Capturing the Effects of Transportation on the Spread of COVID-19 with a Networked SEIR Model

PI Philip E. Paré, Purdue University & PI Raphael Stern, University of Minnesota

RAPID, FY2020

Community Problem

The coronavirus disease (COVID-19) spreads from human contact. A key source of transmission is through travel between communities. With this in mind, our goal is to:

- Collect travel data for the community of interest before and during the COVID-19 pandemic
- Utilize the collected travel data to better understand viral spread of COVID-19
- Integrate travel behavior into traditional viral spread models to better model the evolution of the pandemic
- Use the developed models to design control strategies to mitigate the spread of COVID-19 by limiting travel in targeted interventions, while minimizing broader impact of the interventions

Community of interest: We consider counties in the Northeastern United States as our community of interest.

Figure 1: Community of Interest: Northeast US

Intellectual Merit

This project furthers the state of knowledge on epidemic spread modeling by providing the following scientific contributions:

- Incorporating transportation and travel data into traditional viral spread models
- Collection of a publicly available dataset for air travel during the COVID-19 pandemic that can be used by the broader research community
- Furthering data analysis/parameter estimation techniques to enable mitigation algorithm design

Project Activity: Including transportation networks in the SEIR model

Consider the susceptible-exposed-infected-recovered (SEIR) model

$$\begin{aligned} s_i^{k+1} &= s_i^k - h s_i^k \iota_i^k \\ e_i^{k+1} &= e_i^k + h s_i^k \iota_i^k - h\sigma_i e_i^k \\ p_i^{k+1} &= p_i^k + h(\sigma_i e_i^k - \gamma_i p_i^k) \\ r_i^{k+1} &= r_i^k + h\left(\gamma_i p_i^k\right) \end{aligned} \tag{1}$$

where i refers to the ith community, k is the time step, h is the sampling parameter, σ_i captures the rate at which the exposed become confirmed infected cases, γ_i is the recovery rate, and

$$\iota_i^k = \beta_i^e \sum_{j\in\mathcal{N}_i} a_{ij} e_j^k + \beta_i^p \sum_{j\in\mathcal{N}_i} a_{ij} p_j^k \tag{2}$$

where β_i^p and β_i^e are the infection parameters associated with the p_i^k and e_i^k states, respectively.

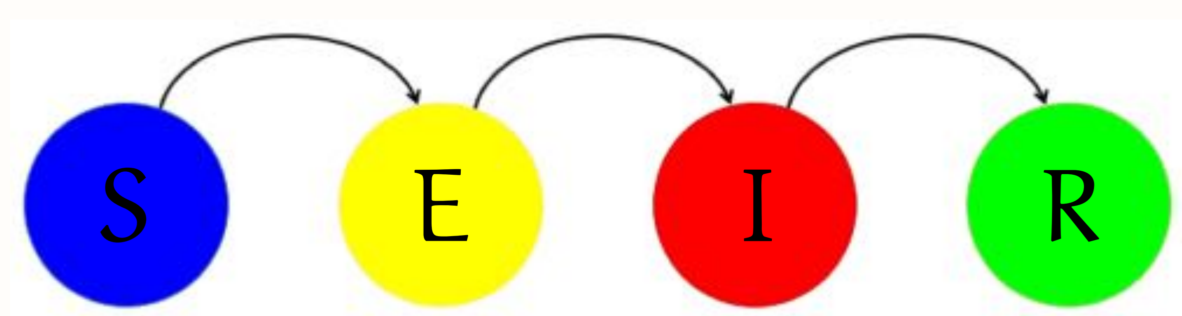


Figure 2: **SEIR Model**–Each node is either susceptible (S), exposed (E), infected (I), or recovered/removed (R).

For the transportation model we extend ι_i^k to be

$$\check{\iota}_i^k = \iota_i^k + \sum_{l\in\mathcal{L}} \left(\check{\beta}_i^{e,l} \sum_{j\in\mathcal{N}_i^l} \check{a}_{ij}^l e_j^k + \check{\beta}_i^{p,l} \sum_{j\in\mathcal{N}_i^l} \check{a}_{ij}^l p_j^k \right) \tag{3}$$

where $\mathcal{L}$ is the set of transportation networks, $\check{a}_{ij}^l$ represents the edge weights, and $\check{\beta}_i^{e,l}$ and $\check{\beta}_i^{p,l}$ are the corresponding infection rates for the lth transportation network.

Assumption 1: For all $i \in [n]$, we have $0 < h\gamma_i < 1$, $0 < h\sigma_i \le 1$, $0 \le h(\beta_i^e + \beta_i^p)\sum_{j\in\mathcal{N}_i} a_{ij} < 1$, and $\beta_i^e, \beta_i^p, a_{ij} \ge 0$, for all $j \in [n]$.

Lemma 1
Consider the model in (1)-(2) under Assumption 1. Suppose $s_i^0, e_i^0, p_i^0, r_i^0 \in [0,1]$, $s_i^0+e_i^0+p_i^0+r_i^0 = 1$ for all $i \in [n]$. Then, for all $k \ge 0$ and $i \in [n]$, $s_i^k, e_i^k, p_i^k, r_i^k \in [0,1]$ and $s_i^k+e_i^k+p_i^k+r_i^k = 1$.

Additionally, we prove linear convergence of the models as well as provide conditions for estimating parameters.

Project Activity: Simulation of viral spread via transportation networks across identified communities

We apply the networked SEIR model to the COVID-19 pandemic in the Northeast US, and incorporate flight mobility data.

A. Study Area

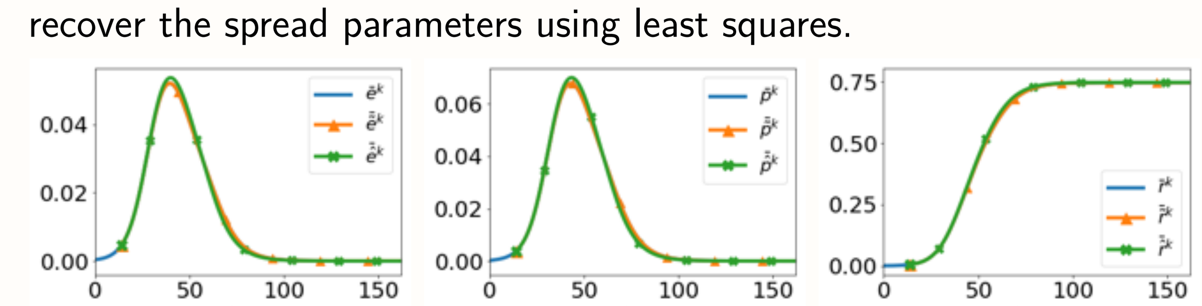
We consider the spread of COVID-19 through five states in the Northeastern US from March through August, 2020, and consider how the underlying air transportation network between the cities in the five-state region propagated the virus. Specifically, we obtain data for New York (NY), New Jersey (NJ), Massachusetts (MA), Rhode Islands (RI), and Connecticut (CT), and consider this five-state region as a closed system (i.e., no virus entering the system).

B. Transportation Data and Network Topologies

Three different types of connections are considered when modeling the network topology in (2) and (3): (*i*) county adjacency (a_{ij}^N) ; (*ii*) self-loops for spread within the county ($a_{ij}^S = I$); (*iii*) flights between airports ($a_{ij}^{F,k}$) to capture long-range links between non-adjacent counties [1]. It should be noted that the flight adjacency matrix $a_{ij}^{F,k}$ is time-varying.

C. Simulations

To simulate the states for the SEIR model we use (1) including the three adjacency matrices described in B. with homogeneous spread parameters. Moreover, we add measurement noise to evaluate the sensitivity of the estimation results and assume that the perturbation on e is greater than that on p and r since it is the most difficult of the three states to measure. In order to emulate the difficulty of measuring the states at the beginning of an outbreak, we start measuring from $k = 14$, and recover the spread parameters using least squares.



(a) Network average exposed state. (b) Network average infected state. (c) Network average removed state.

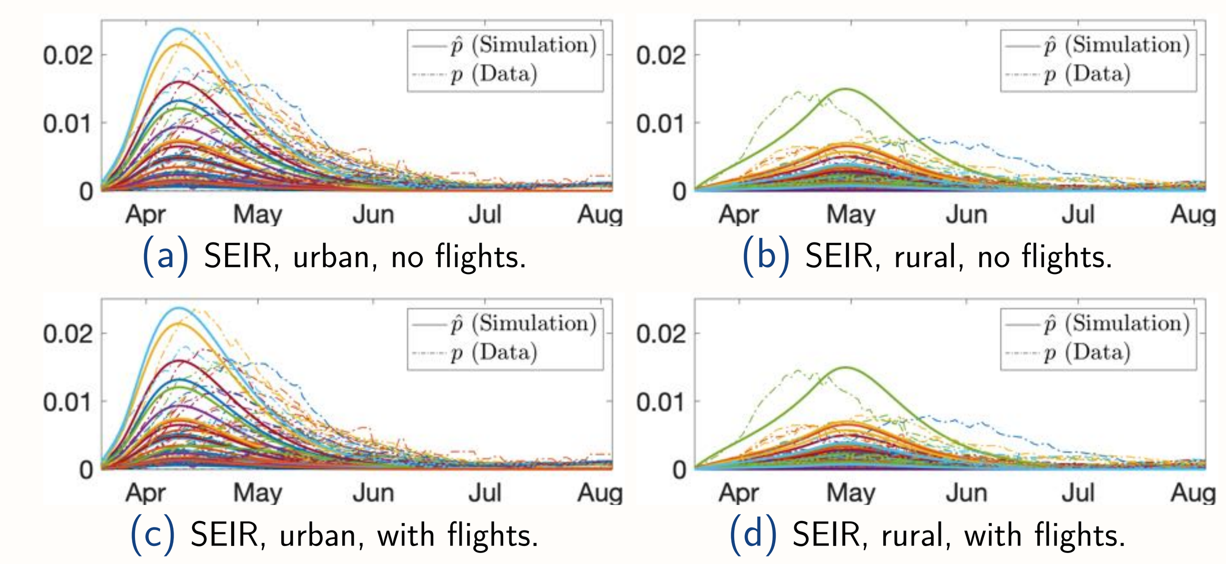
Figure 3: Simulation of a homogeneous SEIR system with three networks where the flights network ($a_{ij}^{F,k}$) is distorted by noise. Shows how well the recovered states captures the average state.

In Figure 3, we see that the averages of the recovered states are fairly close to the averages of the actual states even when accurate flight data is not available.

D. Real COVID-19 Spread Data

We use daily COVID-19 case numbers aggregated by Johns Hopkins University (JHU) [2]. Using this dataset, we are able to approximate e_i^k, p_i^k, and r_i^k in (1). Using (1), we estimate the parameter values by minimizing the least squares' error in the fit while constraining them to be non-negative using the cvx solver [3]. We look at the cases with and without taking transmissions resulting from inter-city travel via flights into account by using (2) and (3) or only (2).

The SEIR model performance is plotted in Figures 4a-4d.



(a) SEIR, urban, no flights. (b) SEIR, rural, no flights. (c) SEIR, urban, with flights. (d) SEIR, rural, with flights.

Figure 4: SEIR model simulations both with and without taking flight data into account. Each curve represents the proportion of infected population p_i^k in a particular county. Dashed lines represent real data, p, solid lines of the same color represent the corresponding simulation results, $\hat{p}$.

Broader Impacts

While the COVID-19 pandemic will end, future viruses may pose a risk for another pandemic. The work conducted as part of this grant will help prepare governments and health agencies to minimize damage from future pandemics. Specifically, the broader impacts of this work are:

- Provide strategies to control transportation network topology to limit viral spread
- Provide advances in viral spread modeling to help decision makers reduce the damage from future pandemics

By achieving these goals, the findings from this grant will help avert future pandemics.

References

[1] Bureau of Transportation Statistics (BTS), "Bureau of transportation statistics," https://www.bts.gov/, 2020, accessed: July 2020.

[2] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, 2020.

[3] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.