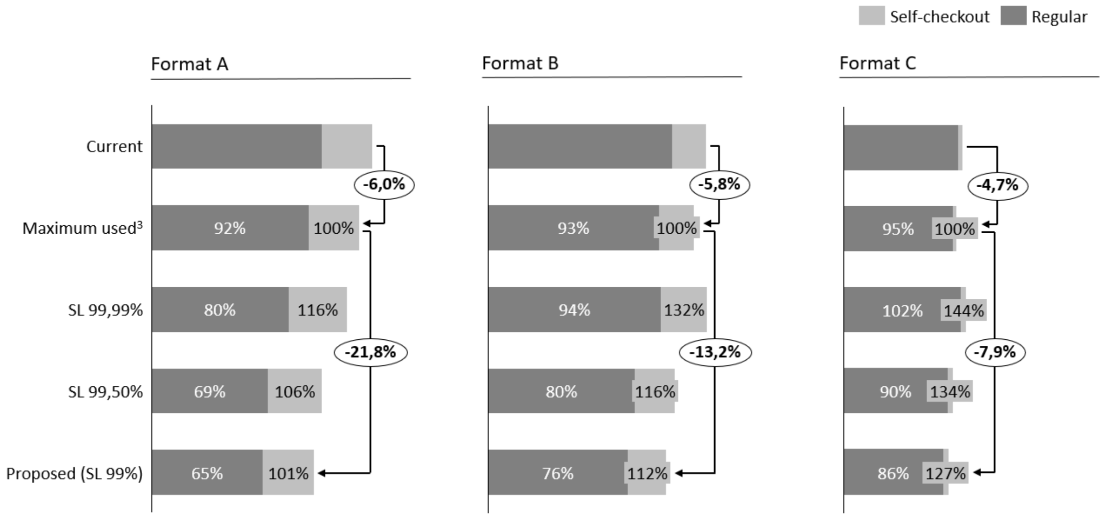


Figure 5.5: Current configuration placed in the stores vs configurations dimensioned to satisfy different service levels

The first key note that can be withdrawn is that, even without the implementation of the proposed methodology and only by analyzing the data available, there are checkouts that were never used in 2017 and, thus, could be removed from the stores. Furthermore, by implementing in each store the proposed configuration for an annual SL of 99%, it would require still less 21,8% of checkouts in format A, 13,2% in format B and 7,9% in format C. However, it is important to notice that the number of self-checkouts actually increases with the new dimensioning, possibly indicating that, in some hours of the year 2017, there were queue sizes over the predefined threshold in this typology. Finally, it is also possible to observe in the Figure 5.5 the recommended number of checkouts of each CT to accomplish an annual SL of 99,99%. Actually, in formats B and C, this number is higher than the maximum used in 2017, once again suggesting that there may have been hours in the year where the queue sizes surpassed the acknowledged threshold of 2 customers.

Following the analysis on the variation of the total number of checkouts between the current and proposed configurations, a relevant variable which should also be taken into account is how the fixed cost is affected by the transformation. Hence, Figure 5.6 outlines the differences between the costs of the current configuration in all stores and the proposed one. Complementary to the total value, a separation is made among the two CTs available (regular and self-checkouts) to provide a higher detail.

The results show that the total cost would be reduced by 14,6% by implementing the configuration that arises from the proposed methodology. This value can be better explained by looking at the variations inside each CT. The cost associated with self-checkouts actually increases in 6,5%,

[3]Maximum number of checkouts of each CT used during 2017.

since, as previously seen, the proposed number is higher than the current one in all three formats, implying an elevated cost with equipment acquisition. However, the reduction in the cost of regular checkouts is far more significant (25%), leading to the achievement of global savings with the new configurations.

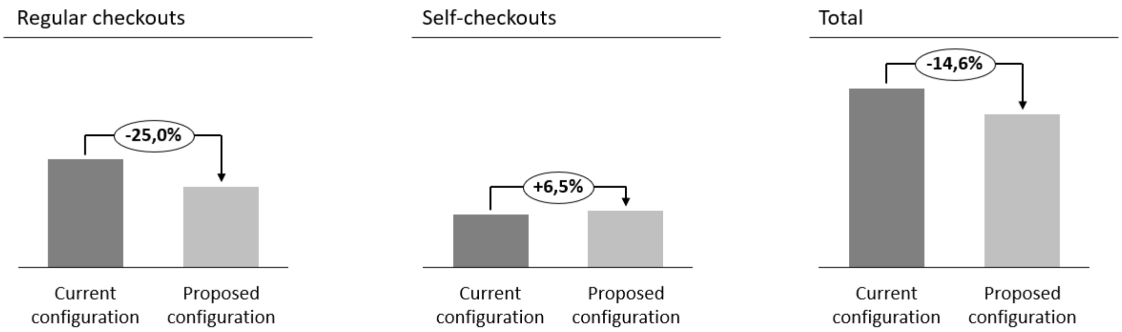


Figure 5.6: Comparison of fixed costs between the current and proposed configurations

Finally, one of the outputs of the methodology is a ranking of the stores with a more urgent reconfiguration need, evaluated through the differential between the current configuration and the proposed one. That said, Table 5.2 presents an example of the top 5 stores of each format with the highest percentual change in the configuration.

Table 5.2: Ranking of the stores of each format with the highest percentual change in the configuration

Rank	Store format		
	A	B	C
1	-42,0%	-38,5%	+44,4%
2	-34,0%	-38,5%	+40,0%
3	-33,3%	-38,5%	+40,0%
4	-29,8%	-17,4%	-40,0%
5	-27,7%	-16,7%	-40,0%

Through the analysis of the results, it is possible to verify that there are stores in a great need for reconfiguration, with a disparity between the current and proposed configurations in the order of 40%. Moreover, we can see that there are stores of format C which are considerably underdimensioned to satisfy the checkouts' demand and, thus, were given priority in the presented ranking. As a final note, the global results presented reflect the number of checkouts suggested by the configurator, which could not all be validated in the simulation model due to time constraints and computational effort required. However, taking into account the validations performed for a set of stores, the results from the configurator and simulation model were rather aligned.

The results obtained with the proposed methodology, and presented in this chapter, received great acceptance by the retailer's team and stores' managers. The suggested configurations are aligned with the retailer's perception prior to this project that the majority of its stores could be over-dimensioned. Also, the recommended increase of the number of self-checkouts was a need that the retailer was already aware of, but lacked an analytical support to grant comfort to the decision-making. Finally, regarding the implementation of the proposed configurations, the retailer has to, *a posteriori*, take into consideration not only the space restrictions but also the previously mentioned emergency response mechanisms. These mechanisms should be capable of providing a complementary support to the stores' during the most demanding hours and also forearm the possibility of checkouts breakdown (e.g. software malfunction).

Chapter 6

Conclusions and Future Work

The developed work focused on the checkout management of a Portuguese food retailer. The main goal was to provide the company with an analytical support to address the checkout dimensioning in its stores. The objective consisted on specifying the number and type of checkouts to be placed in each one of the stores, in order to fulfill the demand while accomplishing a certain service level. The project started by framing the problem at hands, namely identifying the current checkout concepts available in the stores and their specificities. Also, the key parameters that characterize the checkout process of each store and each checkout typology were modeled. Subsequently, a two-fold methodology was developed to provide the analytical support aforementioned. The solution approach is based on the identification of a preliminary ideal configuration for each store through the use of queueing theory models, followed by the development of a simulation model representative of the checkout process. The simulation aims to test the robustness and appropriateness of the proposed configuration when analyzed with a thinner granularity level.

The results enabled the retailer to have an integrated view of the checkout dimensioning of all the stores, raising great opportunities for improvement in the current process. The recommendation lies on a significant change in the total number of checkouts to be available in the stores while maintaining a high service level, a change that would lead to a relevant reduction in the costs associated with the checkouts. Furthermore, this dissertation also contributed to the identification of improvement opportunities related to the management of the in-store checkout process itself. For instance, the encouragement of a higher use of self-service checkouts by customers with small basket sizes could bring benefits for both retailer and customers. A deeper validation of the results is still an ongoing process, namely throughout the execution of a pilot project in a specific store by implementing the proposed configuration and verifying its outcomes.

The simulation basis of the proposed approach allows to provide greater comfort regarding the results. Many aspects of the checkout process, such as queue sizes and utilization ratios, can be analyzed with a much higher detail, allowing the retailer to gather information concerning particular elements that justify the suggested adjustments in the configurations. Also, the stochastic nature of the simulation reassures the decision-making when more radical changes are recommended. In another perspective, the modular structure of the developed methodology ensures that it can be easily extended to other retail contexts with distinct characteristics and different checkout typologies available. When looking at the methodology as a whole, the development of a configurator to guide the simulation model promoted a faster attainment of results, which otherwise could present itself as a time-consuming procedure due to the combinatorial essence of the problem. Also, there is a considerable confidence in the preliminary ideal configurations suggested by the configurator, since they showed to be aligned with the simulation model's results in most of the cases.

Nonetheless, the use of simulation to solve problems of high complexity, like the checkout dimensioning, is still a challenge and up for enhancement. Particularly, when dealing with human behavior, there is an intrinsic unpredictability and uncertainty associated. Thus, a deeper validation of the correct representation of the reality by the simulation model should be conducted in the future. More visits to the stores should be performed, in order to gather data regarding queue sizes and customer behavior, which could then be used to improve the configurator's inputs and to validate the simulation's outputs. However, the collection of this data in the stores is not a trivial matter for the human eye. Hence, a future use of video analytics is also an aspect to take into consideration to support the validation of the simulation model.

Despite the explicit advantages brought up by the presented methodology, some remarks can also be pointed out. In our methodology, the cost component is only analyzed *a posteriori*, essentially due to the difficulty in attributing a cost to the customers' waiting time. Also, the space restrictions are not included in the model, as a result of an inevitable delay in getting access to the data associated with this variable. A suggestion for future work would be to incorporate these variables in the definition of the ideal configuration. This aspect can assume higher importance when considering an even wider diversity of typologies available, where the main differences rest on the cost and space components and not so much on the operational efficiency or customer preference. Moreover, the prior division of the demand between assisted and self-service is based on the data available, which may not unquestionably reflect the customers' primitive preferences. Thus, a broader and more profound research could be developed on the customers' preferences for each checkout typology, identifying the key drivers that guide their choice.

Regarding the current state of the art, this project intends to fill the acknowledged gap and provide an integrated approach to tackle the checkout dimensioning, instead of addressing individual elements of the business scope. Nonetheless, future additional work could bring further benefits to the implementation of this methodology. For example, a possible extension would be the development of a closed loop between the configurator and the simulation model, where the preliminary ideal configuration provided by the former would be automatically adjusted taking into account the outputs of the latter. This upgrade would withdraw a portion of the human intervention in this process, expediting the attainment of results and, consequently, the decision-making. Additionally, with the emergence of new and innovative checkout typologies available in the market, the proposed methodology could be used to conduct a preliminary study on their impact in the checkout performance, as well as on how they tackle the limitations of the current solutions.

Bibliography

Abouee-Mehrizi, H. and Baron, O. (2016). State-dependent m/g/1 queueing systems. *Queueing Syst. Theory Appl.*, 82(1-2):121–148.

Alvarado, J. A. and Pulido, L. M. (2008). Simulation and experimental design applied to sizing supermarket cashiers in colombia. In *2008 Winter Simulation Conference*, pages 1356–1361.

Amorim, M., Lago, A., Moscoso, P., and Prieto, E. (2016). Assisted vs. self-checkout in retail: An empirical analysis of the impact of operational process dimensions on customer satisfaction, recommendation and reuse. *Journal of Service Science Research*, 8(1):1–39.

Baldwin, W. C., Sauser, B., and Cloutier, R. (2015). Simulation approaches for system of systems: Events-based versus agent based modeling. *Procedia Computer Science*, 44:363–372.

Balsamo, S. (1993). Properties and analysis of queueing network models with finite capacities. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 729 LNCS, pages 21–52.

Bar-Lev, S. K., Boxma, O., Stadje, W., Van der Duyn Schouten, F. A., and Wiesmeyr, C. (2009). Two-stage queueing network models for quality control and testing. *European Journal of Operational Research*, 198(3):859–866.

Belli, F., Budnik, C. J., and White, L. (2006). Event-based modelling, analysis and testing of user interactions: approach and case study: Research articles. *Softw. Test. Verif. Reliab.*, 16(1):3–32.

Benvenuto, N. and Zorzi, M. (2011). *Principles of Communications Networks and Systems*. Wiley.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences Proceedings of the National Academy of Sciences*, 99(Supplement 3):7280–7287.

Buell, R. W. (2017). Last place aversion in queues. *Harvard Business School*.

Chan, W. K. V., Son, Y. J., and Macal, C. M. (2010). Agent-based simulation tutorial - simulation of emergent behavior and differences between agent-based simulation and discrete-event simulation. In *Proceedings of the 2010 Winter Simulation Conference*, pages 135–150.

Davison, J. (1993). Electronic point of sale scanning in multiple food retailing: A rebuttal to "scanning in the supermarket: For better or worse?". *Behaviour & Information Technology*, 12(1):54–64.

Demirci Orel, F. and Kara, A. (2014). Supermarket self-checkout service quality, customer satisfaction, and loyalty: Empirical evidence from an emerging market. *Journal of Retailing and Consumer Services*, 21(2):118–129.

Ekman, E. (2016). *Self-Scanning Services in Food Retail : An Investigation on how Self-Scanning Services are adopted in the Swedish Food Retail Environment from a Service Fascination Perspective*. Thesis.

Grigoryev, I. (2015). *AnyLogic 7 in Three Days*. CreateSpace Independent Publishing Platform.

Hao, Q. and Shen, W. (2008). Implementing a hybrid simulation model for a kanban-based material handling system. *Robotics and Computer-Integrated Manufacturing*, 24(5):635–646.

Hartman, T. (2008). A hybrid checkout system. *University of Twente*.

Hillier, F. S. and Lieberman, G. J. (2001). *Introduction to Operations Research*. McGraw-Hill, New York, NY, USA, seventh edition.

J Williams, E., Karaki, M., and Lammers, C. (2003). *Use of simulation to determine cashier staffing policy at a retail checkout*.

Kameda, H. (1984). A property of normalization constants for closed queueing networks. *IEEE Transactions on Software Engineering*, SE-10(6):856–857.

Kendall, D. G. (1951). Some problems in the theory of queues. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(2):151–185.

Kleinrock, L. (1967). Optimum bribing for queue position. *Operations Research*, 15(2):304–318.

Klukin, P. (2014). Single line or multiple lines: Which should you choose? `https://www.lavi.com/en/resources-detail/single-line-multiple-line-queue` [Accessed: 13th April 2018].

Kwak, J. K. (2017). Analysis on the effect of express checkouts in retail stores. *The Journal of Applied Business Research*, 33(4):8.

Lilliefors, H. W. (1969). On the kolmogorov-smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325):387–389.

LineLogic (2011). Explore your queuing system options. `http://blog.linelogic.com/blog/2011/11/17/explore-your-options/` [Accessed: 13th April 2018].

Little, J. D. C. (1961). A proof for the queuing formula. *Operations Research*, 9(3):383–387.

Lu, Y., Musalem, A., Olivares, M., and Schilkrut, A. (2013). Measuring the effect of queues on customer purchases. *Management Science*, 59(8):1743–1763.

Mani, V., Aneja, Y. P., and Singh, N. (1990). Throughput analysis of a two station closed queueing network with blocking and server breakdown. *International Journal of Systems Science*, 21(4):611–619.

Mentzer, J. T. and McCarthy-Byrne, T. M. (2011). Integrating supply chain infrastructure and process to create joint value. *International Journal of Physical Distribution and Logistics Management*, 41(2):135–161.

Nelson, B. L. and Gerhardt, I. (2011). Modelling and simulating non-stationary arrival processes to facilitate analysis. *Journal of Simulation*, 5(1):3–8.

Opara-Nadi, G. E. (2005). *Electronic self-checkout system vs. cashier operated system: A performance based comparative analysis*. Thesis.

Pantano, E. (2014). Innovation drivers in retail industry. *International Journal of Information Management*, 34(3):344–350.

Rossetti, M. D. and Pham, A. T. (2015). Simulation modeling of customer checkout configurations. In *2015 Winter Simulation Conference (WSC)*, pages 1151–1162.

Schimmel, M. (2013). *Deployment of express checkout lines at supermarkets*. Thesis.

Schuller, P. (2018). A machine learning approach to promotional sales forecasting. *Faculdade de Engenharia da Universidade do Porto*.

Schwenke, C., Vasyutynskyy, V., and Kabitzsch, K. (2010). Simulation and analysis of buying behavior in supermarkets. In *2010 IEEE 15th Conference on Emerging Technologies & Factory Automation (ETFA 2010)*, pages 1–4.

Serasinghe, M. and Vasanthapriyan, S. (2017). Intelligent retail checkout management system. In *2016 2nd IEEE International Conference on Computer and Communications, ICCC 2016 - Proceedings*, pages 929–932.

Shannon, R. (1975). *Systems simulation: the art and science*. Prentice-Hall.

Shunko, M., Niederhoff, J., and Rosokha, Y. (2017). Humans are not machines: The behavioral impact of queueing design on service time. *Management Science*, 64(1):453–473.

Terano, T., Kishimoto, A., Takahashi, T., Yamada, T., and Takahashi, M. (2009). Agent-based in-store simulator for analyzing customer behaviors in a super-market. In *Knowledge-Based and Intelligent Information and Engineering Systems*, pages 244–251. Springer Berlin Heidelberg.

Terblanche, N. S. (2018). Revisiting the supermarket in-store customer shopping experience. *Journal of Retailing and Consumer Services*, 40:48–59.

Van Riel, A. C. R., Semeijn, J., Ribbink, D., and Bomert-Peters, Y. (2012). Waiting for service at the checkout: Negative emotional responses, store image and overall satisfaction. *Journal of Service Management*, 23(2):144–169.

Vardon, J. (2015). 360 degree scanning: Saving time, adding value. *UST Global*.

W. Conway, R. and L. Maxwell, W. (1962). A queuing model with state dependent service rate. *Journal of Industrial Engineering*, 12.

Weijters, B., Rangarajan, D., Falk, T., and Schillewaert, N. (2016). Determinants and outcomes of customers' use of self-service technology in a retail setting. *Journal of Service Research*, 10(1):3–21.

Ya. Khinchin, A. (1967). *The Mathematical Theory of a Stationary Queue*, volume 39.

Yamane, S., Ohori, K., Obata, A., Kobayashi, N., and Yugami, N. (2012). *Agent-Based Social Simulation for a Checkout Layout Design of a Specific Supermarket*, pages 153–64.

Zagel, C. (2016). *Service Fascination - Gaining Competitive Advantage through Experiential Self-Service Systems*.

Zeigler, B., Prähofer, H., and Kim, T. G. (2000). Theory of modeling and simulation: Integrating discrete event and continuous complex dynamic systems. 2.

Appendix A

The Decision Support System

Following the development of the presented methodology, it was decided to design an interface to be used by the case study retailer whenever addressing the checkout dimensioning problem. This preliminary Decision Support System (DSS), which is still in an early stage of development, was created in MS Excel®, due to the greater knowledge on the retailer's side while working with this software, as well as its user-friendly environment. The DSS main interface is presented in Figure A.1.

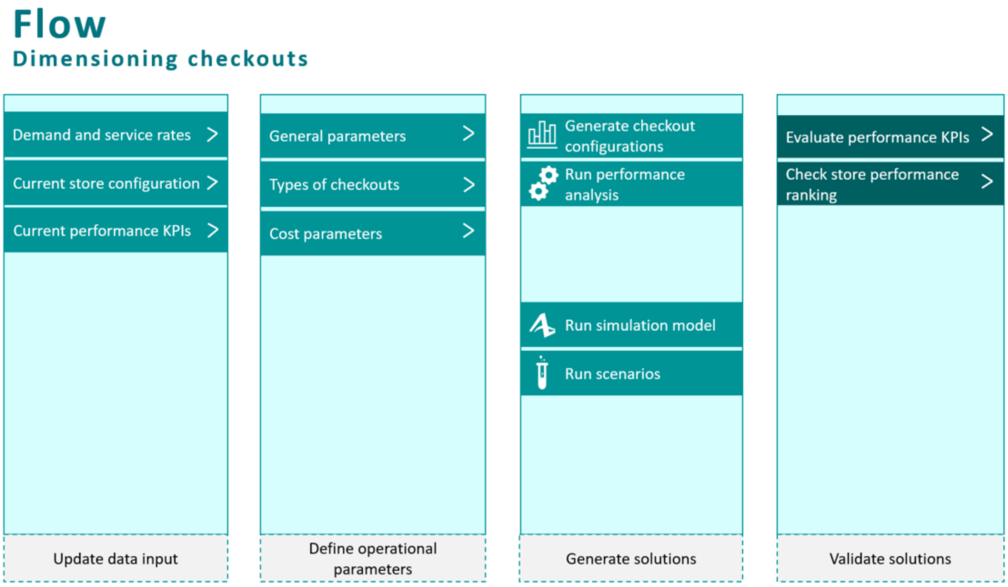


Figure A.1: The main interface of the Decision Support System

The DSS is divided in four main blocks. First of all, the "Update data input" block, where the data regarding the current situation – for instance, the current configurations and their demand – can be analyzed and updated. Secondly, in the "Define operational parameters" block, the input parameters regarding the stores and the various CTs can be conferred, as well as it is where the cost parameters can be defined for a later attainment of the total cost of the solutions. Thirdly, the "Generate solutions" block is considered to be the main module of the DSS. It is where the methodology implemented in the configurator and simulation model can be executed. The "Generate checkout configurations" and "Run performance analysis" are based on the configurator and output the proposed configuration for each store, benchmarking with the current one for a preliminary performance assessment. Then, it is possible to, for a specific store, run the simulation model for the base and alternative scenarios, with the results being imported to the DSS. Finally, the "Validate solutions" block is where the performance KPIs, such as annual service level or operational costs and which result from the previous steps, can be deeply analyzed by the retailer's team. The ranking of the stores' reconfiguration need can also be consulted in this block.

The main objective of the DSS is to diminish the user's interaction with the more intrinsic details of the methodology, which can be difficult to understand in the absence of simulation and queueing theory know-how. This way, the implementation of the proposed methodology can be done in a friendlier and more straightforward manner.