THE UNIVERSITY OF CHICAGO | Department of Public Health Sciences

ASA DATA SCIENCE IN ACTION IN RESPONSE TO THE OUTBREAK OF COVID-19

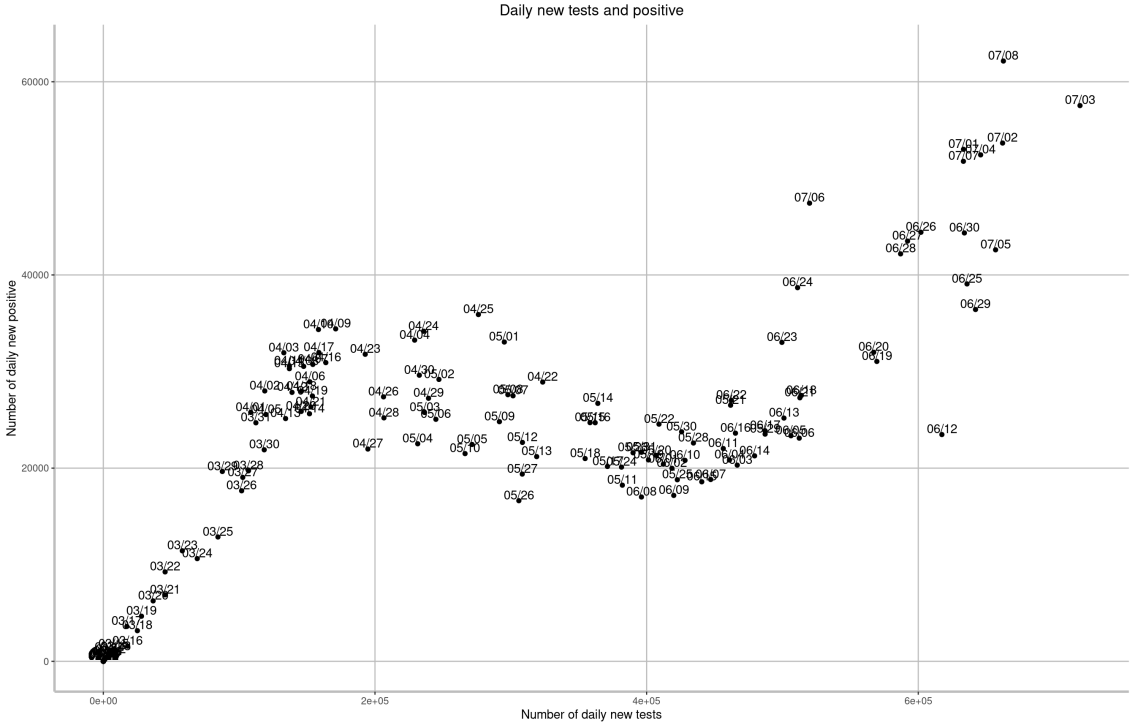
BaySIR: A Bayesian Semiparametric Compartmental Models for Modeling COVID-19 Epidemics

Tianjian Zhou, Ph.D.
Yuan Ji, Ph.D.
Department of Public Health Sciences
The University of Chicago

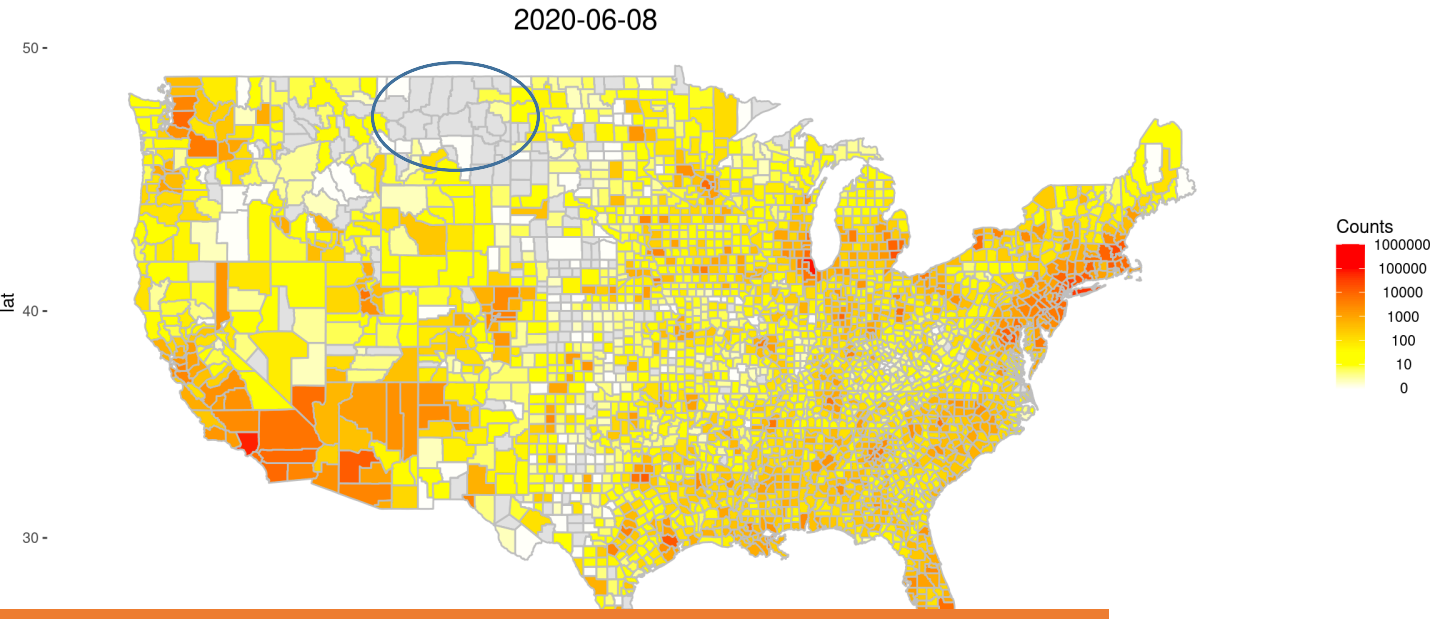
Paper: https://arxiv.org/pdf/2006.05581.pdf
Software: https://github.com/tianjianzhou/BaySIR

Part I

Visualization of COVID-19 Pandemic



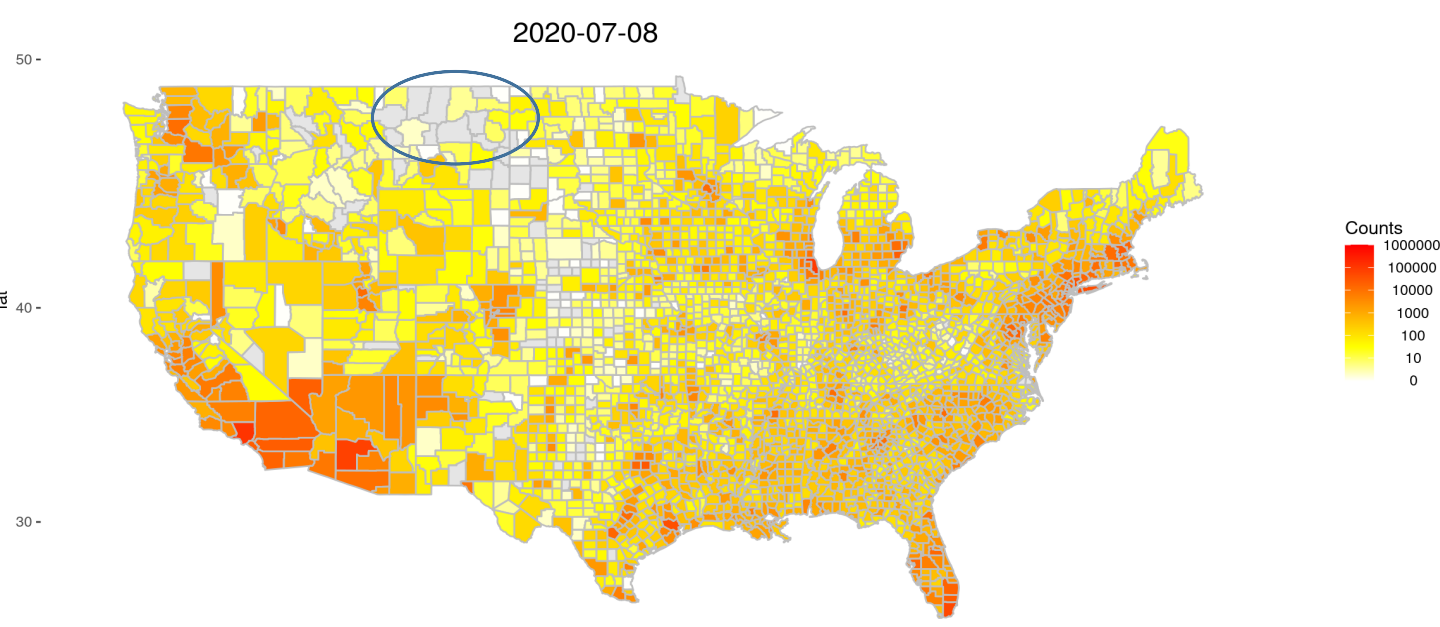
US Case and Test Numbers



US Epidemic

Cumulative cases at the county level of US

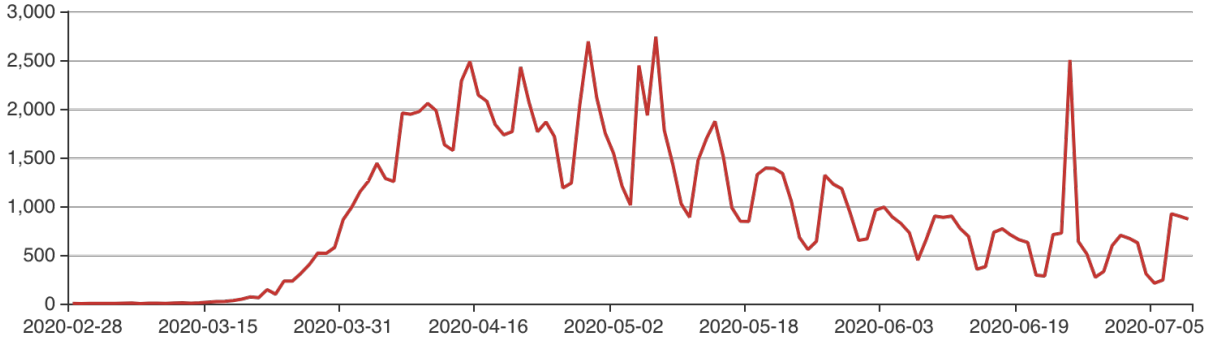
- **About 150 counties have not reported a case (grey blocks)**
- **Epicenters in major metropolitans**
- **We should all go to Montana!**



US Epidemic

Cumulative cases at the county level of US

- **About 150 counties have not reported a case (grey blocks)**
- **Epicenters in major metropolitans**
- **We should all go to Montana!**



US COVID-19 Death Numbers

Advanced Visualization of COVID-19 Pandemics

http://covid19.laiyaconsulting.com/

Part II

Statistical Modeling for COVID-19

Compartmental Models – Classic approaches

- A generative dynamic mathematical model mimicking how infectious diseases spread through the entire life circle
- During an epidemic, the entire population is divided into compartments, corresponding to different stages of a disease
- Individuals in the same compartment have the same characteristics
- Interested in: how individuals flow through compartments over time, i.e., the dynamics of the spread of the disease

Compartmental Models: SIR

SIR model: Susceptible, Infectious and Removed

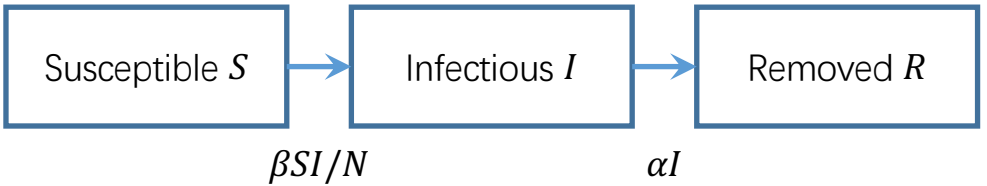
- **S**: Do not have the disease and can be infected
- **I**: Have the disease and can infect others
- **R**: Had the disease but then removed from the possibility of being infected again or spreading the disease (deceased or recovered)

- Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London, Series A*, 115(772):700–721.
- Weiss, H. (2013). The SIR model and the foundations of public health. *Materials Matematics*, 2013(3):1–17.

Deterministic Compartmental Models

- Characterize the flow of individuals through a set of differential equations
- Given initial values & parameters, deterministic trajectory
- **SIR model**:

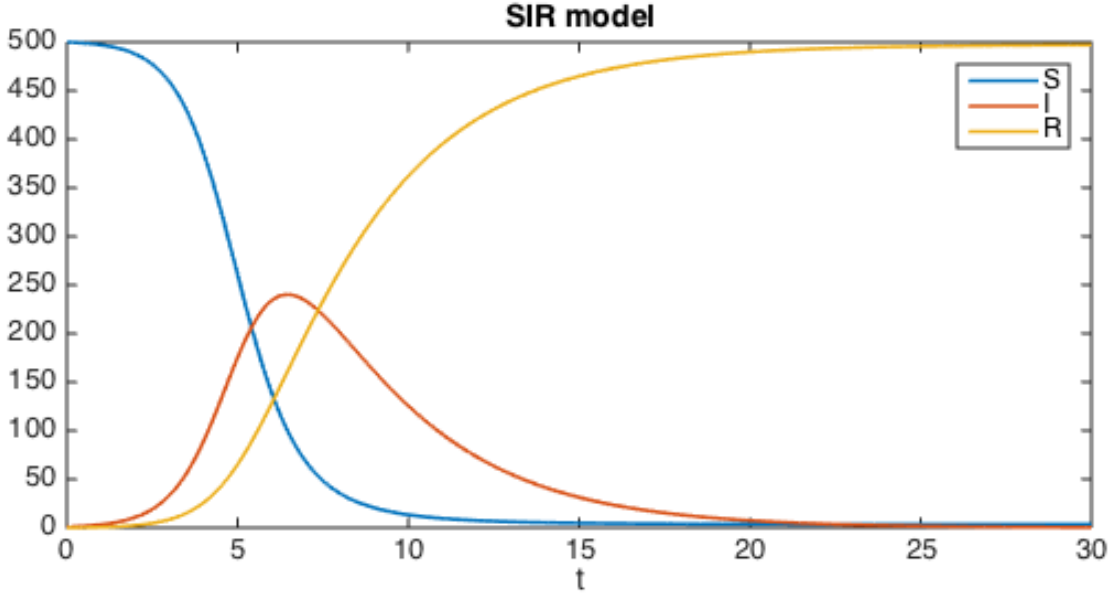
$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = -\frac{\beta}{N} S_t I_t, \qquad \frac{\mathrm{d}I_t}{\mathrm{d}t} = \frac{\beta}{N} S_t I_t - \alpha I_t, \qquad \frac{\mathrm{d}R_t}{\mathrm{d}t} = \alpha I_t$$



$\frac{\beta S}{N}$ → expected number of contacts of susceptible people per unit time

α → expected number of removals per person per unit time

A stylized example of the life circle for an epidemic



Deterministic Compartmental Models

SIR Model - Terminology:

- Disease transmission rate

$$\beta$$

- Removal rate

$$\alpha$$

- Infectious period

$$\alpha^{-1}$$

Deterministic Compartmental Models

SIR Model - Terminology:

- Basic reproduction number

$$\mathcal{R}_0 = \beta/\alpha$$

- Effective reproduction number

$$\mathcal{R}_e = (\beta S_0)/(\alpha N)$$

Theorem 2.1. *1. If $R_e \leq 1$, then $I(t)$ decreases monotonically to zero as $t \to \infty$.*

2. If $R_e > 1$, then $I(t)$ starts increasing, reaches its maximum, and then decreases to zero as $t \to \infty$. We call this scenario of increasing numbers of infected individuals an epidemic.

Discrete time approximation of SIR

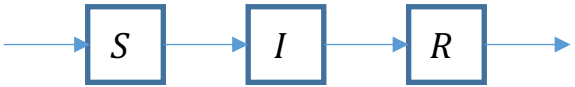
$$S_{t+1} = S_t - \frac{\beta}{N} S_t I_t,$$

$$I_{t+1} = I_t + \frac{\beta}{N} S_t I_t - \alpha I_t,$$

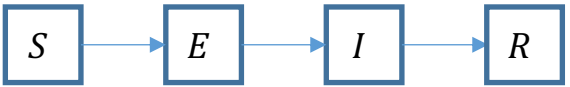
$$R_{t+1} = R_t + \alpha I_t$$

SIR Model - Extensions?

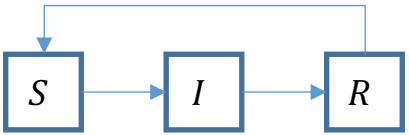
- Add vital dynamics (births and deaths) & demographics



- Add more compartments: Exposed but not infectious yet



- Allow more possible transitions across compartments:
  - Recovered and become susceptible again



Deterministic Compartmental Models

- Gumel, A. B., Ruan, S., Day, T., Watmough, J., Brauer, F., Van den Driessche, P., Gabrielson, D., Bowman, C., Alexander, M. E., Ardal, S., Wu, J., and Sahai, B. M. (2004). Modelling strategies for controlling SARS outbreaks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1554):2223–2232.

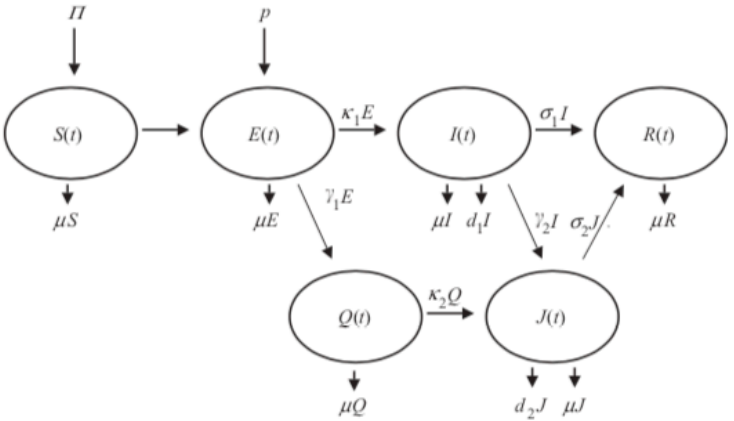


Figure 1. Schematic flow diagram for the SARS model (2.1)–(2.6). The model consists of six sub populations: susceptible $(S(t))$, asymptomatic $(E(t))$, quarantined $(Q(t))$, symptomatic $(I(t))$, isolated $(J(t))$ and recovered $(R(t))$ individuals in a population of $N(t) = S(t) + E(t) + Q(t) + I(t) + J(t) + R(t)$ individuals.

Stochastic Compartmental Models

- Deterministic models are simple, easy to analyze, and reflects the "law of large numbers".
- However, the spread of disease is naturally stochastic: disease transmission between two individuals is random rather than deterministic!

Stochastic Compartmental Models

Stochastic SIR Model:

(S_t, I_t, R_t) is a continuous time Markov chain

- Each infectious individual makes effective contacts (sufficient for disease transmission) with any specific individual in the population at times given by a Poisson process of rate β/N, and assume all Poisson processes are independent of each other. (Expected number of effective contacts made by each infectious individual is β per unit time.)

- The infectious period (length of an infectious individual being infectious before getting removed) follows an exponential distribution with mean α^{-1}.

Stochastic Compartmental Models

Stochastic SIR Model:

(S_t, I_t, R_t) is a continuous time Markov chain

- Within $(t, t+\delta]$, (δ small) each susceptible individual independently avoids infection with probability $\exp(-\beta i\delta/N)$;
- Each infectious individual independently avoids removal with probability $\exp(-\alpha\delta)$
- Then,

$$Pr\{(S_{t+\delta}, I_{t+\delta}) = (s-1, i+1) | (S_t, I_t) = (s,i)\} = \frac{\beta si\delta}{N} + o(\delta)$$

$$Pr\{(S_{t+\delta}, I_{t+\delta}) = (s, i-1) | (S_t, I_t) = (s,i)\} = \alpha i\delta + o(\delta)$$

Stochastic Compartmental Models

Stochastic SIR Model:

(S_t, I_t, R_t) is a continuous time Markov chain

- Usually need individual-level data (time of infection, removal, etc. for each individual, some may be latent) to make inference

Stochastic Compartmental Models

Stochastic SIR Model – Discrete-time Approximation:

(S_t, I_t, R_t) is a discrete-time Markov chain

- Within $(t, t+1]$, each susceptible individual independently avoids infection with probability $\exp(-\beta i/N)$;
- Each infectious individual independently avoids removal with probability $\exp(-\alpha)$

Stochastic Compartmental Models

Stochastic SIR Model – Discrete-time Approximation:

(S_t, I_t, R_t) is a discrete-time Markov chain

- Then,

$$S_{t+1} = S_t - A_t$$
$$I_{t+1} = I_t + A_t - B_t$$
$$R_{t+1} = R_t + B_t$$

where

$$A_t \sim \text{Bin}\left(S_t, 1 - e^{-\beta I_t/N}\right)$$
$$B_t \sim \text{Bin}\left(I_t, 1 - e^{-\alpha}\right)$$

Stochastic Compartmental Models

Stochastic SIR Model – Discrete-time Approximation:

- A similar model was used in

  Lekone, P. E. and Finkenstädt, B. F. (2006). Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. Biometrics, 62(4):1170– 1177.

Biometrics 62, 1170–1177
December 2006

DOI: 10.1111/j.1541-0420.2006.00609.x

Statistical Inference in a Stochastic Epidemic SEIR Model with Control Intervention: Ebola as a Case Study

Phenyo E. Lekone* and Bärbel F. Finkenstädt
Department of Statistics, University of Warwick, Coventry CV4 7AL, U.K.
*email: lekonepe@mopipi.ub.bw

Given initial conditions $S(0) = s_0$, $E(0) = e_0$, $I(0) = a$, and the population size N, the discretized stochastic SEIR model is specified by

$$S(t+h) = S(t) - B(t), \tag{1}$$

$$E(t+h) = E(t) + B(t) - C(t), \tag{2}$$

$$I(t+h) = I(t) + C(t) - D(t), \tag{3}$$

$$S(t) + E(t) + I(t) + R(t) = N, \tag{4}$$

where

$$B(t) \sim \text{Bin}(S(t), P(t)), \quad C(t) \sim \text{Bin}(E(t), p_C),$$
$$D(t) \sim \text{Bin}(I(t), p_R) \tag{5}$$

are random variables with binomial $\text{Bin}(n,\ p)$ distributions with probabilities:

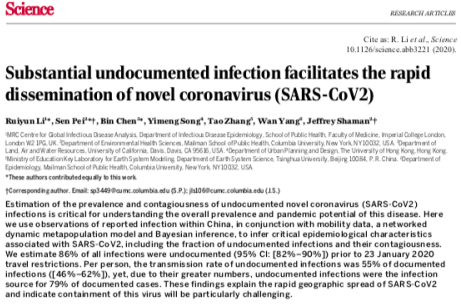
$$P(t) = 1 - \exp\left[-\frac{\beta(t)}{N} h I(t)\right], \quad p_C = 1 - \exp(-\varrho h),$$
$$p_R = 1 - \exp(-\gamma h). \tag{6}$$

Stochastic Compartmental Models

Stochastic SIR Model – Discrete-time Approximation:

(S_t, I_t, R_t) is a discrete-time Markov chain

$$A_t \sim \text{Bin}\left(S_t, 1 - e^{-\beta I_t/N}\right) \approx \text{Pois}(\beta S_t I_t / N)$$
$$B_t \sim \text{Bin}(I_t, 1 - e^{-\alpha}) \approx \text{Pois}(\alpha I_t)$$



Science

RESEARCH ARTICLES

Cite as: R. Li *et al.*, *Science* 10.1126/science.abb3221 (2020).

Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)

Ruiyun Li[1*], Sen Pei[2*†], Bin Chen[3*], Yimeng Song[4], Tao Zhang[5], Wan Yang[6], Jeffrey Shaman[2†]

[1]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Faculty of Medicine, Imperial College London, London W2 1PG, UK. [2]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA. [3]Department of Land, Air and Water Resources, University of California, Davis, Davis, CA 95616, USA. [4]Department of Urban Planning and Design, The University of Hong Kong, Hong Kong. [5]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, P. R. China. [6]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA.

*These authors contributed equally to this work.

†Corresponding author. Email: sp3449@cumc.columbia.edu (S.P.); jls2168@cumc.columbia.edu (J.S.)

Estimation of the prevalence and contagiousness of undocumented novel coronavirus (SARS-CoV2) infections is critical for understanding the overall prevalence and pandemic potential of this disease. Here we use observations of reported infection within China, in conjunction with mobility data, a networked dynamic metapopulation model and Bayesian inference, to infer critical epidemiological characteristics associated with SARS-CoV2, including the fraction of undocumented infections and their contagiousness. We estimate 86% of all infections were undocumented (95% CI: [82%–90%]) prior to 23 January 2020 travel restrictions. Per person, the transmission rate of undocumented infections was 55% of documented infections ([46%–62%]), yet, due to their greater numbers, undocumented infections were the infection source for 79% of documented cases. These findings explain the rapid geographic spread of SARS-CoV2 and indicate containment of this virus will be particularly challenging.

The core model structure (Equations 1-5) was integrated stochastically using a 4th order Runge-Kutta (RK4) scheme. Specifically, for each step of the RK4 scheme, each unique term on the righthand side (rhs) of Equations 1-4 was determined using a random sample from a Poisson distribution, i.e.

$$\text{U1} = \text{Pois}\left(\frac{\beta S_i I_i^r}{N_i}\right)$$
$$\text{U2} = \text{Pois}\left(\frac{\mu\beta S_i I_i^u}{N_i}\right)$$
$$\text{U3} = \text{Pois}\left(\theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r}\right)$$

Stochastic Compartmental Models

Bayesian Analysis (2009) 4, Number 4, pp. 465–496

Bayesian Analysis for Emerging Infectious Diseases

Chris P Jewell*, Theodore Kypraios†, Peter Neal‡ and Gareth O. Roberts§

Abstract. Infectious diseases both within human and animal populations often pose serious health and socioeconomic risks. From a statistical perspective, their prediction is complicated by the fact that no two epidemics are identical due to changing contact habits, mutations of infectious agents, and changing human and animal behaviour in response to the presence of an epidemic. Thus model parameters governing infectious mechanisms will typically be unknown. On the other hand, epidemic control strategies need to be decided rapidly as data accumulate. In this paper we present a fully Bayesian methodology for performing inference and online prediction for epidemics in structured populations. Key features of our approach are the development of an MCMC- (and adaptive MCMC-) based methodology for parameter estimation, epidemic prediction, and online assessment of risk from currently unobserved infections. We illustrate our methods using two complementary studies: an analysis of the 2001 UK Foot and Mouth epidemic, and modelling the potential risk from a possible future Avian Influenza epidemic to the UK Poultry industry.

Stochastic Compartmental Models

Statistical Science
2018, Vol. 33, No. 1, 44–56
https://doi.org/10.1214/17-STS617

Bayesian Nonparametrics for Stochastic Epidemic Models

Theodore Kypraios and Philip D. O'Neill

Abstract. The vast majority of models for the spread of communicable diseases are parametric in nature and involve underlying assumptions about how the disease spreads through a population. In this article, we consider the use of Bayesian nonparametric approaches to analysing data from disease outbreaks. Specifically we focus on methods for estimating the infection process in simple models under the assumption that this process has an explicit time-dependence.

Key words and phrases: Bayesian nonparametrics, epidemic model, Gaussian process.

State-space Compartmental Models

- Can be seen as statistical models built on deterministic compartmental models
- Deterministic model with randomness to account for measurement error & process error (but in a way different from stochastic compartmental models)
- Recently becoming popular in statistics literature

State-space Compartmental Models: Example 1

Tracking Epidemics With Google Flu Trends Data and a State-Space SEIR Model

Vanja DUKIC, Hedibert F. LOPES, and Nicholas G. POLSON

In this article, we use Google Flu Trends data together with a sequential surveillance model based on state-space methodology to track the evolution of an epidemic process over time. We embed a classical mathematical epidemiology model [a susceptible-exposed-infected-recovered (SEIR) model] within the state-space framework, thereby extending the SEIR dynamics to allow changes through time. The implementation of this model is based on a particle filtering algorithm, which learns about the epidemic process sequentially through time and provides updated estimated odds of a pandemic with each new surveillance data point. We show how our approach, in combination with sequential Bayes factors, can serve as an online diagnostic tool for influenza pandemic. We take a close look at the Google Flu Trends data describing the spread of flu in the United States during 2003–2009 and in nine separate U.S. states chosen to represent a wide range of health care and emergency system strengths and weaknesses. This article has online supplementary materials.

KEY WORDS: Flu; Google correlate; Google insights; Google searches; Google trends; H1N1; Infectious Diseases; Influenza; IP surveillance; Nowcasting; Online surveillance; Particle filtering.

State-space Compartmental Models: Example 1

Due to the nature of ILI surveillance data, our observations will consist only of noisily observed weekly counts of ILI visits, $\tilde{I}_t$, which can be thought of as a proxy to the true fraction of infected population, I_t, in each week-long time period $(t-1, t]$. Instead of working directly with $\tilde{I}_t$ however, we will model the observed growth rate of the infectious population, $y_t = (\tilde{I}_t - \tilde{I}_{t-1})/\tilde{I}_{t-1}$. This leads to the following state-space model for the growth rate:

$$y_t = g_t + \varepsilon_t^y \qquad \varepsilon_t^y \sim N(0, \sigma_y^2) \tag{3}$$

$$g_t = -\gamma + \alpha \frac{E_{t-1}}{I_{t-1}} + \varepsilon_t^g \qquad \varepsilon_t^g \sim N(0, \sigma_g^2). \tag{4}$$

We will refer to Equation (3) as the "observation equation" and Equation (4) as the "evolution equation" for the growth rate. The mean component of Equation (4) is derived from the deterministic evolution of I_{t-1} based on the discretized SEIR model (2) above, with the true number of infections I_t related to g_t via $I_t = (1 + g_t)I_{t-1}$. With the infectious state I_t modeled

Journal of the American Statistical Association, December 2012

expressed as follows:

$$\begin{aligned} S_t &= S_{t-1} - \beta S_{t-1} I_{t-1}/N \\ E_t &= (1-\alpha)E_{t-1} + \beta S_{t-1} I_{t-1}/N \\ I_t &= (1-\gamma)I_{t-1} + \alpha E_{t-1} \\ R_t &= R_{t-1} + \gamma I_{t-1}, \end{aligned} \tag{2}$$

State-space Compartmental Models: Example 2

The Annals of Applied Statistics
2017, Vol. 11, No. 1, 202–224
DOI: 10.1214/16-AOAS1000

FORECASTING SEASONAL INFLUENZA WITH A STATE-SPACE SIR MODEL[1]

BY DAVE OSTHUS[*,†], KYLE S. HICKMANN[*], PETRUŢA C. CARAGEA[†], DAVE HIGDON[‡,*] AND SARA Y. DEL VALLE[*]

Los Alamos National Laboratory[*], *Iowa State University*[†], *and Virginia Tech*[‡]

State-space Compartmental Models: Example 2

error) simultaneously. The state-space model we propose, henceforth referred to as the Dirichlet-Beta state-space model (DBSSM), is defined as

$$y_t|\theta_t, \phi \sim Beta(\lambda\theta_t^I, \lambda(1-\theta_t^I)), \tag{4.1a}$$

$$\theta_t|\theta_{t-1}, \phi \sim Dirichlet(\kappa f(\theta_{t-1}, \beta, \gamma)), \tag{4.1b}$$

where y_t is ILI+ at time $t = 1, 2, \ldots, T$, $\theta_t = (\theta_t^S, \theta_t^I, \theta_t^R)'$ represents the true but unobservable susceptible, infectious, and recovered proportions of the population, respectively, $\phi = \{\theta_0, \gamma, \beta, \kappa, \lambda\}$ where $\gamma > 0$ is the recovery rate, $\beta > 0$ is the disease transmission rate, $\kappa > 0$ controls the variance of equation (4.1b), $\lambda > 0$ controls the variance of equation (4.1a), and $f(\theta_{t-1}, \beta, \gamma) \in \mathbb{R}^3$ is defined in detail in the following paragraph. Furthermore, $\theta_t^S + \theta_t^I + \theta_t^R = 1$ and $\theta_t^S, \theta_t^I, \theta_t^R > 0$ for all t. The DBSSM assumes $\theta_{0:T} = (\theta_0, \theta_1, \ldots, \theta_T)$ is a first-order Markov chain (i.e., $[\theta_t|\theta_{0:(t-1)}] = [\theta_t|\theta_{t-1}]$ for all t) and for all $t \neq s$, y_t is independent of y_s, given θ_t.

We define $f(\cdot)$ as the solution to equation (4.2),

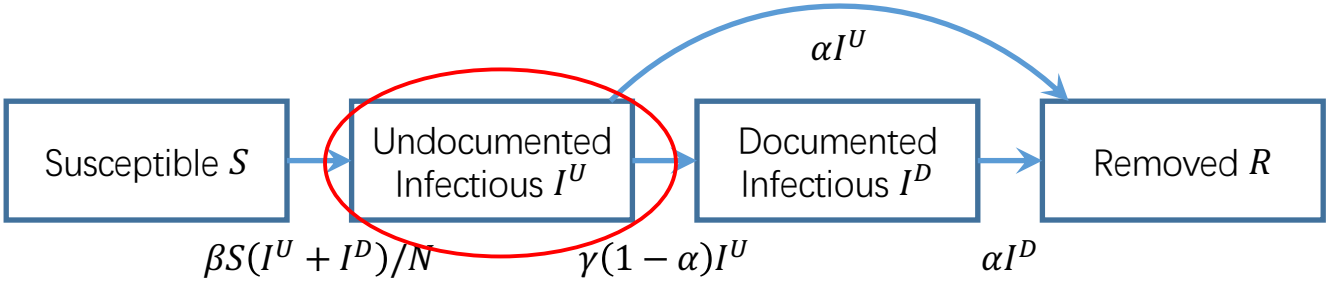
$$\frac{d\theta^S}{dt} = -\beta\theta^S\theta^I, \qquad \frac{d\theta^I}{dt} = \beta\theta^S\theta^I - \gamma\theta^I, \qquad \frac{d\theta^R}{dt} = \gamma\theta^I, \tag{4.2}$$

starting the ODE at θ_{t-1}. The solution to equation (4.2) is not known explicitly,

Our state-space model for COVID-19:

Bayesian Hierarchical Semiparametric modeling

Model for COVID-19: Compartments



Model for COVID-19: Epidemiological Parameters

Some parameters vary over time

- Disease transmission rate: $\beta \rightarrow \beta_t$
- Diagnosis rate: $\gamma \rightarrow \gamma_t$
- Removal rate: α (not time-varying)

Model for COVID-19: Epidemic Process

$$
\begin{aligned}
S_t &= S_{t-1} - \beta_{t-1} S_{t-1} (I^U_{t-1} + I^D_{t-1})/N, \\
I^U_t &= (1-\alpha) I^U_{t-1} + \beta_{t-1} S_{t-1} (I^U_{t-1} + I^D_{t-1})/N - B_{t-1}, \\
I^D_t &= (1-\alpha) I^D_{t-1} + B_{t-1}, \\
R_t &= R_{t-1} + \alpha (I^U_{t-1} + I^D_{t-1}),
\end{aligned}
$$

- B_{t-1}: number of new confirmed cases between day $t-1$ and t, our observed data
- Epidemic process is fully determined by the initial values, parameters, and observations

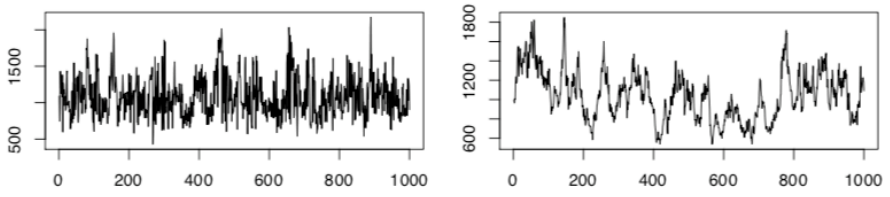
Model for COVID-19: Observed Data

- $B_t = \gamma_t(1-\alpha)I_t^U$
- $\tilde{\gamma}_t = \log(-\log(1-\gamma_t))$
- $\tilde{\gamma}_t \sim N\left(\boldsymbol{y}_t^T\boldsymbol{\eta}, \sigma_\gamma^2\right)$
  For simplicity, $\boldsymbol{y}_t = 1$
  Prior: $\boldsymbol{\eta}$ and σ_γ^2 follow normal-inverse-gamma

Model for COVID-19: Prior

- I_0^D known (set at 100) , $I_0^U / I_0^D \sim$ Gamma,
- $R_0 = 0$, $S_0 = N - I_0^U - I_0^D - R_0$
- $\tilde{\beta}_t = \log(\beta_t)$, $\tilde{\beta}_t \sim$ Gaussian Process
- $\alpha^{-1} \sim$ Gamma

Model for COVID-19: Parallel Tempering



(a) I_0^U using PT, $z_G = 0.38$

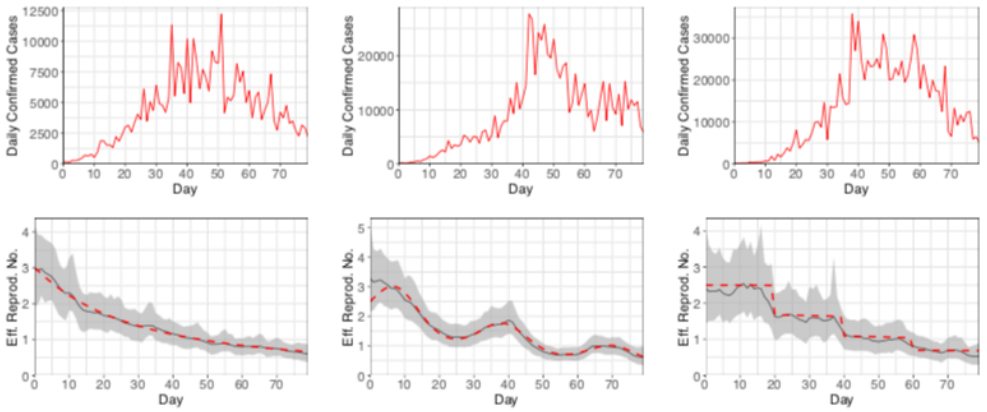
(b) I_0^U not using PT, $z_G = 4.46$

(c) η using PT, $z_G = -0.30$

(d) η not using PT, $z_G = -5.90$

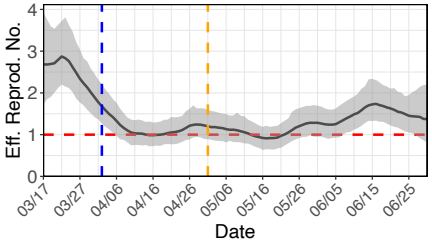
Figure 2: Markov chains for I_0^U and η using (a, c) or not using (b, d) parallel tempering. The posterior correlation of I_0^U and η is -0.82. The value z_G refers to Geweke's z-score for convergence diagnostic. All chains are based on 30,000 iterations (discarding first 10,000 iterations as burn-in and keeping 1 draw every 20 iterations).

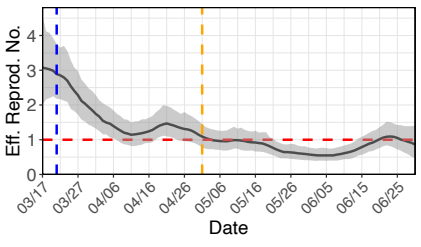
Model for COVID-19: Simulation

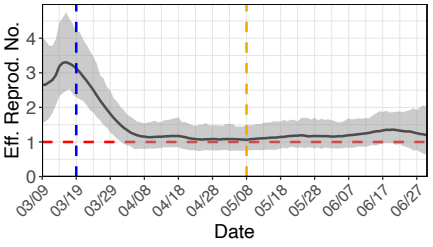


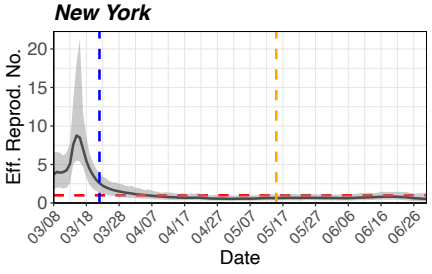
(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

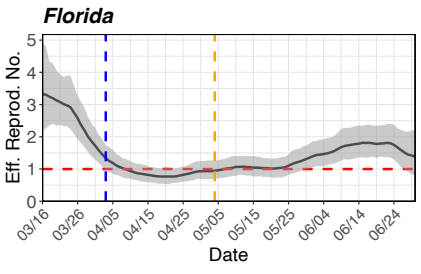
Figure 3: The upper panel shows the simulated daily confirmed cases for the three scenarios. The lower panel shows the estimated time-varying effective reproduction numbers (solid black line), 95% credible intervals (grey band), and simulation truth (dashed red line) for the three scenarios.

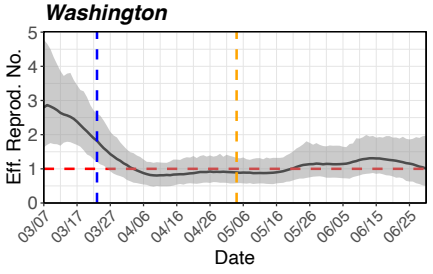












Results: Effective Reproduction Number

Results: Test of Fit

- Johnson (2004): Bayesian chi-square test for goodness-of-fit

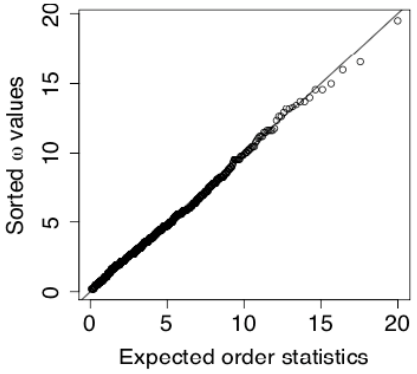
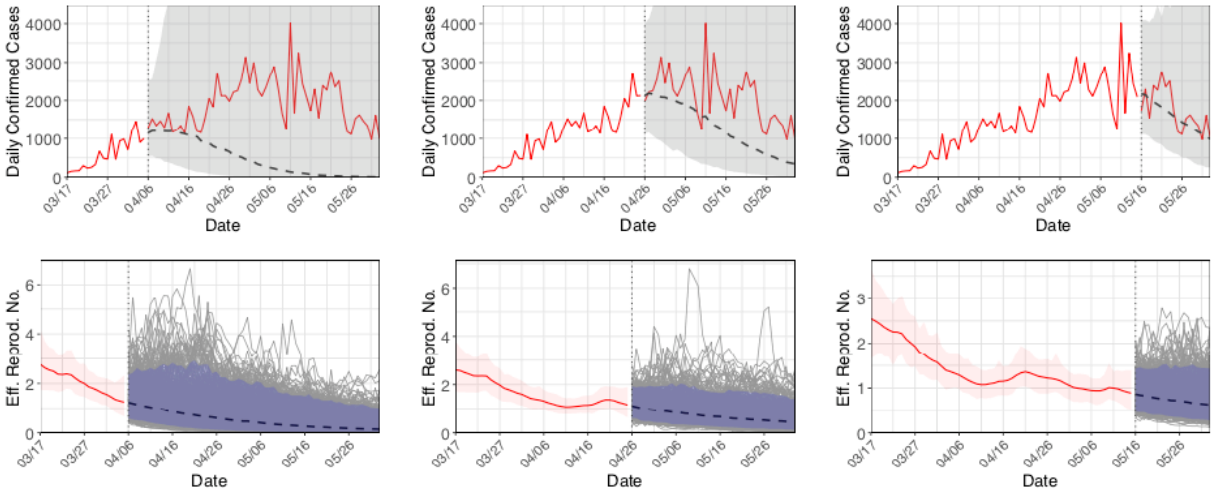


Figure 5: Quantile-quantile plot of posterior samples of the test statistic ω against expected order statistics from a χ^2_4 distribution for the Bayesian χ^2 test

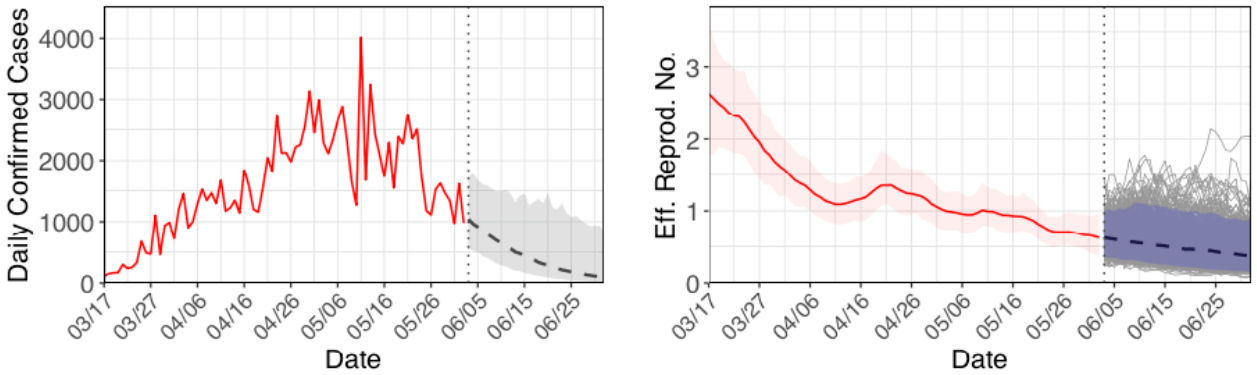
Results: Within-sample Prediction



(a) 20-day training data (b) 40-day training data (c) 60-day training data

Figure 6: Within-sample forecasts for Illinois using 20-day, 40-day or 60-day training data.

Results: Out-of-sample Prediction



(a) Projected daily confirmed cases

(b) Projected $\mathcal{R}_e(t)$

Figure 7: Out-of-sample forecasts for Illinois in the next 30 days. (a) Observed daily

Future Works

- State-space compartmental models built on deterministic models are quite flexible & easier to make inference compared to stochastic models

- Future:

  - Model epidemiological parameters (regression, time-dependent, hierarchical, flexible form using Bayesian nonparametrics like GP)

  - Recurring and seasonal trends

  - Borrow information from other countries and states

References on statistical models

- Aguilar, J. B., Faust, J. S., Westafer, L. M., and Gutierrez, J. B. (2020). Investigating the impact of asymptomatic carriers on COVID-19 transmission. medRxiv
- Chen, X. and Qiu, Z. (2020). Scenario analysis of non-pharmaceutical interventions on global COVID-19 transmissions. arXiv preprint arXiv:2004.04529
- Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Coupland, H., Mellan, T. A., Zhu, H., Berah, T., Eaton, J. W., Guzman, P. N., et al. (2020). Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Technical report.
- Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nature Medicine. Forthcoming.
- Gomez, J., Prieto, J., Leon, E., and Rodriguez, A. (2020). INFEKTA: A general agentbased model for transmission of infectious diseases: Studying the COVID-19 propagation in Bogot´a-Colombia. medRxiv
- Gu, J., Yan, H., Huang, Y., Zhu, Y., Sun, H., Zhang, X., Wang, Y., Qiu, Y., and Chen, S. (2020). Better strategies for containing COVID-19 epidemics—a study of 25 countries via an extended varying coefficient SEIR model. medRxiv.
- He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine, 26(5):672–675
- IHME COVID-19 health service utilization forecasting team and Murray, C. J. (2020). Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. MedRxiv.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020b). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science, 368(6490):489–493.
- Wang, L., Wang, G., Gao, L., Li, X., Yu, S., Kim, M., Wang, Y., and Gu, Z. (2020a). Spatiotemporal dynamics, nowcasting and forecasting of COVID-19 in the United States. arXiv preprint arXiv:2004.14103.

References on statistical models

- Wang, L., Zhou, Y., He, J., Zhu, B., Wang, F., Tang, L., Eisenberg, M. C., and Song, P. X. (2020b). An epidemiological forecast model and software assessing interventions on COVID-19 epidemic in China. Journal of Data Science. Forthcoming
- Woody, S., Tec, M. G., Dahan, M., Gaither, K., Lachmann, M., Fox, S., Meyers, L. A., and Scott, J. G. (2020). Projections for first-wave COVID-19 deaths across the US using social-distancing measures derived from mobile phones. medRxiv.
- Wu, J. T., Leung, K., and Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. The Lancet, 395(10225):689–697
- Zhang, Y., You, C., Cai, Z., Sun, J., Hu, W., and Zhou, X.-H. (2020). Prediction of the COVID-19 outbreak based on a realistic stochastic model. medRxiv.
- Donnat, C and Holmes, S (2020). Modeling the Heterogeneity in COVID-19's Reproductive Number and its Impact on Predictive Scenarios. ArXiv.
- Roy, A and Karmakar, S (2020). Bayesian semiparametric time varying model for count data to study the spread of the COVID-19 cases. ArXiv.

- Many many other preprints on ArXiv now.