Epidemics 38 (2022) 100544

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

# Modeling the effects of social distancing on the large-scale spreading of diseases

Paulo Cesar Ventura [a,*], Alberto Aleta [b], Francisco Aparecido Rodrigues [c], Yamir Moreno [b,d,e]

[a] Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil
[b] ISI Foundation, Turin, Italy
[c] Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, SP, Brazil
[d] Institute for Biocomputation and Physics of Complex Systems (BIFI), 50018 Zaragoza, Spain
[e] Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain

ARTICLE INFO

Dataset link: https://doi.org/10.5281/zenodo.6258445

*Keywords:*
Epidemiology
Social distancing
Metapopulation model
Non-pharmaceutical interventions

ABSTRACT

To contain the propagation of emerging diseases that are transmissible from human to human, non-pharmaceutical interventions (NPIs) aimed at reducing the interactions between humans are usually implemented. One example of the latter kind of measures is social distancing, which can be either policy-driven or can arise endogenously in the population as a consequence of the fear of infection. However, if NPIs are lifted before the population reaches herd immunity, further re-introductions of the pathogen would lead to secondary infections. Here we study the effects of different social distancing schemes on the large scale spreading of diseases. Specifically, we generalize metapopulation models to include social distancing mechanisms at the subpopulation level and model short- and long-term strategies that are fed with local or global information about the epidemics. We show that different model ingredients might lead to very diverse outcomes in different subpopulations. Our results suggest that there is not a unique answer to the question of whether contention measures are more efficient if implemented and managed locally or globally and that model outcomes depends on how the full complexity of human interactions is taken into account.

## 1. Introduction

The spreading of infectious diseases is a complex process involving two main aspects. On the one hand, the spreading capabilities of the pathogen depend on its biological properties, the characteristics of the host, and the many environmental factors that can play an important role. On the other hand, the pathogen (except for vector-borne diseases) can only spread if two hosts have some kind of contact. As such, the behavior of the host is a key element in the study of epidemic spreading (Funk et al., 2010; Read et al., 2012). For this reason, many researches have focused on studying how animals and humans interact, in order to inform mathematical models and produce better forecasts (Wilson-Aggarwal et al., 2019; Mossong et al., 2008; Arregui et al., 2018; Salathé et al., 2010).

Despite having acknowledged the role that behavior plays in the spreading of infectious diseases, epidemic models usually neglect the possibility that hosts will change their behavior due to an ongoing outbreak (Eksin et al., 2019). An exception to this are awareness models, in which both an epidemic and information spread at the same time in the population and host reacts accordingly (see da Silva et al. (2019) and the references therein). These behavioral changes have been observed both in animal (Stroeymeyt et al., 2018) and human societies (SteelFisher et al., 2009). More recently, the spreading of COVID-19 has clearly demonstrated the variety of ways in which humans react to an epidemic (Coletti et al., 2020; Jarvis et al., 2020; Zhang et al., 2020; Feehan and Mahmud, 2021; Woskie et al., 2021). For instance, without forceful government intervention, traffic in Seoul's subway declined sharply following the first deaths in South Korea (Jewel, 2020). Conversely, several communities in other regions of the world defied social distancing measures (Waitzberg et al., 2020; Gollwitzer et al., 2020).

Public health authorities base many decisions on the forecasts produced by epidemic models. It is thus of paramount importance to include the effect of behavioral changes as something inherently attached to the spreading of the epidemic (Eksin et al., 2019). Of particular interest in this regard are metapopulation models. These models provide a simple way of incorporating the spatial heterogeneity of human societies, while keeping them mathematically tractable. In essence, in a metapopulation model hosts are grouped in different subpopulations,

* Corresponding author.
*E-mail address:* paulo.pc.vs@gmail.com (P.C. Ventura).

https://doi.org/10.1016/j.epidem.2022.100544
Received 21 May 2021; Received in revised form 21 December 2021; Accepted 9 February 2022
Available online 21 February 2022

**Table 1**
List of symbols and acronyms used throughout the paper.

| Description | Symbol |
|---|---|
| Individual infection probability/rate (1/day) | $\beta$ |
| Recovery probability/rate (1/day) | $\mu$ |
| Basic reproduction number | $R_0 = \beta/\mu$ |
| Contact reduction coefficient | $a_i$ |
| Response strength exponent | $k$ |
| Memory coefficient | $l$ |
| Mobility coefficient | $\tau$ |
| Threshold for infectious fraction | $\epsilon_I$ |
| Individual counts in each state in subpopulation $i$ | $S_i, I_i, R_i$ |
| Global individual counts in each state | $N_S, N_I, N_R$ |
| Global fractions in each state | $\rho_S, \rho_I, \rho_R$ |
| Number of nodes (subpopulations) on network | $V$ |
| Number of links on network | $M$ |
| Number of individuals in subpopulation $i$ | $N_i$ |
| Total number of individuals in population | $N_{pop}$ |
| Maximum distance for connection on RGN | $d$ |
| Link weight between subpopulations $i$ and $j$ | $T_{ij}$ |
| Long-term memory scheme | LT |
| Short-term memory scheme | ST |
| Network | RGN |

and the exchange of individuals between subpopulations is governed by certain rules. Within each subpopulation it is possible to assume that the spreading from host to host follows the classical homogeneous mixing approach (Lloyd and May, 1996; Colizza and Vespignani, 2008; Aleta et al., 2017), or include more details on individual heterogeneities through different techniques (Ajelli et al., 2010; Apolloni et al., 2014).

These models have been extensively used during the emergence of new pathogens to study how an outbreak might propagate globally (Bajardi et al., 2009). In the particular case of COVID-19, they have also been used to study the early spreading of the disease across countries (Chinazzi et al., 2019), but also within countries (Costa et al., 2020; Aleta et al., 2020; Aleta and Moreno, 2020). Yet, these models usually focus on the effect of varying the rules governing the flux of individuals between subpopulations, or on globally modifying the transmission due to the introduction of public health interventions. In this paper, our objective is to understand the role that social distancing can play in these models and shed some light on the effect that it can have on the spatial spreading of epidemics when the response is heterogeneous across regions.

## 2. Model description

We implement a discrete and stochastic SIR-metapopulation model composed by $V$ subpopulations (Keeling et al., 2004; Ball et al., 2015; Wang and Li, 2014). In each subpopulation, individuals can interact with each other and spread the disease following the classical SIR model. These subpopulations are connected through a certain network, so that an individual may travel to another subpopulation only if it exists a direct link between the source and target subpopulations.

To include social distancing effect, either policy-driven or as a consequence of fear of infection, we consider an additional coefficient that modifies the transmissibility of individuals. This coefficient mimics the behavioral responses of the population, and evolves according to the densities of infectious and recovered individuals (Eksin et al., 2019). In Section 3, we study this model in an homogeneously mixed population. For the remaining sections, we apply the model to a metapopulation system.

### 2.1. Epidemic spreading

For the SIR compartmental model, individuals are assigned compartments according to their infectious status: susceptible (S) if they do not have the disease and can catch it; infectious (I) when they have the disease and can transmit it to susceptibles, and removed (R) when they no longer transmit the disease after being infectious (either by recovery or deceasing). At each model's epidemic update, the following rules determine the transitions between compartments: (See Table 1).

- **S → I**: a susceptible individual in subpopulation $i$ can become infectious with probability $P^i_{S\to I} = 1 - (1 - a_i\beta/N_i)^{I_i}$, where $\beta$ is the individual transmission probability, $N_i$ is the number of individuals in subpopulation $i$ and $I_i$ is the number of infectious individuals in the same subpopulation. The term $a_i$ is called the *coefficient of contact reduction*, and depends on the considered memory scheme as we describe later on.
- **I → R**: an infectious individual is moved to the R (removed) compartment with probability $P^i_{I\to R} = \mu$, which is the inverse of the average infectious period. After this event, the individual no longer participates on the epidemic dynamics.

The number of new infections and removals are both calculated through binomial sampling, with distributions $B(S_i, P_i(S \to I))$ for infections and $B(I_i, \mu)$ for removals. Throughout this paper, we set $\mu$ to $1/4$ and the basic reproductive number $R_0 = \beta/\mu$ to 1.5, which is compatible with the parameters of an influenza-like disease. In the Supplementary Material we report also results with $R_0 = 1.2$ and $R_0 = 2.5$, the latter being compatible with the spreading of COVID-19 in 2020. We define $N_S = \sum_{i=1}^{V} S_i$ as the total number of susceptible individuals in the whole population and, analogously, $N_I$ and $N_R$ for the infectious and removed compartments. The coefficient of contact reduction $a_i$ is what determines the behavioral responses to the epidemics, and we consider different scenarios (memory schemes) for such a response, each one with a different definition of $a_i$:

1. **Social distancing based on global information**
   This scenario (or strategy) is based on the one proposed by Eksin et al. in Eksin et al. (2019), which here we extend to metapopulations. The coefficient of contact reduction emulates the social response to an increase in the number of infections, and it is a function of the total (global) number of infected and recovered individuals, given by:
   $$a_i = \left(1 - \frac{(N_I + l \cdot N_R)}{N}\right)^k \tag{1}$$
   Since $a_i$ only depends on global quantities, its value is the same for all subpopulations. The parameter $k$ is the *response strength*, an adjustable exponent that calibrates the intensity of the response against the number of cases. The coefficient $l$ is called *memory coefficient*, and determines the importance of the R compartment to the behavioral response. Following Eksin et al. (2019), $l = 0$ represents the *short-term memory scheme* (ST), where the response weakens when the incidence drops. The case $l = 1$ is the *long-term memory scheme* (LT), in which the awareness is proportional to the total accumulated number of cases. These two limiting cases differ in many aspects, which we describe throughout the rest of the paper.
2. **Social distancing based on local information**
   In this variant of the previous scenario, we consider that each subpopulation (node) responds individually, according to its own number of cases. The expression for the coefficient of contact reduction is:
   $$a_i = \left(1 - \frac{(I_i + l \cdot R_i)}{N_i}\right)^k \tag{2}$$
3. **Constant response after threshold**
   For comparison, we also implement a more traditional scenario in which the transmissibility is reduced by a constant factor $a_0$ once the overall number of cases $N_I + N_R$ overcomes a given

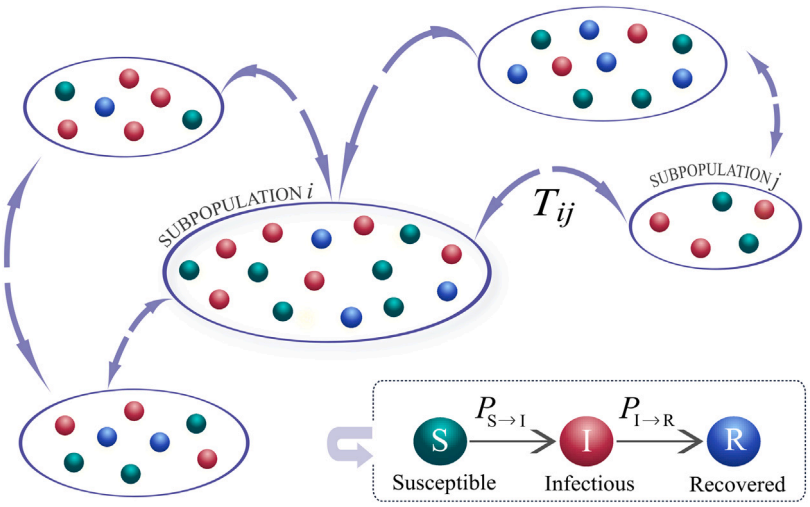


**Fig. 1.** Scheme of the epidemic and mobility models in a metapopulation. The SIR epidemic dynamics occur inside each subpopulation, where homogeneous mixing is assumed. Also at each time step, individuals move between neighboring subpopulations $i$ and $j$ according to a mobility matrix $T_{ij}$.

threshold $N \cdot \varepsilon_I$ (with $0 \leq \varepsilon_I < 1$). Here we use only the global cumulative information, thus $a_i$ is defined as:

$$a_i = \begin{cases} 1 & N_I + N_R < N\varepsilon_I \\ a_0 & N_I + N_R \geq N\varepsilon_I \end{cases} \tag{3}$$

and is the same for all subpopulations. This scenario can be considered as a baseline strategy for modeling, in which social distancing (or governments policies) is constant and independent of the state of the system. This is often assumed as the behavioral response in metapopulation epidemic models. (Perra et al., 2011; Aleta and Moreno, 2020; Calvetti et al., 2020; Manfredi and D'Onofrio, 2013).

The choice of $N_I$ and $N_R$ (or $I_i$ and $R_i$ for local strategies) to guide the intensity of contention measures is justified by the simplicity of this approach. Other metrics, such as the number of new infections or removals, could be used as well, but as these are subject to considerable stochastic fluctuations, additional procedures would be required to prevent the contention measures to unrealistically oscillate over a few steps.

Throughout the paper, we use the words "memory" or "memory scheme" for the different setups of the memory coefficient, namely LT ($l = 1$) and ST ($l = 0$), while the word "strategy" refers to the local or global use of information, as explained above.

### 2.2. Mobility model

Following a standard metapopulation framework, the mobility between subpopulations is modeled as a random walk through the links of a graph of $V$ nodes and $M$ links. It is controlled by a master parameter $\tau$ called the *mobility coefficient*, as well as the weights $T_{ij}$ of existing links between subpopulations (nodes) $i$ and $j$. During a mobility update, each individual in subpopulation $i$ travels to subpopulation $j$ with probability $p_{ij} = \tau \cdot T_{ij}/N_i(0)$, where $N_i(0)$ is the number of individuals attributed to subpopulation $i$ at the initial time step. The number of individuals traveling from $i$ to each neighbor $j$ is calculated through multinomial sampling for each state ($S$, $I$ and $R$), with number of trials given by the numbers of individuals ($S_i$, $I_i$ and $R_i$) and probabilities $p_{ij}$. Within this scheme, we have the following features:

- The average number of individuals expected to travel from $i$ to $j$ is $\tau \cdot T_{ij}$.
- If all links are reciprocal and symmetric (i.e., $T_{ji} = T_{ij}$ for all connected pairs of subpopulations $i, j$), then the net fluxes between the subpopulations are balanced and the population of each one remains approximately constant, fluctuating around $N_i(0)$. We use this configuration throughout the paper.
- The probability that an individual in subpopulation $i$ travels anywhere is $p_i = \tau \sum_j T_{ij}/N_i(0)$, and may vary for each subpopulation. We always set $\tau$ to be small enough so that $p_i < 1$ for every subpopulation $i$.

During a single time step, we first perform the epidemic interactions in each subpopulation, then update the number of individuals that are in each state. After this, we apply the mobility rules to determine how many individuals move through each link, and update the actual numbers only after all fluxes have been calculated. This way, the results do not depend on the order at which we "visit" the subpopulations to perform the calculations. The mobility and epidemic models are schematically represented in Fig. 1.

## 3. Analytical insights for homogeneously mixed populations

For sufficiently low mobility between subpopulations, the local dynamics can be well described by an isolated homogeneously mixed system. Also for sufficiently high number of individuals, we can use rate equations for the expected fractions of the population in each compartment, substituting $S/N$, $I/N$ and $R/N$ by $\rho_S$, $\rho_I$ and $\rho_R$, respectively. The chosen expression for the behavioral response mechanism is simple enough to allow for analytical manipulation. Particularly, in the long-term (LT) memory scheme (i.e., with $l = 1$ in Eqs. (1) and (2)), these rate equations can be integrated to find the trajectory of the dynamical system. In this section, we explore this tractability to extract some insights from the model.

### 3.1. Long-term memory

Considering the dynamics of a single isolated subpopulation, the dynamical equations for the average fractions $\rho_S$, $\rho_I$ and $\rho_R$ of susceptible, infectious and removed individuals can be written as:

$$\dot{\rho_S} = -\mu R_0 \rho_S \rho_I \cdot a(\rho_I, \rho_R) \tag{4}$$

$$\dot{\rho_I} = \mu R_0 \rho_S \rho_I \cdot a(\rho_I, \rho_R) - \mu \rho_I \tag{5}$$

$$\dot{\rho_R} = \mu \rho_I, \tag{6}$$

where $R_0 = \beta/\mu$ is the basic reproduction number and $\mu$ is now interpreted as the recovery rate (rather than probability). For the long-term memory, $a(\rho_I, \rho_R) = (1-(\rho_I+\rho_R))^k = \rho_S^k$, given that $\rho_S+\rho_I+\rho_R = 1$. As described in Eksin et al. (2019), we can divide Eq. (4) by Eq. (6) to obtain a separable differential equation:

$$\frac{\mathrm{d}\rho_S}{\mathrm{d}\rho_R} = -R_0\rho_S^{k+1} \tag{7}$$

For $k > 0$ and assuming that $\rho_R(0) = \rho_0 \geq 0$ and $\rho_I(0) = \delta \to 0$, its solution is given by:

$$\rho_S = \left[\frac{1}{(1-\rho_0)^k} + kR_0(\rho_R - \rho_0)\right]^{-\frac{1}{k}} \tag{8}$$

Alternatively, writing $\rho_I$ as a function of $\rho_R$:

$$\rho_I = 1 - \rho_R - \left[\frac{1}{(1-\rho_0)^k} + kR_0(\rho_R - \rho_0)\right]^{-\frac{1}{k}} \tag{9}$$

Eq. (9) represents the trajectory of the system in a $\rho_I$ *vs* $\rho_R$ diagram, showing how the fraction of infectious individuals peaks as the prevalence increases in the LT memory. Also, the final $\rho_R$ prevalence at the end of the outbreak can be found by solving Eq. (9) for $\rho_I = 0$ besides the trivial solution $\rho_R = \rho_0$. The solid lines in Fig. 2 show such trajectories for different values of the response strength $k$, $\rho_0 = 0$ and $R_0 = 1.5$.

For $k = 0$, Eq. (7) solves as a simple SIR model which, for $\rho_R(0) = \rho_0$ and $\rho_I(0) = \delta \to 0$, leads to a $\rho_I$ *vs* $\rho_R$ trajectory given by:

$$\rho_I = 1 - \rho_R - (1-\rho_0)e^{-R_0(\rho_R-\rho_0)} \tag{10}$$

In the long-term memory with $k > 0$, the contact reduction coefficient $a(\rho_I, \rho_R)$ is a decreasing function of $\rho_R + \rho_I$, which in turn can only increase over time. This means that $a(\rho_I, \rho_R)$ will also always decrease as if the disease contention measures are held forever. This is not reasonable to assume for the long term behavior of a real situation and, for this reason, we can consider that at some point after the main outbreak, $a(\rho_I, \rho_R)$ is set to 1 again, which we call henceforth a *system release* (representing a *release* of the contention measures).

For instance, consider that a system release occurs when the fraction of infectious individuals reaches a small value $\delta > 0$ after the main outbreak initiated. This comprises the fact that, during the time evolution given by Eqs. (4) to (6), $\rho_I(t)$ tends to but never actually reaches zero. For the calculation of the $\rho_I$ *vs* $\rho_R$ trajectories though, we can consider $\delta \to 0$. After the system is released, the trajectory follows Eq. (10) with $\rho_0$ set to the final size of the first outbreak. In Fig. 2, the dashed lines represent the secondary outbreaks produced after the system's release.

As observed in the figure, a system release can generate a second outbreak which, for higher values of $k$, can be greater than the first one. This happens because, although the long-term memory scheme effectively reduces the size of the first outbreak (both in $\rho_R$ and $\rho_I$), it may leave the system under its herd immunity threshold, and thus vulnerable to new outbreaks (Lu et al., 2021). In this homogeneous model, the threshold for herd immunity is given by $\rho_{herd} = 1 - R_0^{-1}$, meaning that for $\rho_R \geq \rho_{herd}$ the fraction of infectious individuals can no longer increase. If the disease dies out before the prevalence reaches that value, the system will experience a new outbreak if the social distancing measures are relaxed and the disease is reintroduced in the population. Note that the critical value of $k$ for reaching herd immunity at the first outbreak is $k = 1$, for which the final outbreak size is exactly solved as $\rho_R^* = 1 - R_0^{-1}$, as noticed by Eksin et al. (2019).

### 3.2. System resets

Besides *releasing* the system after the first outbreak, we propose another way to work with multiple outbreaks in the LT memory scheme. What makes the $a(\rho_I, \rho_R)$ coefficient to be by default strictly decreasing is its dependence on $\rho_R$, which holds a "long-term memory". We can modify and possibly reset such memory by subtracting a constant

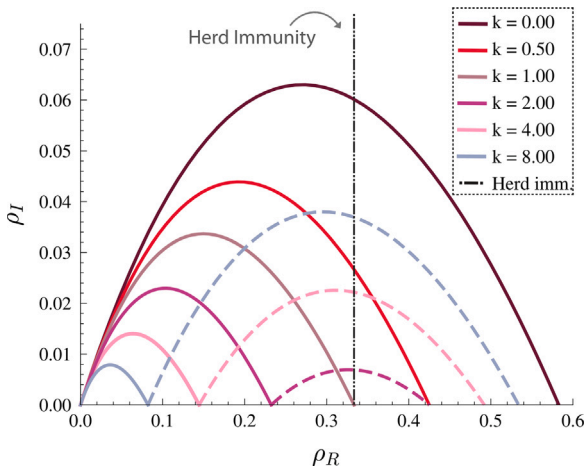


**Fig. 2.** Analytical trajectories of the homogeneously mixed long-term (LT) memory with system release after the first outbreak (i.e., complete lift of the contention measures). The trajectories are combinations of Eqs. (9) (solid lines) and (10) (dashed lines) with $\rho_0$ set to the size of the first outbreak. $R_0$ is set to 1.5.

$\rho_0$ from the prevalence that is considered for the contact reduction coefficient $a(\rho_I, \rho_R)$. If we do this right after the main outbreak and set $\rho_0$ to the recovered fraction at that time, then $a(\rho_I, \rho_R)$ is momentarily reset to 1 (i.e., no contention measures), but the system will still react in the event of another outbreak. We call this a *system reset* (the memory of the population is reset, but the social distancing measures still apply), in contrast with the simpler *system release* described in the previous section (in which the contention measures are completely removed). We are again assuming that $\rho_I$ is arbitrarily small at the end of an outbreak.

During this new round of the dynamics, Eq. (4) divided by Eq. (6) yields the following differential equation:

$$\frac{\mathrm{d}\rho_S}{\mathrm{d}\rho_R} = -R_0 \cdot (\rho_S - \rho_0)^k \rho_S \tag{11}$$

which is still separable, but for $\rho_0 > 0$ solves into a less insightful expression:

$$R_0 \cdot (\rho_R - \rho_0) = P_{\rho_0,k}(\rho_S) \tag{12}$$

where we define:

$$P_{\rho_0,k}(\rho_S) = \int_{\rho_S}^{1-\rho_0} \frac{\mathrm{d}u}{u(u+\rho_0)^k} = \\ = -\frac{1}{k(u+\rho_0)^k}\left(1+\frac{\rho_0}{u}\right)^k {}_2F_1\left(k,k;k+1,-\frac{\rho_0}{u}\right)\Bigg|_{\rho_S}^{1-\rho_0} \tag{13}$$

where ${}_2F_1$ is the Gaussian hypergeometric function. Eqs. (12) and (13) provide $\rho_R$ for any given $\rho_S$. For each of these values, $\rho_I$ is calculated as $\rho_I = 1 - \rho_S - \rho_R$, which finally allows the construction of a $\rho_I$ *vs* $\rho_R$ trajectory.

Fig. 3 shows the trajectories of the model with system resets, for different values of $k$. The first outbreak of each simulation run is represented by solid lines, and are traced using Eq. (9). For the runs that did not achieve the herd immunity threshold (which is shown as a black dot-dashed line), subsequent outbreaks are represented by dashed lines and traced using Eqs. (12) and (13). In this scenario, stricter contention measures (that is, higher $k$) always cause smaller infectious peaks, but can generate more secondary outbreaks, making it more difficult to comply with the measures. The inset of Fig. 3 shows in more detail the secondary outbreaks for $k = 2, 4$ and 8, where we can see that there are multiple outbreaks with decreasing peak size.

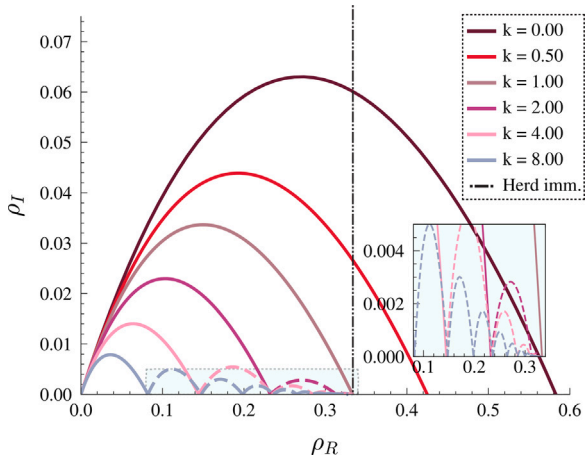


**Fig. 3.** Analytical trajectories of the homogeneously mixed long-term (LT) memory scheme with system resets. The first outbreak (solid lines) follows Eq. (9), while subsequent outbreaks (dashed lines) are traced using Eqs. (12) and (13). The inset shows a sub-region of the main plot (indicated by a blue dotted rectangle) where secondary outbreaks are more easily visible, showing how there may be multiple and successively smaller peaks. $R_0$ is set to 1.5.

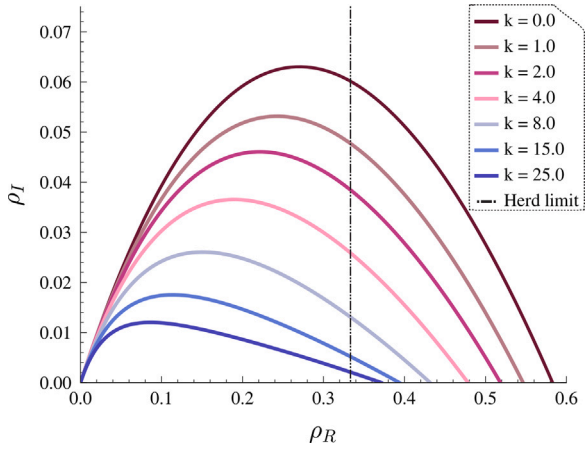


**Fig. 4.** Numerical trajectories of the homogeneously mixed short-term (ST) memory, obtained from Runge–Kutta integration of Eqs. (4) to (6) starting with $\rho_I(0) = 1 \cdot 10^{-5}$ and $\rho_R(0) = 0$. $R_0$ is set to 1.5. Unlike the LT memory, the system always reaches the herd immunity condition in a single outbreak, without considering any modification to the baseline model.

### 3.3. Short-term memory

For the short-term (ST) memory scheme, the coefficient of contact reduction depends only on $\rho_I$, being written as $a(\rho_I) = (1 - \rho_I)^k = (\rho_R + \rho_S)^k$. This breaks the separability of the differential equation obtained by dividing Eqs. (4) and (6), leaving no simple method to solve it for arbitrary values of $k$. We can still use a classic Runge–Kutta of order 4(5) (this is, of order 4 with error correction of order 5) (Dormand and Prince, 1980) to integrate the equations in time, and plot the $\rho_I$ *vs* $\rho_R$ trajectories.

Fig. 4 shows the numerically solved trajectories of the system for different values of $k$. The main difference with respect to the LT memory is that the system eventually reaches the herd immunity condition within the first and only outbreak. Higher values of $k$ reduce the size of the peak in $\rho_I$, at the expense of larger time-span in which the contention measures have to be sustained. This cannot be visualized in Fig. 4 as the time parameter is implicit, but we further explore the interplay between peak size and time span in Section 5.1. Note also that even for very large values of $k$ the final fraction of removed individuals is beyond the herd immunity threshold.

## 4. Global strategies and the heterogeneity of local features

The previous analysis has shown that, even if local extinction might be achievable with social distancing, it renders the system vulnerable to further reintroductions of the pathogen. Thus, if mobility between subpopulations is allowed, there could be spill overs from those with still ongoing outbreaks to those which contained the epidemic. To study this type of events, we now focus on the proposed model in a metapopulation using Monte Carlo simulations, with the algorithm described in Section 2. For the metapopulation, we use a random geometric network (RGN) of $V = 50$ subpopulations, constructed in a square space of length 1 and connecting subpopulations that are closer than $d = 0.25$. At this initial point, all links are unweighted and reciprocal. This gives an expected average degree of $V\pi d^2 \approx 9.82$, though the specific realization we used through this paper has an average degree of 7.2. We choose such a network topology because it emulates the spatial distribution of cities in small to mid scale, where size hierarchy and long-range links are not very present. In this comparison, each subpopulation represents either a city or a small set of cities (such as a province), inside which homogeneous mixing is assumed.

Once this unweighted graph is constructed, we set the initial population of each subpopulation $i$ proportional to its degree $k_i$, according to $N_i(0) = \lfloor N_{pop} k_i/(2M) \rfloor$, where we set $N_{pop} = 10^7$ and $M$ is the number of links of the network. This makes the overall population to be not exactly but very close to $N_{pop}$, only deviated due to truncation. Then we set the weights of existing links between subpopulations $i$ and $j$ as $T_{ij} = N_i(0)N_j(0)/N_{pop} = T_{ji}$. Within this scheme, the local population sizes fluctuate around $N(0)$ over time, as explained before. Moreover, the most connected subpopulations are also most populous ones, though the RGN is reasonably homogeneous in this sense, in comparison with a scale-free topology for example. For consistency of the results, we also use a fixed subpopulation as the seed of the disease, seeding 10 infectious individuals at the beginning of the simulations, with all other subpopulations starting with susceptibles only. The seeded subpopulation was chosen to be around the center of the square space.

### 4.1. Long-term (LT) global strategy

For the long-term (LT) memory with global information strategy, we can compare the outcomes of the simulations with those of a constant response after a threshold (explained in Section 2.1). Fig. 5 shows the outbreak size (total number of infected individuals) in each subpopulation for the two strategies, each one averaged over 1000 independent runs. The constant coefficient of contact reduction $a_0 = 0.78$ was chosen to yield a global outbreak size of $\sim 0.20$, approximately equal to that obtained with the LT global strategy with $k = 2$. We notice that the LT memory provides more heterogeneous outcomes between subpopulations compared to the constant response. Especially, as it can be seen in Fig. 6, subpopulations that are farther from the seed tend to have smaller outbreak sizes. This happens because, in the global LT strategy, the intensity of contention measures due to public awareness is the same for all subpopulations and increases with time, thus subpopulations that are seeded later have smaller effective reproduction numbers.

The fact that our proposed model for behavioral responses with global strategy produces more heterogeneous outbreak sizes is notorious, as this is generally observed after outbreaks of real diseases. In particular, this situation is compatible with the one observed in 2020 during the COVID-19 pandemic in those countries that imposed strict global lockdowns even if only part of the country was severely affected (Pollán et al., 2020; Riccardo et al., 2020; Chinazzi et al., 2019). In contrast, a local strategy would lead to more homogeneous outcomes.

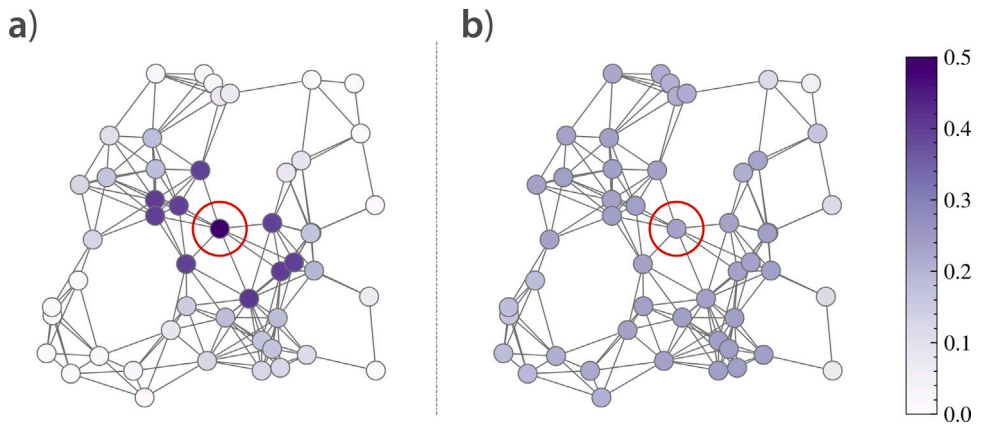


**Fig. 5.** Comparison of the average outbreak size in each subpopulation between (a) LT global strategy and (b) global constant response with threshold. For the LT memory, a response strength $k = 2$ was used, while the constant response uses $a_0 = 0.78$ and a threshold of $\varepsilon_I = 10^{-3}$, meaning that the contact reduction is triggered after $10^4$ overall infections (out of $N_{pop} = 10^7$ individuals) are registered. Both simulations use $\tau = 1 \cdot 10^{-3}$ as the mobility coefficient.

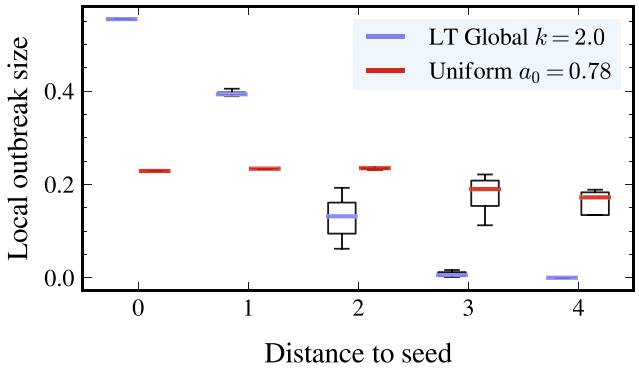


**Fig. 6.** Local outbreak size (given by the maximum of $R_i/N_i$ over time) in each subpopulation as a function of its shortest path length to the seed, for global LT strategy with response strength $k = 2$ (circles) and for constant response with factor $a_0 = 0.78$ (squares). The $y$-axis values represent averages over 1000 independent runs. The boxes show the inner quartiles of each set. The mobility coefficient is set to $\tau = 1 \cdot 10^{-3}$.

## 5. The efficiency of local and global strategies

We now describe a framework to compare local and global strategies in terms of their efficiency, characterized by the costs and benefits of each strategy. Direct comparison of simulations with global and local strategies using the same response strength $k$ is inappropriate, as global strategies require greater values of $k$ to yield similar effects. We therefore define two metrics, one for the cost and another for the effectiveness, and compare local and global strategies in parametric plots of such metrics (with $k$ as an implicit parameter). As long-term and short-term strategies are qualitatively different, we apply different metrics to characterize each one.

### 5.1. Short-term (ST) memory

The short-term memory scheme, in which the contact reduction coefficient only responds to the (either local or global) density of infectious individuals, is characterized by a slow progression of the system towards its herd immunity. A higher value of the response strength $k$ represents a more effective response, which reflects into a smaller prevalence peak (i.e., the maximum $\rho_I$) yet longer outbreak duration. The outbreak size (i.e., maximum of $\rho_R$) however does not vary much with $k$, as it essentially depends on the herd immunity limit of the system. Therefore, the infectious peak size $\rho_I^{\max}$ is a reasonable measure of effectiveness in this case, with smaller peaks attributed to more effective strategies.

Higher values of $k$ also imply more time in which contention measures are active to control disease propagation. This translates into $a_i$ being smaller than 1 during longer times. We can quantify "how much and for how long" the contention measures are applied with the quantity:

$$J_i = \sum_{t=0}^{\infty} (1 - a_i(t)) \tag{14}$$

which accounts for the intensity and time span of the contact reduction. This provides, in a broad sense, a metric for the cost of a strategy in the short-term memory scheme, considering that contention measures typically bring costs to the population. Notice that these metrics must be calculated for each subpopulation.

For the same population and RGN structure described in Section 4, we compare local and global strategies with respect to $\rho_I^{\max}$ and $J$ for simulations with different values of $k$. Fig. 7 shows the plots of these metrics against each other, grouped by sets of subpopulations according to their distance to the seeded subpopulation. The metric values are averaged over 1000 independent runs.

From the plots, it is clear that local strategies outperform global ones for all subpopulations, as the former produces much smaller peaks sizes for similar values of $J$ in the considered range. This means that, for the ST memory, the use of local information about contagions is more efficient than using a unified, global strategy. This happens because using global information may not be well suited for the epidemic situation of a given subpopulation. For instance, subpopulations that are far from the seed effectively apply the contention measures before the epidemic arrives, but as these measures are relaxed at some moment (when the global prevalence decreases), these subpopulations will still undergo an intense outbreak, given that they were only subject to minor outbreaks and the majority of the individuals in these subpopulations are susceptible.

Although we only display the results for a single value of the mobility coefficient $\tau = 10^{-3}$, the same conclusions are obtained for $\tau = 10^{-2}$ and $10^{-4}$, but the advantage of local strategies over global ones is more pronounced with lower mobility rates. This is expected, because with lower $\tau$ values, the dynamics of each subpopulation is less coupled by mobility, making local strategies more practical.

### 5.2. Long-term (LT) memory

In the long-term memory scheme, the overall intensity of contention measures increases with time (though it can locally decrease due to migration), as they are proportional to both I and R densities. As shown in Section 3 for homogeneous mixing, for sufficiently high $k$, the epidemic spreading is halted before herd immunity is achieved,

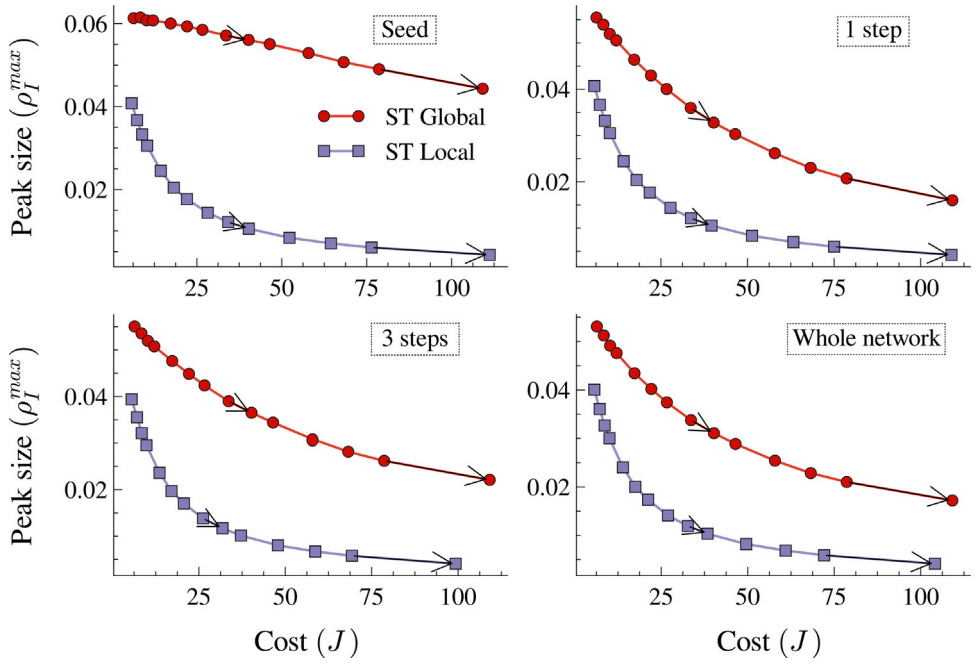


**Fig. 7.** Parametric curves for the effectiveness and cost of local and global ST strategy, for values of $k$ ranging from 3 to 90. The first plot is for the seeded subpopulation, whereas other plots are arithmetic averages over sets of subpopulations according to their shortest path length to the seeded subpopulation. The last plot is an arithmetic average over all subpopulations. The arrows indicate the direction towards which $k$ increases. The mobility coefficient is set to $\tau = 1 \cdot 10^{-3}$. $\rho_I^{\max}$ is the maximum over time of the local infection prevalence $\rho_I$, while $J$ is given by Eq. (14).

leaving the system vulnerable to secondary waves in case of a system release without other immunization policies. We can still characterize the effectiveness of the strategy in this first wave, either by its outbreak size ($\rho_R^{\max}$) or by the peak size ($\rho_I^{\max}$). For consistency with Section 5.1, we chose to work with $\rho_I^{\max}$ as well.

For the cost of the strategy, the metric $J$ (given by Eq. (14)) does not provide a reliable measure, as in this case $a_i$ does not approach 1 again at the end of the main outbreak, making $J$ overly sensitive to the time taken until the epidemic vanishes. We choose instead to simply work with $\bar{a}_{\max} = 1 - \min a(t)$, which is the maximum level of contact reduction adopted by the subpopulation during the simulation. This metric disregards the temporal evolution of $a_i$ (making it inappropriate as a cost of the ST memory), but provides a number that is proportional to the level of contention measures adopted by the subpopulations in the LT memory, being more adequate for this scheme.

Using the same RGN setup, we find that the cost *vs* effectiveness curves are more complex for the LT strategies than for the ST ones. Fig. 8 shows the $\rho_R^{\max}$ *x* $\bar{a}_{\max}$ parametric curves for local and global strategies, and for three values of the mobility master parameter $\tau = 10^{-2}, 10^{-3}$ and $10^{-4}$. In this case, there is not a clear advantage of local strategies over global ones, and this depends both on mobility levels and the distance to the seed. For instance, when $\tau = 10^{-2}$ (Fig. 8(a)), a local strategy is better for the seed, while it is generally worse for more distant subpopulations. For immediate neighbors of the seed (one step away), the curves cross each other, making the optimal strategy to depend on the value of $k$.

For lower mobility values (Fig. 8(b) and (c)), another interesting feature is evident: the curves for the local strategies are not monotonic with the cost, meaning that two strategies (given by two different values of $k$) may have the same cost but notably different effectiveness. This happens because the contention measures in the LT memory scheme can affect the invasion threshold and attack rate of the system, preventing some subpopulations from being reached by the disease for sufficiently high $k$. In this case, as the strategy is based on local prevalences, these subpopulations will not have to implement contention measures, which explains the decrease in the strategy cost.

## 6. Local and global strategies with real metapopulations

The simulations described in Section 5 were performed in a simple RGN topology, which is notably homogeneous. The mobility coefficient $\tau$ is chosen to be not much greater than the invasion threshold of the system, resulting on the interesting relationship between local and global strategies, especially for the LT memory. This has important value to understand the model we propose.

However, for emerging diseases in real world, most systems will be far above their invasion threshold. Moreover, real topologies are hierarchical and differ from an homogeneous topology. To complement our work, we simulate the SIR model with contention measures in two subpopulations based on real mobility and demographic data: one for Spain and another for Brazil. The Spanish demographic data was obtained from the Spanish Statistical Office (2020), while the mobility matrices were provided by the Ministry of Development of Spain (Anon., 2019). The Brazilian metapopulation uses demographic and daily commuting data from the Brasilian Institute of Geography and Statistics (IBGE) (2020), air transport flows from the National Agency of Civil Aviation (ANAC) (2020) and interstate buses from National Agency of Terrestrial Transportation (ANTT) (2020). More details about these data sets are found in the supplementary material.

By comparing the outcomes of the simulations between the Brazilian and Spanish metapopulations, we can better understand the role of the *scale* of the population in the decision between a local and a global strategy. Fig. 9 shows the parametric cost *vs* effectiveness curves of the local and global strategies for the ST memory, for both Spanish (upper panels) and Brazilian (lower panels) populations. In each case, the disease was initiated in the subpopulation with greatest number of individuals: the province of Madrid in Spain and the state of São Paulo in Brazil. In this case, the seeded subpopulation has non-null mobility flows to all other subpopulations, therefore we only plot the curves for the seeded subpopulation, the immediate neighbors (which are all subpopulations except the seed) and the whole network.

As in this more realistic population the mobility is considerably higher than in our synthetic setup from previous sections, the differences between local and global strategies are milder. Nevertheless,

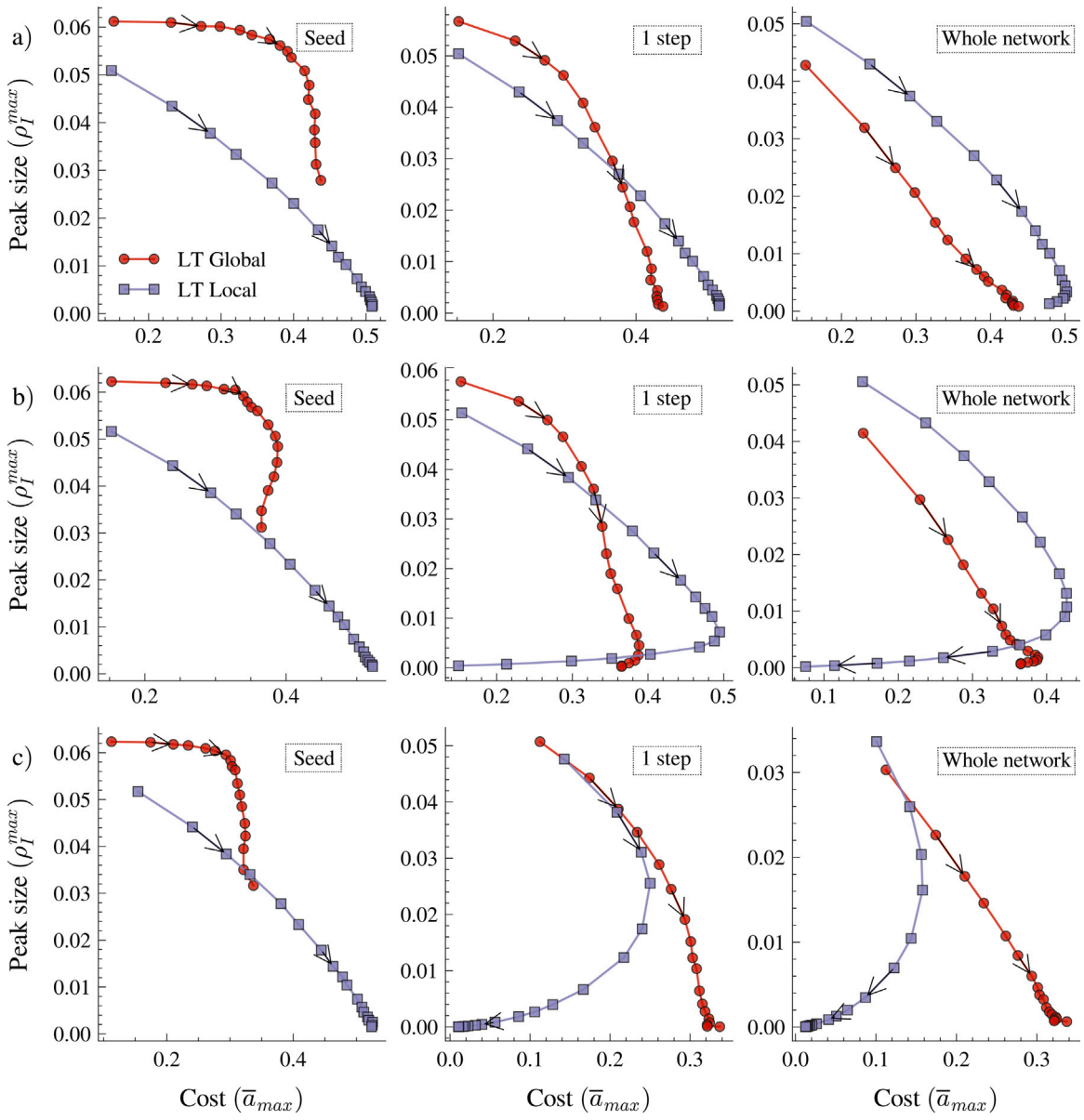


**Fig. 8.** Parametric curves for the effectiveness and cost of local and global LT strategy, for the mobility parameter (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$ and (c) $\tau = 10^{-4}$. The $k$ values range from 0.25 to 50. The leftmost panels represent the seeded subpopulation, the middle panels represent an arithmetic average over its immediate neighbors, and the rightmost panels show the averages over all subpopulations. The arrows indicate the direction towards which $k$ increases. The cost $\overline{a}_{\max}$ is the maximum value of $1 - a_i$ reached during the simulation.

it is clear that a local strategy is always better in cost/benefit than a local one for the ST memory, using $J$ (Eq. (14)) as the strategy cost. The difference is particularly visible for the seeded subpopulation, but also present for the rest of the network. Besides this, it is clear that the Brazilian metapopulation presents a higher difference between the strategies. This happens mainly because the difference in scale: Spain is a densely connected population, with strong mobility flows between provinces. The mobility is also highly centralized around Madrid. Brazil, on the other hand, is a much larger country with less relative mobility between its states. Also notoriously, the most populous state in Brazil (São Paulo) has roughly the same population as the whole Spanish country.

In Fig. 10, we show the same results for the LT memory, using the $\overline{a}_{\max}$ instead of $J$ as the strategy cost. Notice that, again, the Brazilian metapopulation displays greater difference between local and global strategies. However, for the LT memory scheme, the local strategy seems to be better for the seeded subpopulation, while the global strategy is more appropriate for the rest of the network. For high cost strategies, the curves seem also to converge to the same cost as the peak size goes to zero.

By looking at the results from Figs. 9 and 10, we conclude that, for real topologies (which display high mobility flows and heterogeneity between subpopulations), the complicated interplay between local and global strategies is not observed. Yet the main conclusions from Section 5 are still valid: in the ST memory, there is a clear preference for the local strategy, while in the LT memory the choice depends on the circumstances and desired goals for the contention. Notice that we only compare the extreme cases of purely local and global strategies; in reality, "mixed" approaches may be used for optimal disease contention.

## 7. Secondary outbreaks in the global long-term memory with resets

We adapt the metapopulation model with a long-term global strategy to consider system memory resets, as explained in Section 3.2.

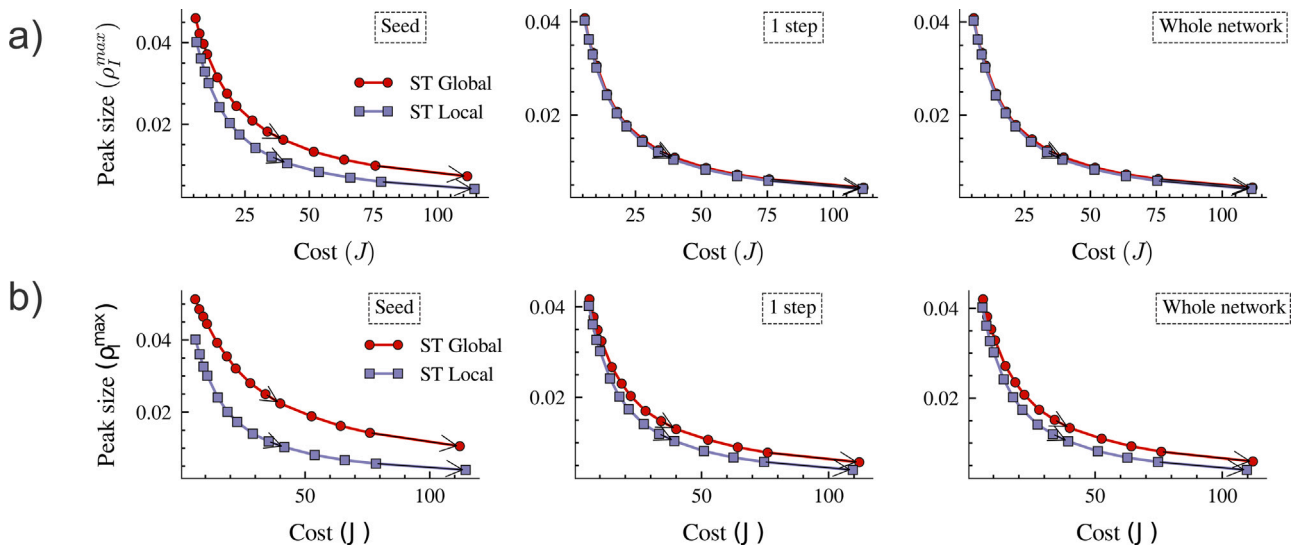


**Fig. 9.** Parametric curves for the effectiveness and cost of local and global ST strategy on metapopulations based on Spanish (a) and Brazilian (b) data. The left panels shows the results for the seeded subpopulation only, the middle panels represent the rest of the network, and the right panels comprise the whole network. Disease and simulation parameters are the same as in Fig. 7.

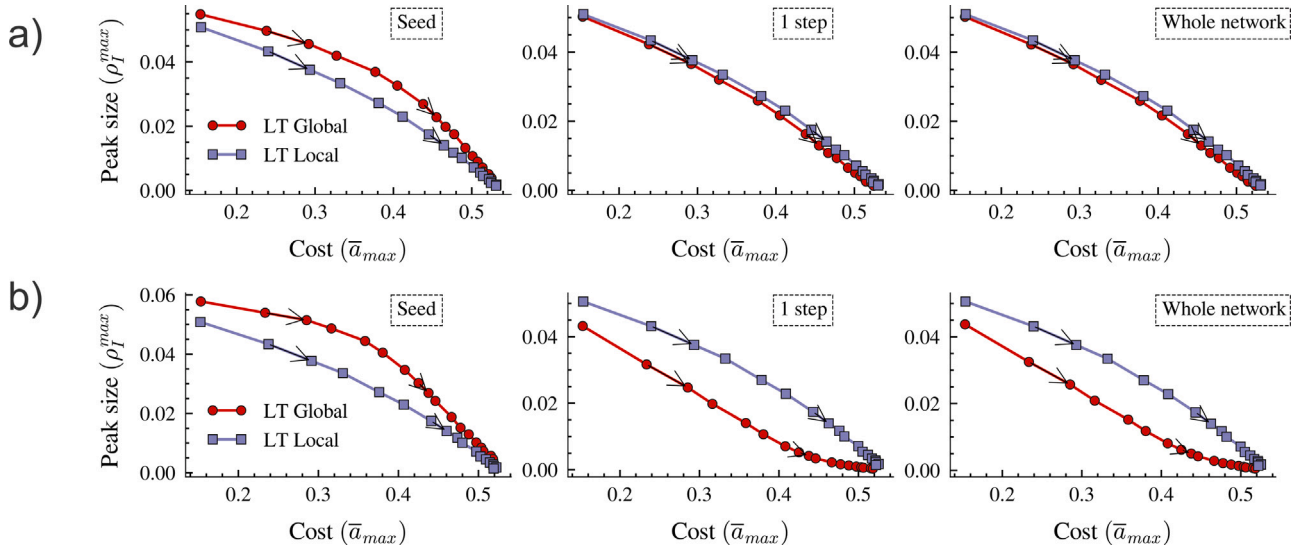


**Fig. 10.** Parametric curves for the effectiveness and cost of local and global LT strategy on metapopulations based on Spanish (a) and Brazilian (b) data. The left panels shows the results for the seeded subpopulation only, the middle panels represent the rest of the network, and the right panels comprise the whole network. Disease and simulation parameters are the same as in Fig. 8.

This allows us to analyze secondary outbreaks caused by a relief in the contention measures, represented here as a reset in the "memory" of the contact reduction mechanism (that is, the value of $\rho_R$ in the $a_i$ coefficient).

The system resets are implemented as follows: the contention measures are activated when the overall fraction of infectious individuals $\rho_I$ surpasses an activation threshold $\varepsilon_I^{\text{in}} = 10^{-4}$ (activation threshold), and then deactivated (possibly after an outbreak) if $\rho_I$ goes under another threshold $\varepsilon_I^{\text{out}} = 0.8 \cdot 10^{-4} < \varepsilon_I^{\in}$. At each activation event, the "memory" of the system is reset, that is, the coefficient of contact reduction is given by $a_i(t) = (1-(\rho_I(t)+\rho_R(t)-\rho_0))^k$, with $\rho_0$ set to the value of $\rho_R$ at the moment of the activation event. This mechanism is slightly different from that used in Section 3.2 for the rate equations model, but the distinction $\varepsilon_I^{\text{out}} < \varepsilon_I^{\text{in}}$ is needed in stochastic simulations so that small fluctuations around $\varepsilon_I$ do not increase the number of outbreaks. To control the number of secondary outbreaks, we also limit the number of system reset events to 10. If this number of reset events is reached before the extinction of the epidemics, the system is then released, that is, $a_i$ is permanently set to 1.

In Fig. 11, we show the time series of the fraction of infectious individuals of each subpopulation (colored thin curves), as well as the global fraction of infectious individuals (black shaded curve) and the value of $a_i$ over time (red dashed curve) for a typical execution of the model. Using the same RGN population as in previous sections, we set the response strength as $k = 5$ and the mobility coefficient to $\tau = 10^{-3}$. Each vertical ascent of the $a_i$ value represents the occurrence of a system reset. We notice that, at each reset, a new set of local outbreaks occurs, raising again the overall incidence and leading the system to readopt contention measures. Differently from the homogeneously mixed case (as in Fig. 3), the secondary outbreaks can be greater in peak size than the first one. This feature is an essential difference between homogeneously mixed populations and structured ones like a metapopulation, and occurs because some (usually many) subpopulations are not reached by the primary outbreak.

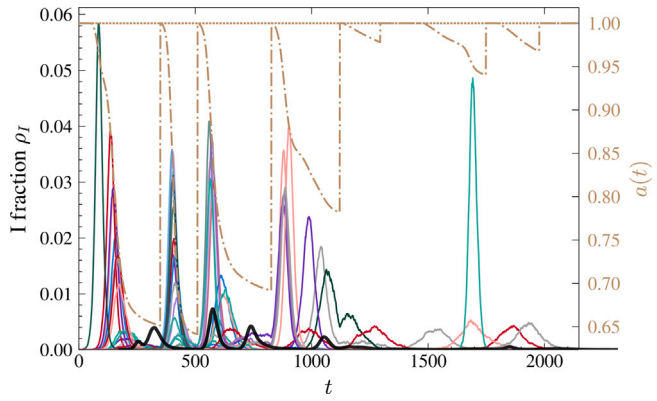


**Fig. 11.** Time series of the fraction of infectious individuals in each subpopulation (colored lines), as well as in the whole population (black line). The global value of $a(t)$ is also shown as a golden dot-dashed line (with the upper limit $a(t) = 1$ as a golden thin horizontal line). The parameters (described in text) are set to: $k = 5$, $\epsilon_I^{\text{in}} = 10^{-4}$, $\epsilon_I^{\text{out}} = 0.8 \cdot 10^{-4}$, $\tau = 10^{-3}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We can better understand this feature by looking at the local incidence. Fig. 12 shows a network map of the greatest peak size (given by $\rho_I^{\max}$) of each subpopulation, along with the time series of the fraction of infectious individuals for some selected subpopulations (lateral panels). The curve of panel (a) corresponds to the seeded subpopulation, displaying a single and pronounced peak in the early stage of the process. Other subpopulations may also present a single main outbreak (as in panel (e)), two (separate or close) outbreaks ((d) and (c)) or even more outbreaks ((b) and (f)). Even though this model is not focused on COVID-19, we can leverage the information collected during the pandemic to put these numbers into perspective. The maximum reductions observed in the figure go up to 40%. Since in this manuscript we have set $R_0 = 1.5$, this leads to an effective reproduction number of $\sim 0.9$. To achieve the same value with the reproduction number of COVID-19 during 2020 ($R_0 = 2.5$), we would require a reduction of the order of 65%. Estimates from several countries have established that the reduction in the number of contacts during the earliest phase of the first wave was around 70%–80%, decaying later on to values in the order of 30%–40% (Coletti et al., 2020; Jarvis et al., 2020; Zhang et al., 2020; Feehan and Mahmud, 2021). As such, the spatial heterogeneity observed in this model is compatible with the one caused by the contention measures during the COVID-19 pandemic (Costa et al., 2020; Sun et al., 2020; Dong et al., 2020; Starnini et al., 2021).

The features of a simulation with system resets are better presented through a single, typical run of the simulation, as we show in Figs. 11 and 12. We still need to ensure that the observed pattern of multiple outbreaks is a solid feature of our model. We do this by counting the number of outbreaks of each time series, then plotting histograms for 1000 independent runs. We use a simple algorithm to determine the number of outbreaks of each local time series, described as follows: every time the fraction of infectious individuals crosses a given threshold $\phi_{\text{in}} = 1 \cdot 10^{-3}$ from below and, posteriorly, another threshold $\phi_{\text{out}} = 0.5\phi_{\text{in}} < \phi_{\text{in}}$ from above, an outbreak is accounted, and the time interval between these two crossings is regarded as a single outbreak. The difference between the two thresholds reduces spurious detection of outbreaks due to stochastic fluctuations around the threshold, though this may still marginally occur. We verified that this choice of parameters prevents most of these spurious detections, while still being able to separate consecutive major outbreaks. The purple shaded areas in each time series of $\rho_I$ (panels a to f) indicate periods of time in which an outbreak was occurring, according to this simple detection algorithm. It can be seen that it performs reasonably well on separating multiple peaks.

Fig. 13 shows the statistics of accounted local outbreaks. On panel (a), the number of outbreaks for each subpopulation and each of the 1000 runs is put into a single histogram, showing that most subpopulations present a single outbreak, but often higher number of outbreaks also occur. On panel (b), we take the average number of outbreaks over all subpopulations in a single simulation run, then plot a histogram for the runs. From it we see that the typical average number of outbreaks is between 1 and 1.4, meaning that most runs present at least some subpopulations that undergo multiple outbreaks.

The main points to be discussed from the results of this section are the following: (i) the occurrence of multiple outbreaks and (ii) the heterogeneity of their features over subpopulations and simulation runs. The first point is a more or less obvious consequence of the mechanism that we introduced: automatic memory resets, which represent a relief to the contention measures. Nevertheless, it is important to demonstrate that our model is capable of producing multiple outbreaks. The second point, heterogeneity of the outbreaks, can be more clearly observed in Fig. 12, as each node can display a considerably different epidemic timeline. Heterogeneity over different runs is visible in Fig. 13.(b), showing that each run can have substantially different average numbers of outbreaks. In Section 5 of the supplemental material, we show how these histograms are affected by the intensity of contention measures and mobility scales.

## 8. Conclusions

We have presented a model that incorporates dynamical behavioral responses to epidemics, which could be driven by governmental policies or by the endogenous response of individuals (e.g. fear of infection), in the context of metapopulations. The emerging features of the model are very rich, showing a complex landscape of outcomes depending on the implemented memory scheme and strategy.

First, we have shown that for isolated populations, strong social distancing measures (LT) are able to ablate the epidemic, but render the system vulnerable. Indeed, as long as the prevalence is below the herd immunity threshold, a reintroduction of the pathogen after the measures are lifted will inevitably lead to a second wave. On the other hand, soft social distancing measures (ST) contain the exponential growth of the epidemic and leave the system in a state in which the prevalence is over the herd immunity threshold, preventing further outbreaks. Note, however, that for a severe disease in which the infection fatality rate is substantial, larger prevalence implies a larger number of deaths. Thus, stronger social distancing policies will reduce the number of deaths at the expense of leaving the system vulnerable, while softer measures will control the spread but increase the number of deaths. These observations are particularly important in the context of metapopulations. Since subpopulations are not isolated, in the absence of additional control measures, the disease might be seeded again in disease-free areas by individuals who where infected in other regions. Thus, besides long and short term strategies, we also need to incorporate whether the contention measures use global or local information.

We compared LT and ST memory schemes of our model, in an attempt to address the question of whether the contention measures are more efficient if implemented and managed locally (at each subpopulation) or globally (equally in the whole population). For the ST memory, we measured its cost by the $J$ metric that quantifies the intensity and duration of the contention measures, and used the maximum number of simultaneous infections (i.e., the peak size) to quantify the strategy's performance. We showed that the local strategy always outperformed the global one, which is valid for any value of the mobility parameter. For the LT memory, we quantified the cost by the maximum intensity reached by the contention measures, and the performance by the peak size. In this case, the cost/benefit relationship was more complex, depending on the mobility rate and the distance to the seeded subpopulation. Global strategies are generally better for

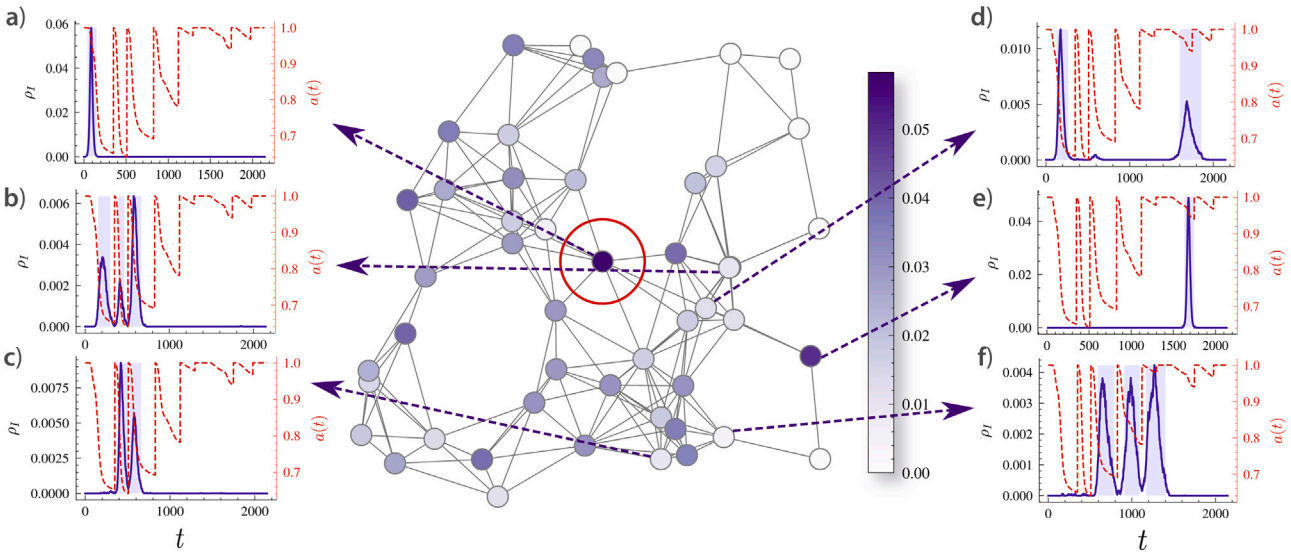


**Fig. 12.** Map of the final outbreak sizes ($\rho_R^{\max}$) of each subpopulation, for a single simulation run with the memory reset mechanism (main panel). The panels show the time series of $I_i/N_i$ for some sample subpopulations, showing that the mechanism introduces diversity in the epidemic trajectory over the population. Shaded areas on each panel show the periods during which an outbreak is taking place, according to an algorithm described in the text. The average number of detected outbreaks for this run was 1.20. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

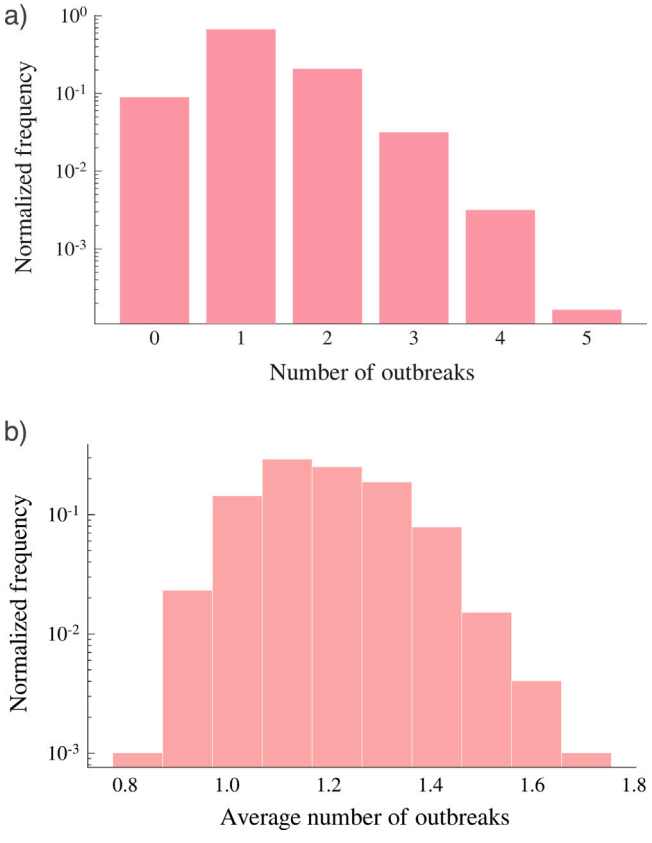


**Fig. 13.** Histograms of (a) the number of local outbreaks (at subpopulation level) for all subpopulations, and (b) of the average number of outbreaks in each simulation run (the average is taken over subpopulations, but not over runs). The parameters of the simulations are set to: $k = 5$, $\varepsilon_I^{\text{in}} = 1 \cdot 10^{-4}$ and $\varepsilon_I^{\text{out}} = 0.8 \cdot 10^{-4}$. The number of system reset events was constrained to at most 10.

distant subpopulations (with respect to the seed). This situation was precisely the one that emerged in the first wave of infections in 2020 during the COVID-19 pandemic in countries that imposed country-wide lockdowns even if only one region was severely affected (Pollán et al., 2020; Riccardo et al., 2020; Chinazzi et al., 2019). Yet, this is reversed if the mobility is low enough so that the local strategy can stop the epidemics before these nodes are reached. For the seeded subpopulation and its neighbors, local strategies are typically preferred. For this type of memory scheme, therefore, the choice between global and local strategies is not trivial and should be addressed appropriately.

We extended the comparison between local and global strategies to data-informed metapopulations, for Spain and Brazil. These differ from the considered RGN by the heterogeneous/hierarchical structure and the greater mobility flows. The comparison with the results in the RGN let us conclude that the local strategies are still better than global ones for the ST memory, while they are only better for the seeded subpopulation in the LT memory scheme. Comparison between the Brazilian and Spanish metapopulations highlight the role of scale: a larger and less connected population (Brazil) will display greater difference between local and global strategies than a smaller, denser population (Spain).

It is important to notice that long-term strategies assume a permanent adoption of contention measures, which leaves the system vulnerable to secondary outbreaks if these measures are lifted. We address this by implementing memory resets which, after outbreaks, instantly lifts the intensity of contention measures but leaves its mechanism active. Due to the residual presence of infectious individuals, the system faces secondary outbreaks. Thus, subpopulations that experienced a milder first wave, might have worse outcomes in subsequent waves. Besides, regardless of the contention measures, once they are lifted the system keeps progressing towards the herd immunity threshold, but each population might experience waves of different intensity at different times. The spatial heterogeneity driven by contention measures depicted in this paper is compatible with the one observed during the COVID-19 pandemic in several regions of the world (Costa et al., 2020; Sun et al., 2020; Dong et al., 2020; Starnini et al., 2021).

In summary, we have shown than even in a simple metapopulation model a very complex scenario emerges. Depending on the mobility and distance to the seed, local vs global strategies yield different results

and introduce different heterogeneities. As such, the full complexity of human gatherings and behaviors should be accounted for to effectively deal with emerging diseases.

Our work sets a new framework to considering social distancing behaviors during emerging epidemics, and we are aware of multiple possibilities to extend it. For example, the inclusion of disease-driven mobility reduction or targeted travel restrictions (Feng and Jin, 2020) could considerably modify the dynamics, possibly enhancing the effectiveness of local strategies. The model (in particular with the LT memory) could also be extended to include vaccines and other pharmaceutical interventions at some point, creating immunity by other means rather than infection and relieving the contention measures. Additionally, the behavioral responses mechanism can be enhanced with opinion dynamics and game theory, regarding the willingness to adopt protection measures. Some of these models for homogeneous mixing (Amaral et al., 2021) could be employed at the subpopulation level. These feature leads to overall weaker responses, but also introduce interesting dynamical features.

## CRediT authorship contribution statement

**Paulo Cesar Ventura:** Methodology, Software, Investigation, Formal analysis, Writing – original draft. **Alberto Aleta:** Conceptualization, Investigation, Writing – original draft. **Francisco Aparecido Rodrigues:** Supervision, Writing – review & editing. **Yamir Moreno:** Project administration, Writing – review & editing.

## Declaration of competing interest

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

## Data availability

The source code used for this work is published in Zenodo, available at https://doi.org/10.5281/zenodo.6258445.

## Acknowledgments

P.C.V. acknowledges the financial support of FAPESP through grants 2016/24555-0 and 2019/11183-5. F.A.R. acknowledges CNPq (grant 309266/2019-0) and FAPESP (grant 19/23293-0) for the financial support given for this research. A.A. and Y.M. acknowledge the support by Soremartec S.A. Y.M. acknowledges partial support from the Government of Aragon, Spain and FEDER funds, Spain through grant E36-20R (FENOL), and by MCIN/AEI and FEDER funds (grant PID2020-115800GB-I00). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epidem.2022.100544.

## References

Ajelli, Marco, Gonçalves, Bruno, Balcan, Duygu, Colizza, Vittoria, Hu, Hao, Ramasco, José J., Merler, Stefano, Vespignani, Alessandro, 2010. Comparing large-scale computational approaches to epidemic modeling: Agent-based versus structured metapopulation models. BMC Infect. Dis. 10 (1), 1–13.

Aleta, Alberto, Hisi, Andreia N.S., Meloni, Sandro, Poletto, Chiara, Colizza, Vittoria, Moreno, Yamir, 2017. Human mobility networks and persistence of rapidly mutating pathogens. R. Soc. Open Sci. 4 (3), 160914.

Aleta, Alberto, Hu, Qitong, Ye, Jiachen, Ji, Peng, Moreno, Yamir, 2020. A data-driven assessment of early travel restrictions related to the spreading of the novel COVID-19 within mainland China. Chaos Solitons Fractals 139, 110068.

Aleta, Alberto, Moreno, Yamir, 2020. Evaluation of the potential incidence of COVID-19 and effectiveness of containment measures in Spain: A data-driven approach. BMC Med. 18 (1), 1–12.

Amaral, Marco A., de Oliveira, Marcelo M., Javarone, Marco A., 2021. An epidemiological model with voluntary quarantine strategies governed by evolutionary game dynamics. Chaos Solitons Fractals 143, 110616.

Anon., 2019. Studying Interprovince Mobility of Passengers using Big Data. Technical report, Ministry of Development of Spain, https://observatoriotransporte.mitma.gob.es/estudio-experimental.

Apolloni, Andrea, Poletto, Chiara, Ramasco, José J., Jensen, Pablo, Colizza, Vittoria, 2014. Metapopulation epidemic models with heterogeneous mixing and travel behaviour. Theor. Biol. Med. Model. 11 (1), 1–26.

Arregui, Sergio, Aleta, Alberto, Sanz, Joaquín., Moreno, Yamir, 2018. Projecting social contact matrices to different demographic structures. PLoS Comput. Biol. 14 (12), e1006638.

Bajardi, Paolo, Poletto, Chiara, Ramasco, Jose J., Tizzoni, Michele, Colizza, Vittoria, Vespignani, Alessandro, 2009. Human mobility networks, travel restrictions, and the global spread of H1N1 pandemic. PLoS One 6 (1), e16591.

Ball, Frank, Britton, Tom, House, Thomas, Isham, Valerie, Mollison, Denis, Pellis, Lorenzo, Tomba, Gianpaolo Scalia, 2015. Seven challenges for metapopulation models of epidemics, including households models. Epidemics 10, 63–67.

Brasilian Institute of Geography and Statistics (IBGE), 2020. https://www.ibge.gov.br. (Accessed 28 May 2020).

Calvetti, Daniela, Hoover, Alexander P., Rose, Johnie, Somersalo, Erkki, 2020. Metapopulation network models for understanding, predicting, and managing the coronavirus disease COVID-19. Front. Phys. 8.

Chinazzi, Matteo, Davis, Jessica T., Ajelli, Marco, Gioannini, Corrado, Litvinova, Maria, Merler, Stefano, Piontti, Ana Pastore y., Mu, Kunpeng, Rossi, Luca, Sun, Kaiyuan, Viboud, Cécile, Xiong, Xinyue, Yu, Hongjie, Elizabeth Halloran, M., Longini, Ira M., Vespignani, Alessandro, 2019. The effect of travel restrictions on the spread of the novel coronavirus (COVID-19) outbreak. Science 368 (6489), 395–400.

Coletti, Pietro, Wambua, James, Gimma, Amy, Willem, Lander, Vercruysse, Sarah, Vanhoutte, Bieke, Jarvis, Christopher I., Van Zandvoort, Kevin, Edmunds, John, Beutels, Philippe, Hens, Niel, 2020. CoMix: Comparing mixing patterns in the Belgian population during and after lockdown - Scientific reports. Sci. Rep. 10 (21885), 1–10.

Colizza, Vittoria, Vespignani, Alessandro, 2008. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. J. Theoret. Biol. 251 (3), 450–467.

Costa, Guilherme S., Cota, Wesley, Ferreira, Silvio C., 2020. Outbreak diversity in epidemic waves propagating through distinct geographical scales. Phys. Rev. Res. 2 (4), 043306.

Dong, Ensheng, Du, Hongru, Gardner, Lauren, 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 20 (5), 533–534.

Dormand, John R., Prince, Peter J., 1980. A family of embedded Runge-Kutta formulae. J. Comput. Appl. Math. 6 (1), 19–26.

Eksin, Ceyhun, Paarporn, Keith, Weitz, Joshua S., 2019. Systematic biases in disease forecasting–The role of behavior change. Epidemics 27, 96–105.

Feehan, Dennis M., Mahmud, Ayesha S., 2021. Quantifying population contact patterns in the United States during the COVID-19 pandemic - Nature communications. Nature Commun. 12 (893), 1–9.

Feng, Shanshan, Jin, Zhen, 2020. Infectious diseases spreading on an adaptive metapopulation network. IEEE Access 8, 153425–153435.

Funk, Sebastian, Salathé, Marcel, Jansen, Vincent A.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: A review. J. R. Soc. Interface 7 (50), 1247–1256.

Gollwitzer, Anton, Martel, Cameron, Brady, William J., Pärnamets, Philip, Freedman, Isaac G., Knowles, Eric D., Van Bavel, Jay J., 2020. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. Nature Hum. Behaviour 4 (11), 1186–1197.

Jarvis, Christopher I., Van Zandvoort, Kevin, Gimma, Amy, Prem, Kiesha, Klepac, Petra, James Rubin, G., John Edmunds, W., 2020. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. BMC Med. 18 (1), 1–10.

Jewel, Park, 2020. Changes in subway ridership in response to COVID-19 in Seoul, South Korea: Implications for social distancing. Cureus 12 (4).

Keeling, Matt J., Bjørnstad, Ottar N., Grenfell, Bryan T., 2004. Metapopulation dynamics of infectious diseases. In: Ecology, Genetics and Evolution of Metapopulations. Academic Press, Cambridge, MA, USA, pp. 415–445.

Lloyd, Alun L., May, Robert M., 1996. Spatial heterogeneity in epidemic models. J. Theoret. Biol. 179 (1), 1–11.

Lu, Dan, Aleta, Alberto, Ajelli, Marco, Pastor-Satorras, Romualdo, Vespignani, Alessandro, Moreno, Yamir, 2021. Data-driven estimate of SARS-CoV-2 herd immunity threshold in populations with individual contact pattern variations. medRxiv, 2021.03.19.21253974.

Manfredi, Piero, D'Onofrio, Alberto, 2013. Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases. Springer, New York, NY, USA.

Mossong, Joël, Hens, Niel, Jit, Mark, Beutels, Philippe, Auranen, Kari, Mikolajczyk, Rafael, Massari, Marco, Salmaso, Stefania, Tomba, Gianpaolo Scalia, Wallinga, Jacco, Heijne, Janneke, Sadkowska-Todys, Malgorzata, Rosinska, Magdalena, John Edmunds, W., 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 5 (3), e74.

National Agency of Civil Aviation (ANAC), 2020. https://www.anac.gov.br. (Accessed 27 May 2020).

National Agency of Terrestrial Transportation (ANTT), 2020. https://www.gov.br/antt. (Accessed 28 May 2020).

Perra, Nicola, Balcan, Duygu, Gonçalves, Bruno, Vespignani, Alessandro, 2011. Towards a characterization of behavior-disease models. PLoS One 6 (8), e23084.

Pollán, Marina, Pérez-Gómez, Beatriz, Pastor-Barriuso, Roberto, Oteo, Jesús, Hernán, Miguel A., Pérez-Olmeda, Mayte, Sanmartín, Jose L., Fernández-García, Aurora, Cruz, Israel, de Larrea, Nerea Fernández, Molina, Marta, Rodríguez-Cabrera, Francisco, Martín, Mariano, Merino-Amador, Paloma, Paniagua, Jose León, Muñoz-Montalvo, Juan F., Blanco, Faustino, Yotti, Raquel, Blanco, Faustino, Fernández, Rodrigo Gutiérrez, Martín, Mariano, Navarro, Saturnin Mezcua, Molina, Marta, Muñoz-Montalvo, Juan F., Hernández, Matías Salinero, Sanmartín, Jose L., Cuenca-Estrella, Manuel, Yotti, Raquel, Paniagua, José León, de Larrea, Nerea Fernández, Fernández-Navarro, Pablo, Pastor-Barriuso, Roberto, Pérez-Gómez, Beatriz, Pollán, Marina, Avellón, Ana, Fedele, Giovanni, Fernández-García, Aurora, Iglesias, Jesús Oteo, Olmeda, María Teresa Pérez, Cruz, Israel, Martinez, Maria Elena Fernandez, Rodríguez-Cabrera, Francisco D., Hernán, Miguel A., Fernández, Susana Padrones, Aguirre, José Manuel Rumbao, Marí, José M. Navarro, Borrás, Begoña Palop, Jiménez, Ana Belén Pérez, Rodríguez-Iglesias, Manuel, Gascón, Ana María Calvo, Alcaine, María Luz Lou, Suárez, Ignacio Donate, Álvarez, Oscar Suárez, Pérez, Mercedes Rodríguez, Sanchís, Margarita Cases, Gomila, Carlos Javier Villafáfila, Saladrigas, Lluis Carbo, Fernández, Adoración Hurtado, Oliver, Antonio, Feliciano, Elías Castro, Quintana, María Noemí González, Fernández, José María Barrasa, Betancor, María Araceli Hernández, Febles, Melisa Hernández, Martín, Leopoldo Martín, López, Luis-Mariano López, Miota, Teresa Ugarte, Población, Inés De Benito, Pérez, María Sagrario Celada, Fernández, María Natalia Vallés, Enríquez, Tomás Maté, Arranz, Miguel Villa, González, Marta Domínguez-Gil, Fernández-Natal, Isabel, Lobón, Gregoria Megías, Bellido, Juan Luis Muñoz, Ciruela, Pilar, Casals, Ariadna Mas i., Botías, Maria Doladé, Angeles Marcos Maeso, M., del Campo, Dúnia Pérez, de Castro, Antonio Félix, Ramírez, Ramón Limón, Retamosa, Maria Francisca Elías, González, Manuela Rubio, Lobeiras, María Sinda Blanco, Losada, Alberto Fuentes, Aguilera, Antonio, Bou, German, Caro, Yolanda, Marauri, Noemí, Blanco, Luis Miguel Soria, González, Isabel del Cura, Pascual, Montserrat Hernández, Fernández, Roberto Alonso, Merino-Amador, Paloma, Castro, Natalia Cabrera, Lizcano, Aurora Tomás, Almagro, Cristóbal Ramírez, Hernández, Manuel Segovia, Elizaga, Nieves Ascunce, Sanz, María Ederra, Baquedano, Carmen Ezpeleta, Bascaran, Ana Bustinduy, Tamayo, Susana Iglesias, Otazua, Luis Elorduy, Benarroch, Rebeca Benarroch, Flores, Jesús Lopera, de la Villa, Antonia Vázquez, 2020. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. Lancet 396 (10250), 535–544.

Read, J.M., Edmunds, W.J., Riley, S., Lessler, J., Cummings, D.A.T., 2012. Close encounters of the infectious kind: Methods to measure social mixing behaviour. Epidemiol. Infect. 140 (12), 2117–2130.

Riccardo, Flavia, Ajelli, Marco, Andrianou, Xanthi D., Bella, Antonino, Manso, Martina Del, Fabiani, Massimo, Bellino, Stefania, Boros, Stefano, Urdiales, Alberto Mateo, Marziano, Valentina, Rota, Maria Cristina, Filia, Antonietta, D'Ancona, Fortunato, Siddu, Andrea, Punzo, Ornella, Trentini, Filippo, Guzzetta, Giorgio, Poletti, Piero, Stefanelli, Paola, Castrucci, Maria Rita, Ciervo, Alessandra, Benedetto, Corrado Di, Tallon, Marco, Piccioli, Andrea, Brusaferro, Silvio, Rezza, Giovanni, Merler, Stefano, Pezzotti, Patrizio, the COVID-19 working group, 2020. Epidemiological characteristics of COVID-19 cases and estimates of the reproductive numbers 1 month into the epidemic, Italy, 28 january to 31 2020. Eurosurveillance 25 (49), 2000790.

Salathé, Marcel, Kazandjieva, Maria, Lee, Jung Woo, Levis, Philip, Feldman, Marcus W., Jones, James H., 2010. A high-resolution human contact network for infectious disease transmission. Proc. Natl. Acad. Sci. USA 107 (51), 22020–22025.

da Silva, Paulo Cesar Ventura, Velásquez-Rojas, Fátima, Connaughton, Colm, Vazquez, Federico, Moreno, Yamir, Rodrigues, Francisco A., 2019. Epidemic spreading with awareness and different timescales in multiplex networks. Phys. Rev. E 100 (3).

Spanish Statistical Office, 2020. https://www.ine.es/. (Accessed February 2020).

Starnini, Michele, Aleta, Alberto, Tizzoni, Michele, Moreno, Yamir, 2021. Impact of data accuracy on the evaluation of COVID-19 mitigation policies. Data Policy 3.

SteelFisher, Gillian K., Blendon, Robert J., Bekheit, Mark M., Lubell, Keri, 2009. The public's response to the H1N1 influenza pandemic. N. Engl. J. Med. 362 (22), e65.

Stroeymeyt, Nathalie, Grasse, Anna V., Crespi, Alessandro, Mersch, Danielle P., Cremer, Sylvia, Keller, Laurent, 2018. Social network plasticity decreases disease transmission in a eusocial insect. Science 362 (6417), 941–945.

Sun, Feinuo, Matthews, Stephen A., Yang, Tse-Chuan, Hu, Ming-Hsiao, 2020. A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: Where geography matters? Ann. Epidemiol. 52, 54–59, e1.

Waitzberg, Ruth, Davidovitch, Nadav, Leibner, Gideon, Penn, Nadav, Brammli-Greenberg, Shuli, 2020. Israel's response to the COVID-19 pandemic: Tailoring measures for vulnerable cultural minority populations. Int. J. Equity Health 19 (1), 1–5.

Wang, Lin, Li, Xiang, 2014. Spatial epidemiology of networked metapopulation: An overview. Chin. Sci. Bull. 59 (28), 3511–3522.

Wilson-Aggarwal, Jared K., Ozella, Laura, Tizzoni, Michele, Cattuto, Ciro, Swan, George J.F., Moundai, Tchonfienet, Silk, Matthew J., Zingeser, James A., McDonald, Robbie A., 2019. High-resolution contact networks of free-ranging domestic dogs canis familiaris and implications for transmission of infection. PLOS Negl. Trop. Dis. 13 (7), 1–19.

Woskie, Liana R., Hennessy, Jonathan, Espinosa, Valeria, Tsai, Thomas C., Vispute, Swapnil, Jacobson, Benjamin H., Cattuto, Ciro, Gauvin, Laetitia, Tizzoni, Michele, Fabrikant, Alex, Gadepalli, Krishna, Boulanger, Adam, Pearce, Adam, Kamath, Chaitanya, Schlosberg, Arran, Stanton, Charlotte, Bavadekar, Shailesh, Abueg, Matthew, Hogue, Michael, Oplinger, Andrew, Chou, Katherine, Corrado, Greg, Shekel, Tomer, Jha, Ashish K., Wellenius, Gregory A., Gabrilovich, Evgeniy, 2021. Early social distancing policies in Europe, changes in mobility & COVID-19 case trajectories: Insights from spring 2020. PLoS One 16 (6), e0253071.

Zhang, Juanjuan, Litvinova, Maria, Liang, Yuxia, Wang, Yan, Wang, Wei, Zhao, Shanlu, Wu, Qianhui, Merler, Stefano, Viboud, Cécile, Vespignani, Alessandro, Ajelli, Marco, Yu, Hongjie, 2020. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science.