Chaos, Solitons and Fractals 160 (2022) 112156

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Assessing the risk of pandemic outbreaks across municipalities with mathematical descriptors based on age and mobility restrictions

Alejandro Carballosa [a,b], José Balsa-Barreiro [c,d], Pablo Boullosa [a,e], Adrián Garea [a,e], Jorge Mira [e,f], Ángel Miramontes [c], Alberto P. Muñuzuri [a,b,*]

[a] Group of Nonlinear Physics, Fac. Physics, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain
[b] Galician Center for Mathematical Research and Technology (CITMAga), 15782 Santiago de Compostela, Spain
[c] Institute IDEGA, Department of Geography, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain
[d] MIT Media Lab, Massachusetts Institute of Technology, 75 Amherst St, Cambridge, MA 02139, USA
[e] Departamento de Física Aplicada, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain
[f] Instituto de Materiais (iMATUS), Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 31 March 2022
Accepted 24 April 2022
Available online 26 May 2022

Keywords:
SEIR pandemic model
Risk factors
Confinement scenarios
Linear stability/growth factors
Multilayer network with mobility and geographical data

ABSTRACT

By March 14th 2022, Spain is suffering the sixth wave of the COVID-19 pandemic. All the previous waves have been intimately related to the degree of imposed mobility restrictions and its consequent release. Certain factors explain the incidence of the virus across regions revealing the weak locations that probably require some medical reinforcements. The most relevant ones relate with mobility restrictions by age and administrative competence, i.e., spatial constrains. In this work, we aim to find a mathematical descriptor that could identify the critical communities that are more likely to suffer pandemic outbreaks and, at the same time, to estimate the impact of different mobility restrictions. We analyze the incidence of the virus in combination with mobility flows during the so-called *second wave* (roughly from August 1st to November 30th, 2020) using a SEIR compartmental model. After that, we derive a mathematical descriptor based on linear stability theory that quantifies the potential impact of becoming a hotspot. Once the model is validated, we consider different confinement scenarios and containment protocols aimed to control the virus spreading. The main findings from our simulations suggest that the confinement of the economically non-active individuals may result in a significant reduction of risk, whose effects are equivalent to the confinement of the total population. This study is conducted across the totality of municipalities in Spain.

1. Introduction

Mathematical models describing the evolution of an epidemic spread have been comprehensively studied in previous years with approaches that cover both deterministic and stochastic natures [1]. Based on the classic SIR family of compartmental models, a large number of simulations have been carried out in order to assess the evolution of the COVID-19 spread [2–6]. Many of them analyzed the SARS-CoV-2 virus spreading on several countries by implementing different impact scenarios, which could help government authorities to adopt non-pharmaceutical interventions and mobility restrictions in order to contain the virus [7]. In this sense, metapopulation models [8] have become very useful and adequate thanks to their ability to incorporate both the age population structure and the human mobility within a region. The interaction between dwellers from neighboring cities and the daily flux of commuters is accounted in order to explicitly model the spatial evolution of the virus, obtaining rich and valuable information that can then be used to assess the impact of different mobility policies, in addition to their efficiency and feasibility [9–11]. Besides, due to the non-uniform impact of COVID-19 on the age stratification [12], separating the population in clearly distinguished age groups has also been reported to be a key feature in successfully tackling the reproduction of the pandemic [9,13–15]. However, increasing refinement and complexity in the model exponentially increases computational costs, whereas data analysis becomes more intricate and intractable. This poses the difficult challenge of finding a suitable descriptor that could quantify all the factors related while minimizing the number of variables.

In this study, we focus on the interplay between the different infectivity patterns caused by socio-spatial distinctions and mobility flows

* Corresponding author at: Group of Nonlinear Physics, Fac. Physics, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain.
E-mail address: alberto.perez.munuzuri@usc.es (A.P. Muñuzuri).

https://doi.org/10.1016/j.chaos.2022.112156

among localities. By using the dynamic parameters of the model and data related to mobility infrastructures, we define a mathematical descriptor that is solely based on analytical calculations [15]. This descriptor is able to evaluate the spatial behavior of the virus across any region by considering different administrative units such as municipalities in the Spanish case. Thus, it is possible to infer what municipalities are more susceptible of suffering an outbreak. Using data related to the pandemic incidence in previous weeks and mobility measures adopted as input, the descriptor returns a specific value unique to each population, that tells us how fast one outbreak could spatially evolve and highlights the critical hotspots by considering the population structure in a certain region. In a previous study, we validated this descriptor by considering the municipalities (lowest-level administration units) in a very particular region, Galicia [16], whereas in this paper we refine its analysis and implement it to the whole country, Spain.

The paper is organized as follows. First, we introduce the study area based on its spatiotemporal context. After that, in Section 3 we describe the mathematical model as well as the tools developed for the analysis. In Section 4 we validate the model by verifying that the model predictions were fulfilled during the second pandemic wave in Spain. In Section 5 we analyze different scenarios with mobility restrictions in order to achieve a significant reduction of the COVID-19 impact within this region. Finally, the discussion of the results and brief conclusions are presented in Section 6.

2. Study area

In this section, we introduce the study area in two successive subsections by considering both the spatial context (Sub-section 2.1) and the COVID-19 impact over time (Sub-section 2.2).

2.1. The spatial context

Administratively, Spain is fragmented into seventeen regions (so-called *Autonomous Communities*) and two autonomous cities with similar governmental status. All of them correspond to the first level of territorial administration, just after the central government. The second administrative level corresponds to provinces (50 in total), while the third one corresponds to municipalities (8131).

In the case of COVID-19 management, we must note that Spain is an administratively decentralized country with most of the relevant competences in health management transferred to the regions. While during the first wave (since COVID-19 emergence to around June 30th, 2020) the national government took charge of imposing the lockdown measures, the regional authorities mostly managed the successive waves considering the general guidelines suggested from the national government [17].

According to the most recent census year (2021), Spain has 47.35 million inhabitants [18]. Uniquely four regions (Andalucia, Catalonia, Madrid and Valencia) concentrate around 58.8% of the population in the 29.7% of the total national area. The population density is about 92 people per km^2, although spatially distributed in a very unbalanced way. Most of the population is concentrated in the coast, especially in the southern and eastern sectors, and some inland urban spots such as the metropolitan area of Madrid, besides a reduced number of cities with more than 250,000 inhabitants such as Zaragoza, Córdoba and Valladolid. In recent decades, the process of urbanization and concentration of the population has been very marked across multiple scales [19], with more than 80% of urban population and 51% of people living in high-density agglomerations [20]. At municipal level, 61.5% of them have fewer than 1000 inhabitants and only 9.2% have more than 10,000 inhabitants.

During the COVID-19 pandemic, the Spanish Statistical Office (in short, INE) periodically collected mobility data nation-wide [21]. A huge dataset with individual locations over time for more than 80% of cellphones was registered. People inflows and outflows were aggregately presented across *mobility areas*. The whole country is spatially fragmented into 3200 mobility areas, whose delimitation partially coincides with the municipal division. These mobility areas contain between 5000 and 50,000 inhabitants. In depopulated regions, these mobility areas will be the sum of tiny, populated municipalities (less than 5000 inhabitants), whereas in urban areas they correspond with urban districts.

In this work, we run the mathematical model in the 3121 most populated municipalities. This corresponds to the 38.4% of the total number of municipalities, but they contain around 99% of the Spanish population concentrated in an area that is equivalent to the 58.3% of the whole Spain. We analyze demographic data by municipalities, whereas the flows of people collected from the INE's mobility areas are estimated by municipalities (grouping or separating those mobility areas that do not territorially match the municipality area) (Fig. 1).

2.2. COVID-19 territorial impact

Up to the present date in Spain, more than 10 million cases and around 100,000 deaths from COVID-19 have been confirmed [22]. The first wave produced officially 28,985 deaths and 130,749 infections, this latter one a clearly underestimated figure that could be up to 10 times higher. During the second wave (until Dec 2020), the number of deaths was about 20,504 people, reaching the sum of 46,646 people since the emergence of the virus. The third wave (until Feb 2021) meant death for 19,200 people. The fourth wave (until May 2021) accounted for more than 5170 deaths and 288,445 infections. The fifth wave (until mid–Sept 2021) added 5800 deaths and about 1.3 million new infections. All the data are estimates, offering important differences with the official mortality monitoring (MoMo) database that shows differences with the officially reported mortality rates [23].

Fatality data related to COVID-19 clearly shows very relevant inequalities across the country. The four most populated regions were the most affected (58.4% of deaths). However, the fatality rate (per 100,000 inhabitants) shows how most of the interior provinces registered much higher fatality rates compared to the coastal and island provinces (i.e. Soria: 497 deaths per 100,000 people; Las Palmas: 35 in September 23rd, 2021) [24,25].

3. Methods

This section is subdivided in the following sub-sections: Age-structured epidemic model (Sub-section 3.1), metapopulation network and mobility map (Sub-section 3.2), and mathematical descriptors and linear stability analysis (Sub-section 3.3).

3.1. Age-structured epidemic model

We use a deterministic SEIR model with individuals separated into different compartments according to infection status: susceptible (S), exposed (E), infected (I) and removed (R); and also according to age: less than 20 years old (AG1), 20–39 (AG2), 40–69 (AG3), and 70 years and older (AG4). Inside each one of the 16 compartments, the main assumption is that the population is distributed homogeneously. Following a similar approach as in [26], social mixing is then included considering different interaction rates between each possible combination of age groups. On the other hand, the infection transmission works as follows: susceptible individuals of any age in compartment l can become exposed at a rate β^{lr} via interaction with infected individuals of any age from compartment r. Exposed ones become infectious with rate ν and the infected ones are either removed or recovered with rate μ. This process is summarized in the flowchart shown in Fig. 2.

We consider that the total population is conserved, so that at the location i the population of age group l will be equal to the sum of all their

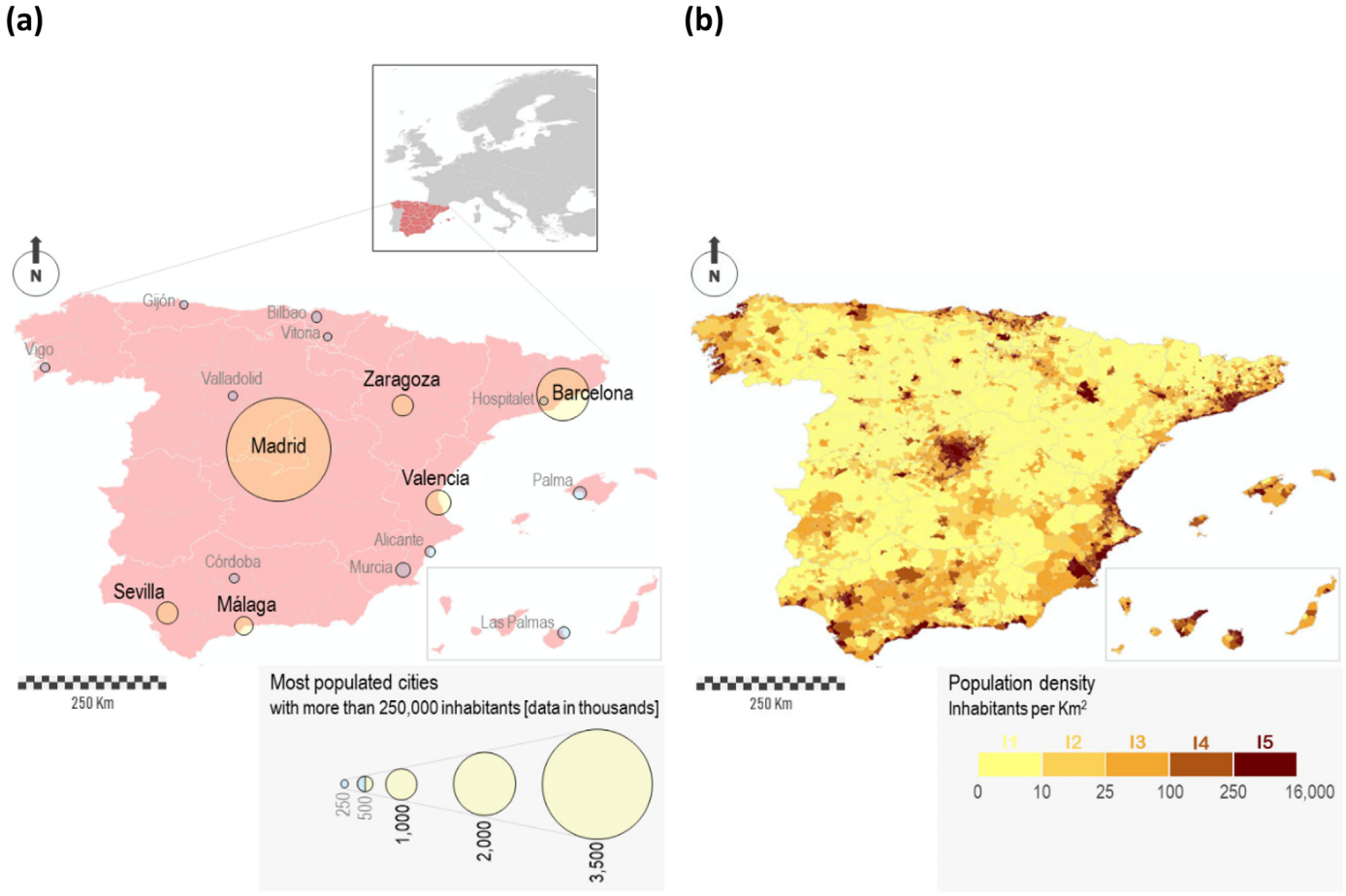


Fig. 1. (a) Location of the study area and characterization of its urban system based on population. Representation of all the cities with more than 250,000 people. (b) Population density by municipalities.

respective compartments, that is $S_i^l + E_i^l + I_i^l + R_i^l = N_i^l$. In mathematical terms, epidemic transitions for a given age group l at any municipality i are described by the equations,

$$\begin{cases} \dfrac{dS_i^l}{dt} = -\sum_{r=1}^{4} \dfrac{1}{2}\beta_i^r I_i^r \left(\dfrac{1}{N_i^l} + \dfrac{1}{N_i^r}\right)\beta_i^l S_i^l \\ \dfrac{dE_i^l}{dt} = \sum_{r=1}^{4} \dfrac{1}{2}\beta_i^r I_i^r \left(\dfrac{1}{N_i^l} + \dfrac{1}{N_i^r}\right)\beta_i^l S_i^l - \nu E_i^l \\ \dfrac{dI_i^l}{dt} = \nu E_i^l - \mu I_i^l \\ \dfrac{dR_i^l}{dt} = \mu I_i^l \end{cases} \tag{1}$$

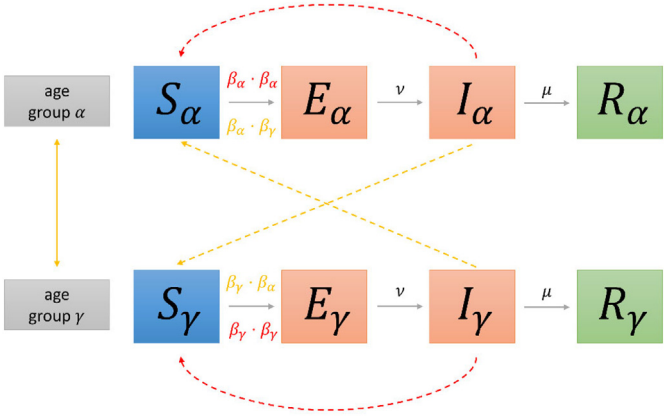


Fig. 2. Compartments of the epidemic model for any two age groups α and γ. The solid grey arrows indicate the transition probabilities between compartments, the red dashed arrows indicate the interactions between infected and susceptible individuals of the same age group and the yellow dashed arrows the interactions among individuals of different age groups.

where the sum in r is performed over the four age compartments and $l = r$ is allowed. At this stage, the spatial distribution of the population is not yet included. This set of differential equations is then fitted using the function *curve_fit* implemented in Python, which is based on non-linear least squares minimization to the regional cumulative incidence data of the pandemic's second wave allowing to estimate the parameters β_i^l (one for each age group). This is, we fit the total cumulative incidence of each community i to $\sum_{l=1}^{4}(I_i^l(t) + R_i^l(t))$, computed by means of a Runge-Kutta integration method. Fitting the curves of each age group to its separated data was pursued unsuccessfully due to small irregularities in the datasets and lack of further parameters that could fit complicated epidemic curves. Besides, the pairwise infective transmission rates β^{lr} were factorized into two separate terms $\beta^{lr} = \beta^l \cdot \beta^r$ to gain simplicity during the fitting process. This way, instead of pursuing a complex estimation of sixteen parameters at a time (the 4×4 elements of the β^{lr} matrix), we only fit four parameters simultaneously. Note that the fitting process is carried out in 3121 municipalities, representing the

99% of the Spanish population. The factorization of the infective transmission rates will also result useful in the linear stability analysis described below. The cumulative incidence data was collected from the Health Institute Carlos III (ISCIII) [27] for the period covering the second wave, which ranges from August 1st, 2020 to November 30th, 2020.

3.2. Metapopulation network and mobility map

Once the epidemic incidence across municipalities is estimated, we introduce a mobility-based coupling between each pair of communities including demographic data obtained from the Spanish Institute of Statistics [18]. The nature of the coupling is based mainly on people commuting among the different regions for whatever reason. This flux of mobile individuals between any two locations i and j is represented in the model by the quantity p_{ij}, then normalized by the local population in each community i, N_i. In order to consider different mobilities among the age groups, we define the matrix q_{ij}^l, whose elements follow,

$$q_{ij}^l = \begin{cases} 0 & i = j \\ r_l \cdot \dfrac{p_{ij}}{N_i^l} & i \neq j \end{cases} \tag{2}$$

where l refers to a specific age group and r_l is a control parameter that allows us to tune the relative mobilities among different age groups such that the age group l holds a percentage r_l of the total mobility. One can think of this whole mathematical abstraction as a weighted metapopulation network [8], where nodes represent the number n of municipalities and the weighted links the amount of mobility between them. These weighted connections are represented mathematically by the adjacency matrix C_{ij}^l, which is directly built from q_{ij}^l,

$$C_{ij}^l = \begin{cases} \dfrac{1}{2}\left(1 + \sum_k^n \left(q_{ki}^l - q_{ik}^l\right)\right) & i = j \\ q_{ij}^l & i \neq j \end{cases} \tag{3}$$

In other words, in the same municipality ($i = j$), C_{ij}^l shows the proportion of people present at node i taking into account the people of age l that leaves node i, $\sum_k^n q_{ik}^l$, and the people of age l that enters i from node k, $\sum_k^n q_{ki}^l$. On the other hand, when $i \neq j$, C_{ij}^l directly represents the normalized flux of individuals of age l from node i to node j. Thus, we can introduce the complete model,

$$\begin{cases} \dfrac{dS_i^l}{dt} = -\sum_{j=1}^n \sum_{r=1}^4 \left(\dfrac{C_{ji}^l}{N_i^l} + \dfrac{C_{ij}^l}{N_j^r}\right) \beta_j^r I_j^r \beta_i^l S_i^l \\ \dfrac{dE_i^l}{dt} = \sum_{j=1}^n \sum_{r=1}^4 \left(\dfrac{C_{ji}^l}{N_i^l} + \dfrac{C_{ij}^l}{N_j^r}\right) \beta_j^r I_j^r \beta_i^l S_i^l - \nu E_i^l \\ \dfrac{dI_i^l}{dt} = \nu E_i^l - \mu I_i^l \\ \dfrac{dR_i^l}{dt} = \mu I_i^l \end{cases} \tag{4}$$

In close analogy to what occurred in the fitting process, the idea behind the complete model is that susceptible individuals from node i and age group l can become exposed when they come into contact with infected individuals of all age groups of node j when they either travel to it (represented by the mobility term C_{ij}^l/N_j^l) or when the infected individuals from j are the ones traveling to i (C_{ji}^l/N_i^l). Population normalization is carried out by considering the population where the infective interactions take place. Although this normalization should consider the local variations due to daily mobility, we will assume that these are negligible regarding the total population in each municipality

$$N_i^l + \sum_j^n \left(p_{ji}^l - p_{ij}^l\right) \approx N_i^l \tag{5}$$

Thus, we can perform the desired mathematical analysis by avoiding intricate terms related to the denominator's derivatives (this will become evident in the next section). Notice that when we only check interactions within the same municipality ($i = j$), the mobility term is repeated and thus we need to introduce the 1/2 factor in Eq. (3). Also notice that, if we do not consider mobility between regions, q_{ij} will be zero for all i and j, leading to $C_{ij} = 1/2$ when $i = j$, turning our mobility-based model (Eq. (4)) into the simple model we used for fitting β (eq. 1).

With the complete model, we can estimate a mathematical descriptor that can give us a preliminary assessment of how quickly an outbreak could develop in the different nodes, based on both the force of infection of each node and the mobility weight of its links. In other words, we can evaluate the risk that each network element holds against the pandemic. Furthermore, we can easily evaluate the evolution of this descriptor in different possible scenarios by tuning either the force of infection of a given age group (β_i^l), the proportion of mobility corresponding to each group (r_l), or the mobility between different communities (q_{ij}). This means isolating a specific age group from the rest or to cut down the mobility across the country either by limiting interactions within the municipalities or selected groups of municipalities.

The mathematical descriptor at hand is based on linear stability theory and network theory [28]. Given a dynamical system in a stationary state, it analyzes how each node of the network will respond to a small perturbation of the state based on its internal dynamics and the interaction pattern existing in the network. Prior studies demonstrated how these descriptors are heavily influenced by both the network topology and the mathematical form of the dynamic system [29]. This was recently tested on an epidemical model by considering a relatively small geographical network of 313 municipalities that corresponds to a particular region of Spain [16,30,31]. In this paper, we come up with the following hypothesis: considering as the unstable stationary state of the system the situation where all individuals belong in the susceptible compartment, we introduce a small perturbation to the system as a small initial number of infected individuals into every node. In this scenario, the epidemic in each node should evolve at different rates, based on its local dynamics and interaction patterns, acquiring a unique positive value for the mathematical descriptor. The higher this value, the more quickly will this small perturbation grow, indicating a stronger impact of the disease. Thus, we obtain a quantitative measure for each node, which we can then compare to reproduce a nation-wide map of infection risk.

3.3. Mathematical descriptors and linear stability analysis

We simplify the index notation and the summatory by combining all $4\,n \times n$ adjacency matrices C^l into one, with dimension $(4 \cdot n) \times (4 \cdot n)$. This means that, in the complex network, each municipality is now split into 4 nodes, each representing a different age group. In other words, we had a network with four interconnected layers (one layer for each age group) and now we have a clearer, more tractable mathematical situation where each pair of nodes has only a single link. With this transformation, Eq. (4) become,

$$\begin{cases} \dfrac{dS_i}{dt} = -\sum_{j=1}^n \left(\dfrac{C_{ji}}{N_i} + \dfrac{C_{ij}}{N_j}\right) \beta_j I_j \beta_i S_i \\ \dfrac{dE_i}{dt} = \sum_{j=1}^n \left(\dfrac{C_{ji}}{N_i} + \dfrac{C_{ij}}{N_j}\right) \beta_j I_j \beta_i S_i - \nu E_i \\ \dfrac{dI_i}{dt} = \nu E_i - \mu I_i \\ \dfrac{dR_i}{dt} = \mu I_i \end{cases} \tag{6}$$

which are easier to work with from the point of view of our scheme. Let us recall then a stationary state of the system, namely $\left(\overline{S_i}, \overline{E_i}, \overline{I_i}, \overline{R_i}\right)$, and let us introduce in each node a small perturbation to the state:

Table 1
Spearman's correlation rank for the four age groups and for different illustrative combinations of the relative mobility r_l.

r_l (AG1-AG2-AG3-AG4)	0–19 (AG1)	20–39 (AG2)	40–69 (AG3)	70+ (AG4)
0.1-0.4-0.4-0.1	0.7272	0.7916	0.7817	0.5345
0.01-0.49-0.49-0.01	0.7222	0.7933	0.7835	0.5342
0.2-0.3-0.3-0-2	0.7270	0.7905	0.7669	0.5363
0.25-0.25-0.25-0.25	0.7283	0.7894	0.7569	0.5370

$$(S_i, E_i, I_i, R_i) = \left(\overline{S_i}, \overline{E_i}, \overline{I_i}, \overline{R_i}\right) + (\delta S_i, \delta E_i, \delta I_i, \delta R_i)$$

After that, we can apply linear stability theory in order to estimate the growth factor as described in [16,28,29] for each node in the network. The mathematical derivation of this analysis is included in the Supplementary Information. These factors describe the growth rate of the perturbation to the non-stable epidemic-free steady, allowing us to estimate the vulnerability of each node.

4. Results

This section is subdivided in the following sub-sections: Fitting the model and validation (Sub-section 4.1) and different scenarios for epidemic control (Sub-section 4.2).

4.1. Fitting the model and validation

All parameters β_i are estimated fitting the model to the incidence data as commented in the *Methods* section, while parameters μ and ν are extracted from [11]. To get the virus transmissibility for each municipality, we made non-linear least squares fits using data extracted from mobility areas. In many cases, mobility areas match with municipalities, but some neighboring small municipalities are grouped into one mobility area, while the bigger ones are split into two or more mobility areas. This becomes useful for the smaller ones, since they do not have enough cases to make the fits properly. The same parameters that are obtained from these fits are used for the municipalities that make up the mobility areas. For the bigger localities, we fit them directly, without splitting them in any way, as the pandemic data do not differentiate the different mobility areas corresponding to a single town. We have, therefore, a set of 4 β_i parameters for each one of the 3121 municipalities we are considering. To have a measure of the goodness of the fits, we use the Pearson's correlation coefficient [32], with values from 0 to 1, to compare the observed curves for each age group with the simulated ones using the model given by Eq. (1). Their distribution, represented in the Supplementary Information (Fig. S2), shows that for most municipalities the obtained β_i parameter is in good agreement with the data, finding correlation values around 0.8 or higher for most of the cases. The rest of the cases for which fitting was unfeasible were usually linked to municipalities with really low incidence or with data not smooth enough for the fitting to converge, achieving underestimated values for the β_i parameters. The latter was especially true for the age group concerning people older than 70 years, where data were the most irregular showing major differences from our simulated infection curves.

Once we obtained the local infectivity transmission rates β_i, we computed the distribution of growth rates (GR from here on) for each age group and compared them with the maximum values of cumulative incidence data of each age group to adjust the relative mobilities r_l from Eq. (2). For this, we use the Spearman's rank correlation coefficient [32], which finds correlations among the monotonic tendencies of two given variables. We opt for this coefficient because GR do not share a linear relation with the cumulated cases and thus the Pearson correlation loses applicability.

For simplification, we shall refer to the group of people with ages between 0 and 19 as AG1 (age group 1), to people with ages 20–39 as AG2, people from 40 to 69 as AG3, and people with more than 70 years as AG4. Following the same trend as with the infectivities (see Fig. S2 of the Supplementary Information), we obtained higher correlation values for age groups 2 and 3 while medium ones for age group 4, as shown on Table 1 for different combinations of r_α. The latter is a result of the less convergent fitting for the infectivities. For this age group, there is a higher scattering in the GR values around the populations of 10^4 than in the incidence data, which induced the lower values in the correlation. The purpose of comparing these two variables is not to find a perfect agreement between the points in both series, since we expect the GR to represent the inherent risk or the probability of suffering an outbreak

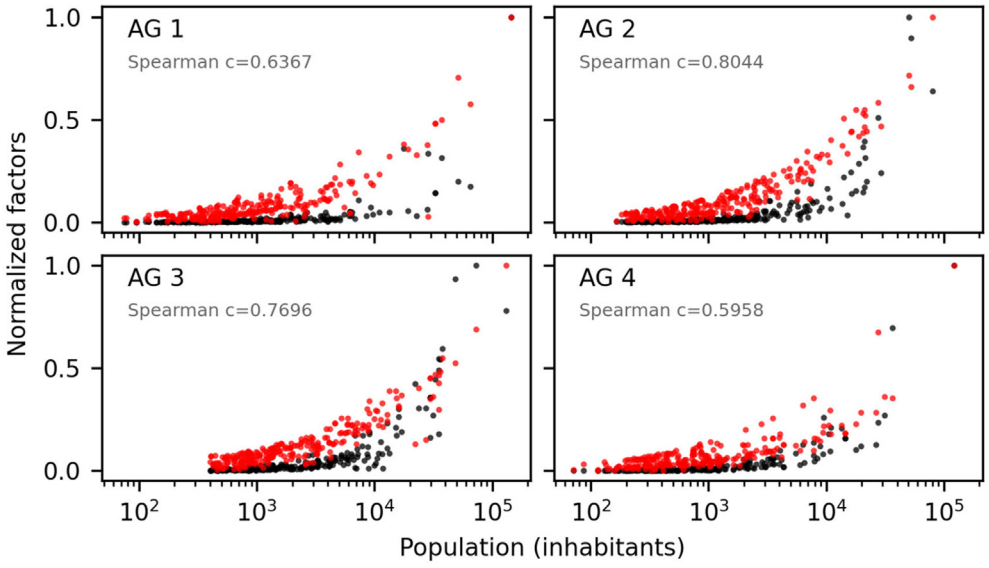
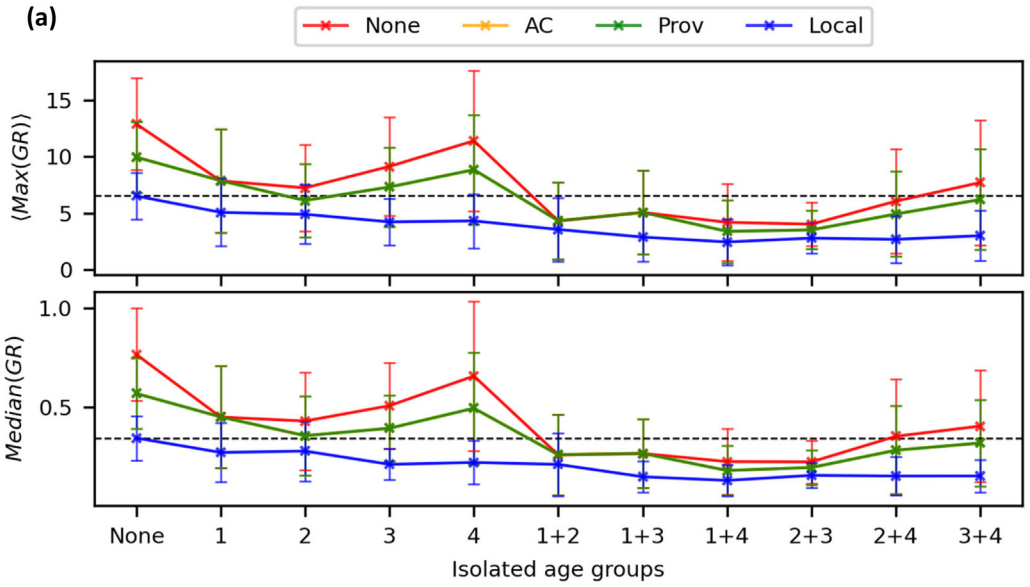
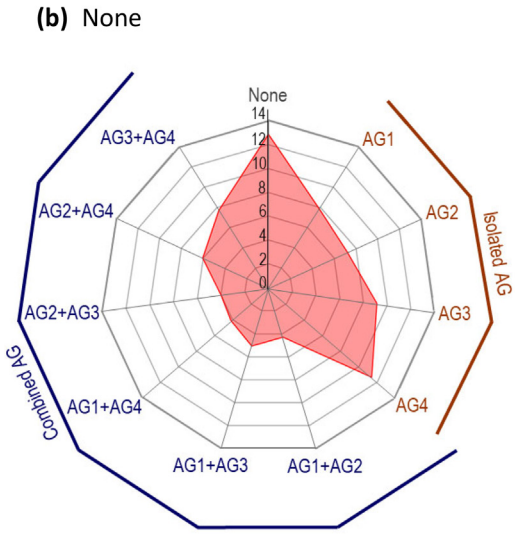
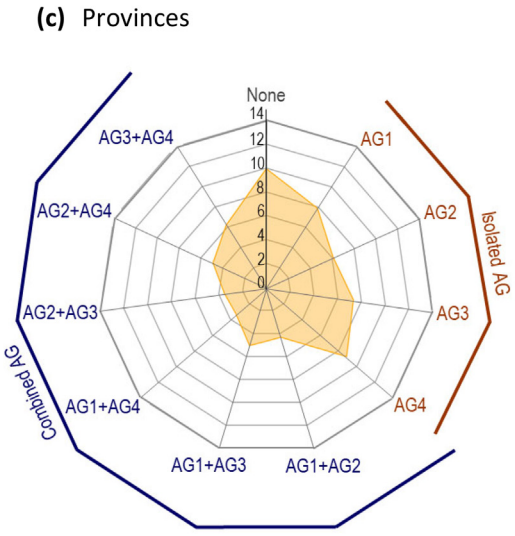
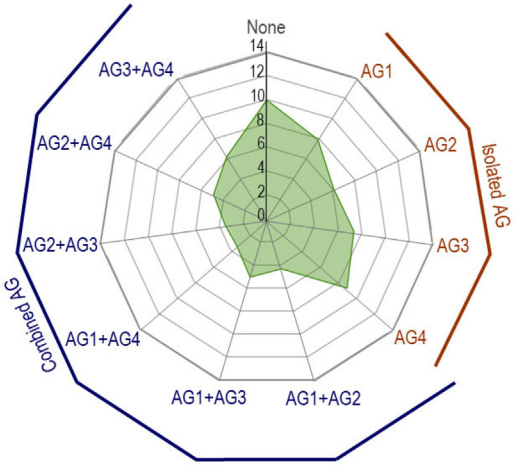


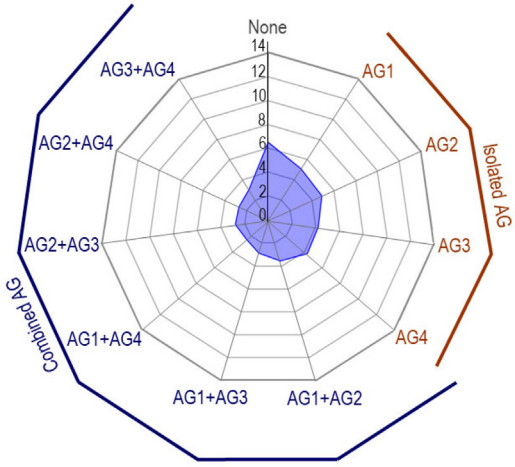
Fig. 3. Sample distribution of Growth Rates (red) and maximum cumulative incidence data (black) against the node's population for the case with relative mobility of (0.1, 0.4, 0.4, 0.1). Each variable was normalized to the maximum value its own series so the comparison can be feasible. Only 10% of the points are shown for visualization purposes.


(a)
None
AC
Prov
Local
⟨Max(GR)⟩
Median(GR)
None
1
2
3
4
1+2
1+3
1+4
2+3
2+4
3+4
Isolated age groups


(b) None
None
AG1
AG2
AG3
AG4
AG1+AG2
AG1+AG3
AG1+AG4
AG2+AG3
AG2+AG4
AG3+AG4
Isolated AG
Combined AG


(c) Provinces
None
AG1
AG2
AG3
AG4
AG1+AG2
AG1+AG3
AG1+AG4
AG2+AG3
AG2+AG4
AG3+AG4
Isolated AG
Combined AG


(d) AC
None
AG1
AG2
AG3
AG4
AG1+AG2
AG1+AG3
AG1+AG4
AG2+AG3
AG2+AG4
AG3+AG4
Isolated AG
Combined AG


(e) Local
None
AG1
AG2
AG3
AG4
AG1+AG2
AG1+AG3
AG1+AG4
AG2+AG3
AG2+AG4
AG3+AG4
Isolated AG
Combined AG

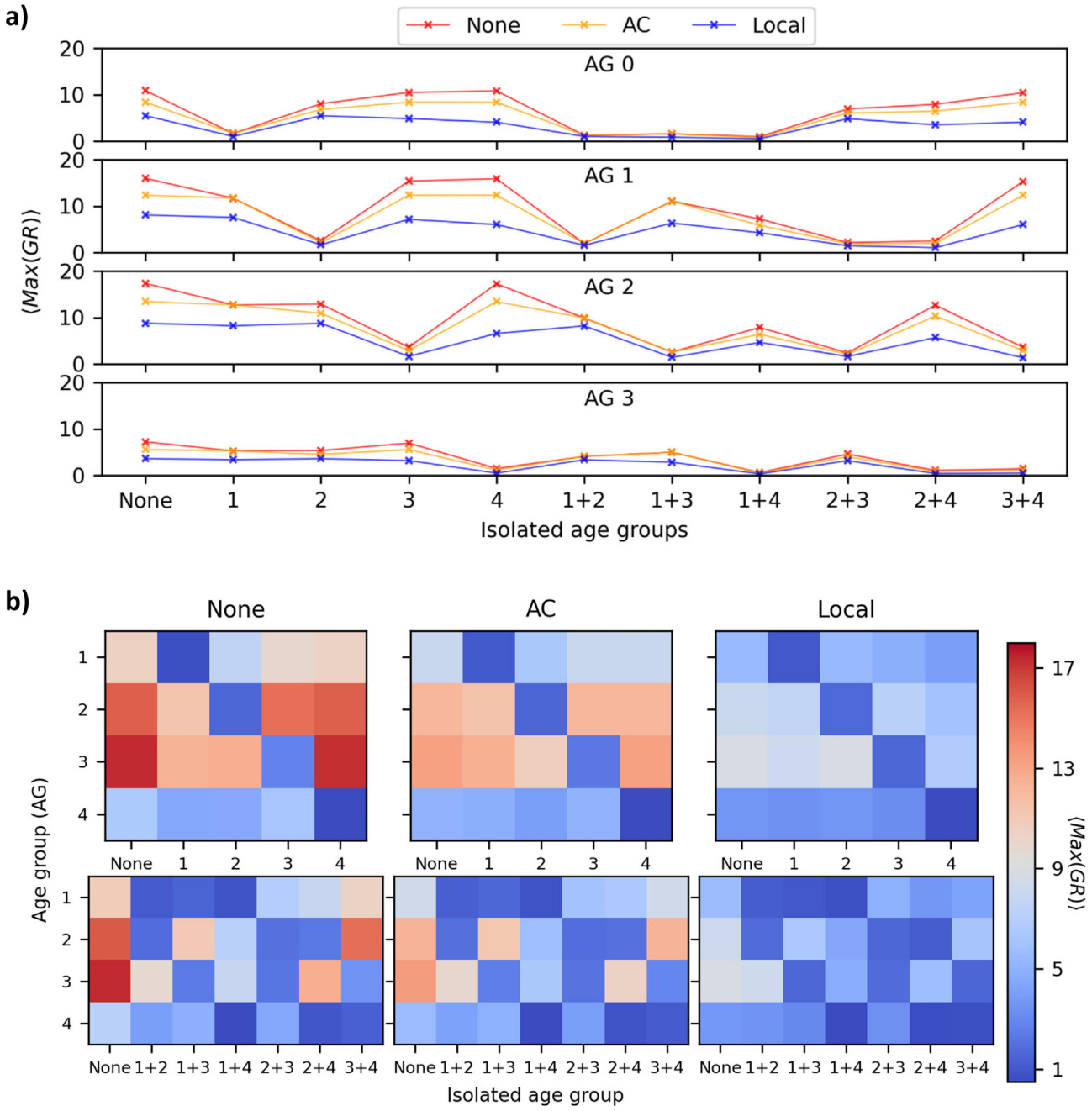


Fig. 5. (a) Average of the 10 maximum values of the GR distributions for each specific age group in each scenario. The numbers on the x labels refer to isolation of a single or a combination of age groups, while restrictions to the mobility are shown in different colors (al) and consecutive columns (b). (b) Summary of the previous results in a color-coded map. Relative mobility between municipalities is fixed to (0.1, 0.4, 0.4, 0.1).

which is in nature a very different quantity than the number of cumulative cases. Thus, any community with non-zero mobility flux may have a non-zero value of the descriptor although its actual incidence was almost null. Regarding the free parameter r_l, there were no significant differences for these combinations attending to the correlation coefficient, as it is shown in Table 1. In the following, we shall choose the combination $r_l = (0.1,0.4,0.4,0.1)$ as the main one for our computations and further analysis, setting out a clear difference between the adult population (groups 2 and 3) which contains the main bulk of working population that is involved in periodic mobility

Fig. 4. (a) Average of the 10 maximum values (a, upper panel) and the median values (a, lower panel) of the growth rates (GR) distributions for different scenarios of isolation and mobility restriction for different age groups. The points represent averaging over the four age groups and the error bars the standard deviation for each case. The black dashed line is drawn to guide the eye when comparing the two control measures at hand, this is, local mobility restrictions with no isolation to no-mobility restrictions with age groups isolation. Panels (b–e) itemize the information of the top upper panel by mobility restriction, showing the impact of isolating different age groups in each situation. Here, only the average of the four age groups is shown. It must be remarked that the mobility restriction by autonomous communities' results in a very similar outcome as the mobility restriction by provinces, as can be grasped from panels (c, d). Thus, in panels (a) these lines appear overlapped. Relative mobility between municipalities is fixed to (0.1, 0.4, 0.4, 0.1).

patterns, from the retired elderly people and the younger individuals who are still attending school. The GR distribution for this configuration is shown in Fig. 3 where they are compared against the data of cumulated cases. Based on this comparison, the correlation values of Table 1 can be better understood. We can conclude that the growth rates calculated as indicated in the Methods section constitute an appropriate parameter to describe the pandemic risk at each location. For the curious reader, the tabled values of the growth rates represented in Fig. 3 are presented at the end of the Supplementary Information.

4.2. Different scenarios for epidemic control

In the following, we will consider different scenarios and using the GR will estimate the best options to minimize the effects of a pandemic wave. The implemented model allows two different approaches to halt a possible epidemic outbreak. The first one consists in severing the mobility across different scales nationwide modifying the mobility fluxes of matrix C_{ij} (and so the set of eigenvalues Λ_α of the network's Laplacian matrix that are used in the computing of the growth rates). The second one consists of removing or isolating a given age group (or a combination of two age groups) by turning their infectivity β_i to almost zero. The latter translates to these groups being unable to become carriers of the disease, so it could be interpreted either as a totally effective protection or simply a home isolation of the given group. Notice that this is different to equating the parameter r_l to zero since the latter means mobility among different municipalities and the former would also affect the infections at a local scale. Regarding the mobility scales, Spain is hierarchically subdivided into three different scales (regional autonomous communities, provinces and municipalities). Following the different levels of restrictions posed by the government during the different virus waves, we will analyze the effect of breaking down the mobility across these three different scales.

Attending to these considerations, we summarized our main findings on Figs. 4 and 5. Since we obtained one distribution of growth rates for each age group and mobility scenario, we define a set of suitable statistics to summarize all the information at hand. We opted for using the upper bound and the median value of each distribution, being the former an average of the ten highest growth rates. This gives us an idea of which are the highest values of the distribution (highest risk factors then) while reducing the bias of the most populated cities, which show exponential growing values. Thus, the upper bound anticipates how fast the outbreaks emerge in the most critical locations, whereas the median value shows the behavior in most of them. Note that we discard the mean value since it is highly biased by the outcome in the upper bounds of the distributions. We consider the following different scenarios with different age groups confined or restricted mobility. Panel (a) of Fig. 4 presents a summary view of the results obtained. Each statistic indicator was weighted over the four age groups for each situation, being included the standard deviation of the average value as a dispersion estimator. Panels (b–e) of Fig. 4 represent the same information itemized by mobility restrictions and in radial shape, while Fig. 5 breaks down all this information by age group, highlighting the impact of the different configurations in each of the groups.

Regarding the mobility restrictions, we observe that cutting the mobility down by provinces has a very small effect on the GR values, while the control measure at local a scale seems to be the most efficient, decreasing the overall risk in more than half in almost every scenario. Note that in Fig. 4 the points corresponding to the restrictions adopted by autonomous communities (AC) are hidden behind, overlapped by the points corresponding to the provinces. We believe that this is related to the mobility data rather than a model prediction, since during the second wave there were still some restrictions to travel even among different provinces and thus the existing variations were not strong enough to induce a different outcome with our model.

Focusing now only on the protection by age groups, it is especially significant the impact that comes from isolating a combination of AG1 with any of the other groups. These combinations without mobility restrictions (red line), decrease the overall GR distributions below the situation with local level mobility restrictions and no age isolation

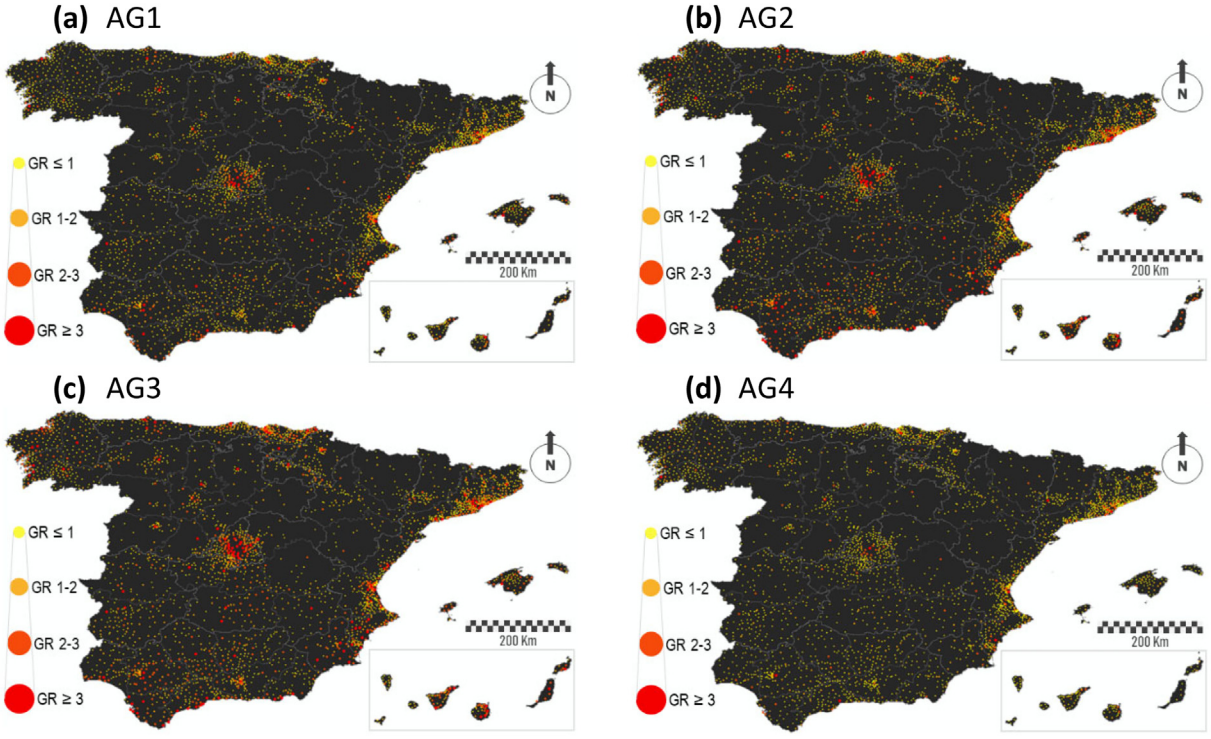


Fig. 6. Spatial maps of the growth rates without mobility restrictions by age group. Relative mobility between municipalities by age groups is fixed to (0.1, 0.4, 0.4, 0.1).

measures (first point of the blue line). This is highlighted on panel (a) of Fig. 4 by the dashed black line, where we can also see that isolating AG2 and AG3 at the same time returns a similar outcome. In particular, a combination of AG1 and AG4 seems to be the one that best reduces the GR in the remaining age groups, as Figs. 4 and 5 suggest. While isolating a substantial portion of the population logically halts the virus spreading, it is interesting that the rest of the combinations with AG4 cannot achieve such improvements on the other age groups. Attending more specifically to Fig. 5, it is noticeable that isolating either AG3 or AG4 keeps the values on AG1 and AG2 almost equivalent with the scenario with no isolations. This is mainly related to the fact that AG4 has in general lower transmission rates in our fittings than the other three groups, but it is also slightly biased with the underestimation of some of these rates from the fitting issues commented above. In the same way, isolating only a single age group of population barely modifies the GR obtained on the remaining groups, suggesting that it would be a pointless measure to adopt solely according to our model.

5. Discussion

The aim of our approach lies in developing a single and efficient tool in which several scenarios can be simulated while keeping the computational cost minimized. We present a mathematical descriptor that is able to incorporate empiric information on commuting flows and the individual infective dynamic to recreate the infective risk map across municipalities. An age stratification in each municipality allows for control measures that aim to slow down the virus impact by reducing the age-structure mixing in the population. The manual modification of both variables, mobility flows and population mixing, allows to minimize the risk of each node while comparing different mobility restrictions in order to keep in operation the active people, which means those between 20 and 69 years old (i.e., AG2 + AG3).

From a purely mathematical perspective, no assumptions were made over the age-specific transmissibility of COVID-19. Thus, the relative rates of infection transmission were obtained from fitting the model to the cumulative infection data series by municipalities. Since our model is minimal and it is based only on four homogeneous compartments, we believe that this local-oriented fitting helped us capture the unique features of each municipality by introducing the necessary heterogeneity to accurately reproduce the pandemic evolution. One of the main results from our analysis shows that isolating individuals younger than 19 (AG1) shows a particular impact on the growth rates (GR), especially when the isolation restrictions are combined with any of the other groups. In fact, the combination of the latter with the eldest people (AG4) seems to be the one to reduce the most the risk factors in the other remaining groups; closely followed by a combined isolation of AG2 and AG3, that is, between 20 and 69 years old. However, the latter would imply a total freezing of the population's active sector that is one of the less desirable outcomes. On the other hand, our mobility analysis shows that, given the mobility traffic allowed during the second wave, only restrictions adopted by municipalities would have significantly reduced the risk of pandemic outbreaks. Furthermore, by combining all the control measures by mobility and age, one of the key implications of our model is that the already discussed combination of AG1 with any other achieves a further reduction in the risk factors than imposing local mobility restrictions across the whole country. This could suggest that efficiently protecting AG1 and AG4 simultaneously while allowing AG2 and AG3 to pursue their normal activity would achieve one of the optimal scenarios with the ongoing pandemic. Although our model does not incorporate a time variable and therefore is not able to predict for how long these restrictions should hold, it does show some interesting new insights on the age-structure mixing that could help to adopt.

In Spain, the distribution of growth rates is mainly determined by the spatial pattern of population distribution. The metropolitan area of Madrid, the Mediterranean coastline (especially in the eastern sector) and certain areas located on the north coast concentrate the largest urban regions. These urban regions show high economic dynamism and young age-structures. Highest growth rates are observed in these regions. The inland provincial cities emerge as isolated foci presenting significantly high growth rates. In a scenario without restrictions, such as the one shown in Fig. 6, growth rates trace this spatial pattern showing highest values in the youth (AG1) and the active population groups (AG2 + AG3).

The scenario shown in Fig. 7 represents the growth rates in the active groups (A2 + A3) by considering mobility restrictions in the youngest and eldest groups (A1 + A4). The effects on the active population describe a very similar effect, showing much lower impact values in growth rates. The yellow hue dominates most of the map, whereas the red tones are restricted to main cities and most populated urban regions.

The rate between growth rates obtained in a free-mobility scenario (Fig. 6) and the one with mobility restrictions in AG1 and AG4 (Fig. 7) is estimated for the intermediate age groups, i.e. AG2 + AG3, in Fig. 8. As we can observe, major reduction rates are visible on the suburban regions, while moderate reduction rates are observed precisely in the main cities and urban regions.

The present methodology could be easily applied to other regions with a population structure heterogeneously distributed and, also, to other kinds of spreading phenomena whose mechanisms can be modelled mathematically. Following the line presented here, further studies will be mandatory in near future to determine how to optimal trade-off between the expected positive effect on public health with the negative impact on freedom of movement, economy impact, and society at large.

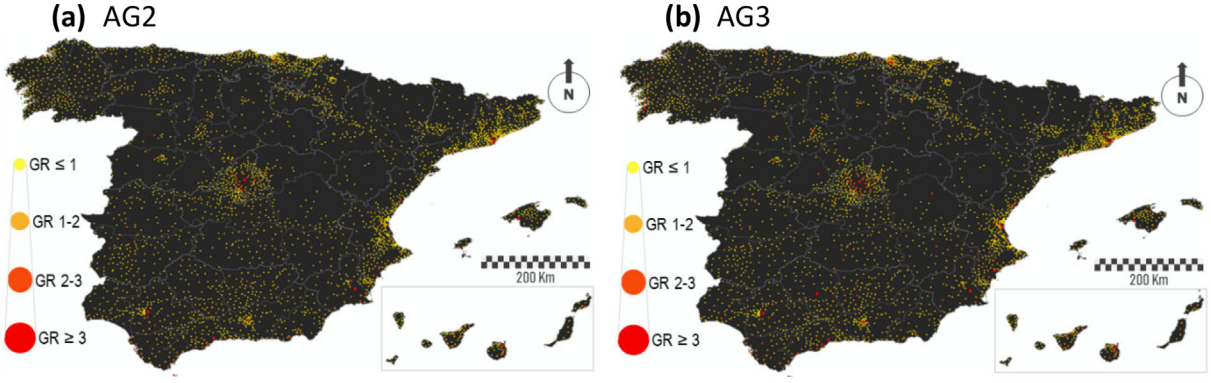


Fig. 7. Growth rates in AG2 and AG3 with mobility restrictions in AG1 and AG4. Relative mobility between municipalities by age groups is fixed to (0.1, 0.4, 0.4, 0.1).

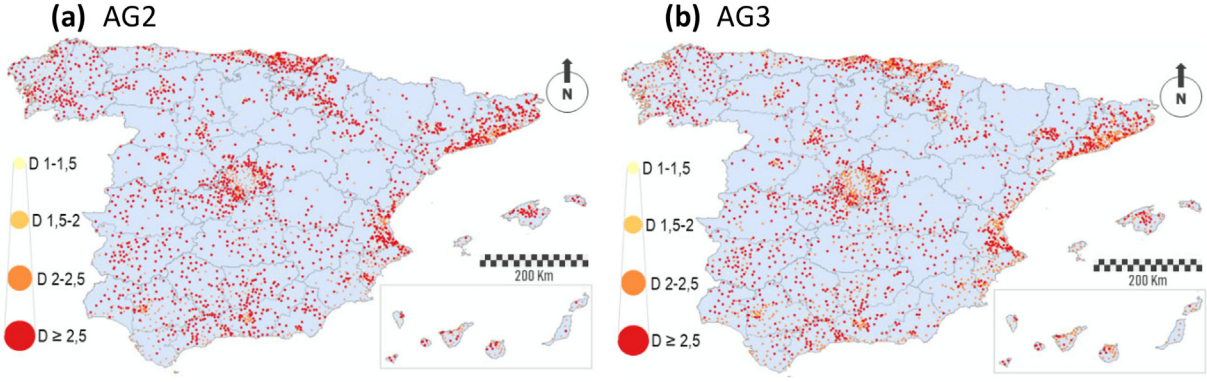


Fig. 8. Reduction in the growth rates in active groups AG2 and AG3 by applying mobility restrictions to AG1 and AG4. Relative mobility between municipalities by age groups is fixed to (0.1, 0.4, 0.4, 0.1).

6. Conclusions

In this article we implement a method based on a SEIR compartmental model for analyzing the COVID-19 impact in combination with mobility flows. This method allows quantifying the potential impact for becoming a hotspot. For the study, we consider different confinement scenarios and containment protocols aimed to control the virus spreading. These are based on mobility restrictions by age and administrative competences. Our results show how the confinement of the economically non-active individuals results in a significant reduction of risk, whose effects are equivalent to the confinement of the total population.

Ethics approval

The research presented in this manuscript does not involve confidential medical data neither involves actual patients. All epidemiological data was freely available, anonymous and is referred along the manuscript as first shown.

CRediT authorship contribution statement

A. Carballosa performed most of the numerical analysis in the text. P. Boullosa and A. Garea contributed to the model fits and with some of the calculations. A. Carballosa, J. Balsa-Barreiro, J. Mira, A. Miramontes and A.P. Muñuzuri designed the project and wrote de manuscript. A.P. Munuzuri coordinated the project. All authors read and approved the final manuscript.

Declaration of competing interest

Authors declare they have no financial interests.

Acknowledgements

We gratefully acknowledge financial support by the Spanish Ministerio de Ciencia e Innovación and European Regional Development Fund under contract RTI2018-097063-B-I00 AEI/FEDER, UE, and by Xunta de Galicia under Research Grant No. 2021-PG036. AC and APM are part of the CITMAga Strategic Partnership and JM is part of iMATUS, both supported by the Galician Government. All these programs are co-funded by FEDER (EU). This research was also supported by the Health Institute Carlos III, within the Project COV20/00617 in the scope of the "Fondo COVID" of the Spanish Ministerio de Ciencia e Innovación, and by the crowd funding program "Sumo Valor" of the Universidade de Santiago de Compostela. We obtained data from the Spanish National Center for Epidemiology (Instituto de Salud Carlos III). AC acknowledges financial support from the Galician Government. A. M. acknowledges the Galician Innovation Agency (GAIN) of the Xunta de Galicia. We also acknowledge the support of the CESGA Supercomputing Center (Galicia, Spain), where a substantial portion of the simulations were run.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chaos.2022.112156.

References

[1] Walters CE, Meslé MMI, Hall IM. Modelling the global spread of diseases: A review of current practice and capability. Epidemics. 2018;25(12):1–8.
[2] Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, Flasche S, Clifford S, Pearson CAB, Munday JD, Abbott S, Gibbs H, Rosello A, Quilty BJ, Jombart T, Sun F, Diamond C, Gimma A, van Zandvoort K, Funk S, Jarvis CI, Edmunds WJ, Bosse NI, Hellewell J, Jit M, Klepac P. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health. 2020;5(5):e261–70.
[3] Giordano G, Blanchini F, Bruno R, Colaneri P, Filippo AD, Matteo AD, Colaneri M. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nat Med. 2020;26(4):855–60.
[4] Hoertel N, Blachier M, Blanco C, Olfson M, Massetti M, Rico MS, Limosin F, Leleu H. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. Nature Medicine. 2020;26(7):1417–21.
[5] Ndaïrou F, Area I, Nieto JJ, Torres DFM. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. Chaos, Solitons & Fractals. 2020;135(6):109846.
[6] Kimathi M, Mwalili S, Ojiambo V, Gathungu DK. Age-structured model for COVID-19: Effectiveness of social distancing and contact reduction in Kenya. Infect Dis Model. 2021;6(1):15–23.
[7] Silva CJ, Cruz C, Torres DFM, Munuzuri AP, Carballosa A, Area I, Nieto JJ, Fonseca-Pinto R, da Fonseca RP, dos Santos ES, Abreu W, Mira J. Optimal control of the COVID-19 pandemic: controlled sanitary deconfinement in Portugal. Sci Rep. 2021;11:3451.
[8] Colizza V, Vespignani A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. J Theor Biol. 2008;251(4):450–67.
[9] Arenas A, Cota W, Gómez-Gardeñes J, Gómez S, Granell C, Matamalas JT, Soriano-Paños D, Steinegger B. Modeling the spatiotemporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions. Phys Rev X. 2020;10(12):041055.
[10] Calvetti D, Hoover AP, Rose J, Somersalo E. Metapopulation network models for understanding, predicting, and managing the coronavirus disease COVID-19. Frontiers in Physics. 2020.;8(6).
[11] Bertuzzo E, Mari L, Pasetto D, Miccoli S, Casagrandi R, Gatto M, Rinaldo A. The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. Nat Commun. 2020;11:8.
[12] CDCMMWR. Severe outcomes among patients with coronavirus disease 2019 (COVID-19) — United States, February 12–March 16, 2020. 2020.;69.
[13] Balabdaoui F, Mohr D. Age-stratified discrete compartment model of the COVID-19 epidemic with application to Switzerland. Sci Rep. 2020;10(12):21306.
[14] Li M, Zu J, Li Z, Shen M, Li Y, Ji F. How to reduce the transmission risk of COVID-19 more effectively in New York City: an age-structured model study. Front Med. 2021;8:641205.
[15] Colombo RM, Garavello M. Optimizing vaccination strategies in an age structured SIR model. Mathematical Biosciences and Engineering. 2020;17:1074–89.
[16] Carballosa A, Balsa-Barreiro J, Garea A, García-Selfa D, Miramontes Á, Muñuzuri AP. Risk evaluation at municipality level of a COVID-19 outbreak incorporating relevant geographic data: the study case of Galicia. Sci Rep. 2021;11(10):21248.

[17] Villalón JC, González RCL, Carballada ÁM, Cantos JO. Le Covid-19 en Espagne. Les défis géopolitiques de la gouvernance d'un problème de santé. Herodote. 2021; 183:197–225.
[18] Spanish Statistical Office. Population and demography data, [Online]. Available: https://www.ine.es. [Accessed 2021].
[19] Balsa-Barreiro J, Morales AJ, Lois-González RC. Mapping population dynamics at local scales using spatial networks. Complexity. 2021;2021:e8632086.
[20] Martínez ER, Gisbert FJG, Martí IC. Delimitación de áreas rurales y urbanas a nivel local. Bilbao: Fundación BBVA; 2016.
[21] Spanish Statistical Office. Mobility statistics during COVID-19, [Online]. Available: https://www.ine.es/en/covid/covid_movilidad_en.htm?L=1.
[22] Expansion - datos macro, [Online]. Available: https://datosmacro.expansion.com/otros/coronavirus/espana.
[23] Mortality Monitoring (MoMo). Institute of Health Carlos III. https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/MoMo/Paginas/Informes-MoMo-2020.aspx; 2020.
[24] ElDiario.es. El coronavirus en España: mapas y últimos datos de los casos y su evolución. Available: https://www.eldiario.es/sociedad/mapa-datos-coronavirus-espana-comunidades-autonomas-diciembre-17_1_1039633.html; 2022.
[25] Zuil MLDOMaría. El Confidencial. Available: https://www.elconfidencial.com/espana/2021-09-27/mapa-covid-municipios-mortalidad-espana_3295460/; 2021.
[26] Zhao Z-Y, Zhu Y-Z, Xu J-W, Hu S-X, Hu Q-Q, Lei Z, Rui J, Liu X-C, Wang Y, Yang M, Luo L, Yu S-S, Li J, Liu R-Y, Xie F, Su Y-Y, Chiang Y-C, Zhao B-H, Cui J-A, Yin L, Su Y-H, Zhao Q-L, Gao L-D, Chen T-M. A five-compartment model of age-specific transmissibility of SARS-CoV-2. Infect Dis Poverty. 2020;9(8):117.
[27] Instituto de Salud Carlos III (ISC-III), [Online]. Available: https://www.isciii.es/Paginas/Inicio.aspx. [Accessed 2021].
[28] Nakao H, Mikhailov AS. Turing patterns in network-organized activator–inhibitor systems. Nat Phys. 2010;6(4):544–50.
[29] Mimar S, Juane MM, Park J, Muñuzuri AP, Ghoshal G. Turing patterns mediated by network topology in homogeneous active systems. Physical Review E. 2019;99(6): 062303.
[30] Miramontes Carballada Á, Balsa-Barreiro J. Territorial impact of the COVID-19 pandemic in Galicia (Spain): a geographical approach. Bol AGE. 2021.;91 Asociación de Geógrafos Españoles.
[31] Miramontes A, Balsa-Barreiro J. Geospatial analysis and mapping strategies for fine-grained and detailed COVID-19 data with GIS. ISPRS Int J Geo Inf. 2021;10:602.
[32] Kokoska DZaS. de CRC standard probability and statistics tables and formulae. Chapman & Hall; 2000. Section 14.7.