

## Research paper

# Novel approach for Monte Carlo simulation of the new COVID-19 spread dynamics

Stavros Maltezos<sup>\*</sup>, Angelika Georgakopoulou

National Technical University of Athens, Physics Department, Greece

## ARTICLE INFO

## Keywords:

COVID-19

Epidemic

Monte Carlo simulation

SIR model

Effective reproductive number

## ABSTRACT

A Monte Carlo simulation in a novel approach is used for studying the problem of the outbreak and spread dynamics of the new COVID-19 pandemic in this work. In particular, our goal was to generate epidemiological data based on natural mechanism of transmission of this disease assuming random interactions of a large-finite number of individuals in very short distance ranges. In the simulation we also take into account the stochastic character of the individuals in a finite population and given densities of people. On the other hand, we include in the simulation the appropriate statistical distributions for the parameters characterizing this disease. An important outcome of our work, besides the generated epidemic curves, is the methodology of determining the effective reproductive number during the main part of the daily new cases of the epidemic. Since this quantity constitutes a fundamental parameter of the SIR-based epidemic models, we also studied how it is affected by small variations of the incubation time and the crucial distance distributions, and furthermore, by the degree of quarantine measures. In addition, we compare our qualitative results with those of selected real epidemiological data

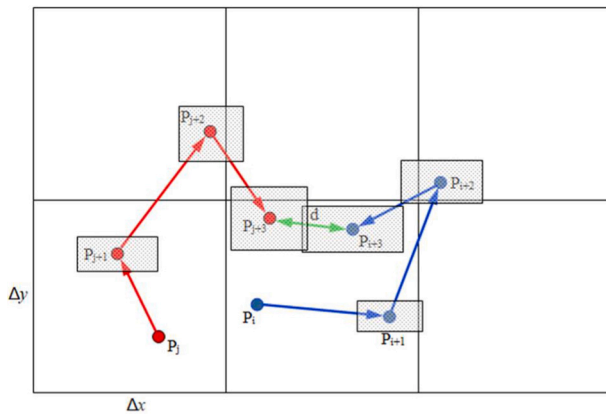
## 1. Introduction

Facing the problem of the spread of the new COVID-19 disease, caused by the new coronavirus SARS-CoV-2 (Dilucca et al., 2020), the standard procedure is to analyse or to fit the epidemiological data with the appropriate mathematical models (Shah et al., 2020; Abdo et al., 2020; Ahmad et al., 2020; Yousaf et al., 2020). This methodology is very useful, not only to determine the particular parameters of the disease, but also to forecast the evolution of its spread by using some typical figures of merit such as the doubling time, the peaking time and the basic and effective reproductive numbers (Kermack and McKendrick, 1927; Siettos and Russo, 2013; Wangping et al., 2020; Cacciapaglia et al., 2020). Another approach is to work in the reverse problem, that is, to generate epidemiological data consistent with the disease under study. By data we mean the fundamental ones, that is, the reported – confirmed “daily new cases” (DNC), expressing the daily rate of infected individuals (<https://www.worldometers.info/coronavirus/>, 2021). The advantage of this methodology, implemented by a Monte Carlo (MC) simulation, is that it allows the use of the stochastic character of the parameters used in the mathematics of epidemics (the known classical SIR model and its extensions), as well as it includes other extrinsic

factors, like quarantine, physical distances and other specific social measures. In this methodology we are studying the set of people in a city or even in a country as a complicate system in which there is no sense in concentrating on the behaviour or habits of individuals or groups of them in detail. Besides, this is surely infeasible. Two factors only must be taken into account, the first is the radius of movement or transport and the second the crucial physical distance for transmitting the disease. The choice of the appropriate range of values, because we are facing random variables with the associated distributions, is one of the main tools for the manipulation of the generated data. To be consistent with the new COVID-19 disease we used medical data which can be found in the available literature. However, the spread of the disease differs widely among the countries and thus we are forced to focus on the particular ones with a similar epidemic “picture”. Another very important and useful outcome of this simulation is the so called “effective reproductive number”,  $R_e$ , which is a function of time. From this we can also determine the fundamental parameter used in SIR-based models, the “basic reproductive number”,  $R_0$ . These quantities are defined in the next Section 3 and present very large uncertainties when they are calculated from real epidemiological raw data (the DNC), even in the case of using parametrization mathematical models (Narison, 2020) and the ideas

<sup>\*</sup> Corresponding author at: National Technical University of Athens, ZIP 15780 Athens, Greece.

E-mail address: [maltezos@central.ntua.gr](mailto:maltezos@central.ntua.gr) (S. Maltezos).



**Fig. 1.** A random movement of two individuals  $P_i$  and  $P_j$  along a 2-D grid during three successive days. Their distance at the 3rd day is equal to  $d$ . The shade regions around any new location represent the uncertainty of the radial translations following Gaussian distribution in  $x$  and  $y$  coordinates.

included in the preprint works (Liu et al., 2020; Vazquez, 2005; Maltezos, 2020; Ziff and Ziff, 2020). The reason for this is, firstly, the large fluctuations of the data and secondly the mathematical procedure for determining  $R_e$  based on variations day-by-day, as we describe in the following Sections. Studying the impact of the main parameters of the disease to  $R_e$  or its time trend is also a basic subject of our work and we suspect that the results can be useful for understanding and facing this serious worldwide pandemic. This pandemic is also called “syndemic”, since it is dangerous to individuals with other subject diseases (especially, autoimmune, cardiovascular and respiratory).

## 2. Simulation methodology

The Monte Carlo (MC) simulation we have developed is based on the idea of simulating the natural mechanism and process which occurred during the spread of the disease (Xie, 2020). In accordance to our claim that we must examine and study the community as a complicated system, we clarify the basic conditions by assuming the following: a) the transmission of the virus occurs only if the physical distance becomes small in the sense of it is in a crucial range where the probability of transmission tends to be equal to one asymptotically, b) the step of movement of a group of people during the day is a free variable of the simulation because it depends on personal or social factors and c) the specific parameters of the particular disease of the new COVID-19 which are mainly, the incubation time and the recovering time, are combined by using a convolution (Hogg and Craig, 2005; Menth et al., 2006; Lloyd, 2001; Feng et al., 2007; Sinkoa et al., 2020; Anastassopoulou et al., 2020). The three aforementioned conditions constitute the core of the simulation and thus their mathematical description is essential. In particular, due to their stochastic character we must determine their probability distribution and the appropriately chosen associated parameters, mainly the mean and standard deviation values. In addition, the population and the corresponding people density which are both related to what is called “size” in the mathematical theory of epidemics, must be specified in the beginning of the simulation. These conditions can be adapted to urban or country cases implemented in real conditions. Below we describe these conditions in more details, where the first three constitute random processes while the fourth one constitutes a set of given constant parameters (Roda et al., 2020; Mohsen et al., 2020) (see also in preprint work (Peng et al., 2020)).

- **Daily step of movement (random process):** the area has been divided in grid shells of squares with side 7.9 m. Along this grid the movement of the dots, because they represent real individuals, must present realistic characteristics. This step represents either a walking

distance or longer distance (by using vehicles). However, because the movements are performed with respect to the previous location, the cumulative distance should be much greater than one step. This parameter has been chosen to be a fraction of the length corresponding to the dimension of a grid shell according to the selected people density and follows a Gaussian distribution. In a typical – baseline run we set a mean step equal to 1/4 of the grid dimensions with a standard deviation (std) equal to 1/12 of the grid dimensions. In Fig. 1 an indicative sequence of random movement by using discrete steps of two individuals during three successive days is presented.

- **Physical distance (random process):** the physical distance is a pivotal issue, in the sense that it is the fundamental parameter causing infection among two individuals in a short distance, if one of them is already infected. The relevant infection mechanism is described analytically in the next bullet. In order to introduce this condition in a realistic way, we must answer two fundamental questions, a) what is the appropriate relationship between the infection probability and the physical distance, and b) what is the statistical distribution and its parameters describing the variance of the probability related to this medical phenomenon. Even if an appropriate distribution is specified, the phenomenon in the real world is very complex: we can imagine a pair of individuals; one of them having a high or low virus load, not only keeping a shorter or longer physical distance, but also wearing a mask or not, or wearing it correctly or incorrectly, facing one another or not and other circumstances playing their particular role in the probability of infection.

Moreover, the environmental conditions (temperature, relative humidity, air ventilation etc.) might contribute to the transmission probability. For this reason, the crucial parameters in the mathematical approach must be considered to be the more effective ones incorporating all the cases described above in a good approximation. The success of this approximation can be verified only in comparison to real epidemiological data analysis. The probability distribution that we assumed is the exponential one, whose the probability density function (PDF) is

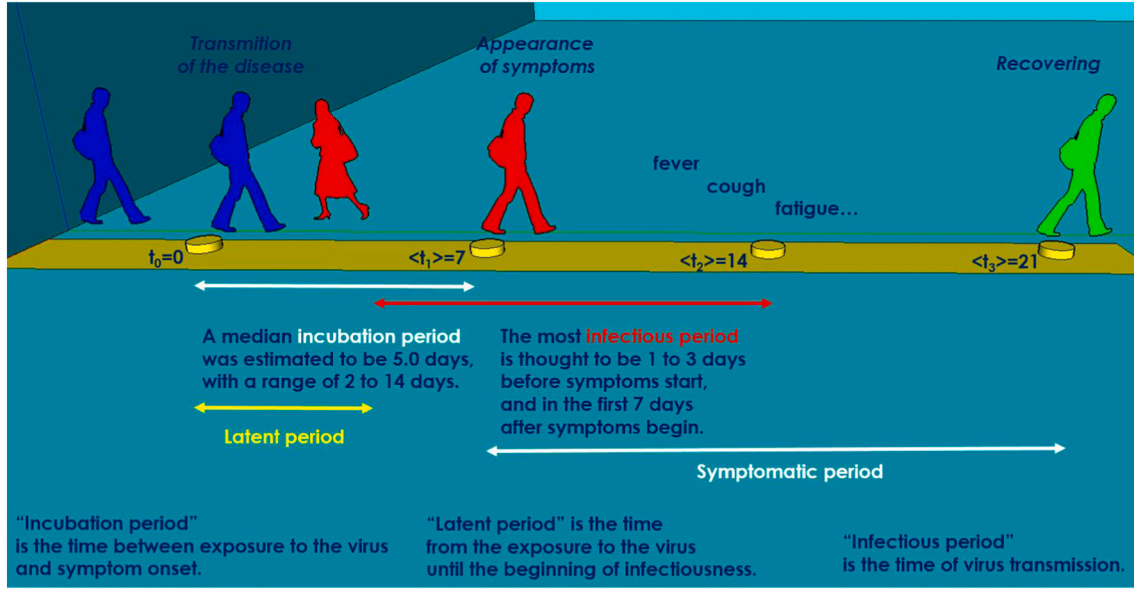
$$f(x; \lambda) = \lambda e^{-\lambda x} \quad (1)$$

where  $x$  is a random variable (assumed  $x > 0$ ) and  $\lambda$  is the “rate parameter”. The mean value is equal to  $1/\lambda$  which is used in the trial runs of our simulation. This probability distribution in our MC simulation is used for randomly generating a crucial physical distance (as a cut-off) in each interaction between two individuals, where  $x$  random variable symbolizes the crucial distance and  $1/\lambda$  is expressed in length units. In particular, in each individual pair we obtain a crucial distance for deciding if an infection has occurred or not.

As an alternative probability distribution we can assume is a sigmoid-shape function. Thus, we created a modified sigmoid function as a function the physical distance,  $d$ , which can take values from maximum equal to 1 (for  $d = 0$ ) to minimum equal to 0 (for  $d = \infty$ ). This presents a saturation for very low values of the distance close to zero. However, this function has been tested by using an algorithm in the simulation code. This modified sigmoid function, assuming  $d \geq 0$  as the physical distance is the following

$$f(d; d_r) = \frac{1}{1 + de^{d-d_r}} \quad (2)$$

where  $d_r$  is a reference parameter specifying in which distance we want to achieve a particular probability of infection. For instance, setting  $d_r = 0.1931$  the probability of infection at a distance of 0.5 m is equal to 0.5. For the typical physical distance of 2 m the probability decreases to 0.0528 (or about only 5.3%). In the present study of the MC simulation we have used the exponential distribution with mean value of 8 m. By using this mean value the probability of infection at the typical (crucial) physical distance of 2 m is 78% which could be considered realistic. We



**Fig. 2.** An abstract design showing the various time periods of the virus transmission mechanism from one infected individual to another. The time moments concern the mean values of the relevant random processes.

pointed out that the use of the sigmoid-shape function causes differences in the results much smaller than the uncertainties of the MC simulation.

- Incubation and recovering time (random process): these parameters are very crucial for a realistic simulation. We are basically forced to use mathematical principles because of their stochastic character, due to the fact that both are related to medical processes inside a human body. It is known that the problems of waiting times are faced using the exponential distribution while in a more advance approach a Gamma distribution is assumed, which is belonging to the same generic distribution family. It is di-parametric including the positive parameters, shape  $\alpha$  and the rate  $\beta$ . The corresponding PDF is the following.

$$f(x; \alpha; \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (3)$$

where  $x$  is a random variable,  $x \in (0, \infty)$ . The mean value is  $\mu = \alpha/\beta$  and the variance  $\sigma^2 = \alpha/\beta^2$ .

In our MC simulation, measuring the random variable of discrete time in days, we have chosen  $\alpha = 6$  (dimensionless) and  $\beta = 2/3 \text{ days}^{-1}$  in order to achieve a mean value equal to 9 days and a rms (square root of variance) equal to 4.9 days. These two parameters lead to a distribution which show characteristics consistent with the most medical observations published in the literature. In Fig. 2 we present an abstract design of the transmission mechanism between two individuals as a function of time, where one of them is infected and the other is considered as susceptible. We have to face a random process and thus we investigated the most realistic assumption for the required distributions.

- Population, Area and Density (given constants): in the simulation, we chose the total number of dots in a finite area. Therefore, the density is specified at the same time. The population density can be considered in large and very dense cities (e.g. like Paris) with around 20,000 people per  $\text{km}^2$  or for typical small size cities with a density around 2000–4000 people per  $\text{km}^2$ . In our trial runs we used the density of 2000 people per  $\text{km}^2$ . As an initial condition we have chosen 10 infected individuals, constituting only 0.5% of the population.

### 3. Determination of the effective reproductive number

The theoretical approach for studying the spread of an epidemic is the SIR model which belongs to the State-Space category of dynamical models. Below we give a brief description of this model based on the following three state equations

$$\frac{dS}{dt} = -\frac{a}{N} SI \quad (4)$$

$$\frac{dI}{dt} = \frac{a}{N} SI - \beta I \quad (5)$$

$$\frac{dR}{dt} = \beta I \quad (6)$$

The function  $S$  represents the number of susceptible individuals,  $I$  the number of infected individuals and  $R$  the number of recovered individuals, all referred to per unit time, usually measured in [days]. The constant factor  $a$  is the transmission rate, the constant factor  $\beta$  is the recovering rate and  $N$  is the size of the system (the total number of individuals assumed constant in time), that is,  $S + I + R = N$ , at every time. The assumptions concerning the initial and final conditions are,  $S(0) \neq 0$ ,  $S(\infty) > 0$ ,  $I(\infty) = 0$  and  $R(0) = 0$ ,  $R(\infty) = N - S(\infty)$ .

The effective reproductive number,  $R_e$  is a fundamental quantity depending on the virus spread rate and the average number of contacts of individuals as a function of time. In particular, it expresses an average of how many secondary individuals one infected individual would transmit the virus to, during the mean infectious period (Narison and Maltezos, 2021; Cori et al., 2013; Hill, 2020) (see also in preprint works (Hebert-Dufresne et al., 2020; Li, 2020)). Mathematically it is defined as

$$R_e = -\frac{dS}{dR} \quad (7)$$

Because  $S + I + R = N$  we have  $S = N - I - R$  and also the differential,  $-dS = d(I + R)$ . Replacing to Eq. (7) we obtain

$$R_e = -\frac{dS}{dR} = \frac{d(I + R)}{dR} = 1 + \frac{dI}{dR} = 1 + \left( \frac{a}{N} SI - \beta I \right) \frac{1}{\beta I} = R_0 \frac{S}{N} \quad (8)$$

Because the condition for creating an epidemic is  $R_e > 1$ , the corresponding condition should be  $S/N > 1/R_0$ , being as a function of time  $t$ . We also observe: at  $t = 0$  we obtain  $R_e(0) \equiv R_0$ , at the peaking time  $t =$

$t_p$  (the time where a maximum of  $I$  is present) we have  $R_e(t_p) = 1$ , while at  $t = \infty$  it takes the threshold value  $R_t = R_0 \frac{S}{N}(\infty) < R_0$ . By using the expressions of Eq. (8) and using the third equation of SIR model,  $\frac{dR}{dt} = \beta I$ , we can determine  $R_e$  at any time  $t$  based only the function  $I$ , as below

$$R_e = 1 + \frac{1}{\beta I} \frac{dI}{dt} \quad (9)$$

The generated data from the MC simulation are not only the DNC corresponding to the theoretical function of  $t$ ,  $I$ , but also that corresponding to the theoretical functions  $S$  and  $R$ . The determination of  $R_e$  requires only the knowledge of  $I$  and as well as the parameter  $\beta$ . This parameter can be considered to be constant because it constitutes an intrinsic characteristic of the disease under study and can be available in the literature. Nevertheless, having the raw epidemiological data of the epidemic spread in a country, the value of  $\beta$  could be fitted or adjusted in order to be consistent with that of the daily recovered cases of individuals, let us call DRC. Even though these reported data are reliable, they contain the relevant delay, in average around 14 days, with respect to DNC ones. For this reason we prefer to adhere to the theory described by SIR-based models.

Below, we present the methodology to estimate the uncertainty of  $R_e$ . According to Eq. (9) the main source of the uncertainty is that of  $I$  due to its statistical fluctuations while the deviation of the parameter  $\beta$  from a standard bibliographic value is very small and is considered constant.

In order to calculate the uncertainty transmission from  $I$  to  $R_e$  we can assume a short time period around  $t = 0$ , equal to  $\tau_R$ , where  $R_e$  can be considered constant. This assumption is realistic, and we can also say necessary, because the mean recovering time in this disease is finite beside its stochastic variations. Being  $R_e$  constant, we can write

$$\frac{dI}{dt} = \beta(R_e - 1)I \quad (10)$$

We integrate from an arbitrary time  $t_{i-1}$  to  $t_i$ .

$$\int_{t_{i-1}}^{t_i} \frac{dI'}{I'} = \beta(R_e - 1)\tau_R \quad (11)$$

Leading to the expression

$$I_{t_i} = I_{t_{i-1}} e^{c(R_e - 1)} \quad (12)$$

where we have set  $c \equiv \beta\tau_R$ , which differs slightly from the unity in practice presenting minor impact to the result. As we can observe, according to this assumption of constant  $R_e$  in finite time slots, its value depends on the logarithm of the ratio of the two successive values of  $I$  at the beginning and the end of this time slot. This expression can also be very useful when a surveillance of the epidemic must be implemented for making a reliable prediction of it, mainly during its mitigation stage. At this point, based on this equation, Eq. (12), we proceed to solve the algebraic equation for  $R_e$  obtaining

$$R_e = 1 + \ln\left(\frac{I_{t_i}}{I_{t_{i-1}}}\right)^{1/c} \quad (13)$$

The values of  $I$  follow the Poisson distribution, presenting the associated uncertainty  $\delta I_{t_i} = k\sqrt{I_{t_i}}$  and  $\delta I_{t_{i-1}} = k\sqrt{I_{t_{i-1}}}$  respectively, where  $k$  is a positive constant which is used for the cases of observing larger fluctuations in the raw data of  $I$  taking a value greater than one. Therefore, the uncertainty of  $R_e$  should be

$$\delta R_e = \sqrt{\frac{1}{I_{t_i}^2}(\delta I_{t_i})^2 + \frac{1}{I_{t_{i-1}}^2}(\delta I_{t_{i-1}})^2} = \frac{k}{c} \sqrt{\frac{1}{I_{t_i}} + \frac{1}{I_{t_{i-1}}}} \quad (14)$$

While the corresponding relative one is

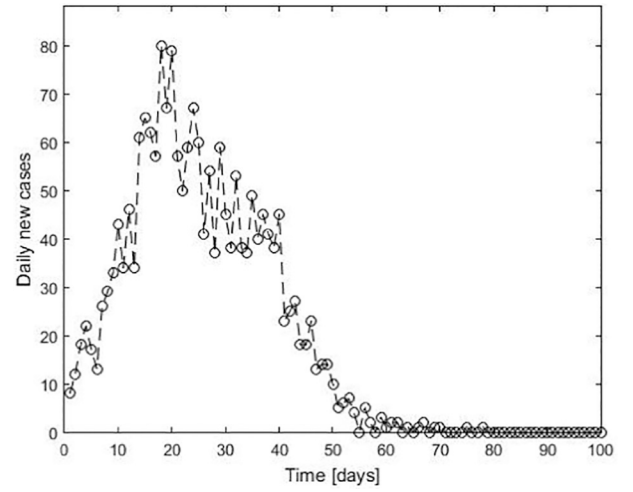


Fig. 3. The generated epidemiological data (DNC) by the MC simulation, counted day-by-day during a single run and cumulated along the whole area of movement.

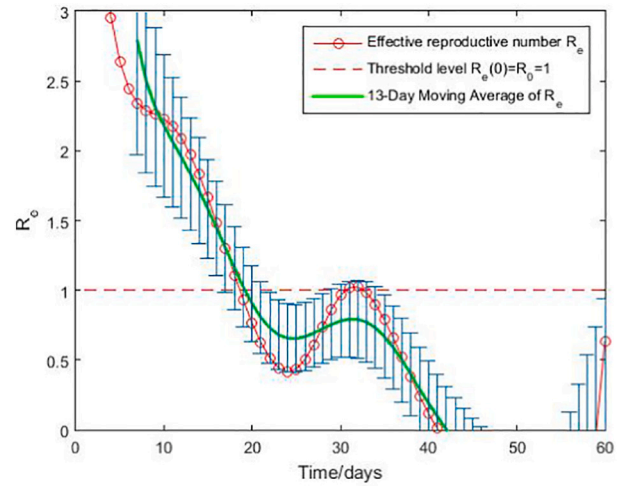


Fig. 4. The effective reproductive number from the MC simulation as a function of time. The error bars correspond to the moving average values.

$$\frac{\delta R_e}{R_e} = \frac{k/c}{1 + \ln\left(\frac{I_{t_i}}{I_{t_{i-1}}}\right)^{1/c}} \frac{\sqrt{I_{t_i} + I_{t_{i-1}}}}{\sqrt{I_{t_i} I_{t_{i-1}}}} \quad (15)$$

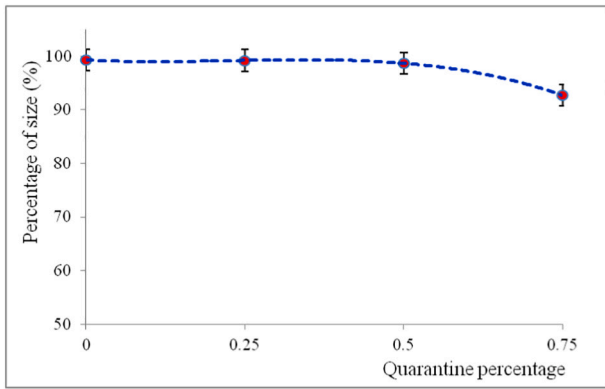
In the segment around the peak of the DNC curve, we theoretically expect  $R_e = 1$ . Indeed, from Eq. (15) we obtain this value because of the maximum of the curve (having an extremum), where  $I_{t_i} \approx I_{t_{i-1}}$ . In this region, the relative uncertainty takes its minimum value and becomes

$$\delta R_e = \frac{k}{c} \sqrt{\frac{2}{I_{t_i}}} \quad (16)$$

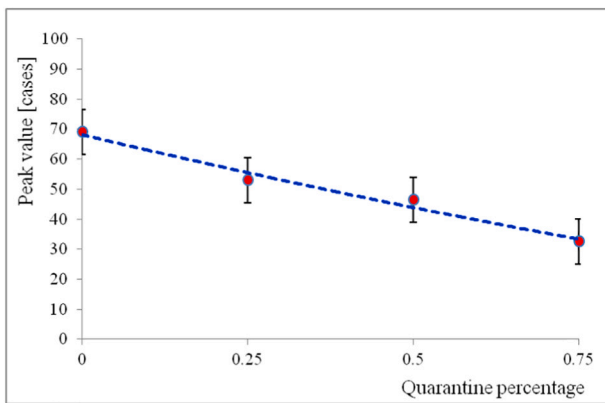
The basic reproductive number,  $R_0$ , can be calculated by the above equations by using the values in the first time slot (typical is of the order of 13 days), that is,  $I_0$  and  $I_{t_1}$ . Definitely,  $R_0$  is the maximum value of  $R_e$ , as it can be easily proved from Eq. (8) where at  $t = 0$  the fraction  $S/N$  takes its maximum value and then is decreasing due to monotonically decreasing of  $S$ .

Implementing the above methodology in the simulation program we use two processes for improving the accuracy of  $R_e$  determination. The first concerns a digital filtering, reducing the statistical fluctuations, and





**Fig. 5.** The size percentage as a function of quarantine degree. The dotted line represents a 2<sup>nd</sup> degree polynomial fit.



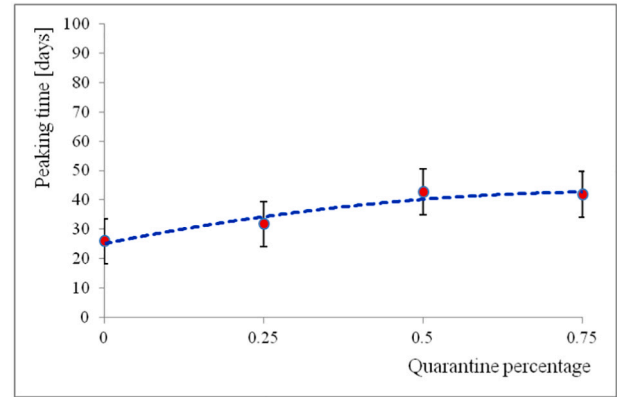
**Fig. 6.** The peak value of DNC as a function of quarantine degree. The dotted line represents a 2<sup>nd</sup> degree polynomial fit.

the second is a moving average algorithm, both realized by functions. The MC simulation and its associated algorithms have been written in Matlab package (Mathworks, 2015) by using also some of its tool boxes.

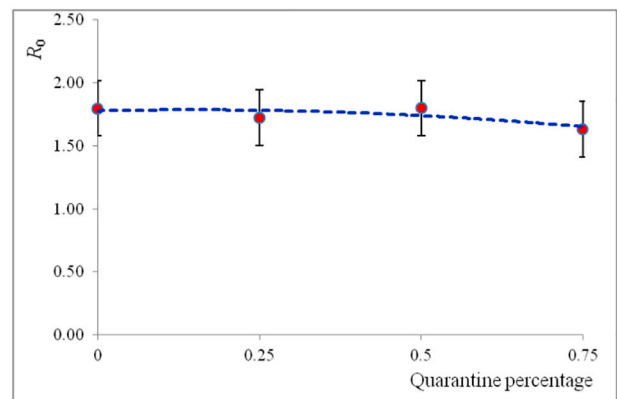
## 4. Implementation and results

### 4.1. Results obtained by the simulation

During a typical run of the MC simulation the  $N$  dots within the assumed area are categorized by using colors: a) the black for the susceptible individuals, b) the red for the infected individuals and c) the green for the recovered individuals. The dots change their location day-by-day representing the random movement of the individuals in the real world. In Fig. 3 we present a diagram with the daily new data (DNC) generated by the MC simulation. In Fig. 4 the corresponding effective reproductive number calculated by the MC simulation as a function of time (red circles and spline solid line) is shown. We also used a smoothed curve (green solid line) by calculating the “moving average” every 13 days, which has been chosen as the optimum time interval. Focusing on the peaking time in both plots, we observe that the value  $R_e = 1$  corresponds to the peaking time, as one might expect according to the SIR-based models. Moreover, apart from the gradually decreasing trend, we observe some slow variations, before and after the value corresponding to the peak, a phenomenon which is hard to be explained in a simple way. These variations are visible also in the obtained  $R_e$  from real epidemiological data (see Section 4.2). The maximum value is around 1.80 and the minimum, relatively adequate accurate value, is at the level of 0.2 around day 40.



**Fig. 7.** The peaking time of DNC as a function of quarantine degree. The dotted line represents a 2<sup>nd</sup> degree polynomial fit.



**Fig. 8.** The basic reproductive number as a function of quarantine degree. The dotted line represents a 2<sup>nd</sup> degree polynomial fit.

**Table 1**

Summary of the MC simulation results related to the effect of the quarantine degree, with the same density and the mean daily step and the std. of daily step was that of the baseline run, that is, 7.90 m and 2.64 m respectively.

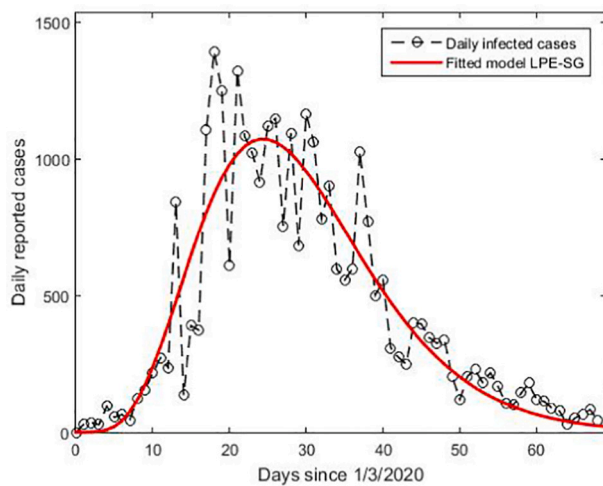
Quarantine degree [%]	Percentage of size %	DNC peak value [cases]	DNC peaking time [days]	$R_0$
0	99.3	69	26	1.80
25	99.2	53	32	1.73
50	98.7	46	43	1.80
75	92.7	33	42	1.63

The density was set 2000 p/km<sup>2</sup>.

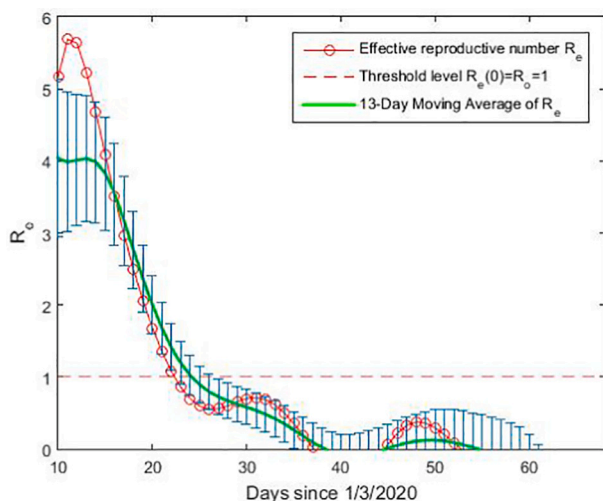
### 4.2. Qualitative discussions and comparisons

By the MC simulation we have studied the effect of the degree of quarantine, by means of the positive effect on the spread dynamic of the epidemic. We have set three degrees of quarantine by reducing the movement range of the individuals and the rms deviation in steps of 0.25, from zero (no quarantine) to 0.75. Because of the fluctuations caused by the stochastic process of the MC simulation we have run it 10 times calculating the mean and rms value for each particular quantity which was adequate to achieve statistically significant results. In Figs. 5–8 these results are plotted with the associated fitted curves. In Table 1 the obtained numerical results are summarized.

Observing these results we find some interesting consequences of the quarantine as follows: The percentage of size (the percentage of the infected population with respect to the initial one) shows a very small decrease of 6% within the time scale under study. Also, the greater the



**Fig. 9.** The epidemiological data (DNC) for the first phase of epidemic outbreak in Switzerland since 1/3/2020, obtained from (<https://www.worldometers.info/coronavirus/>, 2021).



**Fig. 10.** The effective reproductive number as a function of time for the first phase of epidemic outbreak in Switzerland, reproduced as given in (Narison and Maltezos, 2021) used as a case study.

degree of quarantine the greater the decrease in the DNC peak value, while the peaking time extended accordingly. Concerning the basic reproductive number,  $R_0$ , we observe that it shows small decrease compared to the statistical errors which is of the order of 10%. In the approach of the SIR epidemic model  $R_0$  is a constant quantity related to the two basic parameters, the transmission rate and the recovering rate and it is expected to be affected by the quarantine in some degree. The overall variations, from 0 to 75%, due to the quarantine are: the DNC peak value is decreased by 53%, while the DNC peaking time is extended by 61%.

#### 4.3. Results analysing real epidemiological data

To compare the results of the MC simulation with real epidemiological data we have chosen a country as a case study where the raw data and any parametrization of them present a “picture” very close to the theoretical as well as having relatively small fluctuations. This country is Switzerland where, besides the above characteristics, the compliance with the measures is largely secured. Therefore, it can be considered to be a well reference for comparison. In Figs. 9 and 10 the DNC data

parametrized by the Large Population Epidemic – Semi Gaussian (LPE-SG) model and the  $R_e$  are shown respectively (Maltezos, 2021). The results of the MC simulation we developed show similar qualitative consistency with the real epidemic cases, by means of the DNC shape, peaking time appearing, overall time scale and the  $R_e$  drop-off general shape. We must notify that the present MC simulation is a method mainly for generating epidemiological data than performing analysis or fitting to the real ones. However, it can be adapted for being more accurate referring to a particular country under study.

## 5. Conclusions

A Monte Carlo simulation has been developed to study the spread dynamics of the new COVID-19 disease. The novel approach we followed in methodology was based on fundamental mechanisms during the spread of the epidemic. One issue of key importance was the realistic selection of statistical distributions so that the results of the simulation correspond to those previously recorded to date in various countries during the first and second “wave” of the epidemic. This task was faced by studying the appropriate statistical distributions and their associated parameters. Moreover, we described the mathematical process for determining the effective reproductive number with a good approximation, either by the simulated or by the real epidemiological data. This quantity is very important in playing the role of “figure of merit” helping to survey and understand the dynamics of the epidemic spread. We also studied the effect of quarantine at different levels of its implementation on the shape of the epidemic curve and its parameters. In addition, we present some indicative real epidemic results analysed using a specific parametrization model.

## Declaration of competing interest

None.

## Acknowledgements

We gratefully thank Mrs. Barbara Theodoropoulou, experienced English teacher, for her willing support in the language proofreading.

## References

- Abdo, Mohammed S., Shah, Kamal, Wahash, Hanan A., Panchal, Satish K., 2020. On a comprehensive model of the novel coronavirus (COVID-19) under Mittag-Leffler derivative. *Chaos, Solitons Fractals* 135, 109867. <https://doi.org/10.1016/j.chaos.2020.109867>.
- Ahmad, Shabir, Ullah, Aman, Al-Mdallal, Qasem M., Khan, Hasib, Shah, Kamal, Khan, Aziz, 2020. Fractional order mathematical modeling of COVID-19 transmission. *Chaos, Solitons Fractals* 139, 110256. <https://doi.org/10.1016/j.chaos.2020.110256>.
- Anastassopoulou, Cleo, Russo, Lucia, Tsakris, Athanasios, Siettos, Constantinos, 2020. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One*. <https://doi.org/10.1371/journal.pone.0230405>.
- Cacciapaglia, Giacomo, Cot, Corentin, Sannino, Francesco, 2020. Second wave COVID-19 pandemics in Europe: a temporal playbook. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-72611-5>.
- Cori, Anne, Ferguson, Neil M., Fraser, Christophe, Cauchemez, Simon, 2013. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* 178 (9) <https://doi.org/10.1093/aje/kwt133>.
- Dilucca, Maddalena, Forcelloni, Sergio, Georgakilas, Alexandros G., Giansanti, Andrea, Pavlopoulou, Athanasia, 2020. Codon usage and phenotypic divergences of SARS-CoV-2 genes. *Viruses* 12 (5), 498. <https://doi.org/10.3390/v12050498>.
- Feng, Zhilan, Xu, Dashun, Zhao, Haiyun, 2007. Epidemiological models with non-exponentially distributed disease stages and applications to disease control. *Bull. Math. Biol.* 69, 1511–1536.
- Hebert-Dufresne, Laurent, Althouse, Benjamin M., Scarpino, Samuel V., Allard, Antoine, 14 Apr 2020. Beyond  $R_0$ : heterogeneity in secondary infections and probabilistic epidemic forecasting. *arXiv*, 2020.04004v2 [q-bio.PE].
- Hill, Alison L., 2020. The math behind epidemics. *Phys. Today* 73 (11), 28. <https://doi.org/10.1063/PT.3.4614>.
- Hogg, Robert V., Craig, Allen T., 2005. *Introduction to Mathematical Statistics*. Copyright © 1978. Macmillan Publishing Co., Inc., New York. Collier Macmillan Publishers, LONDON. <https://www.worldometers.info/coronavirus/>, 2021.

- Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* 115 (772), 700–721.
- Li, Jun, 7 Feb 2020. A robust stochastic method of estimating the transmission potential of 2019-nCoV. *arXiv*, 2002.03828v1 [q-bio.PE].
- Liu, Z., Magal, P., Seydi, O., Webb, G., 27 Feb 2020. Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. *arXiv*, 2002.12298v1 [q-bio.PE].
- Lloyd, Alun L., 2001. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proc. R. Soc. Lond. B* p985–p993. <https://doi.org/10.1098/rspb.2001.1599>.
- Maltezos, S., 2020. Parametrization model motivated from physical processes for studying the spread of COVID-19 epidemic. *arXiv submit/3123114* [q-bio.QM].
- Maltezos, S., 2021. Methodology for modelling the new COVID-19 pandemic spread and implementation to European countries. *Infect. Genet. Evol.* <https://doi.org/10.1016/j.meegid.2021.104817>.
- Mathworks, 2015. MATLAB R2015a.
- Menth, Michael, Henjes, Robert, Zepfel, Christian, Tran-Gia, Phuoc, 2006. Gamma-approximation for the waiting time distribution function of the M/G/1–1 Queue. In: 2nd Conference on Next Generation Internet Design and Engineering, Valencia, Spain, (April 2006). NGI '06, 2006.
- Mohsen, Ahmed A., AL-Husseiny, Hassan Fadhil, Zhou, Xueyong, Hattaf, Khalid, September 2020. Global stability of COVID-19 model involving the quarantine strategy and media coverage effects. *AIMS Public Health*. <https://doi.org/10.3934/publichealth.2020047>.
- Narison, Stephan, 2020. The first months of COVID-19 in Madagascar. *Infect. Genet. Evol.* 85, 104506.
- Narison, S., Maltezos, S., 2021. Scrutinizing the spread of Covid-19 in Madagascar. *medRxiv*. <https://doi.org/10.1101/2020.09.27.20202556> preprint has posted on.
- Peng, Liangrong, et al., 16 Feb 2020. Epidemic analysis of COVID-19 in China by dynamical modeling. *arXiv*, 2002.06563v1. q-bio.PE].
- Roda, Weston C., Varughese, Marie B., Han, Donglin, Li, Michael Y., 2020. Why is it difficult to accurately predict the COVID-19 epidemic? *Infect. Dis. Model.* 5.
- Shah, Kamal, Abdeljawad, Thabet, Mahariq, Ibrahim, Jarad, Fahd, 2020. Qualitative analysis of a mathematical model in the time of COVID-19. *Biomed. Res. Int.* 2020 <https://doi.org/10.1155/2020/5098598>. Article ID 5098598.
- Siettos, K., Russo, L., 2013. Mathematical modeling of infectious disease dynamics. *Virulence* 4 (4), 295–306. <https://doi.org/10.4161/viru.24041>.
- Sinkoa, Arthur, Turner, Alex, Nikolovab, Silviya, Suttonb, Matt, . Models for Waiting Times in Healthcare: Comparative Study Using Scottish Administrative Data. *www.interscience.wiley.com*. <https://doi.org/10.1002/sim.0000>.
- Vazquez, Alexei, 2005. Polynomial growth in branching processes with diverging reproductive number. *arXiv cond-mat/0505116*.
- Wangping, Jia, et al., 2020. Extended SIR prediction models of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *Front. Med.* <https://doi.org/10.3389/fmed.2020.00169>.
- Xie, G., 2020. A novel Monte Carlo simulation procedure for modelling COVID-19 spread over time. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-70091-1>.
- Yousaf, Muhammad, Zahir, Samiha, Riaz, Muhammad, Hussain, Sardar Muhammad, Shah, Kamal, 2020. Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan. *Chaos, Solitons Fractals* 138, 109926. <https://doi.org/10.1016/j.chaos.2020.109926>.
- Ziff, Anna L., Ziff, Robert M., 2020. Fractal Kinetics of COVID-19 Pandemic (with update 3/1/20). <https://doi.org/10.1101/2020.02.16.20023820>.