Improvement of a Queuing System through Simulation and Lean Methodologies

An applied study of EDP stores

Carlos Guilherme Almeida de Frias

Department of Engineering and Management, Instituto Superior Técnico

Abstract

This work approaches a real problem in EDP Stores, where, as it will be detailed, customers must endure high waiting times which harms their contentedness about the service. The company aspires for a more efficient service of its clients (better KPIs) and, consequently, a higher global satisfaction-level.

A tailored framework is hereby proposed to tackle the identified causes for the system's malfunctioning. The symbiotic knowledge from more than 100 sources of three research areas (Queuing Theory, Simulation, and Lean) was gathered, unveiling improvement opportunities in the stores through a good characterization, understanding and modelling of their queuing system.

After reading this work, queues' crucial impact on the welfare of the customer shall become clear. Mitigating their hazards, while improving their performance, the tailored framework's output is being currently implemented by the firm in Bragança with an estimated reduction of its waiting-time of 76%.

Keywords: Queuing System; Waiting Time Reduction; Queuing Discipline; DES; Lean Improvement

1. Introduction

Being formed whenever the demand surpasses a system's service capacity, queuing has a crucial impact on the efficiency level of any system (Hillier & Lieberman, 2010). Research about Queuing Systems began with Agner Erlang in 1918 and it has, since then, been developed by many other academic fields due to its impact on any system. Responsible for up to 90% of the lead-time in the industry (Ndukwe *et al.*, 2011), for a given demand, queues may be mitigated by either increasing the system's capacity, or the system's efficiency (Gutacker, 2016). Analysing relatable literature, Xu *et al.* (2016) have even assessed that queues performances will impact firms' revenues, as they may incite a customer to buy either at a higher price, or in bigger quantities, when they are properly managed, ultimately becoming loyal and returning more often to the store.

In fact, that is one of the reasons why the Portuguese firm *Energias de Portugal* (EDP) has sought a framework to improve its queues performance. Operating in a recently liberalized Energy market, customers' satisfaction is crucial to continue thriving economically.

2. Case Study

Therefore, EDP seeks to improve the two KPIs (Key Performance Indicators) measured in-stores: the average waiting-time (AWT), and the Quality Indicator (QI) that measures the percentage of customers called under 20 minutes of wait. Nonetheless, as stores' activity is heavily regulated, the customer journey (Figure 1) may not be changed. Any queue system solution must encompass: the printing of one inquiry ticket (five types), the waiting-period in-store, the service of the client, and the digital recording of those services' characteristics.

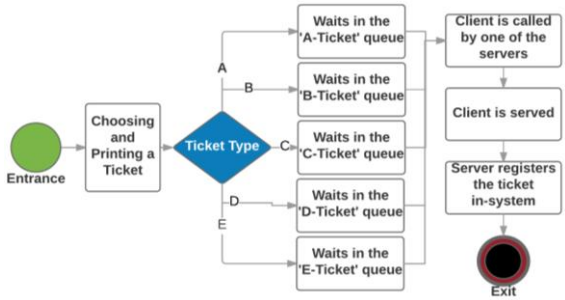

Figure 1: Customer's Journey

Currently, the system's performance is being hampered by the lack of a systematic way to manage its queues. In fact, while EDP has acquired a ticket management system (Inline®), it suspects that Inline®'s priority-levels (PL) are not being correctly implemented. With a national QI of 93.1% in 2016, *ca.* 150 thousand clients waited more than 20 minutes to be called. Besides, three ticket-types (A, C, D) have higher AWTs than the official target (7'30''). The situation aggravates considering that, while in-target, 'B' and 'E' clients have higher AWT than their actual average service-time (AST).

3. Literature Review

To develop a robust framework that would improve EDP's, or any other similar, queuing system, a literature review analysed the three next research areas to, respectively, characterize, model and seek feasible solutions.

3.1. Queuing Theory

Whilst Queuing Theory often lacks pragmatism by developing complex models that search for optimal solutions, when relaxed solutions would often be sufficient, it is still crucial to organizations, as it provides valuable data and understanding of their systems (Newell, 2013).

In a queuing model, entities (*e.g.,* customers) enter the system from a calling population (the *input source)* at a specific rate, remaining in line (the *queue*) until they are called to a service-facility. This entrance in the *service mechanism* follows a specific order (the *queue discipline*) through which the server (human assistant, or machine) calls a client and begins serving him during a certain duration (*service-time*). Upon completion, the customer leaves the system (Foster *et al.*, 2010). Since its characterization often becomes complex, Gautam (2008) developed a five-fields notation A/S/c/k/QD, standardizing the description of the main components of any queuing system.

'A' characterises a crucial parameter, the interarrival-time. As arrivals may be considered random, whenever the source is assumed to be infinite, interarrivals are often modelled by a Poisson distribution. Knowing the average number of hourly arrivals (μ), the interarrival-time (λ) can then be modelled by an exponential distribution following Equation 1, with limited impact on the model's precision (Creemers & Lambrecht, 2010).

$$\lambda = (\mu/60)^{-1} \qquad (1)$$

'S' defines the service-times distribution, that may be similar, distinct, or even dependent on each other. 'c' quantifies the number of servers available. 'k' represents the queue's capacity that may imply a loss of clients due to a certain length of the queue, whenever it is not considered to be infinite. And 'QD' sets the queuing discipline that may range from a pure FIFO (First In First Out) strategy to a priority discipline, where high-prioritized clients are called first, allowing, or not, preemptive calls.

Characterizing EDP's system, it was assumed to: have an infinite population, as the entrance of one customer does not impact the size of its input source; have an infinite queue, since there is no official limit to the number of waiting customers; be visited by patient customers whose call is processed following the non-preemptive Inline®'s PLs; and to have the same number of servers as of service-stations.

The most relatable research-area found was the telephone call-centres, where, with the same characteristics (M/M/c/k/PD), the imposed high-dimensionality on the Markov-chain (MC), due to the several (*m*) priority levels, hampers the solvability of these models. In fact, when solving them, the MC will become infinite in all the '*m*' dimensions, which is intractable. Studies have failed to mathematically solve these models without incurring in inaccuracies. Thus, even though Queuing Theory has characterised the system, Harchol-Balter *et al.* (2005) suggests modelling it by Simulation techniques instead.

3.2. Simulation

Simulation has been used by companies to help improving their decision-making process after applying Queuing Theory. According to Pidd (2004), its benefits outweigh other techniques, as Simulation allows to explore a set of different scenarios through a virtual environment that is easy to understand and that averts interfering with the real system prior to understanding its potential ramifications.

Among the several simulation approaches, Discrete Event Simulation (DES) is the best fit to EDP's stores environment, as it was specially developed to mimic dynamic and stochastic systems whose quantitative variables change at discrete moments (Banks *et al.*, 2005).

DES mimics real environments resorting to two main components: the dynamic inputs (*entities*) and the system's *resources*. While *entities* represent units of traffic (*e.g.,* clients), *resources* simulate anything that is capacity-constrained (*e.g.,* workers). *Entities* have individual *attributes* that are crucial to measure the performance of the simulated system, since they record every interaction between *entities* and *activities* (*i.e.,* a set of actions they may undertake). These interactions will result in *events* that may alter the system's *state* by changing the defined *global variable*. Throughout its journey, an entity may be in five different states. Generally, it is *active* whenever it is moving through the system, *conditioned-delayed* when it is on a queue, *ready* when it is available to be processed, *time-delayed* when it is being served for a certain service-time, and *dormant* whenever it is becoming ready in a specific scheduled moment (Rossetti, 2015).

Detailing two specific dimensions of stochastic and dynamic models, Nova (2008) states that, while ideally statistical inference methods would be applied to fit the sample-data to a statistical model, Simulation does not require that, as their Monte Carlo techniques generate artificial inputs that respect the characteristics of the original statistical model. Moreover, Simulation models have a time-keeping mechanism, the *simulation clock*, that keeps track of the timeline and can pin planned events on specific *clock times* (Rubinstein & Kroese, 2016).

3.3. Lean Theory

Lean's findings suggest a culture of incremental improvements to raise the overall system efficiency. By focusing on quality, costs and service-levels may be revised by redesigning processes and eliminating all non-value adding activities. Decisions must be based on objective data and inter-disciplinary discussions to unveil transversal opportunities which will contribute to the robustness (*i.e.,* decreased variability) of the system (Bicheno & Holweg, 2000).

A Lean framework will also be adopted to identify best-practices and its philosophy will be embedded during this work to foster its implementability through an enhancement of everyone's commitment to be involved in the process of change and to adopt the unveiled and suggested improvement opportunities.

4. Proposed Methodology

Convening four symbiotic frameworks found in the literature, this Methodology developed a tailored framework that ascertains improvement opportunities in any ticket-based store system.

As presented in Figure 2, its sequential stages plan to, first, unveil best-practices and, then, to assess their overall impact on the system's performance (*i.e.,* analysing its KPIs variation).

This goal is accomplished through a robust process where every aspect of the real system is addressed, characterized and tested, hence leading to a realistic set of solutions that may be proposed for implementation.

The framework relies on Engström *et al.*'s (2006) proposal of using the coefficient of variation (CV) to compare the robustness of different performances. The lower it is the value yielded by Equation 2 (quotient between the standard deviation and its mean), the less variable is process 'x'. These values imply, nonetheless, a careful data analysis (collection, sampling and treatment) to guarantee the similarity of performances prior to rank them.

$$CV_x = \sigma_x / \mu_x \qquad (2)$$

Regarding the implementation of DES, three symbiotic works were considered. One proposing an initial conceptual phase to guarantee a complete definition of all the model's components and particularities (Sachidananda *et al.*, 2016), and two others focusing on the simulation phase, where not only the conceptual findings should be implemented, but also the model's validity is asserted (Cardoen *et al.*, 2010; Balci, 2003).

5. Implementation of the Framework

5.1. Model's scope definition

To understand EDP's system, the entire 2016's Inline® database was imported to a MS Access file. To collect workable data, the several tables in the database were merged after identifying their key field (the *ID_Record*, a unique ticket code is attributed to each visitor). 36 fields (out of 48) were also removed, since they contained random sample-data of non-included features in EDP's version of Inline®. Resulting in a Table with 12 fields (Figure 3), each line convened the journey of one client.

Ticket_ID	Store_ID	Server_ID	Service_mode
Ticket_mode	Ticket_type	Service_type	Subservice_type
Start_hour	Ticket_hour	Duration	Service_Station_ID

Figure 3: Collected data-fields from Inline® system

Using these fields, the store's demand, layout, AST, AWT, and the client's journey could be characterized. Indeed, as a customer arrives to an EDP Store (*Store_ID*), he prints a ticket (*Ticket_ID*) choosing the nature of his inquiry (*Ticket_type*). A time-stamp (*Ticket_Hour*) sets the entrance time and, as he will wait until being called (*Start_hour*), his waiting-time (WT) may be computed by subtracting these two parameters. His service-duration will also be registered from the moment he was called by a server (*Server_ID*) to his station (*Service_Station_ID*).

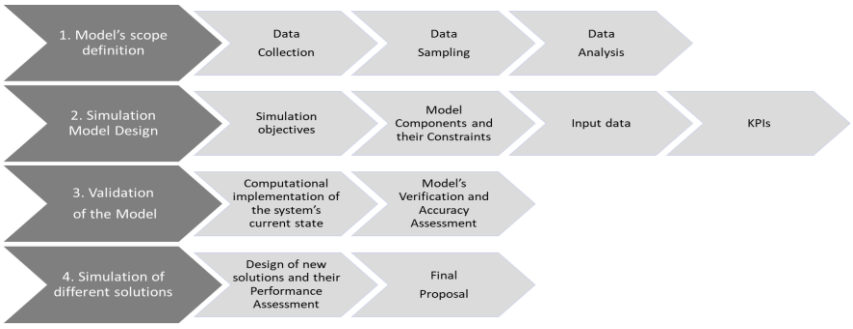

Figure 2: Proposed Methodology - A four-stages framework

Upon closing the client's ticket after serving him, the assistant validates the service type (*Service_type*) and classifies the subservice (*Subservice_type*). Automatically, the system characterises that exact ticket as: having been served (*Service_mode*), differentiating it from no-shows, or failed services; and as having been a printed ticket (*Ticket_mode*).

Containing records from over 2 million clients, the scope had to be restricted by focusing the analysis on two specific stores and by assessing on them the potential impact of this framework. Aggregating the data per store, their KPIs were computed and stores were ranked by their QI. Diametric performances were found, where top-ranking stores (*i.e.,* highest QI) presented a better queue management than the others, constituting a proper benchmark.

However, not all stores could be directly compared. Clusters were created so that stores would only be compared if they had the same layout (number of servers) and a similar demand (within 1-σ_D of the average demand, μ_D).

Within one of the clusters, Guarda's, the best performing store in EDP, was found to be comparable with Bragança's store which, while having 91.2% of Guarda's demand, had almost 7 times the number of clients waiting more than 20 minutes. Their data was, then, further analysed quantitatively and qualitatively.

Comparing their WT histograms (Figure 4), the standard of excellence of Guarda was confirmed. Even though both distributions are right-skewed, Bragança presents a longer tail (*i.e.,* has a larger portion of occurrences far from the mode). These findings suggest a better management of the queue in Guarda where, on average, customers will endure shorter waiting-periods and a lower variation of WT among clients. Furthermore, both stores' demand was compared to assess if they are subjected to the same seasonality peaks per hour and per month. A table in MS Excel was created for each ticket-type (5 tables per store), as exemplified in Figure 5. After analysing their CVs, they were not only found similar, but it was also ascertained that the hourly distribution was the most variable one and, thus, the most propitious to simulate the demand peaks.

The remaining component, the demand-mix, was compared in Table 1, where both stores present a similar demand-mix, with 'E' tickets being the most sought and 'A' and 'D' tickets being the less representative. Moreover, while 'B,C' tickets had different shares, they had similar ASTs.

Table 1: Comparison of both stores' demand and KPIs

Ticket Type	Guarda Store			Bragança Store		
	Demand (%)	AST (min.)	AWT (min.)	Demand (%)	AST (min.)	AWT (min.)
A	19%	12:28	04:27	15%	09:49	08:38
B	21%	05:36	03:06	18%	06:49	08:06
C	19%	06:03	03:16	24%	06:47	08:29
D	9%	07:37	03:55	8%	06:59	08:24
E	31%	03:35	03:49	35%	05:34	08:31
Total	100%	06:31	03:41	100%	06:49	08:26

Validating that both stores were comparable, their KPIs were also confronted. While Bragança's AWT is 2.3 times higher than Guarda's, their AST does not present a significant difference (18'').

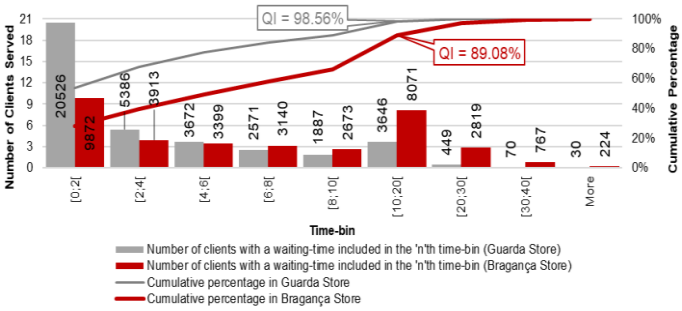

Figure 4: Overlapped Histograms of stores' waiting-time

Month	A 08	09	10	11	12	13	14	15	Grand Total
jan	26	101	137	122	42	45	118	101	699
feb	26	93	105	100	50	42	84	88	588
mar	24	116	87	104	40	51	104	107	635
apr	34	83	102	93	47	47	88	78	575
mai	33	78	116	86	37	33	97	78	558
jun	41	125	130	92	61	46	115	106	716
jul	42	126	107	94	60	36	99	95	660
aug	45	149	155	161	69	35	112	132	861
set	39	82	137	94	27	33	86	79	577
oct	24	82	90	76	38	35	77	67	489
nov	22	56	99	68	26	34	77	51	437
dec	9	59	69	54	33	29	60	51	365
Grand Total	365	1150	1334	1144	530	466	1117	1033	7160

Figure 5: Arrivals seeking A-tickets in Guarda

In fact, even though it is substantially faster (-36%) serving 'E' tickets, Guarda also takes more time serving 'A' (+27%) and 'D' (+9%) tickets. Therefore, this analysis pointing-out the existence of service best-practices (SBP) in both stores, but only one queue management best-practice (QBP) - Guarda's store.

To define the SBP, service performances were compared ranking employees by their CVs, where smaller values would mean more robust processes. Yet, for each subservice-type (33), only servers with ASTs lower than the stores' AST and with more served clients were qualified to this ranking. After all, considering Equation 2, high ASTs, or employees with low experience serving that type (leading to a small variance) could otherwise be qualified as SBPs.

Analysing the outcome, Bragança still had 35% of the identified SBP, showing that, even though this work will focus on assessing the impact on Bragança of adopting the remaining 65%, Guarda's store is also expected to improve its performance if it adopts these SBP (Table 2).

Table 2: Expected AST following SBP

Ticket Type	AST (min.)	Std. Dev. of AST (min.)
A	08:03	03:31
B	05:33	01:59
C	05:55	03:31
D	07:38	03:34
E	03:56	02:17

To define the QBP, qualitative analyses were performed by conducting interviews with both stores managers and with EDP. Neither one of the stores was found to follow Inline®, due to its fixed prioritization of tickets. But, while Bragança replaced this strategy by using a pure FIFO strategy, Guarda has implemented a tailored solution designed by its own staff who, according to their perceived demand-mix and the store's AST per ticket-type, has implemented a *fast-track* through which fast and frequent services (B,C,E) were being prioritized in one of the service-stations. EDP was surprised by none of the stores respecting Inline®, acknowledging the need for reviewing this component. But regarding the two different practices, the focus on reducing the AWT and its distribution tail was stressed. Therefore, the corporate view was leaning towards Guarda's solution. Still, considering Lean's approach of fostering common solutions, Bragança's store manager was asked why he had adopted a solution that was clearly hampering his KPIs. It was, then, understood that he was aiming at soothing his clientele, as it was prone to complain whenever someone was called by a different order than the entrances' order. EDP refuted this standing, explaining that, indeed, some clients should wait more than others, as they are less willing to wait, if they seek faster services (consistent with Nah's (2004) findings). Furthermore, EDP explained that regardless the undertaken sacrifices by Bragança to achieve a lower number of complaints, clients' satisfaction (measured by the Net Promoter Score, NPS) was higher in Guarda, suggesting a higher valorisation from the clients of the waiting-time than what Bragança's manager believed. Thus, Guarda's method was considered the QBP.

5.2. Designing the simulation model

To model Bragança's store, Simul8 a building-blocks based simulator was chosen. Starting by simply identifying the key activities, entities and queues, the customer journey was modelled according to Bragança's store specificities being split into four areas (*cf.* Figure 6).

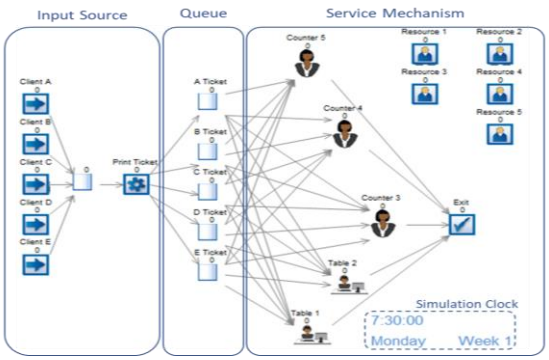

Figure 6: Bragança's store model in Simul8 (as-is)

In the *input source*, the five *start-points*, one *queue*, and one *activity* simulate the entrance of entities/clients and all its specificities, by differing the journey of each ticket-type. Their specific arrival rates were computed (applying Equation 1 to 2016's collected data) and introduced to model an exponential distribution. Two attributes (labels) were also set defining the routing (between the *source* and the *queue*) and the sought service. Since this model simulates the monitored time (*i.e.*, after printing a ticket), the duration of the "Print Ticket" *activity* was set to zero and its capacity to infinite, so that the transition between the entity's entrance and its respective queue would be instantaneous.

In the *queue*, the five *queue blocks* represent each one of the ticket-type queues. Entities will enter these queues according to their label, and wait for his turn in his FIFO ordered queue. Besides, for now, all queues are connected to the five service stations, since currently they are called by all service-stations.

In the *Service Mechanism*, the five service-stations, their five resources/servers, and the system's *exit block* model the customers' service, who are called by a FIFO order (*i.e.,* the entity with the highest WT) and have a service-duration that follows a label-based tailored distribution (the AST histograms were found to not follow any probabilistic distribution). Moreover, servers' efficiency levels (90%) were computed and lunch-breaks were modelled as it would not be plausible to consider they are always available. The efficiency ratio was computed by a tailored algorithm that for each 5-minutes slot considered a server as active, whenever he was calling/serving a ticket. Servers presence was also required in every station to calling an entity, together with a set of rules that mimicked the real system: a station can only server one customer at a time; there are no set-up, nor travel-times, since they are encompassed by the AST; and servers cannot leave the station until the client is fully served.

Finally, to mimic the dynamic facet of the store a *simulation clock* was set to run 5 days a week along the year, abiding by their opening times, and vacating the store at the end of each day (*i.e.,* after serving all customers).

5.3. Validation of the Model

After defining all the KPIs (AWT and QI) the model was set with: zero warm-up period (as there was no stock of entities between days), a data collection period of one year (to be compared with 2016's), and with 4-runs (as suggested by Simul8 calculator for a 95% confidence-level). Tests were performed to guarantee that the model was mimicking the real environment by running simulation subsets and monitoring their behaviour. Once verified, the simulation estimated a AWT offset of -1.63%, and a QI offset of +3.6% comparing with the real ones. Conveying a high accuracy rate (*ca. 97.5%)*, the model was considered to be valid and accurate, especially considering the usual 95% benchmark (Sargent, 2013), while being also acknowledged a slight tendency to underestimate the queuing effect.

5.4. Simulation of different solutions

The simulation phase encompassed 23 scenarios based on either the current Inline® system, or on the identified QBP. As the latter one proved to lead to better results, Guarda's concept (the *fast-track*) was progressively developed until every service station was calling through a triage system. Another addition to the original concept was the *time-triggers*. Showing to be a crucial component, as it improved both KPIs, different sets of *time-triggers* were tested. Still, the one leading to better results was based on Equation 3, where, for each ticket-type 'i', a

ratio for the desired AWT was pre-set based on its: demand weight, AWT_i, AST_i. Hence, one of the ticket-types ('x') would set the scale, being equal to one (the most prioritized) and the others would have a AWT 'n' times higher than x's.

$$Ratio_i = n = \frac{max_{\forall i}[D_i(\%) * AWT_i/AST_i]}{D_i(\%) * AWT_i/AST_i} \quad (3)$$

Confronting their performances (summarised in Figures 7 to 9), a scenario should be considered as being improving the system, if it decreases the AWT and increases the QI. Furthermore, to raise the likelihood of its implementation in Bragança (*cf.* section 5.1), a scenario should not significantly raise the current CV of the store.

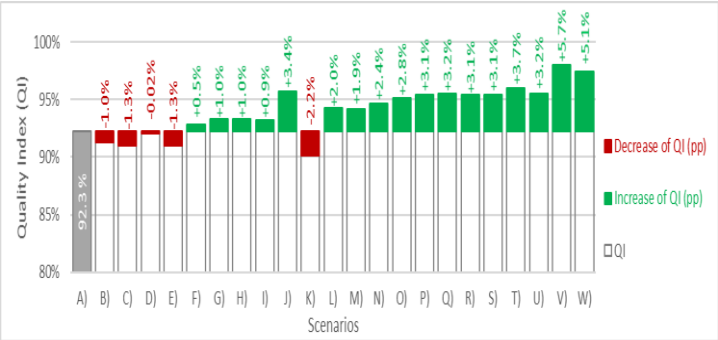

Figure 7: QI Performances of each simulated scenario (1st KPI)

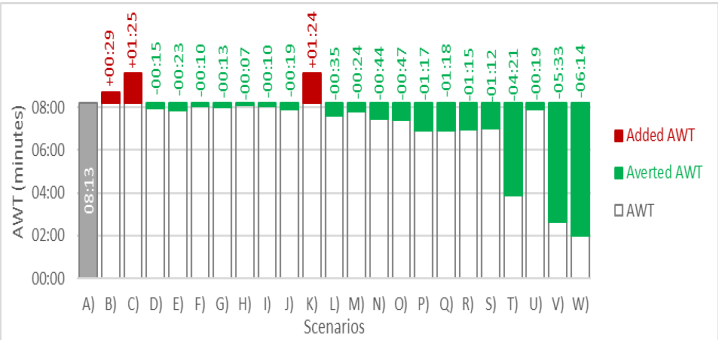

Figure 8: AWT Performances of each simulated scenario (2nd KPI)

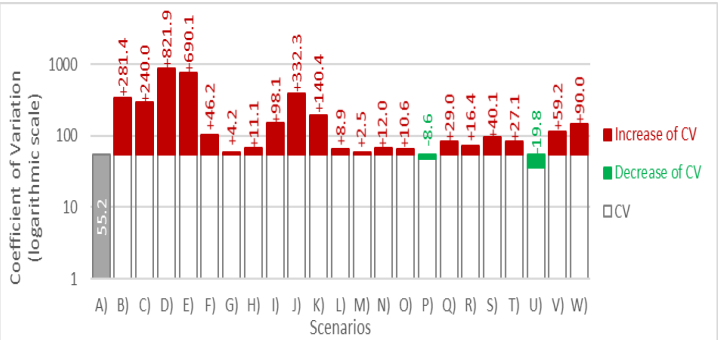

Figure 9: CVs of each simulated scenario (not a KPI)

Following Inline®'s programmed priority levels (scenario B, with Bragança's and C with Guarda's) would have led Bragança store to worse KPIs. Yet, results suggest that there was an effort by Inline®'s team to tailor each solution to each store (as B leads to better KPIs), even if it proved unfruitful. Scenarios D and E still tried to test other rigid prioritization methods as ranking ticket-types per demand-share and calling first expired tickets (7 minutes, near EDP's goal), respectively.

The two first improving scenarios were: F, where Guarda's solution was applied; and G, where the *time-triggers* rationale (*cf.* Equation 3) complemented Guarda's solution, successfully decreasing its CV. Further asserting that the system compelled for a more flexible solution, scenarios H to K simulate different solutions that dedicate either fast-tracks, or servers to specific tickets subsets, which have failed to promote better performances, especially due to a steep increase of the CV, which could make the store manager less likely to adopt the solution.

From scenarios L to P, the number of service-stations increased progressively, leading to the best performance so far, in P, where an improvement of both KPIs and lower CV than in 2016 was witnessed. The three scenarios 'Q,R,S' run P with other time-triggers based on solely the demand, or the AST, and failed to convey better KPIs, which was helpful when proposing scenario P to the manager without getting into mathematical details.

Having ascertained that scenario P was to be proposed, the addition of the SBP to the model was simulated in scenario T, where, with a slight increase of the CV, the QI would increase by 4% and the AWT would decrease 53%.

Three final scenarios complemented this analysis by simulating the system with four and six servers. After adopting both the QBP and

SBP, even with four servers (U), better KPIs would be achieved than the current model (A). Furthermore, if a sixth server was to be hired, results indicate that a plateau may have been reached regarding the *fast-track* impact, as scenario W leads to a better AWT by setting a FIFO discipline in the sixth server. However, scenario V still proves to foster a more robust solution by leading to a better QI (98% *vs.* 97.4%) and a lower CV (114.4 *vs.* 145.2).

6. Conclusions

Applying the proposed Framework, a future state for EDP was proposed (*cf.* Figure 10) that ultimately leads to an increase of both KPIs and, hence, fosters higher customers' satisfaction-levels as similar examples in EDP have proved.

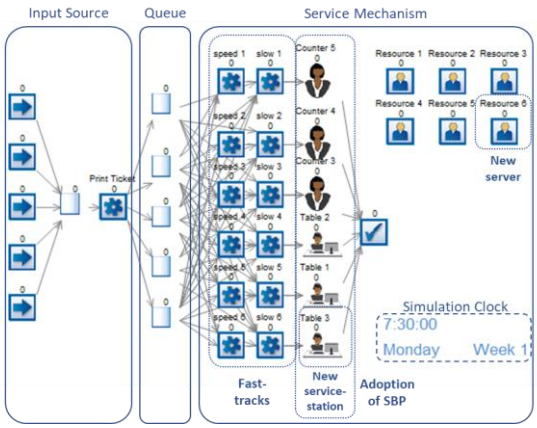

Figure 10: Bragança's store model in Simul8 (to-be)

The system was successfully characterised by Queuing Theory, but became unmanageable to model. Thus, a symbiotic combination of Simulation and Lean methodologies was used to model the impact of unveiled best-practices.

Even though, some opportunities for future-work, as widening the simulation scope, or redesigning the service-processes, arose, the main goal of this work has been achieved.

Despite encompassing some assumptions, the model has reached an accuracy rate of about 97.5% when simulating the real scenario of 2016, with approximately half (*ca.* 2.5%) of the usual 5% error-margin benchmark in most simulations (Sargent, 2013).

Moreover, most improvement opportunities do not imply any significant investment and can be promptly implemented. This is the case of the adoption of a new queuing discipline (QBP), and the acknowledgment of the SBP, as now that they were both identified only one-day training is required. These proposed solutions will not face significant resistance neither, having its phased-implementation already begun, since the store manager concerns were taken into consideration and still can make Bragança's the 3rd ranking store in EDP. On the other hand, the third proposal (*cf.* Table 3) would, in fact, require a considerable investment.

Table 3: Proposed solutions

Recommended Solutions	AWT	%Var. AWT	QI	%Var. QI	CV	Required Investment
(1) Adoption of QBP	06:56	-15.6%	95.4%	+3.4%	46.6	N/A
(2) Adoption of SBP + (1)	03:52	-53.0%	96.0%	+4.0%	82.3	Low (1-day training)
(3) Hiring a sixth server + (2)	02:40	-67.6%	98.0%	+6.2%	114.4	28 000€/year

Still, considering the estimated decrease (2'41'') on the contractual ('A') tickets' AWT and their annual demand (*ca.* 5 700), if, on average, 'A' clients would be willing to spend 1.83€ more per each avoided minute of waiting, EDP would even profit from the hiring of this new server.

Overall, this Framework may improve EDP's 2 million annual visitors' journey, as it can be applied to every store and decrease their clients' AWT solely by adopting suitable best-practices from other comparable stores. Staff concerns can also be included (such as Bragança's, in this work) to ensure their embracement of these proposed solutions. Besides, at a corporate level, the Framework will also help guaranteeing EDP's leading position by improving the stores' QI which is monitored by the Portuguese Energy Market regulator (ERSE). The extension of this work to the remaining 40 stores is therefore expected, as EDP was involved in the project and is already adopting some of its findings.

Thanks to the mitigation of the unintended queuing consequences, customers satisfaction will increase, hence promoting a higher willingness to buy EDP's products according to the literature. Indeed, if they are promptly implemented and comprehended, stores may sustain the benefits of the uncovered solutions in the long-term. But, either way, what seems settled is that, on the short-term, stores will experience its crucial impact by a lowering of their AWTs and by a potentially increasing of its annual revenues in a business where, on average, one residential customer represents a yearly profit after taxes of 50€ (Brandão, 2005).

7. References

Balci, O. (2003). Verification, validation, and certification of modeling and simulation applications. *Proceedings of the 35th conference on Winter simulation* (pp. 150-158).

Banks, J., CARSON II, J. S., & Barry, L. (2005). *Discrete-event system simulation* (4th ed.). Pearson.

Bicheno, J., & Holweg, M. (2000). *The lean toolbox* (Vol. IV). Buckingham: PICSIE books.

Brandão, N. (2005). Impact on profitability of the Energy Market regulation. Lisbon: EDP.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research, 201*(3), 921-932.

Creemers, S., & Lambrecht, M. (2010). Queuing models for appointment-driven systems. *Annals of Operations Research, 178*(1), 155-172.

Daellenbach, H., McNickle, D., & Dye, S. (2012). *Management science: decision-making through systems thinking.* Palgrave Macmillan.

Engström, T., Jonsson, D., Medbo, L., & Engström, T. J. (2006). Production model discourse and experiences from the Swedish automotive industry . *International Journal of Operations & Production Management, 16*(2), 141–158.

Gautam, N. (2008). *Queuing theory, Operations Research and Management Science Handbook.* London: CRC Press, Taylor and Francis Group.

Gutacker, N. S. (2016). Waiting time prioritisation: Evidence from England. *Social Science & Medicine*, 140-151.

Harchol-Balter, M., Osogami, T., Scheller-Wolf, A., & Wierman, A. (2005). Multi-server queuing systems with multiple priority classes. *Queuing Systems, 51*(3-4), 331-360.

Hillier, F.S., & Lieberman, G.J. (2010). *Introduction To Operations Research* (9th ed). McGraw-Hill.

Ndukwe, H. C., Omale, S., & Opanuga, O. O. (2011). Reducing queues in a Nigerian hospital pharmacy. *Pharmacy and Pharmacology, 5*(8), 1020-1026.

Newell, C. (2013). *Applications of queuing theory* (Vol. IV). Springer Science & Business Media.

Nova, A. M. (2008). *Apontamentos de Simulação.* Departamento de Engenharia e Gestão.

Pidd, M. (2004). *Computer simulation in management science* (5th ed.). Chichester: John Wiley & Sons, Ltd.

Rossetti, M. D. (2015). *Simulation modeling and Arena.* John Wiley & Sons.

Rubinstein, R. Y., & Kroese, D. P. (2016). *Simulation and the Monte Carlo method.* John Wiley & Sons.

Sachidananda, M., Erkoyuncu, J., Steenstra, D., & Michalska, S. (2016). Discrete Event Simulation Modelling for Dynamic Decision Making in Biopharmaceutical Manufacturing. *Procedia CIRP, 49*, 39-44.

Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of Simulation, 7*(1), 12–24.

Xu, X., Lian, Z., Li, X., & Guo, P. (2016). A Hotelling queue model with probabilistic service. *Operations Research Letters, 44*(5), 592-597.