

## Title

Hospital preparedness in epidemics by using simulation. The case of COVID-19

## Authors

Daniel Garcia-Vicuña<sup>1</sup>, Laida Esparza<sup>2</sup>, and Fermin Mallor<sup>1</sup>

## Affiliation(s) and address(es) of the author(s)

<sup>1</sup>Institute of Smart Cities, Public University of Navarre

Campus Arrosadia

Pamplona, 31006, SPAIN

<sup>2</sup>Hospital Compound of Navarre

Irunlarrea, 3

Pamplona, 31008, SPAIN

## E-mail address, and telephone number(s) of the corresponding author:

mallor@unavarra.es

+34 948 16 92 15

## 6-digit ORCID of the author(s):

Fermin Mallor: "0000-0001-9800-1498"

## Abstract

This paper presents a discrete event simulation model to support the decision-making concerned with the short-term planning of the necessary hospital resources, especially Intensive Care Unit (ICU) beds, to face outbreaks, as the SARS-CoV-2. Being used as a short-term forecasting tool, the simulation model requires an accurate representation of the current system state and high fidelity in mimicking the system dynamics from that state. The two main components of the simulation model are the stochastic modeling of the admission of new patients and the patient flow through the hospital facilities. For the patient arrival process, we analyze different models based on growth curves of the twenty most affected countries (until June 15) and propose the use of the Gompertz curve. The length of stay is divided into several stages, each one modeled separately. We analyze the starting of the simulation model, which requires different procedures depending on the information available about the patients currently hospitalized. We also report the use of this simulation model during the COVID-19 outbreak in the Autonomous Community of Navarre, in Spain. Every day, the research team informed the regional logistic team in charge of planning the health resources, who programmed the ward and ICU beds based on the resulting predictions.

## Keywords

COVID-19, bed occupancy, hospital resources planning, Gompertz growth model, patient flow modelling

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## Declarations

**Funding:** the grant MTM2016-77015-R (AEI, FEDER EU).

**Conflicts of interest/Competing interests:** Declarations of interest: none.

**Availability of data and material:** Historical data is available on publicly accessible web pages. The patient data used in the case study is not authorized to be made public.

**Code availability:** custom code programming with Python.

**Authors' contributions:** **Daniel Garcia-Vicuña:** Methodology, Programming, Formal Analysis, Investigation, Data Curation, Writing. **Laida Esparza:** Conceptualization, Validation, Resources, Investigation, Data Curation, Writing. **Fermin Mallor:** Conceptualization, Methodology, Formal Analysis, Data Curation, Writing, Supervision, Funding Acquisition.

## Acknowledgments

This paper has been supported by grant MTM2016-77015-R (AEI, FEDER EU). The authors are grateful to the medical staff of the main hospital in Navarre for providing data and validating the simulation model.

# 1 Introduction

COVID-19 pandemic presents an important threat to global health. Since the outbreak in China in early December 2019, the number of patients confirmed to have the disease has exceeded seventeen million, and more than 650,000 people have died from COVID-19 infection (up to 30 July 2020 <https://coronavirus.jhu.edu/map.html>). Regularly updated information on the COVID-19 outbreak is also available on the [European Centre for Disease Prevention and Control's \(ECDC\) website](#) [1], the [European Commission's \(EC\) website](#) [2], and the [World Health Organization's \(WHO\) website](#) [3]. This outbreak has changed how health care is delivered, and it has affected the operation in hospitals, which have experienced an increase in demand. The pandemic has impacted on the Intensive Care Units (ICU), which involve highly specialized personnel and expensive technical sanitary material, that require time to their planning and acquisition. Therefore, an accurate prognosis of the necessary resources is needed to efficiently manage the scarce resources and to provide the best possible care to patients. The precision of the predictions allows preparing the response and helping to save lives. Usually, the hospitalization bed is still widely used as a hospital (ICU) management parameter both at the strategic and operational levels.

Hospitals are complex systems evolving in a stochastic environment whose uncertainty is even higher in pandemic periods because of the lack of knowledge about the spread of the disease and its consequences on patients. In this unsettled context, simulation emerges as a suitable analytical tool, since it can represent the complexity of the analyzed system and the variability and uncertainty of its environment. Besides, simulation can be used in combination with other analytical techniques. The literature contains numerous bibliographical references relating to the use of simulation models for decision making in the healthcare context, as it is exposed in Section 2. Most of the applications use the simulation to support strategic decisions, usually for dimensioning of resources, for their scheduling, or for their management. All these cases require the design of a simulation model to reproduce the performance of the health system in its stationary state and evaluate the resource levels, the patient flow management policies, and the decision making process in the long term. The recommendations obtained from the simulation analysis are meant to be implemented in the health system and last for a certain time horizon.

However, a simulation model designed to help to make tactical decisions related to the provision of specialized health resources during the current outbreak has to focus on the transition period, to project in the future the current state of the hospital. A discrete event simulation (DES) model is presented in this research that combines a dynamic forecast method to predict (simulate) the new incoming patients and the reproduction of the patient flow through the health system. The simulation output provides scenarios of future utilization of resources that keep informed the health authorities of future needs and give them time for their planning. Therefore, the main feature of the simulation model presented here is its capacity of reproducing the evolution of the health system from its current state to be used as a forecasting tool. The correct representation of the current state of the hospital and the methodology used to start the simulation plays a crucial role in the accuracy of the simulation results. Nevertheless, these two aspects strongly depend on the available information about the hospital and its admitted patients. In some cases, there could be data to the patient level, that is, with knowledge of the admission and discharge dates (including transferring dates between facilities or departments inside the hospital) while in other cases only aggregated information is available, for example, the total number of admitted patients each day. The simulation model we have developed is flexible enough to handle the different levels of information.

The main contribution of this paper is the proposal of a new simulation framework to enable the prediction in the short term (from days to one month) the need for critical resources to provide healthcare to COVID-19 patients. The simulation framework can be adapted to be applied to other future outbreaks. To reach this main contribution our research includes:

- A simulation method of patient arrivals based on population growth (PG) models.
- A statistical analysis of four different PG models to analyze their accuracy to represent the expansion of the pandemic and their forecasting capacity.

- The representation of the current state of the health system and the beginning of the simulation from this state by considering four different information levels for the admitted patients.
- A dynamic statistical analysis of the patient flow through the health facilities.
- The combination of all elements in a DES model flexible enough to recreate scenarios based on stochastic models fitted to data (data-driven prediction), scenarios defined by expert judgment, and a mixture of both.
- From a practical point of view, this paper also reports a successful real application of simulation to support an important decision-making process essential for the health of patients in an Autonomous Community of Spain (Navarre).

The rest of the paper is organized as follows. Section 2 presents a literature review of papers dealing with the use of quantitative methods to predict needs in health care systems and improve their management. Section 3 studies the adequacy of PG models to predict the evolution of the spread of a pandemic. The modeling of the patient flow through the hospital is presented in Section 4. The structure of the DES model and the methodology to begin the simulation under different available data are included in Section 5. Section 6 describes the simulator, with its inputs and outputs. Results of the application of the simulation model to the Autonomous Community of Navarre (Spain) are included in Section 7. Finally, Section 8 ends the paper with the conclusions of this work.

## 2 Related literature

Simulation is one of the most suitable analytical tools for the analysis of complex systems, as the health systems are, which is reflected in numerous articles in the specialist literature describing the use of simulation models for decision-making in the healthcare context. DES has been applied to model and analyze all aspects of logistics management in healthcare. In particular, it has been used to improve patient flow management, bed-planning, waiting list management, health service design, medical staff scheduling, etc. For reviews of the use of simulation models in healthcare, we refer to [4–7]. Usually, these simulation models focus on studying the stationary state of the health system to support strategic decisions related to the dimensioning of resources or designing management policies.

The ultimate goal of these models is to reconcile resource availability with demand in order to provide high-quality healthcare to patients while maintaining a reasonable level of human and technological resources. Problems analyzed into this framework are patient flow [8, 9], bed planning [10–12], health service design [13] and medical staff scheduling [14], among others. Although discrepancies between assumptions made in mathematical simulation models and behavior of real health systems reported in the medical literature have been pointed out [15], there is no doubt about the usefulness of simulation models for the analysis of relevant problems in complex health systems.

However, simulation not only helps to ensure that medical staff and facilities are offering the highest quality services but also increases the likelihood of following best practices. Since the pandemic began, all national governments and the World Health Organization have extensively used simulation modelling to decide the best strategies to reduce the impact of COVID-19. Currie et al. [16] identify challenges from this disease and discuss how simulation modelling could help to support decision-makers in making the most informed decisions.

The accuracy of a simulation model to predict the necessary resources during a pandemic depends on an accurate model to predict the arrival of new patients to the health service. Infectious disease prediction models mainly include differential equation prediction models based on population dynamics [17, 18]. These mathematical models are essential to understand the course of the epidemic and to plan effective control strategies [19–21]. One of the most widely used models in the human-to-human transmission is the SIR model [22]. The individuals of the population are divided into different categories, each one considered as a possible state for the

individual: S (Susceptible), I (Infected), and R (Remove). The population in each state is calculated over time from the estimation of the transition rates among these states. By increasing the complexity of this model, it is possible to recreate the spread of specific epidemics. In fact, for the COVID-19 pandemic, extensions of the classical SIR model have been developed [23–27], as well as stochastic transmission models [28, 29]. However, this kind of models are complicated and need strong assumptions and simplifications, because they are based on a set of few differential equations with initial conditions and a number of adaptive parameters [30–33].

Therefore, growth population models suppose a simpler alternative to model the accumulated number of infected people. Growth curves are found in a wide range of areas, such as fishery research [34, 35], biology [36] or other infectious disease outbreaks [37–40]. Specifically, Logistic, Gompertz, Rosenzweig, and Richards models have been already used to model the spread of outbreaks such as A/H1N1 and Ebola in [41]. In relation to COVID-19 disease, several papers have been found in the literature which develop a growth model to predict new positive cases in different countries such as China [42], India [43], Spain [44], and other European countries [45]. These mathematical models present a set of mathematical equations that include some adaptive parameters that can be determined numerically based on available real data [46]. So, the model can be used daily (by updating the number of positive cases) and automatically adapt to the evolution of each parameter.

If all the mathematical models mentioned in the previous paragraphs are capable of fitting well to the real data, it would be possible to predict quite accurately what might happen in the future (e.g., emergency planning, resource allocation) [47–49]. This is very important especially for those resources that are normally scarce in a hospital, such as ICU beds. Manca et al. [50] present and discuss a few regression models developed using historical data of ICU patients and deaths during COVID-19 pandemic. They are capable of reproducing the bed occupancy curve using mathematical models, which can be very useful from the point of view of decision-making and emergency planning in future pandemics.

Besides, in recent decades, simulation in healthcare has become a new way of learning through experiences with advanced technology [51, 52]. With this modern education and training technique, healthcare professionals can learn new cognitive, technical, and behavioral skills. Before working in real-world scenarios, where real patients are treated, both professionals and students can use this experimental learning style to develop their skills and expand knowledge without taking any risk in their decision-making processes [53]. Simulation models, as the one presented in this paper, can serve also for learning the management of health care services in emergencies. One of the most critical decisions in periods of scarcity of resources is their assignment to patients, especially when these decisions can make the difference between recovery and death, as it could happen with the admission of patients to ICUs. This triage is difficult to implement during pandemics where the situations of a scarcity of resources are further aggravated. Different protocols for ICU triage during a pandemic have been suggested in [54–56]. Foreseeing the increasing need for beds is essential to avoid ethical dilemmas [15, 57]. According to Utley et al. [58], “the impact of triage is dependent on the level of demand and on the scale of achievable differences between included and excluded groups in terms of anticipated length of stay and critical care survival”. A simulation model can help better planning of critical resources, during a pandemic; and as an off-line tool, it can be used as a learning tool to test new protocols for triage, because sometimes triage is not as effective as might be expected, and other hard-to-anticipate factors must be considered.

### **3 Modelling the patient arrival pattern**

In this section, the adequacy of population models to predict new cases is studied. First, four different models are compared statistically to select the more suitable model. Then, we described the simulation of patient arrival to the hospital.

### 3.1 Population growth models

The simulation model needs to include a stochastic model representing the pandemic evolution to predict (simulate) the number of patients arriving at the health system. Because of their simplicity, we propose the use of PG models. A number of growth models are found in the literature, such as the models of Gompertz [59], Richards [60], Stannard [61], and logistic model and others [62]. They start with exponential growth but gradually decrease their specific growth rates. The equations that relate the number of cases in the population (infected, patients needing hospitalization, etc..) and the time are shown in Table 1.

Model	Equation
Logistic	$y = \frac{a}{1 + \exp(b - cx)}$
Gompertz	$y = a \cdot \exp[-\exp(b - cx)]$
Richards	$y = a\{1 + v \cdot \exp[k(\tau - x)]\}^{(-1/v)}$
Stannard	$y = a \left\{ 1 + \exp \left[ -\frac{(l + kx)}{p} \right] \right\}^{(-p)}$

Table 1. Models considered for the data fitting and their equations.

In this section, we carry out two statistical analysis for elucidating the adequacy of the PG models to represent and predict the evolution of the pandemic caused by the SARS-CoV-2 virus. The first analysis evaluates the capacity of the PG models to fit complete sets of real data representing positive cases registered in different countries. Specifically, PG models have been fitted to data coming from the 20 most-affected countries by COVID-19 on June 15, as it was recorded in [Worldometer](#) [63]. The parameter estimation of the PG models is done by minimizing the sum of squared errors. There are functions implemented in free software that perform this estimation of parameters, for instance, the *curve\_fit()* function in the *optimize* module of SciPy in Python [64] or the *growthrates* package in R [65]. The fit quality is measured by the Mean Absolute Errors (MAE). Table 2 includes all MAE values calculated for each country and model. The best fits are marked in bold (differences less than 0,1% are not distinguished). Additional information in this table is the total population of each country and the total number of positive cases on June 15.

These results show that similar fitting quality of Gompertz, Richards, and Stannard models, while the Logistic PG model underperforms the other three.

#	Country	Population	Total positive cases (2020-06-15)	Logistic	Gompertz	Richards	Stannard
1	USA	330,922,877	2,094,069	47,128.9	<b>21,461.2</b>	21,464.4	21,466.7
2	Brazil	212,496,348	867,624	3,182.3	3,245.1	<b>2,754.7</b>	<b>2,754.7</b>
3	Russia	145,932,063	528,964	5,667.4	<b>1,688.7</b>	<b>1,689.0</b>	<b>1,689.2</b>
4	India	1,379,418,901	332,424	1,392.9	<b>537.9</b>	<b>538.0</b>	<b>538.2</b>
5	UK	67,871,466	295,889	4,758.0	<b>1,046.6</b>	<b>1,046.9</b>	<b>1,047.2</b>
6	Spain	46,754,084	245,194	5,676.2	<b>2,261.7</b>	<b>2,262.1</b>	<b>2,262.6</b>
7	Italy	60,465,149	236,989	4,928.0	<b>1,061.2</b>	<b>1,061.6</b>	<b>1,061.7</b>

8	Peru	32,951,046	229,736	2,399.4	<b>1,255.3</b>	1,256.6	1,257.6
9	Iran	83,944,885	187,427	8,105.8	<b>6,176.3</b>	<b>6,176.6</b>	<b>6,176.9</b>
10	Germany	83,773,297	186,461	4,091.4	<b>1,519.2</b>	<b>1,519.5</b>	<b>1,519.7</b>
11	Turkey	84,299,464	178,239	5,335.7	<b>2,418.3</b>	<b>2,418.8</b>	<b>2,419.1</b>
12	Chile	19,109,226	174,293	1,132.5	1,601.2	<b>1,068.5</b>	<b>1,068.5</b>
13	France	65,267,844	157,220	3,396.4	<b>1,547.0</b>	<b>1,547.2</b>	<b>1,547.3</b>
14	Mexico	128,873,820	153,507	<b>1,418.9</b>	1,641.0	1,535.5	1,535.5
15	Pakistan	220,685,460	144,478	1,631.5	1,324.5	<b>1,301.0</b>	1,321.6
16	Saudi Arabia	34,788,836	127,541	1,709.0	<b>814.9</b>	<b>814.9</b>	<b>814.9</b>
17	Canada	37,728,057	98,776	1,494.6	<b>331.0</b>	<b>331.1</b>	<b>331.1</b>
18	Bangladesh	164,618,467	87,520	655.6	<b>315.5</b>	<b>315.6</b>	<b>315.7</b>
19	China	1,439,323,776	84,335	1,166.6	1,133.0	<b>1,097.6</b>	<b>1,097.7</b>
20	Qatar	2,807,805	79,602	417.7	369.2	<b>263.6</b>	<b>263.6</b>

Table 2. The 20 most-affected countries by COVID-19 until June 15. The last four columns show the MAE calculated for the fit with each of the applied models. Bold values represent the best scores.

The second statistical analysis is designed to test the predictive capacity in the short-medium term of the PG models. The accurate prediction of future cases of the disease is crucial for the accurate prediction of the resources needed. Three experiments have been carried out to determine which model has a better predictive capacity. For each country, the dates in which the data exceed 25%, 40%, and 65% of positive cases registered on June 15 are selected. The rationale behind this choice is the following: 25% represents an early stage of the pandemic but with enough information to fit the curves; however, when few data are available, they lie in the exponential growth phase and lead usually to overestimations of the maximum of the function. This problem is accentuated in countries where the pandemic has a strong impact at the beginning, such as in Spain or Italy. In Section 3.2 it is explained how we deal with this problem. In addition, the values 40% and 65% are chosen to observe how each model predicts as more information becomes available.

The prediction of the fitted curves for the next 5, 10, and 15 days is assessed by calculating the MAE. These time horizons are considered as sufficient for the hospital managers to adapt extra resources for new needs. As new positive case data is added every day, and predictions are refreshed also every day, the long-term predictive capacity of the model will not be analyzed. The results of these analyzes are collected in Appendix A. On the one hand, we present fits to the curves until the selected days, obtaining an MAE for each model and country. On the other hand, the tables of the MAEs made in the predictions are shown. To facilitate the comparison of results, MAEs are normalized by the total number of positive cases on the selected days. From these results, we can conclude that the Gompertz model outperforms in predictive capacity the other PG models.

Table 3 summarizes the relevant information from all the tables in Appendix A. It indicates the number of countries in which each model is the best in terms of predictive capacity (as before, differences smaller than 0.1% have been considered equal). It is observed that the Gompertz model is the one that more accurately predicts future values, specifically in all the situations analyzed except for one. For this reason, the Gompertz model is recommended to predict new cases of COVID-19.

Model	25%			40%			65%		
	5 days	10 days	15 days	5 days	10 days	15 days	5 days	10 days	15 days
Logistic	4	5	5	4	4	3	2	1	2*
Gompertz	<b>11</b>	<b>12</b>	<b>13</b>	<b>13</b>	<b>14</b>	<b>14</b>	14	<b>15</b>	<b>13*</b>
Richards	<b>11</b>	9	7	<b>13</b>	9	10	<b>15</b>	<b>15</b>	11*
Stannard	<b>11</b>	8	7	<b>13</b>	9	10	<b>15</b>	<b>15</b>	11*

Table 3. The number of countries in which each model is equal or better than the others in terms of predicting new positive cases for the next 5, 10, and 15 days. Bold values represent the best scores. \*These values are of 18 countries because 2 of them have no real data for that period.

## 3.2 Patient arrival pattern simulation

The Gompertz growth model can be used to simulate both the number of positive cases and the entry of patients to the health system. The parameters of the original equation of this model (Table 1) have a more mathematical than biological interpretation ( $a, b, c \dots$ ), like most equations that describe a sigmoidal growth curve. Therefore, before using it in our modeling, it is convenient to carry out a transformation to easily interpret the results obtained in the fit. Zwietering et al. [66] rewrite the Gompertz growth model as it is shown in equation (1).

$$c(t) = A \exp \left( -\exp \left( \frac{\mu_m \cdot e}{A} (\lambda - t) + 1 \right) \right) \quad (1)$$

where,

- $e = \exp(1)$
- $c(t)$  is the cumulative number of positive cases or hospitalized patients until time  $t$ .
- $A$  is a parameter of the growth model that represents the total number of positive cases or hospitalized patients at the end of the outbreak.
- $\mu_m$  is the maximum specific growth rate of the curve.
- $\lambda$  is the lag time, defined as the  $t$ -axis intercept of the tangent through the inflection point.

Once the curve  $c(t)$  is fitted, for example to cumulative hospitalized patients, it is used to simulate the number of new hospitalized patients of each of the following days. Let consider  $t_i$  and  $t_{i+1}$  two consecutive days in the future, the number of expected arrivals for the day  $t_{i+1}$  is calculated as  $c(t_{i+1}) - c(t_i)$ . Then, the simulation model samples the number of hospitalized patients for that day  $t_{i+1}$  from a Poisson distribution with mean  $c(t_{i+1}) - c(t_i)$ . This procedure is applied to simulate the hospitalized patients each future day. The Gompertz curve is fitted each time a new datum is observed.

At the first stages of the outbreak could not be enough data to fit the Gompertz model, overall in terms of hospitalized patients. In this case, we propose two alternatives that rely also on the use of PG models:

The first one consists of fitting the growth model to cumulative positive cases registered in a greater area, for example, the whole country when only a region of it is considered. Two factors are required to transform the predicted positive cases of the greater area into arrivals to hospitals in the subarea: the first one is a population factor that relates the population of both areas (e.g. the region of Navarre in Spain represents the 1.3% of the whole country population), and the second one transforms the new cases into patients requiring hospitalization (at first stages about 40% of detected cases needed hospitalization). Denoting by  $f_g$  and  $f_h$  both geographic and hospitalization factors, and by  $C(t)$  the fitted Gompertz curve to the positive cases in the bigger area, then  $c(t) = f_g f_h C(t)$ . Given that in this first period there is a data shortage,  $f_h$  can be estimated by experts according to data published in the literature referred to the behavior of the disease in countries where it is more spread. This indirect technique could be applied also to the detected cases in the same area and apply only the  $f_h$  factor.

Secondly, when the hospitalization cases begin to accumulate significantly (meaning the observed cumulative cases are starting its exponential growth) then we propose to fix (by expert opinion) the maximum population parameter ( $A$  in equation (1)) and to estimate the other two parameters by minimizing the sum of squared errors.



## 4 Modelling the patient flow

This section focusses on modelling the patient flow through the health system. In the first subsection, we describe all considered paths that a patient can take through the hospital. In the second subsection, we explain how the length stay of each patient is modelled.

### 4.1 Hospital patient path

COVID-19 patients can access the health system in different ways: a person can be diagnosed with COVID-19 illness in a primary healthcare facility, in the hospital emergency department, in a nursing home, after undergoing a SARS-CoV-2 test control (as a *Polymerase chain reaction (PCR)* test), etc. Depending on the severity of the illness the person is kept “isolated” at home or is admitted to the health care system as a COVID-19 patient, either in a hospital ward or in the ICU.

Treatment of COVID-19 patients requires dedicated resources, material, and personnel. We focus our study in the ward beds and ICU beds because they are scarce resources. Other necessary resources, as nurses and physicians, can be calculated from the number of required beds. The COVID-19 patient path through the hospital does not differ from other hospital patients. Fig. 1 shows the patient flow through the health system, highlighting the transitions between the hospital ward and the ICU. Patients can be admitted either into the ICU or into the hospital ward. A patient admitted to a hospital ward can be transferred to the ICU due to a health worsening. The discharge of the hospital ward is because of death or health improvement. Patients are discharged from the ICU to a hospital ward either because of death or because of health improvement.

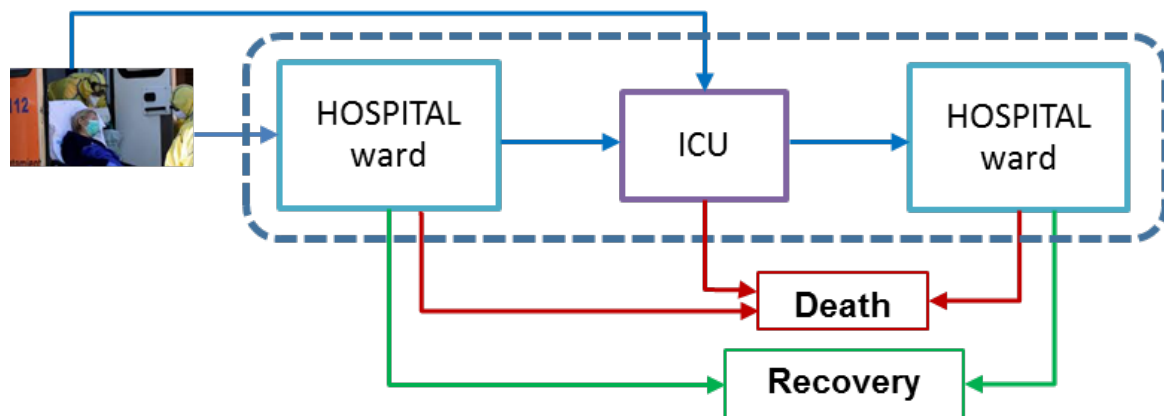


Fig. 1 Representation of the COVID-19 patient flow in the health system.

### 4.2 Stochastic modelling of hospital length of stay

The following variables allow the modeling of the length of stay in the hospital:

- $X_1$ , the length of stay of a patient in the hospital ward not needing the ICU
- $Y$ , the length of stay of a patient in the ICU.
- $W$ , time spent by a patient in the hospital ward before being admitted to the ICU.
- $X_2$ , the length of stay of a patient in the hospital ward after being discharged from the ICU.
- $X$ , the length of stay of a patient from its admission to a hospital ward or to the ICU until the discharge of the hospital facilities.

Besides, the following probabilities determine the patient-path through the healthcare facilities:

- $p_I$ , the probability that a patient is admitted to the ICU.

- $p_{I0}$ , the probability that a patient needing the ICU is admitted just after being detected positive. Then,  $1 - p_{I0}$  is the probability that a patient is admitted to the ICU after being admitted first to a hospital ward.
- $p_{WI}$ , the probability that a patient admitted to a hospital ward ultimately needs to be transferred to the ICU.  $p_{WI} = p_I(1 - p_{I0})/(1 - p_I p_{I0})$
- $p_{IW}$ , the probability that a patient is discharged from the ICU to a hospital ward. Then,  $1 - p_{IW}$  is the probability that a patient dies in the ICU.

The estimation of both the probability distributions and the set of probabilities depends on the available data, which, in turn, depends on the pandemic development stage. During the first stages of the outbreak, when no patient hospitalization data or very few exist the model uses the triangular distribution family. The triangular is a popular family to estimate the time to accomplish a task because it embodies the idea of ‘three-point estimation’ where subjective judgment is used to estimate a minimum, a ‘best guess’, and a maximum value of the variable [67]. In the case of variable  $X_1$ , these three values correspond to the minimum, most probable, and maximum times of hospitalization in a ward for patients not needing ICU, which are provided by expert opinion (hospital managers and the medical staff). Experts can rely on values reported in the literature corresponding to countries where the pandemic spread earlier (for example the China and the Italian situation are described in [68–71]).

As the pandemic progresses and there are more data on hospitalized COVID-19 patients, they can be used to estimate the probability distributions parameters. One main feature of these data is their high level of censorship; that is, only a small percentage of patients that have been admitted so far to the hospital ward and the ICU has been discharged after few weeks of the outbreak. This fact motivates the need for re-estimating the distribution fitting daily by adding the new data collected. The use of probability plots facilitates the selection of the parametric probability distribution family that best fits the data. The parameters of the selected family are estimated by fitting the updated health electronic recorded data on admission and discharge dates by the maximum likelihood method. Weibull and Lognormal distribution families have proved to be good probability models for the length of stay related variables, as we show in Section 7.

At the beginning of a new pandemic, there is insufficient knowledge of the illness and no effective treatments exist, as it has happened with the COVID-19 outbreak. As the medical and biological research progresses, new drugs and therapeutic protocols are discovered, which benefit the care of patients, affecting their length of stay in both hospital ward and ICUs. This observation reinforces the need for gathering every new data about patient admission and discharge to use them to update the estimation of the distribution parameters and bifurcation probabilities.

## 5 The discrete event simulation model

In this section, the mathematical modelling of the hospital dynamics is presented. First, we define a DES model to simulate the arrivals and the length of stay of patients. Then, a methodology is described that allows the simulation to start considering different levels of known information.

### 5.1 Entities, processes, events, and flowchart

DES model is defined by the set of state variables, which provide at any time a complete description of the simulated system, and the set of events, which modify the value of state variables. The simulation model represents the patient flow through the different ways of hospitalization, that is, the part inside the box with dashed lines in Fig. 1. In this subsection, we propose the state variables that describe the health system and a set of events grouped into three different categories.

Only one set of variables is considered in this model, which is composed of two variables,  $H = (H_W, H_I)$ . These variables describe the number of patients hospitalized in hospital wards ( $H_W$ )

and the number of patients in the ICU ( $H_I$ ). Observe that the number of total patients admitted in the hospital ( $N$ ) due to COVID-19 at time  $t$  is the sum of these two state variables ( $N = H_W + H_I$ ).

There are three different types of events that modify through time the value of state variables. They have been classified according to how the value of  $N$  changes, if there is an increase, a decrease, or if its value does not change after the event occurs. The first set of events  $E_A$  are associated with the patient's arrival times classified in hospital arrivals and ICU arrivals. In this group, we only consider external arrivals, i.e., patients who are not in the hospital and directly arrive after being detected with the disease. These arrivals occur as we mentioned in Section 3.2 distinguishing with a percentage what type of patients they are.

The second category of events  $E_B$  produces a decrease in the state variables. They are related to the end of the patient hospitalization time. In wards, we consider both discharges for improvement and death, but in ICU only death cases are included here. In this way, these events only reduce the values of  $H_W$  or  $H_I$ , and therefore the value of  $N$  decreases. In Section 4.2 it is explained how the length of stay is assigned to each patient, and therefore, the discharge times.

The third set of events  $E_C$  is associated with the transfer of patients between the hospital wards and the ICU, which can be bidirectional (see Section 4.2). On the one hand, a patient who is in the hospital may worsen his health of status and have to be transferred to the ICU. On the other hand, ICU patients who improve and are discharged do not go directly home but are first treated in wards. Observe that in both cases, only the values of  $H_W$  and  $H_I$  are affected, but the value of  $N$  remains constant since with these actions neither patients enter nor leave the system. Fig. 2 outlines the simulation model of the health system.

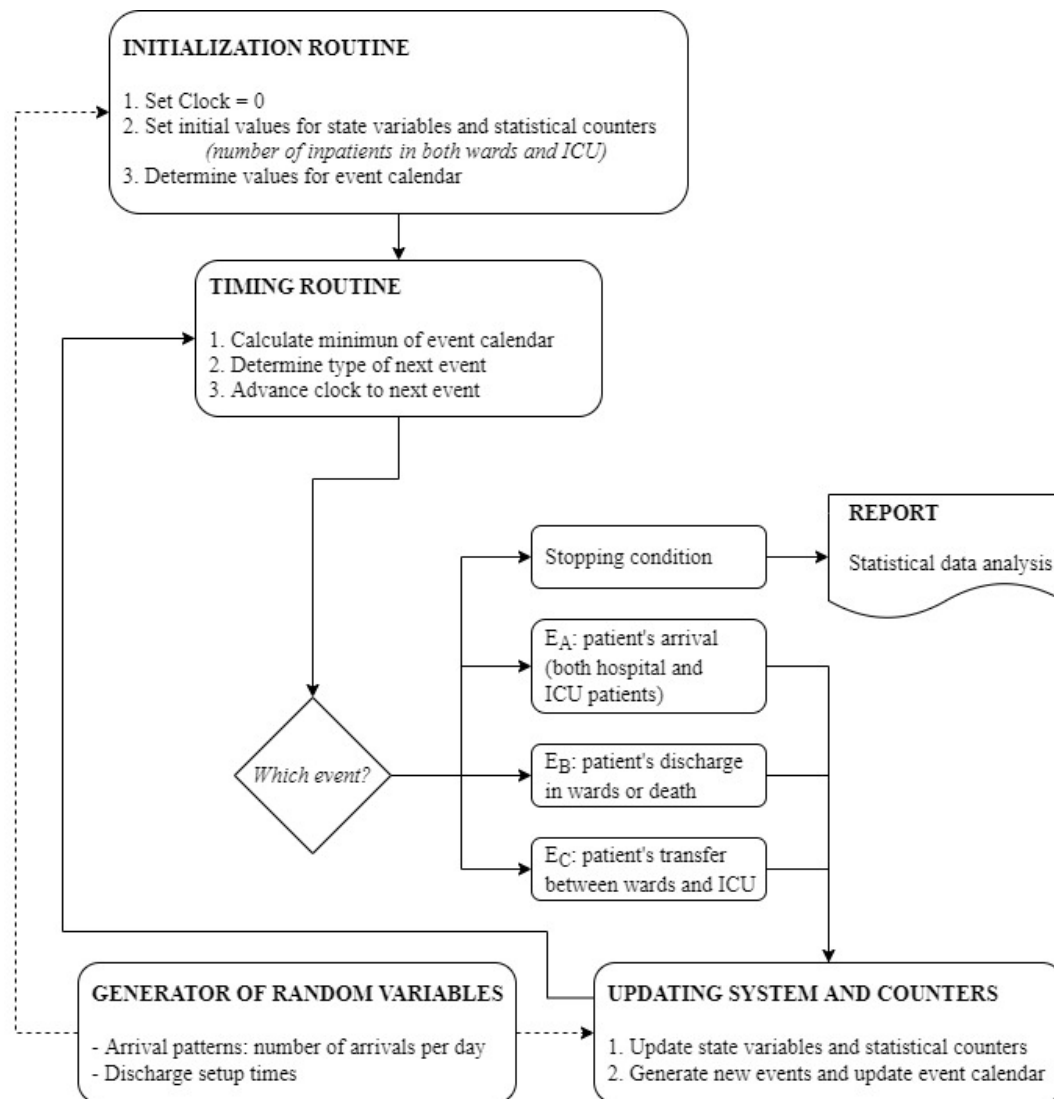


Fig. 2 Flow diagram of the health system simulation model. Three types of events are considered (patient's arrival, patient's discharge, and patient's transfer).

## 5.2 How to get the simulation started?

The purpose of the simulation model is the short-term necessary resources prediction. Therefore, the precision of the predictions strongly depends on the model accuracy in both representing the health system's initial state and the initial time evolution of the already occupied resources. Usually, this last aspect of the mathematical modeling is not such important when the simulation is intended to investigate the behavior of systems in the long term, in its stationary state, which is usually independent of the system's initial state. However, when the simulation is used as a prediction tool to support tactical decisions, the initial state of the simulation model and the system dynamic at the initial moments are the main factors to determine the health system state in the near future. The number of patients in the hospital wards and in the ICU, as well as the available information about the time since their admission, define the initial state of the health-system simulation model.

The simulation clock is set to zero at the time of the EHR-file's last update. Before that time, it is the past that can be reproduced by the simulator by reading the records in the EHR file, and ahead, it is the future that needs to be simulated. The transition from the past to the future is done at the moment the simulation starts, which requires to initialize the event calendar [67]. There are three types of events, admission to hospital, transferring between hospital facilities (ward and ICU) and discharge from hospital. The simulation of the event type "arrival of a new patient" was

explained in Section 3.2. Now we expose how to simulate the other two types of events to those patients already admitted in the hospital at time zero.

The simulation of these events depends on the available information about the patients. Basically, the main difference is whether there is access to patient-level information or only aggregated information about daily admissions and discharges is available. Zero time for the simulation model is considered to be the end of the  $k$ th day of the pandemic. That is, the simulation model begins to simulate what will happen from day  $(k + 1)$ th, using the information collected during the first  $k$  days of the pandemic and considering patients admitted to the health system at the end of the  $k$ th day as the initial state. The following notation, together with the notation introduced in Section 4.2, is used in the subsequent analysis:

$H_{Wk}$ : the number of hospitalized patients in hospital wards at the end of the day  $k$ .

$H_{Ik}$ : the number of admitted patients at the ICU at the end of the day  $k$ .

$I_j$  and  $O_j$  number of new patients admitted, and number of patients discharged the day  $j$ , respectively.

They will refer to the hospital wards or to the ICU depending on the context of the analysis.

$u$ : an observation from a Uniform  $[0,1]$  random variable.

**A. Information at the patient level available: The admission date of each hospitalized patient is known.** Therefore, for each currently hospitalized patient the time  $s_i$  already spent at the hospital ward or the ICU is known.

**Patient admitted to the ICU:** The discharge date is calculated by sampling from the random variable  $Y$  conditioned to a stay longer than  $s_i$ , the number of days already spent at the hospital. Let  $y_i$  be a value sampled from the conditional distribution  $Y/Y > s_i$ , then the value  $y_i - s_i$  is the ICU discharge time simulated for patient  $i$ .

**Patient admitted to a hospital ward:** A patient occupying a ward bed can ultimately be discharged from the hospital or be transferred to the ICU. The probability that a patient hospitalized for  $s_i$  days is transferred in the future to the ICU, denoted by  $p_{ICU/s_i}$ , is calculated by using Bayes theorem:

$$\begin{aligned} p_{ICU/s_i} &= P(\text{patient in ward needs ICU} / X > s_i) \\ &= \frac{P(X > s_i / \text{patient in ward needs ICU}) P(\text{patient in ward needs ICU})}{P(X > s_i)} \\ &= \frac{(1 - F_W(s_i)) p_{WI}}{(1 - F_X(s_i))} \end{aligned}$$

Therefore, for each patient  $i$ , a random number  $u$  is used to classify each patient into one out of two categories: patient to be transferred to ICU, when  $u \leq p_{ICU/s_i}$ , or patient to be discharged from the hospital, when  $u > p_{ICU/s_i}$ .

In the case of a patient transfer to the ICU, the event transfer time is simulated by sampling a value  $w_i$  from the conditional distribution  $W/W > s_i$ . The value  $w_i - s_i$  is the simulated transfer time from hospital ward to the ICU. Similarly, the discharge time from the hospital ward is simulated by sampling a value  $x_i$  from the conditional distribution  $X_1/X_1 > s_i$ , then the value  $x_i - s_i$  is the simulated discharge time.

**B. The series of new admitted patients  $I_j, j = 1, \dots, k$  and the number  $H_{Wk}$  of hospitalized patients are known at the moment of running the simulation model.** However, individual arrival dates are not known for the admitted patients and the method developed for the case A can not be applied.

Let  $p_{ij}$  be the probability that a patient admitted the day  $i$  is discharged the day  $j$ , and  $q_{ij}$  the probability that a patient admitted the day  $i$  is discharged later than day  $j$ . The end of day  $k$  is the present time at which the simulation model is run. The probability that a patient arriving the day  $i$  is still hospitalized is  $q_{ik}$ . Every past day  $i$ ,  $i = 1, \dots, k$ ,  $I_i$  new patients are admitted, which account for  $\sum_{i=1}^k I_i$  patients of which  $H_{Wk}$  are still hospitalized. The method sketched in Fig. 3 selects  $H_{Wk}$  patients out of  $\sum_{i=1}^k I_i$  by a sampling process without replacement that uses the set of probabilities  $\{q_{ik}; i = 1, \dots, k\}$ .

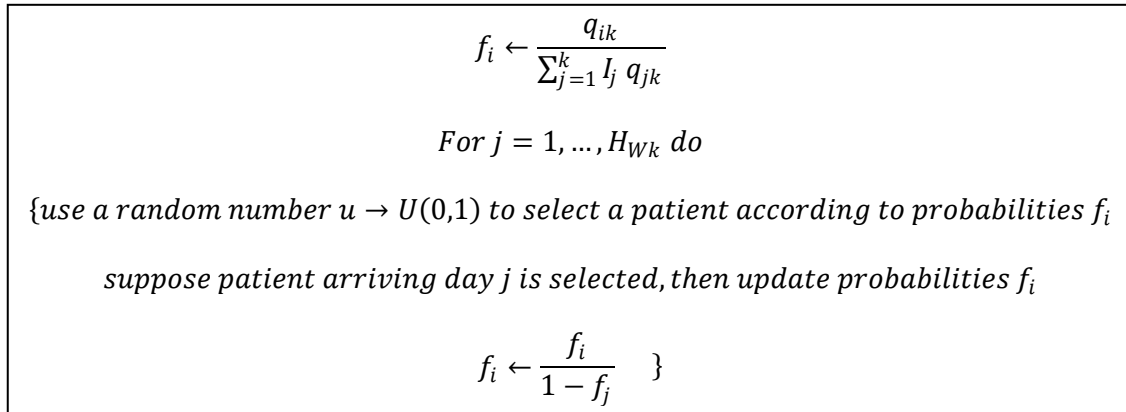


Fig. 3 The algorithm implemented to select which patients are hospitalized at the beginning of the simulation. For the ICU the algorithm is the same but referring all the parameters to the ICU (arrivals, bed occupation...).

The application of this method results in the simulation of the admission date of each admitted patient. Therefore, the simulation of the discharge and transfer events can be performed according to the method of case A.

The probabilities  $q_{ik}$  are calculated for patients admitted in the ICU from the probability distribution of variable  $Y$ , and for patients admitted to a hospital ward from the probability distribution of variable  $X$ .

- C. For each day  $j$ ,  $j = 1, \dots, k$ , the number  $I_j$  of new admitted patients and the number  $O_j$  of discharged patients are known.** We use this information (see Fig. 4) to estimate for each patient the probability of being one of the  $H_{Wk}$  admitted patients at the end of day  $k$  and the day of admission  $i$ . Once these probabilities are calculated, the procedures exposed in cases A and B can be applied.

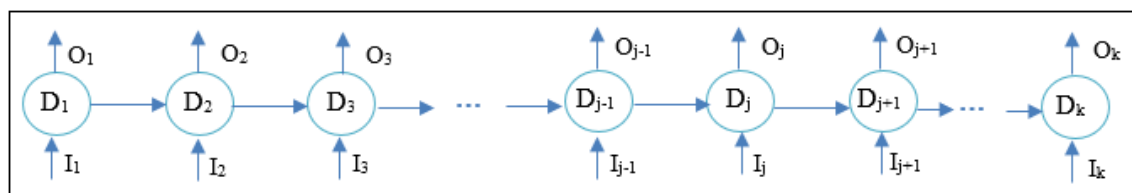


Fig. 4 Diagram that illustrates the number of new admitted patients ( $I_j$ ) and the number of discharged patients ( $O_j$ ) per day ( $D_j$ ).

Using, as before, the notation  $p_{ij}$  for the probability of a patient being admitted the day  $i$  and discharged the day  $j$ , and  $q_{ij}$  for the probability of a patient being admitted day  $i$  and discharged later than day  $j$ , then the expected number of discharge patients the day  $j$  is  $\sum_{i=1}^j p_{ij} I_i$  and the observed number  $O_j$ . Let  $g_j$  be the factor verifying  $g_j \sum_{i=1}^j p_{ij} I_i = O_j$ , so that  $n_{ij} = g_j p_{ij} I_i$  is the updated expected number of patients arriving the day  $i$  that are discharged the day  $j$ .

Similar calculation is performed for the number of admitted patients at the end of day  $k$ :  $\sum_{i=1}^k q_{ik} I_i$  is the expected number of admitted patients at the end of day  $k$  that arrived the day  $i$

and  $H_{Wk}$  the actual number of admitted patients. Let  $g_k^*$  be the factor verifying  $g_k^* \sum_{i=1}^k q_{ik} I_i = H_{Wk}$ , and therefore  $n_{ik}^* = g_k^* q_{ik} I_i$  is the updated expected number of patients arriving the day  $i$  that still are hospitalized at the end of day  $k$ .

The estimated new expected values  $n_{ij}, j = i, \dots, k$  and  $n_{ik}^*$  are used to rescale the probabilities  $q_{ik}$ : let  $\gamma_i$  the factor satisfying that  $\gamma_i (\sum_{j=i}^k n_{ij} + n_{ik}^*) = I_i$ , that is,  $\gamma_i (\sum_{j=i}^k g_j p_{ij} I_i + g_k^* q_{ik} I_i) = I_i$  and then,  $\gamma_i (\sum_{j=i}^k g_j p_{ij} + g_k^* q_{ik}) = 1$ . Therefore,  $\gamma_i g_k^* q_{ik}$  is the updated estimation of the probability that a patient arriving the day  $i$  is still hospitalized at the end of day  $k$ .

**D. Only the number  $H_{Wk}$  of hospitalized patients is known.** This case happens when the forecasting of daily hospitalized patients is performed from the series of aggregated positive cases in the local area,  $c(t)$ , and a hospitalization factor  $f_h$  is applied. In this situation the arrival date of the patients currently admitted at the hospital (end of day  $k$ ) is simulated by first simulating the arrival of patients each day before ( $i = 1, \dots, k$ ) and then by applying the method of paragraph B. The number of hospitalized patients each day  $i = 1, \dots, k$  is simulated by sampling values from a binomial distribution with parameters  $c(i) - c(i - 1)$ , the total number of new positive cases in day  $i$ , and  $f_h$ , the probability that a positive case needs to be hospitalized.

Once a discharge time or a transfer time between ward and ICU has been simulated for each hospitalized patient, the event calendar records those times, together with the arrival time of the next COVID-19 patient. The fitted Gompertz curve provides the number of patients arriving each day. These arrivals can be uniformly distributed during the next 24 hours or following a non-stationary pattern when, for example, a significant decrease of arrivals occurs at night. The clock of the simulation is advanced from zero to the minimum of the times recorded in the event calendar vector.

## 6 The simulator

The simulator developed for this work is presented in this section. The types of inputs required to start the simulator are described first, depending on the level of information known. Subsequently, the results obtained with this tool are shown graphically.

### 6.1 Input data

The developed simulator has a lot of flexibility to simulate a large number of scenarios and in different conditions. This allows medical staff to adapt the model to their particular hospital while testing different results by modifying some parameters. There are two types of input data to the simulator, depending on whether they are input parameters that are configured on the simulator screens before predictions are made or if they are external files to the program that the simulator needs to read to perform the simulation.

#### Simulator parameters

The simulator allows the user to modify the factors mentioned in Section 3.2 ( $f_g$  and  $f_h$ ). As explained in this section, in the early stages of the pandemic, it can be difficult to fit the curve of hospitalized patients directly due to the scarcity of data. In that phase, the cases of a larger region are fitted, and the predictions are scaled with both factors. These factors can be introduced following the experts' criteria (overall  $f_h$ ) or based on the proportion of the population when scaling cases of positives from a larger region to a smaller one. When fitting the hospitalized patient curve, the factors mentioned must be 1.

Another parameter to introduce is the probability that a patient who arrives at the hospital needs special care in the ICU ( $p_{ICU}$ ). The percentage of such patients who are directly admitted is also

indicated ( $p_{ICU_0}$ ), distinguishing them from those who spent a few days in wards before ( $p_{ICU_1}$ ). In turn, it is possible to define the distribution that the length of stay of the COVID-19 patients follows (Weibull, lognormal ...). For each selected distribution, the simulator allows the user to configure its parameters.

Finally, the simulator allows the user to select one of the four types of simulation outlined in Section 5.2 (A, B, C, D), depending on the amount of information available to the user. He or she should consider the simulation type selected to provide the correct input files.

### External files

One of the input files required by the simulator is the accumulated historical series, either of positive cases or patients admitted to the hospital. One series or another will be used depending on the phase of the pandemic or the amount of information available. These series are used to obtain the best fitting Gompertz curve, which allows the patient arrivals pattern to be estimated.

The other input files are related to the type of simulation to be run. If full information is required (A), the simulator is fed by an input data file obtained from the hospital electronic health record (EHR) system. The information needed for each COVID-19 patient arriving at the health system is the dates of hospital admission and discharge and the dates of ICU admission and discharge. If a field is empty, then the associated event has not occurred. For example, an already discharged patient from hospital with empty fields in the ICU fields means that this patient did not need intensive healthcare. A patient will be admitted in the health system at the beginning of the simulation when they have an admission date, but their discharge date field is empty.

The input data file records the information of all COVID-19 patients from the first day of the outbreak to the present day. Therefore, in this modality, the simulator can reproduce the occupancy of the hospital and ICU from the beginning of the outbreak until the present, and at this moment it is known the number of patients in the health system (in all hospitalization modalities) and the date of admission of those patients. This information is used to start the simulation run as it is explained in Section 5.2.

Finally, the simulation could also be run without specifying all the disaggregated information of each patient in the health system, just by knowing the number of patients admitted in wards and in the ICU (D). Or knowing, in addition to this, both the number of daily admissions and patient discharges (C) or only the admissions (B). These series are provided as external input files. In these cases, the simulator is not able to reproduce the past and get a very accurate representation of the health system at time zero of the simulation clock, but it can be estimated.

## 6.2 Simulation output

During the simulation, the program generates arrivals and discharges of patients of both the hospital and the ICU. The system evolves and the number of necessary beds is recorded at any time. In this simulation model, there are two sources of randomness. On the one hand, the number of patients arriving at the system (hospital and ICU), and on the other, the length of stay of patients, according to the distributions mentioned in Section 4.1. In each program run, the configured scenario is simulated many times (thousands), and thus a distribution of the number of beds needed each day is obtained since, in each simulation, the occupancy curve varies. The simulator calculates percentiles, which are stored in an Excel file, specifically the 5<sup>th</sup> percentile (P5), the 50<sup>th</sup> percentile (P50), and the 95<sup>th</sup> percentile (P95). Also, these percentiles, which represent confidence bands, are plotted on a graph.

Fig. 5 shows the graphic outputs that are obtained with the program. On the left side, the necessary hospitalization beds are observed, while on the right side the same graph is shown but for the ICU occupation. The green-colored line represents the real evolution of the occupation, and the black dot indicates the Simulation Starting Point (SSP), that is, the moment from which it is simulated.



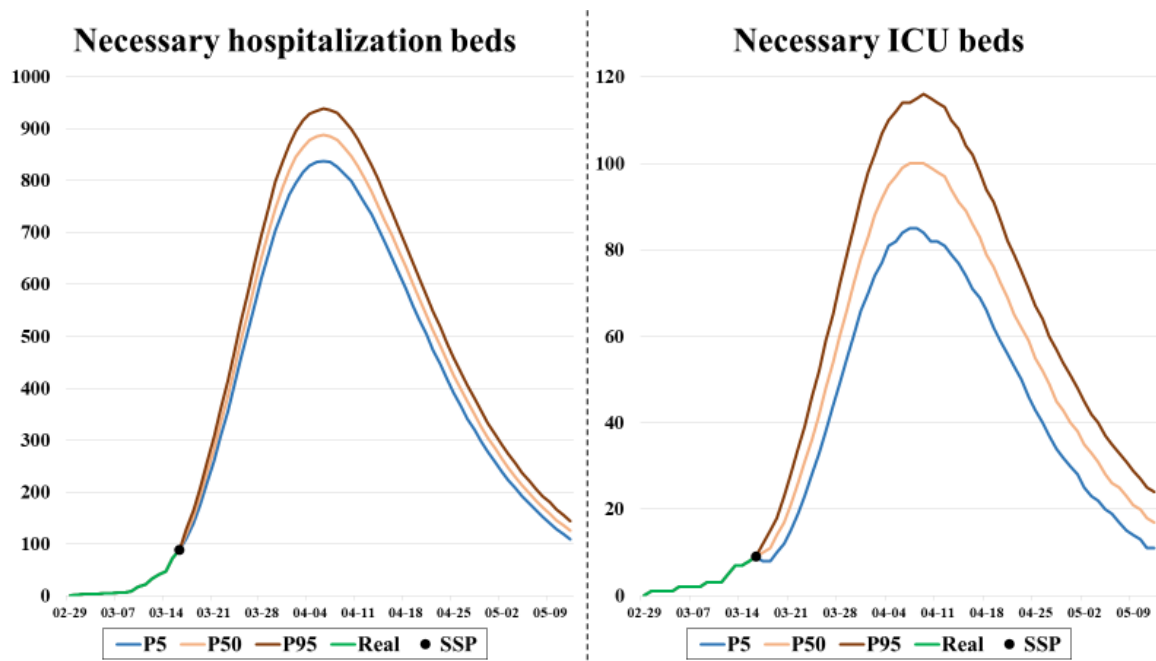


Fig. 5 Number of necessary hospitalization and ICU beds respectively for the following days. From the SSP, 3 lines corresponding to the 5<sup>th</sup>, 50<sup>th</sup>, and 95<sup>th</sup> percentiles are plotted.

## 7 Application to real cases

In this section, a case study is presented, in which the explained methodology has been applied to a region of Spain (Navarre). Firstly, we briefly introduce how the virus has affected Navarre globally, and then, some predictions of beds obtained with the program in different moments are presented and compared with the real bed occupancy curve.

### 7.1 Incidence of COVID-19 disease in Navarre (Spain)

Navarre is a small community of the north of Spain. Its population is around 650,000 inhabitants, with more than a half gathered around the capital (Pamplona) and its surroundings, an area that represents 5% of the total surface. With this distribution of population, the Health Services of Navarre has a main hospital located in Pamplona, with more than 1,000 available hospital beds, and two secondary hospitals in two of the most populated cities (Estella and Tudela). In total, the public health system in Navarra has 1,466 hospital beds and 45 ICU beds, with the possibility of increasing these quantities if necessary.

Navarre was among the five Spanish autonomous communities with a higher cumulative incidence rate of COVID-19 confirmed cases at the beginning of the pandemic, according to the data collected by the [Government of Navarre](#) [72]. By May 12, 7,752 COVID-19 cases (11.9 per 1,000 inhabitants) had been confirmed in Navarre, and among them, 1,704 had been admitted to public hospitals (2.6 per 1,000). Fig. 6 shows four graphs with daily and cumulative data about positive cases and patients who have been admitted to public hospitals in Navarre. It is observed that time trends peaked on March 25 and 27 and decreased after April 1 and 2.

Observe that as of April 28 there seems to be a rebound in the number of positive cases, (see red-colored values), which are not reflected in the number of hospitalized patients. This is because in the first stage of the pandemic the system's ability to perform tests was limited, and symptomatic patients were those who approached the system. However, as cases have decreased, the health system has reached out to patients to prevent further outbreaks, and far more cases of asymptomatic individuals have been detected than previously accounted for. This shows that

when adjusting the Gompertz model based on positive cases, there are parameters, such as the hospitalization factor ( $f_h$ ), that vary dynamically and must be updated.

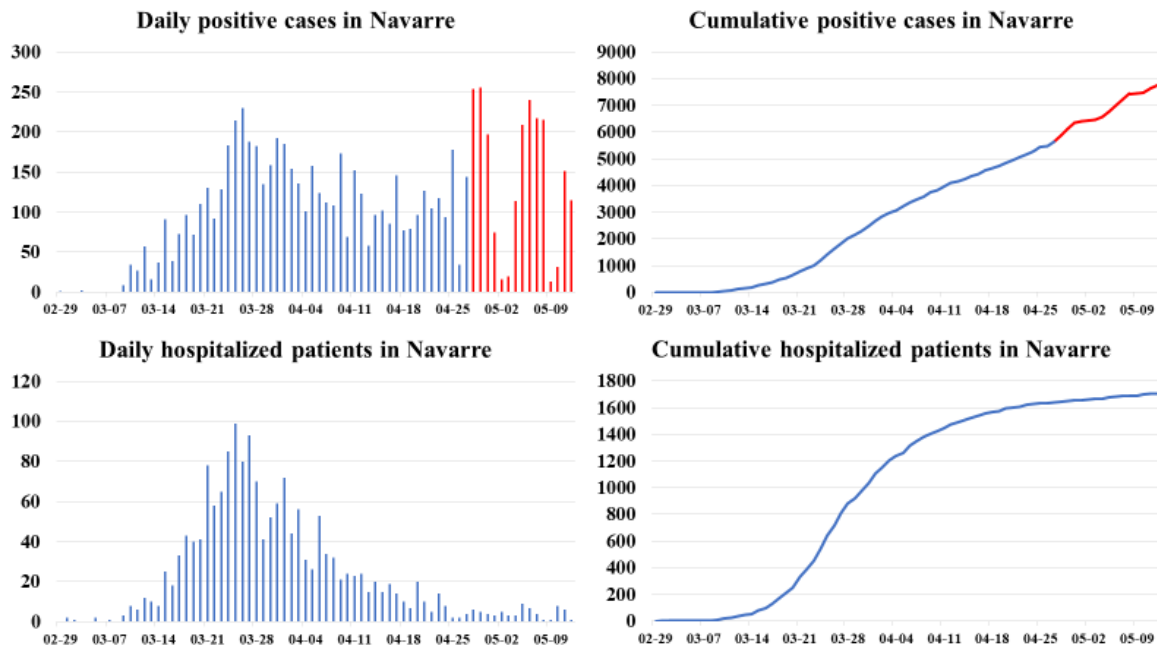


Fig. 6 Daily and cumulative data of positive cases and hospitalized patients recorded in Navarre from February 29 to May 12.

## 7.2 Predictions of both hospitalization and ICU bed occupancy

### 7.2.1 Fitting the cumulative positive cases

In the early stages of a pandemic, there is a scarcity of data on both the number of people infected with the virus and hospitalized patients, as well as in the length of stays observed in the hospital. Therefore, the involvement of experts is very important and necessary to estimate these parameters at these stages. In this way, a first approach can be made to simulate the evolution of the pandemic in the following days.

Navarre was one of the first communities in Spain to be infected, so at the beginning, the historical series of people infected with the virus was used to estimate the patient arrival pattern. Another possibility would be to use the cumulative positive cases of Spain and apply the appropriate geographic factor ( $f_g$ ). The problem that usually appears at the beginning of the outbreak is that the trend may lead to an exponential growth of the series. To avoid this, the total number of expected positive cases ( $A$  in equation (1)) is fixed, to obtain a realistic output from the Gompertz growth model fit.

Fig. 7 shows different results after applying the Gompertz growth model fit to the cumulative positive cases in Navarre. In the early stages of the pandemic (03-16), two curves can be seen, one of them obtained using the cumulative positive cases of Spain (03-16(Spain)). In addition, three more fits made on March 31 (03-31), April 15 (04-15), and April 30 (04-30) are exposed. Note that the last values in the series are not predicted very accurately and this is due to what was discussed in Section 7.1. These results demonstrate that accounting for positive cases has changed.

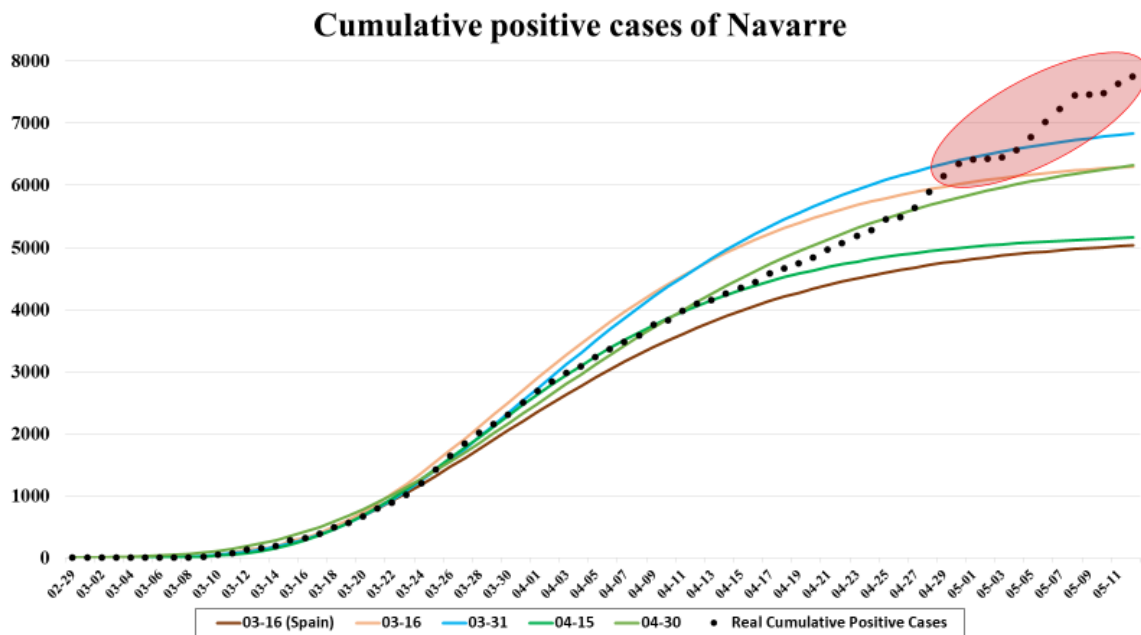


Fig. 7 Cumulative positive cases of Navarre from February 29 to May 12 and different fitted curves obtained from the Gompertz growth model.

The first two curves on March 16 were obtained after setting parameter  $A$ . For the case of Spain,  $A = 175,000$  was established according to the cumulative incidence observed in China and reported in the bibliography, and comparing the evolution with Italy, where the pandemic outbreaked a few weeks earlier. Then  $f_g$  scales the curve. The criterion for choosing its value was the rate of positive cases between Navarre and Spain at that time (3%) rather than the rate of populations between both regions (1.3%). In the case of the cumulative positive cases in Navarra,  $A = 6,500$  was set (1% of the total population of Navarra), since the population in this region is more concentrated than in the rest of Spain, and this makes the infection more aggressive.

At the initial stage of the outbreak when there is not enough data to estimate the length of stay of patients, a triangular distribution is used. Experts fix the minimum, maximum, and most probable time as 10, 18, and 13 days respectively. In addition, a hospitalization factor ( $f_g$ ) must be applied (0.35 for hospitalized patients and 0.04 for ICU patients), because the arrival pattern is based on positive cases. Fig. 8 shows both the prediction of the number of beds occupied in the hospitals and in the ICUs of Navarre simulated on March 16 for the next days. On the upper side, the predictions based on the cumulative positive cases of Navarre are observed, while those on the lower side are based on the cumulative positive cases of Spain. In each graph, the real evolution of bed occupancy in each area is shown in green.

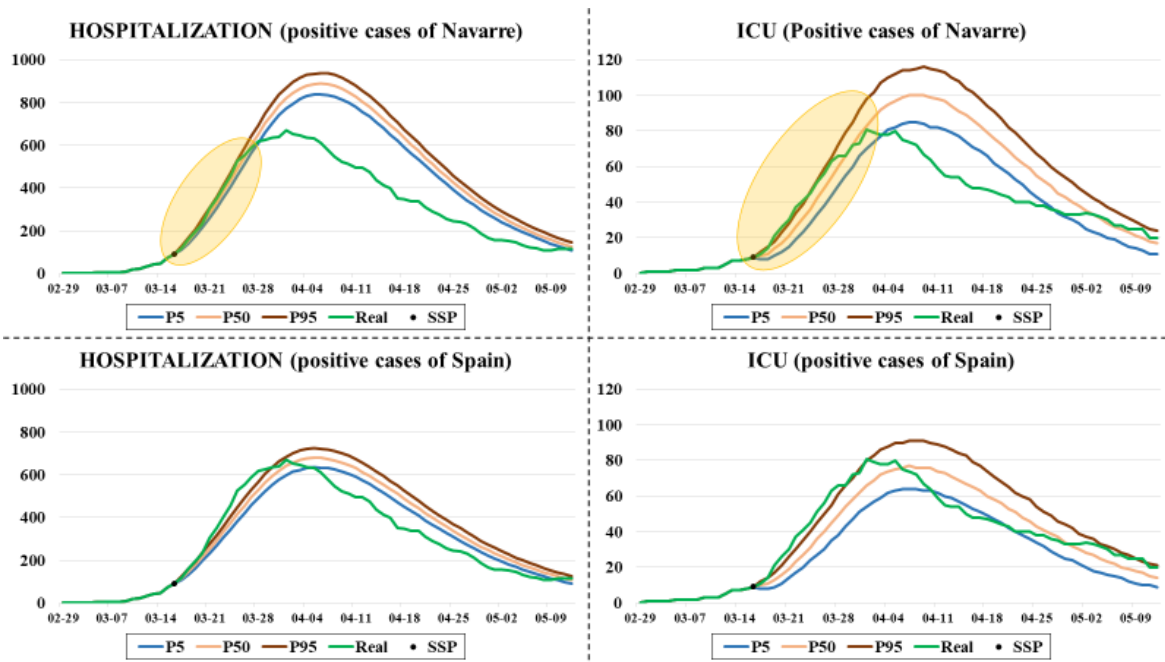


Fig. 8 Comparison between the predictions made on March 16 for the number of beds occupied in the hospitals and the ICUs of Navarre and the real occupancies. The cumulative positive cases of Navarre have been used for predictions in the two graphs at the top, while the cumulative positive cases of Spain in those at the low.

Note that the most important predictions for the medical staff are for the short-medium term (yellow-colored area in Fig. 8), and the results obtained are very similar to the real ones concerning both hospital and ICU occupancies. Fig. 9 shows a zoom of these parts of the graphs. The simulator's ability to obtain accurate predictions for the next 15 days, even in the early stages of the pandemic, is demonstrated here.

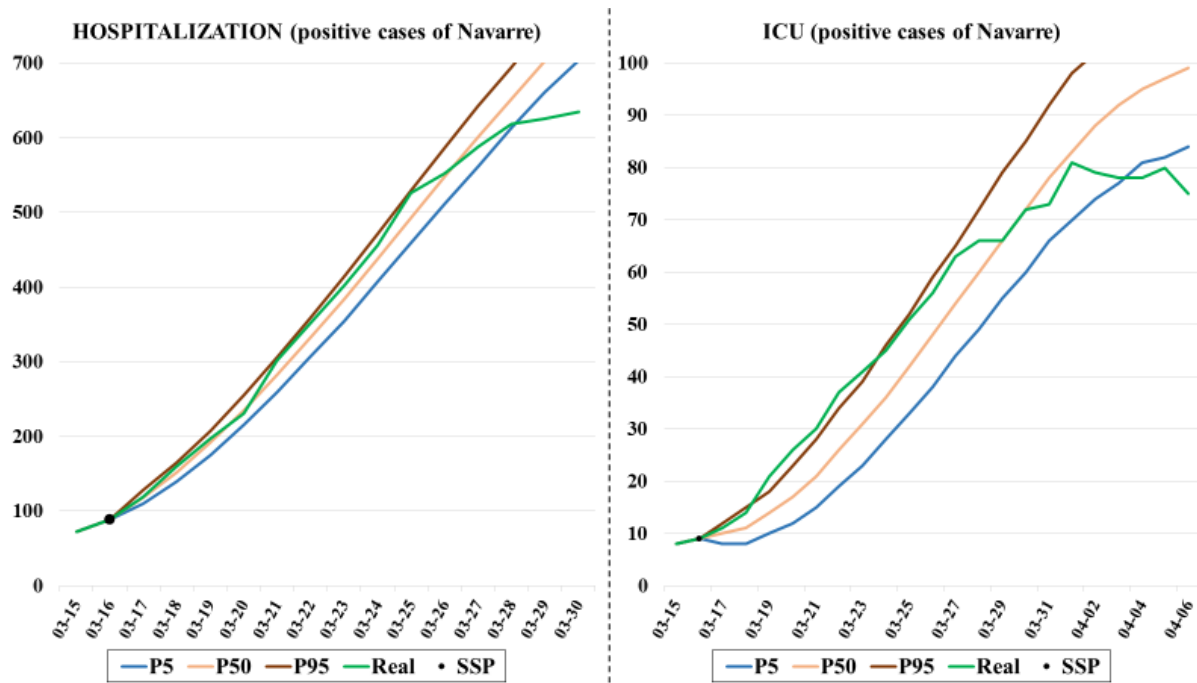


Fig. 9 Expanded part of the predictions (for both hospitalization and ICU) made on March 16 for the next 15 days using the cumulative positive cases of Navarre.

## 7.2.2 Fitting the cumulative hospitalized patients comparing levels of information

As the pandemic spreads, the amount of information increases, which makes it possible to improve the simulation model. Since March 26, the arrival pattern is calculated from the series of patients admitted to the hospital. In the first few days, it was required to fix the total number of expected hospitalized patients ( $A$ ), as with the positive cases, due to the curve fitted by the Gompertz growth model increases exponentially. According to experts, about 2,000 patients were expected to be admitted to public hospitals during the pandemic. Fig. 10 shows different results after applying the Gompertz growth model fit to the cumulative hospitalized patients in Navarre. It can be seen that the curve obtained on March 26 overestimate hospitalizations, so it was replaced in the simulation by that in which the maximum is set at 2,000 (03-26 (2,000)).

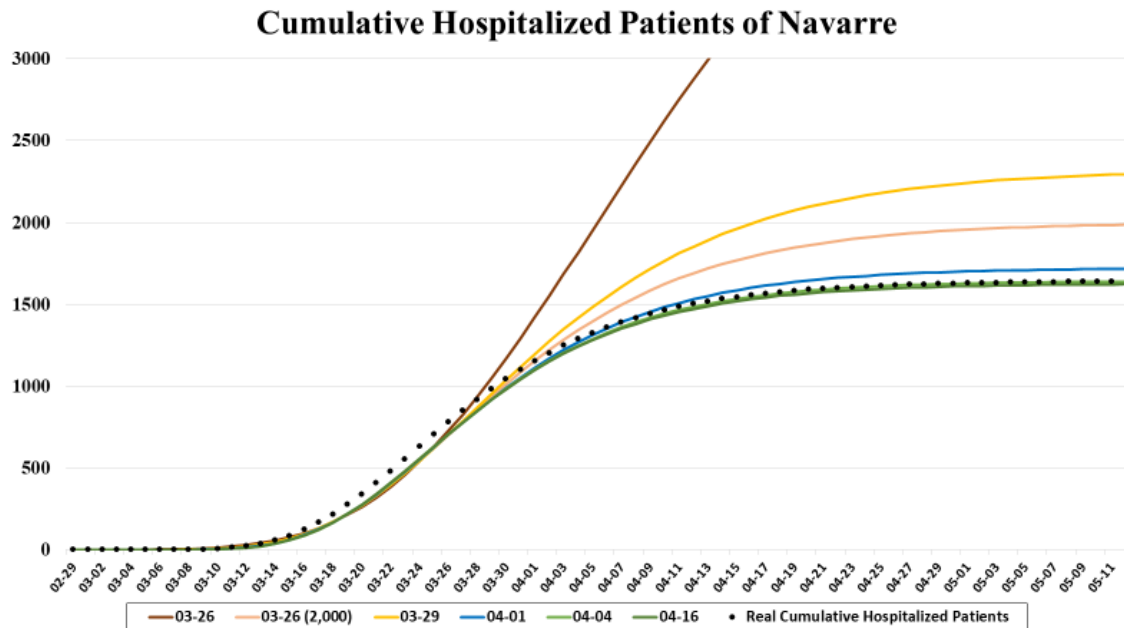


Fig. 10 Cumulative hospitalized patients of Navarre from February 29 to May 12 and different fitted curves obtained from the Gompertz growth model.

Regarding the length of stay of patients both in hospital and ICU, there were already enough data to estimate them every day, as it is explained in Section 4.2. At the beginning of the pandemic, the length of stay of patients showed a reasonable fit to a Weibull distribution, although, as time goes by and more data has been obtained, the lognormal distribution has shown to be a better model. Each time the data was analyzed, the parameters of the distribution were configured for the simulation. Table 4 lists those parameters used in the simulations at different moments of the pandemic, as well as some of the variables entered manually in the simulator defined in Section 4.2.

With all this information updated day by day, and the curves shown in Fig. 10, it is possible to simulate the evolution of the bed occupancy from a specific day. Fig. 11 shows two types of simulations carried out on March 26 and April 4. The difference between them lies in the amount of information used for the simulation. On the one hand, all the information was available, such as series of hospitalized patients and detailed information of the patients admitted (dates of admission and discharge of each patient, admission to the ICU, etc.). On the other hand, only the information about the series of hospital and ICU admissions, and the number of patients admitted at the beginning of the simulation ( $H_{Wk}$  y  $H_{Ik}$ ) were used. These assumptions correspond to scenarios A and B respectively, which are developed in Section 5.2.

Date	$X$ (days)	$\bar{X}$ (days)	$Y$ (days)	$\bar{Y}$ (days)	$Z$ (days)	$\bar{Z}$ (days)	$p_I$	$p_{I0}$	$p_{IW}$
03-26	Weibull (1.223, 18.470)	17.291	Weibull (1.450; 22.707)	20.589	Weibull (2.337, 3.802)	3.369	0.075	0.390	0.444
03-29	Weibull (1.204, 18.489)	17.376	Weibull (1.450; 24.663)	22.363	Weibull (2.424, 4.016)	3.560	0.046	0.310	0.308
04-01	Weibull (1.261, 15.421)	14.334	Weibull (1.450; 25.423)	25.423	Weibull (1.670; 4.410)	3.939	0.043	0.310	0.353
04-04	Weibull (1.030, 15.000)	14.821	Weibull (1.430; 28.677)	26.053	Weibull (1.691; 4.445)	3.967	0.042	0.272	0.36
04-10	Weibull (1.050; 14.914)	14.627	Weibull (1.500; 23.948)	21.619	Weibull (1.652; 4.436)	3.966	0.042	0.272	0.500
04-16	Weibull (1.000; 14.844)	14.844	Weibull (1.400; 24.304)	22.151	Weibull (1.646; 4.385)	3.921	0.040	0.309	0.619
04-26	Lognormal (2.275; 1.070)	17.241	Weibull (1.050; 26.006)	25.507	Weibull (1.464; 4.565)	4.134	0.042	0.331	0.628

Table 4. The parameters used in the simulation at different moments during the pandemic. The distributions of the different lengths of stay considered as well as the probabilities related to ICU patients are shown.

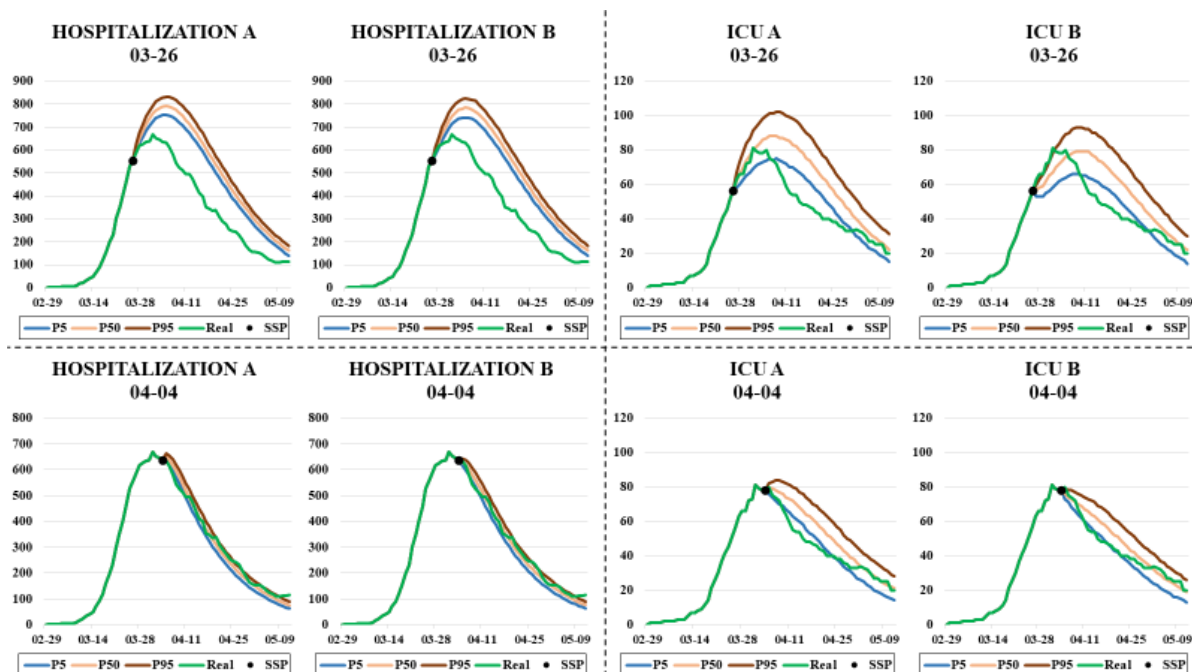


Fig. 11 Comparative graphs on the predictions made with different levels of information (scenarios A and B). Each graph shows the predictions made for the number of beds occupied in the hospitals and the ICUs of Navarre and the real occupancies. The four predictions at the top side were made on March 26, while those four at the low were made on April 4.

In view of the results, it can be seen that there is no great difference between the two types of simulation, which is very encouraging since the usefulness of the simulator is demonstrated even when all the disaggregated information on patients is not available. In addition to this, it is observed that the results are more accurate on the second date because more information is available and it is somewhat easier to predict the occupation of beds once the peak occupancy has been exceeded.

## 8 Conclusions

Healthcare systems are overburdened due to a large demand for healthcare services from COVID-19 patients that leads to strained ICU's capacity and overworked healthcare workers. Having accurate predictions of the resources needed for taking care of these patients are essential for planning in advance the necessary resources and reduce the pressure on the system and stress on the healthcare staff.

Normally, managers can create contingency plans to deploy sufficient resources to meet the increase in demand based on the predictions made about one week in advance. In this paper, we have developed a DES model for predicting the need for hospital resources, particularly ward and ICU beds. The simulation model is fed by the predictions of new hospitalizations made, directly or indirectly, through a PG model. The Gompertz growth model has been selected after analyzing the fitting and forecasting properties of four PG models: logistic, Richards, Stannard, and Gompertz. An ensemble of these models could improve predictions, but this research is out of the scope of this paper and it is left for future work. Besides, other factors influencing the consumption of resources are the age and the Adjusted Morbidity Group (AMG) whose inclusion in the stochastic models for the length of stay can provide more accurate predictions.

The structural simplicity of the simulation model makes it appropriate for general use, i.e., it can be adapted to estimate the bed needs in any geographic area. The growth model is simple enough to fit it to available data or, in the absence of it, to be estimated by educated guesses of experts.

It is worthy to mention the strength of simulation models in this context of uncertainty: their capability to run what-if scenarios that allow decision-makers to explore the consequences of different policy choices, like the location and number of additional healthcare resources needed for COVID-19 patients given the uncertainty in demand. The simulation model is data-driven, patients arrivals and length of stays can be estimated from data, but it has also the flexibility of allowing the simulation from the input determined by the user to explore additional scenarios.

From a technical and methodological point of view, a distinct feature of the simulation model is its focus on the transition period of the health system instead of the stationary state as it is usual in the simulation studies or transition periods but after regeneration points. This transition period reproduced by the simulator is unique as the outbreak evolves with no regeneration points. Therefore, the accurate representation of the initial health system state plays an important role. The simulation of the remaining length of stay of each patient already admitted in the hospital has shown to be a key point to project smoothly the dynamics of the health system and linking it (and mixing it) with the new dynamics obtained from the simulation of the arrivals and stays of the new incoming patients. However, the simulation of the remaining length of stay depends on the information level about the admitted patients. In this paper, we have considered four different levels of information that grades from the complete information to the patient level (knowing exactly the dates of admission and discharge) to the only knowledge of the number of admitted patients. The simulation model can work with any of these information levels, making it a tool of very general application.

This simulation paradigm is most suitable for the realistic representation of processes in health services, which makes it more credible and easier to understand by the managers that will have to rely on their results to make their decisions. The involvement in the development of the simulation model of the person in charge of the hospital system logistic person has been crucial to obtaining a user-centered simulator, which was daily feed with the new data gathered by the hospital administrative information system.

## APPENDICES

### A. Results of positive cases predictions for the following 5, 10, and 15 days, at 25%, 45%, and 65% of total cases detected.

25%

#	Country	Date	Total positive cases	Logistic	Gompertz	Richards	Stannard
1	USA	2020-04-12	529,951	2,233.0	554.9	<b>553.7</b>	<b>553.7</b>
2	Brazil	2020-05-16	218,223	1,497.5	1,099.9	<b>1,098.4</b>	1,100.4
3	Russia	2020-05-04	134,687	564.2	381.8	<b>357.2</b>	<b>357.2</b>
4	India	2020-05-16	85,940	601.4	<b>350.0</b>	<b>350.0</b>	<b>350.1</b>
5	UK	2020-04-12	78,991	320.6	249.7	<b>209.7</b>	<b>209.7</b>
6	Spain	2020-03-26	66,460	297.1	<b>147.8</b>	<b>147.8</b>	<b>147.9</b>
7	Italy	2020-03-24	63,927	269.7	<b>150.8</b>	156.7	156.7
8	Peru	2020-05-08	58,526	608.4	<b>385.0</b>	<b>385.1</b>	<b>385.1</b>
9	Iran	2020-04-02	47,593	1,192.5	<b>969.6</b>	<b>969.6</b>	<b>969.7</b>
10	Germany	2020-03-28	48,582	<b>269.9</b>	315.8	287.4	287.4
11	Turkey	2020-04-11	47,029	639.9	<b>272.4</b>	<b>272.5</b>	<b>272.6</b>
12	Chile	2020-05-18	43,781	741.7	<b>732.2</b>	735.8	735.8
13	France	2020-03-30	40,174	163.2	<b>92.6</b>	<b>93.0</b>	94.6
14	Mexico	2020-05-06	40,186	207.8	130.5	<b>128.3</b>	<b>128.3</b>
15	Pakistan	2020-05-15	37,218	304.3	<b>223.5</b>	<b>223.1</b>	223.5
16	Saudi Arabia	2020-05-07	31,938	<b>194.3</b>	281.4	<b>195.2</b>	<b>195.2</b>
17	Canada	2020-04-14	25,663	97.0	83.7	<b>57.3</b>	<b>57.3</b>
18	Bangladesh	2020-05-18	22,268	270.9	<b>134.5</b>	<b>134.5</b>	<b>134.5</b>
19	China	2020-02-05	24,320	164.3	<b>114.1</b>	<b>114.1</b>	<b>114.1</b>
20	Qatar	2020-05-09	20,201	147.6	206.0	<b>139.2</b>	<b>139.2</b>

Table 5. MAE calculated for the fit of each model at 25% of total cases detected.

#	Country	L 5	G 5	R 5	S 5	L 10	G 10	R 10	S 10	L 15	G 15	R 15	S 15
1	USA	6.840%	2.234%	<b>2.090%</b>	<b>2.091%</b>	15.771%	1.732%	<b>1.558%</b>	<b>1.559%</b>	26.066%	<b>2.473%</b>	2.646%	2.645%
2	Brazil	5.154%	<b>1.565%</b>	<b>1.596%</b>	<b>1.567%</b>	12.003%	<b>3.820%</b>	<b>3.899%</b>	<b>3.827%</b>	20.594%	<b>4.942%</b>	5.080%	<b>4.956%</b>
3	Russia	14.866%	<b>7.418%</b>	8.145%	8.145%	27.301%	<b>12.043%</b>	13.585%	13.585%	40.983%	<b>15.807%</b>	18.525%	18.525%
4	India	6.100%	<b>2.053%</b>	<b>2.055%</b>	<b>2.056%</b>	13.805%	<b>5.244%</b>	<b>5.247%</b>	<b>5.248%</b>	24.003%	<b>8.760%</b>	<b>8.765%</b>	<b>8.768%</b>
5	UK	8.099%	1.903%	<b>1.089%</b>	<b>1.089%</b>	16.215%	3.978%	<b>1.343%</b>	<b>1.343%</b>	26.289%	6.670%	<b>2.387%</b>	<b>2.387%</b>
6	Spain	<b>3.502%</b>	10.295%	10.283%	10.284%	<b>9.853%</b>	31.939%	31.899%	31.898%	<b>19.188%</b>	65.792%	65.700%	65.691%
7	Italy	<b>1.865%</b>	10.643%	7.881%	7.881%	<b>5.990%</b>	24.930%	17.171%	17.171%	<b>13.465%</b>	44.691%	28.294%	28.295%
8	Peru	<b>1.806%</b>	3.482%	3.480%	3.480%	5.851%	<b>5.624%</b>	<b>5.619%</b>	<b>5.618%</b>	12.624%	<b>8.101%</b>	<b>8.092%</b>	<b>8.089%</b>
9	Iran	2.635%	<b>1.778%</b>	<b>1.777%</b>	<b>1.777%</b>	<b>2.157%</b>	7.030%	7.027%	7.028%	<b>3.204%</b>	16.013%	16.006%	16.007%
10	Germany	10.915%	3.847%	<b>2.470%</b>	<b>2.470%</b>	30.272%	<b>3.495%</b>	12.258%	12.258%	50.039%	<b>4.967%</b>	23.066%	23.066%
11	Turkey	15.075%	<b>4.756%</b>	<b>4.761%</b>	<b>4.764%</b>	29.363%	<b>6.250%</b>	<b>6.261%</b>	<b>6.268%</b>	45.514%	<b>7.633%</b>	<b>7.652%</b>	<b>7.666%</b>
12	Chile	5.868%	8.082%	<b>5.183%</b>	<b>5.183%</b>	10.310%	13.948%	<b>8.985%</b>	9.253%	12.393%	18.030%	<b>9.969%</b>	20.650%
13	France	9.720%	4.424%	<b>4.396%</b>	<b>4.302%</b>	<b>15.968%</b>	17.090%	16.910%	16.324%	<b>26.496%</b>	37.173%	36.735%	35.309%
14	Mexico	6.396%	<b>0.895%</b>	2.187%	2.187%	13.397%	<b>2.498%</b>	5.044%	5.044%	23.518%	<b>5.093%</b>	9.494%	9.494%
15	Pakistan	<b>0.886%</b>	4.092%	4.063%	4.089%	<b>1.968%</b>	5.827%	5.779%	5.823%	<b>3.048%</b>	10.311%	10.236%	10.303%
16	Saudi Arabia	5.616%	<b>2.142%</b>	6.813%	6.814%	12.669%	<b>2.608%</b>	14.917%	14.917%	24.360%	<b>2.470%</b>	27.770%	27.770%
17	Canada	11.915%	<b>3.656%</b>	6.871%	6.871%	24.599%	<b>10.482%</b>	16.277%	16.277%	38.866%	<b>19.003%</b>	27.603%	27.603%
18	Bangladesh	13.722%	<b>8.747%</b>	<b>8.748%</b>	<b>8.750%</b>	25.468%	<b>15.608%</b>	<b>15.610%</b>	<b>15.613%</b>	40.500%	<b>24.279%</b>	<b>24.283%</b>	<b>24.288%</b>
19	China	13.144%	<b>5.681%</b>	<b>5.669%</b>	<b>5.663%</b>	44.477%	<b>12.518%</b>	<b>12.528%</b>	<b>12.534%</b>	75.439%	<b>11.828%</b>	<b>11.796%</b>	<b>11.775%</b>
20	Qatar	8.920%	<b>1.879%</b>	11.294%	11.294%	19.084%	<b>5.653%</b>	23.360%	23.360%	31.546%	<b>9.641%</b>	37.870%	37.870%

Table 6. Normalized MAEs obtained for each prediction and model at 25% of total cases detected.



40%

#	Country	Date	Total positive cases	Logistic	Gompertz	Richards	Stannard
1	USA	2020-04-23	842,629	6,372.8	<b>1,421.9</b>	<b>1,422.8</b>	1,433.3
2	Brazil	2020-05-24	347,398	2,085.6	1,422.1	<b>1,413.6</b>	1,422.7
3	Russia	2020-05-12	221,344	1,316.9	<b>521.2</b>	<b>521.3</b>	<b>521.5</b>
4	India	2020-05-25	138,845	882.2	<b>480.3</b>	<b>480.4</b>	<b>480.5</b>
5	UK	2020-04-20	120,067	617.4	320.3	<b>257.3</b>	<b>257.3</b>
6	Spain	2020-03-31	104,267	351.9	438.7	<b>343.9</b>	<b>343.9</b>
7	Italy	2020-03-30	97,689	374.7	343.5	<b>205.8</b>	<b>205.8</b>
8	Peru	2020-05-18	92,273	759.3	483.2	<b>482.1</b>	<b>482.1</b>
9	Iran	2020-04-16	76,389	1,071.3	1,201.7	<b>1,043.0</b>	<b>1,043.0</b>
10	Germany	2020-04-03	79,696	486.0	339.2	<b>338.2</b>	<b>338.2</b>
11	Turkey	2020-04-17	74,193	795.4	<b>345.4</b>	<b>345.5</b>	<b>345.7</b>
12	Chile	2020-05-26	73,997	<b>854.4</b>	858.2	<b>854.4</b>	<b>854.4</b>
13	France	2020-04-04	64,338	260.8	<b>200.4</b>	219.2	219.2
14	Mexico	2020-05-15	62,527	438.7	<b>151.8</b>	<b>151.8</b>	<b>151.9</b>
15	Pakistan	2020-05-27	59,151	335.9	<b>276.6</b>	285.4	285.4
16	Saudi Arabia	2020-05-17	52,016	360.6	273.4	<b>256.6</b>	<b>256.6</b>
17	Canada	2020-04-23	40,179	405.1	<b>175.7</b>	<b>175.7</b>	175.7
18	Bangladesh	2020-05-26	35,585	434.7	<b>270.4</b>	<b>270.4</b>	270.5
19	China	2020-02-08	34,625	185.9	130.2	<b>126.6</b>	<b>126.6</b>
20	Qatar	2020-05-18	32,604	274.1	<b>216.8</b>	<b>216.8</b>	<b>216.8</b>

Table 7. MAE calculated for the fit of each model at 40% of total cases detected.

#	Country	L 5	G 5	R 5	S 5	L 10	G 10	R 10	S 10	L 15	G 15	R 15	S 15
1	USA	9.702%	<b>3.501%</b>	<b>3.503%</b>	<b>3.525%</b>	15.921%	<b>6.527%</b>	<b>6.531%</b>	<b>6.567%</b>	22.590%	<b>10.112%</b>	<b>10.117%</b>	<b>10.167%</b>
2	Brazil	<b>1.165%</b>	3.163%	2.984%	3.161%	<b>2.858%</b>	3.924%	3.617%	3.919%	6.209%	6.041%	<b>5.547%</b>	6.032%
3	Russia	3.420%	<b>2.303%</b>	<b>2.301%</b>	<b>2.299%</b>	<b>5.698%</b>	6.093%	6.090%	6.085%	<b>9.245%</b>	10.728%	10.721%	10.714%
4	India	4.669%	<b>1.331%</b>	<b>1.334%</b>	<b>1.333%</b>	9.212%	<b>2.274%</b>	<b>2.278%</b>	<b>2.277%</b>	16.059%	<b>3.695%</b>	<b>3.701%</b>	<b>3.700%</b>
5	UK	6.330%	1.005%	<b>0.710%</b>	<b>0.711%</b>	11.831%	<b>0.913%</b>	2.210%	2.210%	19.098%	<b>1.746%</b>	5.344%	5.344%
6	Spain	6.632%	<b>5.630%</b>	5.905%	5.906%	13.015%	11.927%	<b>11.725%</b>	<b>11.725%</b>	20.529%	<b>18.121%</b>	18.794%	18.795%
7	Italy	3.211%	6.294%	<b>1.096%</b>	<b>1.096%</b>	8.347%	9.938%	<b>1.571%</b>	<b>1.571%</b>	14.945%	13.137%	<b>4.334%</b>	<b>4.334%</b>
8	Peru	6.317%	<b>0.817%</b>	1.035%	1.035%	11.975%	<b>1.669%</b>	2.145%	2.146%	22.254%	<b>6.073%</b>	6.872%	6.873%
9	Iran	<b>0.582%</b>	6.097%	1.628%	1.628%	<b>1.757%</b>	8.452%	3.338%	3.337%	<b>3.372%</b>	10.640%	5.413%	5.413%
10	Germany	11.428%	<b>2.083%</b>	2.224%	2.220%	19.849%	<b>2.095%</b>	2.390%	2.381%	27.383%	<b>1.755%</b>	<b>1.802%</b>	<b>1.801%</b>
11	Turkey	7.319%	<b>1.436%</b>	<b>1.432%</b>	<b>1.429%</b>	13.559%	<b>3.472%</b>	<b>3.463%</b>	<b>3.457%</b>	19.637%	<b>6.500%</b>	<b>6.486%</b>	<b>6.476%</b>
12	Chile	<b>1.645%</b>	<b>1.636%</b>	<b>1.645%</b>	<b>1.645%</b>	5.823%	<b>5.639%</b>	5.823%	5.817%	14.435%	14.059%	14.432%	<b>7.991%</b>
13	France	<b>2.158%</b>	12.580%	6.941%	6.940%	<b>3.202%</b>	22.970%	9.453%	9.451%	<b>7.545%</b>	35.881%	10.903%	10.900%
14	Mexico	7.023%	<b>2.499%</b>	<b>2.501%</b>	<b>2.503%</b>	12.126%	<b>3.558%</b>	<b>3.563%</b>	<b>3.566%</b>	19.255%	<b>5.141%</b>	<b>5.149%</b>	<b>5.154%</b>
15	Pakistan	3.195%	1.697%	<b>1.489%</b>	<b>1.489%</b>	11.675%	<b>4.830%</b>	7.837%	7.837%	23.909%	<b>11.558%</b>	17.595%	17.595%
16	Saudi Arabia	9.921%	<b>4.418%</b>	5.822%	5.822%	15.880%	<b>5.955%</b>	8.499%	8.499%	20.963%	<b>5.429%</b>	9.512%	9.512%
17	Canada	12.339%	<b>6.475%</b>	<b>6.477%</b>	<b>6.478%</b>	19.244%	<b>9.745%</b>	<b>9.749%</b>	<b>9.751%</b>	27.788%	<b>14.521%</b>	<b>14.526%</b>	<b>14.529%</b>
18	Bangladesh	4.289%	<b>1.024%</b>	<b>1.025%</b>	<b>1.025%</b>	11.261%	<b>3.951%</b>	<b>3.955%</b>	<b>3.954%</b>	20.791%	<b>7.544%</b>	<b>7.550%</b>	<b>7.550%</b>
19	China	11.511%	10.730%	<b>10.225%</b>	<b>10.225%</b>	37.831%	<b>12.943%</b>	16.756%	16.756%	52.069%	<b>10.778%</b>	15.505%	15.506%
20	Qatar	5.565%	<b>1.188%</b>	<b>1.204%</b>	<b>1.211%</b>	9.960%	<b>1.348%</b>	<b>1.379%</b>	<b>1.392%</b>	16.349%	<b>1.774%</b>	<b>1.827%</b>	<b>1.849%</b>

Table 8. Normalized MAEs obtained for each prediction and model at 40% of total cases detected.

65%

#	Country	Date	Total positive cases	Logistic	Gompertz	Richards	Stannard
1	USA	2020-04-13	1,369,964	20,036.0	<b>7,343.8</b>	7,345.7	7,346.8
2	Brazil	2020-05-31	584,016	2,654.8	<b>1,857.5</b>	1,908.6	1,908.4
3	Russia	2020-05-19	344,481	1,469.0	1,097.0	<b>846.3</b>	<b>846.3</b>
4	India	2020-04-13	216,919	1,113.0	<b>511.4</b>	<b>511.6</b>	<b>511.7</b>
5	UK	2020-04-30	194,990	2,136.8	<b>593.2</b>	<b>593.2</b>	<b>593.3</b>
6	Spain	2020-06-04	163,472	917.7	639.9	<b>382.2</b>	<b>382.2</b>
7	Italy	2020-04-15	156,363	1,420.8	585.3	<b>526.1</b>	<b>526.1</b>
8	Peru	2020-05-27	155,671	1,757.1	<b>792.6</b>	<b>792.9</b>	<b>793.1</b>
9	Iran	2020-06-06	122,492	3,124.1	<b>1,912.6</b>	<b>1,912.7</b>	<b>1,912.8</b>
10	Germany	2020-05-31	123,016	907.9	419.2	<b>400.3</b>	<b>400.3</b>
11	Turkey	2020-05-08	117,589	1,222.6	528.2	<b>463.3</b>	<b>463.3</b>
12	Chile	2020-06-05	113,628	1,011.6	1,038.2	<b>844.2</b>	<b>844.2</b>
13	France	2020-02-13	103,573	495.0	525.4	<b>350.4</b>	<b>350.4</b>
14	Mexico	2020-05-30	101,238	746.9	<b>237.1</b>	<b>237.2</b>	<b>237.3</b>
15	Pakistan	2020-04-13	93,983	918.8	<b>565.8</b>	<b>565.8</b>	<b>566.0</b>
16	Saudi Arabia	2020-05-31	83,384	683.4	453.3	<b>430.7</b>	<b>430.7</b>
17	Canada	2020-05-19	64,922	836.0	<b>346.9</b>	<b>346.9</b>	<b>347.0</b>
18	Bangladesh	2020-04-13	57,563	534.4	<b>310.3</b>	<b>310.4</b>	<b>310.4</b>
19	China	2020-04-30	59,865	841.5	<b>556.6</b>	<b>556.6</b>	<b>556.7</b>
20	Qatar	2020-06-04	52,907	352.0	199.7	<b>199.6</b>	<b>199.6</b>

Table 9. MAE calculated for the fit of each model at 65% of total cases detected.

#	Country	L 5	G 5	R 5	S 5	L 10	G 10	R 10	S 10	L 15	G 15	R 15	S 15
1	USA	7.840%	<b>3.355%</b>	<b>3.356%</b>	<b>3.356%</b>	10.822%	<b>4.910%</b>	<b>4.911%</b>	<b>4.911%</b>	13.993%	<b>6.696%</b>	<b>6.698%</b>	<b>6.698%</b>
2	Brazil*	2.720%	<b>1.445%</b>	<b>1.396%</b>	<b>1.397%</b>	3.898%	3.799%	<b>3.153%</b>	<b>3.156%</b>	4.007%	4.631%	<b>3.863%</b>	<b>3.867%</b>
3	Russia	4.285%	1.705%	<b>0.901%</b>	<b>0.902%</b>	7.670%	<b>1.598%</b>	2.524%	2.524%	11.790%	<b>1.254%</b>	4.822%	4.822%
4	India*	4.293%	<b>0.859%</b>	<b>0.861%</b>	<b>0.861%</b>	7.276%	<b>0.607%</b>	<b>0.609%</b>	<b>0.610%</b>	8.062%	<b>0.575%</b>	<b>0.577%</b>	<b>0.578%</b>
5	UK	9.641%	<b>4.441%</b>	<b>4.443%</b>	<b>4.444%</b>	12.914%	<b>5.616%</b>	<b>5.618%</b>	<b>5.620%</b>	16.050%	<b>6.681%</b>	<b>6.684%</b>	<b>6.687%</b>
6	Spain	6.365%	<b>0.354%</b>	2.477%	2.477%	9.902%	<b>0.485%</b>	4.430%	4.430%	13.199%	<b>1.001%</b>	6.451%	6.451%
7	Italy	7.789%	<b>2.131%</b>	3.037%	3.037%	11.523%	<b>3.678%</b>	4.993%	4.993%	15.241%	<b>5.390%</b>	7.116%	7.116%
8	Peru	10.445%	<b>6.119%</b>	<b>6.121%</b>	<b>6.122%</b>	12.525%	<b>5.208%</b>	<b>5.210%</b>	<b>5.212%</b>	15.847%	<b>4.811%</b>	<b>4.815%</b>	<b>4.818%</b>
9	Iran	14.856%	<b>10.700%</b>	<b>10.702%</b>	<b>10.702%</b>	18.765%	<b>13.960%</b>	<b>13.962%</b>	<b>13.963%</b>	23.212%	<b>17.807%</b>	<b>17.809%</b>	<b>17.810%</b>
10	Germany	4.053%	2.415%	<b>1.864%</b>	<b>1.864%</b>	7.360%	2.354%	<b>1.473%</b>	<b>1.473%</b>	10.611%	2.080%	<b>1.118%</b>	<b>1.118%</b>
11	Turkey	4.858%	1.079%	<b>0.484%</b>	<b>0.484%</b>	7.419%	<b>0.972%</b>	1.333%	1.333%	10.097%	<b>0.817%</b>	2.504%	2.504%
12	Chile*	<b>6.750%</b>	9.179%	7.516%	7.517%	<b>11.535%</b>	16.538%	17.398%	17.400%	<b>12.420%</b>	18.133%	19.826%	19.827%
13	France	5.234%	<b>1.690%</b>	1.813%	1.813%	7.530%	2.840%	<b>2.534%</b>	<b>2.534%</b>	9.905%	3.784%	<b>3.530%</b>	<b>3.530%</b>
14	Mexico	5.207%	<b>1.082%</b>	<b>1.084%</b>	<b>1.085%</b>	8.908%	<b>1.584%</b>	<b>1.587%</b>	<b>1.588%</b>	11.785%	<b>2.536%</b>	<b>2.537%</b>	<b>2.538%</b>
15	Pakistan**	11.997%	<b>9.348%</b>	<b>9.351%</b>	<b>9.351%</b>	18.651%	<b>14.243%</b>	<b>14.246%</b>	<b>14.247%</b>	-	-	-	-
16	Saudi Arabia	<b>0.598%</b>	4.724%	3.395%	3.395%	4.029%	4.289%	<b>2.612%</b>	<b>2.612%</b>	10.167%	<b>3.939%</b>	4.547%	4.547%
17	Canada	5.661%	<b>0.861%</b>	<b>0.862%</b>	<b>0.863%</b>	7.801%	<b>0.674%</b>	<b>0.676%</b>	<b>0.677%</b>	10.379%	<b>0.791%</b>	<b>0.793%</b>	<b>0.794%</b>
18	Bangladesh	5.516%	<b>2.097%</b>	<b>2.099%</b>	<b>2.099%</b>	9.975%	<b>3.061%</b>	<b>3.064%</b>	<b>3.065%</b>	-	-	-	-
19	China	14.044%	<b>5.492%</b>	<b>5.495%</b>	<b>5.495%</b>	14.165%	<b>8.987%</b>	<b>8.976%</b>	<b>8.973%</b>	<b>14.158%</b>	16.806%	16.781%	16.776%
20	Qatar	4.045%	<b>0.780%</b>	<b>0.777%</b>	<b>0.775%</b>	5.683%	<b>2.302%</b>	<b>2.293%</b>	<b>2.286%</b>	7.766%	<b>4.616%</b>	<b>4.600%</b>	<b>4.589%</b>

Table 10. Normalized MAEs obtained for each prediction and model at 65% of total cases detected.

## REFERENCES

1. European Centre for Disease Prevention and Control (ECDC). COVID-19 2020. <https://www.ecdc.europa.eu/en/novel-coronavirus-china>. Accessed 30 Jul 2020
2. European Commission (EC). COVID-19. [https://ec.europa.eu/health/coronavirus\\_en](https://ec.europa.eu/health/coronavirus_en). Accessed 30 Jul 2020
3. World Health Organization (WHO). Coronavirus disease (COVID-19) outbreak. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Accessed 30 Jul 2020
4. Brailsford SC, Harper PR, Patel B, Pitt M (2009) An analysis of the academic literature on simulation and modelling in health care. *J Simul* 3:130–140. <https://doi.org/10.1057/jos.2009.10>
5. Günal MM, Pidd M (2010) Discrete event simulation for performance modelling in health care: A review of the literature. *J Simul* 4:42–51. <https://doi.org/10.1057/jos.2009.25>
6. Katsaliaki K, Mustafee N (2011) Applications of simulation within the healthcare context. *J Oper Res Soc* 62:1431–1451. <https://doi.org/10.1057/jors.2011.20>
7. Mielczarek B, Uziąłko-Mydlikowska J (2012) Application of computer simulation modeling in the health care sector: A survey. *Simulation* 88:197–216. <https://doi.org/10.1177/0037549710387802>
8. Kolker A (2009) Process modeling of ICU patient flow: Effect of daily load leveling of elective surgeries on ICU diversion. *J Med Syst* 33:27–40. <https://doi.org/10.1007/s10916-008-9161-9>
9. Shahani AK, Ridley SA, Nielsen MS (2008) Modelling patient flows as an aid to decision making for critical care capacities and organisation. *Anaesthesia* 63:1074–1080. <https://doi.org/10.1111/j.1365-2044.2008.05577.x>
10. Ridge JC, Jones SK, Nielsen MS, Shahani AK (1998) Capacity planning for intensive care units. *Eur J Oper Res* 105:346–355. [https://doi.org/10.1016/S0377-2217\(97\)00240-3](https://doi.org/10.1016/S0377-2217(97)00240-3)
11. Rodrigues F, Zaric GS, Stanford DA (2018) Discrete event simulation model for planning Level 2 “step-down” bed needs using NEMS. *Oper Res Heal Care* 17:42–54. <https://doi.org/10.1016/j.orhc.2017.10.001>
12. Zhu Z, Hen BH, Teow KL (2012) Estimating ICU bed capacity using discrete event simulation. *Int J Health Care Qual Assur* 25:134–144. <https://doi.org/10.1108/09526861211198290>
13. Mallor F, Azcárate C, Barado J (2016) Control problems and management policies in health systems: application to intensive care units. *Flex Serv Manuf J* 28:62–89. <https://doi.org/10.1007/s10696-014-9209-8>
14. Erhard M, Schoenfelder J, Fügenger A, Brunner JO (2018) State of the art in physician scheduling. *Eur J Oper Res* 265:1–18. <https://doi.org/10.1016/j.ejor.2017.06.037>
15. Azcarate C, Esparza L, Mallor F (2020) The problem of the last bed: Contextualization and a new simulation framework for analyzing physician decisions. *Omega* 96:102120. <https://doi.org/10.1016/j.omega.2019.102120>
16. Currie CSM, Fowler JW, Kotiadis K, et al (2020) How simulation modelling can help reduce the impact of COVID-19. *J Simul* 14:83–97. <https://doi.org/10.1080/17477778.2020.1751570>
17. Grassly NC, Fraser C (2008) Mathematical models of infectious disease transmission. *Nat Rev Microbiol* 6:477–487. <https://doi.org/10.1038/nrmicro1845>
18. Brauer F, Castillo-Chavez C (2012) *Mathematical models in population biology and epidemiology*, 2nd ed. Springer, New York
19. Anderson RM, May RM (1991) *Infectious diseases of humans: dynamics and control*.

- Oxford University Press, New York
20. Diekmann O, Heesterbeek JAP (2000) *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. John Wiley, Chichester
  21. Hethcote HW (2000) The mathematics of infectious diseases. *SIAM Rev* 42:599–653. <https://doi.org/10.1137/S0036144500371907>
  22. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A* 115:700–721. <https://doi.org/10.1098/rspa.1927.0118>
  23. Lin Q, Zhao S, Gao D, et al (2020) A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int J Infect Dis* 93:211–216. <https://doi.org/10.1016/j.ijid.2020.02.058>
  24. Anastassopoulou C, Russo L, Tsakris A, Siettos C (2020) Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* 15:e0230405. <https://doi.org/10.1371/journal.pone.0230405>
  25. Casella F (2020) Can the COVID-19 epidemic be controlled on the basis of daily test reports? Prepr <https://arxiv.org/abs/200306967>
  26. Wu JT, Leung K, Bushman M, et al (2020) Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* 26:506–510. <https://doi.org/10.1038/s41591-020-0822-7>
  27. Giordano G, Blanchini F, Bruno R, et al (2020) Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Springer US
  28. Hellewell J, Abbott S, Gimma A, et al (2020) Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Heal* 8:e488–e496. [https://doi.org/10.1016/S2214-109X\(20\)30074-7](https://doi.org/10.1016/S2214-109X(20)30074-7)
  29. Kucharski AJ, Russell TW, Diamond C, et al (2020) Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis* 20:553–558. [https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/10.1016/S1473-3099(20)30144-4)
  30. Li CH, Tsai CC, Yang SY (2014) Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks. *Commun Nonlinear Sci Numer Simul* 19:1042–1054. <https://doi.org/10.1016/j.cnsns.2013.08.033>
  31. Li GH, Zhang YX (2017) Dynamic behaviors of a modified SIR model in epidemic diseases using nonlinear incidence and recovery rates. *PLoS One* 12:e0175789. <https://doi.org/10.1371/journal.pone.0175789>
  32. Magal P, Seydi O, Webb G (2016) Final size of an epidemic for a two-group SIR model. *SIAM J Appl Math* 76:2042–2059. <https://doi.org/10.1137/16M1065392>
  33. Xia CY, Sun SW, Liu ZX, et al (2009) Epidemics of sirs model with nonuniform transmission on scale-free networks. *Int J Mod Phys B* 23:2203–2213. <https://doi.org/10.1142/S021797920905211X>
  34. Oliveira Zardin AM da S, de Oliveira CAL, de Oliveira SN, et al (2019) Growth curves by Gompertz nonlinear regression model for male and female Nile tilapias from different genetic groups. *Aquaculture* 511:734243. <https://doi.org/10.1016/j.aquaculture.2019.734243>
  35. Oribe-Pérez IA, Velázquez-Abunader I, Poot-López GR (2020) Age and multi-model growth estimation of white grunt, *Haemulon plumieri*, in the southern Gulf of Mexico from otolith macrostructure analysis. *Reg Stud Mar Sci* 34:101069. <https://doi.org/10.1016/j.rsma.2020.101069>
  36. Sun K, Ren JS, Bai T, et al (2020) A dynamic growth model of *Ulva prolifera*: Application in quantifying the biomass of green tides in the Yellow Sea, China. *Ecol Modell* 428:109072. <https://doi.org/10.1016/j.ecolmodel.2020.109072>
  37. Horimoto K, Hofuku I, Oshima K (1997) Population dynamics of HIV positive by

- application of a locally modified Gompertz curve: A case study for Japan. *Nonlinear Anal Theory, Methods Appl* 30:985–993. [https://doi.org/10.1016/S0362-546X\(97\)88553-7](https://doi.org/10.1016/S0362-546X(97)88553-7)
38. Roberts MG, Saha AK (1999) The asymptotic behaviour of a logistic epidemic model with stochastic disease transmission. *Appl Math Lett* 12:37–41. [https://doi.org/10.1016/S0893-9659\(98\)00123-2](https://doi.org/10.1016/S0893-9659(98)00123-2)
  39. Viboud C, Simonsen L, Chowell G (2016) A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics* 15:27–37. <https://doi.org/10.1016/j.epidem.2016.01.002>
  40. Ghazvini K, Mansouri S, Shakeri M, et al (2019) Prediction of Tuberculosis Using a Logistic Regression Model. *Rev Clin Med* 6:108–112. <https://doi.org/10.22038/RCM.2019.41778.1284>
  41. Liu W, Tang S, Xiao Y (2015) Model Selection and Evaluation Based on Emerging Infectious Disease Data Sets including A/H1N1 and Ebola. *Comput Math Methods Med* 2015:207105. <https://doi.org/10.1155/2015/207105>
  42. Shen CY (2020) Logistic growth modelling of COVID-19 proliferation in China and its international implications. *Int J Infect Dis* 96:582–589. <https://doi.org/10.1016/j.ijid.2020.04.085>
  43. Malavika B, Marimuthu S, Joy M, et al (2020) Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clin Epidemiol Glob Heal*. <https://doi.org/10.1016/j.cegh.2020.06.006>
  44. Sánchez-Villegas P, Codina AD (2020) Modelos predictivos de la epidemia de COVID-19 en España con curvas de Gompertz. *Gac Sanit*. <https://doi.org/10.1016/j.gaceta.2020.05.005>
  45. Cássaro FAM, Pires LF (2020) Can we predict the occurrence of COVID-19 cases? Considerations using a simple model of growth. *Sci Total Environ* 728:138834. <https://doi.org/10.1016/j.scitotenv.2020.138834>
  46. Panovska-Griffiths J (2020) Can mathematical modelling solve the current Covid-19 crisis? *BMC Public Health* 20:551. <https://doi.org/10.1186/s12889-020-08671-z>
  47. He H, Hu C, Xiong N, et al (2020) How to transform a general hospital into an “infectious disease hospital” during the epidemic of COVID-19. *Crit Care* 24:145. <https://doi.org/10.1186/s13054-020-02864-z>
  48. Poston JT, Patel BK, Davis AM (2020) Management of critically ill adults with COVID-19. *JAMA - J Am Med Assoc* 323:1839–1841. <https://doi.org/10.1001/jama.2020.4914>
  49. Steinberg E, Balakrishna A, Habboushe J, et al (2020) Calculated decisions: COVID-19 calculators during extreme resource-limited situations. *Emerg Med Pract* 22:CD1–CD5
  50. Manca D, Caldiroli D, Storti E (2020) A simplified math approach to predict ICU beds and mortality rate for hospital emergency planning under Covid-19 pandemic. *Comput Chem Eng* 140:106945. <https://doi.org/10.1016/j.compchemeng.2020.106945>
  51. Almagooshi S (2015) Simulation Modelling in Healthcare: Challenges and Trends. *Procedia Manuf* 3:301–307. <https://doi.org/10.1016/j.promfg.2015.07.155>
  52. Persson J (2017) A review of the design and development processes of simulation for training in healthcare – A technology-centered versus a human-centered perspective. *Appl Ergon* 58:314–326. <https://doi.org/10.1016/j.apergo.2016.07.007>
  53. Palominos E, Levett-Jones T, Power T, Martinez-Maldonado R (2019) Healthcare students’ perceptions and experiences of making errors in simulation: An integrative review. *Nurse Educ Today* 77:32–39. <https://doi.org/10.1016/j.nedt.2019.02.013>
  54. Zhang J, Zhou L, Yang Y, et al (2020) Therapeutic and triage strategies for 2019 novel coronavirus disease in fever clinics. *Lancet Respir Med* 8:E11–E12. [https://doi.org/10.1016/S2213-2600\(20\)30071-0](https://doi.org/10.1016/S2213-2600(20)30071-0)

55. Christian MD, Sprung CL, King MA, et al (2014) Triage: Care of the critically ill and injured during pandemics and disasters: CHEST consensus statement. *Chest* 146:e61S-e74S. <https://doi.org/10.1378/chest.14-0736>
56. Cheung WK, Myburgh J, Seppelt IM, et al (2012) A multicentre evaluation of two intensive care unit triage protocols for use in an influenza pandemic. *Med J Aust* 197:178–181. <https://doi.org/10.5694/mja11.10926>
57. Garcia-Vicuña D, Mallor F, Esparza L, Mateo P (2019) A Management Flight Simulator of an Intensive Care Unit. In: *In Proceedings of the 2019 Winter Simulation Conference*. pp 1196–1207
58. Utley M, Pagel C, Peters MJ, et al (2011) Does triage to critical care during a pandemic necessarily result in more survivors? *Crit Care Med* 39:179–183. <https://doi.org/10.1097/CCM.0b013e3181fa3c3b>
59. Gompertz B (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc London B Biol Sci* 182:513–585. <https://doi.org/10.1098/rstl.1825.0026>
60. Richards FJ (1959) A flexible growth function for empirical use. *J Exp Bot* 10:290–301. <https://doi.org/10.1093/jxb/10.2.290>
61. Stannard CJ, Williams AP, Gibbs PA (1985) Temperature/growth relationships for psychrotrophic food-spoilage bacteria. *Food Microbiol* 2:115–122. [https://doi.org/10.1016/S0740-0020\(85\)80004-6](https://doi.org/10.1016/S0740-0020(85)80004-6)
62. Ricker WE (1979) Growth rates and models. *Fish Physiol* 8:677–743. [https://doi.org/10.1016/S1546-5098\(08\)60034-5](https://doi.org/10.1016/S1546-5098(08)60034-5)
63. Worldometer - real time world statistics. <https://www.worldometers.info/coronavirus/>. Accessed 15 Jun 2020
64. Python Software Foundation. Python Language Reference, version 3.7.3. <http://www.python.org>
65. Petzoldt, T. 2019. growthrates: Estimate Growth Rates from Experimental Data. R package version 0.8.1. <https://cran.r-project.org/package=growthrates>
66. Zwietering MH, Jongenburger I, Rombouts FM, van 't Riet K (1990) Modeling of the bacterial growth curve. *Appl Environ Microbiol* 56:1875–1881
67. Law AM (2014) *Simulation modeling and analysis*, 5th ed. McGraw-Hill, New York
68. Grasselli G, Pesenti A, Cecconi M (2020) Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. *JAMA*. <https://doi.org/10.1001/jama.2020.4031>
69. Guan W, Ni Z, Hu Y, et al (2020) Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. <https://doi.org/10.1056/nejmoa2002032>
70. Young BE, Ong SWX, Kalimuddin S, et al (2020) Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. *JAMA* 323:1488–1494. <https://doi.org/10.1001/jama.2020.3204>
71. Zhou F, Yu T, Du R, et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)
72. Government of Navarre. <https://gobiernoabierto.navarra.es>. Accessed 30 Jul 2020