Proceedings of the 16th Conference on Computer Science and Intelligence Systems pp. 393–402
DOI: 10.15439/2021F78
ISSN 2300-5963 ACSIS, Vol. 25

Mass Vaccine Administration under Supply Uncertainty

Salvatore Foderaro
Università di Tor Vergata,
via del Politecnico 1, Roma, Italy
Email: salvatorefoderaro@gmail.com

Maurizio Naldi
Università LUMSA
via Marcantonio Colonna 19, Roma, Italy
Email: m.naldi@lumsa.it

Gaia Nicosia
Università Roma Tre,
via della Vasca Navale 79, Roma, Italy
Email: gaia.nicosia@uniroma3.it

Andrea Pacifici
Università di Tor Vergata,
via del Politecnico, 1, Roma, Italy
Email: andrea.pacifici@uniroma2.it

***Abstract*—The insurgence of COVID-19 requires fast mass vaccination, hampered by scarce availability and uncertain supply of vaccine doses and a tight schedule for boosters. In this paper, we analyze planning strategies for the vaccination campaign to vaccinate as many people as possible while meeting the booster schedule. We compare a conservative strategy and q-days-ahead strategies against the clairvoyant strategy. The conservative strategy achieves the best trade-off between utilization and compliance with the booster schedule. Q-days-ahead strategies with $q < 7$ provide a larger utilization but run out of stock in over 30% of days.**

I. Introduction

DUE to the global COVID-19 pandemic emergency, mass vaccinations are taking place all over the world. Mass vaccinations have been held in the past, starting with the vaccination days and the mass campaign to eradicate smallpox in the early 19th century [1]. However, two centuries after, mass vaccination is still a challenge [2]. The challenge is particularly severe for situations where the need for mass vaccination arises while vaccines are being developed. In that case, time constraints conflict: herd immunity calls for fast action, but the need for wide vaccine availability slows down the campaign deployment. As a consequence, a great variety of logistic problems connected to this enormous task arise. In this paper, we focus on planning a vaccination campaign, i.e, determining a day-by-day prescription on the number of doses that have to be administered in the presence of uncertainties in the distribution provided by the vaccine suppliers. Here, we are not looking at allocation options concerning the priority order of specific population segments or areas in the territory. Instead, we focus on the *downstream* problem connected to the effective and efficient delivery of the vaccines to eligible individuals. In designing such a decision support tool, several issues must be considered, mainly due to different characteristics of the administered vaccines.

This work is partially supported by MIUR PRIN Project AHeAD (Efficient Algorithms for HArnessing Networked Data).

- Inventory Management issues: vaccine products are temperature-sensitive and must be stored and handled correctly to ensure efficacy and maximize shelf life. Proper storage and handling practices are critical to minimize vaccine loss and limit the risk of administering the COVID-19 vaccine with reduced effectiveness. Expiration dates have also to be taken into account. It is also to be noted that timely management of inventories requires the fast updating and integration of hospital information systems [3] as well as their reliability [4].
- Booster shot: whether or not a second dose is required and the prescribed time interval between the first and the second doses is a component that greatly affects any planning model for mass vaccination.
- Overall vaccination capacity: we must consider the maximum number of vaccinations (independently of whether they are first or second doses) that the system can administer every day. We are considering this information as given and deterministic, to be set as a function of the number of operators and vaccination sites capacity;
- Trade-off between different procurement and vaccination strategies under a limited budget and time horizon constraints, which we do not consider here, but should enter the more general analysis framework [5].

The paper is organized as follows.

Hereafter we briefly report on a few related works in the literature and describe the context and notation for the addressed problem.

In Section IV, we illustrate how the arrival process is modelled and, consequently, how the input data of the experiments are generated.

Optimization models and approaches to the problem are presented in Section V, where two main points of view are considered. We first address the problem as if it were a deterministic (*off-line*) one, i.e., all data are assumed to be known and given in advance. The output of such a phase is a point of reference or benchmark to assess the quality of different non-clairvoyant methods presented in the same

section. We then consider models where supply is regarded as a random process over time, are designed to tackle the real-world stochastic problem effectively.

Section VI reports the results of an extensive computational campaign aimed at testing and assessing the effectiveness of the different proposed approaches.

Finally, in Section VII, some conclusions are drawn.

II. Related literature

Due to the current pandemic situation, mass vaccination logistics problems came (overwhelmingly indeed!) to the attention of researchers only very recently. As a consequence, the literature concerning this specific area is still relatively scarce. There are, however, several papers dealing with various problems arising in the event of a sudden burst of infections caused by a pathogen in a population.

A comprehensive introduction to the mathematical modelling of infectious diseases can be found in the book by Keeling and Rohani [6] as an essential tool in public health planning and response. Several techniques are illustrated to model basic epidemiological processes, such as the propagation of infectious diseases. Such techniques range from differential equations to computer simulations.

The effectiveness of mass vaccination against other policies (such as trace vaccination) is discussed in [7] for the hypothetical case of a smallpox bioterrorist attack in a large U.S. city.

Among the few papers dealing with mass vaccination logistics in the most recent literature (not necessarily related to the COVID-19 pandemic), the following ones present some appreciable connections with the problem at hand.

In [8] the authors formulate a bi-objective model for planning vaccination campaigns that aim at minimising both control costs and the number of infected individuals.

In the context of mass vaccination, some papers in the literature have addressed different types of problems. In a very recent paper, the authors address the problem of allocating vaccines across geographical regions to utilise available vaccines as effectively as possible [9].

Both the above papers base their analysis on the epidemiological conditions of a population and/or a geographical area.

Unlike the above studies, our aim in this work is not to identify who should get vaccinated first. Instead, we want to establish how to optimally exploit the uncertain supplies of doses to speed up the vaccination process and rapidly immunise the largest possible fraction of the population.

In fact, the problem we address here resembles the so-called lot-sizing in production planning, which has been extensively studied since the seminal work of Wagner and Whitin in the Fifties [10]. In [10], the authors propose a forward dynamic programming algorithm for a generalised version of the uncapacitated economic lot-sizing model with dynamic demand under a general concave cost function. The latter model has been extended in [11] by considering the possibility of backlogging. These prototypical models can be viewed as special fixed charge network problems. Several variants have been investigated and still are an important topic of research, including single-item and multi-item, uncapacitated and capacitated lot-sizing problems. However, differently from lot-sizing problems, as we are discussing below, in our model, we are not considering inventory costs as a main component of the decision criteria. Though stock expenses are not negligible, we prioritise the average vaccination time since a fast immunisation of the largest possible audience is of much greater importance in this situation.

III. Context and problem definition

The purpose of our study is to design a support tool providing the decision-maker with a suggestion about the number of doses of vaccines to be administered every day along a given planning horizon. The set of vaccine types is $\mathcal{V}$, and the planning horizon consists of a number T of periods (days):

$$\mathcal{W} = \{1, \ldots, T\}.$$

This information specifies the type of vaccine, a first or a booster dose, and the daily inventory level for each vaccine. These decisions are based on imperfect information concerning the supply of doses during the planning time window: The number b_t^i of doses of vaccine $i \in \mathcal{V}$ delivered at day $t \in \mathcal{W}$ is considered as a random variable and, in Section IV, the arrival process is thoroughly described using suitable probability density functions.

We wish to (i) determine benchmarks against which the algorithms for the non-deterministic case can be confronted, and (ii) to design a tool that can be safely used when the information about the arrivals will be more reliable, as it can be expected in the (hopefully near) future. For those reasons, in Section V-A, we are also considering an *off-line* version of the problem, in which the amount of daily provision b_t^i for each type of vaccine are given as deterministic data.

Another critical input parameter is the capacity limit of the system, i.e., the maximum number of vaccine doses that may be administered at time t. This limit could be independent of the administered vaccine type. However, in our algorithms, we are considering upper bounds k_t^i on the number of each single vaccine type i that can be administered in day t as given input, for all $i \in \mathcal{V},\ t \in \mathcal{W}$. Of course, the size of these variables (so that $\sum_{i \in \mathcal{V}} k_t^i$ is a constant or slightly variable over time) can be suitably tuned, depending on the available supply.

The algorithms we are proposing are basically models that return a prescriptive vaccination plan over the next T days. This is especially true of off-line algorithms. In particular, the output of the algorithm is the number x_t^i, resp. y_t^i, of people receiving the first, resp. the second, dose of vaccine i on day t. As a consequence, an additional output of these procedures is the stock level of each vaccine i at (the end of) day $t \in \mathcal{W}$. Clearly, the algorithms could also be used in a *rolling-horizon'* fashion, i.e., re-optimizing every single or one-in-n day, taking into account current inventory levels and new estimated future dose arrivals.

For the developments to follow, we refer to the total number of supplied doses of vaccine i until day t by:

$$B^i(t) = \sum_{\theta=1}^{t} b_\theta^i$$

Observe that, if Δ^i is the recommended time interval, expressed in number of periods, between the first and the mandatory booster (or second) dose of vaccine $i \in \mathcal{V}$, since $\sum_{t=1}^{T-\Delta^i} x_t^i = \sum_{t=\Delta^i+1}^{T} y_t^i \leq \frac{1}{2}B^i(T)$ and $\sum_{t=1}^{T-\Delta^i} x_t^i \leq B^i(T-\Delta^i)$, then there is a feasible solution with $s_T^i = 0$ if and only if $B^i(T-\Delta^i) \geq \frac{1}{2}B^i(T)$. In fact, any feasible solution has $s_T^i \geq B^i(T) - 2B^i(T-\Delta^i)$. (With no loss of generality, we assume that the right-hand side of the latter inequality is not positive. Otherwise, we may subtract this quantity from the arrivals of the last $T-\Delta^i$ days and then apply the algorithm. So doing we eventually have $s_T^i = B^i(T) - 2B^i(T-\Delta^i)$.)

IV. The Arrival Process

Vaccine administration is fed by the availability of vaccine doses. A smooth and regular procurement and delivery process of vaccine doses (and the vaccine supply chain in general) is essential to the correct planning of the administration phase [12], [13]. However, the delivery of doses has been hampered by repeated delays, well reported in the general press [1]. As a consequence, the delivery of doses to nations, and subsequently to vaccination centers, appears as largely random. In this section, we provide a stochastic model for the arrival of vaccine doses, considering a whole nation as the recipient.

For this purpose, we rely on the datasets provided for Italy under an OpenData agreement at https://github.com/italia/covid19-opendata-vaccini. The datasets are updated daily and include the number of doses received for each supplier (which are AstraZeneca, Moderna, Pfizer-BioNTech, and Johnson&Johnson). For the time being, the observation interval considered for this study is from December 27, 2020, to April 2, 2021. The sample record of arrivals for Pfizer-BioNTech is shown in Fig. 1. We can notice two major features of the time series:

- the arrivals exhibit a growing trend;
- arrivals do not take place each day, and on most days, there are no arrivals.

The first feature is probably a consequence of the current transient nature of the process. Pharmaceutical companies are scaling up their production to meet the growing demand of nations to vaccinate their citizens. As seen at this stage, the resulting stochastic process of arrivals would be non-stationary, calling, e.g., for the use of a non-homogeneous Poisson model [14]. However, we are more interested in the steady-state version of the process since we imagine an ongoing massive vaccination after facing today's initial phase.

[1]See, e.g., just the recent headlines "Covid vaccine: UK supply hit by India delivery delay" at https://www.bbc.com/news/uk-56438629 and "Covid: What is happening with the EU vaccine rollout?" at https://www.bbc.com/news/explainers-52380823.

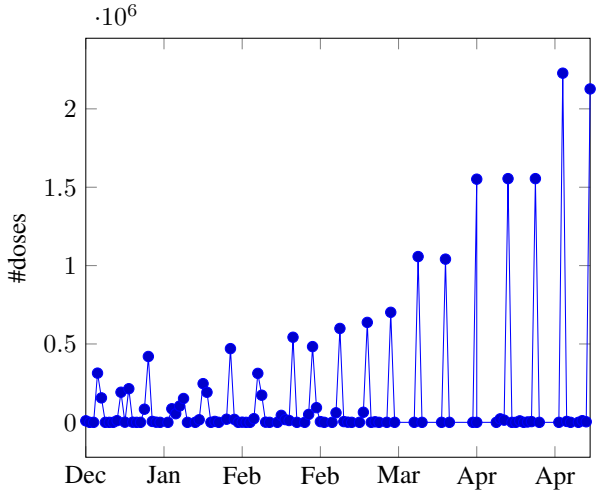

Fig. 1. Daily arrivals of Pfizer-BioNTech doses over Dec 2020 - May 2021

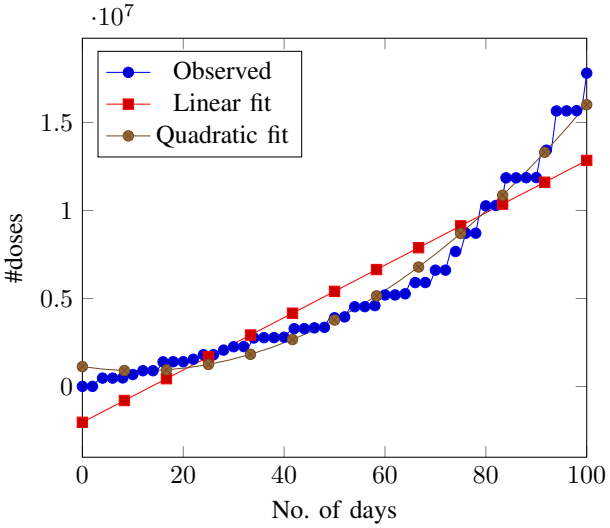

Fig. 2. Cumulative arrivals of Pfizer-BioNTech doses (Dec 2020 - May 2021)

We can, however, examine the growing transient to compare a homogeneous against a non-homogeneous Poisson process (as adopted, e.g. in [15]). In Fig. 2, we have reported the cumulative number of dose arrivals. If that process followed a homogeneous Poisson model, the expected number of arrivals over any period would be proportional to that period (i.e., linear in time). On the same Fig. 2, we have also reported the linear trend that would better fit the observed data. As we can see, the linear trend is a poor approximation of the real growth of arrivals. A quadratic function, also shown in the picture, would be a better fit.

As to the second feature, delivery days are interspersed with relatively long intervals of no-delivery. For that reason, we are led towards a zero-inflated model, where zeros occur more often. In zero-inflated models, the occurrence of zero arrivals is superimposed with a model assuming a larger domain (in our case, the domain $\mathbf{N}$) [16], [17]. The latter

model may be a Poisson, negative binomial, binomial, beta-binomial or hypergeometric. Here we have opted for a zero-inflated Poisson model, also known as ZIP. We recall that we are considering the steady-state scenario with a homogeneous ZIP model. Should we wish to examine the transient phenomenon, we could consider a quadratic approximation for the expected number of arrivals in a non-homogeneous zero-inflated Poisson model, as shown earlier.

In the ZIP model, the probability distribution for the number X of dose arrivals is

$$\mathbf{P}[X=k] = \begin{cases} \pi + (1-\pi)e^{-\lambda} & \text{if } k=0, \\ (1-\pi)e^{-\lambda}\lambda^k/k! & \text{if } k=1,2,\ldots. \end{cases} \quad (1)$$

where $0 \leq \pi \leq 1$ and $\lambda \geq 0$.

This model is then represented by two parameters, π and λ. We can adopt several methods to estimate those parameters [18]. In particular, if we indicate the sampling mean and variance respectively as m and s^2, the method of moments provides us with the following estimates:

$$\begin{aligned} \hat{\pi} &= \frac{s^2 - m}{s^2 + m^2 - m} \\ \hat{\lambda} &= \frac{s^2 + m^2}{m} - 1. \end{aligned} \quad (2)$$

V. Vaccination Planning Approaches

In this section, we describe the algorithms that can be used as a decision support tool for designing a vaccination plan over the next T days. The plan consists of establishing the number of (first and second) doses of vaccines that shall be administered every day.

A. Off-line optimization model

Hereafter, we present a linear programming model providing a solution to maximize the number of vaccinated people per day while exploiting all the doses supplied during the planning time window $\mathcal{W} = \{1, \ldots, T\}$.

The set of available vaccine types is $\mathcal{V} = \mathcal{V}' \cup \mathcal{V}''$. Vaccine type $\mathcal{V}'$ requires a single dose while vaccine type $\mathcal{V}''$ requires a double shot.

In our model, we consider the following decision variables:

- x_t^i and y_t^i, $t \in \mathcal{W}$, $i \in \mathcal{V}$ indicate the number of first and second doses of vaccine i administered during time t;
- s_t^i is the amount of doses of vaccine i remaining in stock at (the end) of period t.

We also assume that, in the initial period $t = 1$, a given inventory $s_0^i \geq 0$ of doses is available in stock and that a given maximum inventory level $\bar{s}_F^i \geq 0$ is required at the end of the planning horizon. We discuss possible feasible values for $\bar{s}_F^i$ later on.

The LP model is

$$\min f(y) \quad (3)$$

$$s.t. \quad x_t^i + y_t^i + s_t^i - s_{t-1}^i = b_t^i \qquad i \in \mathcal{V}, t \in \mathcal{W} \quad (4)$$

$$x_t^i + y_t^i \leq k_t^i \qquad i \in \mathcal{V}, t \in \mathcal{W} \quad (5)$$

$$x_t^i = y_{t+\Delta^i}^i \qquad i \in \mathcal{V}'', t \in \mathcal{W} \quad (6)$$

$$x_t^i = 0 \qquad i \in \mathcal{V}', t \in \mathcal{W} \quad (7)$$

$$s_T^i \leq \bar{s}_F^i \qquad i \in \mathcal{V}, t \in \mathcal{W} \quad (8)$$

$$x_t^i, y_t^i, s_t^i \geq 0 \qquad i \in \mathcal{V}, t \in \mathcal{W} \quad (9)$$

In this model the objective function (3) is the average vaccination time that may be expressed as

$$f(y) = \frac{\sum_{t \in \mathcal{W}} (t \sum_{i \in \mathcal{V}} y_t^i)}{\sum_{t \in \mathcal{W}} \sum_{i \in \mathcal{V}} y_t^i} \quad (10)$$

Note that, each time period t is "weighted" by the number of people $\sum_i y_t^i$ receiving the final dose at t. Equation (10) can be linearized by approximating the denominator with $\frac{1}{2}\sum_{i \in \mathcal{V}} B^i(T)$, i.e., the maximum number of second doses that can be administered in the face of certain dose-supplies b^i: In presence of null final stocks $s_T^i = 0$, the two expressions have equal values.

Equations (4) are simple continuity constraints expressing the obvious relationship among the variables and the supplied number of doses. Constraints (5) bound the total number of doses administered in each period. Constraints (6) ensure that the second dose of the vaccination is given after the recommended time interval, while constraints (7) refer to the single-dose vaccines[2] Constraints (8) impose that the final inventory level is at most a given quantity $\bar{s}_F^i$. Without such constraints, the optimal solution would be not to administer any vaccine (and obtain a null valued objective). The values $\bar{s}_F^i$ are given as an input to the LP model and can be chosen small enough to guarantee that we are using as much as possible of the supplied doses. In any case, it is clear that

$$\bar{s}_F^i \geq \max\{0, B^i(T) - 2B^i(T - \Delta^i)\} \quad (11)$$

should hold.

The above model computes an optimal solution of our planning problem under the assumption that the exact amount of supply b_t^i of each vaccine i is given, for each period t, i.e., the LP solves a deterministic (off-line) version of the actual stochastic problem. Such solutions can be used as a benchmark to assess the quality of blind heuristic algorithms providing prescriptive information on the number of doses that can be administered each day without a perfect knowledge on the doses supplied in the future. As it is expected that the supply process will become steady and the data about arrival dates trustworthy enough, the above linear programming models could be used as a reliable decision support system.

[2]While we show that single-dose vaccines may be easily included in our models, due to scarcity of data about this type of immunization, in the remainder of the paper we only present algorithms and experiments concerning two-doses vaccines.

With regard to the off-line version of the problem, in which the supply b_t^i is deterministic and given for all $i \in \mathcal{V}, t \in \mathcal{W}$, one may ask if a simpler mechanism than the LP-based one described above would determine an optimal (or close to optimal) solution without recurring to a mathematical program. A greedy strategy seems a viable tool due to the simple continuity relations binding the different quantities together (similar, for instance, to those of the classical lot-sizing problem). Therefore, limiting ourselves to the uncapacitated case, we tested a natural heuristic that guarantees the maximum consumption of all the supplied doses and tries to schedule as early as possible the vaccinations while keeping inventory levels nonnegative. Note that the problem is decomposable, and the heuristic provides a separate administration plan for each vaccine $i \in \mathcal{V}$. Such a heuristic algorithm is sketched hereafter.

We first define an initial feasible solution $\bar{x}$ as:

$$\bar{x}_t^i = \begin{cases} \frac{1}{2}B^i(T) & t = T - \Delta^i \\ 0 & t \in \mathcal{W} \setminus \{T - \Delta^i\} \end{cases} \tag{12}$$

This is a feasible solution in which all the first [second] doses are administered in the last possible time slot, namely $T - \Delta^i$ [T]. Moreover, no stock is left at the end of the planning horizon, i.e., $s_T^i = 0$.

Starting from $\bar{x}^i$, our algorithm tries to move vaccinations earlier in a greedy fashion while always keeping the following relation true

$$x_t^i + y_t^i \leq B^i(t). \tag{13}$$

Algorithm 1 Greedy off-line heuristic

```
for all i ∈ V do
    x^i := x̄^i;
    for t = T − Δ^i down to 2 do
        c := min{s(t + Δ^i − 1), x_t, B(t − 1)};
        Augment x^i_{t−1} and y^i_{t+Δ^i−1} by c;
        Decrease x^i_t, y^i_{t+Δ^i}, s^i_{t−1} and s^i_{t+Δ^i−1} by c
    end for
    return x^i
end for
```

It is not hard to see that, at each step, c is the maximum amount of doses that can be moved earlier without violating the constraint $s_t \geq 0$.

In Section VI, we will report about the performance of the above described Algorithm 1.

B. Blind Algorithms

Hereafter we illustrate different greedy approaches to the decision on the number of doses to administer each day t, when no clairvoyance can be assumed on the future arrivals of doses. Note, however, that the plan output at t exploits the knowledge of the supply b_t^i on that day, for each vaccine i.

A simple idea consists in imposing that the amount of stock at the end of each day has to be equal at least to the number of second doses to be administered the next day (day-by-day no-out-of-stock condition).

The balance equation at the end of day t for the i-th vaccine is

$$\begin{aligned} s_t^i &= s_{t-1}^i - x_t^i - y_t^i + b_t^i \\ &= s_{t-1}^i - x_t^i - x_{t-\Delta^i}^i + b_t^i, \end{aligned} \tag{14}$$

due to the delay introduced between the first and second dose.

If we do not want to run out of stock and guarantee the administration of the second dose at day $t+1$, we must impose the following no-out-of-stock condition, which assumes that no doses arrive on day $t+1$ is equivalent to the following relation

$$s_t^i \geq y_{t+1}^i = x_{t+1-\Delta^i}^i. \tag{15}$$

Equation (15) embodies the day-by-day administration strategy. Since new arrivals may arrive on day t the number of first doses that can be safely administered at time t while still meeting the condition of Equation (15) is

$$x_t^i \leq s_t^i - x_{t-\Delta^i}^i - x_{t+1-\Delta^i}^i + b_t^i. \tag{16}$$

Again, in Equation (16), three quantities are known at the end of time $t-1$ (s_{t-1}^i, $x_{t-\Delta^i}^i$, and $x_{t+1-\Delta^i}$), while the fourth one (b_t^i) is random but will be known at the beginning of day t.

If the sum in the right-hand term of Equation (16) is not positive, then we will not be able to administer any first dose on day t. This unfortunate situation takes place if the number of arrivals on day t is

$$b_t^i \leq z_t^i = x_{t-\Delta^i}^i + x_{t+1-\Delta^i}^i - s_{t-1}^i \tag{17}$$

The risk of incurring the no-first-doses situation is then $\mathbb{P}[b_t^i \leq \gamma_t^i]$. For the ZIP process, this risk $\mathbb{R}_{\text{no-1}}^{(1)}$ is

$$\mathbb{R}_{\text{no-1}}^{(1)} = \begin{cases} 0 & \text{if } z_t^i < 0, \\ \pi + (1-\pi)\sum_{i=0}^{\gamma_t^i} \frac{\lambda^i}{i!} e^{-\lambda} & \text{if } \gamma_t^i \geq 0 \end{cases} \tag{18}$$

On the other hand, if no arrivals take place for a long period of time, we cannot even guarantee that the day-by-day strategy is applicable. In that case, we would not even be able to administer all the second doses. This very unfortunate case takes place if the following condition holds

$$s_{t-1}^i + b_t^i < y_t^i \rightarrow b_t^i < \eta_t^i = y_t^i - s_{t-1}^i \tag{19}$$

The risk of incurring the no-second-doses situation is then $\mathbb{P}[b_t^i \leq \eta_t^i]$. For the ZIP process, this risk $\mathbb{R}_{\text{no-2}}^{(1)}$ is

$$\mathbb{R}_{\text{no-1}} = \begin{cases} 0 & \text{if } \eta_t^i < 0, \\ \pi + (1-\pi)\sum_{i=0}^{\eta_t^i} \frac{\lambda^i}{i!} e^{-\lambda} & \text{if } \eta_t^i \geq 0 \end{cases} \tag{20}$$

Extending the above arguments, we may derive a prescription on the amount of first doses to be safely administered so that we are guaranteed that the stock we have at time $t-1$ is enough to cover the overall demand over the next q days, establishing a sliding window that is shifted each day. This more conservative administration strategy may be called the q-days-ahead strategy (it is readily seen that the day-by-day

strategy represents a special case of the q-days-ahead one when $q = 1$.)

In this case, the no-out-of-stock condition (15) clearly becomes

$$s_t^i \geq \sum_{\ell=1}^{q} y_{t+\ell}^i \tag{21}$$

As a consequence, if we include capacity constraints on the maximum number of doses that can be administered per day, the number of first doses that can be safely administered at time t would satisfy the following inequality:

$$x_t^i \leq \max\{0, \min\{k_t - y_t^i, s_{t-1}^i - \sum_{\ell=1}^{q} y_{t+\ell}^i + b_t^i\}\}. \tag{22}$$

The above relation directly suggests a simple myopic algorithm in which the risk of running out-of-stock is inversely proportional to the value of q.

Since we can rely on new arrivals to meet the constraint on the stock at the end of the day, we will not be able to safely administer first doses if the following condition holds

$$\begin{aligned} &s_{t-1}^i - x_{t-\Delta^i}^i + b_t^i - \sum_{\ell=1}^{q} y_{t+\ell}^i \leq 0 \rightarrow \\ &b_t^i \leq \phi_t^i = \sum_{\ell=0}^{q} y_{t+\ell}^i + x_{t-\Delta^i}^i - s_{t-1}^i. \end{aligned} \tag{23}$$

In the case of the ZIP process, the risk of not being able to administer any first dose is then

$$\mathbb{R}_{\text{no-1}}^{(q)} = \begin{cases} 0 & \text{if } \phi_t^i < 0, \\ \pi + (1-\pi) \sum_{i=0}^{\phi_t^i} \frac{\lambda^i}{i!} e^{-\lambda} & \text{if } \phi_t^i \geq 0 \end{cases} \tag{24}$$

Similarly, the risk of not even meeting the full demand for second doses equals that of not satisfying the inequality (19). Though the resulting risk expression equals that expressed for the day-by-day strategy in Equation (20), we must consider that the condition applied for the q-days-ahead strategy in Equation (21) clearly includes the one for the day-by-day strategy in Equation (15). The conclusion is that the risk of not being able to administer either first or second doses in the q-days-ahead strategy is surely lower than that suffered in the day-by-day strategy.

Though a suitable choice of q avoids the risk of dose shortages with a reasonable degree of confidence, it does not rule out the possibility of such an undesirable situation. As a consequence, in our experiments, we will consider (and assess the performance) of a *conservative* algorithm that guarantees that an adequate level of inventory is always available for second doses — with no knowledge on the future supplies. As one may expect, this conservative attitude of the algorithm has, of course, its disadvantages in terms of residual inventory and average vaccination times.

The basic idea is that whenever a number b of doses becomes available at t, one can immediately administer $b/2$ first doses and reserve the remainder to administer the corresponding second doses after the prescribed period. The output plan is obtained by augmenting the current solution each time a positive supply of new doses arrives. If the capacity (at t or $t + \Delta^i$) limits the number of doses that could be administered now, the possible excess of available doses at t gets transferred to the next period, and the procedure is iterated.

The conservative Algorithm 2 is illustrated hereafter. For each vaccine i and day t, we store the amounts x_t^i of the first doses to be administered. The current amount of supply available for the first doses is stored in a_t^i. Moreover, the parameter k_t^i is an upper bound set on the number of doses of vaccine i (capacity) that can be administered on the day t.

Algorithm 2 Conservative algorithm

1: **for** $i \in \mathcal{V}$ **do**
2: Initialize x, y and a as null vectors;
3: **for** $t = 1, 2, \ldots, T$ **do**
4: $a_t^i := a_t^i + b_t^i$;
5: $\delta := \max\{x_t^i + y_t^i + \frac{1}{2}a_t^i - k_t^i,\ x_{t+\Delta^i}^i + y_{t+\Delta^i}^i + \frac{1}{2}a_t^i - k_{t+\Delta^i}^i\}$; {Excess is computed}
6: **if** $\delta > 0$ **then**
7: $a_{t+1}^i := a_{t+1}^i + \frac{1}{2}a_t^i + \delta$;
8: $a_t^i := a_t^i - 2\delta$ {Excess is transferred}
9: **end if**
10: **if** $a_t^i > 0$ **then**
11: $x_t^i := x_t^i + \frac{1}{2}a_t^i$;
12: $y_{t+\Delta^i}^i := y_{t+\Delta^i}^i + \frac{1}{2}a_t^i$
13: **end if**
14: **end for**
15: **end for**

The idea is to compute the excess availability with respect to the capacity values for days t and $t + \Delta^i$. The quantity δ represent such an excess: If it is positive, we are not allowed to administer all the available $a_t^i/2$ doses plus the previously planned amount of $x_t^i + y_t^i$ vaccines. In this case, we are reducing the availability at the current day t and reserve the exceeding quota of vaccines as available for the next day $t+1$ (Steps 7 and 8.)

Note that, when $\delta > 0$, the maximum number of *additional* doses that can be administered at day t is the minimum between $k_t^i - x_t^i - y_t^i$ and $k_{t+\Delta^i}^i - x_{t+\Delta^i}^i - y_{t+\Delta^i}^i$. One can easily see that such a minimum is given by $\frac{1}{2}a_t^i - \delta$ and hence that the actual excess is $\frac{1}{2}a_t^i + \delta$. The latter quantity is therefore added to the availability a_{t+1}^i of the next day while it is cut from a_t^i. (For instance, if $\delta = x_t^i + y_t^i + \frac{1}{2}a_t^i - k_t^i > 0$, then a_t^i becomes $a_t^i - 2\delta$ and x_t^i is augmented to the maximum value allowed by the capcity limit, i.e., $x_t^i := x_t^i + \frac{1}{2}a_t^i = k_t^i - y_t^i$. Similarly, if $\delta = x_{t+\Delta^i}^i + y_{t+\Delta^i}^i + \frac{1}{2}a_t^i - k_{t+\Delta^i}^i > 0$ then we obtain that $y_{t+\Delta^i}^i = k_{t+\Delta^i}^i - x_{t+\Delta^i}^i$.)

VI. Experiments and results

In this section, we describe our computational experience to compare the different approaches described above that design daily vaccination plans over a time horizon of T days. We first define a set of performance metrics and then report the results of our simulation experiments.

As a model for the arrival process, we have employed a ZIP model where the parameters are respectively $\pi = 0.85$ and $\lambda = 10^7$. Those parameters give us an average daily number of doses equal to $1.5 \cdot 10^6$, which is consistent with the current trend in Europe. In Fig. 4, we see that European countries currently lie in the 0.5-1 million bracket, but are following a growing trend[3]. We have run 1000 simulations to generate as many realistic arrival scenarios, each corresponding to an instance of the problem. The capacity is set as a multiple (through the capacity factor) of the average number of arrivals.

In the first set of experiments, we compared the performance of the linear program (3)-(9) against the greedy heuristic Algorithm 1. The empirical probability density function of the relative gap $\frac{H-LP}{H}\%$ between the objective function values LP and H obtained by the linear program and the heuristic, respectively, is shown in Fig. 3. We have employed the Gaussian kernel method [19], with a data-driven kernel bandwidth set as in [20], equal to 0.3719 in our case. As can be seen, the mode is around 2%. As discussed above, both algorithms

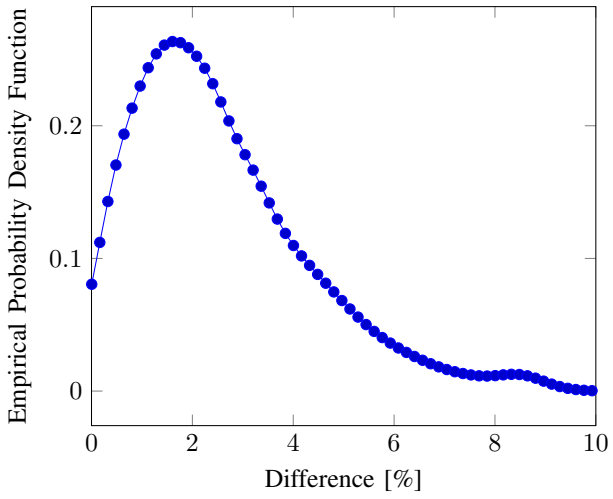

Fig. 3. Distribution of objective function percentual gap for Algorithm 1.

provide an off-line benchmark to measure the performance of the blind algorithms. The tests show that the heuristic algorithm is quite effective in most cases: The average gap is 2.57% with more than 90% of the instances with a relative gap below 5%. Though the greedy heuristic finds a solution much faster than the linear programming solver and we need to compute an optimal off-line solution several times, it is still definitely compatible with our experiments to use the optima provided by the linear programming solver, as, on the average, an optimum of a single instance is computed by Gurobi in around 60 milliseconds. (Indeed, running all five heuristics on a single instance requires around six milliseconds.) It is clear that in a different context, e.g., larger instances (greater T values), a larger number of scenarios, additional constraints, etc., due to the effectiveness of the greedy heuristic witnessed by the low gap values, the proposed algorithm might be a good alternative approach to provide the necessary off-line benchmarks.

Recall that the capacity values k_t^i measure the maximum number of doses of vaccine i that the system is able to dispense in day t, for all $i \in \mathcal{W}$. In our tests, for each arrival scenario, we have considered 14 runs of the algorithms each corresponding to a specific capacity value:

$$k_t^i = (1 + \frac{\alpha}{2})c^i \quad \alpha = 0, 1, 2, \ldots, 13;\ i \in \mathcal{V};\ t \in \mathcal{W} \tag{25}$$

in which, for all $i \in \mathcal{V}$, $c^i = \frac{1}{t}\sum_{t\in\mathcal{W}} b_t^i$ is a base-step capacity[4] (corresponding to the number of average per day arrival for vaccine i) and $(1 + \frac{\alpha}{2})$ is a *capacity factor*.

All the experiments here reported were run on a PC with CPU Intel i5 − 5300U 64bit 2.30GHz clock, and 8GB RAM. The algorithms were coded in Python 3, v3.8.5. The implementation of the LP model solver makes use of the Python Gurobipy library 9.1.2 [21].

In our experiments we measure the following metrics:

- *Average vaccination time*. When $s_T^i = 0$ the average vaccination time is exactly equal to the expression given in Equation (10). In order to compare the results among experiments in which there are different levels of unused stocks at the end of the planning period, we use the following expression:

$$\frac{\sum_{i\in\mathcal{V}}(s_T^i(T + \Delta^i) + 2\sum_{t\in\mathcal{W}} t y_t^i)}{\sum_{i\in\mathcal{V}}\sum_{t\in\mathcal{W}} b_t^i}. \tag{26}$$

  As in Equation (10), the denominator counts (twice) the total number of administered vaccines. The numerator is the sum of the vaccination days $t = 1, \ldots, T$, each weighted by (twice) the number of second doses administered that day. The additive term $s_T^i(T + \Delta^i)$ accounts for those doses that remain unused at the end of the day T and are then stocked for the next days. Basically, we are considering that half of those doses will be administered on the same day T and, as booster vaccinations, after Δ^i days.
- *Utilization*. It measures how efficiently the supplied doses have been utilized at the end of the planning period: This figure is simply the ratio between the total number of planned final doses $\sum_i \sum_t y_t^i$ over *half* of the total number of supplied doses $B(T)$ (which is the theoretical maximum). Clearly, low levels of utilization mean that at the end of the planning period, the proposed approach has in stock a significant number of doses. It is important to stress that these data include possible backlogs in the planned number of delivered vaccines. That is, a positive y_t^i value may exist in correspondence to a negative value of inventory level s_t^i. In such a case, this is equivalent to assume that s_t^i vaccines would take place on a day later

[3]See the latest data published on https://ourworldindata.org/grapher/daily-covid-19-vaccination-doses

[4]The idea is that a system with a base-step capacity c^i for all $i \in \mathcal{V}$ would be able to consume all the arrived doses of vaccine i, at the end of the planning period, only if it would deliver doses at its capacity level, each single day.

than t, the planned day. In this regard, we also report the following two figures.

- *Number of out-of-stock days*. The number of days in which there would be a shortage of doses, i.e., $s_t^i < 0$. In the graphics, we report these numbers as a percentage over the number T of days of the planning period.
- *Average backlog*. It is the absolute value of the average negative stocks $\frac{1}{T}\sum_i \sum_t |\min\{0, s_t^i\}|$ and it is also a measure of the average number of doses that would be administered after their prescribed date. This quantity might be of help in sizing an adequate level of safety stocks. For ease of readability, in the graphic of Fig. 7 this measure is reported as a percentage over the total number of supplied doses:

$$\frac{\sum_i \sum_t |\min\{0, s_t^i\}|}{T \sum_i \sum_t b_t^i}\%.$$

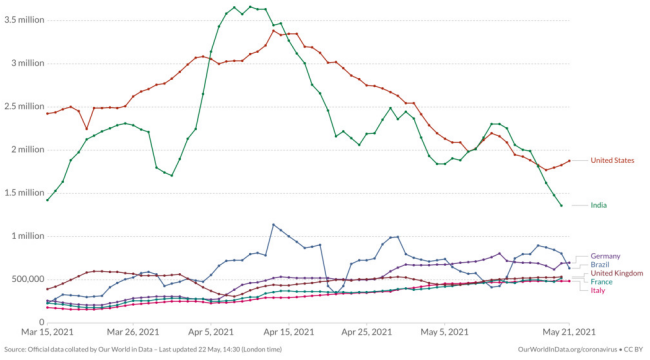

Fig. 4. Daily vaccination rates

We first take a look at the average vaccination time in Fig. 5. Of course, we aim at the lowest vaccination time possible. We see that the clairvoyant strategy achieves the best performance, as expected: the average vaccination time is roughly four months away from the start of the vaccination campaign. All the look-ahead strategies perform worse, with times getting longer as we lengthen our look-ahead horizon, playing safe against long periods of no dose arrivals. While guaranteeing that all second doses are administered on time, the conservative approach achieves an average vaccination time, which is not the worst in the group, ranking between the 7-days-ahead and the 14-days-ahead strategy. All the curves flatten out as the capacity factor grows. It appears then useless to have a capacity factor larger than 2 in most cases.

While both the clairvoyant and the conservative strategy guarantee that all second doses needs are met on the very same day, that's not the case for other strategies. We see that the vaccination system may incur a dose debt, where vaccination has to be postponed because no doses are available. The larger the fraction of days with no stocks, the worse the situation. In Fig. 6, we see that we may not meet the second doses needs as often as 40%. Assuming longer periods of no arrivals (i.e., Lengthening the look-ahead period) acts a hedge against worse periods and strongly reduces the fraction of out-of-stock days.

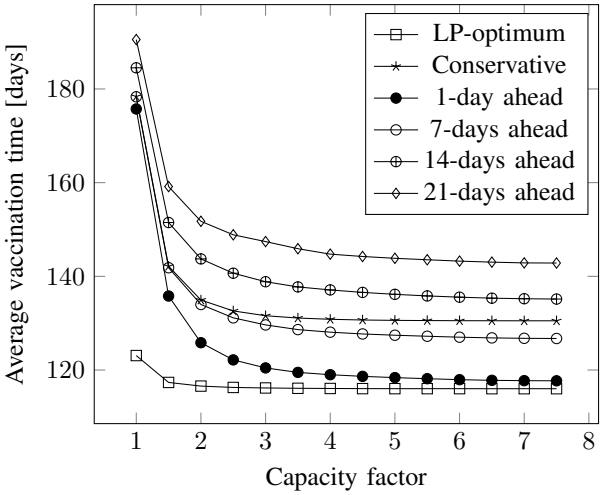

Fig. 5. Average vaccination time

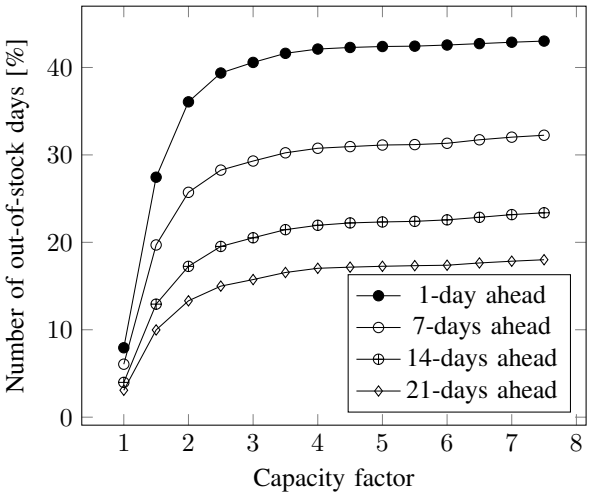

Fig. 6. Percentage of out-of-stock days

However, the relevance of out-of-stock days depends on how many doses we miss. For that reason, we also take a look at the actual backlog, which is shown in Fig. 7. The worst case takes place with the 1-day-ahead strategy, where the doses needed amount to 3% at most of the overall number of arrivals. Here, being limited by vaccination capacity helps since doses exceeding the daily capacity must be kept for use in the following days, thus acting as a reserve for days of no arrivals. Also, longer look-ahead strategies are less impacted since a longer planning period allows to override no-arrivals periods.

Since the final aim is to exploit the delivered doses as most as possible, we can analyze the utilization rate, which gives us the percentage of first-plus-second doses that have actually been administered (i.e., the percentage of vaccination cycles that have been completed) over the whole number of doses. In Fig. 8, we see that being limited by the vaccination capacity may strongly lower the capability to vaccinate as many people

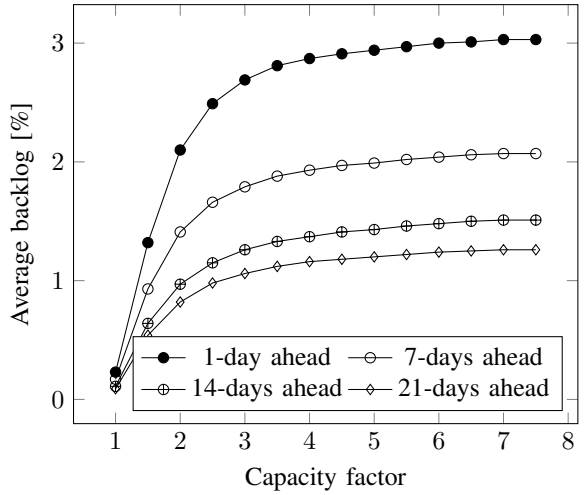

Fig. 7. Average backlog

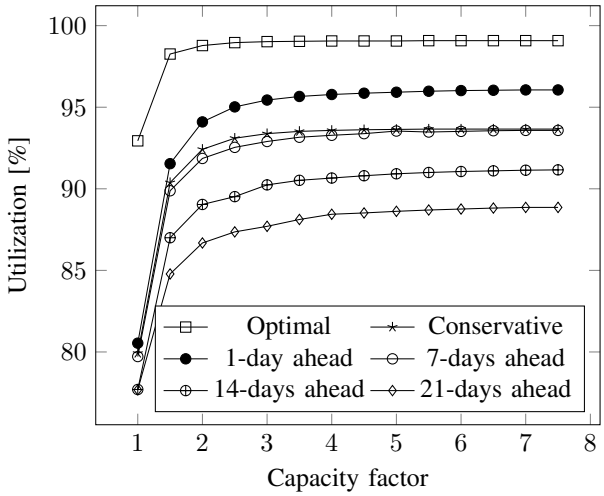

Fig. 8. Completion of vaccination cycle [%]

as possible. The curve steeply grows as we increase the capacity factor from 1 to 2. Though the clairvoyant strategy is again the best in class, as expected, with the 1-day ahead-strategy as a not-too-close runner up, the conservative strategy achieves a higher utilization than all the look-ahead strategies with a look-ahead period longer than seven days.

VII. Conclusion

In this work, we have presented a computational study to compare alternative strategies for massive vaccination under uncertain supply. The aim is to size the vaccination center capacity adequately.

The clairvoyant strategy can be set as the benchmark, leading to an optimal linear programming solution. Among the strategies examined, the best trade-off is achieved by the conservative strategy, where the administration of second doses is guaranteed on time since second doses are kept in stock as soon as the pertaining first doses are administered. Its average vaccination time is just 11% longer than what the clairvoyant strategy offers, with a utilization rate that is just 6% lower.

Among the look-ahead strategies, the 7-days-ahead strategy has very close performance regarding vaccination times and utilization, but lack of stocks is incurred in 30% of time (while this never happens with the conservative strategy). The 1-day-ahead strategy would allow reaching vaccination times just slightly longer than the clairvoyant strategy with a utilization ratio just 4% lower, but results in an unacceptable 40% fraction of out-of-stock days.

As to the sizing of the vaccination center, the capacity appears as a critical factor as long as it is too low. Capacity factors between 2 and 3 (i.e., the capability to vaccinate daily as many as 2-3 times the average number of arrivals on a day) are enough to achieve performances very close to those obtainable under infinite capacity.

As to the latter issue, we might improve the vaccination strategy by including the possibility to allow the decision-maker to change the capacities dynamically for each vaccine and each day. In practice this is not always doable, or it could be limited by a number of constraints, e.g., the severe requirements that may be imposed on the stocking devices.

Any future work will also benefit from the growing availability of data about vaccine delivery, which will allow for more accurate modelling of the process of dose arrivals. In addition, strategies envisaging mixed vaccination, i.e., adopting a different vaccine for the second dose, could incorporate knowledge about the joint distribution of arrivals for the two vaccines.

In addition, strategies that have been considered so far do not exploit any information on future supply, though uncertain, future arrivals could be categorized into a set of scenarios. In this regard, a robust optimization approach [22], [23], [24] is suitable to be devised for our mass vaccination planning problem. Several criteria have been adopted in the robust optimization literature. For instance, a widely used robustness criterion is the so-called maximin criterion, according to which the best worst-case performance has to be sought. In our case, such a robust optimization approach would maximize the system performance while guaranteeing the feasibility of the prescribed vaccine administration along the whole planning horizon for any possible future scenario.

Another important topic to be considered in future extended models for mass vaccination is the design of inducement policies to encourage the largest possible fraction of the population to uptake the vaccine. From a methodological point of view, Stackelberg game approaches (see, e.g., [25], [26]) appear as a natural direction to deal with these issues. Of course, from a different perspective, also communication strategies play a crucial role, as discussed in [27] in which the effectiveness of different health communication frames is assessed.

Finally, the reliability of the implemented decision support tool must be evaluated: methods as the one illustrated in [28] are an essential appliance to investigate the influence of any system component failure on the system functioning.

REFERENCES

[1] E. J. Edwardes, *A concise history of small-pox and vaccination in Europe*. HK Lewis, 1902.

[2] D. L. Heymann and R. B. Aylward, "Mass vaccination: when and why," *Mass Vaccination: Global Aspects—Progress and Obstacles*, pp. 1–16, 2006. doi: http://dx.doi.org/10.1007/3-540-36583-4

[3] P. Homayounfar, "Process mining challenges in hospital information systems," in *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 2012, pp. 1135–1140.

[4] E. Zaitseva, V. Levashenko, and M. Rusin, "Reliability analysis of healthcare system," in *2011 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 2011, pp. 169–175.

[5] M. Naldi, G. Nicosia, A. Pacifici, and U. Pferschy, "Profit-fairness trade-off in project selection," *Socio-Economic Planning Sciences*, vol. 67, pp. 133–146, 2019. doi: http://dx.doi.org/10.1016/j.seps.2018.10.007

[6] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2011. ISBN 9781400841035. doi: http://dx.doi.org/10.1515/9781400841035

[7] E. H. Kaplan, D. L. Craft, and L. M. Wein, "Emergency response to a smallpox attack: The case for mass vaccination," *Proceedings of the National Academy of Sciences*, vol. 99, no. 16, pp. 10 935–10 940, 2002. doi: http://dx.doi.org/10.1073/pnas.162282799

[8] A. R. da Cruz, R. T. N. Cardoso, and R. H. C. Takahashi, "Multiobjective synthesis of robust vaccination policies," *Applied Soft Computing*, vol. 50, pp. 34–47, 2017. doi: http://dx.doi.org/10.1016/j.asoc.2016.11.010

[9] D. Bertsimas, J. Ivanhoe, A. Jacquillat, M. Li, A. Previero, O. S. Lami, and H. T. Bouardi, "Optimizing Vaccine Allocation to Combat the COVID-19 Pandemic," *medRxiv*, 2020. doi: http://dx.doi.org/10.1101/2020.11.17.20233213

[10] H. M. Wagner and T. M. Whitin, "Dynamic version of the economic lot size model," *Management Science*, vol. 5, pp. 89–96, 1958. doi: http://dx.doi.org/10.1287/mnsc.5.1.89

[11] W. I. Zangwill, "A backlogging model and a multi-echelon model of a dynamic economic lot size production system-a network approach," *Management Science*, vol. 15, no. 9, pp. 506–527, 1969. doi: http://dx.doi.org/10.1287/mnsc.15.9.506

[12] S. S. DeRoo, N. J. Pudalov, and L. Y. Fu, "Planning for a COVID-19 vaccination program," *Jama*, vol. 323, no. 24, pp. 2458–2459, 2020. doi: http://dx.doi.org/10.1001/jama.2020.8711

[13] R. Shretta, N. Hupert, P. Osewe, and L. J. White, "Vaccinating the world against COVID-19: getting the delivery right is the greatest challenge," *BMJ Global Health*, vol. 6, no. 3, 2021. doi: http://dx.doi.org/10.1136/bmjgh-2021-005273

[14] L. Bell and R. Wagner, "Modeling Emergency Room Arrivals Using the Poisson Process," *The College Mathematics Journal*, vol. 50, no. 5, pp. 343–350, 2019. doi: http://dx.doi.org/10.1080/07468342.2019.1662710

[15] M. Naldi, "Measurement-based modelling of internet dial-up access connections," *Computer networks*, vol. 31, no. 22, pp. 2381–2390, 1999. doi: http://dx.doi.org/10.1016/S1389-1286(99)00091-2

[16] D. K. Agarwal, A. E. Gelfand, and S. Citron-Pousty, "Zero-inflated models with application to spatial count data," *Environmental and Ecological statistics*, vol. 9, no. 4, pp. 341–355, 2002. doi: http://dx.doi.org/10.1023/A:1020910605990

[17] A. F. Zuur, E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith, *Zero-truncated and zero-inflated models for count data*. Springer, 2009, pp. 261–293. doi: http://dx.doi.org/10.1023/10.1007/978–0–387–87 458–6_11.

[18] S. Beckett, J. Jee, T. Ncube, S. Pompilus, Q. Washington, A. Singh, and N. Pal, "Zero-inflated Poisson (ZIP) distribution: Parameter estimation and applications to model data from natural calamities," *Involve, a Journal of Mathematics*, vol. 7, no. 6, pp. 751–767, 2014. doi: http://dx.doi.org/10.2140/involve.2014.7.751

[19] B. W. Silverman, *Density estimation for statistics and data analysis*. Routledge, 2018. ISBN 9780412246203 doi: http://dx.doi.org/10.1201/9781315140919

[20] S. J. Sheather and M. C. Jones, "A reliable data-based bandwidth selection method for kernel density estimation," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 53, no. 3, pp. 683–690, 1991. doi: http://dx.doi.org/0.1111/j.2517-6161.1991.tb01857.x

[21] Gurobi, "Gurobi Optimizer Reference Manual," https://www.gurobi.com/documentation/9.1/refman/index.html, 2020, [Online; accessed 24-May-2021].

[22] P. Detti, G. Nicosia, A. Pacifici, and G. Zabalo Manrique de Lara, "Robust single machine scheduling with a flexible maintenance activity," *Computers and Operations Research*, vol. 107, pp. 19–31, 2019. doi: http://dx.doi.org/10.1016/j.cor.2019.03.001

[23] A. Parnianifard, A. S. Azfanizam, M. K. A. Ariffin, and M. I. S. Ismail, "An overview on robust design hybrid metamodeling: Advanced methodology in process optimization under uncertainty," *International Journal of Industrial Engineering Computations*, vol. 9, no. 1, pp. 1–32, 2018. doi: http://dx.doi.org/10.5267/j.ijiec.2017.5.003

[24] D. S. Yamashita, V. A. Armentano, and M. Laguna, "Robust optimization models for project scheduling with resource availability cost," *Journal of Scheduling*, vol. 10, no. 1, pp. 67–76, 2007. doi: http://dx.doi.org/10.1007/s10951-006-0326-4

[25] U. Pferschy, G. Nicosia, A. Pacifici, and J. Schauer, "On the Stackelberg knapsack game," *European Journal of Operational Research*, vol. 291, no. 1, pp. 18–31, 2021. doi: http://dx.doi.org/10.1016/j.ejor.2020.09.007 Cited By 0.

[26] U. Pferschy, G. Nicosia, and A. Pacifici, "A Stackelberg knapsack game with weight control," *Theoretical Computer Science*, vol. 799, pp. 149–159, 2019. doi: http://dx.doi.org/10.1016/j.tcs.2019.10.007

[27] M. Motta, S. Sylvester, T. Callaghan, and K. Lunz-Trujillo, "Encouraging COVID-19 Vaccine Uptake Through Effective Health Communication," *Frontiers in Political Science*, vol. 3, p. 1, 2021. doi: http://dx.doi.org/10.3389/fpos.2021.630133

[28] E. Zaitseva, J. Kostolny, M. Kvassay, V. Levashenko, and K. Pancerz, "Failure analysis and estimation of the healthcare system," in *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems*, M. P. M. Ganzha, L. Maciaszek, Ed. IEEE, 2013, pp. pages 235–240.