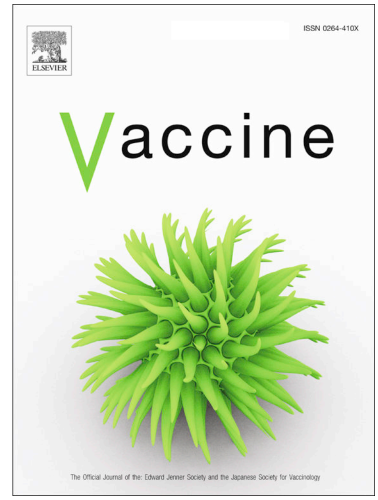
Journal Pre-proofs



A queueing Network approach for capacity planning and patient Scheduling: A case study for the COVID-19 vaccination process in Colombia

Carlos Franco, Nilson Herazo-Padilla, Jaime Andrés Castañeda

PII: S0264-410X(22)01199-9
DOI: https://doi.org/10.1016/j.vaccine.2022.09.079
Reference: JVAC 24382

To appear in: *Vaccine*

Received Date: 11 August 2022
Revised Date: 20 September 2022
Accepted Date: 26 September 2022

A Queueing Network Approach for Capacity Planning and Patient Scheduling: A Case Study for the COVID-19 Vaccination Process in Colombia

Carlos Franco*
Associate Professor
School of Management and Business
Universidad del Rosario
CL 12 C 6 25
Bogotá, D.C. 111711, Colombia
+57 (1) 297-0200 ext. 3942
carlosa.franco@urosario.edu.co
ORCID: 0000-0002-8288-8044
* Corresponding author

Nilson Herazo-Padilla
Assistant Professor
School of Management and Business
Universidad del Rosario
CL 12 C 6 25
Bogotá, D.C. 111711, Colombia
+57 (1) 297-0200 ext. 3991
nilson.herazo@urosario.edu.co
ORCID: 0000-0001-8936-342X

Jaime Andrés Castañeda
Associate Professor
School of Management and Business
Universidad del Rosario
CL 12 C 6 25
Bogotá, D.C. 111711, Colombia
+57 (1) 297-0200 ext. 3901
jaime.castaneda@urosario.edu.co
ORCID: 0000-0002-3458-7685

Abstract: This paper considers the problem of patient scheduling and capacity planning for the vaccination process during the COVID-19 pandemic. The proposed solution is based on a non-linear mathematical modeling approach representing the dynamics of an open Jackson Network and a Generalized Network. To test these models, we proposed three objective functions and analyzed different configurations of the process corresponding to various levels of the models' parameters as well as the conditions present in the case study. To assess the computational performance of the models, we also experimented with larger instances in terms of number of steps or stations used and number of patients scheduled. The computational results show how parameters such as the minimum percentage of patients served, the maximum occupation allowed per station and the objective functions used have an impact on the configuration of the process. The proposed approach can support the decision-making process in vaccination centers to efficiently assign human and material resources to maximize the number of patients vaccinated while ensuring reasonable waiting times, number of patients in queue and servers' utilization rates, which in turn are key to avoid overcrowding and other negative conditions in the system that could increase the risk of infections.

Keywords: Jackson Network Model, Generalized Queueing Network, Non-Linear Mixed Integer Programming, COVID-19 Vaccination, Capacity Planning.

1. INTRODUCTION

Vaccination centers are locations set up to immunize large numbers of people within short time frames. The final aim of setting up vaccination centers is to accelerate disease control through a rapid increase in vaccination coverage. However, this goal may not be attainable without proper vaccine distribution to vaccination centers and effective vaccine application (Gianfredi et al., 2021). In this context, planning and managing the capacity of vaccination centers is vital to achieve the necessary high vaccination rates to control disease spread.

Planning and managing vaccination centers may include deciding on several aspects such as the location and layout of the center, the number and type of staff and the required performance to achieve given vaccination quantity targets, among others (Gianfredi et al., 2021). The implementation of COVID-19 vaccination centers attempts to address some of these issues. For example, Signorelli et al. (2021) presented the implementation of a “vaccine islands” model where COVID-19 vaccines were applied in up to 10 “islands”, each with two vaccination servers, divided into two parallel rows of five islands separated by a central corridor. This setup allowed to administer up to 6,000 shots per day while reducing the overall time a patient spent in the whole process and crowds. Similarly, Yogev et al. (2021) presented the implementation of a COVID-19 vaccination site where six vaccination servers were organized in a C shape with the vaccine preparation server at the heart of the C. This setup allowed a productivity of 147 shots per hour while improving control and supervision of vaccine administration, patient flow, dose preparation burnout and safety.

While these and other experiences on vaccination centers can provide some insight into the challenges of planning and managing these centers (Gianfredi et al., 2021), they are not suited for determining how to best plan and manage a vaccination center. In this regard, modeling and simulation approaches can be used to test a range of design alternatives for vaccination centers and determine their best operating parameters. For example, Pryor et al. (2021) developed a resource calculator based on scheduling models to determine the resource allocation needed (staff and room capacity) to maximize the number of COVID-19 vaccines applied per hour. Similarly, Wood et al. (2021) developed a discrete event simulation model to determine arrival and service rates for accommodating practical service and queueing levels in a COVID-19 vaccination center with up to six immunizers.

Our work develops a mathematical modeling framework that generalizes patient scheduling and resource allocation decision-making for vaccination centers considering congestion and workload constraints. These are decisions that hospitals and vaccination centers must make under new conditions brought by the pandemic, which make the vaccination process different from typical vaccination processes such as seasonal flu vaccination. The starting point is a three-step COVID-19 vaccination process observed in a healthcare provider in Bogotá, Colombia where, in total the city has to vaccinate in the first phase approximately 5 million of people, more details in https://saludata.saludcapital.gov.co/osb/index.php/datos-de-salud/enfermedades-trasmisibles/covid-19-vacunometro/. In this process, the personal information of a patient is verified, she is vaccinated and, finally, her personal and vaccination information is registered in a governmental platform. All COVID-19 vaccination centers in Colombia follow this same three-step vaccination process, though in different scales.

The remainder of this paper is organized as follows. Section 2 presents a review of related literature. Section 3 describes the COVID-19 vaccination center case that serves as the starting point of our modeling effort. Section 4 formulates the Jackson Network model and the G/G/m model that generalizes the problem. Section 5 reports the computational experiments, their results and the theoretical and managerial implications. Finally, Section 6 draws conclusions and opportunities for future research.

2. LITERATURE REVIEW

This section does not intend to provide a comprehensive literature review but to point to the kind of studies that are relevant to our work. Sections 2.1 and 2.2 highlight contextual factors that are relevant to understand some of the challenges for the operation of vaccination centers, which we capture through the way we model demand and some of our constraints. And Section 2.3 points to modeling approaches relevant to analyze vaccination processes to help position our work.

2.1. Vaccine distribution challenges

Given the particular conditions of the current COVID-19 pandemic caused by the outbreak of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), many countries face multiple challenges for the distribution of COVID-19 vaccines to their population such as the heterogeneity of vaccines, the different cold chains needed for each vaccine type and the mismatches between the rate of vaccine production, the supply of logistics infrastructure, the medical and technical staff at vaccination centers and the people to be vaccinated. Additionally, misinformation, hesitancy and unwillingness of part of the population to vaccination also pose challenges for accurately assessing the demand for a given distribution time window. Other challenges for the distribution of these vaccines are the need for mass vaccination across the entire territory of the country, the volume of patient enrollment and scheduling efforts needed, perishability of doses and social distancing in vaccination centers, among others (Goralnick et al., 2021).

For example, Mills & Salisbury (2021) mention that, as of the end of 2020 it was unclear how resources will be split between different COVID-19 vaccination sites in UK (e.g., primary care networks and mass vaccination centers) and that vaccination programs would be constrained by vaccine availability. This was the case specially in developing countries (Sheikh et al., 2021). They mention that reported vaccination rates in the UK under normal conditions are 10.7 million doses per year (for seasonal flu), where it is estimated that vaccination sites report around 200-500 doses per day per site, which adds up to just over 3 million doses per week. These rates were not sufficient to quickly vaccinate the 70% of the population needed to reach herd immunity, which raised the need for suitable strategies to accelerate and optimize vaccination.

Eshun-Wilson et al. (2021) conducted a survey in US to inform vaccine distribution strategies that are aligned with public preferences for COVID-19 vaccination. Their data suggested that making vaccination easy (e.g., single over two doses and reduce waiting times at vaccination sites) and promoting autonomy by offering choices of vaccine brands and locations may increase vaccine uptake, and that vaccine mandates could reduce vaccination in those who are hesitant. Many of this information is important to manage the demand in a vaccination process and indeed proved to be of use when implemented for mass vaccination (Wang et al., 2021). It is also key when used in different operations research modeling strategies to optimize resource allocation and vaccination rates.

Finally, the analysis presented by Weintraub et al. (2020) compiles lessons learned from past pandemics and vaccine campaigns for successful COVID-19 vaccine delivery. They mention that demand must be generated through communication campaigns tailored to specific populations. They suggest that qualified organizations and expert coalitions must determine evidence-based vaccine allocation strategies.

2.2. Vaccine allocation challenges

Goralnick et al. (2021) claim that a hybrid approach using conventional vaccination sites and high-throughput, large-venue mass-vaccination sites such as stadiums and convention centers is essential for COVID-19 vaccine delivery to keep up with vaccine distribution. By January 11, 2021, for instance, the US had distributed 22.1 million doses but administered only 6.7 million vaccinations since vaccination plans relied on leveraging conventional healthcare sites. Additionally, they mention that partnerships that draw on innovation and expertise from outside healthcare are valuable to identify bottlenecks and accelerate vaccine delivery. More recent studies also demonstrated the effectiveness of this approach (Lemaitre et al., 2022; Roy et al., 2021; Wrigley-Field et al., 2021). These findings highlight a need for planning strategies for vaccination centers with the capacity to quickly adapt their operations according to the demand.

The study presented by Roy et al. (2021) suggests that COVID-19 vaccine allocation strategies that incorporate epidemiological factors that account for the vulnerability of zones to the epidemic spread may enhance recommendations and aid policy making, as their analysis on real demographic and infection data from NY state shows that a small fraction of zones tends to exhibit a high resource demand due to their vulnerability to the pandemic spread. Similarly, (Lemaitre et al., 2021) propose COVID-19 vaccine allocation strategies that account for spatial heterogeneities in transmission rate and disease history among Italian provinces under constraints on vaccine supply and distribution logistics. These strategies significantly outperform simple allocation strategies based on incidence, population distribution or prevalence of susceptible in each province. From these studies we can recognize that COVID-19 vaccination centers not only need to be highly efficient but also very flexible to adapt to changing demand conditions.

Finally, Zhou et al. (2021) show through a simulation study with data from Guangzhou, China, that herd immunity is heterogeneously distributed in space, for which reason COVID-19 vaccine allocation strategies should be spatialized. Accordingly, their simulations show that an allocation strategy based on space and age is more effective to control the epidemic duration than a random strategy and strategies focused solely on space or age. These findings also highlight that vaccination sites should have their own configuration and capacity to adapt. This was further demonstrated in a case study involving several Italian provinces conducted by (Lemaitre et al., 2022).

2.3. Vaccination center planning and management

Some experiences on COVID-19 vaccination center design have shed light on some challenges of planning and managing these centers. For example, Yogev et al. (2021) presented two setups for the arrangement of vaccination centers. One setup had the vaccine preparation module in the backstage, while the other had a C-shaped arrangement with the vaccine preparation module at the center of the C. Comparing both setups allowed the authors to share some insights into patient flow, personnel burnout, and safety, among others. Although these experiences shed some light into the design of vaccination sites for Covid-19, the authors, however, did not take into account capacity planning considerations.

Signorelli et al. (2021) emphasize that sharing best practices in immunization delivery is fundamental to achieve population health during health emergencies. As a contribution to this, they implemented a “vaccine islands” setup, where the anamnestic (medical history) and inoculation areas are unified in “islands” each with four and two anamnestic and inoculation modules, respectively. Although this setup was efficient for the chosen location, it is unclear whether similar setups will adapt to different locations.

2.3.1. Operations Research strategies

Pryor et al. (2021) use scheduling models to develop a resource calculator to increase COVID-19 vaccination rates. They stress that vaccination efforts may be hampered by supply, delivery, storage, patient prioritization, administration infrastructure or logistics problems, which support the development of Operations Research (OR) strategies. These constraints for rapid vaccination were specially observed in developing countries (Sheikh et al., 2021). For example, Palmer et al. (2017) present a systematic literature review on the application of OR methods for modeling patient flow, which is a relevant issue in COVID-19 vaccination processes.

2.3.2. Simulation strategies

Simulation models have also been proposed to manage and maximize the throughput of COVID-19 vaccination centers. For example, Wood et al. (2021) developed a discrete event simulation-based approach to plan the demand (arrivals) and capacity (service rates) of vaccination centers while accommodating practical service and queueing levels. Similarly, Asgary et al. (2021) developed a hybrid discrete event and agent-based simulation to model a drive-thru mass vaccination facility. They also coupled it with a machine learning model to assess the potential outcomes of the simulation much faster than the simulation itself, which can help authorities quickly compare different scenarios.

2.3.3. Mathematical programming and optimization models for the appointment scheduling and capacity planning problem

The capacity planning problem for a COVID-19 vaccination center can be modeled as a special case of the largely studied appointment scheduling problem. The main characteristics of this problem include two conflicting objectives that are considered in three objective functions in our study. These objectives are the minimization of waiting times for patients and the maximization of the utilization of resources (such as admin and medical staff). Additional characteristics of this problem related to a typical COVID-19 vaccination process are the presence of stochastic interarrival and service times and the use of multiple often sequential steps or stations through which patients need to go through. Early studies such as the one presented by Anparasan & Lejeune (2019) consider the resource deployment problem for epidemic outbreaks, however, this would only be a part of the problem in the context of the COVID-19 pandemic.

Kuiper et al. (2021) highlight an important gap between the mathematical optimization methods presented in the literature for the patient appointment scheduling problem and healthcare practice. The main reasons for this as found in Kuiper et al. (2021) are the uncertain and dynamic nature of demand, the dependencies in the process structure and the variety in services and in patient and resource characteristics, among others. This reveals the need for further work on suitable optimization approaches for appointment scheduling problems in healthcare, such as the capacity planning problem for COVID-19 vaccination centers.

Several approaches have been presented in the literature to study appointment scheduling problems in clinics. For example, Nguyen et al. (2018) develop a linear programming model that determines the required number of physicians over a finite horizon to achieve targets on appointment lead-times. Moreno & Blanco (2018) develop an integer linear programming model to plan the admission of patients for receiving a set of prescribed clinical services minimizing admission dates and length of stay. Dogru & Melouk (2019) propose a simulation-optimization approach that sequentially schedules appointments that minimize the weighted expected costs of patient waiting time and physician idle time and overtime. Alizadeh et al. (2020) solve a problem with a similar objective function using mixed integer linear programming. Instead of minimizing a typical expected cost function (e.g., patient waiting time and physician idle time and overtime), Sang et al. (2021) minimize any quantile of the cost distribution, which could be a better approach if service level and risk measures are considered. Finally, Yan et al. (2021) formulate a stochastic overbooking model that maximizes expected profit (revenue generated from accepted appointment requests minus the cost from patient waiting time and physician overtime).

Maximizing profit is not characteristic of vaccination processes. Additionally, the stochastic nature of a COVID-19 vaccination process requires stochastic approaches for the solution, which are not always considered (cf. Nguyen et al., 2018; Yan et al., 2021). Several approaches assume a fixed and known demand (e.g., Moreno & Blanco, 2018; Alizadeh et al., 2020; Sang et al., 2021), while the demand is usually unknown in a COVID-19 vaccination process. Several approaches also discretize time (e.g., slots, hours) to facilitate the assignment of appointments and services (e.g., Alizadeh et al., 2020; Dogru & Melouk, 2019; Moreno & Blanco, 2018; Nguyen et al., 2018; Yan et al., 2021), which does not usually apply in a COVID-19 vaccination process. Finally, most of these studies do not consider capacity planning but rather optimize the use of available resources (cf. Nguyen et al., 2018). Thus, several configurations and/or characteristics commonly considered in existing studies are not similar to those observed in a COVID-19 vaccination process.

2.3.4. Contributions of this paper

Our work contributes to modeling strategies for patient scheduling and planning and managing the capacity of a sequential COVID-19 vaccination process according to local needs. It also considers the stochastic nature of the vaccination process by incorporating stochastic parameters in both the demand and the service provision, and analyzes different scenarios of operation that could be produced by external conditions (e.g., staff unavailability, unforeseen delays, sudden increases in arrivals). These cases can be analyzed using the different objective functions proposed (Server assignment (SA), Patient scheduling (PS) and SA + PS) and setting different servers' utilization rates. These modelling strategies on one hand, are essential for determining the necessary additional capacity required in such cases. On the other hand, the models proposed can be adapted to different healthcare processes such as other type of vaccinations campaigns or different related processes, for which, one of the main objectives is to determine the capacity of each station and to schedule the patients arrivals given some operational constraints. They can also be essential for cases where new healthcare processes need to be designed or the performance of some KPIs needs improvement.

3. CASE STUDY

3.1. Foundation overview

The Foundation was founded in the 1970s as a private non-profit organization focused on treating cardiovascular disease in children from low-income families. In the 1980s, the Foundation formally established a social program to pursue its goal, and, to date, more than 5000 patients have been successfully treated through this program. In that same period, the Foundation built the facilities that house its current hospital with the help of private donations and expanded its scope to also treat adults with cardiovascular disease. Nowadays, the Foundation has become established as a private non-profit teaching and research hospital for cardiovascular and tertiary care, and it is recognized as one of the best Latin-American hospitals by the América Economía ranking. As a renowned healthcare provider in Colombia, they joined Colombia's governmental effort of massively vaccinating its population against COVID-19.

3.2. COVID-19 vaccination process

Colombia started its COVID-19 vaccination program on February 2021. At the time, since vaccines were applied according to prioritized population groups, the vaccination program worked by scheduling vaccination appointments and involved the following overall activities. First, upon receiving vaccine batches, the Colombian government allocated vaccines to departmental authorities according to the information on the prioritized population groups that these authorities shared with the government. Following, departmental authorities together with healthcare providers scheduled vaccination appointments for prioritized population groups. Next, based on the scheduled appointments, departmental authorities allocated the necessary vaccines to healthcare providers. Finally, on vaccination day, healthcare providers vaccinated all people that showed up for their appointment.

After all prioritized population groups were vaccinated, the allocation of vaccines to departmental authorities is done based on their demographic weight and unvaccinated population. In this new phase of the vaccination program and for large urban areas, departmental authorities together with healthcare providers set up bigger vaccination centers, and departmental authorities allocate the necessary vaccines to these vaccination centers according to the requests of the different healthcare providers operating in a given vaccination center and to the storage and handling capacity of the center. For other areas, departmental authorities allocate vaccines to individual healthcare providers according to the latter's requests. In this new phase, vaccination centers could schedule appointments, reserve the right to refuse entrance according to vaccine availability or work with a mix of these two approaches. Regardless of the approach, patient admission is contingent on vaccine availability, which guarantees a normal operation of the centers.

Regardless of whether the vaccination is scheduled or not, an individual goes through a three-step vaccination process.[1] First, when the individual arrives, her personal, consent and symptom screening information is verified at the verification step or station. At the time, three staff members, each serving one individual, verified this information at the Foundation. Second, the individual goes to a vaccination server to get her shot. At the time, the Foundation had 10 vaccination servers, operating most of the time with eight servers and enabling the remaining two as needs arose. Finally, the individual goes to the registration station where her personal and vaccination information is registered in a governmental web platform. At the time, ten staff members, each serving one individual, registered this information at the Foundation. Figure 1 shows a flowchart of this vaccination process at the Foundation with performance data based on a time study.

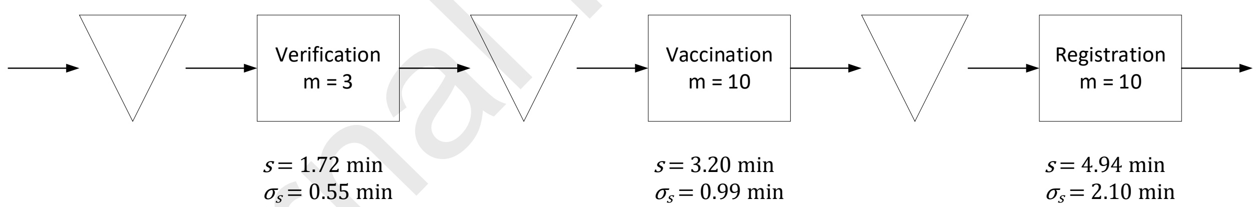


Figure 1. Vaccination process flowchart at the Foundation.

This general vaccination process constitutes the basis of our optimization model to support decisions on the number of staff members (servers) to be assigned to each station, and the number of patients to be scheduled in a day considering congestion and workload constraints. Our initial

[1] The process is the same in all healthcare providers and vaccination centers authorized to apply covid-19 vaccines in the country. However, there could be differences in the number of staff members assigned to each step of the process.

analysis of the optimization model assumes the exponential distribution is a good fit to model the three service times. Thus, we use an open Jackson Network model for this analysis. As we mentioned above, we collected data on service times and, according to an input analysis performed in Stat::Fit® Version 3 by Geer Mountain Software Corp., the exponential distribution is a good fit to model vaccination and registration times. Thus, we perform an additional analysis where we relax the exponential distribution assumption and model the three service times with a general distribution.

4. PROBLEM DEFINITION AND MATHEMATICAL FORMULATION

According to the case study described in Section 3, a corresponding queueing network of the system is presented in Figure 2.

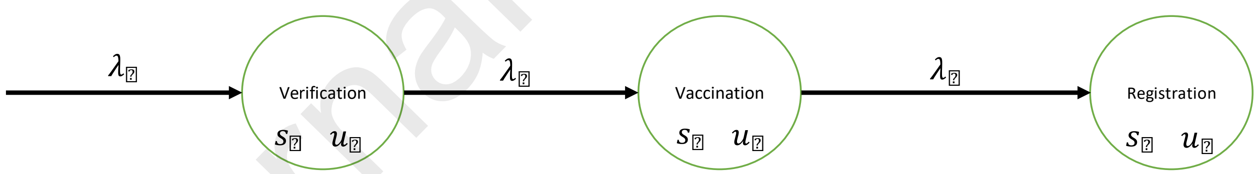


Figure 2. Queueing network of a typical vaccination process.

Patients arrive to the first step or station at an arrival rate λ_1 according to a Poisson process where the no-show probabilities are not considered. The number of parallel servers in each station is denoted with s_j and the service time is exponentially distributed with mean $\frac{1}{u_j}$. Given that the model does not consider deferral or no-show probabilities, it follows that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. The following are the assumptions used for the mathematical model:

- Arrival rate λ_1 follows an exponential distribution and it is considered as a variable of the model
- Servers has a process time s_j which follows an exponential distribution
- Rework is not presented in the vaccination process neither in registration
- Capacity of queue and system is infinite
- The discipline of queue is FIFO
- The preparation of vaccines is not considered as a station given that we are studying the process where patients are involved

4.1. Jackson Network model

The first model (see Appendix A) corresponds to a nonlinear mathematical Jackson Network composed of a set O of different stations, where each station represents a service provider. There are totally n patients to be scheduled in a day of H hours. The decision maker needs to define the following parameters: the minimum percentage of patients served per day (denoted by α), the minimum and maximum occupation or utilization (in percentage) of each station, denoted by c_o and b_o, respectively, and the maximum patient waiting time in queue and number of patients in queue in each station, denoted by d_o and e_o, respectively. The random parameters are related to the service time and service rate in each station, denoted by ti_o and μ_o, respectively.

The decision variables are the frequency of patients scheduled per hour λ, the number of servers assigned to each station s_o, the probability of idle time in a specific station $\pi 0_o$ and the average number of patients in queue in each station Lq_o.

Three different objective functions are evaluated individually (Equations 1-3). The first one minimizes the total number of servers assigned while the second one maximizes the number of patients scheduled. The third objective function considers the previous two objectives together and normalizes them by dividing each of them by its best result.

Equations 4 and 5 define constraints ensuring that patients are scheduled within a day's working hours, defining the time at which each patient must arrive to the vaccination center. It can be noted that depending on the objective function, this constraint can generate different results even if the parameters n and H are constant. Equations 6 and 7 define the constraints for the utilization of each station. The Jackson Network metrics are calculated in Equations 8 and 9, where the former is the proportion of idle time of each station and the latter is the average number of patients in queue in each station. Equations 10 and 11 define the constraints for the maximum average number of patients in queue in each station and the maximum average patient waiting time in queue in each station, respectively. Finally, the constraints in Equations 12-15 represent the types of variables.

4.2. Mathematical reformulation

The mathematical model presented in Section 1.1 is unsolvable since decision variables are in a summation (Equations 1 and 3), in factorial numbers (Equations 8 and 9) and as super indexes (Equations 8 and 9). To solve this problem, we propose a reformulation based on the considerations presented in Appendix B.

4.3. Generalized Network Model (G/G/m)

Given that Model 1 follows the M/M/s queueing model assumptions, we propose a second model with nonexponential interarrival and process time distributions, i.e., a G/G/m model. We refer to this model as Model 2. It uses the same notation of Model 1 and we add the elements included are presented in Appendix C.

In Model 2, we use the same three objective functions presented for RF Model 1 and they are presented in Equations 39-41. Equations 42-46 replicate Equations 22-26 from RF Model 1, i.e., the working hours and utilization constraints. Equation 47 defines the interarrival time squared coefficient of variation for the first station, which depends on the number of servers assigned by the binary variable. We assume that such coefficient remains the same regardless of whether the arrival rate changes. This assumption is grounded on the fact that we are modeling a flow process and is further formalized in Equations 48 and 49. Equations 50 and 51 define the Kingman approach (Hopp & Spearman, 2011) to the average waiting time. Equations 52 and 53 represent the bounds of the average patient waiting time and the average number of patients waiting, respectively. Finally, Equations 54-59 represent the type of variables used.

5. COMPUTATIONAL EXPERIMENTS AND RESULTS

5.1. Case study results

To assess the performance of RF Model 1 and the objective functions over the case study data, we analyzed the impact of different values of the parameters on the configuration of each station and the performance metrics. Table 1 shows the values we tested:

Table 1. Parameter values to test RF Model 1's performance.

Parameter	Values						
α	0.80	0.85	0.90	0.95	1	--	--
b_o	0.70	0.75	0.80	0.85	0.90	0.95	0.99
n	500	1000	1500	…	…	…	$Max. solvable\ value$

As an example, the results obtained with a single parameter configuration ($\alpha = 0.8$, $b_o = 0.7$ and $n = 500$) are presented in Figures 3, 4 and 5: the first shows three configurations of the process, each corresponding to the solution of one of the objective functions (SA, PS and SA + PS, in that order), and the remaining two show the performance metrics for each station along the three objective functions.

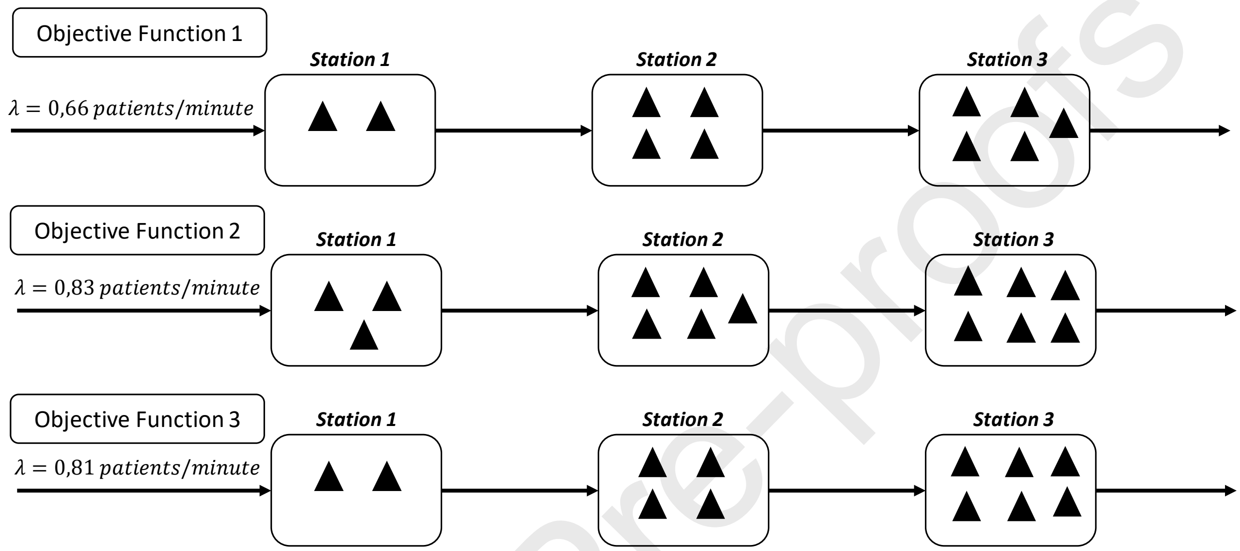


Figure 3. Configuration results for the three objective functions.

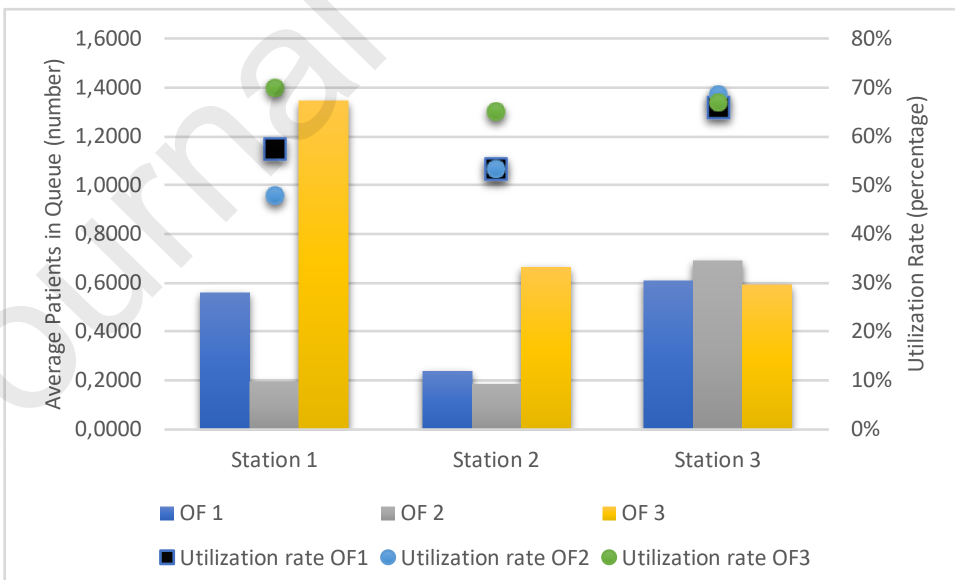


Figure 4. Average number of patients in queue and utilization rates.

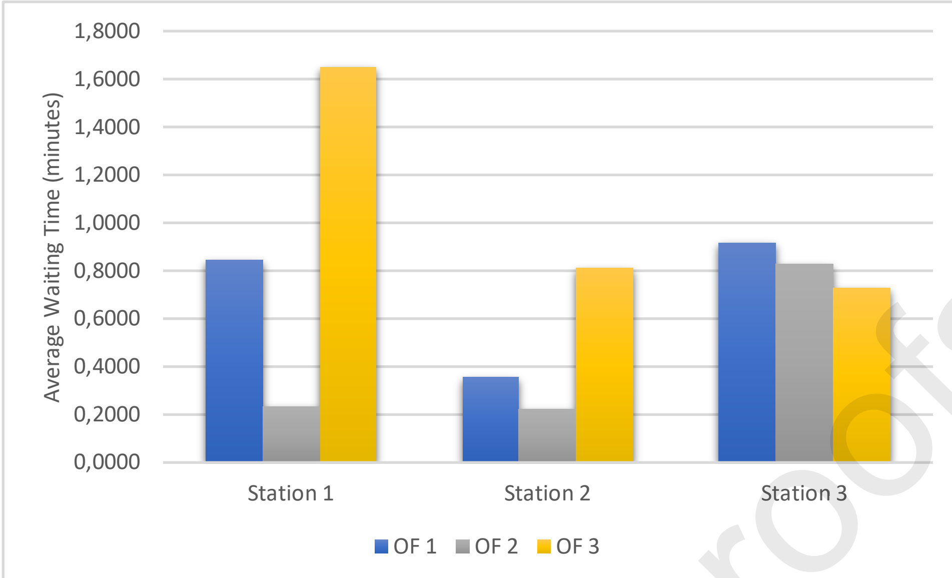


Figure 5. Average patient waiting time in queue (min).

As Figure 1 shows, the number of servers for Station 1 is the same among objective functions (OF) 1 and 3 (2 servers), while OF2 leads to 3 servers. OF2 also leads to the best performance metrics. It is natural to conclude that increasing the number of servers, as OF2 does for Station 1, reduces the average utilization rate, the average number of patients in queue and their average waiting times, as shown by the results of Station 1 in Figures 2 and 3. The model solved with OF3 leads to the highest values for the performance metrics (i.e., worst results) due to the balance it seeks between the first two objective functions.

For Station 2, even though OF1 and OF3 generate the same number of servers (4 servers), the performance metrics are different among them. Specifically, the performance metrics are better with OF1 since it leads to a lower λ (average number of patients scheduled per hour). An additional server is obtained with OF2 (5 servers), though its results are similar to those with OF1 since OF2 leads to a greater λ. Nonetheless, a closer look shows that the waiting times with OF1 are somewhat higher than those with OF2, since OF1 aims at minimizing the number of servers. Finally, as with Station 1, OF3 leads to the worst performance metrics.

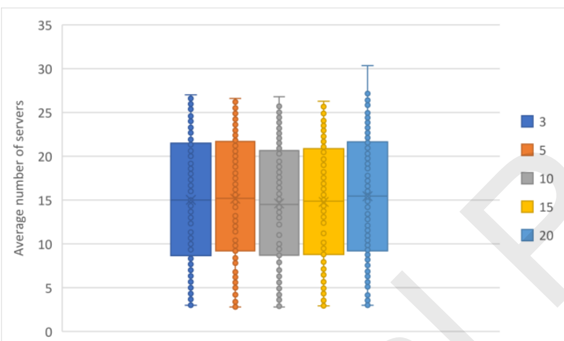
Finally, in Station 3, OF2 and OF3 lead to 6 servers, while OF1 leads to 5 servers. This is explained by the minimization of the number of servers, which increases the average patient waiting time but satisfying the waiting time constraint from Equation10. Unlike Stations 1 and 2, the results of the performance metrics are similar over the three objective functions. The computational time required for solving the three objective functions is 0.35s, 0.16s and 0.89s, respectively, all lower than a second, which can be a reasonable computational time for a planning model.

As final remarks, in this section we have analyzed the results obtained by the mathematical model contrasting the three different objective functions proposed for the case study. The results allow us to describe the benefits of different configurations. If the decision maker is focused on minimizing the number of servers, OF1 can obtain the best results over this performance metric, but if the decision maker is focused on minimizing the utilization rate, OF2 is more suitable. Conversely, if the purpose is to use the resources at maximum extent, OF3 should be used. Finally, if the main focus of the operation is to reduce the average number of patients in queue and their average waiting times, OF2 should be selected.

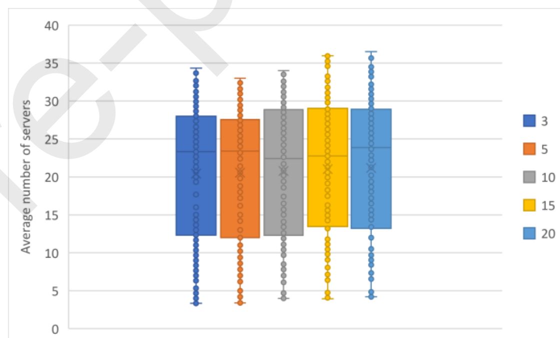
5.2. Performance of the Jackson Network model

To test the mathematical model proposed and the different objective functions analyzed, we performed an experimentation assessing different scenarios with the parameters previously described (α, b_o and n) and considering different number of stations: 3 (case study), 5, 10, 15 and 20, which allows us to run 1023 instances for each size of the number of stations.

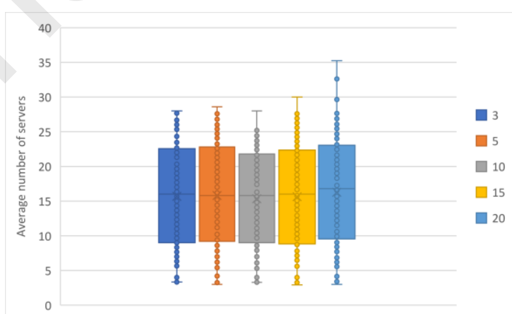
The first analysis performed corresponds to the number of servers for each objective function and for each number of stations. These results are shown in boxplots in Figure 6. Within each objective function, the results of the average number of servers are similar over the number of stations. It can also be observed that the boxplots of OF1 (a) are shorter than those of OF2 (b) and OF3 (c), which is expected since OF1 considers the minimization of the number of servers. Also, a larger number of servers is obtained with OF2 (b) because of the direct relationship between the maximization of the number of patients scheduled and the number of stations that is captured in the utilization constraint from Equation (6).



(a) OF1



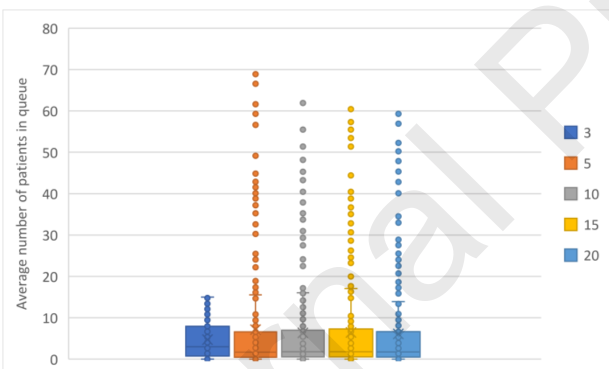
(b) OF2



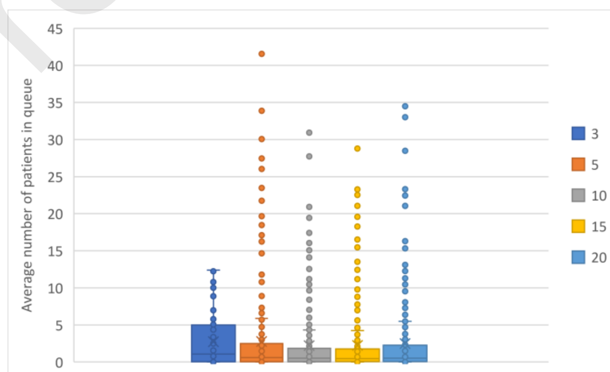
(c) OF3

Figure 6. Average number of servers for each objective function over the number of stations: (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

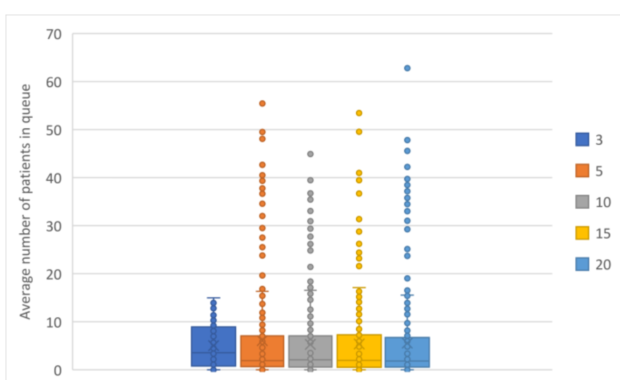
We also analyzed the performance of the system for each objective function and for each number of stations in terms of the average number of patients in queue in the network and their average waiting times. The boxplots of the average number of patients in queue in the network are shown in Figure 7. It can be concluded that for the three objective functions, the highest value of patients in queue is obtained by the first configuration analyzed (3 stations). Nevertheless, the outliers of the rest of the models are higher than the ones obtained by the first configuration. The results obtained with OF2 (b) show the lowest values of patients in queue over the three objective functions. This is expected considering the results obtained for the number of servers analyzed in Figure 7(b), which shows the highest values of number of servers over the three objective functions.



(a) OF1



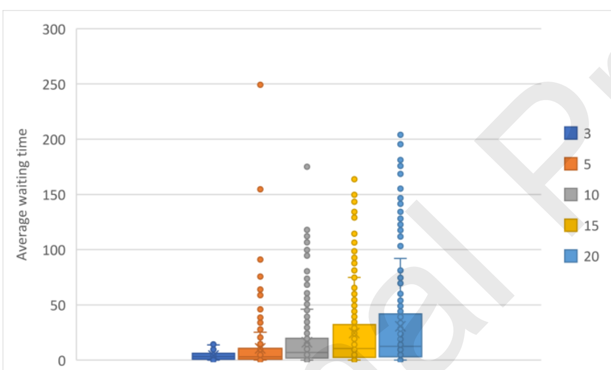
(b) OF2



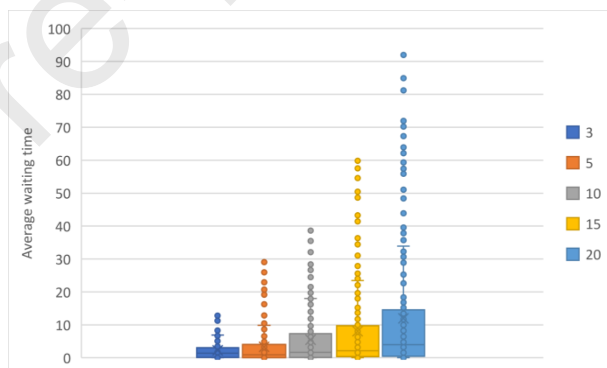
(c) OF3

Figure 7. Average number of patients in queue for each objective function over the number of stations: (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

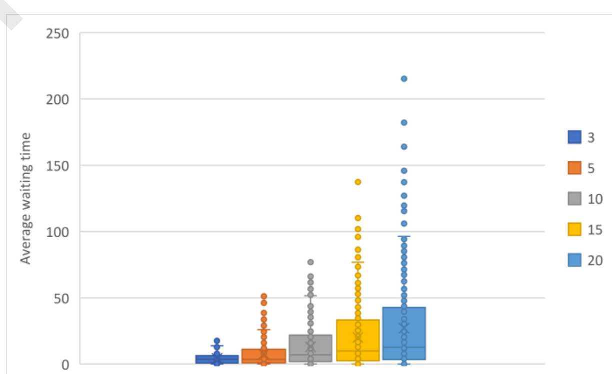
The boxplots of the average patient waiting times in queue in the network are presented in Figure 8. In contrast to the results obtained for the number of patients in queue (see Figure 4), the average waiting times increase with the number of stations. It can also be observed that OF2 (b) leads to the lowest average waiting times over the three objective functions, which is consistent with the higher number of servers observed with OF2 compared to OF1 and OF3 (see Figure 4).



(a) OF1



(b) OF2



(c) OF3

Figure 8. Average patient waiting time (in minutes) in queue for each objective function over the number of stations: (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

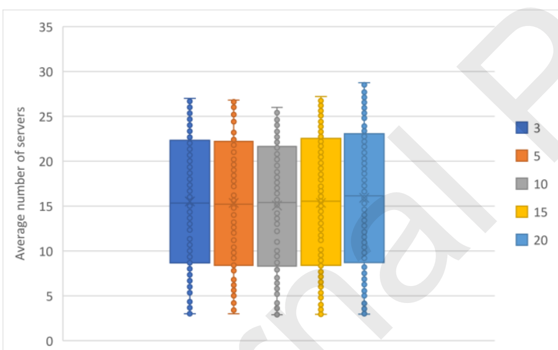
If we compare the results presented in Figures 4, 5 and 6, it can be concluded that OF1 can be used to obtain lower values for the number of servers. In contrast, OF2 leads to higher number of servers. If the objective of the decision maker is minimizing the average number of patients in queue and their average waiting times, she should select OF2.

5.3. Performance of the Generalized Network model

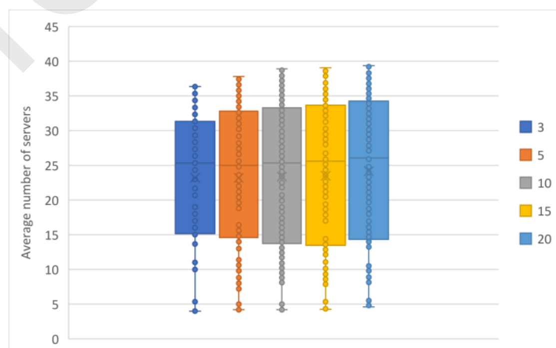
Similar to the previous analyses performed, in this section we describe the results obtained with the Generalized Network model in the same way that we described the results of the Jackson Network model. The following are the assumptions used for the mathematical model:

- Arrival rate λ_1 follows an General distribution and it is considered as a variable of the model
- Servers has a process time s_j which follows a General distribution
- Rework is not presented in the vaccination process neither in registration
- Capacity of queue and system is infinite
- The discipline of queue is FIFO
- The preparation of vaccines is not considered as a station given that we are studying the process where patients are involved

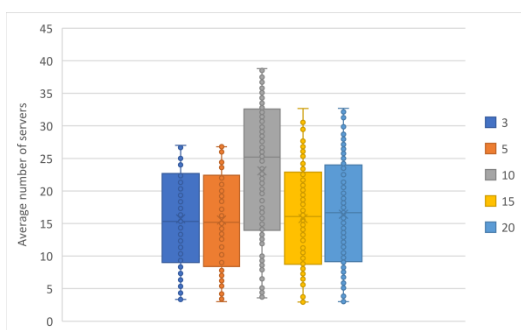
The first part of the results corresponds to the average number of servers for each objective function and for each number of stations. These results are shown in boxplots in Figure 9. In general, the results obtained with OF1 (a) and OF3 (c) show a lower number of servers compared to OF2 (b), and the ranges with OF1 are the smallest. Within OF1 (a) and OF2 (b), the average number of servers is similar over the number of stations, which is a similar behavior to that observed in Figure 6 (a) and (b). For OF3 (c), the number of servers is quite similar over the number of stations, except for the 10-station instances whose values are higher. If we contrast these results with the results from Figure 4 (RF Model 1), the ranges with Model 2's OF2 increase with the number of stations. In addition, the results with Model 2's OF3 show some heterogeneity given the peak in the number of servers for the 10-station instances.



(a) OF1



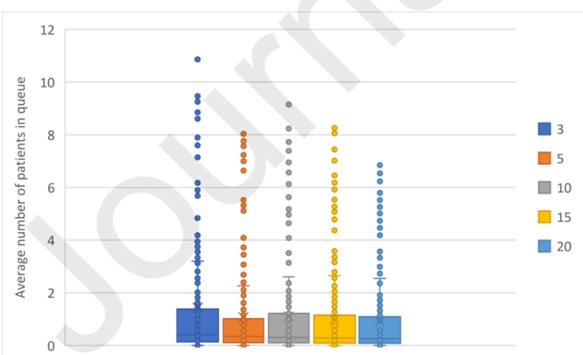
(b) OF2



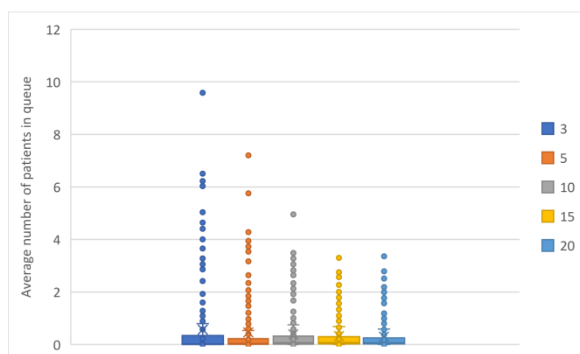
(c) OF3

Figure 9. Average number of servers for each objective function over the number of stations (Model 2): (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

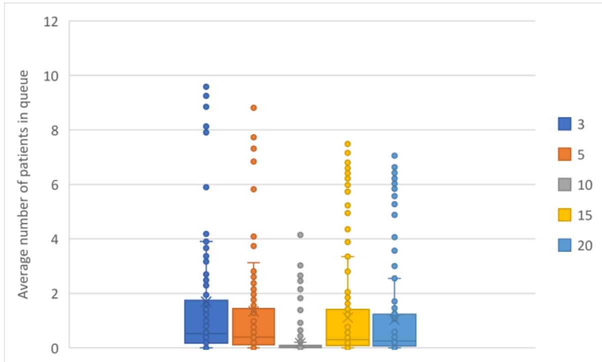
On the other hand, the average number of patients in queue in the network and their average waiting times are presented in boxplots in Figures 10 and 11, respectively. From Figure 10, it can be concluded that a high number of patients in queue is obtained with both OF1 (a) and OF3 (c), while the lowest with OF2 (b). This is due to the main objectives pursued (for example, OF1 minimizes the number of servers, while OF2 maximizes the scheduled patients, which generates higher number of servers and, thus, less patients in queue). For OF3 (c), the number of patients in queue over the number of servers is quite similar, except for the 10-station instances, which show lower values. These OF3 results are consistent with the previous results on the number of servers (see Figure 7 (c)). In contrast to the results presented in Figure 7, the average number of patients in queue is lower for the three objective functions analyzed in Model 2.



(a) OF1



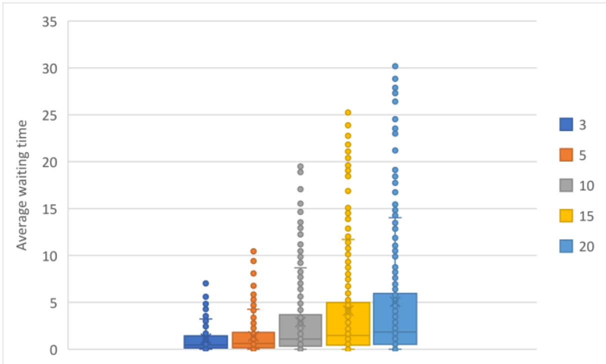
(b) OF2



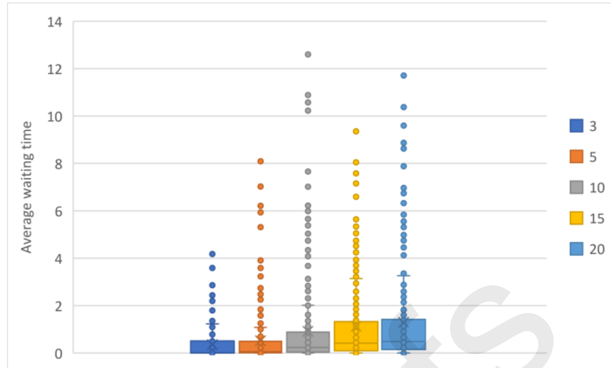
(c) OF3

Figure 10. Average number of patients in queue for each objective function over the number of stations (Model 2): (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

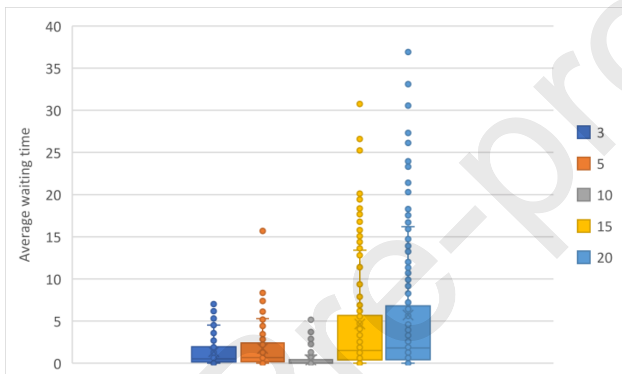
From Figure 11, a general conclusion that can be drawn over the three objective functions is that the patient waiting time increases with the number of stations (except for the 10-station instances with OF3 (c)). In addition, the lowest waiting times are obtained with OF2 (b), while the highest with OF3 (c) due to the peaks reached. Overall, the increased waiting times observed in this figure are consistent with the results presented in Figure 8, where waiting times also increase with the number of stations. However, the ranges obtained with Model 2 are smaller than those obtained with RF Model 1.



(a) OF1



(b) OF2



(c) OF3

Figure 11. Average patient waiting time (in minutes) in queue for each objective function over the number of stations (Model 2): (a) Objective Function 1, (b) Objective Function 2 and (c) Objective Function 3.

If we analyze the results obtained with Model 2, we can conclude that, on average, OF1 is appropriate for obtaining low values for the number of servers. In contrast, OF2 is more appropriate for obtaining better results for the average number of patients in queue and their average waiting times. Overall, we can observe some differences between the results of Model 2 and RF Model 1. This is because of two major reasons: (i) Model 2 does not use the assumption of the exponential times and (ii) Model 2 also considers the coefficients of variation among the stations.

5.4. Theoretical and managerial insights

From a theoretical perspective, we highlight two main aspects. First, our approach combines the modeling of both patient scheduling and capacity planning, giving the decision maker the option to optimize the system for different occupation levels, e.g., staff totally or partially assigned to the vaccination operation, as well as for different maximum allowed number of patients in queue, which helps manage space availability and crowds in vaccination centers. And second, unlike other approaches presented for each of these problems, our approach also considers the stochastic behavior of arrival rates and service times.

Our numerical experiments allow us to highlight two major managerial insights. First, the analysis shows that over each objective function analyzed, the average number of servers required for each station remains stable independent of the number of stations. By implementing our model as a support tool, decision makers can thus efficiently define the capacity of a system in terms of staff required (and other associated resources) regardless of how many steps the process has, e.g., registry, paperwork validation, instructions, vaccination, medical assessment. They can also use this tool to plan patient scheduling, thus combining tactical and operational decisions while optimizing either the rate of patients served, the number of medical staff used or a combination of these depending on the chosen objective function. Second, the inclusion of the parameter λ (average number of patients scheduled per hour) as a variable in the mathematical models allows to reduce the average number of patients in queue and their average waiting times, which are key to avoid overcrowding and reduce the risk of infections. All in all, the capacity planning of stations and patient scheduling can be translated into a staffing decision that can support a cost-effective planning and managing of vaccination (and similar) processes by healthcare providers.

6. Conclusions

Our work studied the joint problem of (i) capacity planning (with multiple and sequential stations and multiple servers within each station) and (ii) patient scheduling, where the service times and patient arrivals were assumed to be stochastic. The problem was modeled through two different approximations: an open Jackson Network model and a Generalized Network model, and for each of them we tested three different objective functions that allowed us to analyze different system configurations. The first model has the hard assumption of exponential times (Poisson process), while the second one relaxes it. Both models were solved using a commercial solver. The experimentation performed allowed us to analyze a COVID-19 vaccination case study, whose results can be used by the Foundation and other healthcare providers as a tool to plan the configuration of their vaccination processes including the patient scheduling according to their specific needs and goals. We also tested the proposed models over different instances analyzing different parameter configurations and increasing the number of stations. This allowed us to optimize the system in each case for different occupation levels as well as for different maximum allowed number of patients in queue, which helps manage space availability and crowds in vaccination centers.

This work can be extended in several ways. For example, the mathematical models could include several other aspects such as the probability of no-shows, reneging and/or deferrals, which can increase the models' flexibility according to different patient behaviors. In addition, another objective function that considers the costs of servers (or workers modeled as servers) could be included, helping contrast the balance between service level and financial costs.

Acknowledgements

We thank Fair Isaac Corporation (FICO) for providing us with Xpress-MP licenses under the Academic Partner Program subscribed with Universidad del Rosario. This study was partially supported by Universidad del Rosario through its COVID-19 projects fund. We also thank Prof. Catalina Latorre and Prof. Carlos E. Trillos from the Universidad del Rosario's School of Medicine and Health Sciences for helping us understand the COVID-19 vaccination process from the medical perspective.

7. References

Alizadeh, R., Rezaeian, J., Abedi, M., & Chiong, R. (2020). A modified genetic algorithm for non-emergency outpatient appointment scheduling with highly demanded medical services considering patient priorities. *Computers & Industrial Engineering*, *139*, 106106. https://doi.org/https://doi.org/10.1016/j.cie.2019.106106

Anparasan, A., & Lejeune, M. (2019). Resource deployment and donation allocation for epidemic outbreaks. *Annals of Operations Research*, *283*(1), 9–32. https://doi.org/10.1007/s10479-016-2392-0

Asgary, A., Valtchev, S. Z., Chen, M., Najafabadi, M. M., & Wu, J. (2021). Artificial Intelligence Model of Drive-Through Vaccination Simulation. In *International Journal of Environmental Research and Public Health* (Vol. 18, Issue 1). https://doi.org/10.3390/ijerph18010268

Dogru, A. K., & Melouk, S. H. (2019). Adaptive appointment scheduling for patient-centered medical homes. *Omega*, *85*, 166–181. https://doi.org/https://doi.org/10.1016/j.omega.2018.06.009

Eshun-Wilson, I., Mody, A., Tram, K. H., Bradley, C., Sheve, A., Fox, B., Thompson, V., & Geng, E. H. (2021). Strategies that make vaccination easy and promote autonomy could increase COVID-19 vaccination in those who remain hesitant. *MedRxiv*, 2021.05.19.21257355. https://doi.org/10.1101/2021.05.19.21257355

Gianfredi, V., Pennisi, F., Lume, A., Ricciardi, G. E., Minerva, M., Riccò, M., Odone, A., & Signorelli, C. (2021). Challenges and Opportunities of Mass Vaccination Centers in COVID-19 Times: A Rapid Review of Literature. *Vaccines*, *9*(6). https://doi.org/10.3390/vaccines9060574

Goralnick, E., Kaufmann, C., & Gawande, A. A. (2021). Mass-Vaccination Sites — An Essential Innovation to Curb the Covid-19 Pandemic. *New England Journal of Medicine*, *384*(18), e67. https://doi.org/10.1056/NEJMp2102535

Hopp, W. J., & Spearman, M. L. (2011). *Factory physics*. Waveland Press.

Kuiper, A., de Mast, J., & Mandjes, M. (2021). The problem of appointment scheduling in outpatient clinics: A multiple case study of clinical practice. *Omega*, *98*, 102122. https://doi.org/https://doi.org/10.1016/j.omega.2019.102122

Lemaitre, J. C., Pasetto, D., Zanon, M., Bertuzzo, E., Mari, L., Miccoli, S., Casagrandi, R., Gatto, M., & Rinaldo, A. (2021). Optimizing the spatio-temporal allocation of COVID-19 vaccines: Italy as a case study. *MedRxiv*, 2021.05.06.21256732. https://doi.org/10.1101/2021.05.06.21256732

Lemaitre, J. C., Pasetto, D., Zanon, M., Bertuzzo, E., Mari, L., Miccoli, S., Casagrandi, R., Gatto, M., & Rinaldo, A. (2022). Optimal control of the spatial allocation of COVID-19 vaccines: Italy as a case study. *PLOS Computational Biology*, *18*(7), e1010237. https://doi.org/10.1371/JOURNAL.PCBI.1010237

Mills, M. C., & Salisbury, D. (2021). The challenges of distributing COVID-19 vaccinations. *EClinicalMedicine*, *31*. https://doi.org/10.1016/j.eclinm.2020.100674

Moreno, M. S., & Blanco, A. M. (2018). A fuzzy programming approach for the multi-objective patient appointment scheduling problem under uncertainty in a large hospital. *Computers & Industrial Engineering*, *123*, 33–41. https://doi.org/https://doi.org/10.1016/j.cie.2018.06.013

Nguyen, T. B. T., Sivakumar, A. I., & Graves, S. C. (2018). Capacity planning with demand uncertainty for outpatient clinics. *European Journal of Operational Research*, *267*(1), 338–348. https://doi.org/https://doi.org/10.1016/j.ejor.2017.11.038

Palmer, R., Fulop, N. J., & Utley, M. (2017). A systematic literature review of operational research methods for modelling patient flow and outcomes within community healthcare and other settings. *Health Systems*. https://doi.org/10.1057/s41306-017-0024-9

Pryor, G. E., Marble, K., Velasco, F. T., Lehmann, C. U., & Basit, M. A. (2021). COVID-19 Mass Vaccination Resource Calculator. *Appl Clin Inform*, *12*(04), 774–777.

Roy, S., Dutta, R., & Ghosh, P. (2021). Optimal Time-Varying Vaccine Allocation Amid Pandemics With Uncertain Immunity Ratios. *IEEE Access*, *9*, 15110–15121. https://doi.org/10.1109/ACCESS.2021.3053268

Sang, P., Begen, M. A., & Cao, J. (2021). Appointment scheduling with a quantile objective. *Computers & Operations Research*, *132*, 105295. https://doi.org/https://doi.org/10.1016/j.cor.2021.105295

Sheikh, A. B., Pal, S., Javed, N., & Shekhar, R. (2021). COVID-19 Vaccination in Developing Nations: Challenges and Opportunities for Innovation. *Infectious Disease Reports 2021, Vol. 13, Pages 429-436*, *13*(2), 429–436. https://doi.org/10.3390/IDR13020041

Signorelli, C., Odone, A., Gianfredi, V., Capraro, M., Kacerik, E., Chiecca, G., Scardoni, A., Minerva, M., Mantecca, R., Musarò, P., Brazzoli, P., Basteri, P., Bertini, B., Esposti, F., Ferri, C., Alberti, V. A., & Gastaldi, G. (2021). Application of the “immunization islands” model to improve quality, efficiency and safety of a COVID-19 mass vaccination site. *Annali Di Igiene : Medicina Preventiva e Di Comunita*, *33*(5), 499–512. https://doi.org/10.7416/ai.2021.2456

Wang, B., Nolan, R., & Marshall, H. (2021). COVID-19 Immunisation, Willingness to Be Vaccinated and Vaccination Strategies to Improve Vaccine Uptake in Australia. *Vaccines 2021, Vol. 9, Page 1467*, *9*(12), 1467. https://doi.org/10.3390/VACCINES9121467

Weintraub, R. L., Subramanian, L., Karlage, A., Ahmad, I., & Rosenberg, J. (2020). COVID-19 Vaccine To Vaccination: Why Leaders Must Invest In Delivery Strategies Now. *Health Affairs*, *40*(1), 33–41. https://doi.org/10.1377/hlthaff.2020.01523

Wood, R. M., Murch, B. J., Moss, S. J., Tyler, J. M. B., Thompson, A. L., & Vasilakis, C. (2021). Operational research for the safe and effective design of COVID-19 mass vaccination centres. *Vaccine*, *39*(27), 3537–3540. https://doi.org/https://doi.org/10.1016/j.vaccine.2021.05.024

Wrigley-Field, E., Kiang, M. v., Riley, A. R., Barbieri, M., Chen, Y. H., Duchowny, K. A., Matthay, E. C., van Riper, D., Jegathesan, K., Bibbins-Domingo, K., & Leider, J. P. (2021). Geographically targeted COVID-19 vaccination is more equitable and averts more deaths than age-based thresholds alone. *Science Advances*, *7*(40). https://doi.org/10.1126/SCIADV.ABJ2099/SUPPL_FILE/SCIADV.ABJ2099_SM.PDF

Yan, C., Huang, G. G. Q., Kuo, Y.-H., & Tang, J. (2021). Dynamic appointment scheduling for outpatient clinics with multiple physicians and patient choice. *Journal of Management Science and Engineering*. https://doi.org/https://doi.org/10.1016/j.jmse.2021.02.002

Yogev, D., Burlak, K., Tabak, N., Fink, N., Karp, E., & Segal, D. (2021). An Insight on the Arrangement of a Remote Military COVID-19 Vaccination Site. *Military Medicine*, *186*(7–8), 201–202. https://doi.org/10.1093/milmed/usab089

Zhou, S., Zhou, S., Zheng, Z., & Lu, J. (2021). Optimizing Spatial Allocation of COVID-19 Vaccine by Agent-Based Spatiotemporal Simulations. *GeoHealth*, *5*(6), e2021GH000427. https://doi.org/https://doi.org/10.1029/2021GH000427

Appendix A

Notation:

Sets

O: $set\ of\ stations$

Parameters

n: $number\ of\ patients\ to\ be\ scheduled\ per\ day$

H: $hours\ per\ day$

α: $mininum\ percentage\ of\ number\ of\ patients\ served\ per\ day$

b_o:$maximum\ ocupation\ for\ each\ station\ o \in O$

c_o: $minimum\ ocupation\ for\ each\ station\ o \in O$

d_o: $maximum\ average\ waiting\ time\ in\ each\ station\ o \in O$

e_o: $maximum\ average\ number\ of\ patients\ in\ queue\ in\ each\ station\ o \in O$

Random parameters

ti_o: $service\ time\ in\ station\ o \in O$

μ_o: $service\ rate\ \left(\frac{1}{ti_o}\right)\ in\ station\ o \in O$

Variables

λ:$average\ number\ of\ patients\ scheduled\ per\ hour$

s_o: $number\ of\ servers\ assigned\ to\ each\ station\ o \in O$

$\pi 0_o$: $probability\ that\ a\ station\ o \in O\ is\ empty$

Lq_o: $average\ number\ of\ patients\ in\ queue\ in\ station\ o \in O$

The mathematical model is as follows:

$$\min \sum_{o \in O} s_o \tag{1}$$

$$\max \lambda \tag{2}$$

$$\min \frac{\sum_{o \in O} s_o}{best\ server} - \frac{\lambda}{best\ \lambda} \tag{3}$$

Subject to:

$$\lambda \leq \frac{n}{H} \tag{4}$$

$$\lambda \geq \frac{n}{H}\alpha \tag{5}$$

$$\lambda \leq s_o \mu_o b_o \quad \forall\, o \in O \tag{6}$$

$$\lambda \geq s_o \mu_o b_o \quad \forall\, o \in O \tag{7}$$

$$\pi 0_o = \left(\frac{1}{1 + \sum_{k=1}^{s_o - 1} \frac{\left(\frac{\lambda}{\mu_o}\right)^k}{k!} + \frac{\lambda^{s_o} j \mu_o}{\mu_o^{s_o} s_o! (\mu_o s_o - \lambda)}} \right) \quad \forall\, o \in O \tag{8}$$

$$Lq_o = \frac{\lambda^{s_o} ti_o^{s_o} \pi 0_o \left(\frac{\lambda}{\mu_o s_o}\right)}{s_o! \left(1 - \frac{\lambda}{\mu_o s_o}\right)^2} \quad \forall\, o \in O \tag{9}$$

$$\frac{Lq_o}{\lambda} \leq d_o \quad \forall\, o \in O \tag{10}$$

$$Lq_o \leq e_o \quad \forall\, o \in O \tag{11}$$

$$s_o \in Z \quad \forall\, o \in O \tag{12}$$

$$\lambda \geq 0 \tag{13}$$

$$\pi 0_o \geq 0 \quad \forall\, o \in O \tag{14}$$

$$Lq_o \geq 0 \quad \forall\, o \in O \tag{15}$$

Appendix B

- To limit the number of servers for each station, we use lb_o and ub_o as lower and upper bounds for each station, respectively.
- We use a new constraint that determines the bounds for each station:

$$lb_o \leq s_o \leq ub_o \tag{16}$$

- We introduce a new binary variable $z_{k,o}$ that indicates the number of servers assigned to each station, thus the number of servers for each station is defined as:

$$s_o = \sum_{k = lb_o}^{up_o} k z_{k,o} \tag{17}$$

The binary variable is limited to:

$$z_{j,o} \in \{0, 1\} \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\} \tag{18}$$

Thus, the modified mathematical model is as follows:

$$\min \sum_{o \,\in\, O} \sum_{j = lb_o}^{ub_o} j * z_{j,o} \tag{19}$$

$$\max \lambda \tag{20}$$

$$\min \frac{\sum_{o \,\in\, O} \sum_{j = lb_o}^{ub_o} j * z_{j,o}}{best\ Server} - \frac{\lambda}{best\ \lambda} \tag{21}$$

Subject to:

$$\lambda \leq \frac{n}{H} \tag{22}$$

$$\lambda \geq \frac{n}{H}\alpha \tag{23}$$

$$\lambda \leq \sum_{j = lb_o}^{ub_o} j * z_{j,o} \mu_o b_o \quad \forall\, o \in O \tag{24}$$

$$\lambda \geq \sum_{j = lb_o}^{ub_o} j * z_{j,o} \mu_o c_o \quad \forall\, o \in O \tag{25}$$

$$\sum_{j = lb_o}^{ub_o} z_{j,o} = 1 \quad \forall\, o \in O \tag{26}$$

	$\pi 0_{o,j} = \left(\frac{1}{1 + \sum_{k=1}^{j-1} \frac{\left(\frac{\lambda}{\mu_o}\right)^k}{k!} + \frac{\lambda^j j \mu_o}{\mu_o^j j! (\mu_o j - \lambda)}} \right) z_{j,o} \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\}$	(27)

	$\pi 0 f_o = \sum_{j=lb_o}^{ub_o} \pi 0_{o,j} \quad \forall\, o \in O$	(28)

	$Lqi_{o,j} = \frac{\lambda^j ti_o^j \pi 0_{o,j} \left(\frac{\lambda}{\mu_o j}\right)}{j! \left(1 - \frac{\lambda}{\mu_o j}\right)^2} \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\}$	(29)

	$Lqf_o = \sum_{j=lb_o}^{ub_o} Lqi_{o,j} \quad \forall\, o \in O$	(30)

	$\frac{Lqf_o}{\lambda} \leq d_o \quad \forall\, o \in O$	(31)

	$Lqf_o \leq e_o \quad \forall\, o \in O$	(32)

	$z_{j,o} \in \{0, 1\} \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\}$	(33)

	$\lambda \geq 0$	(34)

	$\pi 0_{o,j} \geq 0 \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\}$	(35)

	$\pi 0 f_o \geq 0 \quad \forall\, o \in O$	(36)

	$Lqi_{o,j} \geq 0 \quad \forall\, o \in O, j = \{lb_o \ldots ub_o\}$	(37)

	$Lqf_o \geq 0 \quad \forall\, o \in O$	(38)

We refer to these models together as Model 1 and will distinguish between the original model (OR Model 1) and the reformulated model (RF Model 1) as needed. Table B1. summarizes how the original and reformulated models relate to each other.

Table B1. Relationship between OR Model 1 and RF Model 1.

	OR Model 1	**RF Model 1**
Objective functions		
Server assignment (SA)	Equation 1	Equation 19
Patient scheduling (PS)	Equation 2	Equation 20
SA + PS	Equation 3	Equation 21
Constraints		
Working hours	Equations 4-5	Equations 22-23
Utilization	Equations 6-7	Equations 24-26
Idle time probability[a]	Equation 8	Equations 27-28
Patients in queue[a]	Equation 9	Equations 29-30
Maximum patients in queue	Equation 10	Equation 31
Maximum waiting time	Equation 11	Equation 32
Types of variables		
Servers	Equation 12	Equation 33
Scheduled patients	Equation 13	Equation 34
Idle time probability	Equation 14	Equations 35-36
Patients in queue	Equation 15	Equations 37-38

[a] Jackson Network metrics

Appendix C

Notation:

Parameters

Ca^2i: $interarrival\ time\ squared\ coefficient\ of\ variation$

Var_o: $variance\ of\ service\ time\ in\ station\ o \in O$

Variables

$Ca^2{}_{o,j}$: $interarrival\ time\ squared\ coefficient\ of\ variation\ of\ station\ o \in O,\ if\ in\ station\ o-1\ j\ servers\ are\ used$

Ca^2f_o: $interarrival\ time\ squared\ coefficient\ of\ variation\ for\ station\ o \in O$

$CTqI_{o,j}$: $average\ waiting\ time\ in\ station\ o \in O\ if\ j\ servers\ are\ assigned$

$CTqF_o$: $average\ waiting\ time\ in\ station\ o \in O$

The mathematical model is defined as:

$$\min \sum_{o\,\in\,O} \sum_{j\,=\,lb_o}^{ub_o} j * z_{j,o} \tag{39}$$

$$\max \lambda \tag{40}$$

$$\min \frac{\sum_{o\,\in\,O} \sum_{j\,=\,lb_o}^{ub_o} j * z_{j,o}}{best\ Server} - \frac{\lambda}{best\ \lambda} \tag{41}$$

Subject to (new equations are presented in **bold**):

$$\lambda \leq \frac{n}{H} \tag{42}$$

$$\lambda \geq \frac{n}{H}\alpha \tag{43}$$

$$\lambda \leq \sum_{j\,=\,lb_o}^{ub_o} j * z_{j,o}\mu_o b_o \quad \forall\, o \in O \tag{44}$$

$$\lambda \geq \sum_{j\,=\,lb_o}^{ub_o} j * z_{j,o}\mu_o c_o \quad \forall\, o \in O \tag{45}$$

$$\sum_{j\,=\,lb_o}^{ub_o} z_{j,o} = 1 \quad \forall\, o \in O \tag{46}$$

$$\boldsymbol{Ca^2{}_{1,j} = Ca^2 i * z_{j,1} \quad \forall\, j = \{lb_1 \ldots ub_1\}} \tag{47}$$

$$\boldsymbol{\lambda Ca^2{}_{o,j} - z_{j,o-1}\lambda\left(1 - \left(\frac{\lambda}{j * \mu_{o-1}}\right)^2\right)\sum_{k\,=\,lb_{o-1}}^{ub_{o-1}} Ca^2{}_{o-1,k} = \lambda\left(\frac{\lambda}{j * \mu_{o-1}}\right)^2\left(\frac{Var_{o-1}}{ti_{o-1}^2}\right) z_{j,o-1} \quad \forall\, o = \{lb_1 \ldots ub_1\} \mid o \geq 2} \tag{48}$$

	$$\boldsymbol{Ca^2f_o = \sum_{j=lb_o}^{ub_o} Ca^2{}_{o,j} \quad \forall\, o \in O}$$	**(49)**
	$$\boldsymbol{CTqI_{o,j} = \left(\frac{Ca^2f_o + \left(\frac{Var_o}{ti_o^2}\right)}{2}\right)\left(\frac{\left(\frac{\lambda}{j\mu_o}\right)^{\sqrt{2j+2}-1}}{j\left(1-\left(\frac{\lambda}{j\mu_o}\right)\right)}\right) ti_o z_{j,o}}$$	**(50)**
	$$\boldsymbol{CTqF_o = \sum_{j=lb_o}^{ub_o} WqI_{o,j} \quad \forall\, o \in O}$$	(51)
	$CTqF_o \leq d_o \quad \forall\, o \in O$	(52)
	$\lambda CTqF_o \leq e_o \quad \forall\, o \in O$	(53)
	$z_{j,o} \in \{0,1\} \quad \forall\, o \in O, j = \{lb_o ... ub_o\}$	(54)
	$\lambda \geq 0$	(55)
	$Ca^2{}_{o,j} \geq 0 \quad \forall\, o \in O, j = \{lb_o ... ub_o\}$	(56)
	$Ca2f_o \geq 0 \quad \forall\, o \in O$	(57)
	$CTqi_{o,j} \geq 0 \quad \forall\, o \in O, j = \{lb_o ... ub_o\}$	(58)
	$CTqf_o \geq 0 \quad \forall\, o \in O$	(59)

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: