Multiplex mobility network and metapopulation epidemic simulations of Italy based on open data

To cite this article: Antonio Desiderio *et al* 2022 *J. Phys. Complex.* **3** 04LT01

View the article online for updates and enhancements.

Journal of Physics: Complexity

LETTER

Multiplex mobility network and metapopulation epidemic simulations of Italy based on open data

Antonio Desiderio[1,2,*] , Gaetano Salina[3] and Giulio Cimini[1,2,3]

1 Physics Department, University of Rome 'Tor Vergata', 00133 Rome, Italy
2 Centro Ricerche Enrico Fermi, 00184 Rome, Italy
3 Istituto Nazionale di Fisica Nucleare, Sezione Roma 'Tor Vergata', 00133 Rome, Italy
* Author to whom any correspondence should be addressed.

E-mail: antonio.desiderio@cref.it

Abstract

The patterns of human mobility play a key role in the spreading of infectious diseases and thus represent a key ingredient of epidemic modeling and forecasting. Unfortunately, as the Covid-19 pandemic has dramatically highlighted, for the vast majority of countries there is no availability of granular mobility data. This hinders the possibility of developing computational frameworks to monitor the evolution of the disease and to adopt timely and adequate prevention policies. Here we show how this problem can be addressed in the case study of Italy. We build a multiplex mobility network based solely on open data, and implement an susceptible-infected-recovered (SIR) metapopulation model that allows scenario analysis through data-driven stochastic simulations. The mobility flows that we estimate are in agreement with real-time proprietary data from smartphones. Our modeling approach can thus be useful in contexts where high-resolution mobility data is not available.

1. Introduction

Facing the SARS-CoV-2 outbreak has required an unprecedented effort for mankind. Governments around the world have applied containment measures in multiple occasions to hinder the spread of the virus. In particular, as human interactions and mobility are known to deeply affect the patterns of epidemic spreading [1–3], non-pharmaceutical interventions (NPIs) have been widely adopted to reduce contacts between individuals [4–6].

NPIs can be generally divided in two main categories [7]. Bottom-up NPIs are self-initiated measurements and behavioral changes due to increase risk perception of the populations [8–10], which can be measured via surveys or mobile data [11–13]. Top-down NPIs are measurements imposed by governments whose effectiveness can be measured via changes in the mobility [14–16] and reduction of the disease prevalence [17]. Making informed decisions, both at the top-down and bottom-up levels, requires timely and quality data about the current stage of the outbreak and contact patterns of individuals [18, 19]. In particular, these data serve to inform computational epidemic models [20], which produce detailed scenario analyses that can be fundamental to inform strategies of response and mitigation of a disease [3, 21–24] and have been widely exploited to face Covid-19 [21, 25–37].

Epidemic models vary in complexity [20], ranging from graph-based [38], agent-based [39] and structured meta-population models [40], each with different data requirements. A common ingredient though is the complex network modeling of social contacts, through which contagion can spread at different scales [41–43]. Mobility data are crucial in this respect [18]. Unfortunately, in many cases the health authorities had to deal with the scarcity and incompleteness of this type of data, which especially in the early stages of the pandemic prevented them from understanding the spreading patterns of the disease and the effects of the various containment measures implemented.

Here we are interested in the case study of Italy, which was the first outbreak of Covid-19 among Western countries and consequently has been the subject of many research studies that tried to model and predict the

evolution of the epidemics. In addition to the development of improved compartment models calibrated on the characteristics of Covid-19 [44–50], much attention has been devoted to the cross-analysis of epidemic and mobility data, in order to evaluate the direct and indirect effect as well as the effectiveness of NPIs.

For instance, Parino *et al* [51] developed a mean-field approach that allows to relate the amount of daily cases to the average number of people that an individual meets per day. The empirical studies have instead used mobile phone data from either Facebook Data For Good [52–54] or telecom operators [55–57]. These data are available as aggregates over territorial units, therefore they have been used to study mobility at the geographic level of regions or provinces. In particular, the authors of [53] have considered the Italian regions as separate entities, due to low level of commuting between regions, while arguing that such inter-regional mobility had a crucial role in the long-range propagation of epidemics [58]. In addition, these proprietary data represent only proxies of the mobility and therefore are affected by systematic effects that cannot be controlled by researchers. Other works [51, 59] defined network models using commuting flows form the 2011 census database of ISTAT (the Italian Institute of Statistics), which however is obsolete and lacks information on long-range mobility.

Through this work, we attempt to overcome the drawbacks of the data sources described above, by building a granular mobility network model of Italy based on data sources that are openly accessible to researchers, as well as up-to-date and complete. The model is defined at the detailed geographical scale of individual municipalities, and is based on a multiplex structure encompassing various means of mobility—which can be tuned independently to mimic the implementation of different lockdown measures. We remark that the use of open data has benefits and drawbacks: they are readily accessible by anyone, can easily be updated in the future, and can be tested against possible systematic errors; however proprietary data typically have higher granularity, especially on the temporal dimension. Besides open data, we also use a gravity-like model as described below to build the mobility network, which thus inherits all the assumption (and possible shortcomings) associated with that model [60]. Nevertheless we show that the flows of population moving among provinces estimated by our multiplex model are quantitatively in agreement with those computed with Facebook Data for Good [61]. We then show how our model can be used to simulate epidemic dynamics and containment measures at the national scale. Here we consider a disease dynamics modeled via an SIR meta-population approach, but the simulation framework can be easily extended to more complex dynamics. We realize representative epidemic scenarios by varying the geographical position and size of the initially infected population, as well as by implementing a different combination of travel limitation measures. In particular we confirm that limiting long range travel is the main effective strategy to slow down the large-scale expansion of the disease. Additionally we show that the effective network distance can be properly used to infer the actual seed of an epidemic outbreak.

2. Methods

2.1. Mobility network

The mobility network is defined by representing each of the $M = 7897$ municipalities in Italy as a node. We use ISTAT data (www.istat.it/it/archivio/6789) to collect information on identification number, reference province, population, surface, latitude and longitude of each municipality. We consider four different mobility layers: intra-province and inter-province layers, representing short range mobility modeled with gravity-like interactions, and train and flight layers, representing long range travels extracted from the origin-destination matrices obtained by government agencies.

To model the mobility of individuals, at each time step of the dynamics we sample a fraction $p_M = 0.1$ of the population of each municipality and move it to other municipalities. This moving population is distributed over the four layers with the following probabilities: $p_{\text{Intra}} = 0.4$, $p_{\text{Inter}} = 0.3$, $p_{\text{Train}} = 0.2$, $p_{\text{Flight}} = 0.1$. These values have been extracted from the origin-destination matrices of individual movements (www.dati.lombardia.it/Mobilit-e-trasporti/Matrice-OD2020-Passeggeri/hyqr-mpe2), as fractions of the total moving population that uses a given transportation method or moves between different provinces. The destinations of the moving population depend on the chosen layer as described in the following. At the end of the current time step, the moving population returns to the municipality of departure.

2.2. Short range networks: intra-province and inter-province layers

The connectivity of the intra-province and inter-province layers is set by the territorial contiguity matrix collected by ISTAT (www.istat.it/it/archivio/157423): $c_{ij} = 1$ if two municipalities i and j are geographically adjacent and $c_{ij} = 0$ otherwise. We use the reference province of a municipality to split the connections into the two layers: if the province of municipality i is different from the province of municipality j, then the connection is placed in the inter-province layer; otherwise it belongs to the intra-province layer. Connections in these layers are weighted using gravity-like assumptions [60]: for the link between nodes i and j we pose

$$w_{ij}^{L} = \frac{\rho_i \rho_j}{r_{ij}} c_{ij}^{L} \qquad (1)$$

where ρ_i is the population density of node i (population over surface), r_{ij} is the distance between the two municipalities and $L \in \{\text{Intra, Inter}\}$ denotes the layer. In order to transform such weights into Markovian connection probabilities we use the following normalization:

$$\tilde{w}_{ij}^{L} = \frac{w_{ij}^{L}}{\sum_{l \in n_i^{L}} w_{il}^{L}} c_{ij}^{L} \qquad (2)$$

where n_i^{L} are the territorial neighbors of node i in the corresponding layer.

2.3. Long range networks: trains and flights layers

To build the train layer, we use the RESTful API of Viaggiatreno (www.viaggiatreno.it/) to retrieve information of the train stations and the connections between them. Each station is identified by a unique id, latitude and longitude, while for each train departing from a given station we have the id of the next stop station and the type of connection (Frecciarossa, Intercity, Metropolitana, EuroNight, Regionale, EuroCity). To map the railway stations on the municipalities, we associated each station with the municipalities that are located no more than 5 km away (in this way, on average a single station corresponds to three municipalities). Overall the weight of a link from node i to node j in the train layer is built as $w_{ij}^{\text{Train}} = \sum_e \tau_{ij}^e g^e$, where τ_{ij}^e is a Boolean indicator for the presence of a connection from i to j by train type e, while g^e is the fraction of total travelers for train type e, obtained from ISTAT (www.istat.it/it/archivio/13995). Finally, as in the short range network, to obtain connection probabilities each w_{ij}^{Train} is normalized over the total weight of the links from i to all possible destinations reachable by train in one stop.

To build the flight layer, we collect IATA and ICAO codes, latitudes and longitudes of airports from OpenFlights (https://openflights.org/data.html). From the annual ENAC report (www.enac.gov.it/pubblicazioni/dati-di-traffico-2019) we collected the air traffic data: origin and destination airports, identified by the IATA and ICAO codes, and total passengers moving between airports. To map the airports on the municipalities, we associated each airport with all municipalities located no more than 15 km away (on average a single airport is mapped on ten municipalities). Overall the weight of a link from node i to node j in the flight layer, w_{ij}^{Flight}, is defined as the fraction of total travelers from i to j. As for the other layers, weights are normalized to obtain connection probabilities.
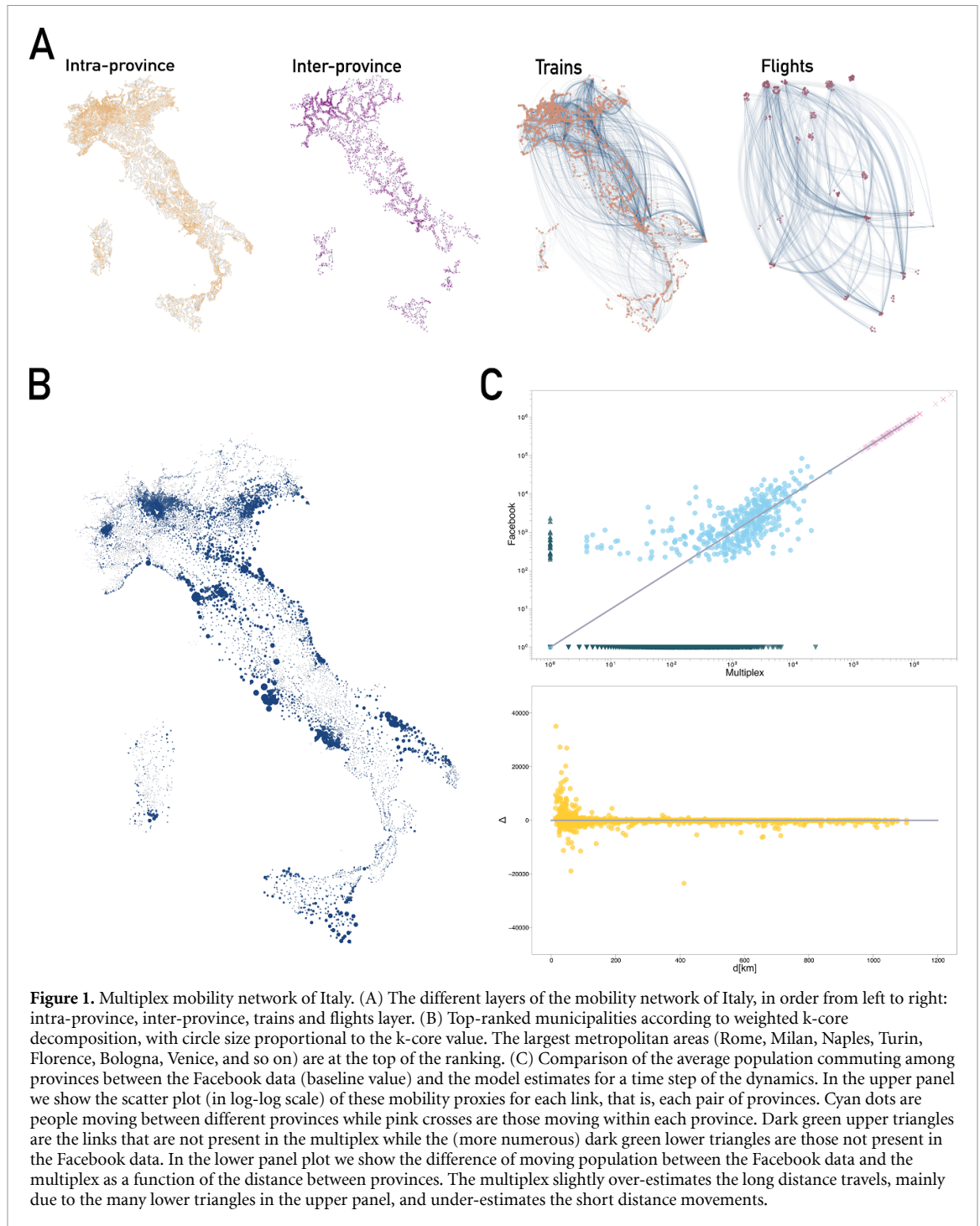
3. Results and discussion

3.1. Comparison of simulated mobility fluxes with Facebook data

Overall the multiplex network model is build using a combination of a gravity-like approach for short distance travels [60] and of a data-driven approach for long range trains and flights. We have derived connection (transition) probabilities between nodes that are stationary and pairwise-independent, which allow representing mobility as a simple Markovian dynamics. In particular the probability that an individual moves from node i to node j in a single time step, $p_{ij} = p_M \sum_L p_L \tilde{w}_{ij}^{L}$ where $L \in \{\text{Intra, Inter, Train, Flight}\}$, is such that $\sum_j p_{ij} = 1 - p_{ii}$ where $p_{ii} = 1 - p_M$ is the probability to remain on node i. Extensive probability fluxes (i.e. the moving population) between any two nodes i, j can be then computed as $\omega_{ij} = P_i p_{ij}$, where P_i is the population of the starting node i. We remark that our model is defined at the geographical resolution of individual municipalities, contrarily to other works that take provinces or regions as minimal spatial units [29, 53]. Figure 1(A) depicts the different layers that compose the network. Clearly the intra- and inter-province layers exhibit a short distance nature of the connections, whilst the trains and the flights layers represent the long range connections.

Given the mobility network, we identify the most important (or 'vital') nodes via k-core decomposition [62] based on the weighted out-degree, which for a generic node i reads

$$k_i^{W} = (k_i^a s_i^b)^{\frac{1}{a+b}}, \qquad (3)$$

with $s_i = \sum_j \omega_{ij}$ and $k_i = \sum_j \Theta(\omega_{ij})$, where $\Theta(x)$ equals 1 for $x > 0$ and 0 otherwise. We use $a = 1/10$ and $b = 1$ to give more weight to the strength in the decomposition since it represents the extent of the mobility. Figure 1(B) shows that the top-ranked nodes are those corresponding to the largest metropolitan areas. Notably, the weighted k-core ranking of a node is derived using only the structural properties of the network, yet it is consistent with a ranking obtained from the size of the epidemic outbreak starting from that node (Kendall's τ for the ranking restricted to region capitals equal to $\tau = 0.4$ at $t = 60$ time steps of the simulations and to $\tau = 0.6$ at $t = 80$).

Figure 1. Multiplex mobility network of Italy. (A) The different layers of the mobility network of Italy, in order from left to right: intra-province, inter-province, trains and flights layer. (B) Top-ranked municipalities according to weighted k-core decomposition, with circle size proportional to the k-core value. The largest metropolitan areas (Rome, Milan, Naples, Turin, Florence, Bologna, Venice, and so on) are at the top of the ranking. (C) Comparison of the average population commuting among provinces between the Facebook data (baseline value) and the model estimates for a time step of the dynamics. In the upper panel we show the scatter plot (in log-log scale) of these mobility proxies for each link, that is, each pair of provinces. Cyan dots are people moving between different provinces while pink crosses are those moving within each province. Dark green upper triangles are the links that are not present in the multiplex while the (more numerous) dark green lower triangles are those not present in the Facebook data. In the lower panel plot we show the difference of moving population between the Facebook data and the multiplex as a function of the distance between provinces. The multiplex slightly over-estimates the long distance travels, mainly due to the many lower triangles in the upper panel, and under-estimates the short distance movements.

We then check whether our mobility model can well represent real data. We thus compare to Facebook Data for Good information; in particular we use the 'baseline' indicator of the Movement Maps. This value represents the daily average of the count of movements between two administrative regions over a time window of 45 days, preceding day 0 of data collection (23 February 2020 in our case). As the minimal spatial unit of the Facebook data is the province, to perform the comparison we compute inter-province fluxes in the multiplex model as the number of people commuting among provinces. Figure 1(C) shows that such mobility fluxes are comparable to Facebook data, the main difference being the under-estimation of short distance mobility by our model and the shortage of long distance travels by Facebook data.

3.2. Disease dynamics: different starting points
We now couple the mobility with an epidemic dynamics (see figure 2(A)) in order to simulate realistic scenarios of the disease diffusion. We choose to use the SIR compartmental model (simulated with a
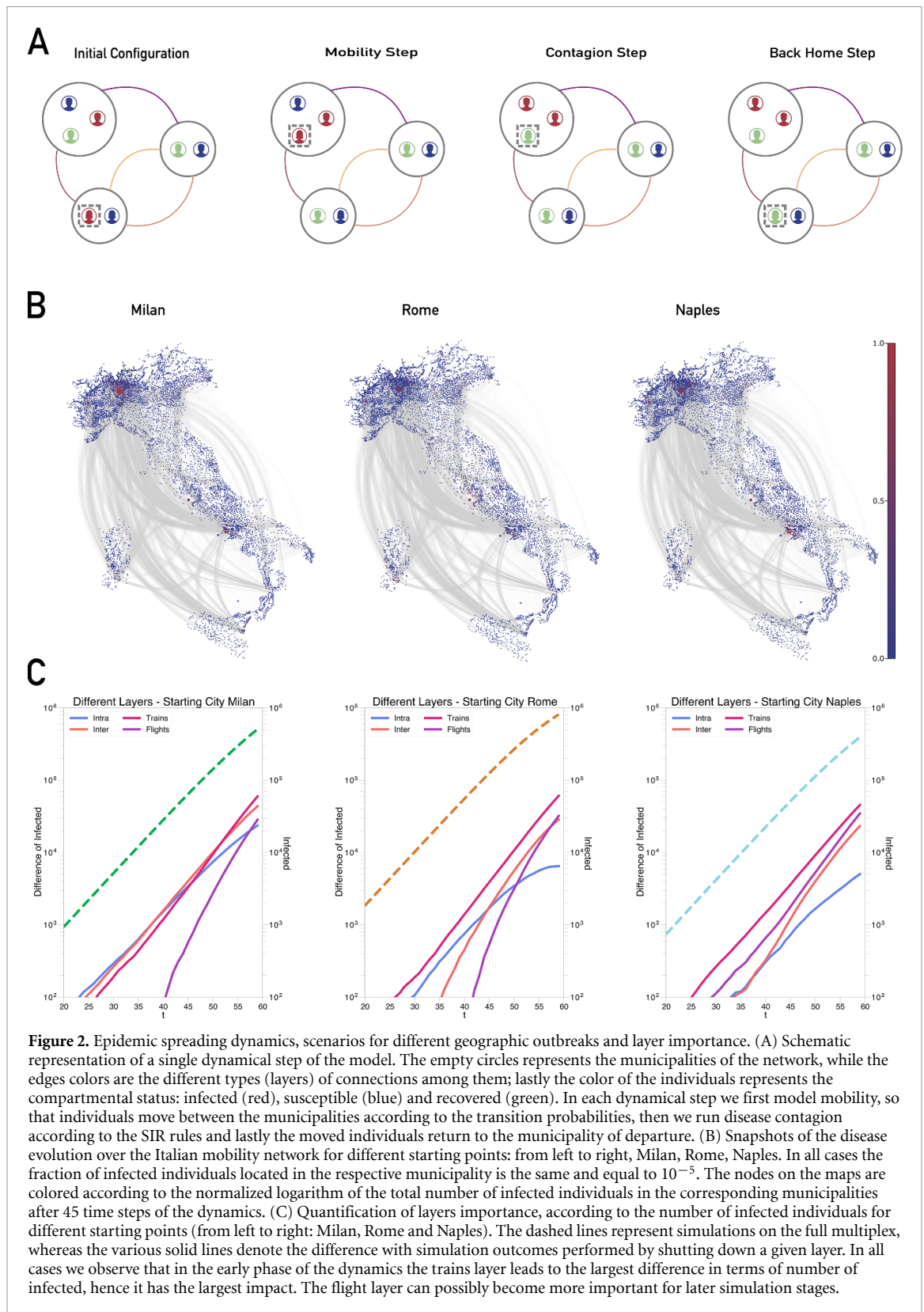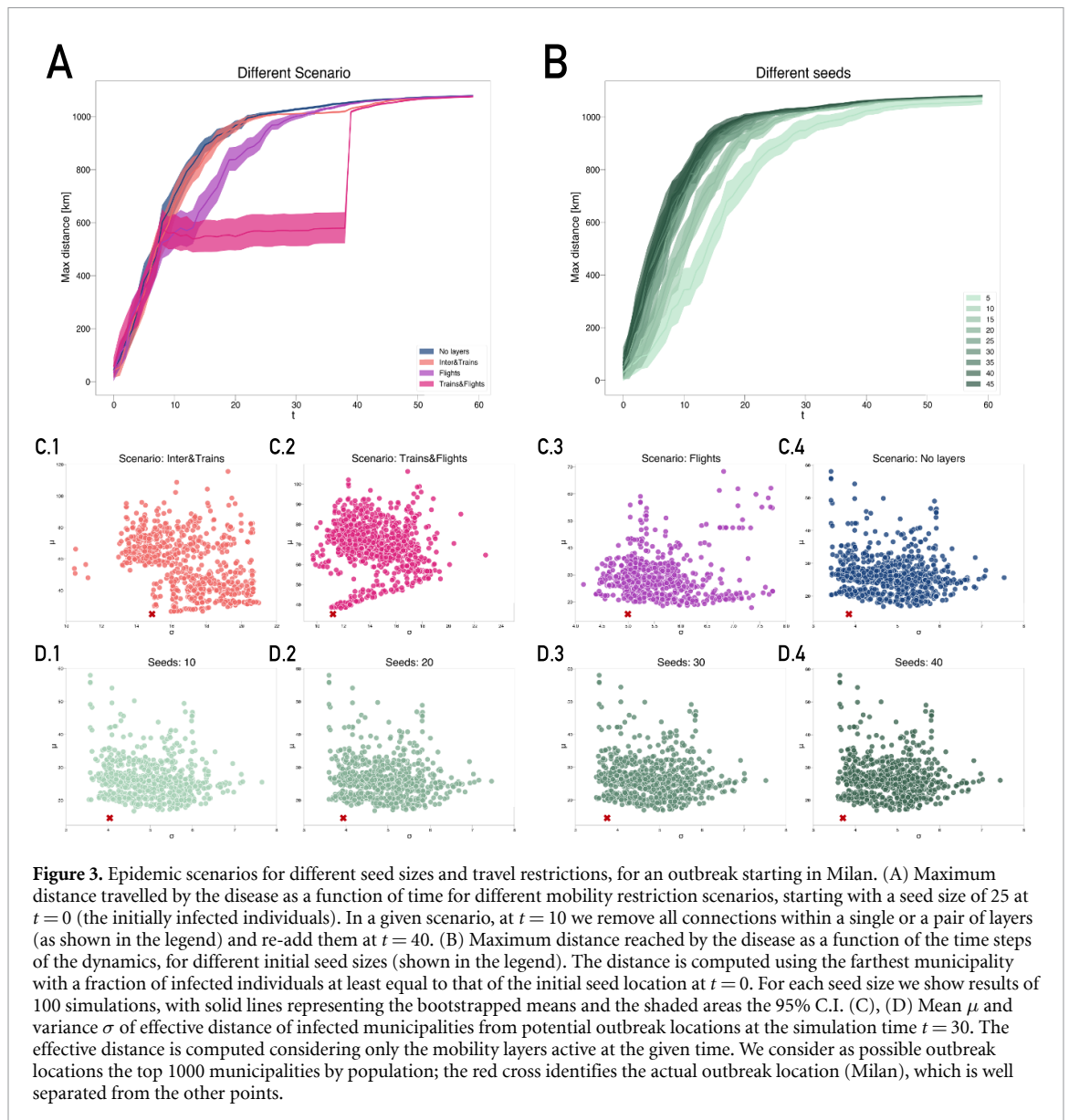
Figure 2. Epidemic spreading dynamics, scenarios for different geographic outbreaks and layer importance. (A) Schematic representation of a single dynamical step of the model. The empty circles represents the municipalities of the network, while the edges colors are the different types (layers) of connections among them; lastly the color of the individuals represents the compartmental status: infected (red), susceptible (blue) and recovered (green). In each dynamical step we first model mobility, so that individuals move between the municipalities according to the transition probabilities, then we run disease contagion according to the SIR rules and lastly the moved individuals return to the municipality of departure. (B) Snapshots of the disease evolution over the Italian mobility network for different starting points: from left to right, Milan, Rome, Naples. In all cases the fraction of infected individuals located in the respective municipality is the same and equal to 10^{-5}. The nodes on the maps are colored according to the normalized logarithm of the total number of infected individuals in the corresponding municipalities after 45 time steps of the dynamics. (C) Quantification of layers importance, according to the number of infected individuals for different starting points (from left to right: Milan, Rome and Naples). The dashed lines represent simulations on the full multiplex, whereas the various solid lines denote the difference with simulation outcomes performed by shutting down a given layer. In all cases we observe that in the early phase of the dynamics the trains layer leads to the largest difference in terms of number of infected, hence it has the largest impact. The flight layer can possibly become more important for later simulation stages.

discrete-time reactive process). The reason is that the simple SIR features the key dynamical ingredients that trigger the exponential trend and thus it can be used to study the territorial expansion of the disease (namely the branching process at the basis of the S→I transition). Therefore this approach suffices for our purposes of simulating different scenarios on short temporal scales, without fitting of real data. Note however that our framework can be easily implemented with any of the more refined epidemiological models for Covid-19 proposed in the literature (see [7] for a review). These approach try to overcome some limitations of the SIR

Figure 3. Epidemic scenarios for different seed sizes and travel restrictions, for an outbreak starting in Milan. (A) Maximum distance travelled by the disease as a function of time for different mobility restriction scenarios, starting with a seed size of 25 at $t = 0$ (the initially infected individuals). In a given scenario, at $t = 10$ we remove all connections within a single or a pair of layers (as shown in the legend) and re-add them at $t = 40$. (B) Maximum distance reached by the disease as a function of the time steps of the dynamics, for different initial seed sizes (shown in the legend). The distance is computed using the farthest municipality with a fraction of infected individuals at least equal to that of the initial seed location at $t = 0$. For each seed size we show results of 100 simulations, with solid lines representing the bootstrapped means and the shaded areas the 95% C.I. (C), (D) Mean μ and variance σ of effective distance of infected municipalities from potential outbreak locations at the simulation time $t = 30$. The effective distance is computed considering only the mobility layers active at the given time. We consider as possible outbreak locations the top 1000 municipalities by population; the red cross identifies the actual outbreak location (Milan), which is well separated from the other points.

framework, adding for instance more compartments and dynamic values of R_0, as well as considering an aged-structured population.

Following the meta-population approach used in [29], within a given municipality i the transition rules between compartments are:

$$P(S_i \rightarrow I_i) = 1 - \left(1 - \frac{R_0 \mu}{N_i} \right)^{I_i} \tag{4}$$

$$P(I_i \rightarrow R_i) = \mu \tag{5}$$

where N_i is the total population in municipality i while S_i, I_i and R_i are the number of individuals in i that are susceptible, infected and recovered, respectively. R_0 is the mean reproduction number and μ is the inverse of mean infectious period (here we consider values of disease parameters taken from [29]: $R_0 = 2.5$ and $\mu = 1/8$).

In figure 2(B) we plot the results of the model evolution, averaged over 100 numerical simulations, after 45 time steps of the dynamics. We consider different representative starting seeds of the disease (the seed is the municipality of the initially infected individuals) and observe markedly different spreading patterns, both regarding the short and long distances. We also investigate the impact of each individual layer on the spreading dynamics, by removing it from the mobility network (for example, removing the flight layer means that air travels become forbidden): individuals who are chosen to use the removed layer simply remain in

their departure location. Figure 2(C) shows, for the different starting seeds, the difference of infected individuals between simulations with the complete multiplex and those where a given layer is removed. We consistently observe that in the early phase of the spreading the impact of the trains layer is the largest, with the flight layer eventually becoming more important at later stages.

3.3. Different scenarios of seed sizes and mobility restrictions

Here we further illustrate how the multiplex model can be used to simulate epidemic scenarios under mobility restrictions. As representative example, we analyzed the maximum distance reached by the disease from the seed location. This distance is computed, for a given time step of the dynamics, considering the farthest municipality with a fraction of infected individuals at least equal to that of the initial seed location at $t = 0$.[4] Figure 3(B) shows that this observable linearly depends on the size of the starting seed, whereas, figure 3(A) illustrates the effect of different restrictions on the mobility—simulated by removing a combination of layers from the network. This simple exercise confirms the effectiveness of long range travels restrictions in slowing down the territorial expansion of the disease [41, 58].

Besides geographical distance, we also consider the effective network distance as introduced in [63]: $d_{ij} = 1 - \log(p_{ij})$ for nodes that are directly connected, while $D_{ij} = \min_\Gamma \lambda(\Gamma)$ otherwise, where $\lambda(\Gamma)$ is the length of a directed path from i to j, obtained by summing the effective lengths along the legs of the path. Figures 3(C) and (D) shows the mean and the variance of the effective distance of infected municipalities from potential outbreak locations (among the top 1000 municipalities by population). The real outbreak location (represented as a red cross) is markedly separated from the other points. Panels (C) shows different mobility restriction scenarios: the mean effective distances of scenarios C.1 and C.2 (more severe restrictions) are a factor of 2 bigger than those of scenarios C.3 and C.4, and also the shapes of the point clouds are markedly different. Panels (D) show instead the case of different sizes of the starting seed: the variances of the effective distance, both for the actual and potential outbreak locations, decreases as the seed size increases.

4. Conclusions

In this work we developed a multiplex mobility network of Italy, defined at the fine-grained level of municipalities. We then showed how the model can be used to perform different scenario analysis. The Covid-19 outbreak has highlighted the lack of such a tool, at least for Italy, that would certainly have helped policy makers to monitor the disease evolution and to promptly adopt adequate prevention policies. We built the networked data structure using only open data collected by government agencies. However the estimated mobility flows are well compatible with empirical Facebook geo-localization data from mobile phones. Nevertheless, the comparison with Facebook data has shown a limitation of our approach, namely the absence of a true high quality data source for short distance travels. While we used territorial contiguity to inform a gravity-like approach, considering (at the moment unavailable) open data on road and highway connections certainly represent interesting directions to further extend our approach. Overall, we believe our modeling framework can be a good starting point to build a detailed simulator for Italy to inform response strategies for future epidemics outbreaks.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://zenodo.org/record/7050931#.YyR7Wi8QNZI.

Code availability statement

The code that support the findings of this study are openly available at the following url https://github.com/RiegelGestr/multiplex_mobility_Italy_municipalities.

Author contributions

A D wrote the simulation code, performed the analysis and realized the figure. G C and G S designed the analysis and supervised the project. All the authors gathered the data, wrote the paper and discussed the results.

---

[4] We consider such an intensive indicator to mark a municipality as 'infected' in order to fairly compare small and large population sizes. Other definitions are of course possible, such as those based on extensive thresholds.

Conflict of interest

The authors declare no competing interests.

ORCID iD

Antonio Desiderio ⓘ https://orcid.org/0000-0002-9971-450X

References

[1] Eubank S, Guclu H, Kumar V S A, Marathe M V, Srinivasan A, Toroczkai Z and Wang N 2004 Modelling disease outbreaks in realistic urban social networks *Nature* **429** 180

[2] Hufnagel L, Brockmann D and Geisel T 2004 Forecast and control of epidemics in a globalized world *Proc. Natl Acad. Sci.* **101** 15124

[3] Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco J J and Vespignani A 2009 Multiscale mobility networks and the spatial spreading of infectious diseases *Proc. Natl Acad. Sci.* **106** 21484

[4] Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, Pinior B, Thurner S and Klimek P 2020 Ranking the effectiveness of worldwide COVID-19 government interventions *Nat. Hum. Behav.* **4** 1303

[5] Bo Y *et al* 2021 Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020 *Int. J. Infect. Dis.* **102** 247

[6] Mendez-Brito A, Bcheraoui C E and Pozo-Martin F 2021 Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19 *J. Infect.* **83** 281

[7] Perra N 2021 Non-pharmaceutical interventions during the COVID-19 pandemic: a review *Phys. Rep.* **913** 1–52

[8] Wise T, Zbozinek T D, Michelini G, Hagan C C and Mobbs D 2020 Changes in risk perception and self-reported protective behaviour during the first week of the COVID-19 pandemic in the United States *R. Soc. Open Sci.* **7** 200742

[9] Bae S Y and Chang P-J 2021 The effect of coronavirus disease-19 (COVID-19) risk perception on behavioural intention towards 'untact' tourism in South Korea during the first wave of the pandemic (March 2020) *Curr. Issues Tour.* **24** 1017

[10] Lüdecke D and von dem Knesebeck O 2020 Protective behavior in course of the COVID-19 outbreak—survey results from Germany *Front. Public Health* **8** 572561

[11] Muto K, Yamamoto I, Nagasu M, Tanaka M and Wada K 2020 Japanese citizens' behavioral changes and preparedness against COVID-19: an online survey during the early phase of the pandemic *PLoS One* **15** e0234292

[12] Bönisch S, Wegscheider K, Krause L, Sehner S, Wiegel S, Zapf A, Moser S and Becher H 2020 Effects of coronavirus disease (COVID-19) related contact restrictions in Germany, March to May 2020, on the mobility and relation to infection patterns *Front. Public Health* **8** 568287

[13] Sun S *et al* 2020 Using smartphones and wearable devices to monitor behavioral changes during COVID-19 *J. Med. Internet Res.* **22** e19992

[14] Santana C, Botta F, Barbosa H, Privitera F, Menezes R and Di Clemente R 2022 Changes in the time-space dimension of human mobility during the COVID-19 pandemic (arXiv:2201.06527)

[15] Woskie L R *et al* 2021 Early social distancing policies in Europe, changes in mobility & COVID-19 case trajectories: insights from Spring 2020 *PLoS One* **16** 1–12

[16] Gozzi N, Comini N and Perra N 2022 The adoption of non-pharmaceutical interventions and the role of digital infrastructure during the COVID-19 pandemic in Colombia, Ecuador, and El Salvador (arXiv:2202.12088)

[17] Schlosser F, Maier B F, Jack O, Hinrichs D, Zachariae A and Brockmann D 2020 Covid-19 lockdown induces disease-mitigating structural changes in mobility networks *Proc. Natl Acad. Sci.* **117** 32883

[18] Oliver N *et al* 2020 Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle *Sci. Adv.* **6** eabc0764

[19] Starnini M, Aleta A, Tizzoni M and Moreno Y 2021 Impact of data accuracy on the evaluation of COVID-19 mitigation policies *Data Policy* **3** e28

[20] Pastor-Satorras R, Castellano C, Van Mieghem P and Vespignani A 2015 Epidemic processes in complex networks *Rev. Mod. Phys.* **87** 925

[21] Ferguson N M, Cummings D A T, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S and Burke D S 2005 Strategies for containing an emerging influenza pandemic in Southeast Asia *Nature* **437** 209

[22] Tizzoni M, Bajardi P, Poletto C, Ramasco J J, Balcan D, Gonçalves B, Perra N, Colizza V and Vespignani A 2012 Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm *BMC Med.* **10** 165

[23] Longini I M, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings D A T and Halloran M E 2005 Containing pandemic influenza at the source *Science* **309** 1083

[24] Germann T C, Kadau K, Longini I M and Macken C A 2006 Mitigation strategies for pandemic influenza in the United States *Proc. Natl Acad. Sci.* **103** 5935

[25] Broeck W V d, Gioannini C, Gonçalves B, Quaggiotto M, Colizza V and Vespignani A 2011 The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale *BMC Infect. Dis.* **11** 37

[26] Hollingsworth T D, Klinkenberg D, Heesterbeek H and Anderson R M 2011 Mitigation strategies for pandemic influenza a: balancing conflicting policy objectives *PLoS Comput. Biol.* **7** e1001076

[27] Zlojutro A, Rey D and Gardner L 2019 A decision-support framework to optimize border control for global outbreak mitigation *Sci. Rep.* **9** 2216

[28] Aleta A *et al* 2020 Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scenarios of the COVID-19 epidemic *medRXiv Preprint* https://doi.org/10.1101/2020.05.06.20092841 (Accessed 10 October 2022)

[29] Aleta A and Moreno Y 2020 Evaluation of the potential incidence of COVID-19 and effectiveness of contention measures in Spain: a data-driven approach *medRxiv Preprint* https://doi.org/10.1101/2020.03.01.20029801 (Accessed 10 October 2022)

[30] Chang S, Pierson E, Koh P W, Gerardin J, Redbird B, Grusky D and Leskovec J 2021 Mobility network models of COVID-19 explain inequities and inform reopening *Nature* **589** 82

[31] Barrat A, Cattuto C, Kivelä M, Lehmann S and Saramäki J 2021 Effect of manual and digital contact tracing on COVID-19 outbreaks: a study on empirical contact data *J. R. Soc. Interface* **18** 20201000

[32] Susswein Z, Valdano E, Brett T, Rohani P, Colizza V and Bansal S 2021 Ignoring spatial heterogeneity in drivers of SARS-CoV-2 transmission in the US will impede sustained elimination *medRXiv Preprint* https://doi.org/10.1101/2021.08.09.21261807 (Accessed 10 October 2022)

[33] Aleta A *et al* 2021 Quantifying the importance and location of SARS-CoV-2 transmission events in large metropolitan areas *medRxiv Preprint* https://doi.org/10.1101/2020.12.15.20248273 (Accessed 10 October 2022)

[34] Domenico L D, Pullano G, Sabbatini C E, Boëlle P-Y and Colizza V 2021 Modelling safe protocols for reopening schools during the COVID-19 pandemic in France *Nat. Commun.* **12** 1073

[35] Abueg M *et al* 2021 Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state *npj Digit. Med.* **4** 49

[36] Mazzoli M, Pepe E, Mateo D, Cattuto C, Gauvin L, Bajardi P, Tizzoni M, Hernando A, Meloni S and Ramasco J J 2021 Interplay between mobility, multi-seeding and lockdowns shapes COVID-19 local impact *PLoS Comput. Biol.* **17** e1009326

[37] Contreras D A, Colosi E, Bassignana G, Colizza V and Barrat A 2022 Impact of the representation of contact data on the evaluation of interventions in infectious diseases simulations *medRxiv Preprint* https://doi.org/10.1101/2022.02.28.22271600 (Accessed 10 October 2022)

[38] Pastor-Satorras R and Vespignani A 2001 Epidemic spreading in scale-free networks *Phys. Rev. Lett.* **86** 3200

[39] Liu Q-H, Ajelli M, Aleta A, Merler S, Moreno Y and Vespignani A 2018 Measurability of the epidemic reproduction number in data-driven contact networks *Proc. Natl Acad. Sci.* **115** 12680

[40] Wang L and Li X 2014 Spatial epidemiology of networked metapopulation: an overview *Chin. Sci. Bull.* **59** 3511

[41] Colizza V, Barrat A, Barthélemy M and Vespignani A 2006 The role of the airline transportation network in the prediction and predictability of global epidemics *Proc. Natl Acad. Sci.* **103** 2015

[42] Moreno Y, Pastor-Satorras R and Vespignani A 2002 Epidemic outbreaks in complex heterogeneous networks *Eur. Phys. J. B* **26** 521

[43] Kraemer M U G *et al* 2020 The effect of human mobility and control measures on the COVID-19 epidemic in China *Science* **368** 493

[44] Parolini N, Dede' L, Ardenghi G and Quarteroni A 2021 Modelling the COVID-19 epidemic and the vaccination campaign in Italy by the SUIHTER model (arXiv:2112.11722)

[45] Giordano G, Blanchini F, Bruno R, Colaneri P, Filippo A D, Matteo A D and Colaneri M 2020 Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy *Nat. Med.* **26** 855

[46] Rossa F D *et al* 2020 A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic *Nat. Commun.* **11** 5106

[47] Arcede J P, Caga-anan R L, Mentuda C Q and Mammeri Y 2020 Accounting for symptomatic and asymptomatic in a SEIR-type model of COVID-19 *Math. Model. Nat. Phenom.* **15** 34

[48] De Falco I, Della Cioppa A, Scafuri U and Tarantino E 2020 Coronavirus COVID-19 spreading in Italy: optimizing an epidemiological model with dynamic social distancing through differential evolution (arXiv:2004.00553)

[49] Guzzetta G *et al* (COVID-19 Working Group) 2020 The impact of a nation-wide lockdown on COVID-19 transmissibility in Italy (arXiv:2004.12338)

[50] Pizzuti C, Socievole A, Prasse B and Mieghem P V 2020 Network-based prediction of COVID-19 epidemic spreading in Italy *Appl. Netw. Sci.* **5** 91

[51] Parino F, Zino L, Porfiri M and Rizzo A 2021 Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading *J. R. Soc. Interface* **18** 20200875

[52] Galeazzi A, Cinelli M, Bonaccorsi G, Pierri F, Schmidt A L, Scala A, Pammolli F and Quattrociocchi W 2021 Human mobility in response to COVID-19 in France, Italy and UK *Sci. Rep.* **11** 13141

[53] Scala A, Flori A, Spelta A, Brugnoli E, Cinelli M, Quattrociocchi W and Pammolli F 2020 Time, space and social interactions: exit mechanisms for the COVID-19 epidemics *Sci. Rep.* **10** 13764

[54] Smolyak A, Bonaccorsi G, Flori A, Pammolli F and Havlin S 2021 Effects of mobility restrictions during COVID-19 in Italy *Sci. Rep.* **11** 21783

[55] Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C and Tizzoni M 2020 COVID-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown *Sci. Data* **7** 230

[56] Cintia P *et al* 2020 The relationship between human mobility and viral transmissibility during the COVID-19 epidemics in Italy (arXiv:2006.03141)

[57] Iacus S M, Santamaria C, Sermi F, Spyratos S, Tarchi D and Vespe M 2020 Human mobility and COVID-19 initial dynamics *Nonlinear Dyn.* **101** 1901

[58] Gross B, Zheng Z, Liu S, Chen X, Sela A, Li J, Li D and Havlin S 2020 Spatio-temporal propagation of COVID-19 pandemics *Europhys. Lett.* **131** 58003

[59] Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R and Rinaldo A 2020 Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures *Proc. Natl Acad. Sci.* **117** 10484

[60] Barbosa H, Barthelemy M, Ghoshal G, James C R, Lenormand M, Louail T, Menezes R, Ramasco J J, Simini F and Tomasini M 2018 Human mobility: models and applications *Phys. Rep.* **734** 1–74

[61] Maas P 2019 Facebook disaster maps: aggregate insights for crisis response & recovery *Proc. 25th ACM SIGKDD Int. Conf. on Knowledge Discovery Data Mining*

[62] Lü L, Chen D, Ren X-L, Zhang Q-M, Zhang Y-C and Zhou T 2016 Vital nodes identification in complex networks *Phys. Rep.* **650** 1

[63] Brockmann D and Helbing D 2013 The hidden geometry of complex, network-driven contagion phenomena *Science* **342** 1337