

Article

Modeling Epidemic Spread among a Commuting Population Using Transport Schemes

Daniela Calvetti ^{1,*} , Alexander P. Hoover ² , Johnie Rose ³  and Erkki Somersalo ¹ ¹ Department of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University, Cleveland, OH 44106, USA; ejs49@case.edu² Department of Mathematics, University of Akron, Akron, OH 44325, USA; ahoover1@uakron.edu³ Center for Community Health Integration, Case Western Reserve University, Cleveland, OH 44106, USA; jxr109@case.edu

* Correspondence: dxc57@case.edu

Abstract: Understanding the dynamics of the spread of COVID-19 between connected communities is fundamental in planning appropriate mitigation measures. To that end, we propose and analyze a novel metapopulation network model, particularly suitable for modeling commuter traffic patterns, that takes into account the connectivity between a heterogeneous set of communities, each with its own infection dynamics. In the novel metapopulation model that we propose here, transport schemes developed in optimal transport theory provide an efficient and easily implementable way of describing the temporary population redistribution due to traffic, such as the daily commuter traffic between work and residence. Locally, infection dynamics in individual communities are described in terms of a susceptible-exposed-infected-recovered (SEIR) compartment model, modified to account for the specific features of COVID-19, most notably its spread by asymptomatic and presymptomatic infected individuals. The mathematical foundation of our metapopulation network model is akin to a transport scheme between two population distributions, namely the residential distribution and the workplace distribution, whose interface can be inferred from commuter mobility data made available by the US Census Bureau. We use the proposed metapopulation model to test the dynamics of the spread of COVID-19 on two networks, a smaller one comprising 7 counties in the Greater Cleveland area in Ohio, and a larger one consisting of 74 counties in the Pittsburgh–Cleveland–Detroit corridor following the Lake Erie’s American coastline. The model simulations indicate that densely populated regions effectively act as amplifiers of the infection for the surrounding, less densely populated areas, in agreement with the pattern of infections observed in the course of the COVID-19 pandemic. Computed examples show that the model can be used also to test different mitigation strategies, including one based on state-level travel restrictions, another on county level triggered social distancing, as well as a combination of the two.

Keywords: metapopulation model; COVID-19 pandemic; parameter estimation

Citation: Calvetti, D.; Hoover, A.P.; Rose, J.; Somersalo, E. Modeling Epidemic Spread among a Commuting Population Using Transport Schemes. *Mathematics* **2021**, *9*, 1861. <https://doi.org/10.3390/math9161861>

Academic Editor: Francesca Pitolli

Received: 6 July 2021

Accepted: 29 July 2021

Published: 5 August 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the outbreak of SARS-CoV-2, the virus which causes COVID-19, in Wuhan, China at the end of 2019 and its rapid spread to nearly every country in the world within a few months, there has been great interest in understanding the transmission of the disease and the dynamics of the epidemic to design and implement suitable control and containment measures. There are many challenges when it comes to making predictions about COVID-19 spread. The novelty of the virus makes the population level data collected during previous epidemics with similar characteristics, e.g., SARS and MERS, obsolete. In addition, the potential for contact-based, droplet, and airborne transmission of SARS-CoV-2 [1], and the significant, though not fully known, extent of asymptomatic transmission contribute to the high infection rate and add to the difficulty of containing the epidemic without resorting to rather drastic measures. It is in this context that mathematical models

play a significant role in understanding the pattern of spread of the disease and evaluating the effectiveness of different measures for its containment and mitigation. A factor of interest is the human mobility and, in particular, certain persistent and recurring mobility patterns such as commuter traffic which may be relatively well understood and predictable, and likely targets of mitigation measures.

Mathematical models have long been used to investigate and predict the dynamics of infectious diseases. Classical population models for the spread of epidemics [2–4] partition the population into cohorts, comprising susceptible (S), infected (I) and recovered (R) individuals, leading to SIR models. A fourth compartment is often included for exposed (E) individuals undergoing an incubation period. The many modifications to the SIR and SEIR type models reveal the inherent flexibility of this modeling paradigm. These compartment models, while useful for understanding the overall dynamics of an epidemic and providing fairly realistic predictions for a homogeneous population, may not provide sufficient resolution when the population consists of communities with different socio-urban characteristics or demographics. In fact, compartment models assume that the population is well mixed, in the sense that all individuals in each cohort are similar, and changes to the state of a compartment affect its entire population equally and simultaneously [5–7]. In many cases this uniformity assumption is not satisfied; indeed, the transmission rate can depend on a number of factors that change according to geographic location, e.g., degree of social interaction, density of the population, and mobility. One way to account for the heterogeneity of the population is by introducing dynamic network metapopulation models. Here we propose a novel network metapopulation model specifically designed to study the spread of the COVID-19 epidemic through regional recurrent travel patterns.

The limitations of compartment models due to the homogeneous assumptions are well understood and have been addressed in the context of metapopulation models, see, e.g., [4] and references therein. Metapopulation models, initially introduced for insect pests [8] and later used in conjunction with networks to include a spatial dimension in modeling the transmission of infectious diseases, can be used to track the geographic patterns of infectious disease spread. Metapopulation network models can be viewed as a framework to connect crowded urban communities and sparsely populated rural ones, and the local uniformity justifies the use of mean field models to track the spread of the epidemics within each individual community. In that sense, metapopulation network models are a compromise between the more granular agent-based models and mean field models; for a comparison of large scale computational approaches to epidemic modeling, in particular the agent-based approach versus structured metapopulation models, we refer to [9]. Metapopulation network models have been used to understand the spread of previous epidemics. In [10], a network metapopulation model was employed to follow the spreading of SARS and the outbreaks of A(H1N1) and A(H7N9) influenza, known as avian flu, three epidemics characterized by a rapid, though less rapid than SARS-CoV-2, spread over a wide region. Based on the modeling, the effectiveness of the suppression of international flights from countries with high prevalence and severe restrictions of local traffic to contain a pandemic have been questioned [11–14]. For a recent dynamic metapopulation network model for COVID-19 spread in a network of cities in China, see, e.g., [15].

Metapopulation models for infectious diseases assume a discrete population structure and are typically based on adjacency matrices: The infection may spread from one node to another if the nodes are adjacent to each other in the infection network, and the underlying hypothesis of spread mechanism is based on unspecified contacts between individuals at different nodes. Commuter traffic, on the other hand, is characterized by relatively regular temporal redistribution of the population, following a well understood pattern. Importantly, unlike the adjacency-based contacts that are hard to quantify, there are relatively solid census-based data describing the commuter traffic in quantitative terms. The main novelty of the proposed metapopulation model is its adherence to transport schemes, commonly encountered in optimal transport theory. A transport scheme describes

a lossless redistribution of a given density in a prescribed way. This framework is ideal to describe human mobility in a closed system, in which the population, originally distributed according to the census data into different communities, is temporarily redistributed as described, e.g., by traffic patterns due to daily commute. The same framework is also ideal for describing the redistribution of the population as inferred by anonymous cell phone data. A common feature is that the connectivity between different communities can be modeled in terms of a lossless transformation of a population density map. In addition, the description of the mobility in terms of population distributions can be extended to the continuous case, that is, the formalism can be developed without specifying a priori a discretization level. This flexibility may be useful to describe sparsely populated areas in which well defined population aggregates such as cities or towns are not descriptive. Unlike in reaction-diffusion-based models (see, e.g., [16]), the propagation process of the infection is non-local in the sense that infection can travel long distances along the highways used by commuters, rather than propagating to the immediate neighborhood due to the local increase in infected population.

Our network metapopulation model is rather general and can be adapted to different scales, as illustrated in the computed examples section where it is used to investigate the geographical pattern of the spread of the COVID-19 epidemic. The network structure in our examples utilizes publicly available mobility data from the US Census database [17] and can be customized for arbitrary groups of counties in a straightforward manner. The dynamics of the epidemic in each node follow the SE(A)IR model proposed in [18], which differs from a standard SEIR model by including in the exposed (E) compartment a subcohort (A) of infectious asymptomatic individuals. In the first computed example, we consider a small network consisting of seven counties in the Greater Cleveland area in Northeast Ohio, where the central metropolitan community constitutes a natural hub linking the surrounding communities through commuter traffic. The dynamic infection rate is estimated with a Bayesian particle filter algorithm from county-level daily infection count data. The second, larger and more challenging example considers a network of 74 counties constituting a corridor south of lake Erie, connecting the Pittsburgh (PA), Cleveland (OH) and Detroit (MI) metropolitan areas. The urban centers are linked by major highways that also determine the commuter traffic, which in turn is encoded in the transport scheme. These models, which can be used as templates for studying the traffic-based spread of COVID-19, provide a realistic test bench for the planning of different mitigation scenarios to contain the virus. A pilot version of the metapopulation network models, whose theoretical and computational details are the main theme of the current contribution, was used in a preliminary study [19] to understand spatial patterns of the spread of COVID-19 in a network of 18 counties in Northeast Ohio, including the greater Cleveland area, and in a network of 19 counties in Southeast Michigan, including the greater Detroit area, respectively. The present article provides a detailed mathematical derivation and description of the model. Finally, the capabilities of an extension to a continuous spatially distributed scheme is demonstrated by a simulated model.

2. Discrete Community Metapopulation Model

The foundation of our model for the spread of the epidemic in the individual communities is the well known SEIR model of Kermack and McKendrick [2], which we briefly review below to establish the notation used in the rest of the paper. In this well-mixed lumped compartment model, the population is partitioned into four non-overlapping cohorts: susceptibles (S), exposed (E), infected (I) and recovered (R), satisfying the system of differential equations

$$\begin{aligned}\frac{dS}{dt} &= -\beta\rho\frac{I}{N}S, & \frac{dE}{dt} &= \beta\rho\frac{I}{N}S - \eta E, \\ \frac{dI}{dt} &= \eta E - (\gamma + \mu)I, & \frac{dR}{dt} &= \gamma I,\end{aligned}\tag{1}$$

where $\beta > 0$ is the infection rate per contact, or infection rate for short, ρ is the number of contacts per day, $\eta = 1/T_{\text{inc}}$ is the incubation rate with expected time of incubation of T_{inc} days, $\gamma = 1/T_{\text{rec}}$ is the expected recovery rate with expected recovery time T_{rec} days, and μ is the death rate. Standard SEIR models assume that exposed individuals are not infective. Two key parameters in SEIR models are the infection rate β that depends not only on the characteristics of the pathogen, but also on the type of contacts between individuals, and the contact rate ρ that, although often modeled as constant, is sensitive to a number of factors, e.g., social distancing and other mitigation measures.

One of the major limitations of the above model is its lack of resolution, in particular when the underlying population is partitioned into several complex interconnected communities, with different contact rates, population sizes and densities. To address this heterogeneity and connectivity, we first propose a network version of the SEIR model with a discrete set of subpopulations, and customize it to include the most salient features of the COVID-19 epidemic which affect transmission. Subsequently, we present a generalization of the model suitable for continuous population distributions.

2.1. Commuter Traffic as a Transport Scheme: Discrete Case

Consider a community C of N individuals distributed over L communities, denoted by C_j , $1 \leq j \leq L$. The population size needs not be fixed, and it may indeed change due to the mortality rate of the disease. For simplicity, we do not include population changes due to immigration, emigration, or births.

We want our model of the community network to take into account regular commuter traffic between the nodes C_j , as illustrated schematically in Figure 1. Although a variety of different factors contribute to the traffic between communities, e.g., travel to restaurants, sporting events, movies, theater and music, and other leisure or professional attractions, we simplify the picture here by focussing on the consistent travel associated with commuting between home and work.

Let $N \in \mathbb{R}^{L \times L}$ be the contact matrix describing the commuter traffic, with entries

$$n_{jk} = \# \text{individuals residing in } C_j \text{ and working in } C_k, \quad 1 \leq j, k \leq L. \quad (2)$$

In the derivation of our network-based infection model, we relate the matrix N to two probability densities. More specifically, let H and W be two random variables taking on values in the state space $\Phi = \{1, 2, \dots, L\}$, with probability mass functions

$$\pi_j^{(H)} = P\{H = j\} = \text{probability that an individual resides in } C_j, \quad (3)$$

$$\pi_j^{(W)} = P\{W = j\} = \text{probability that an individual works in } C_j. \quad (4)$$

These probabilities are interpreted in terms of frequencies and can be expressed in terms of the contact matrix as scaled marginals,

$$\pi_j^{(H)} = \frac{\# \text{ individuals residing in } C_j}{N} = \frac{1}{N} \sum_{k=1}^L n_{jk}, \quad (5)$$

$$\pi_j^{(W)} = \frac{\# \text{ individuals working in } C_j}{N} = \frac{1}{N} \sum_{k=1}^L n_{kj}, \quad (6)$$

where the term residing means, as in the US Census, with usual residence. Likewise, we may define the joint probability mass function in terms of the contact matrix,

$$\pi_{jk}^{(HW)} = P\{H = j \& W = k\} = \frac{\# \text{ individuals residing in } C_j, \text{ working in } C_k}{N} = \frac{n_{jk}}{N}, \quad (7)$$

and denote the corresponding matrix by $\Pi^{(HW)}$. By swapping the order of W and H , we observe that $\pi_{jk}^{(WH)} = \pi_{kj}^{(HW)}$, thus defining the matrix $\Pi^{(WH)} = [\Pi^{(HW)}]^T$.

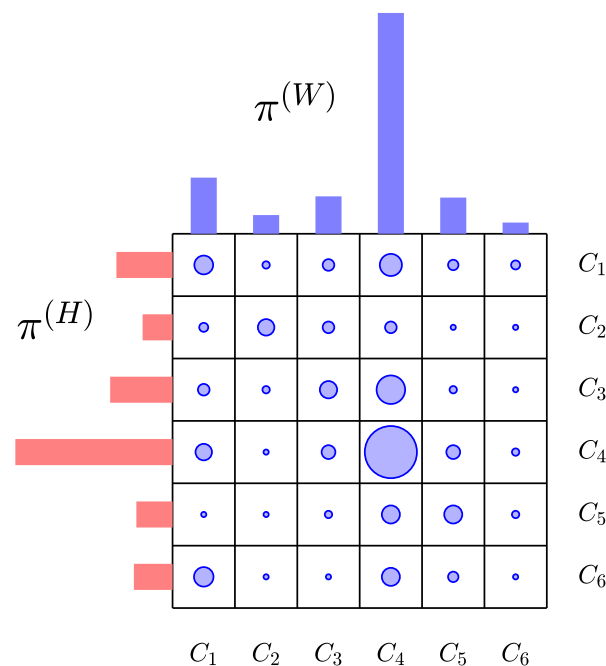


Figure 1. Schematic representation of the transport scheme across six communities. In this cartoon, the dots represent entries of the matrix of the transport scheme, with the size proportional to the radius of the dot. The red histogram $\pi^{(H)}$, which is the marginal distribution of the contact matrix with respect to the row index, represents the distribution of the population according to the census definition of usual residence, and the blue histogram $\pi^{(W)}$, the marginal with respect to the row index, is the population density during the work hours. Here, C_4 is an example of a center attracting population from other communities, while, for instance, C_6 is a bedroom community with low work place density, the principal work places being in C_1 and C_4 .

The conditional probability mass functions, expressing the probability that individuals reside in community j given that they work in community k , or, vice versa, that individuals work in community j given that they reside in community k , can be expressed in terms of the joint and marginal densities as

$$\pi_{jk}^{(H|W)} = P\{H = j \mid W = k\} = \frac{\pi_{jk}^{(HW)}}{\pi_k^{(W)}}, \quad \pi_{jk}^{(W|H)} = P\{W = j \mid H = k\} = \frac{\pi_{jk}^{(HW)}}{\pi_k^{(H)}}. \quad (8)$$

We denote the corresponding matrices by $\Pi^{(W|H)}, \Pi^{(H|W)} \in \mathbb{R}^{L \times L}$. The matrix $\Pi^{(HW)} \in \mathbb{R}^{L \times L}$ with entries $\pi_{jk}^{(HW)}$ defines a transport scheme between the densities $\pi^{(H)} \in \mathbb{R}^L$ and $\pi^{(W)} \in \mathbb{R}^L$. This naturally establishes a connection between our metapopulation network model and optimal transport theory ([20,21]).

2.2. Network SEIR Based on Transport Schemes

In this subsection we begin outlining the structure of the proposed metapopulation network model by first assuming that the epidemic spreads in each community according to a classical SEIR model. Subsequently the model will be modified to suit the dynamics of COVID-19. In a spatially homogenous SEIR model, we can describe the spread of the infection as a transmission flux from the susceptible to the exposed compartment,

$$S \xrightarrow{\varphi} E, \quad \varphi = \beta \rho \frac{I}{N} S, \quad (9)$$

where β is the infection rate that can be interpreted as the probability of contagion, ρ is the contact rate in the population, the frequency I/N can be interpreted as the probability

of the contact being infectious, and S is the size of the susceptible cohort. In the network setting of interest to us, we assume that the population residing in each community C_j is partitioned into the four SEIR cohorts (S_j, E_j, I_j, R_j) of susceptible, exposed, infected, and recovered individuals. Adhering to the probabilistic interpretation of the frequencies, we introduce

$$v_j^I = \frac{I_j}{N_j} = \text{probability that an individual in } C_j \text{ is infected,} \quad (10)$$

$$v_j^S = \frac{S_j}{N_j} = \text{probability that an individual in } C_j \text{ is susceptible,} \quad (11)$$

where N_j is the number of individuals in C_j , or, equivalently, the sum of the entries of the j th row of the contact matrix, $N_j = \sum_{k=1}^L n_{jk}$.

In a heterogeneous network the communities may reflect different social environments, population densities and dynamics determining the rate of daily contacts ρ_j . The contagion probability β , on the other hand, is assumed to be mostly pathogen-specific, and therefore identical in each community.

In our model, we take into account two transmission processes: Exposure to contagion at the workplace during working hours, and exposure in the home residence community when off work. Hence, the total transmission flux of susceptible individuals in the j th community S_j is the sum of the flux φ_j corresponding to at work transmission, and ψ_j corresponding to at home transmission, see Figure 2. In the discussion below, we tacitly assume that the joint probability densities do not change in time. In particular, while the death rate of COVID-19 is a significant factor from the human and public health point of view, we assume that its effect on the population densities can be neglected in the model.

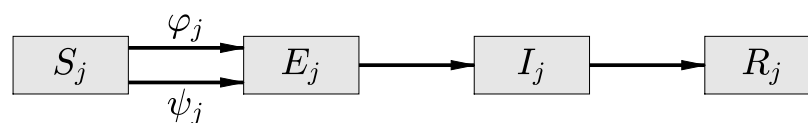


Figure 2. Susceptible resident of the community C_j in the cohort S_j may be exposed and infected either in the work environment (flux φ_j) or off work in the home community (flux ψ_j). The total transmission flux out of the susceptible population is a convex combination of these two fluxes, with coefficients that depend on the number of hours spent at work. In this diagram, the flux corresponding to a positive death rate is omitted.

To model the contagion of the residents of community C_j at the workplace, we further subdivide the susceptible cohort S_j into L sub-cohorts, denoted by $S_j^{(1)}, \dots, S_j^{(L)}$, according to the community where the workplace is located. The sizes of the sub-cohorts can be expressed in terms of conditional probabilities,

$$S_j^{(k)} = \pi_{kj}^{(W|H)} S_j, \quad 1 \leq k \leq L. \quad (12)$$

The individuals in the k th sub-cohort are exposed to infectious individuals working in the k th community. Since the number of contacts in C_k is ρ_k , the transmission flux in this sub-cohort is

$$S_j^{(k)} \xrightarrow{\varphi_j^{(k)}} I_j, \quad \varphi_j^{(k)} = \beta \rho_k S_j^{(k)} \times P\{\text{contact in } C_k \text{ is infectious}\}. \quad (13)$$

Recalling that v_ℓ^I is the probability that a contact originating from C_ℓ is infectious, and $\pi_{\ell k}^{(H|W)}$ the probability that the contact in C_k is a resident of the community C_ℓ , we can express for the probability of an infectious contact in the workplace community as

$$P\{\text{contact in } C_k \text{ is infectious}\} = \sum_{\ell=1}^L \pi_{\ell k}^{(H|W)} v_{\ell}^I, \quad (14)$$

so that

$$\varphi_j^{(k)} = \beta \rho_k S_j^{(k)} \sum_{\ell=1}^L \pi_{\ell k}^{(H|W)} v_{\ell}^I = \beta \left(\rho_k \pi_{kj}^{(W|H)} \sum_{\ell=1}^L \pi_{\ell k}^{(H|W)} v_{\ell}^I \right) S_j. \quad (15)$$

Since the total outflux from the susceptible cohort S_j of C_j is the sum of the contributions over all workplace communities, we can express it as

$$\varphi_j = \sum_{k=1}^L \varphi_j^{(k)} = \beta \sum_{k=1}^L \left(\rho_k \pi_{kj}^{(W|H)} \sum_{\ell=1}^L \pi_{\ell k}^{(H|W)} v_{\ell}^I \right) S_j, \quad (16)$$

or, in matrix form,

$$\varphi_j = \beta \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T v^I \right)_j S_j, \quad (17)$$

where $P = \text{diag}(\rho_1, \dots, \rho_L)$, the diagonal entries being the number of daily contacts in the respective community.

When modeling the contagion of susceptible individuals in their home community, we assume that they are exposed only to infectious individuals also residing in the community through, e.g., schools, local sports events, religious gatherings etc. The effect of exposure to infected individuals residing in other communities is accounted for in the term for contagion in the workplace, which includes the option of the workplace being in the home residence community. Therefore, in this case, the matrices of conditional densities $\Pi^{(H|W)}$ and $\Pi^{(W|H)}$ are replaced by identities, and the expression for the flux becomes

$$\psi_j = \beta \left(P v^I \right)_j S_j. \quad (18)$$

In the process of combining workplace and home contagion fluxes, we assume that individuals spend a fraction p of their time as commuters and the remaining fraction $1 - p$ in the community where they reside. Therefore, the SEIR model for the j th node of the network is described by a system of differential equations of the form

$$\begin{aligned} \frac{1}{N_j} \frac{dS_j}{dt} &= -p\varphi_j - (1-p)\psi_j \\ &= -\beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T v^I \right)_j + (1-p) \left(P v^I \right)_j \right\} v_j^S, \end{aligned} \quad (19)$$

$$\begin{aligned} \frac{1}{N_j} \frac{dE_j}{dt} &= p\varphi_j + (1-p)\psi_j - \eta v_j^E \\ &= \beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T v^I \right)_j + (1-p) \left(P v^I \right)_j \right\} v_j^S - \eta v_j^E, \end{aligned} \quad (20)$$

$$\frac{1}{N_j} \frac{dI_j}{dt} = \eta v_j^E - \gamma v_j^I - \mu v_j^I, \quad (21)$$

$$\frac{1}{N_j} \frac{dR_j}{dt} = \gamma v_j^I. \quad (22)$$

Introducing the diagonal matrix $N_0 \in \mathbb{R}^{L \times L}$ whose nonzero entries are the population sizes of the L communities in the network, $N_0 = \text{diag}(N_1, \dots, N_L)$, we can write the system of equations for all the communities in the network in the form

$$N_0^{-1} \frac{dS}{dt} = -\beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T v^I \right) + (1-p) (P v^I) \right\} \odot v^S, \quad (23)$$

$$N_0^{-1} \frac{dE}{dt} = \beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T v^I \right) + (1-p) (P v^I) \right\} \odot v^S - \eta v^E, \quad (24)$$

$$N_0^{-1} \frac{dI}{dt} = \eta v^E - \gamma v^I - \mu v^I, \quad (25)$$

$$N_0^{-1} \frac{dR}{dt} = \gamma v^I. \quad (26)$$

where “ $\odot$ ” denotes component-wise vector multiplication. If we assume that the population sizes N_j do not change significantly due to the mortality, we may scale the cohorts by N_0 ,

$$N_0^{-1} \frac{dX}{dt} = \frac{dv^X}{dt}, \quad X \in \{S, E, I, R\}, \quad (27)$$

and write the system in terms of the frequencies v^X . This approximation is used in the following discussion.

2.3. Network-Based SEAIR and SE(A)IR Models for COVID-19

A feature of COVID-19 that makes its spread extremely efficient and hard to track is the significant but unknown fraction of infected individuals who shed the virus and are either completely asymptomatic or develop only mild symptoms. While we expect individuals with clear symptoms to either isolate themselves or, in more serious cases, be hospitalized, thus sharply reducing the contact with susceptible individuals, asymptomatic and oligosymptomatic individuals may unknowingly shed the virus. Compartment models taking this cohort into consideration can be found in the literature: see, e.g., [22]. A fairly straightforward way to account for the infectiousness of the asymptomatic cohort is to disaggregate the cohort I into “infected, infectious, and asymptomatic” A , and “infected, infectious, and symptomatic” I , and split the flow from E to I into two separate contributions, as illustrated in the left panel of Figure 3. The resulting model, while rather straightforward to implement, requires information about the ratio of exposed individuals that remain asymptomatic and those who develop symptoms. To compensate for the unavailability of that type of information in the current understanding of COVID-19, in [18] we merged the exposed and asymptomatic cohorts into one compartment, denoted by $E(A)$, see the right panel of Figure 3. This approximation allows us to estimate dynamically the state and the parameters from the newly infected count data.

In the approximate SE(A)IR model, it is assumed that the asymptomatic individuals who are infective represent a portion $\xi < 1$ of the cohort $E(A)$, and that, due to self isolation and potential hospitalization, the contribution of symptomatic individuals to new infections is a fraction $q < 1$ of the contribution of the total cohort. Therefore, we modify the classical SEIR model (1) replacing the infection flux term with

$$\beta \frac{I}{N} S \rightarrow \beta \frac{\xi E(A) + qI}{N} S = (\beta \xi) \frac{E(A) + (q/\xi)I}{N}. \quad (28)$$

Denoting $f = q/\xi$, and letting β denote the product of the parameters β and ξ , the governing equations of the network model of connected heterogeneous communities become

$$\begin{aligned} \frac{dv^S}{dt} = & -\beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T (v^{E(A)} + f v^I) \right) \right. \\ & \left. + (1-p) (P (v^{E(A)} + f v^I)) \right\} \odot v^S, \end{aligned} \quad (29)$$

$$\frac{dv^{E(A)}}{dt} = \beta \left\{ p \left([\Pi^{(W|H)}]^T P [\Pi^{(H|W)}]^T (v^{E(A)} + f v^I) \right) + (1-p) \left(P(v^{E(A)} + f v^I) \right) \right\} \odot v^S - \eta v^E, \quad (30)$$

$$\frac{dv^I}{dt} = \eta v^{E(A)} - \gamma v^I - \mu v^I, \quad (31)$$

$$\frac{dv^R}{dt} = \gamma v^I. \quad (32)$$

We point out that while the approximate model does not explicitly include the incubation delay induced by the exposed compartment of the SEIR model, it retains some of its important properties, allowing the passage of asymptomatic cases directly into the recovered pool without developing symptoms, as schematically indicated in the right diagram in Figure 3. Moreover, this model allows us to estimate the size of the asymptomatic cohort, and produces robust and reasonable short term predictions based on county level data of new daily infected individuals. Most importantly, the model prediction of the size of the asymptomatic cohort ([18]) fully agrees with the CDC estimate that about one third of all infections are asymptomatic [23] during periods of relatively slow spread, while indicating higher ratio during aggressive spreading. We refer to the cited article for further discussion about this modeling choice. For further discussion of various modifications of the classical model with applications to COVID-19 modeling, we refer to [24].

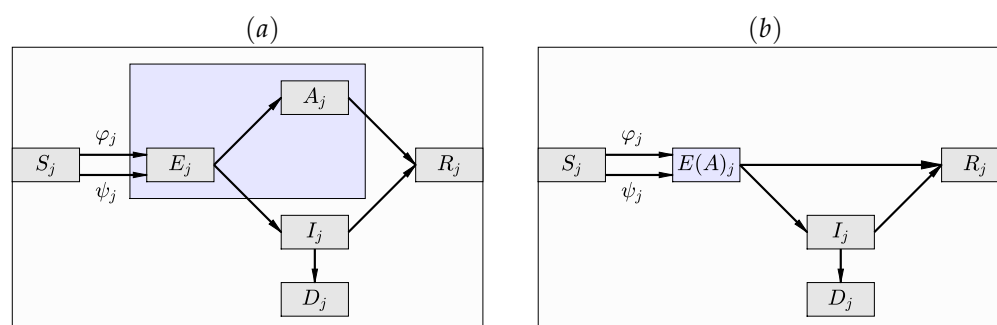


Figure 3. (a) The compartment diagram of the SEAIR model in which the infected and infectious compartment has been split into the asymptomatic (A_j) and symptomatic (I_j) compartments. (b) The SE(A)IR compartment diagram in which the exposed non-infective cohort E_j and the infective, infectious and asymptomatic compartments A_j are merged into a hidden compartment $E(A)_j$ of which only indirect inference is available, as long as the data consist of daily numbers of newly infected lab-confirmed cases.

3. Continuously Distributed Model

Following the construction of the discrete network model, it is possible to derive a continuously distributed model in terms of transport schemes. Let $\Omega \in \mathbb{R}^2$ denote a geographical area with N inhabitants, and define a joint probability density $\pi_{HW}(x, y)$ over Ω such that, for $A, B \subset \Omega$,

$$\int_A \left(\int_B \pi_{HW}(x, y) dy \right) dx = \text{probability that a person lives in } A \text{ and works in } B. \quad (33)$$

We will assume that the density π_{HW} is independent of time, thus neglecting the effect of the deaths on the densities. The marginal densities define the probability densities of individuals working or residing in a given location,

$$\pi_H(x) = \int_{\Omega} \pi_{HW}(x, y) dy, \quad \pi_W(y) = \int_{\Omega} \pi_{HW}(x, y) dx, \quad (34)$$

so that the number $N_H(A)$ of individuals living in $A \subset \Omega$, and the number $N_W(B)$ of individuals working in $B \subset \Omega$ are given by

$$N_H(A) = N \int_A \pi_H(x) dx, \quad N_W(B) = N \int_B \pi_W(y) dy, \quad (35)$$

respectively. The density π_{HW} thus defines a transport scheme between π_H and π_W . Similarly, we define the conditional densities of residing in x given that a known work location, and vice versa, the densities of the work space, given that the residing location is known,

$$\pi_{W|H}(y | x) = \frac{\pi_{HW}(x, y)}{\pi_H(x)}, \quad \pi_{H|W}(x | y) = \frac{\pi_{HW}(x, y)}{\pi_W(y)}. \quad (36)$$

We introduce the functions $v^X(t, x)$, $X \in \{S, E, I, R\}$, $0 \leq v^X \leq 1$, defining the frequencies of individuals in the four cohorts, and satisfying

$$v^S(t, x) + v^E(t, x) + v^I(t, x) + v^R(t, x) \leq 1, \quad (37)$$

where the discrepancy from one of the sum comes from of non-vanishing mortality rate that may decrease the total population size. The actual size of each cohort X in $A \subset \Omega$ is

$$X_A(t) = N \int_A v^X(t, x) \pi_H(x) dx, \quad X \in \{S, E, I, R\}. \quad (38)$$

Finally, let $\rho(x)$ denote the contact rate at a given place $x \in \Omega$. With these notations, we can now set up the infection model following the lines of the network model. As before, we start with the SEIR model. The conditional density of susceptible individuals residing at x , working at y is $\pi_{H|W}(x | y) v^S(x)$, while the conditional density of infected individuals originating from z working at y is $\pi_{W|H}(y | z) v^I(z)$, so the work-related infection flux for individuals at x is

$$\varphi(t, x) = \beta \int_{\Omega} \rho(y) \left(\int_{\Omega} \pi_{W|H}(y | z) v^I(t, z) dz \right) \pi_{H|W}(x | y) v^S(t, x) dy. \quad (39)$$

To account for infections at home, we replace the transport scheme with a diagonal one,

$$\pi_{HH}(x, y) = \pi_H(x) \delta(x - y), \quad (40)$$

for which the conditional densities are $\delta(x - y)$, yielding the flux

$$\psi(t, x) = \beta \rho(x) v^I(t, x) v^S(t, x). \quad (41)$$

The extension of both of the fluxes above are modified to the SE(A)IR model, leading to the following system of integro-differential equations,

$$\begin{aligned} \frac{\partial v^S}{\partial t}(t, x) = & -\beta \left\{ p \int_{\Omega} \rho(y) \left(\int_{\Omega} \pi_{W|H}(y | z) (v^{E(A)}(t, z) + f v^I(t, z)) dz \right) \pi_{H|W}(x | y) dy \right. \\ & \left. + (1 - p) \rho(x) (v^{E(A)}(t, x) + f v^I(t, x)) \right\} \odot v^S(t, x), \end{aligned} \quad (42)$$

$$\begin{aligned} \frac{\partial v^{E(A)}}{\partial t}(t, x) = & \beta \left\{ p \int_{\Omega} \rho(y) \left(\int_{\Omega} \pi_{W|H}(y | z) (v^{E(A)}(t, z) + f v^I(t, z)) dz \right) \pi_{H|W}(x | y) dy \right. \\ & \left. + (1 - p) \rho(x) (v^{E(A)}(t, x) + f v^I(t, x)) \right\} \odot v^S(t, x) - \eta v^{E(A)}(t, x) - \gamma v^{E(A)}(t, x), \end{aligned} \quad (43)$$

$$\frac{\partial v^I}{\partial t}(t, x) = \eta v^{E(A)}(t, x) - \gamma v^I(t, x) - \mu v^I(t, x), \quad (44)$$

$$\frac{\partial v^R}{\partial t}(t, x) = \gamma v^{E(A)}(t, x) + \gamma v^I(t, x). \quad (45)$$

While the continuous model may seem less intuitive than the community-based discrete model, it can help elucidate some aspects of the spreading dynamics in a non-homogeneously distributed population, as the numerical examples show.

4. Results

In this section, we test and analyze two network models based on county level census data. The network in the first example comprises seven counties in the greater Cleveland area in Northeast Ohio. The larger network of the second example comprises 74 counties in the Pittsburgh–Cleveland–Michigan corridor through the four states of Pennsylvania, West Virginia, Ohio and Michigan. Of particular interest in both examples is the role and importance of highways in the spread of the virus.

4.1. Construction of the Transport Scheme

The key constituent of the metapopulation model is the joint probability density matrix Π^{HW} . The construction is based on Formula (7), using the publicly available data provided by the Census Bureau's American Community Survey (ACS) of 2015. The ACS provides a snapshot of regional commuting patterns by quantifying the number of people who report commuting from one county to another for work. The values of the entries n_{jk} quantifying the commuting traffic from j th to k th community in the network were assigned according to this data set for each pair of communities included in the network. A modeling decision is required to truncate the network to exclude commuter traffic to counties outside the network. In our model, we set no traffic to workplace locations outside the network, which amounts to treating commuters to outside communities as workers at the residence community. We remark that the effects of this truncation could be significant if a major nearby workplace center is not part of the network. The population of each community is set according to the Census Bureau's 2019 estimates. The census data, including the commuter traffic data, correspond to the pre-pandemic situation, and it is fair to ask to what extent the numbers should be adjusted to correspond to the decrease due to COVID-19. While precise information of the reduction in the traffic is hard to obtain, some partial estimates are available, with a 15% overall decrease in traffic after reopening noted by the Ohio Department of Transportation [25], with similar decreases noted for Pennsylvania and Michigan [26,27]. While significant, this study's intent is to initialize with the spatiotemporal dynamics brought on by the commuter matrices prior to an imposed lockdown. Finally, we point out that the transport scheme contains no additional model parameters.

4.2. Network Correction: Greater Cleveland Region

The Greater Cleveland network considered in the first example comprises counties in the Cleveland–Elyria Metropolitan Statistical Area (Cuyahoga County containing Cleveland, and Geauga County, Lake County, Lorain County, and Medina County) plus the two bordering counties of the Akron Metropolitan Statistical Area (Summit County containing Akron, and Portage County). The total population in the network is 2.76 million, with 1.25 million in Cuyahoga county alone, which constitutes the central node and the principal generator of commuter traffic in the area. Indeed, according to the census data, about 81,000 of residents of Cuyahoga county, or 6.5% of the population, commute out of the county, compared to an influx of about 211,000 commuters, equivalent to almost 17% of the county's population. The second largest node, Summit County, has a population of 540,000 residents, 13% of which work in a different county, while the influx of workers

from other counties corresponds to about 12% of the residents. The information about commuting in and out of each of the seven counties is summarized in Figure 4, which shows also the graphs of the conditional densities $\pi_{jk}^{(W|H)}$ county by county. Bordering the cluster of seven counties there is Lake Erie to the north, and mostly rural counties to the west, neither of which are likely to cause any boundary effects. On the other hand, Stark county with the city of Canton, right outside the southern border, may be causing some boundary effects. Similar considerations hold for the east and southeast border, where the neighboring Trumbull and Mahoning counties, home of some large industrial plants, may generate commuting patterns not included in the network.

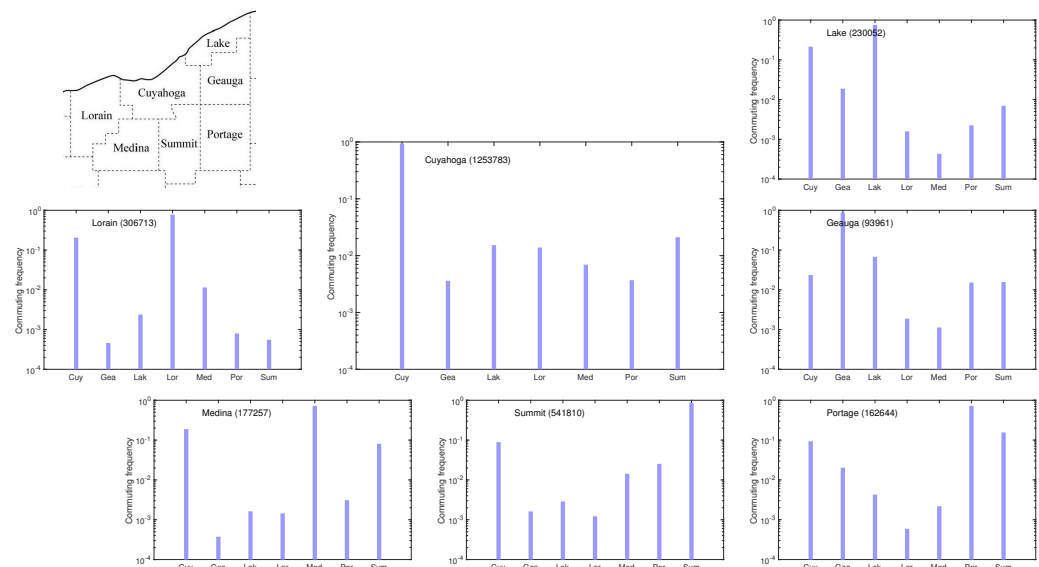


Figure 4. The conditional densities $\pi^{(W|H)}$ county by county in the northeast Ohio network comprising the seven counties shown in the map. The plots are on a logarithmic scale, and the total population of each county is indicated in parenthesis.

In the first numerical simulation, we test whether accounting for the network structure may result in some discrepancy with the Bayesian estimate of the parameters in the SE(A)IR model for the individual counties computed according to [18]. For completeness, below we include a brief description of the Bayesian parameter estimation process.

4.2.1. Bayesian Parameter Estimation

In [18], a suitably designed Bayesian particle filtering (PF) algorithm was used to dynamically estimate the model parameters for individual counties based on data consisting of the daily count of new infections, without taking the population mobility into account. In the algorithm, the infection rate including the contact rate, denoted by $B = \beta \times \rho$, is assumed to be time-dependent, while the parameters η , γ and μ are assumed to be constants. The static parameters are assumed to be either known, or they become part of the estimation problem as well. In the PF algorithm, at each time instance $t = 1, 2, \dots$ measured in units of a day, each particle is a random realization of the vector of the current state, i.e., $X^j(t) = (S^j(t), E^j(t), I^j(t), R^j(t), B^j(t))$, where $1 \leq j \leq n$, n is the number of particles. At time step t , each particle $X^j(t)$ is propagated using the dynamical model to yield a prediction $\hat{X}^j(t+1)$ (prediction step) at time $t+1$, and the predictions are scored according to how well they explain the observed data at time $t+1$ according to a probabilistic likelihood model, in which the daily count of new infections is modeled as a Poisson distributed random variable with a mean value corresponding to the model prediction. The particles are then resampled using the scores as weights (correction step), and the resampled particles are randomly perturbed by an innovation process to enrich the particle cloud. The random innovation can be thought of as a way to account for various

uncertainties, including the mismatch between the dynamic model and reality. After the randomization, the counter t is advanced by one, and the process is repeated. At each time instance, the particle cloud is used to compute an estimate for the state, the parameters and posterior uncertainty. At the beginning of the process, we assume a few asymptomatic or infected cases in each county. Typical outputs comprising the posterior mean and posterior standard deviation envelopes of the infection rates,

$$\bar{B}(t) = \frac{1}{n} \sum_{j=1}^n B^j(t), \quad \sigma(t) = \left(\frac{1}{n-1} \sum_{j=1}^n (B^j(t) - \bar{B}(t))^2 \right)^{1/2}, \quad t = 1, 2, \dots, \quad (46)$$

as shown in Figure 5.

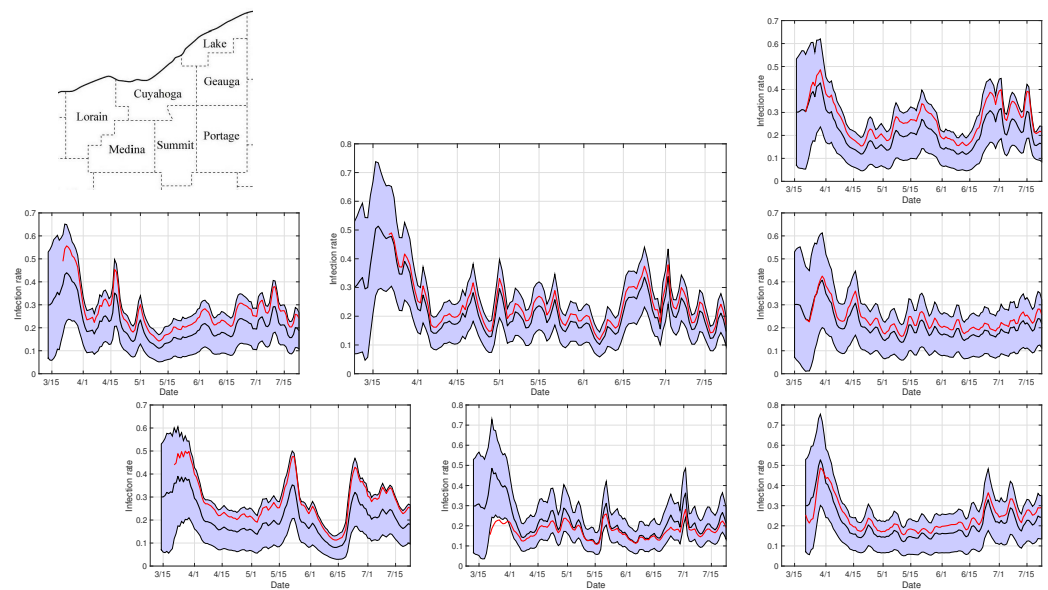


Figure 5. Infection rates $B^{(\ell)}(t)$ in the seven-county network during the first 100 days after the first recorded case in Cuyahoga county. The panels show the ensemble average (black curve) and the envelopes of two standard deviations of the independent county-by-county estimates computed by the PF algorithm, and the weighted least squares estimates (red curve) corresponding to the value of $\delta = 0.02223$ that minimizes the residual over the diagonal entries. Observe that the estimation process of $X(t)$ (red curve) starts as soon as each of the seven counties has had at least one COVID-19 case (Portage county was the last one) to guarantee that the vector of individual county estimates for the infection rate is defined.

4.2.2. Network Correction to Parameter Estimates

In [18], estimates for the infection rate $B(t)$ were computed independently for each of the seven counties in the network, without taking into account the underlying connections. For each county the PF algorithm provides a time-evolving ensemble of particles from which we can estimate their posterior mean and posterior standard deviation (46). In the following, we denote by $(\bar{B}^{(\ell)}(t), \sigma^{(\ell)}(t))$ the estimated posterior mean and posterior standard deviation of the infection rate in the county ℓ , $1 \leq \ell \leq 7$, at the time t , the unit of time being days, with $t = 1$ corresponding to 7 March 2020, when the first COVID-19 case was recorded in Cuyahoga County. Since the infection history started at different times in different counties, we restrict t to $t \geq t_0 > 1$, where t_0 corresponds to the date when all seven counties had at least one reported case, allowing an estimate of $B^{(\ell)}(t)$ in each county.

The question that we want to address in this section is the effect, if any, of accounting for the commuter traffic on the estimated infection rates. To answer this question, given the independent estimates $\bar{B}^{(\ell)}(t)$ for each county in the network, $1 \leq \ell \leq 7$, we seek

to compute a correction to the estimates based on the network model. More precisely, the question can be formulated as follows: If the commuter traffic is taken into account, which infection rates in each county would explain the data with approximately the same accuracy as the infection rates that were estimated without taking the mobility between counties into account? To formalize the question further, let

$$\bar{B}(t) = \begin{bmatrix} \bar{B}^{(1)}(t) \\ \vdots \\ \bar{B}^{(7)}(t) \end{bmatrix} \in \mathbb{R}^7 \quad (47)$$

be the vector of infection rates in the seven counties estimated without accounting for the network structure, and denote by $X(t) \in \mathbb{R}^7$ the unknown effective infection rates $X^{(\ell)}(t)$, $1 \leq \ell \leq 7$ that take the network structure into account. To build a model between $\bar{B}(t)$ and $X(t)$, we posit that if the sizes of the susceptible, asymptomatic, and infectious cohorts in each county are unaltered, the two models should produce, within the margin of uncertainty, the same model predictions. The two quantities that we compare here are the flux vectors from the susceptible cohorts to the exposed/asymptomatic cohorts with and without accounting for the network structure, that is,

$$\begin{aligned} \varphi_{\text{network}} &= -\beta(t) \left\{ p \left([\Pi^{(W|H)}]^T P(t) [\Pi^{(H|W)}]^T (v^{E(A)} + f v^I) \right) \right. \\ &\quad \left. + (1-p) \left(P(t) (v^{E(A)} + f v^I) \right) \right\} \odot v^S \end{aligned} \quad (48)$$

$$= - \left\{ \left[p [\Pi^{(W|H)}]^T \text{diag}(X(t)) [\Pi^{(H|W)}]^T + (1-p) \text{diag}(X(t)) \right] (v^{E(A)} + f v^I) \right\} \odot v^S, \quad (49)$$

$$\varphi_{\text{no network}} = - \left\{ \text{diag}(\bar{B}(t)) (v^{E(A)} + f v^I) \right\} \odot v^S. \quad (50)$$

Above, we merged the parameter $\beta(t)$ and the diagonal matrix $P(t)$ into a diagonal matrix with unknown diagonal entries $X^{(\ell)}(t)$. Therefore, for the two models to yield comparable predictions, up to a level of uncertainty, the matrices defining the transmission rates should match, that is

$$p [\Pi^{(W|H)}]^T \text{diag}(X(t)) [\Pi^{(H|W)}]^T + (1-p) \text{diag}(X(t)) = \text{diag}(\bar{B}(t)) + E(t), \quad (51)$$

where the noise matrix $E(t)$ accounts for the uncertainty in the PF estimate and possibly inaccuracies in the model. Observe that the above matrix model is linear in $X(t)$, and by taking into account the symmetry of the diagonal matrix constituting the data, we may reformulate the problem in terms of $7 \times (7+1)/2 = 28$ scalar equations, one for each entry of the upper triangular portion of the 7×7 matrices. Hence, the matrix Equation (51) can be reorganized as

$$AX(t) = \underbrace{\begin{bmatrix} \bar{B}(t) \\ 0_{21 \times 1} \end{bmatrix}}_{\hat{B}(t)} + \varepsilon(t), \quad \text{subject to } X^{(\ell)}(t) > 0, \quad (52)$$

where $A \in \mathbb{R}^{28 \times 7}$, $\varepsilon(t) \in \mathbb{R}^{27}$ is a noise vector containing the components of the upper triangular part of the matrix $E(t)$, and the entries of $X(t)$ must be nonnegative. Therefore, to estimate $X(t)$ we must solve the weighted non-negative least squares problem (52) for

each t . Assume that the noise $\varepsilon(t)$ has zero mean, and is Gaussian with uncorrelated components, i.e., $\varepsilon(t) \sim \mathcal{N}(0, \Sigma(t))$, and $\Sigma(t)$ is diagonal with entries

$$\Sigma_{\ell\ell}(t) = \begin{cases} (\sigma^{(\ell)}(t))^2 + \delta^2, & 1 \leq \ell \leq 7, \\ \delta^2, & 7 < \ell \leq 28, \end{cases} \quad (53)$$

where $(\sigma^{(\ell)}(t))^2$ is the posterior variance of the PF estimate $\bar{B}^{(\ell)}(t)$, and δ^2 represents the variance of the modeling error. The solution of the weighted least squares problem with non-negativity constraint,

$$\text{minimize } \left\| \Sigma(t)^{-1/2} (\hat{B}(t) - AX(t)) \right\|^2 \quad \text{subject to } X^{(\ell)}(t) > 0, \quad (54)$$

is also the maximum a posteriori estimate of the problem (52). The result is sensitive to the choice of the ill-defined parameter δ : if δ is very small, the off-diagonal entries in (51) are given a large weight, biasing the components of $X(t)$ toward zero, while a δ significantly larger than $\sigma^{(\ell)}(t)$ will lead to solutions that poorly reproduce the data. To find a reasonable value for δ , denote by $X_\delta(t)$ the least squares solution of (54) corresponding to a given value δ , let $T \in \mathbb{R}^{7 \times 28}$ denote the orthogonal projection onto the first seven coordinates, and choose the value of δ that minimizes

$$r(\delta)^2 = \sum_{t=t_0}^N \|B(t) - TAX_\delta(t)\|^2, \quad \delta > 0. \quad (55)$$

Figure 6 shows the graph of the function $\delta \mapsto r(\delta)$ for δ in the interval $[10^{-4}, 10^{-1}]$. The results of the estimation process corresponding to the choice of δ advocated above are shown in Figure 5. Note that the estimation of the $X(t)$ begins at $t = t_0$ when each of the seven counties had at least one COVID-19 case. The last county to report a positive case was Portage County.

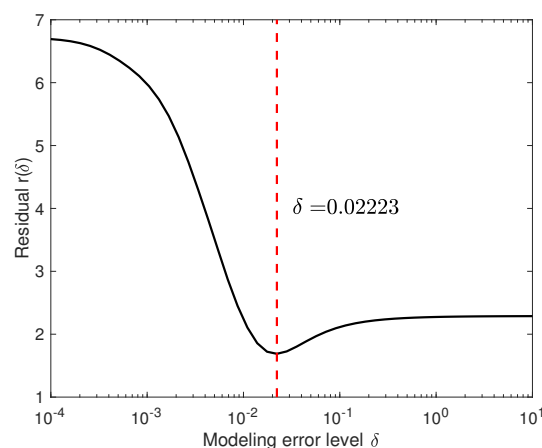


Figure 6. The residual $r(\delta)$ defined by the Formula (55), plotted over the interval $[10^{-4}, 10]$. The minimizer is indicated by the dashed line.

The results show that accounting for the network structure has an effect on the estimated infection rate, although the new estimate is usually within the envelope of two standard deviations around the posterior mean single county estimate. The discrepancy between individual and network-based estimates carries some information about the infection dynamics. For example, the infection rate in Summit county at the first peak of the reported positive cases estimated from the county data only is significantly higher than the network-based estimate. Since in the network model the new infections are explained as a convex combination of local and global infections, a significantly lower local infection rate is sufficient to explain the new cases when taking into account the commuting traffic,

suggesting that the first peak in Summit county may have been the result of workplace contacts. Conversely, in Lorain and Medina counties, the network-based estimates of infection rates are higher than the individual county estimates, suggesting that a higher local infection rate is needed to explain the infections, since the time commuters spent in the workplace was not enough to explain the new daily infections. Therefore in those two communities, the spread may have been due predominantly to local activities, and less to commuters. Confirming these hypotheses would require more information about the local infection dynamics. Finally, we caution that the results reported in this section are preliminary, and a full understanding of the effect of the network structure on the estimates requires the extension of the particle filtering algorithm to the network model, a task that is underway but beyond the scope of this article.

4.3. Pittsburgh–Cleveland–Detroit Corridor

The larger network considered in this example encompasses 74 counties from four different states (Pennsylvania, Ohio, West Virginia and Michigan) covering the corridor connecting three metropolitan areas around Pittsburgh, Cleveland and Detroit. To reduce boundary effects, the network was set up so as to not share any border with outside counties with major metropolitan areas. Instead, the corridor is buffered by mostly rural counties, unlikely to generate much commuting. The reference parameters for these simulations are listed in Table 1.

Table 1. Parameter settings in the network simulation.

Parameter	Symbol	Value
infection rate	β	0.01
daily contacts	r	20, 40, 60
incubation rate	η	1/7
recovery rate	γ	1/14
death rate	μ	0.0035

In our simulations, we classify the 74 counties in the network as having high ($\geq 800,000$), medium (between 800,000 and 150,000), or low ($< 150,000$) populations, using the population size as a proxy for density, and we set the contact rate ρ to 60, 40, and 20 for the high, medium, and low population counties, respectively, to reflect the effect of urbanization on contact rates. The top left panel in Figure 7 displays the contact matrix between pairs of counties in the network under normal pre-pandemic conditions.

The role of the initial condition: Our first simulation starts with a single infected individual in the low population density Huron County, MI (population: 31,166). Figure 8 shows the results of this simulation with our metapopulation models for the 74 counties in the corridor over a period of 21 weeks. In the first few weeks the infection spreads slowly to other counties, while never becoming highly prevalent in the county where it initiated, as shown in the top row. However, once the infection reaches the three-county core of Macomb, Oakland, and Wayne constituting the Detroit metropolitan region, the high population density allows the virus to quickly spread within the region. The three major population centers in the corridor, comprising Macomb/Oakland/Wayne, Cuyahoga (Cleveland), and Allegheny (Pittsburgh) counties, are tightly connected via the contact matrix. Therefore, once the virus takes hold in any one of the three metropolitan areas, it quickly reaches the other two centers, where it spreads rapidly due to the high population density and large number of susceptible individuals. After having run its course in the high population counties, the infection then grows dramatically in the surrounding counties, facilitated by the presence of interstates and major highways, eventually reaching also rural and more remote counties. As the infection hotspots move to rural counties, regions with high population density begin to recover.

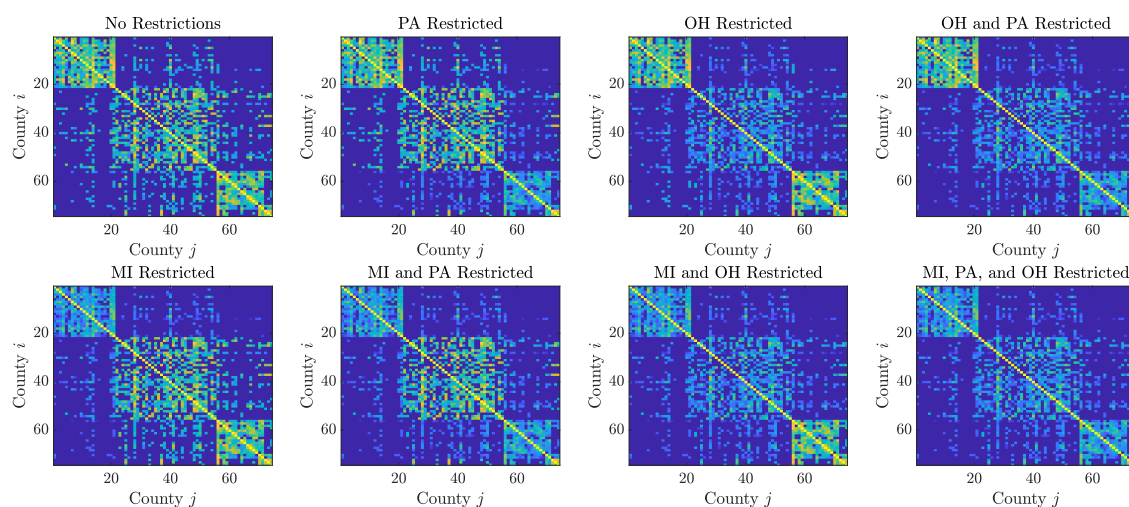


Figure 7. Contact matrices for the 74 counties in the Cleveland–Detroit–Pittsburgh corridor, plotted in a logarithmic scale. Each county corresponds to a row /column of the matrix, with the counties grouped by state, in the order: Michigan, Ohio, Pennsylvania and West Virginia, the latter being represented by only a few counties in the northern panhandle. The intensity of (i, j) th pixel is proportional to the number of individuals commuting from the j th country to the i th county. The top left panel shows the commuting patterns prior to the pandemic: the counties in the states of Michigan and Pennsylvania are more tightly internally connected than those in Ohio, as indicated by the intensity of the pixels in the first and third diagonal blocks, and it is clear that there is substantial commute among the three larger metropolitan areas. The remaining panels show how the communication network changes when the mobility to/from counties in specific states is reduced by 95%.

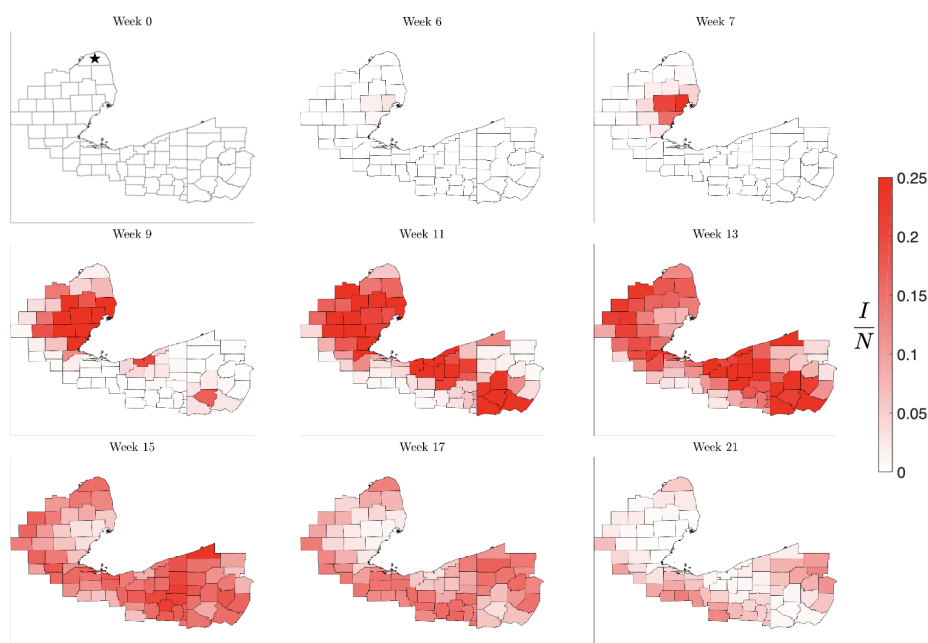


Figure 8. Simulation of the spread of the infection starting in the low-density Huron County, MI, marked by a star in Week 0. The color code indicates the number of infected individuals per 100,000 inhabitants. Observe that the infection level remains low until it reaches the nearest high density metropolitan area, but afterwards propagating quickly to the other centers. The pattern of the evolution of the infection shows that metropolitan areas work as amplifiers, rendering the second wave in the surrounding areas much stronger than the initial infection. The rural areas experience the second wave when the metropolitan areas are already recovering, in agreement with what was observed in the United States in the late summer of 2020.

A similar geographic pattern of spread was predicted by our model when starting with a single infected individual in a medium population density county. Figure 9 shows the model predictions for 21 weeks starting with a first infected individual in the medium density Stark County, OH (population: 370,606). In the first few weeks the infection spreads to nearby counties connected via interstate highways, with Cuyahoga County being the nearest high population density county. Once the infection has reached the densely populated Cleveland area, it spreads to the other two major metropolitan centers, Pittsburgh and Detroit, following a pattern similar to that observed when the infection started in a low population county: at first there is a substantial spike in the large population centers then the infection spreads to the surrounding regions following the interstate systems.

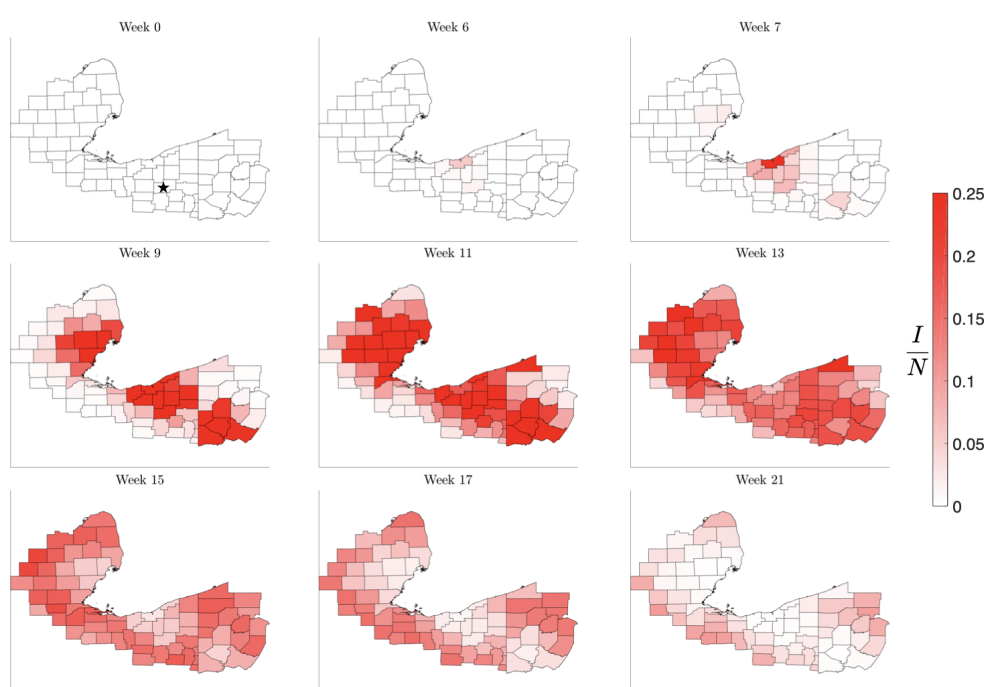


Figure 9. Simulation starting spread beginning in the medium population Stark County, OH, marked by a star in Week 0. As in the previous simulation, the initial infection remains almost unnoticed until it reaches the nearest metropolitan area, from which it quickly spreads to the other high density communities. Again, the metropolitan centers act like amplifiers, rendering the second wave in the initial point more intense than the initial infection. The timeline is similar to the previous simulation.

When the simulation is started with a single infected individual in the high population Allegheny County, PA (population: 1,216,045), the prevalence rapidly increases in the Pittsburgh metropolitan area, as shown in the second panel of the top row of Figure 10, spreading to the other two major metropolitan areas more quickly than in the two previous scenarios. In all three protocols, viral spikes occur in the large population centers before spreading to the surrounding metro area and onward to farther geographic distances, thus suggesting that densely populated regions effectively act like amplifiers for the spread of the infection. The high population density of the region portends that the number of infected will grow very rapidly once an infection has taken hold. As the number of infected quickly rise in the densely populated counties, the consequence of high commuter in- and outflows mean that susceptible commuters will more likely interact with infectious individuals and the infection will likely spread throughout the metropolitan region.

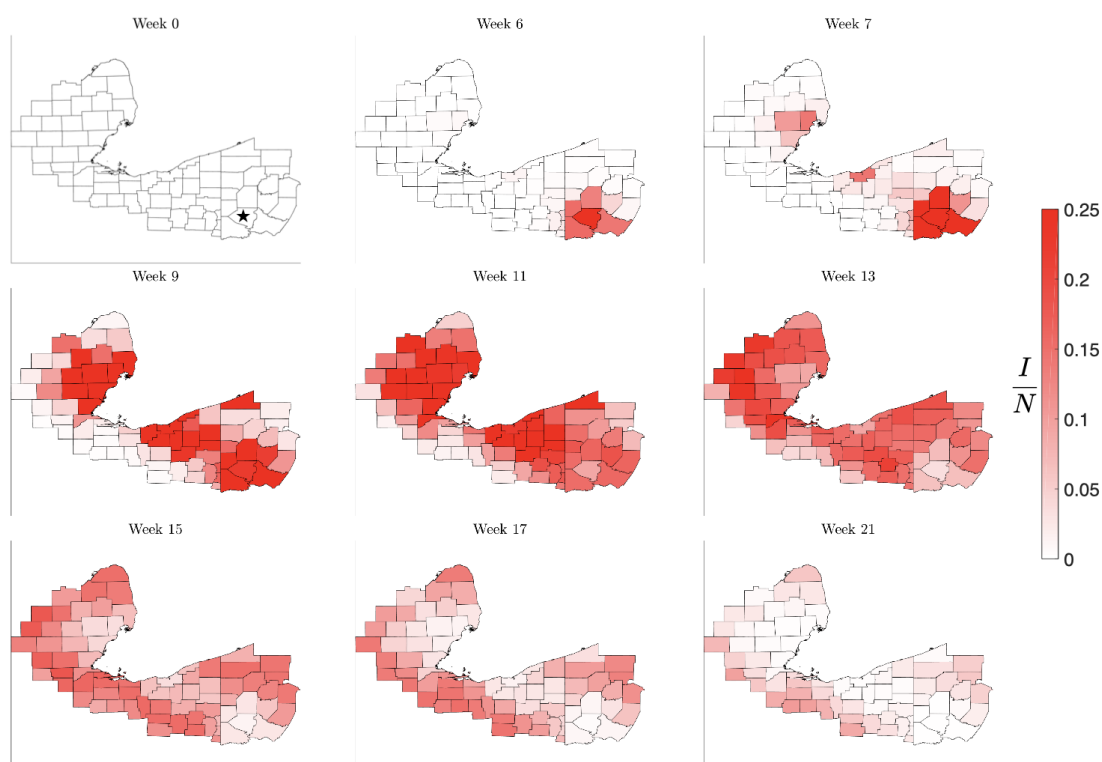


Figure 10. Simulation of the spatial pattern of the spread when starting with a single infected individual in the large population Allegheny County, PA (Pittsburgh area). The infection quickly spreads to the neighboring communities, which concurrently reaches the other two metropolitan areas. While the geographic pattern is similar to the two previous simulations, the timeline for the spread of the infection is faster.

Mitigation by restricting travel to/from counties in a state with high prevalence:

One of the state-level strategies considered to slow down the spread of COVID-19 was to reduce traffic between neighboring states and counties. To see how, according to our model, such measures affect the spread of the infection, we dynamically modified the contact matrix N to reflect the reduction in traffic in and out of the counties in states experiencing spikes in infections. More specifically, when the prevalence of the infected population in the counties of interest in Michigan, Ohio, or Pennsylvania reaches 10% of the total population, traffic into and out of every county in the state is decreased by 95% of its original level, simulating a ‘lockdown’ scenario. Figure 7 displays how the contact matrix changes in response to state shutdowns. Figure 11 shows the model predictions for an initial infection in Stark County, OH, with the described mitigation measures. In the first few weeks, while the prevalence is still low, the pattern is the same as in the unmitigated case, but by week 6 the spread in Ohio slows down after the high prevalence in Cuyahoga County has triggered the reduced traffic scenario statewide. Generally, if the mitigation measures are at the state level, the infection must reach the more densely populated counties of the three states before triggering traffic reductions which slow the transmission of the disease, though it does little to reduce the prevalence of the infections in the individual counties once the disease has taken hold.

Local mitigation measures in counties with high prevalence: In this simulation, we consider a mitigation strategy targeting selected local hotspots, in which strict social distancing measures are introduced in counties with prevalence greater than or equal to 15%. To understand how the spread pattern is affected by the triggered local mitigation measures, we run a simulation starting with a single infected individual in Stark County, dynamically reducing the internal contact frequency by 50% in a county as soon as the infected population is more than 15% of its population. The results, displayed in Figure 12,

show that the infection initially spreads in a similar manner to what was observed in the unrestricted traffic mitigation. However, because the increased prevalence triggers the local mitigation measures, the peaks of infections are not as high as in the unmitigated case, as fewer people are infected at the same time. This in turn leads to a longer period of elevated incidence, as opposed to a quick spike in the number of infected. Overall this results in more counties having an elevated number of infected people, but without the sharp spikes in individual counties observed in the unmitigated cases.

Combined triggered local mitigation and travel restriction: The effects of local mitigation by triggered reduced mobility and contact frequency can be seen in Figure 13, illustrating how these components together change the temporal dynamics of the virus over the 74 counties in the corridor. Noticeable are the lower peaks and longer period of elevated incidence due to the reduced contact frequency, as well as the slower spread of the virus over the 74 county regions due to reduced traffic to and from high prevalence counties. This highlights how the combination of the two mitigation strategies has the dual effect of softening the increase in new infections in each individual county by reducing local contacts, while simultaneously slowing down the spread across the region by reducing commuting traffic.

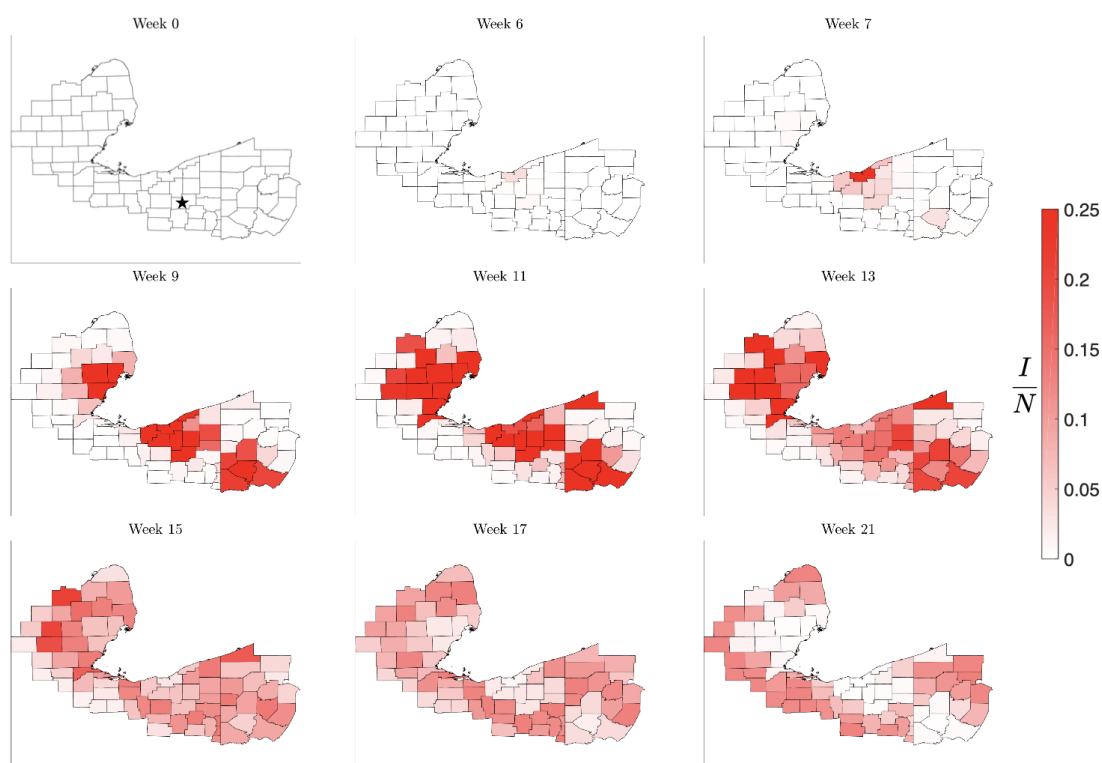


Figure 11. Time course of the spread of the infection under mitigation by drastically reducing traffic to and from the counties in a state with 10% prevalence. This simulation was started with a single infected individual in Stark County, OH (marked with a star in the first panel).

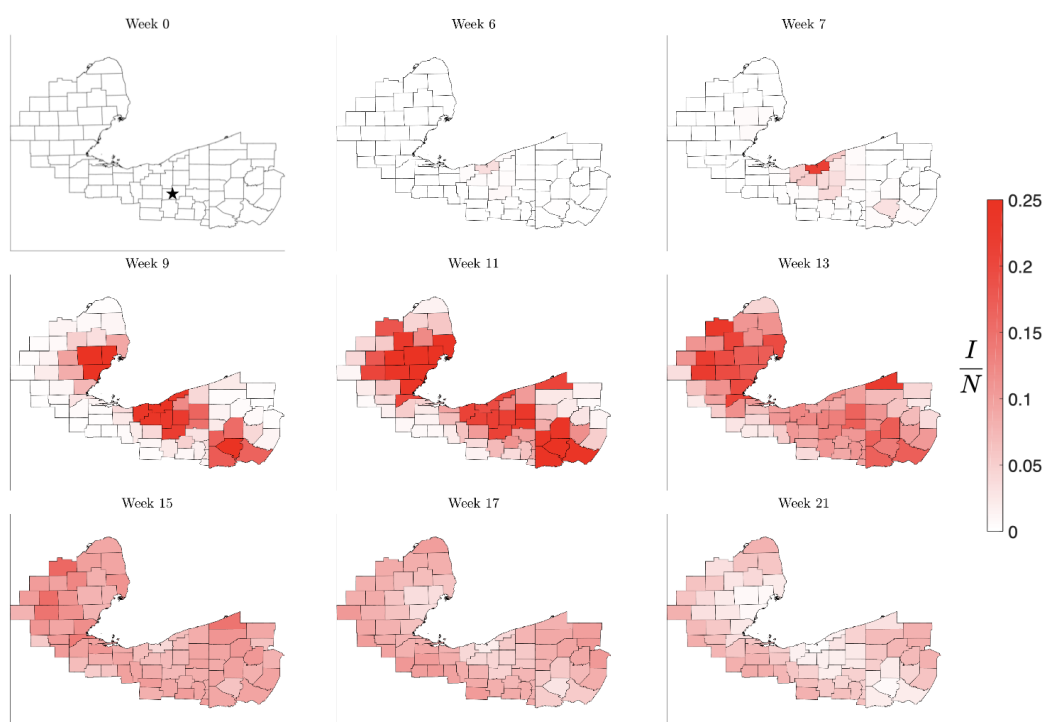


Figure 12. Time course of the spread of the infection under mitigation by triggered contact reduction in counties as soon as the prevalence reaches 15% of the county population. This simulation was started with a single infected individual in Stark County, OH (marked with a star in the first panel).

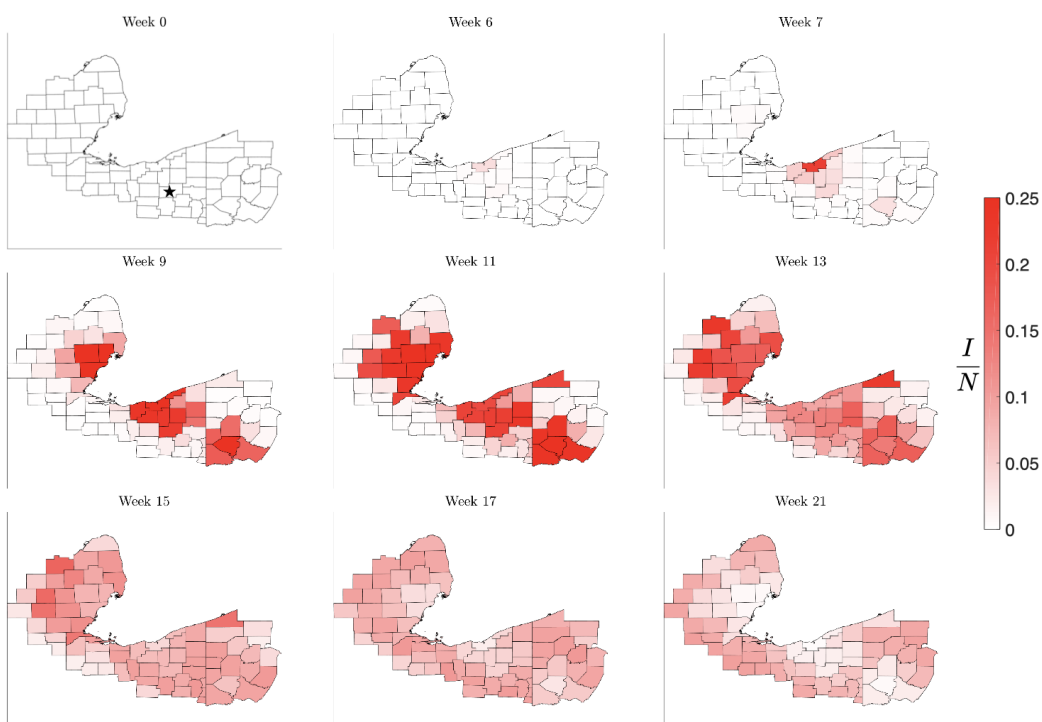


Figure 13. Time course of the spread of the infection under combined mitigation by triggered contact reduction in individual counties as soon as the prevalence reaches 15% of the county population accompanied by reduction in traffic to and from the affected states. This simulation was started with a single infected individual in Stark County, OH (marked with a star in the first panel).

4.4. Simulation with Continuous, Spatially Distributed Model

The network model proposed above can be extended naturally to a continuous spatially distributed setting. To do so, we need to define a continuous transport scheme π_{HW} . Let $\pi_H(x)$ denote the presumably known population density over the area Ω that we may base, e.g., on the census data. We begin by building a kernel K as the sum of a term accounting for local commuter traffic in the proximity of the residence, and a non-local part, accounting for distant commuting to larger, more densely populated centers,

$$K(x, y) = K_{\text{loc}}(x, y) + K_{\text{dist}}(x, y). \quad (56)$$

Denoting by $\lambda > 0$ the distance parameter defining the radius of the the local activities, we write

$$K_{\text{loc}}(x, y) \propto e^{-|x-y|/\lambda}, \quad (57)$$

where “ $\propto$ ” stands for proportionality. The non-local term account for the fact that densely populated centers provide job opportunities for the population residing in less dense nearby locations. The kernel model assumes that workplaces close to home are more attractive than distant ones, and that densely populated centers are surrounded by a host of commuter residential communities. To quantify these assumptions, we introduce a communication weight function ψ ,

$$\psi(s) = \kappa \frac{R^2}{R^2 + s^2}, \quad (58)$$

where $\kappa > 0$ is a coupling constant, and $R > 0$ is an effective radius of influence and define

$$K_{\text{dist}}(x, y) = \frac{\pi_H(y)}{\pi_{\text{max}}} \psi(|x - y|), \quad (59)$$

where $\pi_{\text{max}} = \max \pi_H(y)$. Finally, we write

$$\pi_{HW}(x, y) = \frac{\pi_H(x)}{\int_{\Omega} K(x, y) dy} K(x, y), \quad (60)$$

with the normalization chosen to satisfy the marginal condition

$$\int_{\Omega} \pi_{HW}(x, y) dy = \pi_H(x). \quad (61)$$

Observe that the communication weight function above is chosen as in the Cucker–Smale flocking model [28,29], the parameter R indicating the distance at which the weight function is one half of value at zero distance. The coupling constant κ determines the proportion of the community consisting of the commuter work force.

Figure 14 shows a fictitious density map, and five snapshots of the frequency distribution $\nu^I(x, t)$ of the infected population. The infection starts with a single exposed individual at time $t = 0$ at the location indicated by a red dot. At time $t = 15$, there is a small increase in the number of infections in the surrounding area, but most significantly, the infection has reached the most densely populated areas, with a maximum in the nearby center of the initial point. At $t = 28$, the infection is peaking in the urbanized high density areas, while the prevalence in the rural communities is still low. At $t = 35$, the infection starts to ramp up in the urban cluster at the lower left corner in the map, reaching its peak at $t = 40$. When the rural areas reach their peaks at $t = 52$, the urbanized areas have already achieved sufficient immunity to exit from the infection.

Figure 15 shows the time courses of the frequency distributions ν^S , ν^E and ν^I at three selected locations chosen to represent three essentially different population densities: Using the terminology of the US Census, they are labeled as “Urbanized” for high density, “Urban cluster” for medium density, and “Rural” for low density. The plots show how the peaking times of the infection depend on the population density, but the percentage of individuals

moved from the susceptible to the recovered cohort is higher in the rural community than in the urban cluster. This example demonstrates that the extent of the infection cannot be deduced only from the local population density, but the connectivity structure of the full population needs to be considered.

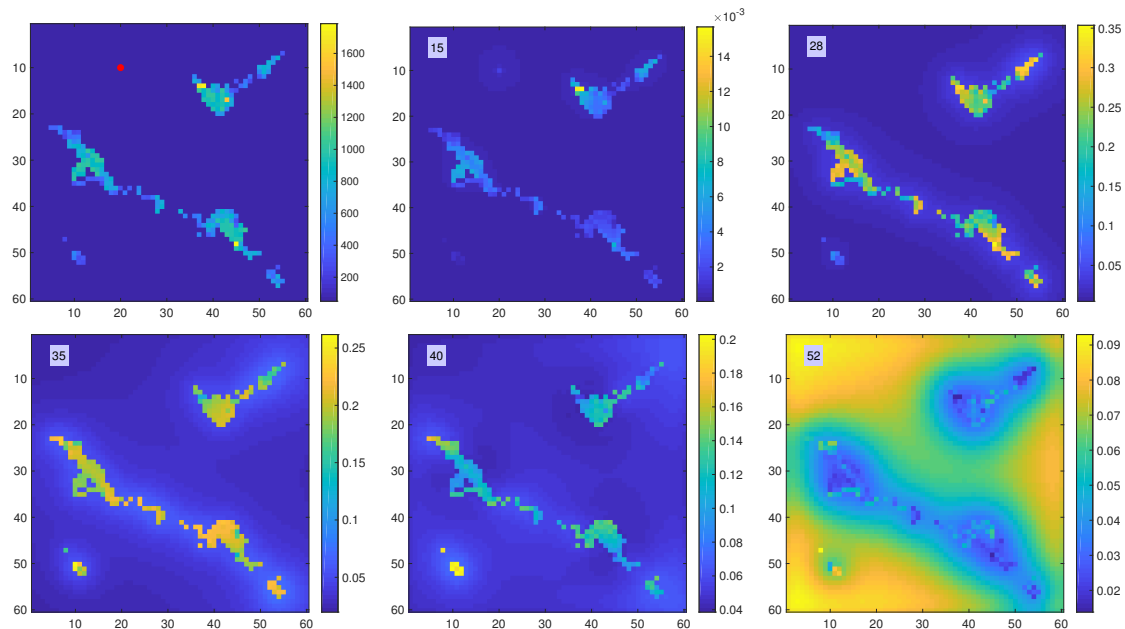


Figure 14. Propagation of the infection. The top left panel shows the population density, in arbitrary units, with the location of the first infection marked by a red dot. Snapshots of the time history of the function $v^I(x, t)$ are shown in lexicographical order. Observe the different color scale in the pictures. At $t = 15$, the infection has reached the nearby urban area, and at $t = 28$ it peaks in all densely populated centers. At $t = 40$, the small center located in bottom left of the map reaches its peak, and at $t = 52$, when the infection in the urban areas has already passed, the rural areas peak.

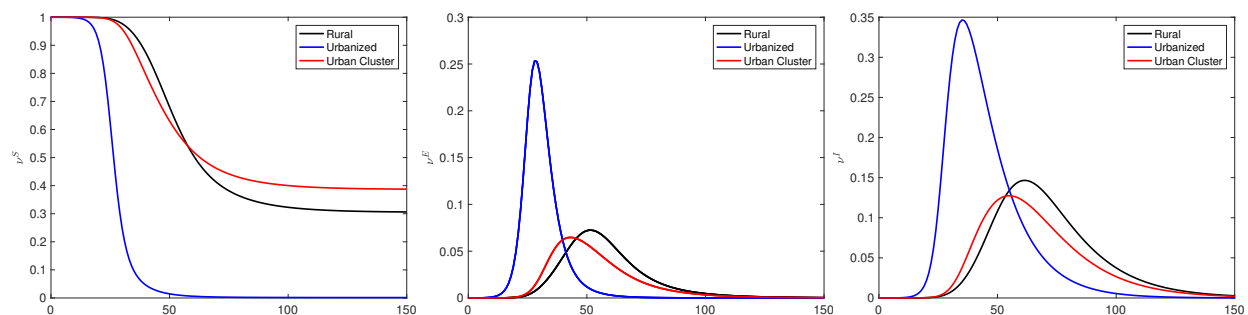


Figure 15. Time traces of the frequencies v^S (left), v^E (middle) and v^I (right) at three selected locations: Referring to the coordinates in Figure 14 as matrix entries with row and column indices (i, j) , the point labeled as “Rural” is at $(i, j) = (10, 10)$, the point labeled as “Urbanized” is at $(i, j) = (30, 13)$, and the point “Urban Cluster” is at $(i, j) = (52, 11)$. We observe that the infection peak times increase as the local population density decreases. Interestingly, in the urban cluster, the percentage of eventually infected individuals is lower than in the rural community, indicating that the dynamics depends on the connectivity and not only on the local density.

One remarkable feature of the curve of the total number of new infection cases over Ω predicted by our model,

$$\# \text{ new infections at time } t = \int_{\Omega} \pi(x) \varphi_2(x, t) dx = \eta \int_{\Omega} \pi(x) v^E(x, t) dx, \quad (62)$$

shown in Figure 16 is its asymmetry. In fact, while the increase in the daily number of new infections follows the characteristic exponential rate, towards the right tail the distribution shows the almost linear decrease observed in the actual COVID-19 case counts.

This asymmetry can be understood by noticing that the initial increase is dominated by the exponential growth in the densely populated centers, while in the decrease phase, the delayed contributions of the less densely populated communities elongate the tail of the infection curve. This suggests that the distribution of new cases does not obey Farr's law of the symmetric rise and fall pattern that has been used in predicting the evolution of, e.g., HIV epidemics [30] as well as the COVID-19 epidemics [31]. The limitations of Farr's law were pointed out recently in connection with the discrepancies between the reported data and the COVID-19 predictions of certain popular statistical models based on it [32]. Our spatially distributed model simulations provide a qualitative explanation of why lumped homogenized models fitted to data at the beginning of the pandemic tend to predict poorly the duration of the epidemics and the shape of the tail. This is yet another example of the limited prediction power of homogenized compartment models applied to situations where the underlying population is rather diverse and there is an underlying significant spatial structure [33].

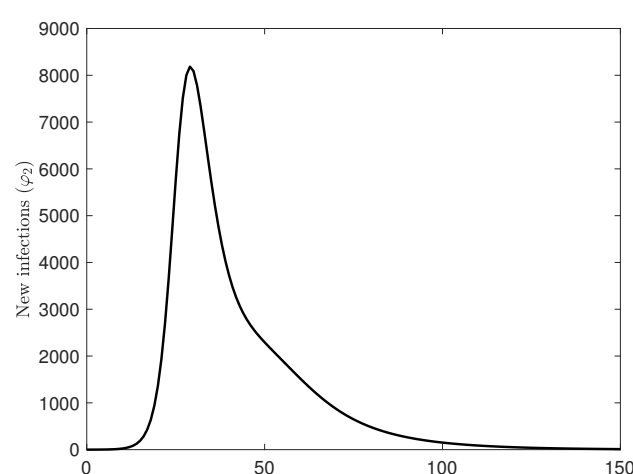


Figure 16. New case count over the entire area. Observe that the new case count increases exponentially, but the decay rate of new cases is lower, with an almost linear segment which is well in line with the observed case count curves, in disagreement with the classical Farr's law stating that the rise and fall behavior is symmetric. The different behavior of the rise and the fall comes from the spatial component of the model and the fact that the population is not well-mixed.

5. Discussion

Metapopulation-based infection models are quintessential for understanding of the infection dynamics in interconnected non-homogenous populations, whether the non-homogeneity is due to geographical, socioeconomic or demographic variability. In this article, we focus on the differences due to geographic characteristics of mutually connected communities and propose a transport scheme based model comprising contact rates that depend on the population densities of the different communities, where the contact between subpopulations is modeled based on commuter traffic. The proposed approach can be extended to include a dynamic redistribution of the population, based, e.g., on cell phone data. The computed examples show how the proposed framework can be used for testing different mitigation scenarios, and sheds some light on the general dynamics of the spread patterns. As an example, the simulations shown in this article elucidate how the infections oscillate between high and low density areas. An infection started in a low density area quickly spreads to the nearby densely populated center and subsequently to the other metropolitan centers in the network, causing an infection flare-up, then returns to the low population areas with an increased intensity, thus demonstrating how the densely populated centers effectively function as amplifiers. The speed at which the infection reaches all three metropolitan centers in the Pittsburgh–Cleveland–Detroit network is noteworthy.

The simulations show the versatility of the proposed model for comparing the effect of statewide travel restrictions versus other measures that aim at lowering the transmission rate within the communities. In particular, the simulations show that even rather drastic travel restrictions (down to 5% of the normal) yield rather limited results, with the densely populated centers retaining their amplifying role, and travel restrictions cause only a slight delay. The reason for the relatively low efficacy of travel restrictions has been discussed in the literature, see, e.g., [13], where the authors showed with a simple model that at the beginning of an outbreak, travel restrictions, at best, slow down the invasion of infection from infected community to a non-infected one, without completely preventing it, while the spread of the infection is predominantly community spread. This can explain why the local mitigation measures are needed to slow down the spread, which cannot be contained solely with travel restrictions. An interesting question for future research is whether the effect of the pandemic on commuter traffic can be estimated from the infection data. This problem is out of the scope of the current study.

Author Contributions: D.C., A.P.H., J.R. and E.S. participated in the conceptualization of the study and the analysis of the results. D.C., A.P.H. and E.S. designed and implemented the algorithms used for the simulations. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by NSF grants DMS-1522334 and DMS-195144 to D.C., and by NSF grant DMS-1714617 to E.S.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Morawska, L.; Cao, J. Airborne transmission of SARS-CoV-2: The world should face the reality. *Environ. Int.* **2020**, *139*, 105730. [\[CrossRef\]](#)
2. Kermack, W.O.; McKendrick, A. A Contribution to the Mathematical Theory of Epidemic. *Proc. R. Soc. Lond.* **1927**, *115*, 700–721.
3. Anderson, R.M.; Anderson, B.; May, R.M. *Infectious Diseases of Humans: Dynamics and Control*; Oxford University Press: Oxford, UK, 1992.
4. Brauer, F.; Castillo-Chavez, C.; Feng, Z. *Mathematical Models in Epidemiology*; Springer: New York, NY, USA, 2019.
5. Anderson, R.M.; May, R.M. Directly transmitted infectious diseases: control by vaccination. *Science* **1982**, *215*, 1053–1060. [\[CrossRef\]](#)
6. Anderson, R.M.; May, R.M. Vaccination and herd immunity to infectious diseases. *Nature* **1985**, *318*, 323–329. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Keeling, M.J.; Rohani, P. *Modeling Infectious Diseases in Humans and Animals*; Princeton University Press: Princeton, NJ, USA, 2011.
8. Levins, R. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Am. Entomol.* **1969**, *15*, 237–240. [\[CrossRef\]](#)
9. Ajelli, M.; Gonçalves, B.; Balcan, D.; Colizza, V.; Hu, H.; Ramasco, J.J.; Merler, S.; Vespignani, A. Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC Infect. Dis.* **2010**, *10*, 190. [\[CrossRef\]](#)
10. Wang, L.; Li, X. Spatial epidemiology of networked metapopulation: An overview. *Chin. Sci. Bull.* **2014**, *59*, 3511–3522. [\[CrossRef\]](#)
11. Hollingsworth, T.D.; Ferguson, N.M.; Anderson, R.M. Will travel restrictions control the international spread of pandemic influenza? *Nat. Med.* **2006**, *12*, 497–499. [\[CrossRef\]](#)
12. Cooper, B.S.; Pitman, R.J.; Edmunds, W.J.; Gay, N.J. Delaying the international spread of pandemic influenza. *PLoS Med.* **2006**, *3*, e212. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Tomba, G.S.; Wallinga, J. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Math. Biosci.* **2008**, *214*, 70–72. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Bajardi, P.; Poletto, C.; Ramasco, J.J.; Tizzoni, M.; Colizza, V.; Vespignani, A. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLoS ONE* **2011**, *6*, e16591. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* **2020**, *368*, 489–493. [\[CrossRef\]](#)
16. Viguerie, A.; Lorenzo, G.; Auricchio, F.; Baroli, D.; Hughes, T.J.; Patton, A.; Reali, A.; Yankeelov, T.E.; Veneziani, A. Simulating the spread of COVID-19 via a spatially-resolved susceptible-exposed-infected-recovered-deceased (SEIRD) model with heterogeneous diffusion. *Appl. Math. Lett.* **2021**, *111*, 106617. [\[CrossRef\]](#)

17. Bureau, U.C. American Community Survey. 2015. Available online: <https://www.census.gov/programs-surveys/acs/data.html> (accessed on 26 May 2020).
18. Calvetti, D.; Hoover, A.P.; Rose, J.; Somersalo, E. Bayesian parameter estimation of an SE(A)IR COVID-19 spread model. *arXiv* **2020**, arXiv:2005.04365.
19. Calvetti, D.; Hoover, A.P.; Rose, J.; Somersalo, E. Metapopulation Network Models for Understanding, Predicting, and Managing the Coronavirus Disease COVID-19. *Front. Phys.* **2020**, *8*, 261. [[CrossRef](#)]
20. Villani, C. *Optimal Transport: Old and New*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008; Volume 338.
21. Santambrogio, F. Optimal transport for applied mathematicians. *Birkhäuser* **2015**, *55*, 94.
22. Giordano, G.; Blanchini, F.; Bruno, R.; Colaneri, P.; Di Filippo, A.; Di Matteo, A.; Colaneri, M. A SIDARTHE model of COVID-19 epidemic in Italy. *arXiv* **2020**, arXiv:2003.09861.
23. Center for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. Available online: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html> (accessed on 26 May 2020).
24. Kyurkchiev, N.; Kyurkchiev, V.; Iliev, A.; Rahnev, A. *Another Look at SIR/SIRD/SEIR/GSEIR Models: New Trends, Methodological Aspects*; Plovdiv University Press: Plovdiv, Bulgaria, 2021.
25. Mitchell, O. COVID-19 Means Fewer Commuters on the Road during ODOT Snow Plowing Season. Available online: <https://www.cleveland.com/news/2020/12/covid-19-means-fewer-commuters-on-the-road-during-odot-snow-plowing-season.html> (accessed on 30 November 2020).
26. Shortell, T. Number of Pennsylvania road deaths kept climbing even as millions stayed home. *The Morning Call*, 12 December 2020.
27. Feather, A. Traffic fatalities up in 2020 despite fewer cars on Michigan roads amid COVID-19 pandemic. *News Channel 3*, 3 December 2020.
28. Cucker, F.; Smale, S. Emergent behavior in flocks. *IEEE Trans. Autom. Control* **2007**, *52*, 852–862. [[CrossRef](#)]
29. Cucker, F.; Smale, S. On the mathematics of emergence. *Jpn. J. Math.* **2007**, *2*, 197–227. [[CrossRef](#)]
30. Bregman, D.J.; Langmuir, A.D. Farr’s law applied to AIDS projections. *JAMA* **1990**, *263*, 1522–1525. [[CrossRef](#)]
31. Pacheco-Barrios, K.; Cardenas-Rojas, A.; Giannoni-Luza, S.; Fregni, F. COVID-19 pandemic and Farr’s law: A global comparison and prediction of outbreak acceleration and deceleration rates. *PLoS ONE* **2020**, *15*, e0239175. [[CrossRef](#)] [[PubMed](#)]
32. Jewell, N.P.; Lewnard, J.A.; Jewell, B.L. Caution warranted: using the Institute for Health Metrics and Evaluation model for predicting the course of the COVID-19 pandemic. *Ann Intern Med.* **2020**, *173*, 226–227. [[CrossRef](#)] [[PubMed](#)]
33. Calvetti, D.; Somersalo, E. Life sciences through mathematical models. *Rend. Lincei* **2015**, *26*, 193–201. [[CrossRef](#)]