Physica A 596 (2022) 127092

Contents lists available at ScienceDirect

Physica A

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

Community-distributed compartmental models

G. Hernández [a,*], A. Martín del Rey [b]

[a] Universidad de Salamanca, BISITE Research Group, Salamanca, Spain
[b] Universidad de Salamanca, Institute of Fundamental Physics and Mathematics and Department of Applied Mathematics, Salamanca, Spain

ARTICLE INFO

Article history:
Received 2 June 2021
Received in revised form 10 February 2022
Available online 3 March 2022

Keywords:
Compartmental models
Community structure
Epidemic control
Short-term forecast
COVID-19

ABSTRACT

A framework that allows the incorporation of community structure into epidemiological compartmental models has been developed. The models resulting from this process are compartmental models as well, which are related to the base models. This work includes an existence and uniqueness theorem, showing that, under certain conditions on the mobility, epidemiological models in which $\mathbf{f}(t, \mathbf{X})$ is continuous in time and Lipschitz continuous on the compartments induce unique community models; and a homogeneous mixing limit, showing that under high mobility conditions the base model is recovered in the global population. Applications of the SIR model and the impact of the community structure on the estimation of their effective parameters are discussed in detail. An open computational implementation of this framework is available to the scientific community. It allows modeling community distribution using mobility data, as shown with Spain data during the 2020 state of alarm.

1. Introduction

The COVID 19 pandemic has transformed the world in many areas, starting with the epidemiological field, where many efforts have been made to extend classical dynamic models to reproduce the behaviors found in the real world while maintaining the advantages of their interpretability.

One of the main flaws which are found in the usual compartmental models is the lack of community structure within the population they represent. Topologically, individuals can be understood as being arranged in a complete graph in which any interaction is possible. However, it is evident that this situation does not reflect reality, and that human communities are arranged in a structure that is worth capturing.

Several precedents on models that take into account the structure of communities can be found in the literature. In [1] the propagation in a discrete network of individuals with a randomly obtained structure following the free-scale-network model is analyzed. Moreover, in [2] a particular case of propagation in two communities is analyzed theoretically, obtaining results on the reproductive number and on the endemic equilibrium point, including immunization scenarios. In [3] an SIS model with two communities with subpopulation structure is analyzed, also with interest in the limiting case of endemic equilibrium, as well as in carrying out simulations of the dynamic evolution of the system. In the work [4], a compartmental model is used to study how outbreaks arise as a consequence of seasonality in the spread of diseases.

The main goal of this work is to introduce a new computational framework for predicting the spatiotemporal propagation of a certain biological agent taking into account a model for the population structure. Specifically, a set of

* Corresponding author.
E-mail address: guillehg@usal.es (G. Hernández).

https://doi.org/10.1016/j.physa.2022.127092

population communities connected through a complex network is considered and specific compartmental models are implemented in each one taking into account population dynamics with exchange terms. The SIR model is used as a paradigmatic example to illustrate the influence of community structure on the resulting global models.

The main contribution of this work with respect to previous ones is that it allows transforming a community dynamics model (or a family of models) and a time-dependent transition matrix into a single model, providing computational tools for it, as well as including theoretical results on the existence and uniqueness of solutions in simple cases. An application of the model with real mobility data is also presented.

2. Methodology

Let $c_1, c_2, \ldots, c_n$ be the communities through which the epidemic propagates. The population in each community can be modeled as belonging to a series of compartments $\{X_i^k(t)\}$, where $X_i^k(t)$ is the population in community $i \in \{1, \ldots, n\}$ with state $k \in \{1, \ldots, m\}$ (e.g., "susceptible", "infected", ...) at time t.

If individuals in these communities are isolated from each other, the evolution of individual's states in each community is defined by the following system of ordinary differential equations:

$$\frac{\mathrm{d}\mathbf{X}_i}{\mathrm{d}t}(t) = \mathbf{f}_i(\mathbf{X}_i(t), t), \quad 1 \leq i \leq n\,, \tag{1}$$

where $\mathbf{X}_i = (X_i^k)_k$ and $\mathbf{f}_i = (f_i^k)_k$ is a community-dependent evolution function, identified with the community model. In the following, bold letters are used for vectors on these states or operators returning such vectors, while superscripts in each letter are used to represent the possible states of individuals in communities.

Example 2.1 (*SIR*)**.** The SIR model is a bi-parametric 3-compartment model defined by

$$\mathbf{SIR}_{\beta,\gamma}(S, I, R) = \begin{pmatrix} -\frac{\beta SI}{S+I+R} \\ \frac{\beta SI}{S+I+R} - \gamma I \\ \gamma I \end{pmatrix}, \tag{2}$$

where β is the infection rate and γ is the recovery rate. Both of these parameters could be treated as functions of time. The compartments have been denoted as $X_i^1 = S, X_i^2 = I, X_i^3 = R$ for the sake of interpretability (Susceptible, Infected, Recovered/Removed).

Note $\sum_k \mathrm{SIR}_{\beta,\gamma}^k = 0$, and hence the total population $S + I + R$ is constant if the evolution is given by (1). This is an intended simplification of the model, which can be extended to account for population dynamics.

Definition 2.1 (*Conservative Model*)**.** A model $\mathbf{f}$ is conservative (with respect to its total population) if

$$\sum_k f^k = 0\,. \tag{3}$$

This definition is a particular case of the general notion of conservative dynamic systems where the conserved quantity is the total population $\sum_k X^k$.

Note many of the models used in current epidemics research are conservative [5–9]. Non-conservative models can also be found, the typical case including birth and/or death rates [10].

Now let $M_{ij}^k(t) \geq 0$ be the rate of movement of individuals in the compartment k from community i to community j at time t. To account for these population dynamics, the evolution of the system originally described by (1) is now given instead by

$$\frac{\mathrm{d}\mathbf{X}_i}{\mathrm{d}t} = \mathbf{f}_i(\mathbf{X}_i, t) - \sum_j \mathbf{M}_{ij}(t) + \sum_j \mathbf{M}_{ji}(t)\,, \quad 1 \leq i \leq n\,, \tag{4}$$

where $\mathbf{M}_{ij} = (M_{ij}^k)_k = \left(M_{ij}^1, M_{ij}^2, \ldots, M_{ij}^m\right)$.

Observation 2.1. *The overall model (4) is also of the form (1) and it is, thus, a compartmental model as well. This can be seen as an operator $\mathcal{C}$ defined as*

$$\mathcal{C}\left(\left(\mathbf{f}_i\right)_i, \left(\mathbf{M}_{ij}\right)_{ij}\right) : \left(\left(\mathbf{X}_i\right)_i, t\right) \mapsto \left(\mathbf{f}_i(\mathbf{X}_i, t) - \sum_j \mathbf{M}_{ij}(t) + \sum_j \mathbf{M}_{ji}(t)\right)_i, \tag{5}$$

which transforms an indexed family of models with the same compartments and their connection rules into a single global compartmental model.

In many cases we are interested in community-transition models where the population of each of the communities is constant, assuming the epidemiological model is also conservative.

Definition 2.2 (*Conservative Transition Model*)**.** A community transition model defined by (4) is conservative (with respect to every community population) if

$$\sum_{j,k} M_{ij}^k(t) - M_{ji}^k(t) = 0 \quad \forall i, t\,. \tag{6}$$

When this condition holds, the total population of each of the communities is constant.

To study the dynamics of (4), it is interesting to define a simpler formalism where the number of parameters is reduced. If the transition between communities is independent of the compartment, it can be described without loss of generality as

$$M_{ij}^k(t) = \frac{X_i^k(t)}{\sum_l X_i^l(t)} m_{ij}(t)\,, \tag{7}$$

where $m_{ij}(t) > 0$ does not depend on any $X_i^k(t)$. Note $m_{ij}(t)$ can be thought as a labeled directed graph with no loops and with time-dependent edge weights, where the communities are identified with the vertices and edges exist if and only if $m_{ij} \neq 0$.

To study the effect of the topology of the communities connection, it is natural to consider a compartment-independent constant transition rate between the connected communities, that is, $m_{ij} = m_{ji} \in \{0, m\}$ for all $1 \leq i, j \leq n$. In that case

$$\frac{\mathrm{d}X_i^k}{\mathrm{d}t} = \mathcal{C}^k(\mathbf{f}_i, G, m) \stackrel{\text{def}}{=} f_i^k(\mathbf{X}_i, t) + m \sum_{j \in \mathcal{N}_G(i)} \left(\frac{X_j^k(t)}{\sum_l X_j^l(t)} - \frac{X_i^k(t)}{\sum_l X_i^l(t)} \right)\,, \tag{8}$$

where $\mathcal{N}_G(i)$ is the open neighborhood of i (i.e., not including i) in the labeled undirected graph with no loops G describing the connections between communities. The last component of (8) can be regarded as a tendency to level the relative population of individuals in each of the compartments of the connected communities.

The operator $\mathcal{C}((\mathbf{f}_i)_i, G, m)$ defined by the components in k of (8) transforms an indexed family of models, a connection graph, and a transition rate into a global compartmental model. The notation $\mathcal{C}$ has been kept as in (5) for simplicity since the meaning can be distinguished by the nature of the arguments.

A further simplification is to assume the models are community-independent. In that case, the system of differential equations is defined by

$$\mathcal{C}(\mathbf{f}, G, m) = \mathcal{C}((\mathbf{f}, \mathbf{f}, \ldots), G, m)\,, \tag{9}$$

where the $\mathcal{C}$ notation has been kept again for simplicity.

Example 2.2 (*Common Graph Topologies*)**.** Some common graph topologies, which will be used below, are the following:

- The complete graph K_n.
- The cycle graph C_n.
- The linear (or path) graph P_n, with edges $\big\{(1,2),(2,3),\ldots,(n-1,n)\big\}$.
- The star graph S_n, with edges $\big\{(1,2),(1,3),\ldots,(1,n)\big\}$.

Example 2.3 (*3 Communities Linear SIR Model*)**.** The evolution of the 3-communities linear SIR model $\mathcal{C}(\mathbf{SIR}_{\beta,\gamma}, P_3, m)$ is given by

$$\begin{array}{rcl}
\dot{S_1} &=& -\beta I_1 S_1/N_1 - mS_1/N_1 + mS_2/N_2 \\
\dot{I_1} &=& \beta I_1 S_1/N_1 - \gamma I_1 - mI_1/N_1 + mI_2/N_2 \\
\dot{R_1} &=& \gamma I_1 - mR_1/N_1 + mR_2/N_2 \\
\dot{S_2} &=& -\beta I_2 S_2/N_2 + mS_1/N_1 - 2\,mS_2/N_2 + mS_3/N_3 \\
\dot{I_2} &=& \beta I_2 S_2/N_2 - \gamma I_2 + mI_1/N_1 - 2\,mI_2/N_2 + mI_3/N_3 \\
\dot{R_2} &=& \gamma I_2 + mR_1/N_1 - 2\,mR_2/N_2 + mR_3/N_3 \\
\dot{S_3} &=& -\beta I_3 S_3/N_3 + mS_2/N_2 - mS_3/N_3 \\
\dot{I_3} &=& \beta I_3 S_3/N_3 - \gamma I_3 + mI_2/N_2 - mI_3/N_3 \\
\dot{R_3} &=& \gamma I_3 + mR_2/N_2 - mR_3/N_3
\end{array}\,, \tag{10}$$

where

$$N_i = S_i + I_i + R_i \tag{11}$$

is the population of the ith community.

To provide a solution for the system, an initial value problem has to be defined by adding a vector of initial conditions $\mathbf{X}_0$. Let $\mathcal{S}(\mathbf{f}, \mathbf{X}_0)$ be the solution of the initial value problem so defined.

The Picard–Lindelöf theorem is commonly used to show the existence and uniqueness of the solution $\mathcal{S}(\mathbf{f}, \mathbf{X}_0)$, as long as $\mathbf{f}(t, \mathbf{X})$ is continuous on t and Lipschitz continuous on $\mathbf{X}$. This theorem can be applied to the extended problem $\mathcal{S}(\mathcal{C}(\mathbf{f}, G, m), \mathbf{X}'_0)$.

Theorem 2.1 (*Existence and Uniqueness*)**.** *Let* $\mathbf{f}$ *be continuous on the first variable and Lipschitz continuous on the rest of them in an open connected set* $D \subseteq \mathbb{R}^n$*. If* $\mathbf{f}$ *is conservative and* $\mathbf{M}$ *is conservative , compartment-independent and continuous in time, then there is unique solution* $\mathcal{S}(\mathcal{C}(\mathbf{f}, \mathbf{M}), \mathbf{X}'_0)$*, where* $X'_0 \in D^{|G|}$*, on some time interval* $(-t', t')$*.*

Proof. Since M is compartment-independent, the system that governs the dynamics is of the form (8). Since the model and the transitions are conservative, the denominators on its right-hand side are constant. Hence, the right-hand side is a summation of Lipschitz continuous and thus Lipschitz continuous. The result follows from the usual theorem. □

It is interesting to compare the solution of the community to aggregate summing in the communities general, the solution $\mathcal{S}(\mathcal{C}(\mathbf{f}, G, m), \mathbf{X}_0)$, which can obtain by addition of the compartments from the different communities. To introduce that, consider the operator

$$\sum_G : \left(X_i^k\right)_i^k \mapsto \left(\sum_i X_i^k\right)^k , \tag{12}$$

which adds the community component while keeping the state compartments separated.

A directed comparison of $\sum_G \mathcal{S}(\mathcal{C}(\mathbf{f}, G, m), \mathbf{X}_0)$ and $\mathcal{S}(\mathbf{f}, \sum_G \mathbf{X}_0)$ might be used to see the impact of the topology in the model induced by G and m. Examples are shown later in Section 3.

A limit case for the values of the exchange rate m can be stated as follows:

Theorem 2.2 (*Homogeneous Mixture Limit*)**.** *Let* f *be a linear compartment model and* G *a connected graph. Then, for any* $\mathbf{X}_0$

$$\lim_{m\to\infty} \sum_G \mathcal{S}(\mathcal{C}(\mathbf{f}, G, m), \mathbf{X}_0) = \mathcal{S}(\mathbf{f}, \sum\nolimits_G \mathbf{X}_0) , \tag{13}$$

i.e., if m *is large enough the evolution of the community-based model is equivalent to the one resulting from applying the model to the global population. This can be denoted informally as* $\mathcal{C}(\mathbf{f}, G, \infty) = \mathbf{f} \circ \sum_G$*.*

Proof. The evolution of the compartments is given by (8). A series on $\varepsilon = m^{-1}$ will be used. Let $T = t$, $\tau = t/\varepsilon$ be the "slow" and "fast" time parameters. Using the series

$$X_i^k(t, \varepsilon) = {X_i^k}_0(\tau, T) + \varepsilon {X_i^k}_1(\tau, T) + O(\varepsilon^2) \tag{14}$$

on (8),

$$\begin{aligned} \frac{\partial {X_i^k}_0}{\partial T} + \varepsilon^{-1}\frac{\partial {X_i^k}_0}{\partial \tau} + \frac{\partial {X_i^k}_1}{\partial T} = {} & f^k(\mathbf{X}_{i0}(\tau, T), T) \\ & + \varepsilon^{-1} \sum_{j\in\mathcal{N}_G(i)} \left(\frac{{X_j^k}_0(\tau, T)}{N_j} - \frac{{X_i^k}_0(\tau, T)}{N_i} \right) \\ & + \sum_{j\in\mathcal{N}_G(i)} \left(\frac{{X_j^k}_1(\tau, T)}{N_j} - \frac{{X_i^k}_1(\tau, T)}{N_i} \right) \\ & + O(\varepsilon) , \end{aligned} \tag{15}$$

where the rule of the chain was applied to the LHS and

$$N_i = \sum_l X_i^l(t) , \tag{16}$$

which is constant for all i due to the conservative mobility.

The $O(\varepsilon^{-1})$ term in (15) states that

$$\frac{\partial {X_i^k}_0}{\partial \tau} = \sum_{j\in\mathcal{N}_G(i)} \left(\frac{{X_j^k}_0(\tau, T)}{N_j} - \frac{{X_i^k}_0(\tau, T)}{N_i} \right) , \tag{17}$$

which is a linear ODE system. Hence, its solution are described by

$${X^k}_0(\tau, T) = \exp(A\tau) \cdot c(T) , \tag{18}$$

where

$$A_{ij} = \begin{cases} -\frac{|\mathcal{N}_G(i)|}{N_i} & i = j \\ \frac{\delta_{j\in\mathcal{N}_G(i)}}{N_j} & i \neq j \end{cases}, \tag{19}$$

and the "constants" of integration $c(T)$ are functions of T.

Since the system (17) comes from the potential

$$V\left((X^k_{i\,0})_i\right) = \frac{1}{2}\sum_{\{i,j\}\in E(G)}\left(\frac{X^k_{j\,0}}{N_j} - \frac{X^k_{i\,0}}{N_i}\right)^2, \tag{20}$$

it solution must converge to a minimum of V, where

$$\frac{X^k_{j\,0}(\infty,T)}{N_j} = \frac{X^k_{i\,0}(\infty,T)}{N_i} \quad \forall i \in V(G), j \in \mathcal{N}_G(i). \tag{21}$$

Since the graph G is connected, the equality holds for all $i, j \in V(G)$ by transitivity, which determines a unique fixed point as the asymptotic solution.

Note that if equilibrium

$$\sum_{j\in\mathcal{N}_G(i)}\left(\frac{X^k_{j\,0}(\tau,T)}{N_j} - \frac{X^k_{i\,0}(\tau,T)}{N_i}\right) = 0, \tag{22}$$

holds for a given T, the solution of (17) would be constant with τ.

Now, the $O(1)$ term in (15) states

$$\frac{\partial X^k_{i\,0}}{\partial T} + \frac{\partial X^k_{i\,1}}{\partial \tau} = f^k(\mathbf{X}_{i0}(\tau,T),T) + \sum_{j\in\mathcal{N}_G(i)}\left(\frac{X^k_{j\,1}(\tau,T)}{N_j} - \frac{X^k_{i\,1}(\tau,T)}{N_i}\right). \tag{23}$$

This can be seen as a linear superposition on the independent problems on $X^k_{i\,0}$ and $X^k_{i\,1}$. The second one defines the fast-time behavior of the next order of approximation, which is not needed here. Considering the first one,

$$\frac{\partial X^k_{i\,0}}{\partial T} = f^k\left(\mathbf{X}_{i0}(\tau,T),T\right), \tag{24}$$

each community $\mathbf{X}_{i0}$ evolves following dynamics like those of the model $\mathbf{f}$, but perhaps for the dependence on the fast scale τ. However, if the equilibrium condition (22) holds at a given T, it is preserved for all T. To see this consider

$$\frac{\partial X^k_{i\,0}/N_i}{\partial T} = \frac{f^k\left(\mathbf{X}_{i0}(\tau,T),T\right)}{N_i} = f^k\left(\frac{\mathbf{X}_{i0}(\tau,T)}{N_i},T\right), \tag{25}$$

where linearity was used in the last step. Applying $\frac{\partial}{\partial T}$ on (22) and substituting (25) there, the proof follows from linearity.

Finally, adding by communities

$$\frac{\partial \sum_i X^k_{i\,0}}{\partial T} = f^k\left(\sum_i \mathbf{X}_{i0}(\tau,T),T\right), \tag{26}$$

where the linearity of f and $\frac{\partial}{\partial T}$ was used. □

Observation 2.2. *In the limit $m \to \infty$, the relative population of compartments in adjacent communities must be equal at all times. This might not be compatible with the initial conditions, so the presence of the fast time component shown above is actually needed for the problem to be consistently defined.*

Observation 2.3. *There is an obvious analog of Theorem 2.2 for general graphs replacing $\sum_G$ by an operator performing the summation by connected components of G.*

A different point of view might be adopted by considering a family of models $\mathbf{f}_\alpha$ which depends on a set of parameters α. The family of solutions to $\sum_G \mathcal{S}(\mathcal{C}(\mathbf{f}_\alpha, G, m), \mathbf{X}_0)$ will not be in correspondence to those of $\mathcal{S}(\mathbf{f}_\alpha, \sum_G \mathbf{X}_0)$ in general. However, an approximation might be provided using a least-square fit of such parameters. In some particular cases, a direct approach might be followed if the parameters can be uniquely determined from $(t, \mathbf{X}, \dot{\mathbf{X}}, \ddot{\mathbf{X}}, \dots)$, taking that as a definition for an effective parameter in the general case.

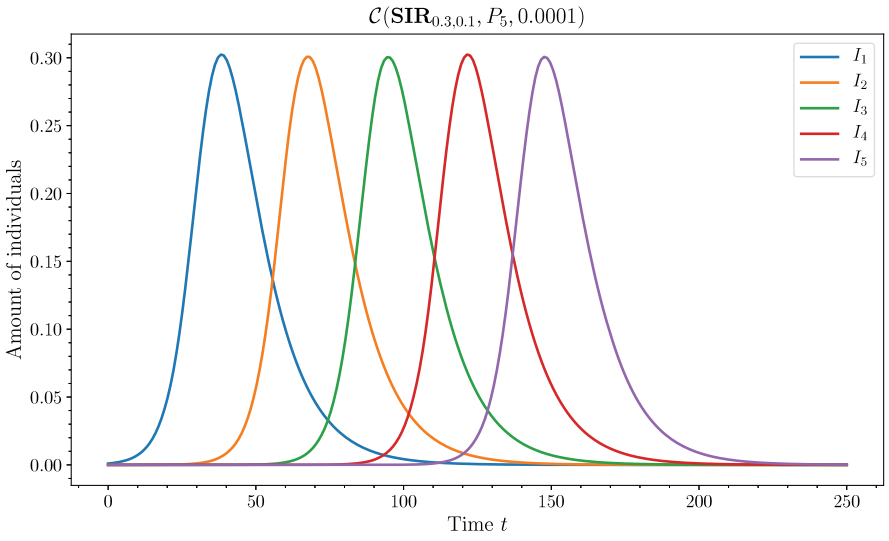

Fig. 1. Infected individuals by communities in a linear model.

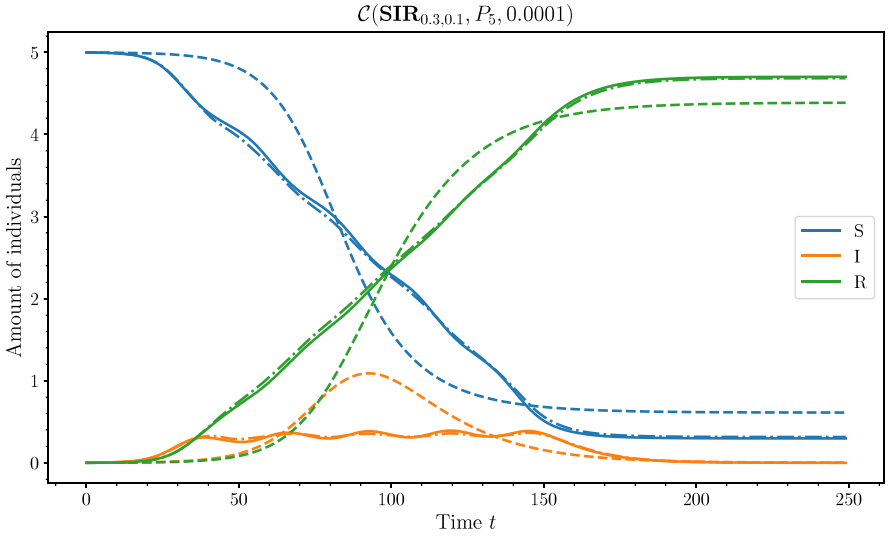

Fig. 2. Global individuals in a linear community model. Best fit with constant parameters is shown as the dashed lines, while sliding window fits are shown as the dot-dashed lines.

Example 2.4. The parameters of $\mathbf{SIR}_{\beta,\gamma}$ can be solved from (2),

$$\beta = -\frac{\dot{S}(S+I+R)}{SI}, \tag{27}$$

$$\gamma = \frac{\dot{R}}{I}. \tag{28}$$

Another parameter that can be extracted to describe the system is the so-called *basic reproductive number*, R_0, which is a crucial epidemiological coefficient that can be defined in a general way as the expected number of secondary infections caused by a unique and "typical" infectious individual during their infectious period in an entirely susceptible population. It depends on (1) the length of the infectious period, (2) the probability that contact between a susceptible individual and an infectious one leads to infection, and (3) the total number of susceptible individuals contacted per unit of time. Note that R_0 can vary significantly when different infectious diseases are considered or, even, when the same biological agent spreads over different susceptible populations.

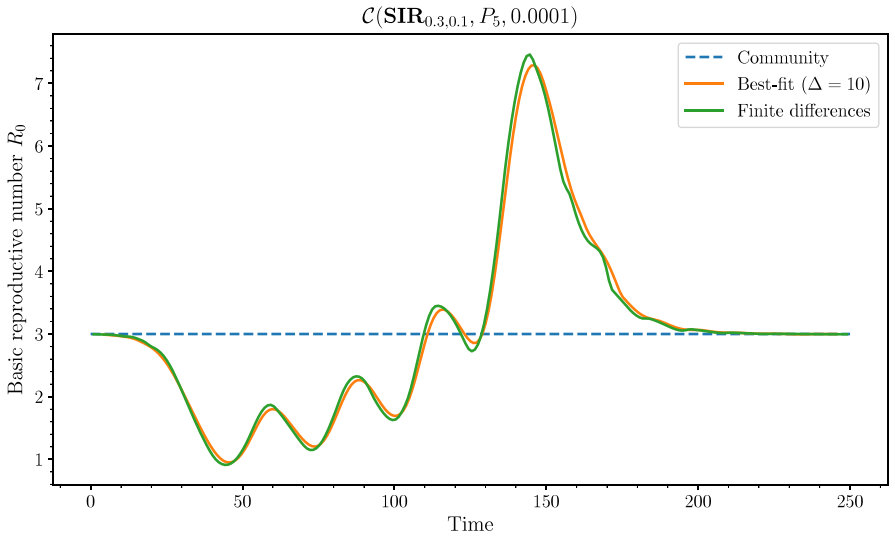

Fig. 3. R_0 coefficient best approximating the model.

In the deterministic paradigm considered in this work, the most simple way to define R_0 is the so-called *survival function method*, according to which

$$R_0 = \int_0^\infty \alpha(t)F(t)dt, \tag{29}$$

where $\alpha(t)$ stands for the average number of newly infected individuals produced per unit of time by the unique infectious individual for total time t, and $F(t)$ denotes the probability that a newly infected individual remains infectious for at least time t.

In a $\mathbf{SIR}_{\beta,\gamma}$ model this parameter takes the value

$$R_0 = \frac{\beta}{\gamma} = -\frac{\dot{S}(S+I+R)}{S\dot{R}}. \tag{30}$$

When $R_0 > 1$ the infection is able to spread in a population, but not so if $R_0 < 1$.

3. Experiments

This section presents a series of studies carried out employing numerical calculation. For the resolution of the initial value problems, the Explicit Runge–Kutta method of order 5(4) has been used [14] through the implementation available in scipy [15]. Best fits were calculated in the sense of the least-squares, adding the errors in each of the compartments, using the Levenberg–Marquardt algorithm [16], also provided by scipy as a wrapper over the MINPACK library.

3.1. Simple graph topologies

A first example with the evolution of the 5-communities linear SIR model $\mathcal{C}(\mathbf{SIR}_{0.3,0.1}, P_5, 0.0001)$ as defined by (8) is given below. The parameters were chosen to allow for an easy interpretation of the recovery time scale ($\gamma = 0.1$, so $\gamma^{-1} = 10$), to model a supercritical infection ($\beta = 0.3$, so $R_0 = 3 > 1$) and a slow but perceptible flow rate ($m = 0.0001$, so a 0.1% of population is exchanged per unit of time). The influence of this values is studied later in this work.

The individual community components are shown in Fig. 1, which illustrates how the infection propagates across the linear network P_5.

The global result is shown in Fig. 2 with best fits to $\mathbf{SIR}_{\beta,\gamma}$, both as constant parameters (dashed lines) and as functions of time (dot-dashed lines) in sliding windows of 10 units of time. On the one hand, as it can be seen in the figure, the adjustment to constant parameters of the SIR model is unable to approximate qualitatively the general behavior of the system. On the other hand, while fitting with time-varying coefficients is able to get closer to the form of the composite system. However, the interpretation of the parameters of the original SIR model is unsatisfactory, as shown below.

The time-dependent of the "effective" R_0 parameter extracted numerically from (30) is shown in Fig. 3. The effective γ not shown here is approximately constant, so, for purpose of interpretation of the figure, β can be regarded as proportional to the R_0 shown.

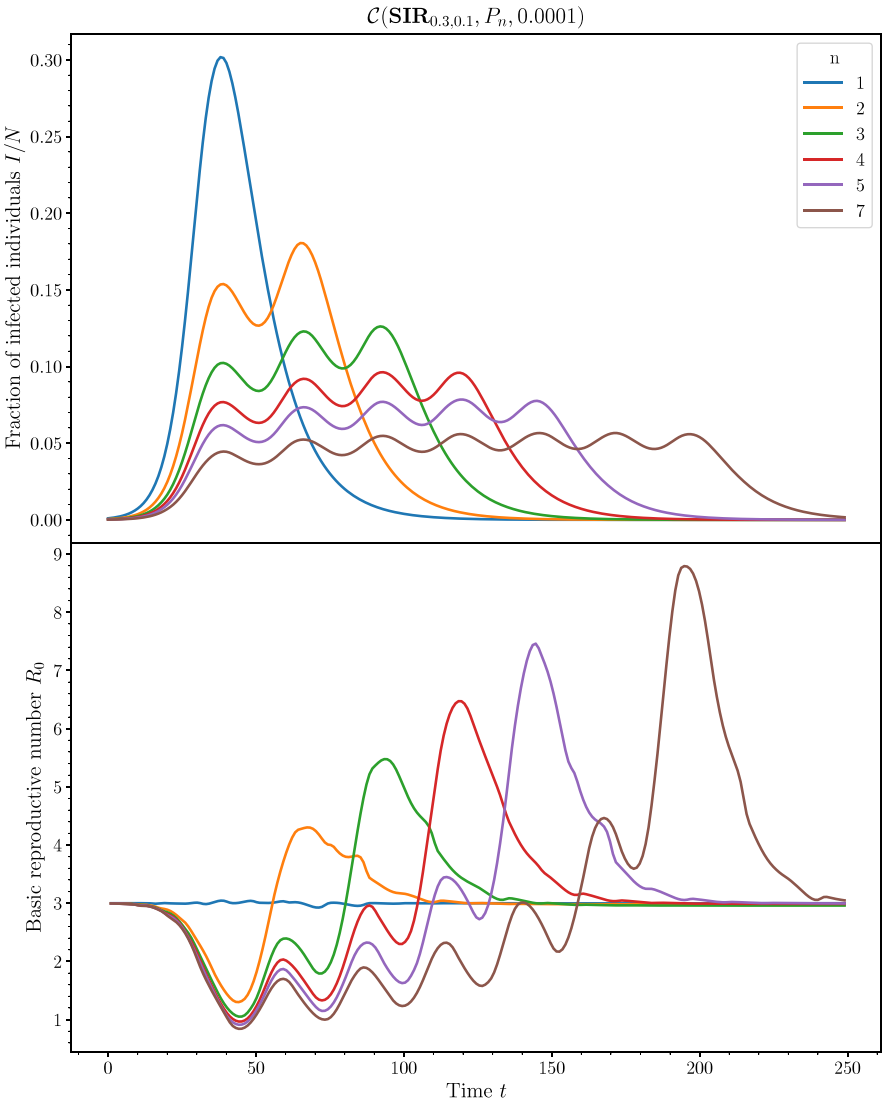

Fig. 4. Effect of the number of communities in $\mathcal{C}(\mathbf{SIR}_{\beta,\gamma}, P_n, m)$.

The reason for the peaked behavior at the end of the epidemics is due to the dividing S found in (30). While the number of overall susceptible individuals has been substantially reduced, the fraction of susceptible individuals in the remaining community is higher. This illustrates the fact that the description provided by a parameter R_0 alone is not sufficient to define the dynamics of the system, and that the model itself is important for this purpose. In this case, there is an internal structure (the communities) which the global model is unable to reproduce.

The effect of the number of communities is shown in Fig. 4. Note the vertical axis in the upper part is scaled by the total population N. Hence, this can be interpreted as a quantitative description of how an imposed limitation on the maximum size of groups can help to "flatten" the curve of infections, as well as allowing for control when a peak of infections is detected.

The effect of the transmission coefficient β is shown in Fig. 5. When $\beta < 0.1$, $R_0 < 1$ and no spread occurs.

The effect of exchange coefficient m is shown in Fig. 6. A flattening effect can be observed again as m decreases. This kind of study can be useful to determine the level of movement-restriction measures needed to prevent the advance of an infection. The homogeneous mixture limit (13) is also seen in the figure.

Finally, some other topologies are shown for comparison in Fig. 7. The peaks of each of the communities can be easily identified taking into account the nature of the network. In the case of the cyclic graph C_5, the peak for the community where the infection begins can be seen, followed by two overlapping peaks of the neighbor communities, and finally other overlapping two from the remaining communities. For the star graph S_5 a small peak after the one of

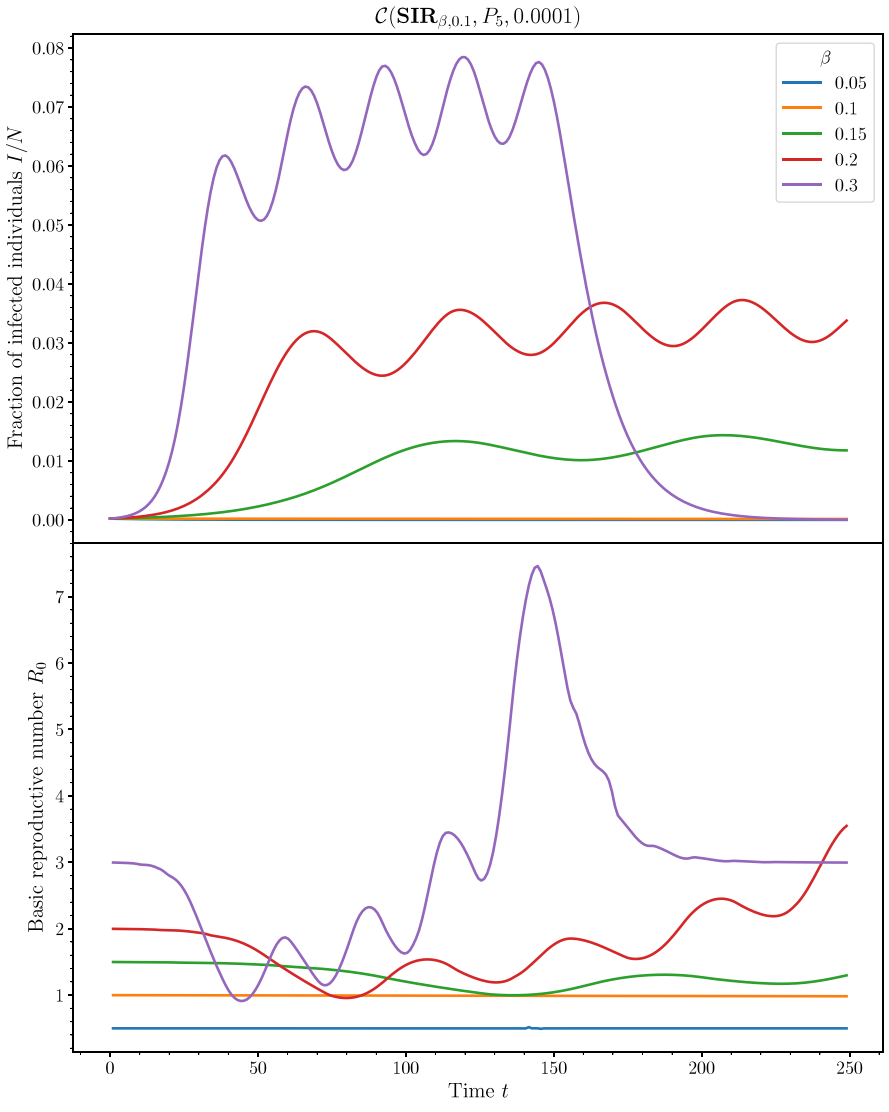

Fig. 5. Effect of the transmission coefficient β in $\mathcal{C}(\mathbf{SIR}_{\beta,\gamma}, P_n, m)$.

the first community, corresponding to the central node of the network can be seen. After that, 3 overlapping peaks for the remaining communities occur. Finally, in the complete graph K_5, four overlapping peaks succeed the first one.

3.2. Realistic example

A realistic example can be build using data modeling transitions from a set of communities. To that purpose, a model has been built using data collected from more than 80% of the mobile phones of Spain [11] during its state of alarm.

For this study, the data between 1 April and 1 June 2020, which represents a situation of reduced mobility due to the measures taken in response to the COVID-19 crisis, have been averaged. A graphical depiction of the transitions so detected is shown in Fig. 8.

An example of the evolution of an infection assumed to follow a $\mathbf{SIR}_{\beta,\gamma}$ model, starting at the province of Albacete with a 0.01% of the infected population is shown in Fig. 9. The time location of the peaks can be related to the distance of the nodes in Fig. 8, which, of course, also resembles the geographical distance.

Best fits of the overall components are shown in Fig. 10. The behavior is similar to that found with simpler graph models from the previous sections. In particular, the increase in the effective number of the R_0 is also found.

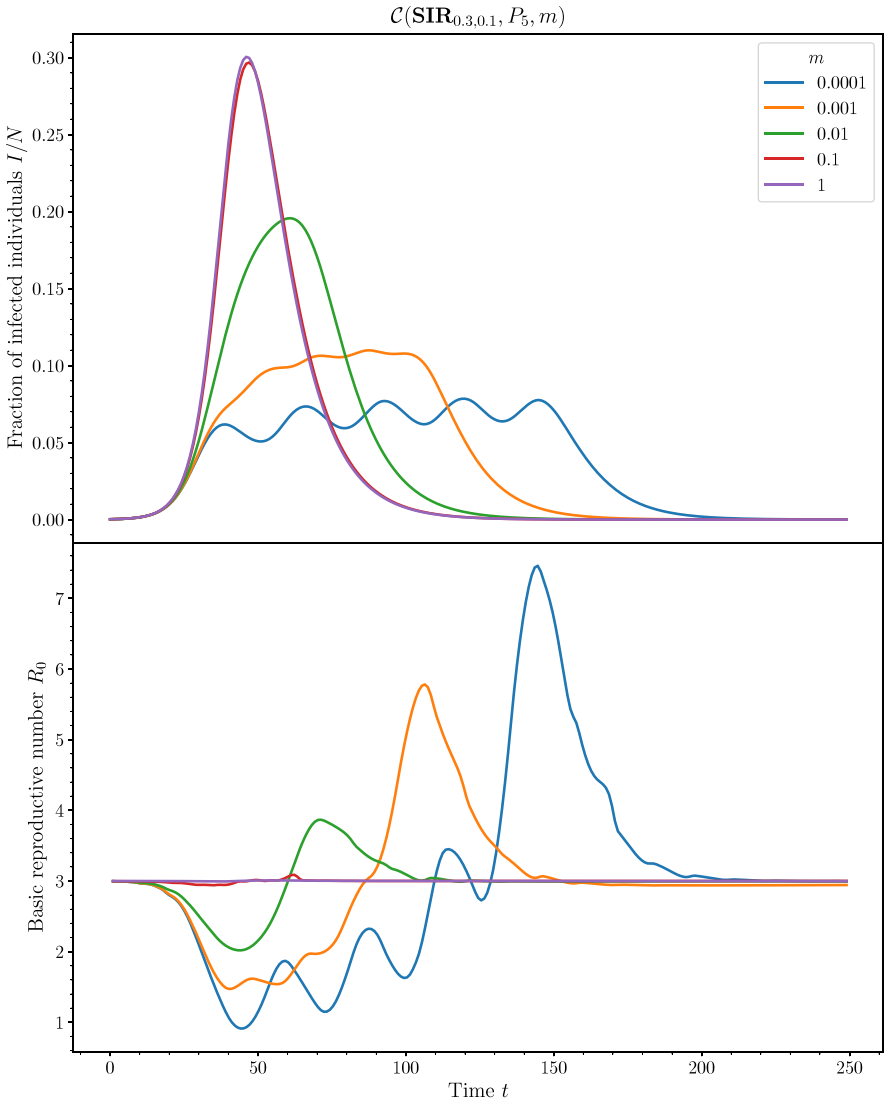

Fig. 6. Effect of the exchange coefficient m in $\mathcal{C}(\mathbf{SIR}_{\beta,\gamma}, P_n, m)$.

4. Conclusions

A mathematical framework to introduce a community structure described by exchange terms within compartmental epidemiological models has been developed. The resulting models are compartmental models as well. Some properties relating the community models with the base epidemiological model have been studied, including an existence and uniqueness theorem and a homogeneous mixture limit which recovers the base model in case of high mobility.

This framework can be used to model the impact that containing measures such as community size restrictions or reduced mobility can have on the spread of an infection. A case study was also built using mobility data from Spain during the state of alarm in 2020.

The computational tools implemented to solve community-based models have been integrated into an open-source Python package. The code generating the simulations shown in this work is available online [17].

The results show that despite a model with time-dependent parameters may be used to approximate a community-based model, the real causes of the dynamics are lost. As a consequence, it is worth including geographical information in epidemiological datasets to allow for community-based modeling.

Future work includes the study of non-conservative models and random graphs.

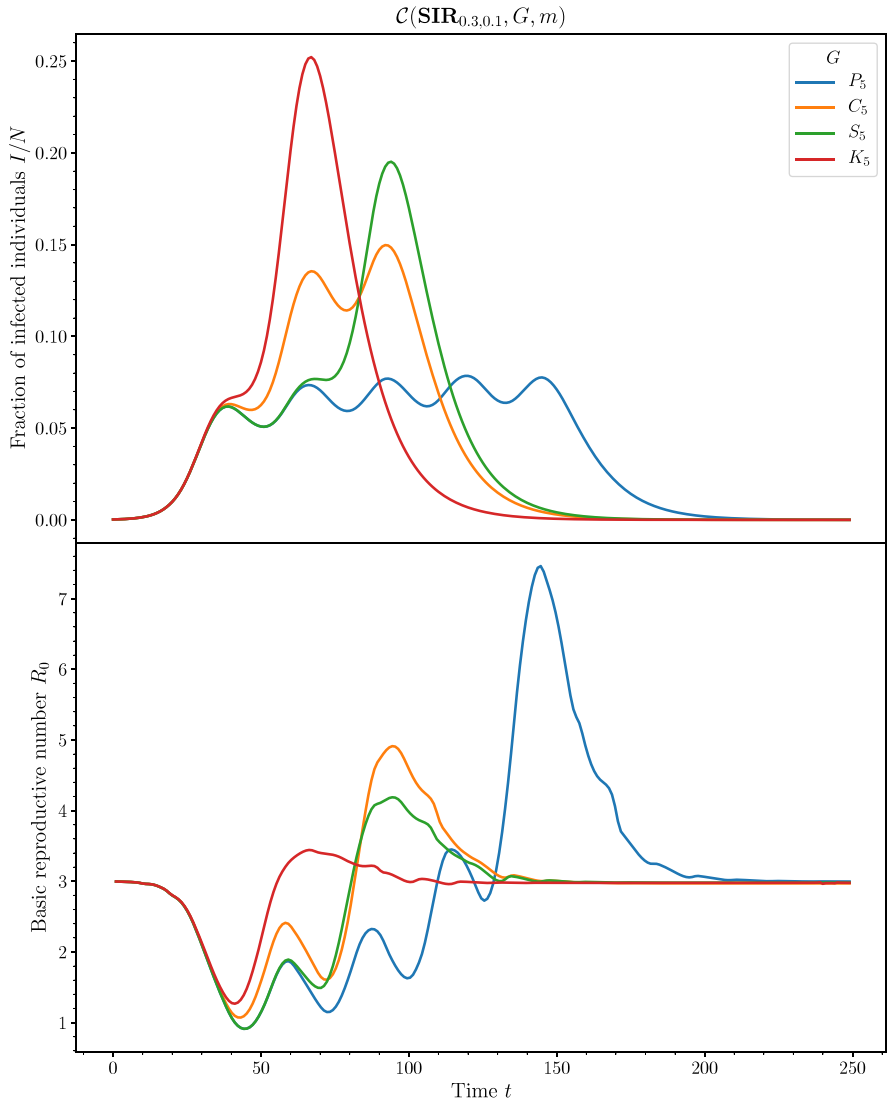

Fig. 7. Effect of the topology of communities in $\mathcal{C}(\mathbf{SIR}_{\beta,\gamma}, G, m)$.

CRediT authorship contribution statement

G. Hernández: Conceptualization, Methodology, Writing – original draft, Visualization. **A. Martín del Rey:** Methodology, Writing – review & editing, Supervision.

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.

Acknowledgments

One of the authors (G.H.) gratefully acknowledges the Junta de Castilla y León, Spain and the European Regional Development Fund for financial support.

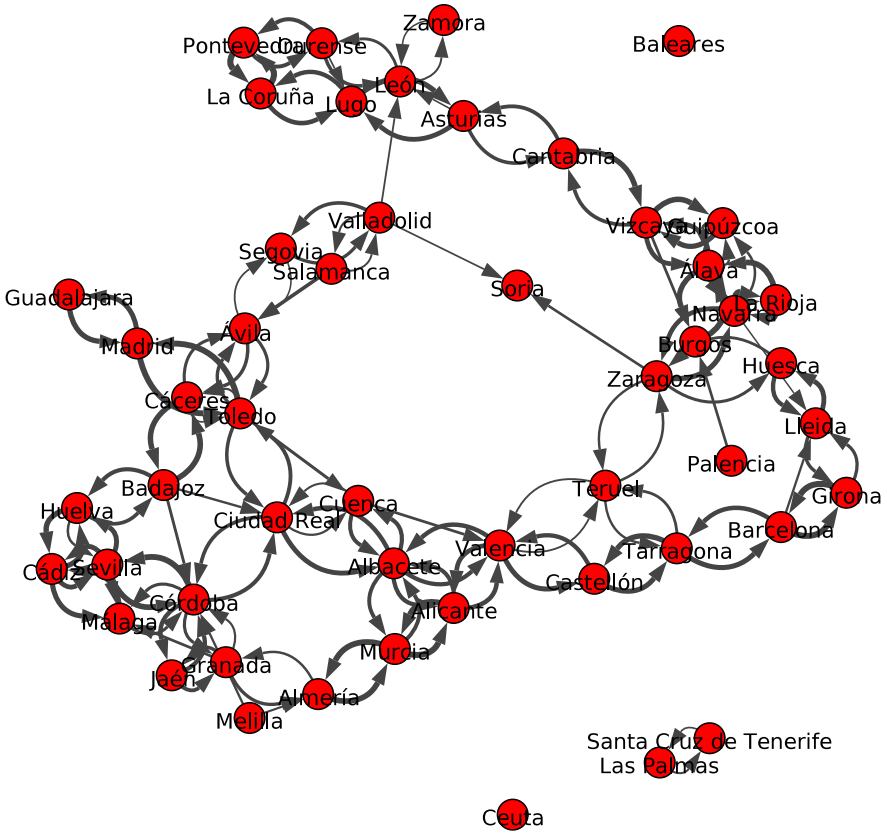

Fig. 8. Graphical depiction of the mobility model M_{Spain} extracted from [11]. The size of the edges is proportional to the quantiles of the values in its matrix. The node layout was obtained by the Fruchterman–Reingold algorithm [12], using the igraph implementation [13].

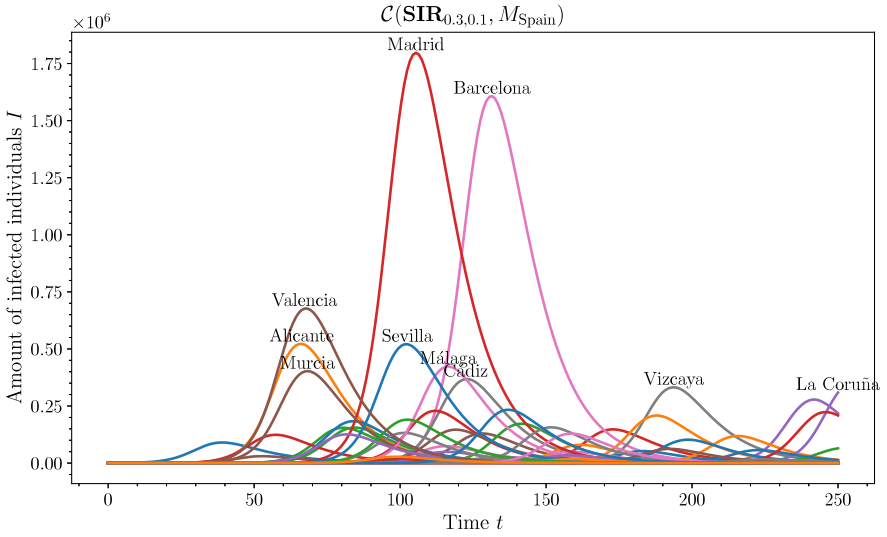

Fig. 9. Infected population on a infection following a $\mathbf{SIR}_{\beta,\gamma}$ model with the mobility data from [11] starting from Albacete. Labels were added to the most prominent peaks, which correspond to higher population areas.

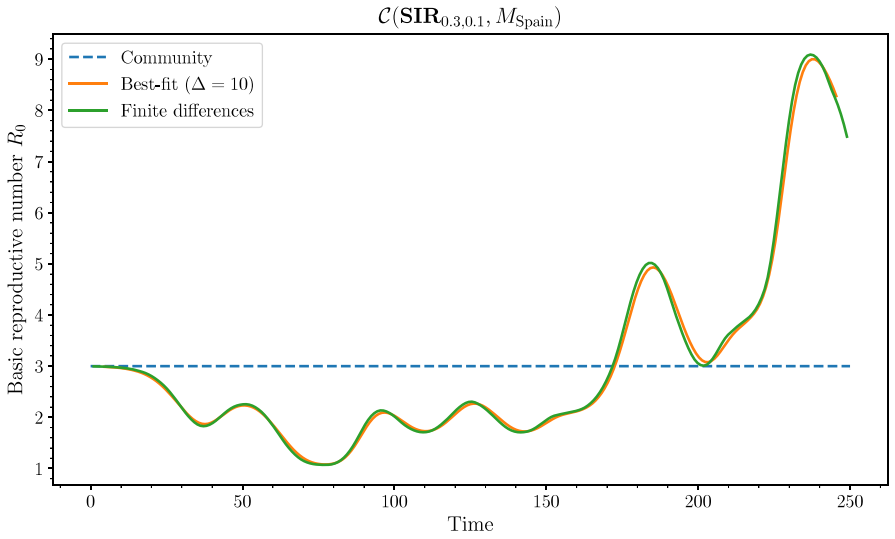

Fig. 10. Effective values of the basic reproductive number R_0 in the $\mathcal{C}(\mathbf{SIR}_{0.3,0.1}, M_{\text{Spain}})$ model.

References

[1] X. Wu, Z. Liu, How community structure influences epidemic spread in social networks, Physica A 387 (2–3) (2008) 623–630.
[2] J.-p. Zhang, Z. Jin, Epidemic spreading on complex networks with community structure, Appl. Math. Comput. 219 (6) (2012) 2829–2838.
[3] W. Pan, G.-Q. Sun, Z. Jin, How demography-driven evolving networks impact epidemic transmission between communities, J. Theoret. Biol. 382 (2015) 309–319.
[4] K.F. Nipa, L.J. Allen, Disease emergence in multi-patch stochastic epidemic models with demographic and seasonal variability, Bull. Math. Biol. 82 (12) (2020) 1–30.
[5] N.A. Kudryashov, M.A. Chmykhov, M. Vigdorowitsch, Analytical features of the SIR model and their applications to COVID-19, Appl. Math. Model. 90 (2021) 466–473.
[6] M.N. Alenezi, F.S. Al-Anzi, H. Alabdulrazzaq, Building a sensible SIR estimation model for COVID-19 outspread in Kuwait, Alexandria Eng. J. 60 (3) (2021) 3161–3175.
[7] P. Singh, A. Gupta, Generalized SIR (GSIR) epidemic model: An improved framework for the predictive monitoring of COVID-19 pandemic, ISA Trans. (2021) in press, corrected proof.
[8] D. Prodanov, Analytical parameter estimation of the SIR epidemic model. applications to the COVID-19 pandemic, Entropy 23 (1) (2021) 59.
[9] B. Malavika, S. Marimuthu, M. Joy, A. Nadaraj, E.S. Asirvatham, L. Jeyaseelan, Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models, Clin. Epidemiol. Global Health 9 (2021) 26–33.
[10] D. Chen, On the integrability of the SIR epidemic model with vital dynamics, Adv. Math. Phys. 2020 (2020).
[11] Instituto Nacional de Estadística, Evolución de la movilidad por ámbito geográfico durante el estado de alarma por COVID-19, 2020, https://www.ine.es/covid/covid_movilidad.htm. (Accessed 12 August 2020).
[12] T.M. Fruchterman, E.M. Reingold, Graph drawing by force-directed placement, Softw. - Pract. Exp. 21 (11) (1991) 1129–1164.
[13] G. Csardi, T. Nepusz, The igraph software package for complex network research, Inter. J. Complex Syst. (2006) 1695.
[14] J.R. Dormand, P.J. Prince, A family of embedded Runge-Kutta formulae, J. Comput. Appl. Math. 6 (1) (1980) 19–26.
[15] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, I. Polat, Y. Feng, E.W.V. Moore, J. erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, Contributors, SciPy 1. 0, SciPy 1.0: Fundamental algorithms for scientific computing in python, Nature Methods 17 (2020) 261–272.
[16] J.J. Moré, The Levenberg-Marquardt algorithm: Implementation and theory, Lect. Not. Math. 630 (1978) 105–116.
[17] G. Hernández, Community-models repository, 2021, https://github.com/Dih5/community-models. (Accessed 10 February 2021).