Connecting models of disease spread over realistic population networks and over the mathematical "fully-mixed" society

Magdalena Georgieva[1], Shweta Bansal[2], Vittoria Colizza[3,4]

[1]University of Oxford, Oxford, UK; [2]Department of Biology, Georgetown University, USA;
[3]Sorbonne Université, Faculté de Mèdecine, Paris, France; [4]INSERM, Paris, France

Abstract

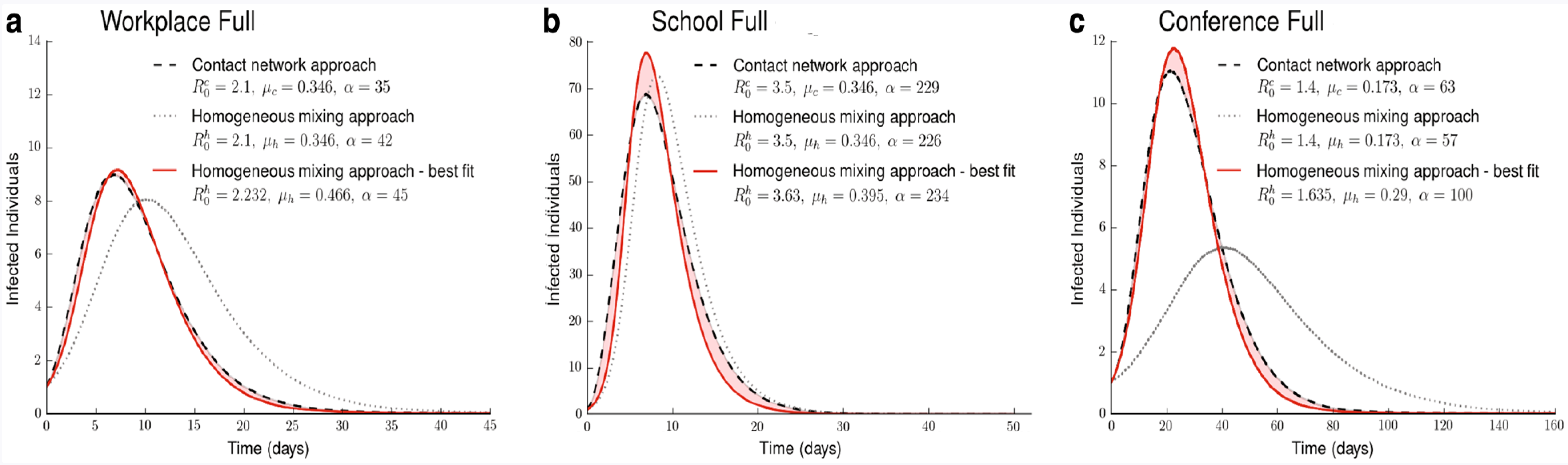
This project links well-researched basic modelling approaches for infection spread to the latest attempts in incorporating complexity for small networks. This connection can create opportunities for easier property inference and further refinements in epidemiology, including extensions to large populations. There is some recent research on the topic but the linear connection assumed in it between infectious individuals and the occurrence of new cases cannot capture the dynamics for all network configurations.

We look into the case of a non-linear relationship which can overcome this problem. We use computer-generated populations controlling for modularity (i.e. existence of communities within the network) to simulate disease propagation for different sets of transmissibility and recovery parameters corresponding to different types of diseases.

A fitting procedure is then run for the obtained epidemics to find the equivalent infections (characterised by their parameters) over the simple population in which everyone communicates with everyone else at random and these are analysed to elicit useful patterns.

Background

There exist well-established methods to study epidemic spread over populations as long as we assume that they are homogeneous and well-mixed, meaning that everyone can communicate with everyone else at random. New research by Bioglio et al. has suggested the possibility of mapping real world scenarios (which are complicated because of time-variablity and heterogenety of mixing) to the known well-mixed ones.



Fitting a homogeneous linear model to results from simulations of epidemics on networks (the parameters are fitted to minimise the least squares difference between the two curves).
Three real small network datasets are used (workplace, school and conference).
Source: L. Bioglio, M Génois, C.L. Vestergaard, C. Poletto, A. Barrat, V. Colizza. Recalibrating disease parameters for increasing realism in modeling epidemics in closed settings. BMC Infectious Diseases. 2016 . 16(1)

While promising, these results suggested that clustered networks influence disease propagation in a way which makes the fit to a linear well-mixed models poor. So, we decided to consider such networks and to introduce non-linearity in the model before fitting to realistic simulated scenarios.

Methods

Step 1: Static Network Simulation
Constructing an algorithm which generates networks with specific modularity.

Step 2: Simulate disease propagation over the network
This is averaged over hundreds of runs.

Step 3: Build a non-linear homogeneous mixing model

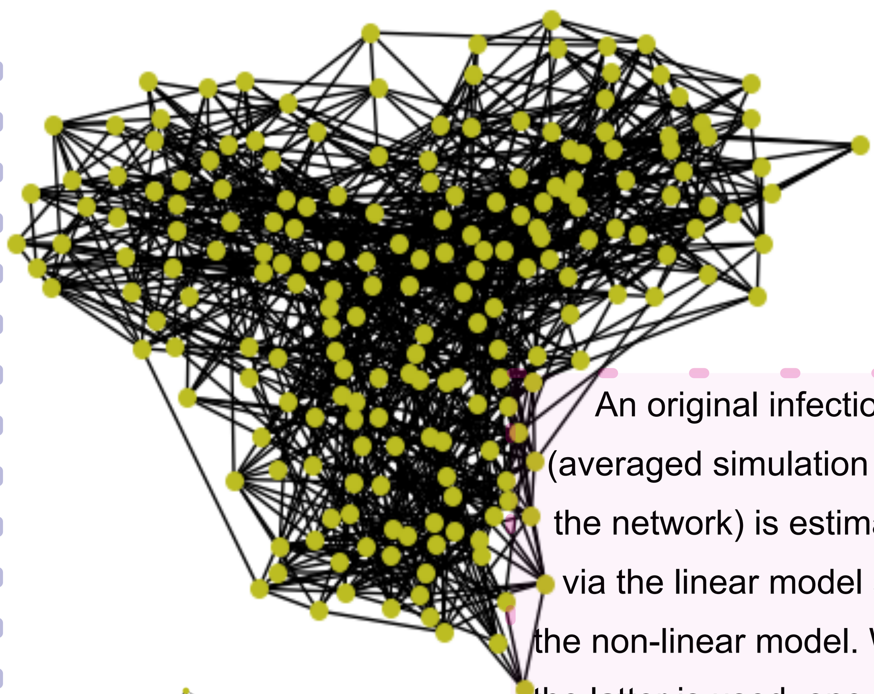
Step 4: Fit the parameters of the non-linear model

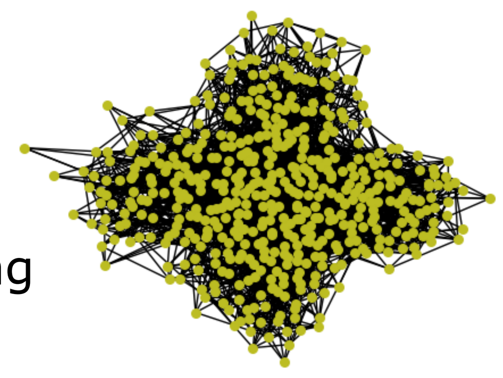
Non-linear (well-mixed) model

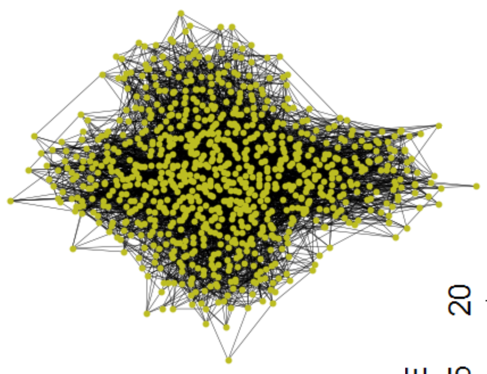
$$\frac{dS}{dt} = -\beta \times S(t)^p \times I(t)^q$$

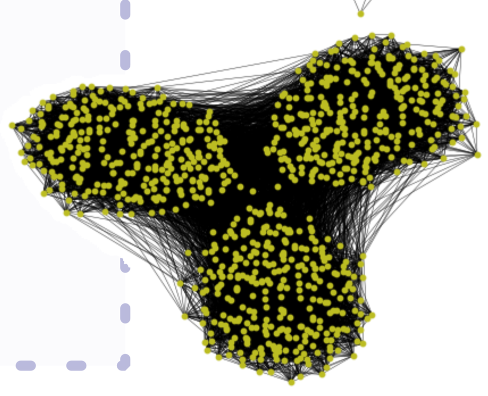
$$\frac{dI}{dt} = \beta \times S(t)^p \times I(t)^q - \gamma \times I(t)$$

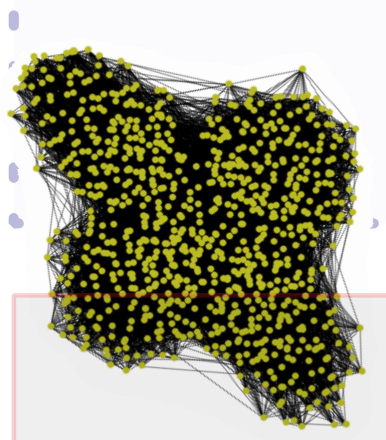
$$\frac{dR}{dt} = \gamma \times I(t)$$







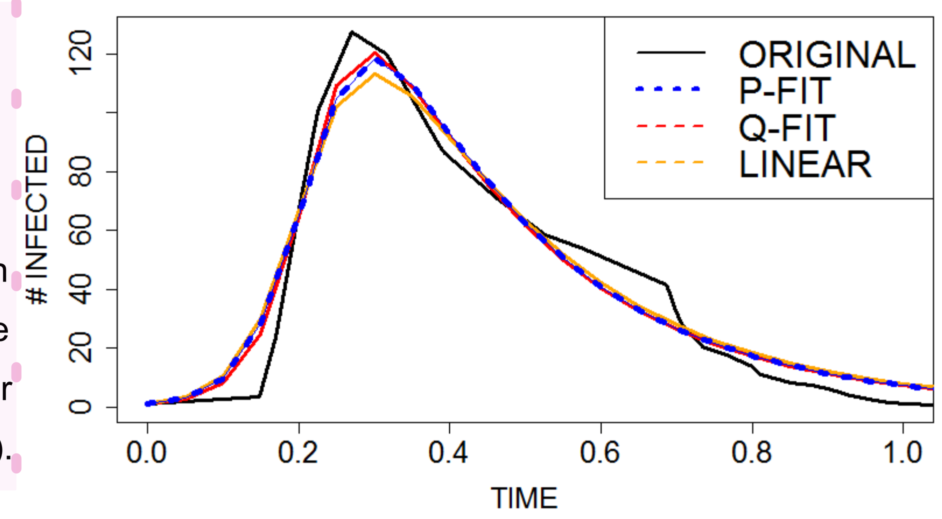


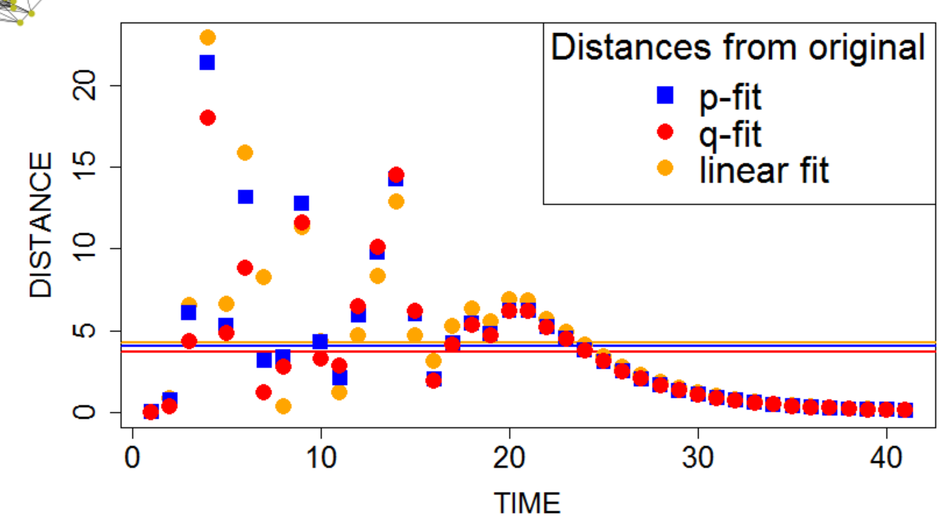


Results

Simulations and Fitting

An original infection (averaged simulation over the network) is estimated via the linear model and the non-linear model. When the latter is used, one of the constants is fit and the other is kept constant (equal to 1).





The figure on the left plots the absolute distances between fitted epidemic curves and original simulation results (for the plot above). Horizontal lines mark the average difference for each respective fit.

Discussion and future directions

We have simulated only networks with between two hundred and nine hundred nodes which were simulated such that to comprise three or four modules with medium degree of separation between them. Modularity was prefered as a feature of the simulated networks (instead of clustering) due to recent research results about the topic. However, there is no sufficient evidence yet to confirm that modularity is indeed an inherent characteristic of real networks of face-to-face interactions so this research is still limited to the synthetic scenarios until more conclusive evidence is found regarding the structure of real-world networks.
The results from simulations and consequent fitting for the networks and epidemic parameters which we have examined, show a slightly better fit of the non-linear model. This is assessed in the observed fit in peak times as well as overall least squares fit. Nevertheless, the average difference (between a fit and the simulated samples) can still exceed 10% (for the example printed here it is 11.6% excluding the times with very little infected numbers).
We have done optimisations in 3 dimensions only since we always had one of the non-linearity constants fixed. Further research can look into the interaction between the two non-linearity constants and whether one of them leads to evidently better results in most cases.

Main references

[1] L. Bioglio, M. Génois, C. Vestergaard, C. Poletto, A. Barrat, V. Colizza. (2015). Recalibrating disease parameters for increasing realism in modeling epidemics in closed settings. BMC Infectious Diseases. 16.
[2] I. Kiss, J. Miller, P. Simon. (2017). Mathematics of Epidemics on Networks. 10.1007/978-3-319-50806-1.
[3] P. Sah, L. Singh, A. Clauset, S. Bansal. (2014). Exploring community structure in biological networks with random graphs. BMC bioinformatics. 15. 220. 10.1186/1471-2105-15-220.

This project was completed thanks to the Laidlaw Scholars Programme at the University of Oxford.

Laidlaw Scholars
Undergraduate Research and Leadership Programme

UNIVERSITY OF OXFORD