Natural Computing (2022) 21:449–461
https://doi.org/10.1007/s11047-022-09893-3

Estimates of the collective immunity to COVID-19 derived from a stochastic cellular automaton based framework

Isaías Lima[1] · Pedro Paulo Balbi[2]

Accepted: 10 May 2022 / Published online: 18 June 2022

Abstract
In the context of the propagation of infectious diseases, when a sufficient degree of immunisation is achieved within a population, the spread of the disease is ended or significantly decreased, leading to collective immunity, meaning the indirect protection given by immune individuals to susceptible individuals. Here we describe the estimates of the collective immunity to COVID-19 from a stochastic cellular automaton based model designed to emulate the spread of SARS-CoV-2 in a population of static individuals interacting only via a Moore neighbourhood of radius one, with a view to analyze the impact of initially immune individuals on the dynamics of COVID-19. This impact was measured by comparing a progression of initial immunity ratio—the percentage of immunised individuals before patient zero starts infecting its neighbourhood—from 0 to 95% of the initial population, with the number of susceptible individuals not contaminated, the peak value of active cases, the total number of deaths and the emulated pandemic duration in days. The influence of this range of immunities over the model was tested with different parameterisations regarding the uncertainties involved in the model such as the durations of the cellular automaton states, the contamination contributions of each state and the state transition probabilities. A collective immunity threshold of $55\% \pm 2.5\%$ on average was obtained from this procedure, under four distinct parameterisations, which is in tune with the estimates of the currently available medical literature, even increasing the uncertainty of the input parameters.

Keywords Stochastic cellular automaton · COVID-19 · SARS-CoV-2 · Contagious disease dynamic · collective immunity

1 Introduction

The Coronavirus Disease 19 (COVID-19) is a contagious disease which can be transmitted via interactions between human beings, by means of particles emanated from the respiratory system of a contaminated person to the mouth or nose of another person. Contaminated individuals present symptoms like fever, body aches and shortage of breath, which are caused by the *Severe Accute Respiratory Syndrome-Coronavirus* (SARS-CoV-2) (Ezhilan et al. 2021). This virus was initially found in animals, and the origin of its transition to human beings is still unknown (Velavan and Meyer 2020).

Generally, the mass contamination caused by this virus began in December of 2019 in the province of Hubei, located in the People's Republic of China, initally diagnosed as a type of pneumonia. In the subsequent months, the World Health Organisation (WHO) declared a pandemic state and authorities from each country began to adopt health and safety protocols trying to contain COVID-19's transmission (Velavan and Meyer 2020). These authorities have been relying on different types of quarantine, from social distancing to lockdowns, so that the effects of these decisions during the last year and a half have led to worldwide decrease in economy, poverty raise and human lives (Ciotti et al. 2020). Consequently, many

✉ Isaías Lima
isaiahlima18@gmail.com

Pedro Paulo Balbi
pedrob@mackenzie.br

[1] Pós-Graduação em Engenharia Elétrica e Computação, Universidade Presbiteriana Mackenzie, Rua da Consolação 896, Consolação, 01302-907 São Paulo, SP, Brazil

[2] Faculdade de Computação e Informática & Pós-Graduação em Engenharia Elétrica e Computação, Universidade Presbiteriana Mackenzie, Rua da Consolação 896, Consolação, 01302-907 São Paulo, SP, Brazil

economic activities were affected and damaged (Ezhilan et al. 2021), prompting political and sanitary agents, and the population in general, to start asking when these restrictions could be removed without causing a new wave of infections.

Understanding when to start removing the restrictions involves essentially understanding when and how *collective immunity* can be achieved (Britton et al. 2020), that is, the individual protection from infection given by a sufficiently large proportion of immune individuals in a society, decreasing the infection rate, or the average number of individuals infected by one infected person, to values under 1.0 (Randolph and Barreiro 2020).

Studying the behaviour of diseases in general, specially in planning prevention measures for authorities can cost time and human lives. But it is possible to resort to theories and models related to epidemiological studies, to develop mathematical and computational approaches to help predict and test control strategies, as well as understand the impact of immune individuals in a disease. In this context, cellular automata based models constitute a paradigmatic example, as their use in disease spreading studies have been shown to be both viable and reliable (Mikler et al. 2005; Monteiro et al. 2020; White et al. 2007; Schimit 2021).

Cellular automaton (CAs) are discrete dynamical systems defined by a regular lattice of cells, where each one is defined by a finite number of discrete states. The dynamics of cellular automata is defined locally, with each cell having its state changed according to its current state and those of its neighbouring cells. And even simple individual cell behaviours, cellular automata may lead to arbitrarily complex global behaviours. On their part, stochastic cellular automata are the ones whose cells define their next state based on a probability distribution.

Here, a computational model based on stochastic cellular automata is proposed to aid in understanding the COVID-19 dynamics, with a focus on collective immunity. We describe the principles of a stochastic cellular automaton which allows to emulate the spreading of COVID-19 in a population, so as to obtain the proportion of the population that has to be vaccinated, or become naturally immunised, for collective immunity to be established. Even though one might argue that the regular topology of cellular automata might not be adequate to reliably support the dynamics of a disease spreading, this has been the case in various successful efforts in the literature (Yakowitz et al. 1990; Fuentes and Kuperman 1999; Schimit and Monteiro 2009; Holko et al. 2016; Schimit and Monteiro 2011; Ahmed and Agiza 1998; Sirakoulis et al. 2000). So, in tune with those, our main goal here is to probe whether such an abstract and minimal model could capture a fundamental aspect of the disease like the collective immunity.

In Sect. 2, the computational model based on the proposed cellular automaton is discussed, and is followed by a section on the methodology underlying the computational experiments carried out. In Sect. 4, the results obtained are presented and discussed, according to four distinct parameterisations setups, hereby called Experiments. Finally, in Sect. 5 the results obtained are discussed and compared to related works.

2 The reference model and its computational implementation

2.1 The reference model

During the ongoing COVID-19 pandemic, the following four conditions were observed in contaminated individuals, related to the duration and the intensity of the disease, according to COVID-19 related literature at the beginning of the pandemic (Ferguson et al. 2020; Lauer et al. 2020; Liu et al. 2020; Wölfel et al. 2020):

- **No symptoms** Individuals who do not present any symptoms of the disease but can contaminate others through social interactions;
- **Moderate symptoms** Describing the ones who present a few symptoms but quickly become immune, but still with the possibility of contaminating others through social interactions;
- **Severe symptoms** Those who have to be hospitalised and present a moderate death risk;
- **Critical symptoms** Individuals in need of hospitalisation in an Intensive Care Unit (ICU), maintained by forced ventilation, and presenting a high death risk.

Figure 1 presents the reference model we rely upon (due to Rahmé 2020), based on Ferguson et al. (2020), Lauer et al. (2020), Liu et al. (2020). It illustrates the four previous possible dynamics of the disease after contagion and divides COVID-19 in terms of the stages of the disease, as it progresses. On the left-hand side, the figure displays the proportions of the population following each dynamics, which can be viewed as a trail followed over time, from contagion to full recovery (or death), with their durations, and correlated to the severity of the symptoms. Also, as severity rises, the recovery is not a single destination but a branch with two possibilities, one of them being the death of an individual, which occur at the different death rates expressed on the rightmost part of the figure. After recovery from the symptoms, the figure also shows the periods where the remaining viruses in an individual can still contaminate others, while their antibodies combat the disease. Each trail shown in the figure has internal segments regarding incubation period, manifestation of

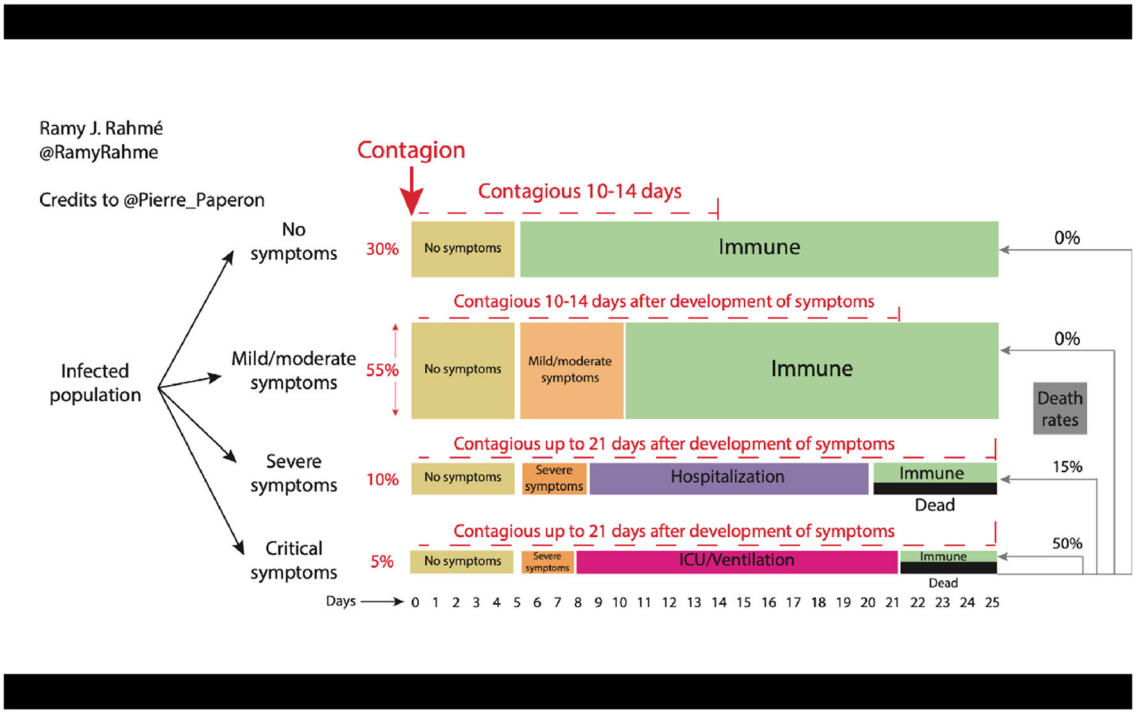

Fig. 1 Timeline, progression and mortality of COVID-19, from Rahmé (2020), based on Ferguson et al. (2020), Lauer et al. (2020), Liu et al. (2020), Wölfel et al. (2020). This chart describes the general scheme supporting the present work. The values were taken by the disease references at the time the chart was created, in 2020

symptoms, hospitalisation and recovery, with their respective durations; for example, the period associated with *No Symptoms*, which describes the incubation period of the disease, takes around five days to end.

As patient zero starts interacting with other members of a population, the disease will start spreading and, after some time, every person will be contaminated and undergo COVID-19 effects. Some will become immune, others will die, and others will not even be contaminated. The dynamics proposed and described by this reference model are experienced individually.

The model can be couched in terms of a stochastic cellular automaton (White et al. 2007; Mikler et al. 2005), by representing the variety of possibilities elicited in Fig. 1 as the states of a cell of the CA, the probabilities therein as those governing the state transitions among the states and each day of the timeline expressed as a time step of the automaton. The details about how we go about it are described in the next section.

2.2 Representing COVID-19 dynamics with a stochastic cellular automaton

2.2.1 The states

Let us consider a closed population with $n \times n$ individuals allocated over a two-dimensional lattice, each one interacting only within those in its direct neighbourhood of the eight surrounding individuals (i.e., in a Moore neighbourhood). Also, let us assume the population has fixed boundaries, assuming a *no interactions* state to missing neighbours. In other words, the interactions will only be among the cells in a Moore nighbourhood of unitary radius and, if a cell is at the boundary of the lattice, its neighbourhood will only include those inside the lattice.

Consider that one of the individuals is arbitrarily contaminated with SARS-CoV-2 and becomes patient zero for COVID-19, thus entering one of the trails mentioned above (no symptoms, moderate symptoms, severe symptoms or critical symptoms).

Let us represent the cell states of the cellular automaton by the notation $\boldsymbol{S}_{id}^{t}$, with t being the state duration in terms of the number of time steps of the automaton's temporal evolution; *id* being just a state identificator; and $\boldsymbol{S}$ standing for the type of state class it belongs to, out of the possibilities **S**usceptible, **I**nfected and **R**emoved. Notice that these are the same alternatives that define a SIR model of epidemic dynamics (Blavatska and Holovatch 2021).

Based on the previous notation, the reference model of Fig. 1 can now be expressed as follows. The durations of specific states with time dependent transition are obtained from the reference model, also being the mean values of probability distributions (further on this will be addressed in Sect. 3.3). These mean times are obtained from the time

axis at the bottom of the chart expressed in days. For example, the duration 13 for state $\mathbf{I}_6^{13}$ corresponds to duration of the magenta ICU/VENTILATION segment, just above the time axis in Fig. 1; the segment starts at day 8 and ends at day 21, thus resulting in the state duration of 13 days (to be translated to 13 times steps in the computational model).

- **S** type state:
  - SUSCEPTIBLE ($\mathbf{S}_0^{f(\mathcal{P}_c)}$): Individuals (cells) capable of being contaminated, during as long the neighbourhood contamination probability $\mathcal{P}_c$ is enough to cause a change to the next state.
- **I** type states:
  - NO SYMPTOMS ($\mathbf{I}_1^5$): Absence of symptoms during the incubation period;
  - MODERATE SYMPTOMS ($\mathbf{I}_2^5$): Moderate symptoms manifestation;
  - SEVERE SYMPTOMS ($\mathbf{I}_3^3$): Severe symptoms manifestation;
  - CRITICAL SYMPTOMS ($\mathbf{I}_4^3$): Critical symptoms manifestation;
  - HOSPITALISED ($\mathbf{I}_5^{12}$): Hospitalisation period after severe symptoms;
  - ICU/VENTILATION ($\mathbf{I}_6^{13}$): Hospitalisation in intensive care unit, after the manifestation of critical symptoms;
  - CONTAGIOUS WITHOUT SYMPTOMS ($\mathbf{I}_7^{10}$): Disease effects ceased, but it is still possible to contaminate others;
  - CONTAGIOUS WITH MODERATE SYMPTOMS ($\mathbf{I}_8^{11}$): Disease effects ceased, but it is still possible to contaminate others after moderate symptoms;
  - CONTAGIOUS WITH SEVERE SYMPTOMS ($\mathbf{I}_9^4$): Disease effects ceased, but it is still possible to contaminate others after hospitalisation;
  - CONTAGIOUS WITH CRITICAL SYMPTOMS ($\mathbf{I}_{10}^5$): Disease effects ceased, but it is still possible to contaminate others after intensive care unit internation.
- **R** type states:
  - DECEASED ($\mathbf{R}_{11}^{\infty}$): Individuals who did not resist the disease effects and were permanently removed from the population, thus no longer interacting;
  - IMMUNE ($\mathbf{R}_{12}^{\infty}$): Immunised individuals who cannot be contaminated again and do not contaminate others.

2.2.2 Contamination contributions

Notice that each state has its own **contamination contribution**, which indicates how much a cell can infect a neighbouring cell in the SUSCEPTIBLE state, at each timestep. For example, an individual who did not have symptoms, after the incubation period (NO SYMPTOMS state) would stay for some time contagious, even possibly unaware of it. For this state, the contamination contribution could be considered 100%. The following provides the whole picture we assumed:

- SUSCEPTIBLE, DECEASED and IMMUNE states were assumed to have 0% contamination contribution;
- HOSPITALISED and ICU/VENTILATION states were assigned 10% contamination contribution (a small rate, since the individuals in these states are under the more controlled conditions of a hospital);
- The remaining states were considered with 100% contamination contribution.

2.2.3 State transitions

It is necessary to define and describe the **state transitions** between states. If any individual in a population gets infected, independently of the dynamics followed, he/she is capable of affecting the neighbourhood, with the contributions expressed in Sect. 2.2.2. In the same way, the neighbourhood is capable of afecting the individual, indicating that there is a correlation between the size of the neighbourhood and the probability of getting ill.

Accordingly, the transitions associated to the SUSCEPTIBLE state are defined by the neighbourhood contamination probability $\mathcal{P}_c$, which is drawn from the average of the contamination contributions of each neighbour in a Moore neighbourhood of unitary radius (the sum of the contamination contributions defined by the current states of the eight neighbourhood cells divided by eight). For the other states, the state transitions are governed by the **durations of the states** (defined in days in the reference model and in time steps of the automaton).

So, we can define the state transitions as follows:

- If a cell is in the SUSCEPTIBLE state, the next state will be NO SYMPTOMS with probability $\mathcal{P}_c$;
- The NO SYMPTOMS state will remain for five time steps. The next possible states are CONTAGIOUS WITHOUT SYMPTOMS with 30% probability, MODERATE SYMPTOMS with 55%, SEVERE SYMPTOMS with 10% and CRITICAL SYMPTOMS with 5%;
- The transitions for the remaining states are defined just by the duration in time steps and can be better observed in Fig. 1. As stablished earlier, these durations are

defined by probability distributions, as addressed in Sect. 3.3);
- The transition to the IMMUNE or DECEASED states is also defined by the previous state durations but with a probability for each one. If the cell is in CONTAGIOUS WITHOUT SYMPTOMS or CONTAGIOUS WITH MODERATE SYMPTOMS states, the probability of moving to IMMUNE state is 100%; if the cell is in HOSPITALISED state, its probability to change to the IMMUNE state is 85% and to DECEASED is 15%; if the cell is in ICU/VENTILATION state the probability of turning to the IMMUNE state is 50% and to DECEASED is 50%.

The resulting state transition diagram of the model describing each probability and each state transition can be seen in Fig. 2.

3 Methodology of the computational experiments

3.1 Preliminaries

With the state transitions and probabilities defined, the proposed cellular automaton can be implemented and simulated, ending its iterations when a **global convergence** is achieved. This steady state condition is defined by a global configuration with all the cells in the states SUSCEPTIBLE, DECEASED or IMMUNE. The initial configuration of the lattice is assumed to contain a randomly chosen single cell in the NO SYMPTOMS state, surrounded by cells in the SUSCEPTIBLE state or the IMMUNE state also randomly distributed (this will be better described in Sect. 3.2).

The characterisation of the dynamics of the temporal evolution of the cellular automaton can then be obtained by monitoring data such as:

- The number of cells in SUSCEPTIBLE state, representing the number of individuals not contaminated;
- The number of cells in DECEASED state, representing the number of deceased individuals;
- The number of cells in IMMUNE state, representing the number of recovered individuals;
- The number of cells at the other states, representing the infected individuals.

Consequently, the previous data provides what is needed to study the collective immunity effect related to COVID-19 in a population, specially the estimate of the necessary immunity ratio to achieve collective imunity.

3.2 Analysing collective immunity

With the computational model, the impact of immune individuals can then be studied by varying the initial global

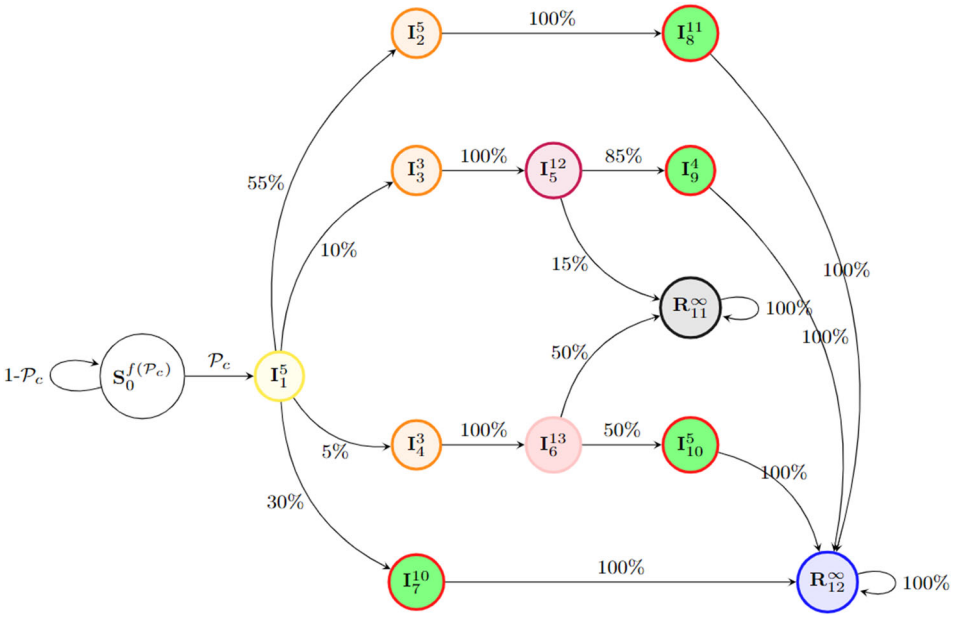

Fig. 2 State transition diagram of the proposed cellular automaton. The durations of each state, superscripted in each state identification, are directly obtained from the reference model, but they really correspond in the computational experiments to the mean value of a probability distribution. At the time a state should transit to the next, this occurs with 100% probability, unless multiple different states might follow; in these cases, the transition probabilities are divided among the possible next states. States involved: SUSCEPTIBLE ($\mathbf{S}_0^{f(\mathcal{P}_c)}$); NO SYMPTOMS ($\mathbf{I}_1^5$); MODERATE SYMPTOMS ($\mathbf{I}_2^5$); SEVERE SYMPTOMS ($\mathbf{I}_3^3$); CRITICAL SYMPTOMS ($\mathbf{I}_4^3$); HOSPITALISED ($\mathbf{I}_5^{12}$); ICU/VENTILATION ($\mathbf{I}_6^{13}$); CONTAGIOUS WITHOUT SYMPTOMS ($\mathbf{I}_7^{10}$); CONTAGIOUS WITH MODERATE SYMPTOMS ($\mathbf{I}_8^{11}$); CONTAGIOUS WITH SEVERE SYMPTOMS ($\mathbf{I}_9^4$); CONTAGIOUS WITH CRITICAL SYMPTOMS ($\mathbf{I}_{10}^5$); DECEASED ($\mathbf{R}_{11}^{\infty}$); and IMMUNE ($\mathbf{R}_{12}^{\infty}$)

configuration of the lattice, which is done by supposing that a vaccination campaign has been performed, that leads to the increase of the initial number of IMMUNE cells; its effect is then monitored by four control variables, as mentioned below.

Initially, a lattice with $n \times n$ cells in the SUSCEPTIBLE state is created. Then, a single cell is randomly chosen to be in the NO SYMPTOMS state, therefore becoming patient zero and a proportion of the remaining cells is randomly chosen to be in the IMMUNE state. In tune with the latter, we define *immunity ratio* of the lattice as the number of cells changed to the IMMUNE state initially, at time step zero, in comparison with the number of cells of the lattice. The rest of them stays in the SUSCEPTIBLE state. Since the automaton is stochastic, a number of 10 simulations is performed, each one running until convergence; along the process, the variables below are monitored, and their values averaged over the 10 simulations:

- The proportion of cells in SUSCEPTIBLE state at final convergence;
- The maximum proportion of infected cells during a run;
- The proportion of cells in DECEASED state at final convergence;
- The number of time steps counted at final convergence.

In the sequence, this procedure is repeated with increasing values of immunity ratio, i.e., the inital proportion of cells in the IMMUNE state. We rely on an immunity ratio range from 0% to 95%, with 5% steps. This procedure should be capable of allowing the analysis of the impact of the immunity ratio in the mentioned variables. Also, it enables the possibility of observing how each variable performs when immunity rises.

3.3 Estimating collective immunity with different parameters

Aiming at strengthening the model's flexibility, the contamination contributions, the durations of each state and the transition probabilities were defined as input parameters of the model, not constants. This approach was adopted to allow testing the impact of these parameters on the estimates of collective immunity. So, additionally to the procedure proposed, related to varying the initial immunity, four parameterisations are adopted to make a proper comparison of collective immunities obtained from each one.

The first one—that we refer herein as defining the computational **Experiment I**—is the same as the one we used previously in Lima and Balbi (2021) and considers the contamination contributions for each state as constants with the values proposed in Sect. 2.2.2. Also, the durations of each state follows a uniform distribution around the values shown in Fig. 2, here considered as means, $\pm 20\%$.

The second setting—defining computational **Experiment II**—increases the degree of uncertainty of the automaton. The contamination contributions are then defined by a Gaussian distribution truncated by three standard deviations with the mean value equal to the contamination contribution (item 2.2.2). The deviation considered was 3.33%, so the probabilities of contamination for each state can variate by $\pm 10\%$. The durations follow the same logic, being defined by a Gaussian distribution truncated by three standard deviations with the mean value equal to the duration given by Fig. 2. The standard deviation considered was 6.67%, so the probabilities of contamination for each state could vary by $\pm 20\%$.

The third parameterisation—related to computational **Experiment III**—expands the range of values applied by the previous one, by doubling their variances. The contamination contributions were considered with a variation of $\pm 20\%$ and the durations were considered with a variation of $\pm 40\%$.

The fourth parameterisation—related to computational **Experiment IV**—turns the variation of the remaining parameters involved in the reference model also as stochastic. So, it improves the previous experiment by defining the transition probabilities ($\mathcal{P}_t$) from NO SYMPTOMS state also as Gaussian distributions truncated by three standard deviations with the mean value equal do the probabilities shown in Fig. 2. The deviation considered as 3.33%, so the variance was $\pm 10\%$.

Table 1 presents a briefing of each parameterisation, also with the **range** of possible values regarding the distribution and variance of each experiment.

4 Results

The proposed procedure was executed and monitored, comparing the average value of each variable between the ten simulations for each immunity ratio. Considering the range of immunity proposed and the simulations performed for each immunity, as well the four experiments, a total number of 800 simulations was obtained from the model. This implementation was written in Wolfram language, native of the software *Mathematica*.

4.1 Experiment I

Figure 3 displays the impact of the immunity ratio on the population main metrics (remaining susceptibles, infection peak, deceaseds and total time steps) considering **Experiment I** (Table 1). First, this impact can be seen in the number of susceptibles at the final convergence (in other

Table 1 Parameterisations employed in the experiments

Parameter		Parameterisations							
		Experiment I		Experiment II		Experiment III		Experiment IV	
Name	Symbol	Range	PD	Range	PD	Range	PD	Range	PD
T	$\mathbf{S}_0^{f(\mathcal{P}_c)}$	∞	**U**	∞	**G**	∞	**G**	∞	**G**
	$\mathbf{I}_1^5$	[4,6]		[4,6]		[3,7]		[3,7]	
	$\mathbf{I}_2^5$	[4,6]		[4,6]		[3,7]		[3,7]	
	$\mathbf{I}_3^3$	[2,4]		[2,4]		[2,6]		[2,6]	
	$\mathbf{I}_4^3$	[2,4]		[2,4]		[2,4]		[2,4]	
	$\mathbf{I}_7^{10}$	[8,12]		[8,12]		[7,17]		[7,17]	
	$\mathbf{I}_5^{12}$	[10,14]		[10,14]		[8,20]		[8,20]	
	$\mathbf{I}_6^{13}$	[10,16]		[10,16]		[6,14]		[6,14]	
	$\mathbf{I}_8^{11}$	[9,13]		[9,13]		[9,21]		[9,21]	
	$\mathbf{I}_9^4$	[3,5]		[3,5]		[3,7]		[3,7]	
	$\mathbf{I}_{10}^5$	[4,6]		[4,6]		[2,6]		[2,6]	
	$\mathbf{R}_{11}^{\infty}$	∞		∞		∞		∞	
	$\mathbf{R}_{12}^{\infty}$	∞		∞		∞		∞	
$\mathcal{P}_c$	$\mathbf{S}_0^{f(\mathcal{P}_c)}$	0%	**n**	0%	**G**	0%	**G**	0%	**G**
	$\mathbf{I}_1^5$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_2^5$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_3^3$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_4^3$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_7^{10}$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_5^{12}$	10%		[0%,20%]		[0%,40%]		[0%,40%]	
	$\mathbf{I}_6^{13}$	10%		[0%,20%]		[0%,40%]		[0%,40%]	
	$\mathbf{I}_8^{11}$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_9^4$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{I}_{10}^5$	100%		[80%,100%]		[60%,100%]		[60%,100%]	
	$\mathbf{R}_{11}^{\infty}$	0%		0%		0%		0%	
	$\mathbf{R}_{12}^{\infty}$	0%		0%		0%		0%	
$\mathcal{P}_t$	$\mathbf{I}_1^5 \rightarrow \mathbf{I}_2^5$	30%	**n**	30%	**n**	30%	**n**	[20%,40%]	**G**
	$\mathbf{I}_1^5 \rightarrow \mathbf{I}_3^3$	55%		55%		55%		[45%,65%]	
	$\mathbf{I}_1^5 \rightarrow \mathbf{I}_4^3$	10%		10%		10%		[0%,20%]	
	$\mathbf{I}_1^5 \rightarrow \mathbf{I}_7^{10}$	5%		5%		5%		[0%,15%]	

Key to the table: **T**: Duration of the states (1 time step = 1 day); $\mathcal{P}_\mathbf{c}$: Neighbourhodd contamination probability; $\mathcal{P}_\mathbf{t}$: State transition probability; **PD**: Type of probability distribution; **U**: Uniform probability distribution; **G**: Gaussian probability distribution; **Range**: The interval from the values are chosen according to the distribution adopted, with notation [min, max]; **n**: Fixed value

words, the ones who did not get ill). In a first and brief look, it is predictable that more immune individuals means less susceptibe individuals, which describes a negative linear correlation. This effect becomes perceptible after around 60% immunity. Before 50%, the final number of susceptibles is near zero, and starts to increase as immunity rises.

Second, the infection peak is also affected by immunity, describing the peak value of active cases. There is a negative correlation between immunity ratio and the infection peak which comes near to zero when immunity achieves 65%. Last, the death rate behaves similarly to the infection peak, coming down close to zero at 65% of immunity.

Figure 4 gives an idea about the emulated pandemic duration according to immunity ratio. Apparently, ratios

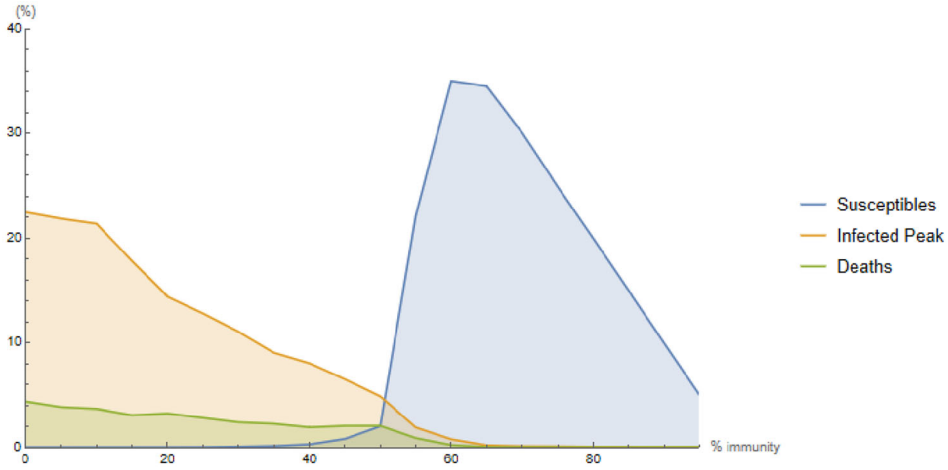

Fig. 3 Remaining susceptibles, deaths and infection peak as a function of the immunity ratio considering **Experiment I**

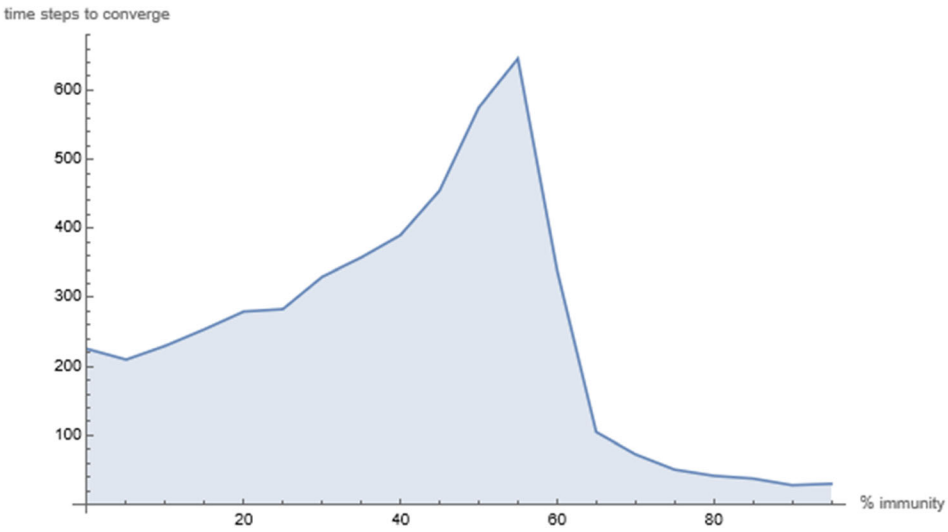

Fig. 4 Variation in the number of time steps necessary to final convergence as a function of the immunity ratio considering **Experiment I**

under 55% lead to a positive correlation, or a tendency of the pandemic to take more time to end as the immunity rises. In other words, the remaining number of susceptibles do not receive protection against contamination and the the disease takes longer to spread. The maximum average time steps taken to convergence are near 640 steps at 55%, starting to decrease for bigger immunities.

4.2 Experiment II

Figure 5 depicts the impact of immunity ratio on the population main metrics considering **Experiment II** (Table 1). The effect previously mentioned related to the decrease of susceptibles with the increase of immunes becomes perceptible after precisely 60% immunity. Before 50%, the final number of susceptibles is near zero, starting to increase as immunity rises.

The infection peak presents the same negative correlation between immunity ratio and the infection peak which comes near to zero when immunity achieves 60%. The death rate behaves similarly to the infection peak, approaching zero at 60% immunity rate.

Figure 6 presents the amount of time steps taken to convergence, with ratios under 55% presenting a positive correlation. The maximum average time steps taken to convergence are near 720 steps at 55%, starting to decrease for bigger immunities.

4.3 Experiment III

Figure 7 presents the impact of immunity ratio on the population main metrics considering **Experiment III** (Table 1). The effect previously mentioned related to the decrease of susceptibles with the increase of immunes becomes perceptible after precisely 60% immunity rate. Before 45%, the final number of susceptibles is near zero, starting to increase as immunity rises, but rising faster than in the previous experiments.

The infection peak shows the same negative correlation between immunity ratio and the infection peak, which

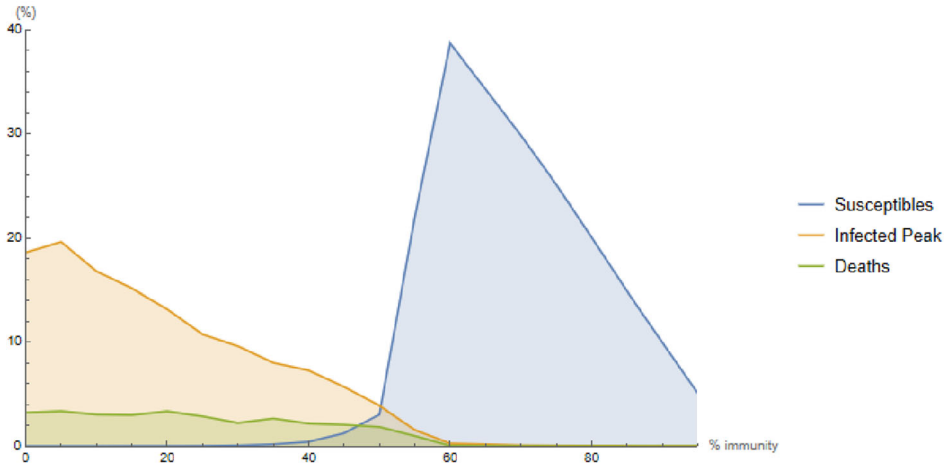

Fig. 5 Remaining susceptibles, deaths and infection peak as a function of the immunity ratio considering **Experiment II**

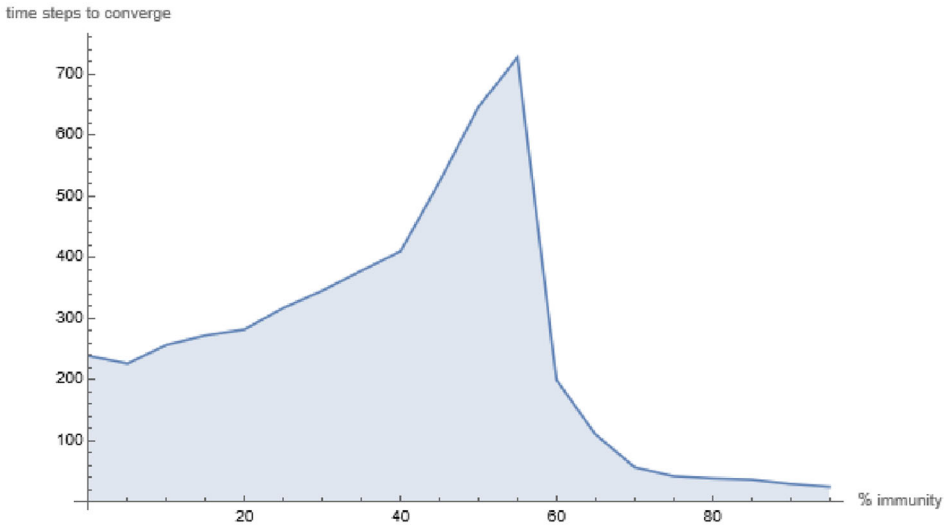

Fig. 6 Variation in the number of time steps necessary to final convergence, as a function of the immunity ratio, for **Experiment II**

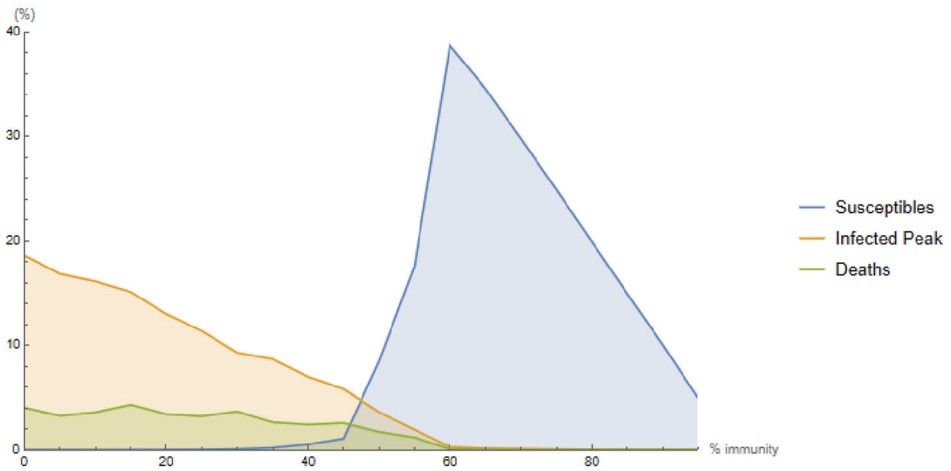

Fig. 7 Remaining susceptibles, deaths and infection peak, as a function of the immunity ratio, for **Experiment III**

comes near to zero when immunity achieves 60%. The death rate behaves similarly to the infection peak, reaching near zero at 60% immunity.

Figure 8 presents the amount of time steps taken to convergence, with ratios under 55% presenting a positive correlation. The maximum average time steps taken to

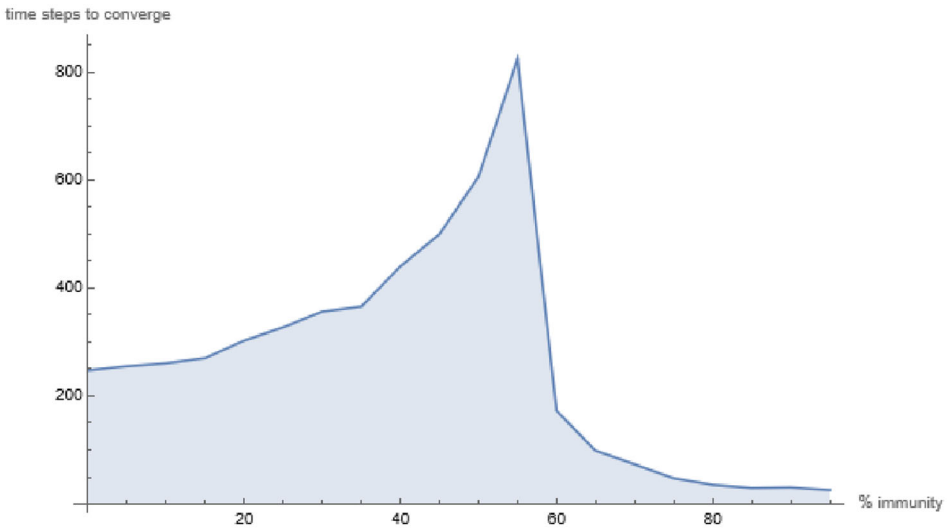

Fig. 8 Variation in the number of time steps necessary to final convergence as a function of the immunity ratio, for **Experiment III**

convergence are near 820 steps at 55%, starting to decrease for bigger immunities.

4.4 Experiment IV

Figure 9 presents the impact of immunity ratio on the population main metrics considering **Experiment IV** (Table 1). The effect previously mentioned related to the decrease of susceptibles with the increase of immunes becomes perceptible after precisely 60% immunity rate. Before 50%, the final number of susceptibles is near zero, starting to increase as immunity rises.

The infection peak shows the same negative correlation between immunity ratio and the infection peak, which approaches zero when immunity achieves 60%. The death rate behaves similarly to the infection peak, getting close to zero at 60% immunity.

Figure 10 displays the number of time steps taken to convergence, with ratios under 50% showing positive correlation. The maximum average time steps taken to convergence are near 720 steps at 50%, starting to decrease for bigger immunities.

5 Conclusions

A stochastic cellular automaton based framework was built to perform estimates of the collective immunity of COVID-19 in a given population of static individuals interacting socially only with their neighbours in a Moore neighbourhood of unitary radius. The main idea was to evaluate the impact of immune individuals, from zero to 95% of the population, in the dynamics of propagation of SARs-CoV-2, by observing the remaining susceptibles, the infection peak, the number of deaths and the time taken for the disease spread to end, thus estimating the collective immunity of COVID-19.

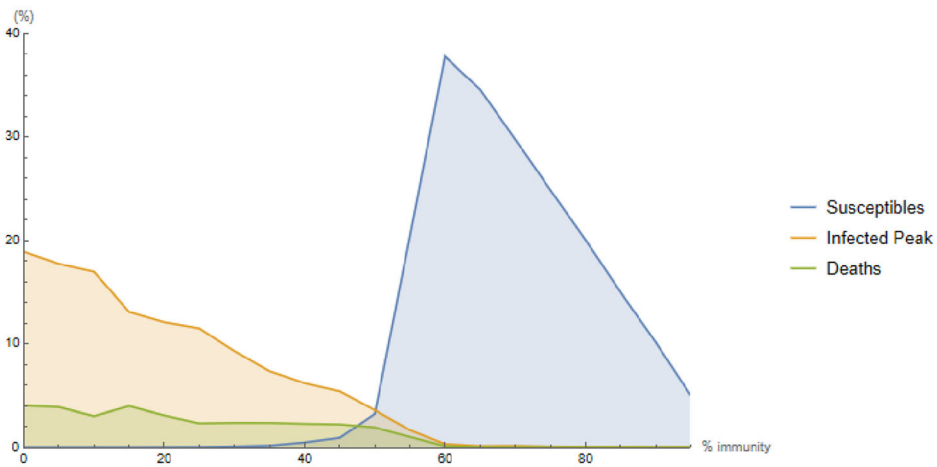

Fig. 9 Remaining susceptibles, deaths and infection peak, as a function of the immunity ratio, for **Experiment IV**

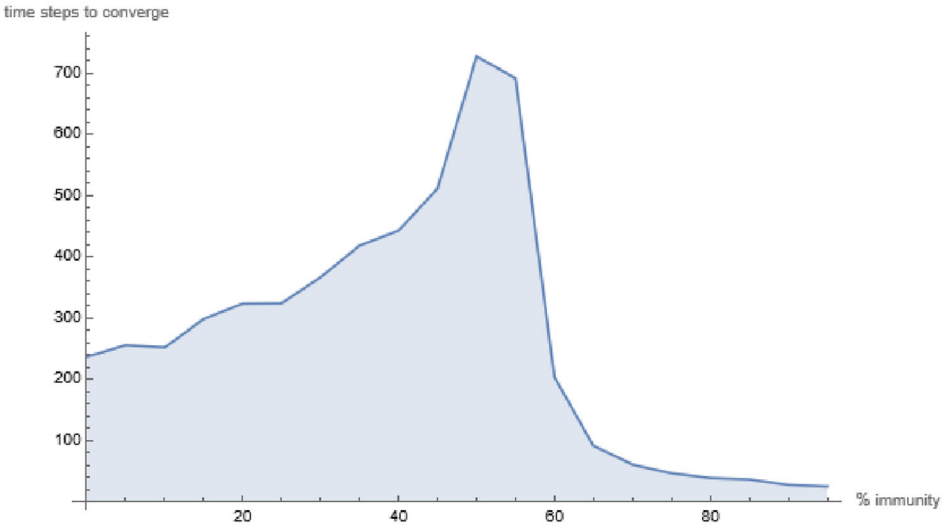

Fig. 10 Variation in the number of time steps necessary to final convergence as a function of the immunity ratio, for **Experiment IV**

Additionally, the proposed model was implemented considering some input parameters regarding the duration of the infected states of COVID-19 and the contamination contributions for these states. Four different parameterisations were applyed to the model, which allowed to compare the collective immunities:

- **Experiment I** led collective immunity to lie between 50% and 65% immunity rate. The peak of remaining susceptibles was found at 60% but the major shift in their increase began at 50%. The infection peak and the death rate approached zero at 65%. The big picture, in this case, shows collective immunity of 57.5% on average;
- **Experiment II** entailed collective immunity in the slightly smaller range between 50% and 60% immunity rate. The peak of remaining susceptibles was found at 60%, with the major shift in their increase also beginning at 50%. The infection peak and the death rate came down close to zero at 60%. Overall, this case also points collective immunity of 55% on average;
- **Experiment III** induced collective immunity to stay between 45% and 60% immunity rate. The peak of remaining susceptibles was found at 60% but the major shift in their increase started earlier than in the previous experiments, at 45%. The infection peak and the death rate went close to zero at 60%. Once again, the present case also makes it evident the onset of collective immunity of 52.5% on average.
- **Experiment IV** induced collective immunity to stay between 50% and 60% immunity rate. The peak of remaining susceptibles was found at 60%, with the major shift in their increase also beginning at 50%. The infection peak and the death rate went close to zero at 60%. So, the present case also makes it evident the onset of collective immunity at 55% on average.

So, considering the four experiments performed, the average collective immunity is 55% with a variance of $\pm 2.5\%$. The estimation of the collective immunity seems to be suitable to this framework, since it is little affected by the input parameters.

A difference between the parameterisations adopted was related to the maximum number of time steps taken to final convergence. In terms of collective immunity, all four led to a peak duration at 55% with a big decrease at 60%, but as randomness increased in the parameterisations, the emulated pandemic took more time to end, except for the last experiment, which presented a minor duration than the previous one. This could be explained by the combinations between lower contamination contributions and lower state durations, which would make it difficult for the virus to spread, while maintaining the main characteristics of the pandemic (infection peaks, death rates and maximum remaining susceptibles), but for the last experiment an additional explanation could be that the variations in the transition probabilities from the NO SYMPTOMS state somehow compensate the variations in durations and contamination contributions. So, the estimation of the pandemic duration is more affected by the input parameters, thus rendering its estimation an unsuitable application to this framework.

The calculated collective immunity at 55% $\pm 2.5\%$ rate is close to the estimated value of 60% or little more found in the literature (Aguas et al. 2020; Jones 2020), also very close to the results in Lima and Balbi (2021), whose value was 60%. But it is important to remember that the collective immunity of COVID-19 is becoming more uncertain as new variants of SARS-CoV-2 appear (García-García et al. 2022).

So, the stochastic cellular automaton architecture defined showed a fair degree of robustness regarding the

parameter values we tested. In order to further probe our approach, it is appealing to perform additional experiments considering larger variations for the contamination contributions and durations, as well some degree of asymmetry in the Gaussian distributions to improve the model sensibility in terms of time steps taken to final convergence. On the other hand, the model could be improved to cope with new variants of SARS-CoV-2 infection capability by modifying the cells neighbourhood size and distribution.

Another aspect that can be improved in the model is changing the IMMUNE ($\mathbf{R}_{12}^{\infty}$) state from a sink state to a time-dependent state, considering that COVID-19 immunity has been proven not to be permanent, so that re-infections are possible, approximating the approach to a SIRS model (Fuentes and Kuperman 1999). This could be addressed by creating new state transitions back to the SUSCEPTIBLE ($\mathbf{S}_0^{f(\mathcal{P}_c)}$) state.

So, considering all above, it is fairer and more realistic to consider the results we presented as more related to the original, pre-delta variants of the SARS-CoV-2. However, as we argued, the model itself can be easily modified so as to account also for the more recent and important variants, such as omicron (Callaway 2022; Willyard 2022).

The chosen number of simulations per immunity ratio, 10, also seemed few to average the sthocastic simulations, but this choice is directly related to the computational effort to perform all the procedure (10 simulations for each immunity ratio), which takes almost four hours to be completed. More simulations to average the outputs will certainly require more computational capability or optimization of the procedure, such like a larger immunity interval, for example 10% instead of 5%.

Finally, the cellular atomaton has regularities which decreases its reliability in emulating the spreading of COVID-19, such like the size and distribution of the neighbourhood and the immobility of the cells. This could be adapted by performing random connections between cells, instead of the rigid Moore neighbourhood with unitary radius, like small-world type networks (Schimit and Monteiro 2009); and representing the mobility of the individuals in the population, thus increasing the parameters of the model (Sirakoulis et al. 2000). These are the directions we are taking.

Acknowledgements This work was supported by research grants provided to P.P.B. by CAPES (STIC-AmSud CoDANet project, number 88881.197456/2018-01) and by CNPq-PQ 305199/2019-6. We also thank Ramy J. Rahmé for providing us with the chart in Fig. 1, which is an improved version of the one he originally posted in March 2020 in Twitter, which, in turn, is an improved version of another chart due to Pierre Paperon, that had just been published then in LinkedIn.

References

Aguas R, Corder RM, King JG, Goncalves G, Ferreira MU, Gomes MGM (2020) Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. https://doi.org/10.1101/2020.07.23.20160762

Ahmed E, Agiza H (1998) On modeling epidemics including latency, incubation and variable susceptibility. Physica A 253(1):347–352. https://doi.org/10.1016/S0378-4371(97)00665-1

Blavatska V, Holovatch Y (2021) Spreading processes in post-epidemic environments. Physica A: Stat Mech Appl 573:125980. https://doi.org/10.1016/j.physa.2021.125980

Britton T, Ball F, Trapman P (2020) A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 369(6505):846–849. https://doi.org/10.1126/science.abc6810

Callaway E (2022) Why does the omicron sub-variant spread faster than the original? Nature 602(7898):556–557. https://doi.org/10.1038/d41586-022-00471-2

Ciotti M, Ciccozzi M, Terrinoni A, Jiang WC, Wang CB, Bernardini S (2020) The COVID-19 pandemic. Crit Rev Clin Lab Sci 57(6):365–388. https://doi.org/10.1080/10408363.2020.1783198

Ezhilan M, Suresh I, Nesakumar N (2021) SARS-CoV, MERS-CoV and SARS-CoV-2: a diagnostic challenge. Measurement 168:108335. https://doi.org/10.1016/j.measurement.2020.108335

Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunuba Perez Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell L, Van Elsland S, Thompson H, Verity R, Volz E, Wang H, Wang Y, Walker P, Winskill P, Whittaker C, Donnelly C, Riley S, Ghani A (2020) Report 9: impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. https://doi.org/10.25561/77482

Fuentes M, Kuperman M (1999) Cellular automata and epidemiological models with spatial dependence. Physica A 267(3):471–486. https://doi.org/10.1016/S0378-4371(99)00027-8

García-García D, Morales E, Fonfría ES, Vigo I, Bordehore C (2022) Caveats on COVID-19 herd immunity threshold: the spain case. Sci Rep 12(1). https://doi.org/10.1038/s41598-021-04440-z

Holko A, Mdrek M, Pastuszak Z, Phusavat K (2016) Epidemiological modeling with a population density map-based cellular automata simulation system. Expert Syst Appl 48:1–8. https://doi.org/10.1016/j.eswa.2015.08.018

Jones F (2020) The uncertainties about herd immunity (https://revistapesquisa.fapesp.br/as-incertezas-sobre-a-imunidade-coletiva). FAPESP Research (In Portuguese.)

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J (2020) The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application. Ann Intern Med 172(9):577–582. https://doi.org/10.7326/M20-0504 (**PMID: 32150748**)

Lima I, Balbi PP (2021) Preliminaries on a stochastic cellular automaton based framework for studying the population dynamics of covid-19. In: Gwizdałła TM, Manzoni L, Sirakoulis GC, Bandini S, Podlaski K (eds) Cellular automata, pp 265–273. Springer, Cham. https://doi.org/10.1007/978-3-030-69480-7_27

Liu Y, Yan LM, Wan L, Xiang TX, Le A, Liu JM, Peiris M, Poon LLM, Zhang W (2020) Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis 20(6):656–657. https://doi.org/10.1016/s1473-3099(20)30232-2

Mikler AR, Venkatachalam S, Abbas K (2005) Modeling infectious diseases using global stochastic cellular automata. J Biol Syst 13(04):421–439. https://doi.org/10.1142/s0218339005001604

Monteiro L, Gandini D, Schimit P (2020) The influence of immune individuals in disease spread evaluated by cellular automaton and genetic algorithm. Comput Methods and Programs Biomed 196:105707. https://doi.org/10.1016/j.cmpb.2020.105707

Rahmé RJ (2020) COVID19 chart: timeline, progression and mortality by severity

Randolph HE, Barreiro LB (2020) Herd immunity: understanding COVID-19. Immunity 52(5):737–741. https://doi.org/10.1016/j.immuni.2020.04.012

Schimit P (2021) A model based on cellular automata to estimate the social isolation impact on covid-19 spreading in brazil. Comput Methods Programs Biomed 200:105832. https://doi.org/10.1016/j.cmpb.2020.105832

Schimit P, Monteiro L (2009) On the basic reproduction number and the topological properties of the contact network: An epidemiological study in mainly locally connected cellular automata. Ecol Model 220(7):1034–1042. https://doi.org/10.1016/j.ecolmodel.2009.01.014

Schimit P, Monteiro L (2011) A vaccination game based on public health actions and personal decisions. Ecol Model 222(9):1651–1655. https://doi.org/10.1016/j.ecolmodel.2011.02.019

Sirakoulis G, Karafyllidis I, Thanailakis A (2000) A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. Ecol Model 133:209–223. https://doi.org/10.1016/S0304-3800(00)00294-5

Velavan TP, Meyer CG (2020) The COVID-19 epidemic. Trop Med Int Health 25(3):278–280. https://doi.org/10.1111/tmi.13383

White SH, del Rey AM, Sánchez GR (2007) Modeling epidemics using cellular automata. Appl Math Comput 186(1):193–202. https://doi.org/10.1016/j.amc.2006.06.126

Willyard C (2022) What the omicron wave is revealing about human immunity. Nature 602(7895):22–25. https://doi.org/10.1038/d41586-022-00214-3

Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, Niemeyer D, Jones TC, Vollmar P, Rothe C, Hoelscher M, Bleicker T, Brünink S, Schneider J, Ehmann R, Zwirglmaier K, Drosten C, Wendtner C (2020) Virological assessment of hospitalized patients with covid-2019. Nature 581(7809):465–469. https://doi.org/10.1038/s41586-020-2196-x

Yakowitz S, Gani J, Hayes R (1990) Cellular automaton modeling of epidemics. Appl Math Comput 40(1):41–54. https://doi.org/10.1016/0096-3003(90)90097-M