COLUMBIA UNIVERSITY | MAILMAN SCHOOL of PUBLIC HEALTH

ENVIRONMENTAL HEALTH SCIENCES

Transmission Dynamics of Influenza and SARS-CoV-2: Environmental Determinants, Inference and Forecast

Jeffrey Shaman

May 18, 2020

Funders

NIH (NIGMS)/NSF (DMS) joint initiative to support research at the interface of the biological and mathematical sciences

MIDAS

Models of Infectious Disease Agent Study

Funded by the National Institutes of Health

DARPA

DEFENSE THREAT REDUCTION AGENCY UNITED STATES OF AMERICA

BIOMEDICAL ADVANCED RESEARCH AND DEVELOPMENT AUTHORITY BARDA

NSF

Collaborators

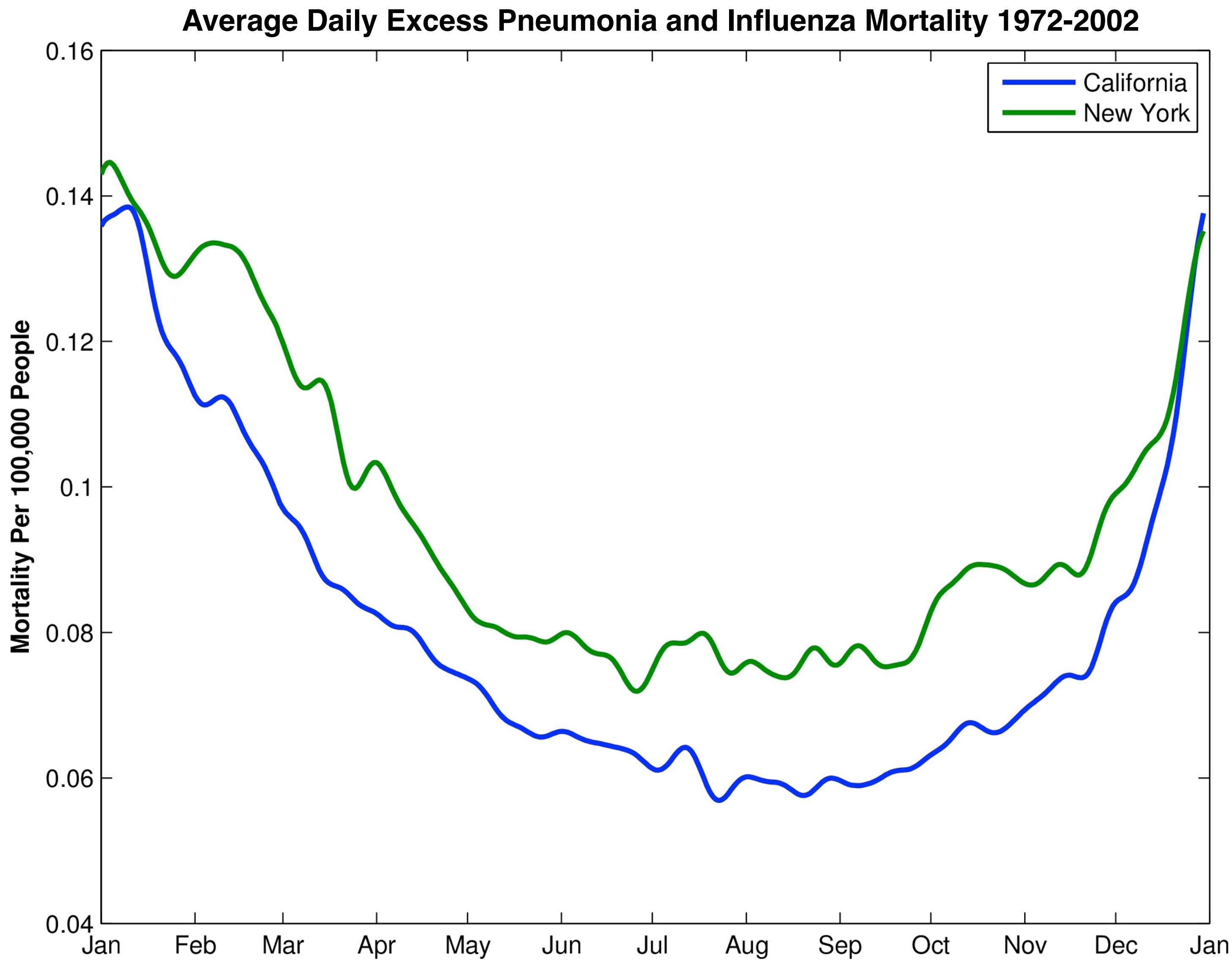
Columbia/Mailman

Wan Yang
Sasikiran Kandula
Teresa Yamana
Sen Pei
Marta Galanti
Nicholas Defelice
Haruka Morita
Alexandra Heaney (UC Berkeley)
Ruiyin Li (Imperial)
James Tamerius (U. Iowa)

Other

Alicia Karspeck
Marc Lipsitch (Harvard)
Cecile Viboud (NIH/Fogarty)
Virginia Pitzer (Yale)
Bryan Grenfell (Harvard)
Bin Chen (UC Davis)
Yimeng Song (U Hong Kong)
Tao Zhang (Tsinghua)

Seasonality of Influenza



Modes of Influenza Transmission

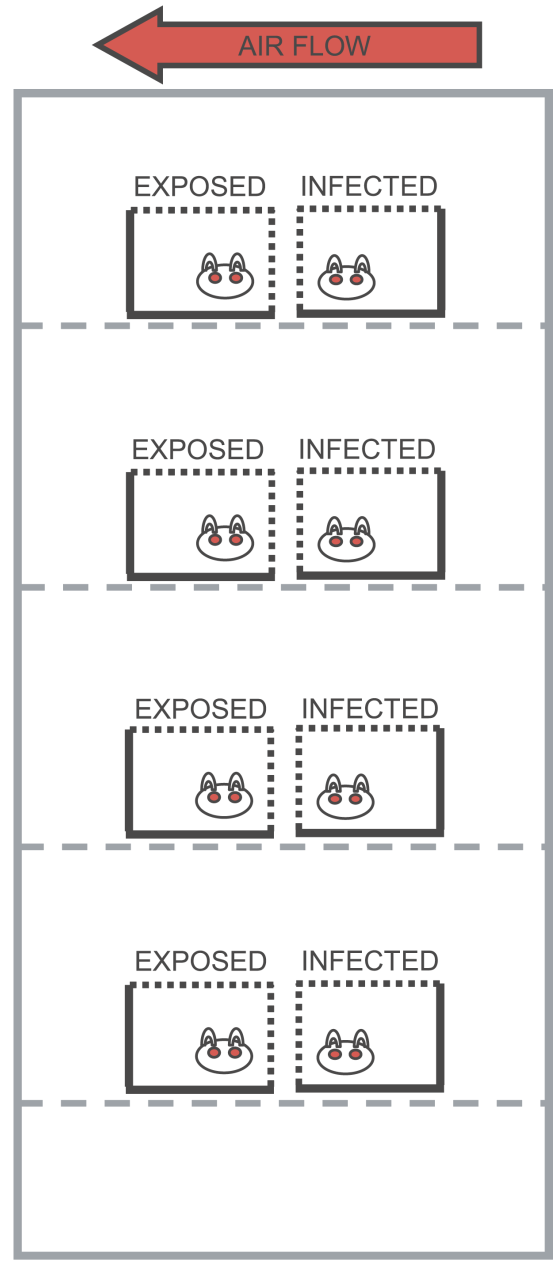
- Direct Contact
- Indirect Contact

(Fomites)

- Droplet
- Airborne

Guinea Pig Experiment

- Ran this chamber experiment 20 times at different temperature and relative humidity (RH) conditions
- Found marginally statistically significant effects
- Colder temperatures and lower RH favored transmission



Lowen et al., 2007

Measures of Humidity

Relative Humidity (RH) is not a well-constrained variable

RH varies as both a function of air water vapor content and temperature

$$RH = \frac{e}{e_s(T)} \times 100\%$$

e is the vapor pressure - a measure of the actual water vapor content of the air

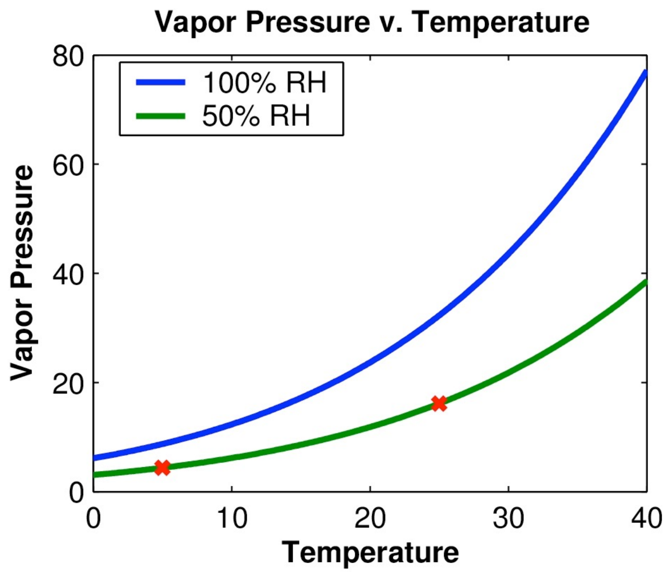
$e_s(T)$ is the saturation vapor pressure, the point at which rates of condensation and evaporation are equivalent. This quantity varies strongly as a function of temperature.

Measures of Humidity

Why explore Relative Humidity (RH)?

Saturation vapor pressure (100% RH) rises exponentially with increasing temperature

Why not use a measure of absolute humidity?



Air with 50% RH at 25ºC has nearly 4 times as much water vapor as air with 50% RH at 5ºC

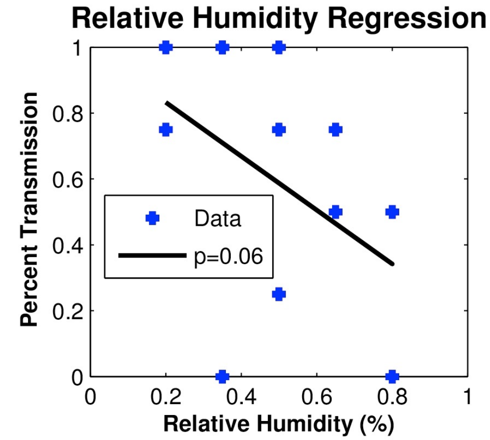
Testing Absolute Humidity

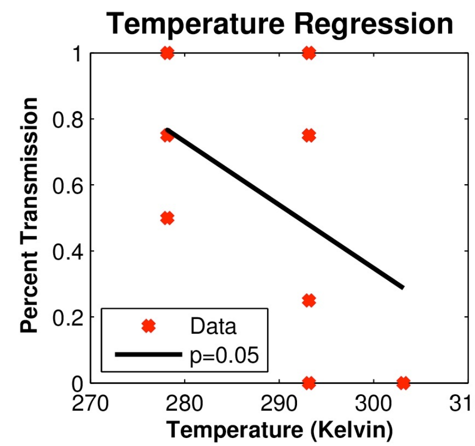
Calculated the vapor pressure from the temperature and RH

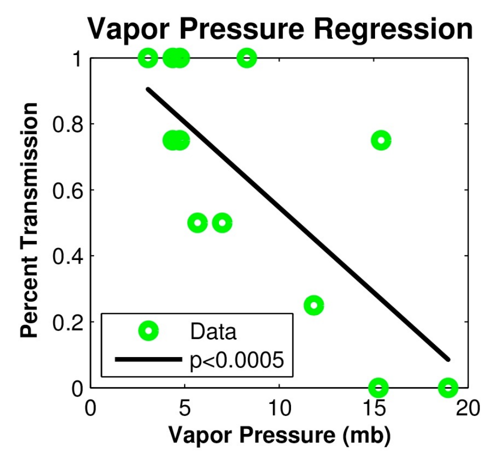
$$e_s(T) = e_s(T_0) \times \exp\left(\frac{L}{R_v}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right)$$

$$e = e_s(T) \times \frac{RH}{100}$$

Found influenza transmission possesses a much more statistically significant association with vapor pressure than either RH or temperature







Hypothesis 1

Virus-laden aerosols (droplet nuclei) are more efficiently produced at lower humidity due to increased evaporation of expelled droplet particles, such that more virus remains airborne longer;

Whether an expelled droplet remains airborne or settles to the surface depends on rates of sedimentation and evaporation

$$v = \frac{dz}{dt} = \frac{2\rho_w g r^2}{9\eta}$$

$$\frac{dr}{dt} = \frac{D\left(e - e_s\right)}{\rho_w R_v T r}$$

Hypothesis 1

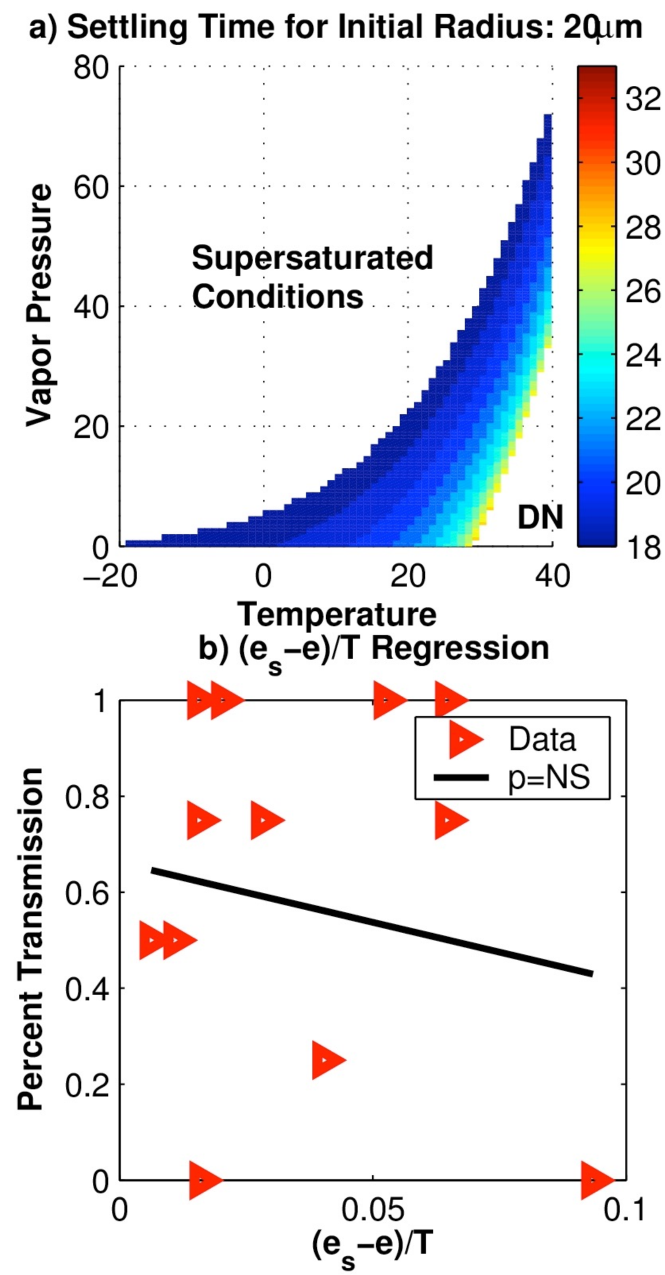
Evaporation will produce more droplet nuclei at lower humidity

However, evaporation really proceeds as:

$$\frac{dr}{dt} = \frac{D(e - e_s)}{\rho_w R_v T r}$$

$$\Rightarrow \frac{dr}{dt} \propto \frac{e - e_s}{T} \qquad (1)$$

If evaporation is the means through which humidity affects influenza transmission, then a strong relationship between influenza transmission and (1) should also exist.



Hypothesis 2

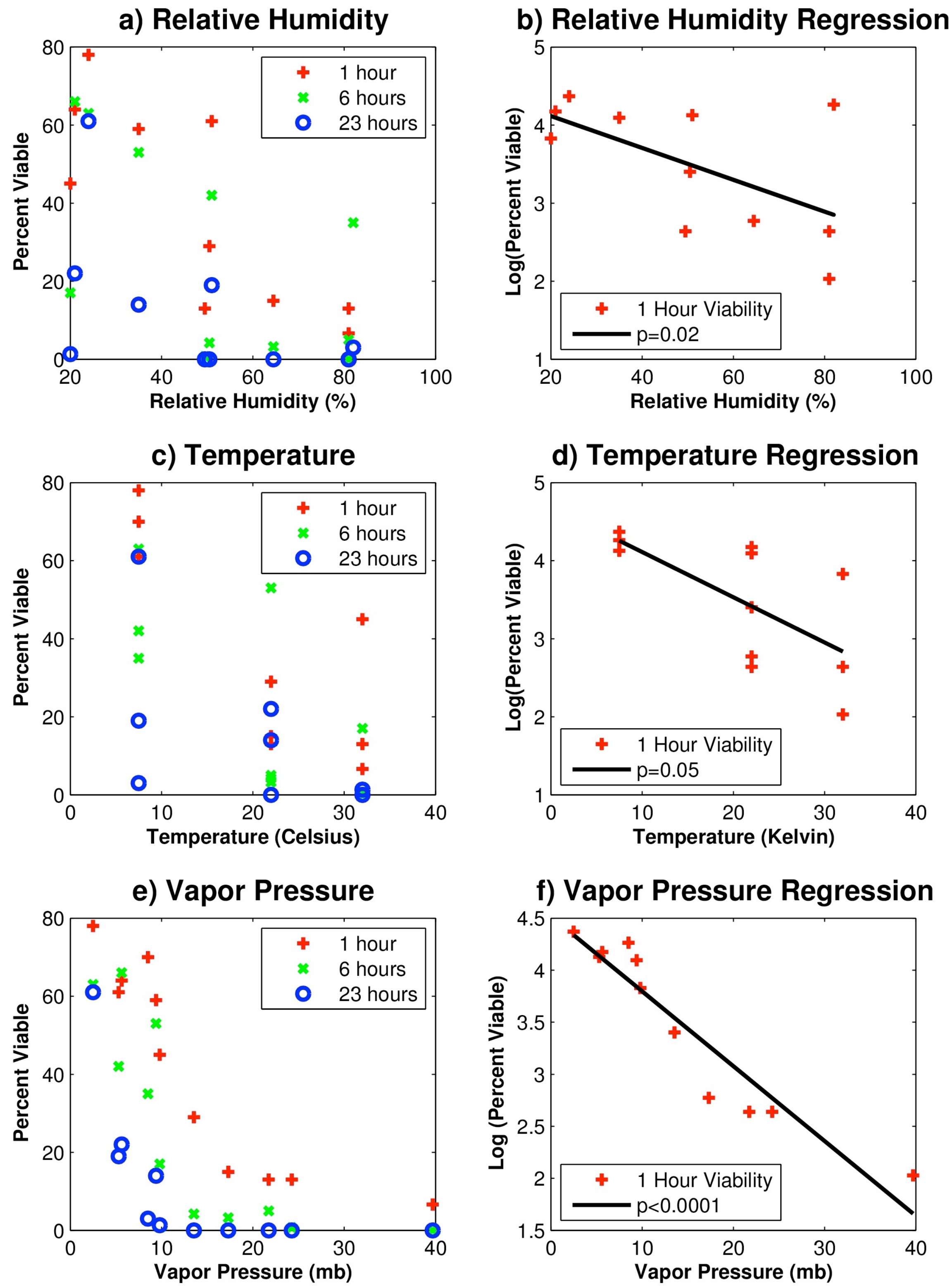
Influenza virus survival (IVS) increases as humidity decreases, such that the airborne virus remains viable longer at lower humidity

Many studies of IVS response to RH and temperature. No studies of IVS response to absolute humidity.

Hypothesis 2

Data of Harper (1961)

90% of 1-hour influenza virus survival variance is explained by absolute humidity

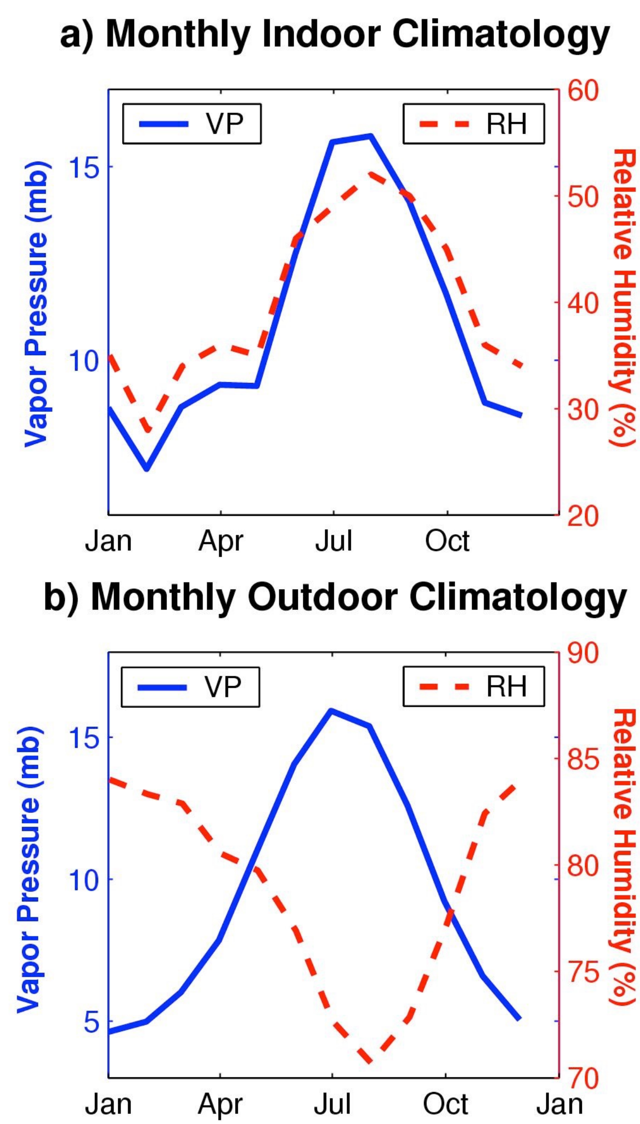


Seasonality of Influenza

The relationships presented indicate that low humidity levels favor influenza survival and transmission

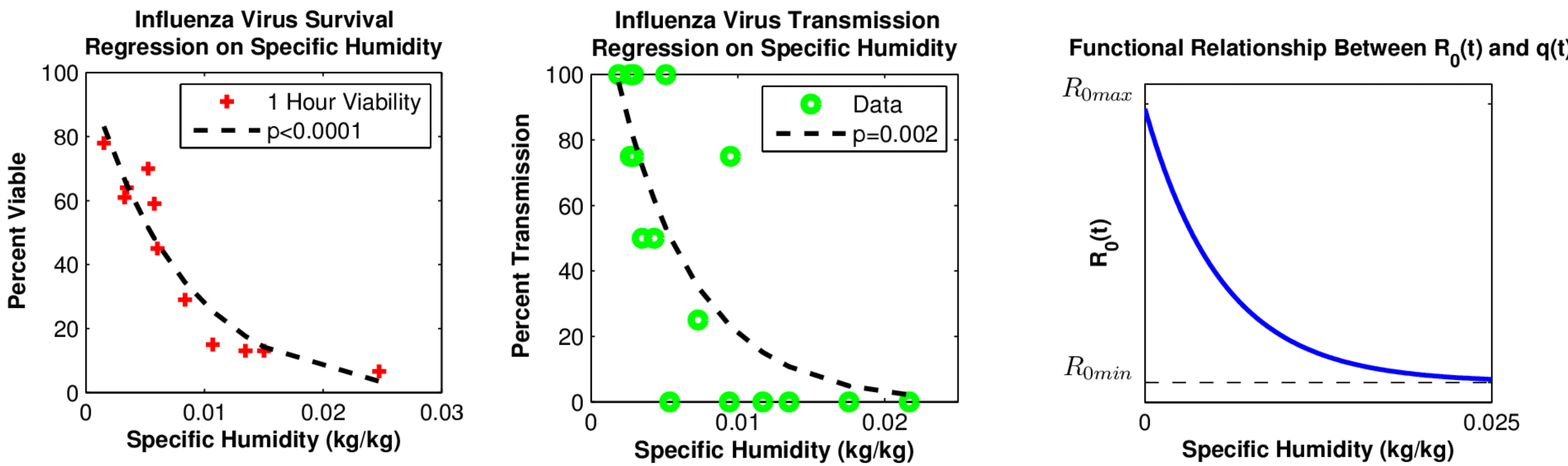
Absolute humidity (VP) is minimal, both indoor and outdoor, in winter

Can we use observed AH conditions to simulate influenza?



Can we use observed AH conditions to simulate influenza?

Functional Humidity Relationship



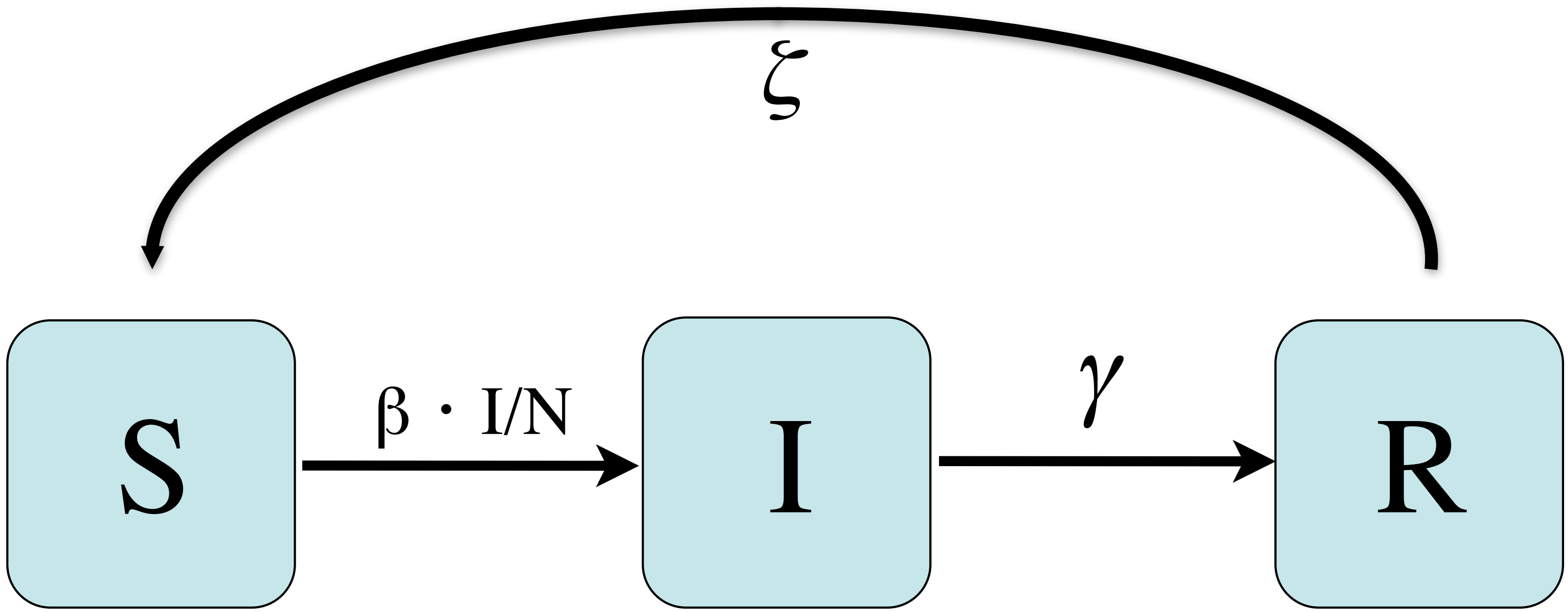
$$R_0(t) = \exp(a \times q(t) + b) + R_{0\,\min}$$

$$a = -180$$

$$b = \log(R_{0\,\max} - R_{0\,\min})$$

γ

Humidity-forced SIRS Model



Here β is a function of observed daily specific humidity, a measure of absolute humidity

$$\frac{dS}{dt} = \frac{N - S - I}{L} - \frac{\beta(t)IS}{N}$$

$$\frac{dI}{dt} = \frac{\beta(t)IS}{N} - \frac{I}{D}$$

Assessed fit to excess weekly P&I mortality via a conversion factor cases->lagged deaths

1972-2002 Model Simulations

5000 1972-2002 simulations run for 5 states (AZ, FL, IL, NY and WA)

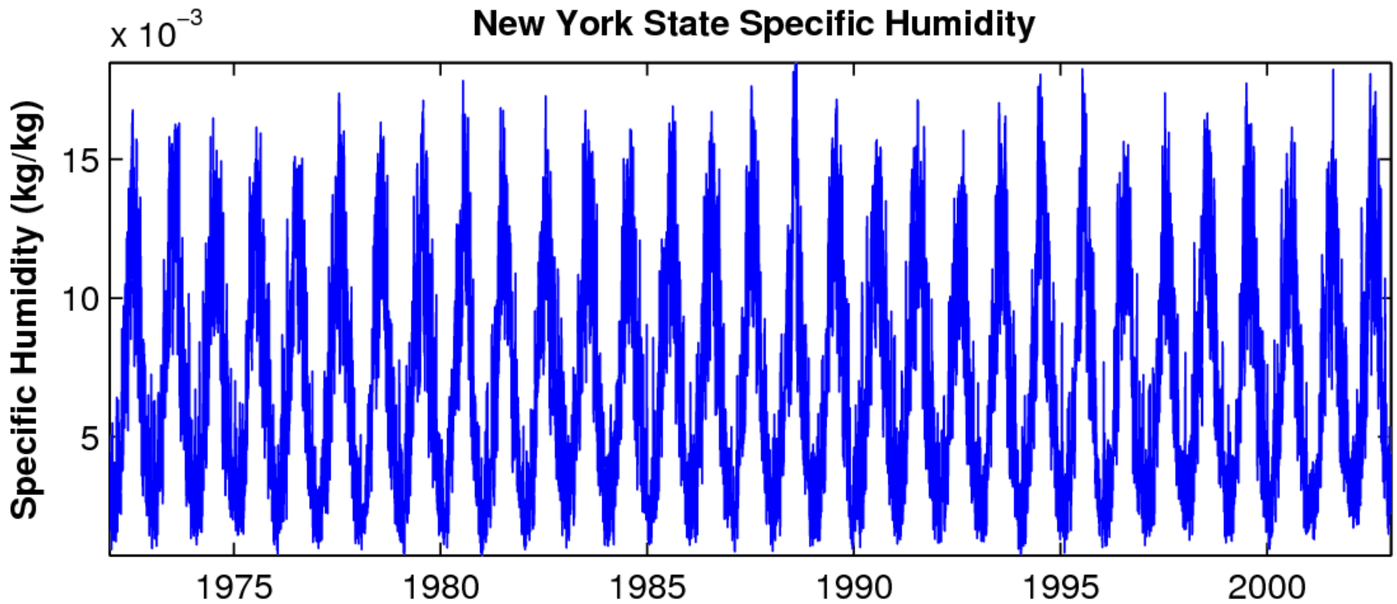
Each run uses a different combination of the parameters:

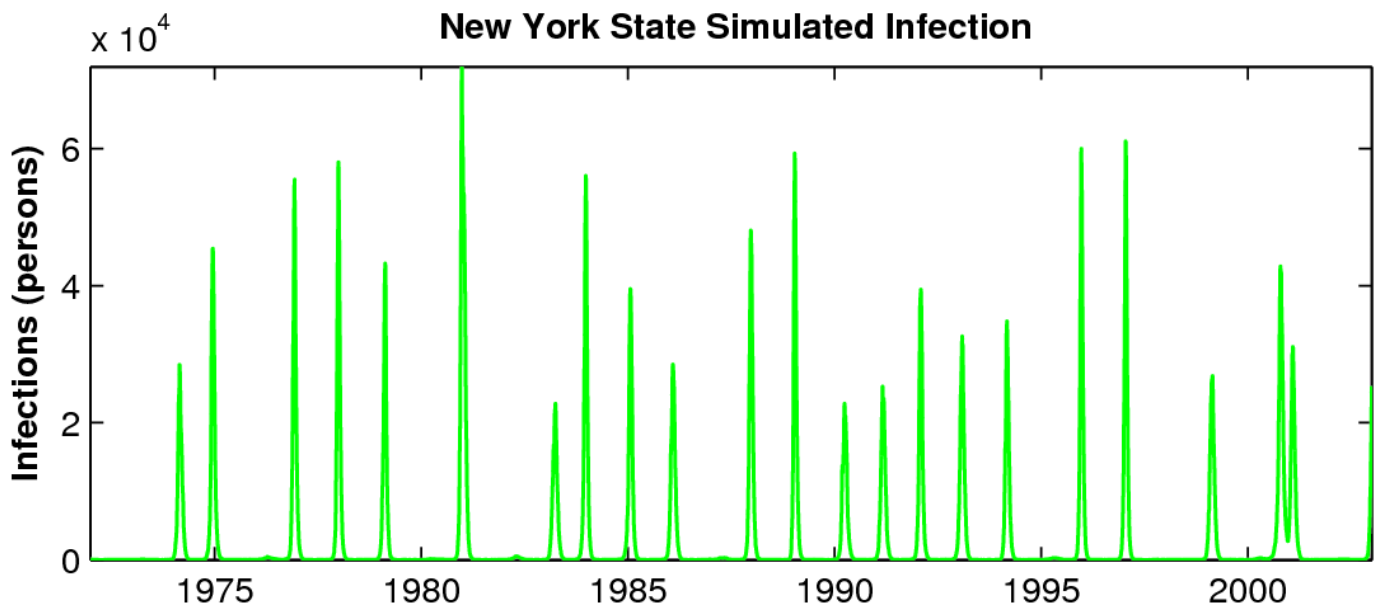
L = 2-10 years

D = 2-7 d

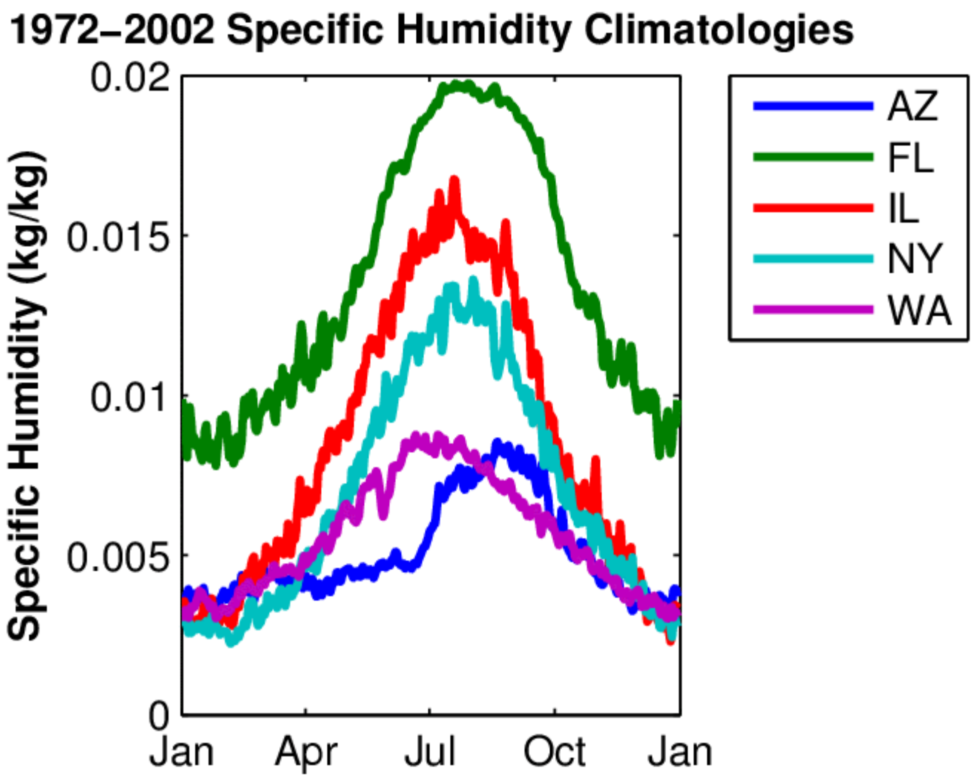
R_{0max} = 1.3-4

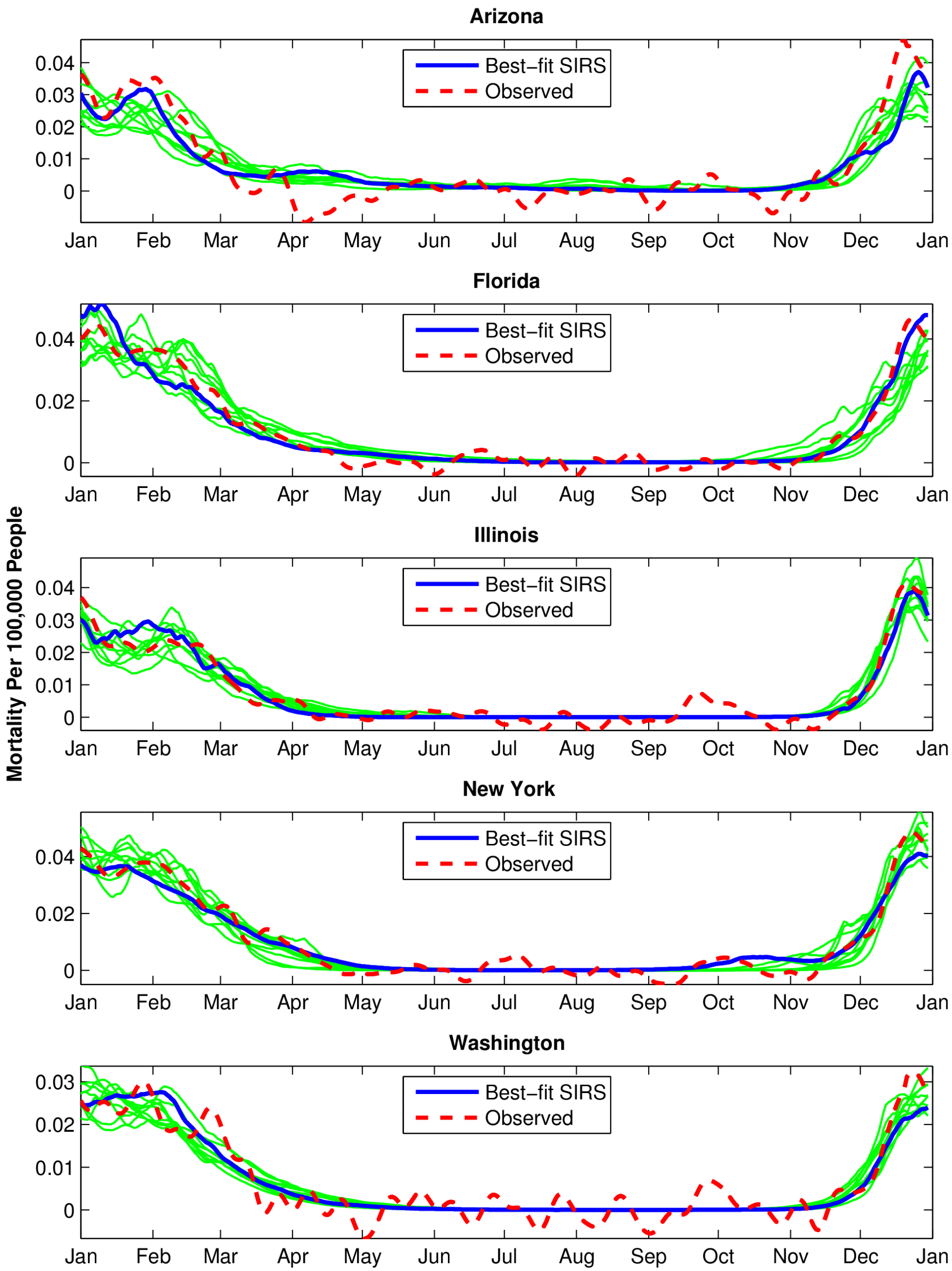
R_{0min} = 0.8-1.3





Modeling the Seasonal Cycle



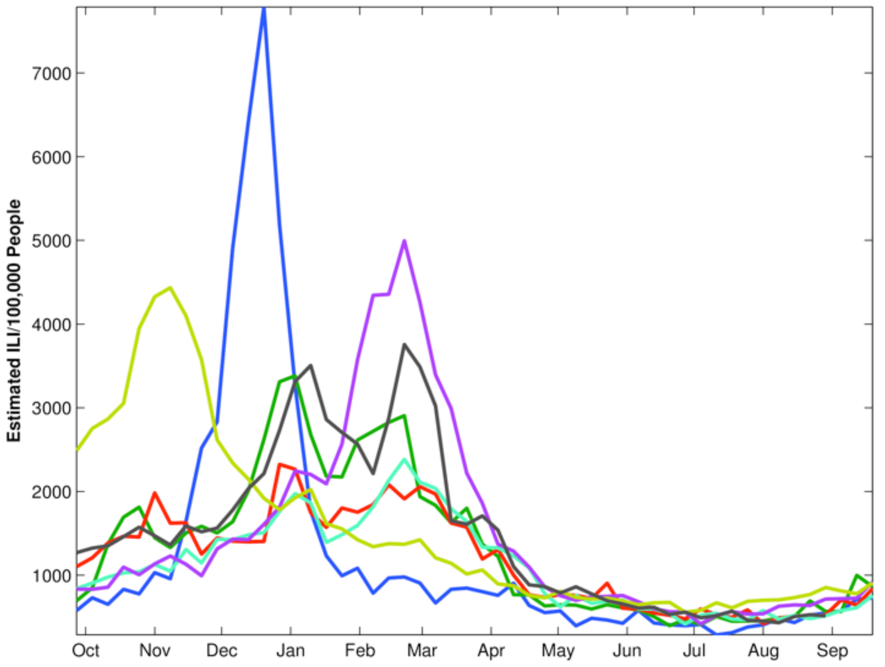


Can We Predict Individual Outbreaks?

Can We Predict Individual Outbreaks?

• Seasonal flu dynamics are nonlinear and irregular

• Outbreaks, though in winter, vary enormously from year-to-year

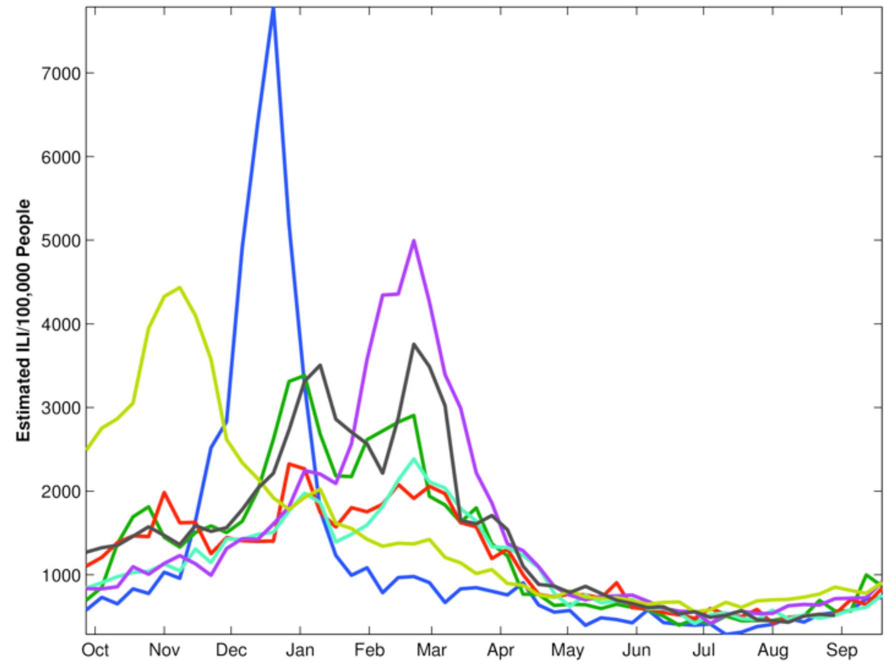


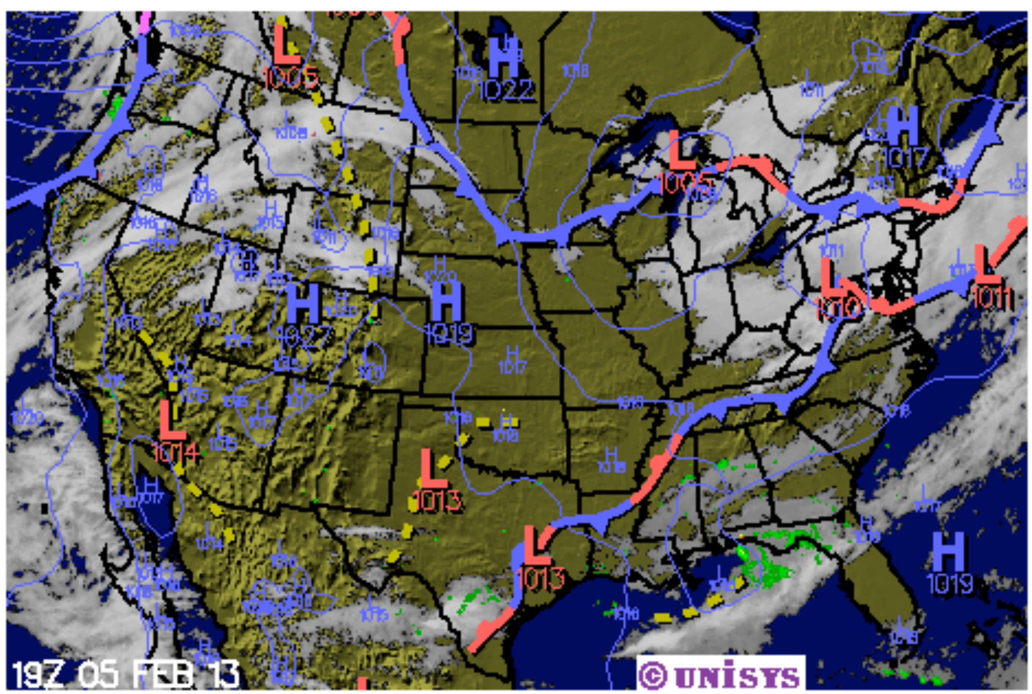
Can We Predict Individual Outbreaks?

- Seasonal flu dynamics are nonlinear and irregular

- Outbreaks, though in winter, vary enormously from year-to-year

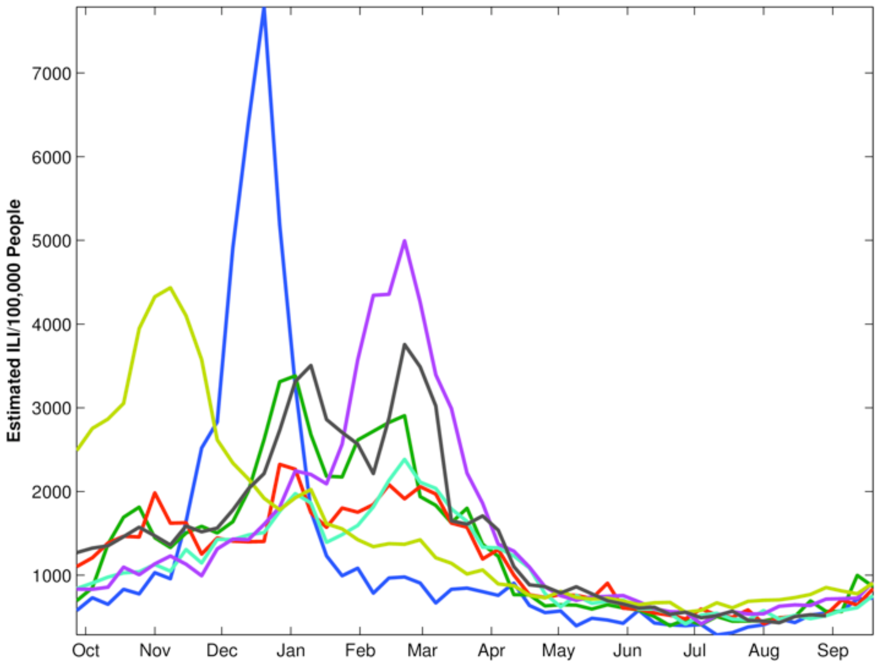
- There are other systems with similar issues that are predicted





Model-Inference Forecasting Approach

- Seasonal flu dynamics are nonlinear and irregular
- Outbreaks, though in winter, vary enormously from year-to-year



To predict influenza, we mimic strategies used in numerical weather prediction

Requires 3 ingredients:

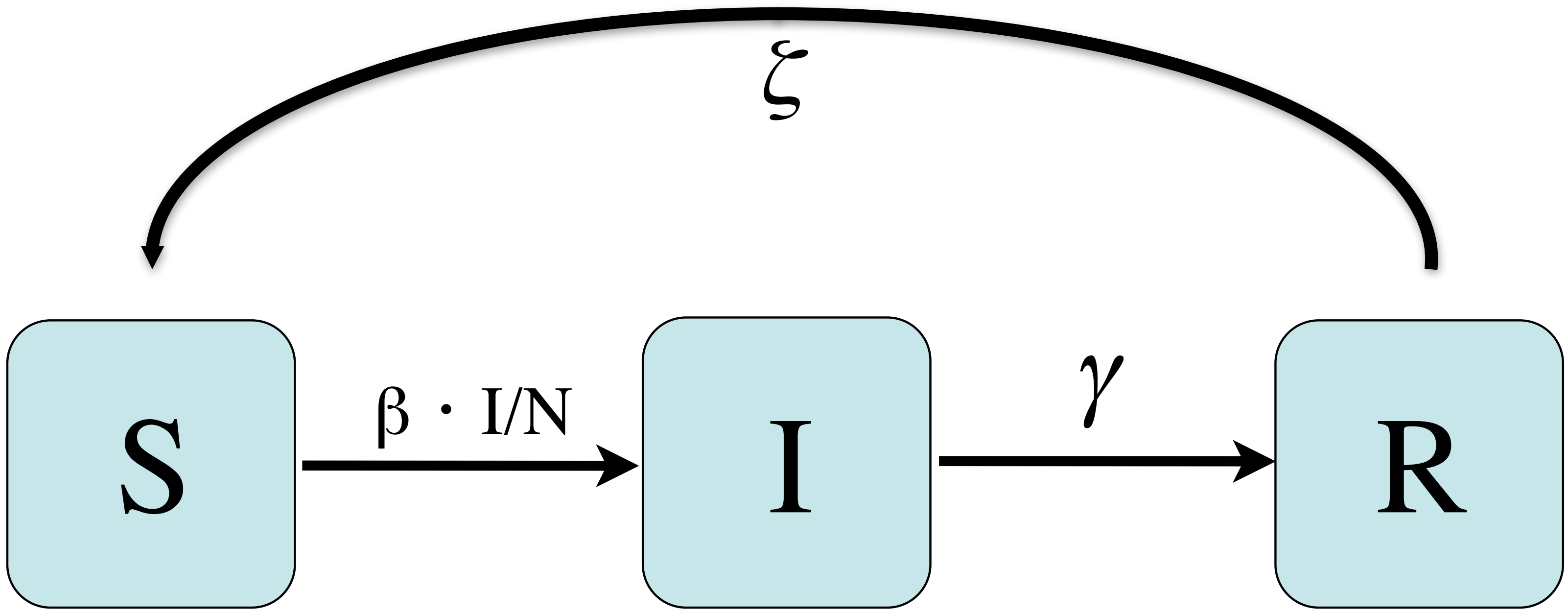
1) Observationally-validated model of influenza transmission dynamics

2) Real-time estimates of influenza infection rates (i.e. observations)

3) Data assimilation method to rigorously combine #1 and #2.

γ

Humidity-forced SIRS Model



Here β is a function of observed daily specific humidity, a measure of absolute humidity

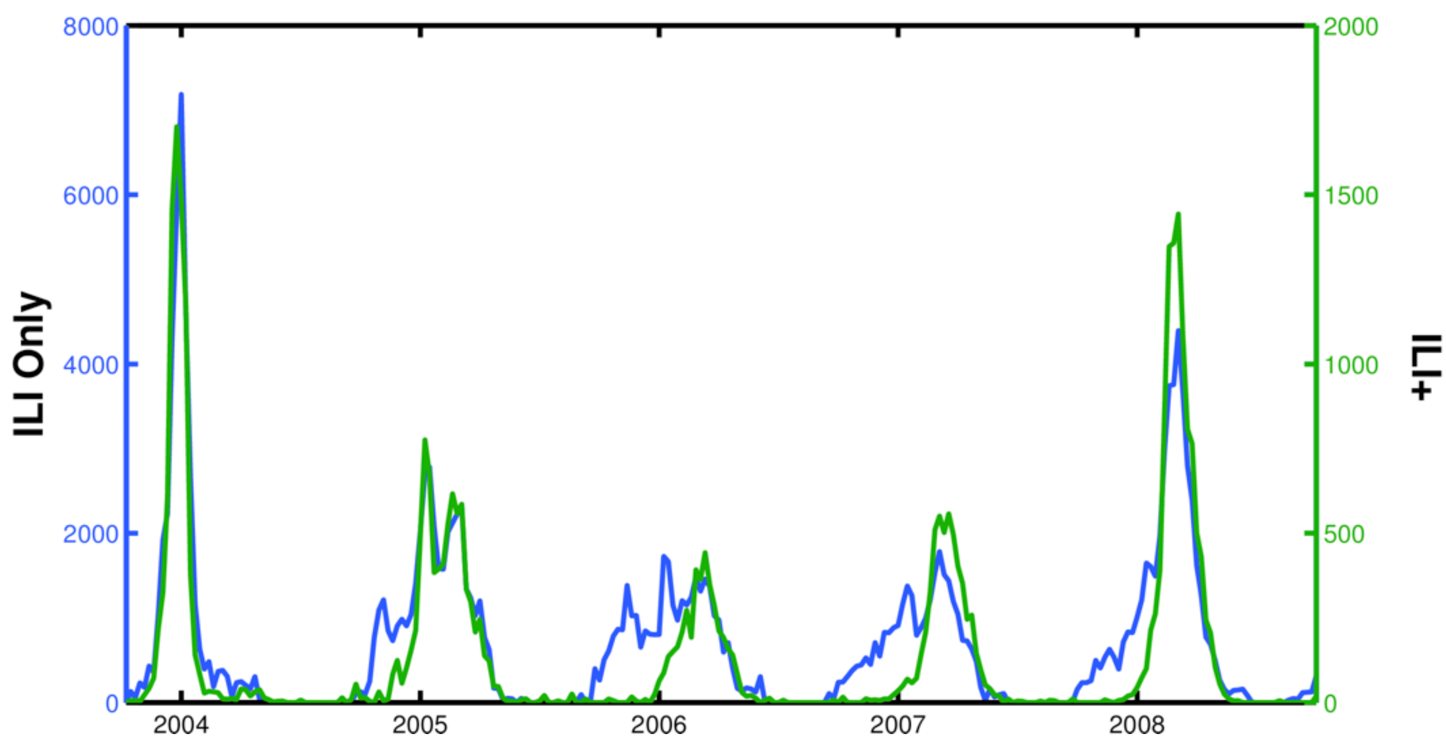
$$\frac{dS}{dt} = \frac{N - S - I}{L} - \frac{\beta(t)IS}{N}$$

$$\frac{dI}{dt} = \frac{\beta(t)IS}{N} - \frac{I}{D}$$

Describes seasonal cycle of influenza (excess weekly P&I mortality)

ILI+

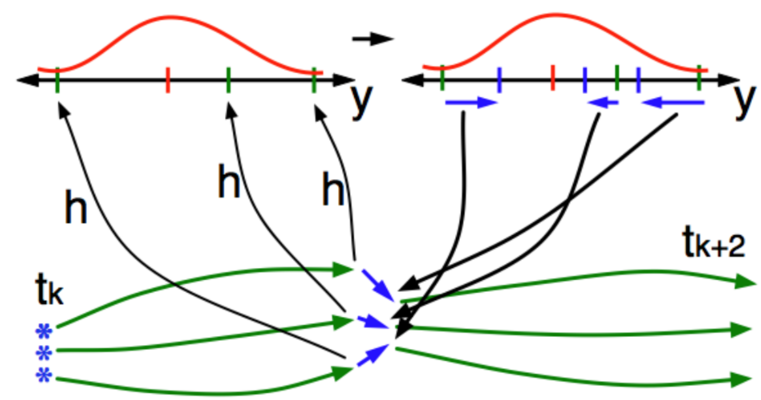
- For municipal forecasting, we often use a more specific estimate of influenza incidence
- We multiply municipal ILI estimates by influenza positive test proportions
- The resulting metric (ILI+) eliminates signal from other respiratory infections, such as rhinovirus



Data Assimilation

Recursive (iterative) filtering of observations in a statistically rigorous fashion into an evolving model construct

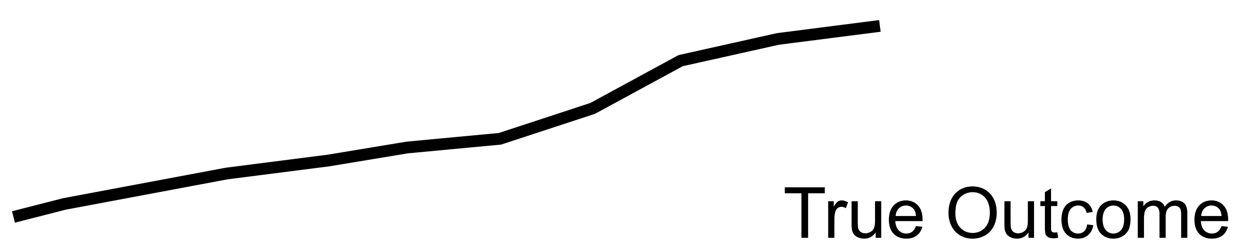
- Particle Filtering
- Kalman Filtering
- Variational Methods



Methods used in many disciplines, including numerical weather prediction where it is used to generate improved forecasts

Prior to Forecast: Training the Model

- Errors in the model structure, model parameters and initial model state amplify through time
- Left to its own devices the model forecast will deviate from reality

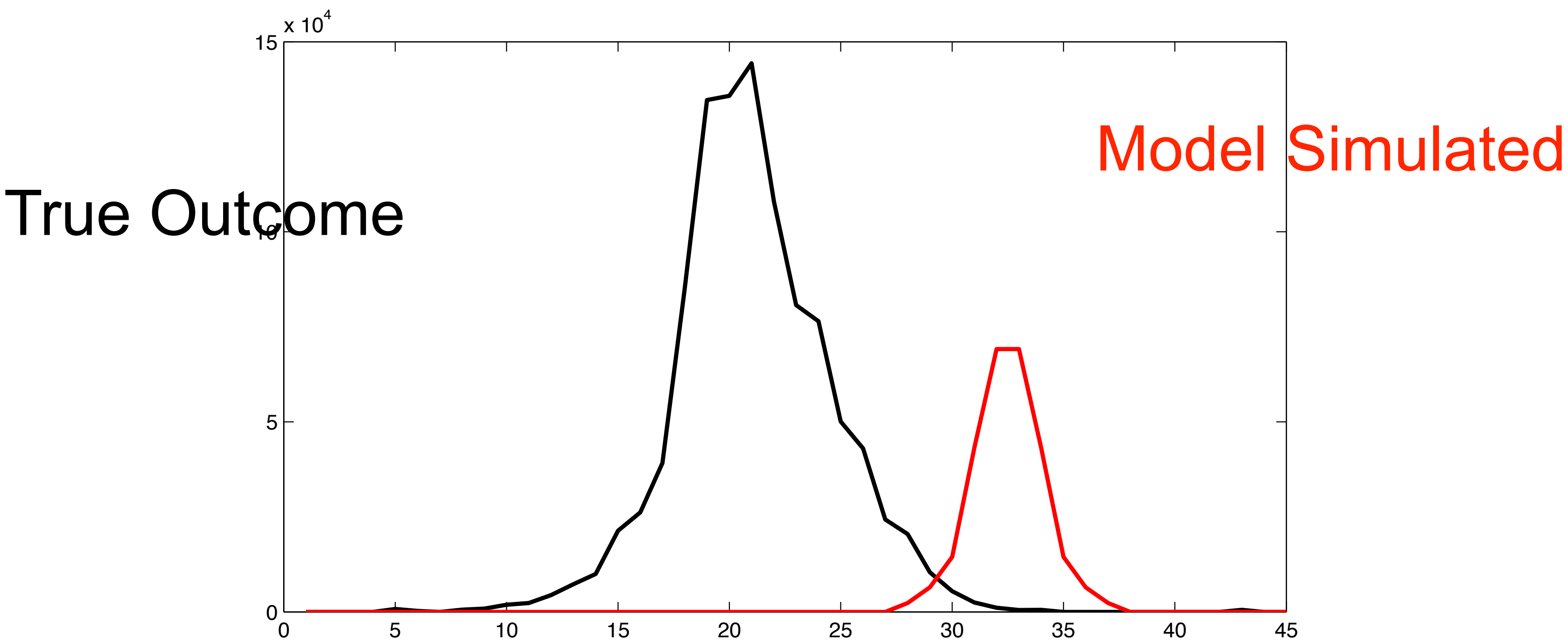


Prior to Forecast: Training the Model

- Errors in the model structure, model parameters and initial model state amplify through time
- Left to its own devices the model forecast will deviate from reality

Prior to Forecast: Training the Model

- Errors in the model structure, model parameters and initial model state amplify through time
- Left to its own devices the model forecast will deviate from reality



Prior to Forecast: Training the Model

- The real-time observations and data assimilation methods are used to recursively adjust and optimize the mathematical model
- It is an inference problem - estimating unobserved state variables and parameters

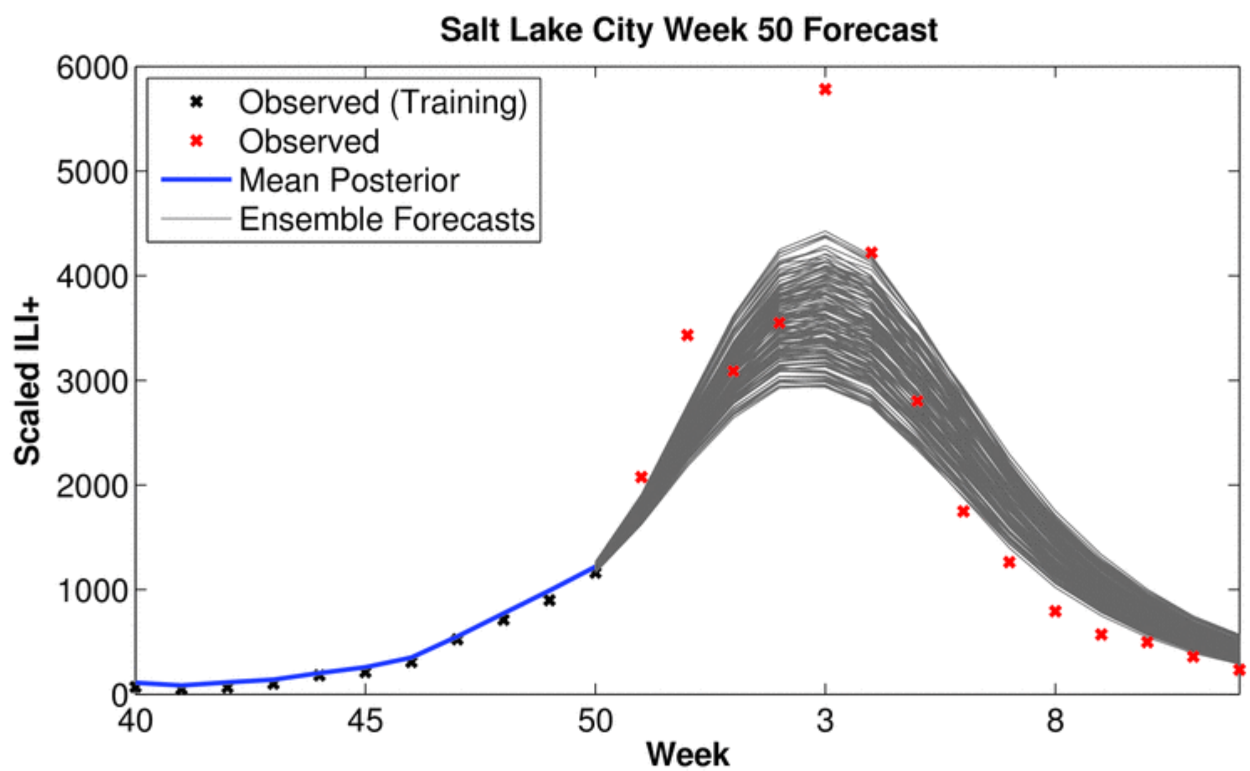
$$P(Z_t|y_t, y_{t-1}, ...) \propto P(y_t|Z_t)P(Z_t|y_{t-1}, ...)$$

- If the data are rich enough for a given system, the state variables and parameters should be identifiable
- By simulating the past to present well, the system has a higher probability of forecasting the future accurately
- The ensemble forecast itself is run following assimilation of the latest observation

Example Real-Time Forecast During 2012-2013

Forecasts (grey lines) made with an SIRS model

Model recursively trained using real-time observations (black ‘x’) and data assimilation methods up to the point of forecast (Week 50)



Observed estimates of influenza incidence that were in the future at the time of forecast are shown as red ‘x’.

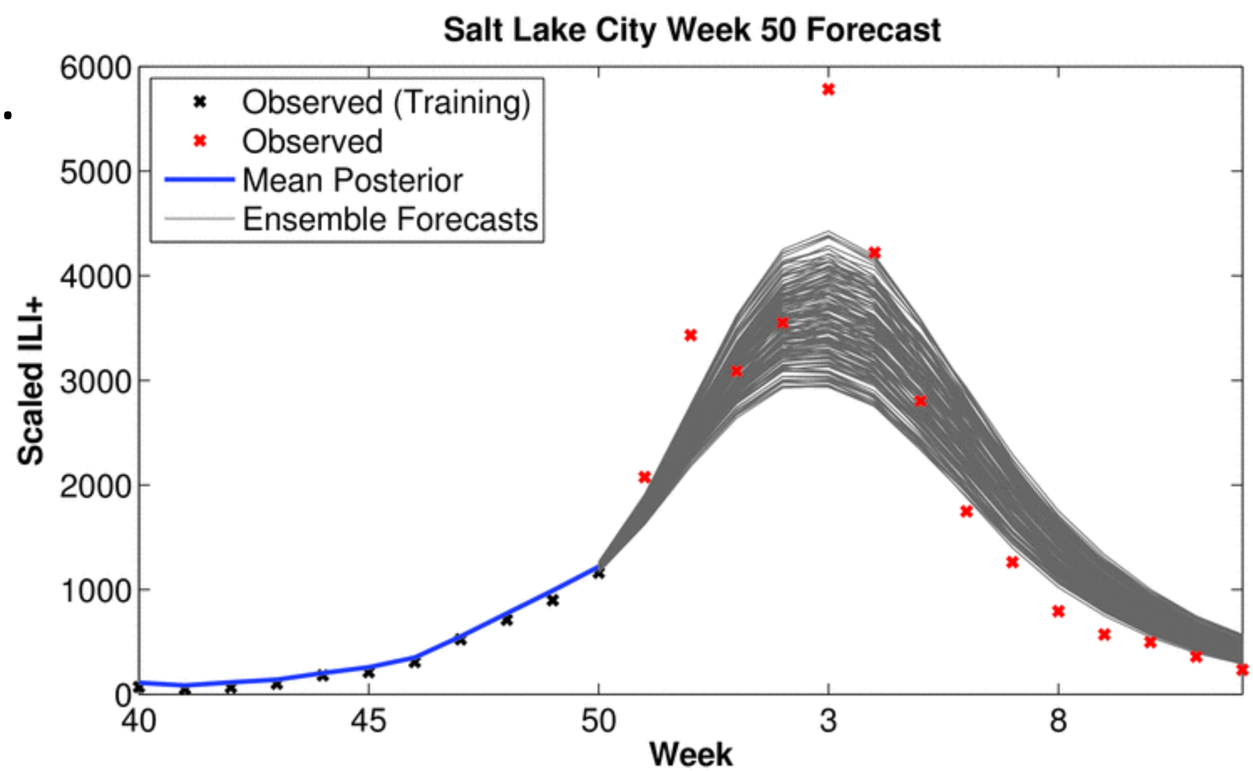
A Calibrated Forecast

Do not simply want to predict an outcome (e.g. the peak will occur in 5 week)

Want to know the certainty of the forecast as it is made

Is there a 90% chance the peak will occur in 5 weeks?

Is there a 20% chance?

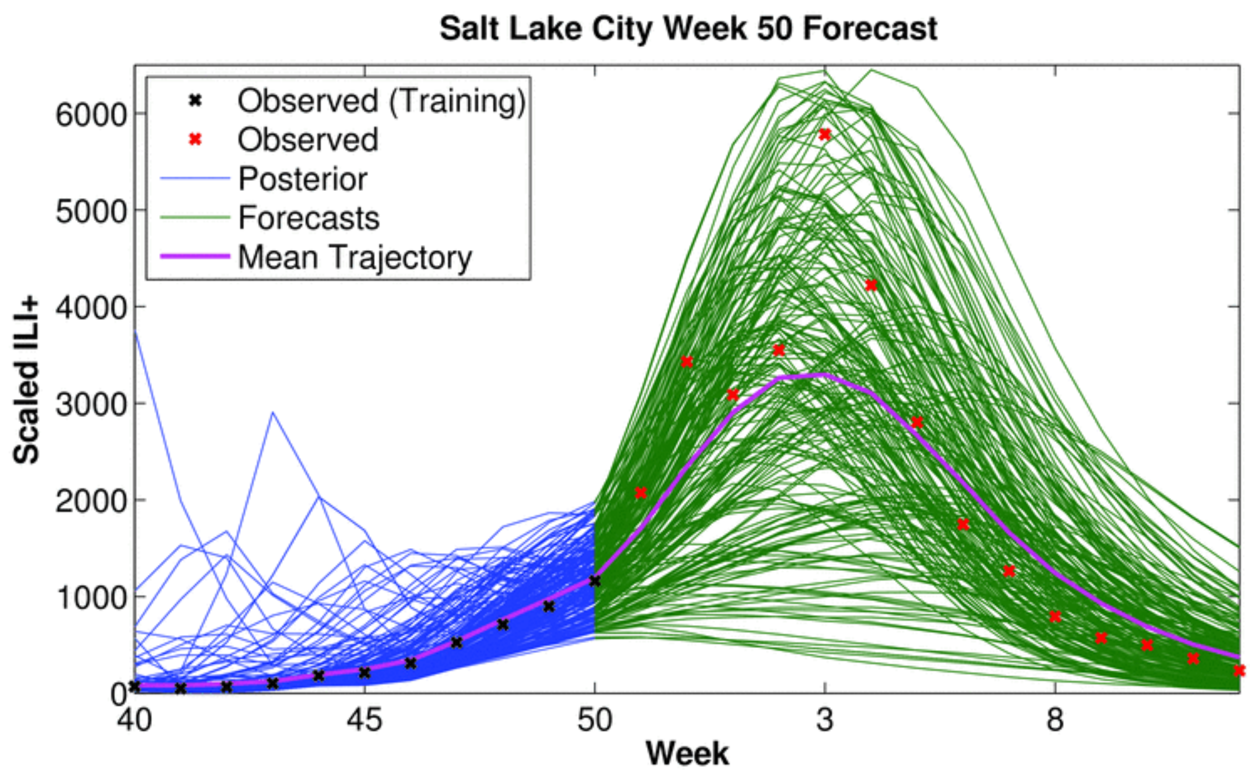


Accurate ascription of forecast certainty provides the public health user a much richer, more actionable prediction

A Calibrated Forecast

It turns out, we can use the spread of each ensemble of predictions to estimate the certainty of a forecast

The relationship between that spread (variance) and accuracy for past forecasts can be used to calibrate forecasts made in real time



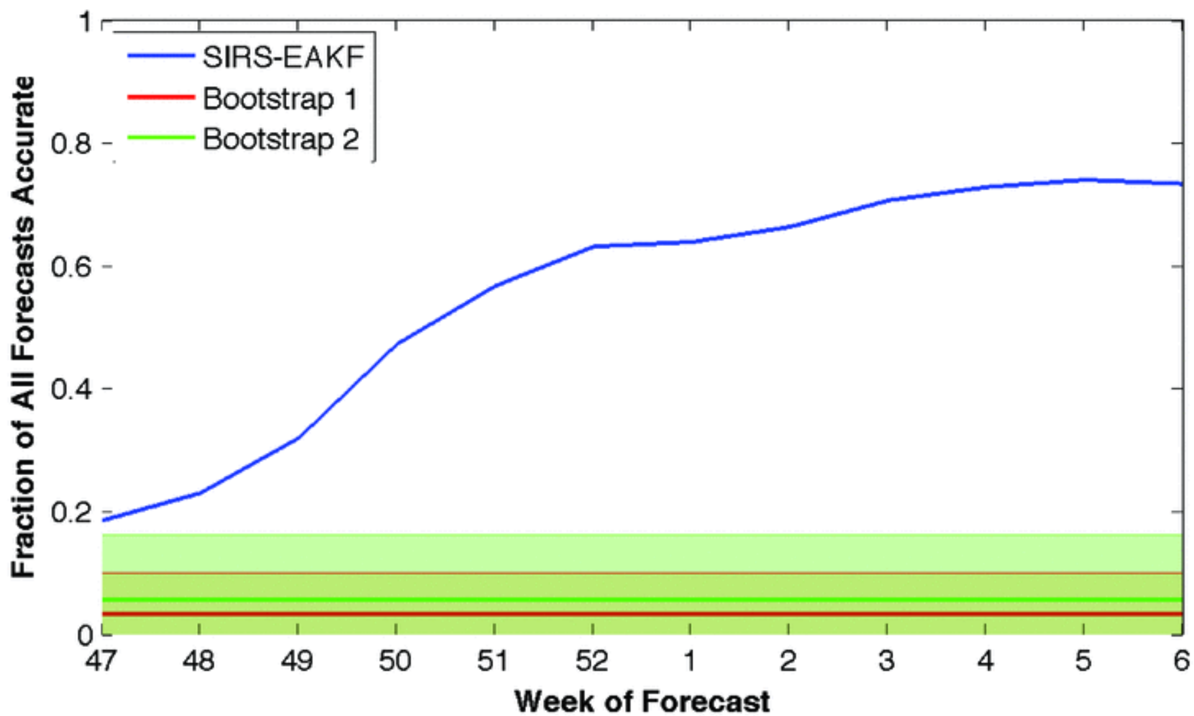
Above plot now shows the individual trajectories within a *single* ensemble forecast

Real-Time Forecast for 108 US Cities 2012-2013 Season

- A number of issues to be verified:
  - The accuracy of the forecasts (are the forecasts superior to climatological expectance)
  - The expected accuracy of the forecasts (does the ensemble spread provide good information on the quality of individual forecasts)
  - The forecast lead (are accurate forecasts of peak timing engendered in advance of the peak)

Predicting Peak Timing

- A number of issues to be verified:
  - The accuracy of the forecasts -- by Week 52 of the 2012-2013 season 63% of forecasts for 108 cities were accurately forecast (84% of cities peaked Week 2 or later)



Much Work Remains

- Can we build a more reliable forecast model?

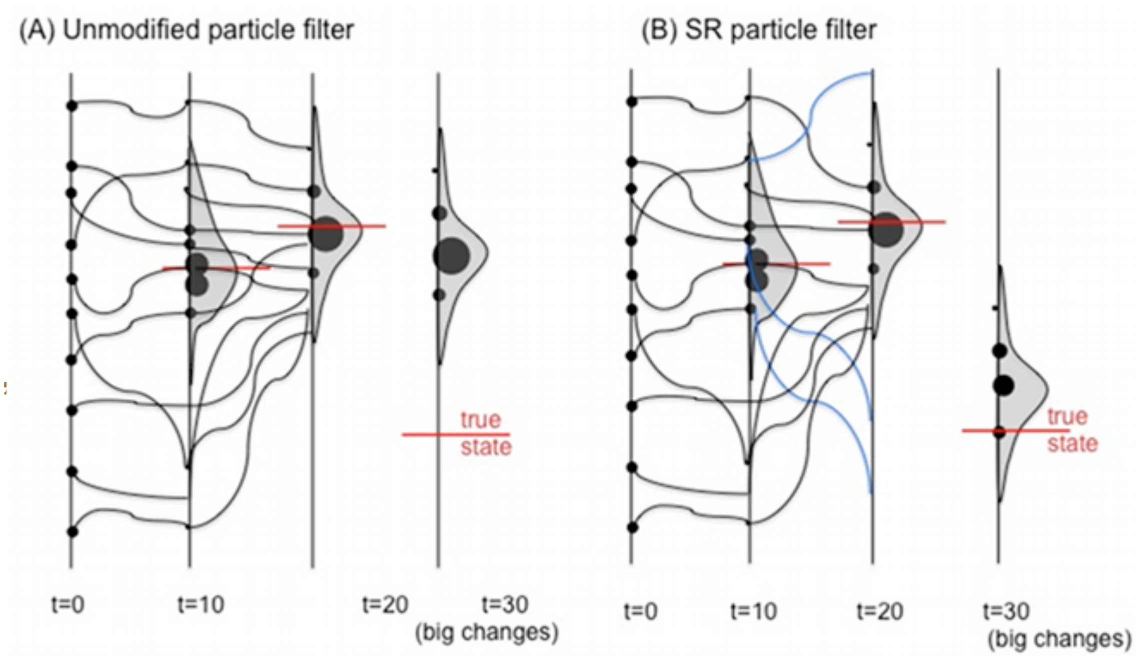
Testing Alternate Model Forms (age-stratified, stochastic v. deterministic, multiple strains, spatially explicit)

- Can we improve model optimization?

Testing and creating different data assimilation methods (ensemble filters, particle filters)

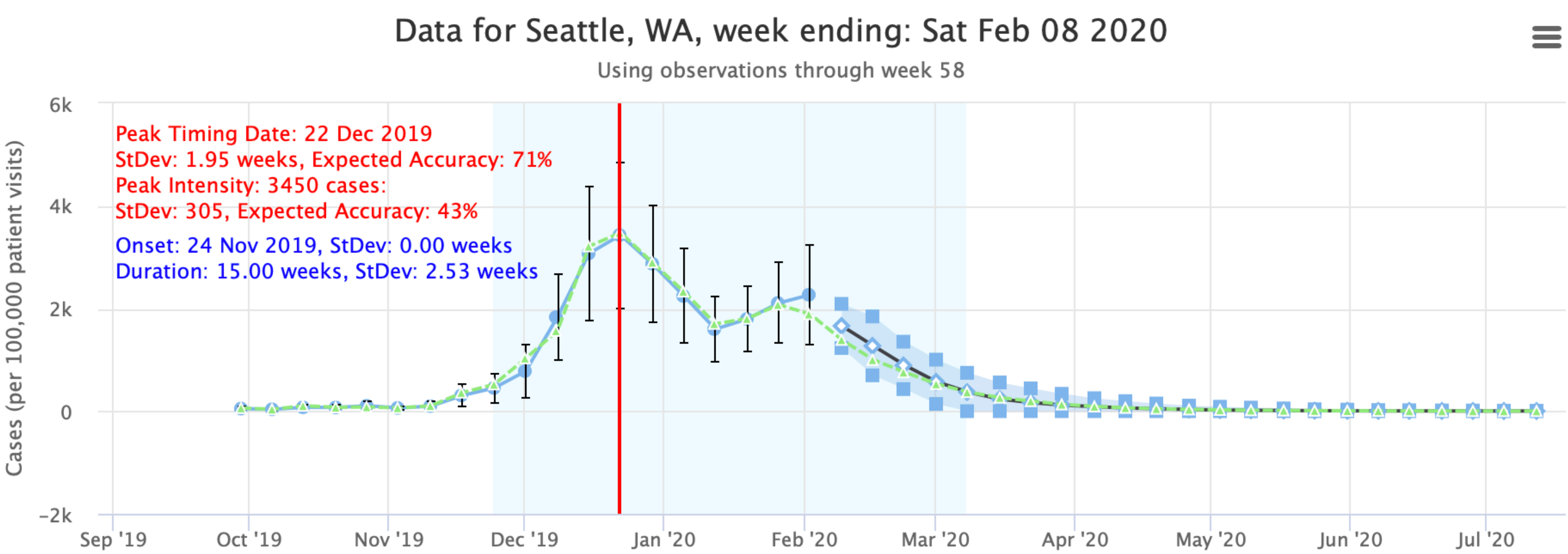
- Can we provide forecasts for local public health use?

Testing different observations of influenza (Google, CDC, Twitter, Wikipedia, WHO)

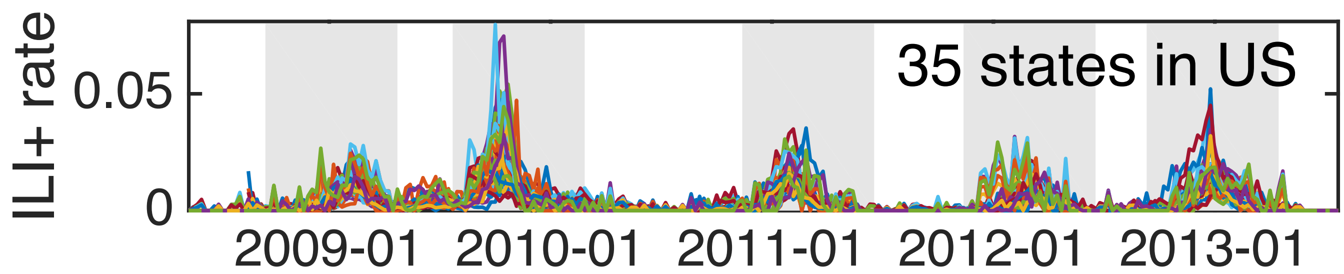


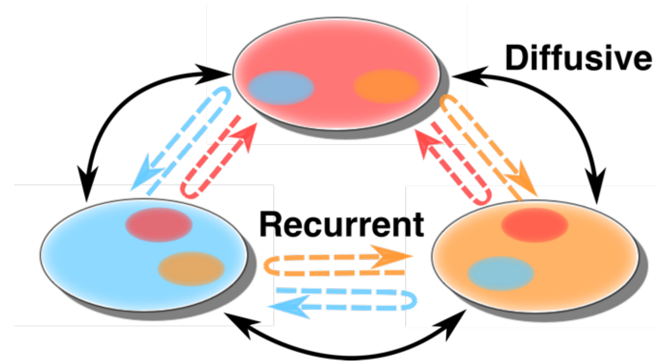
Yang and Shaman, 2014

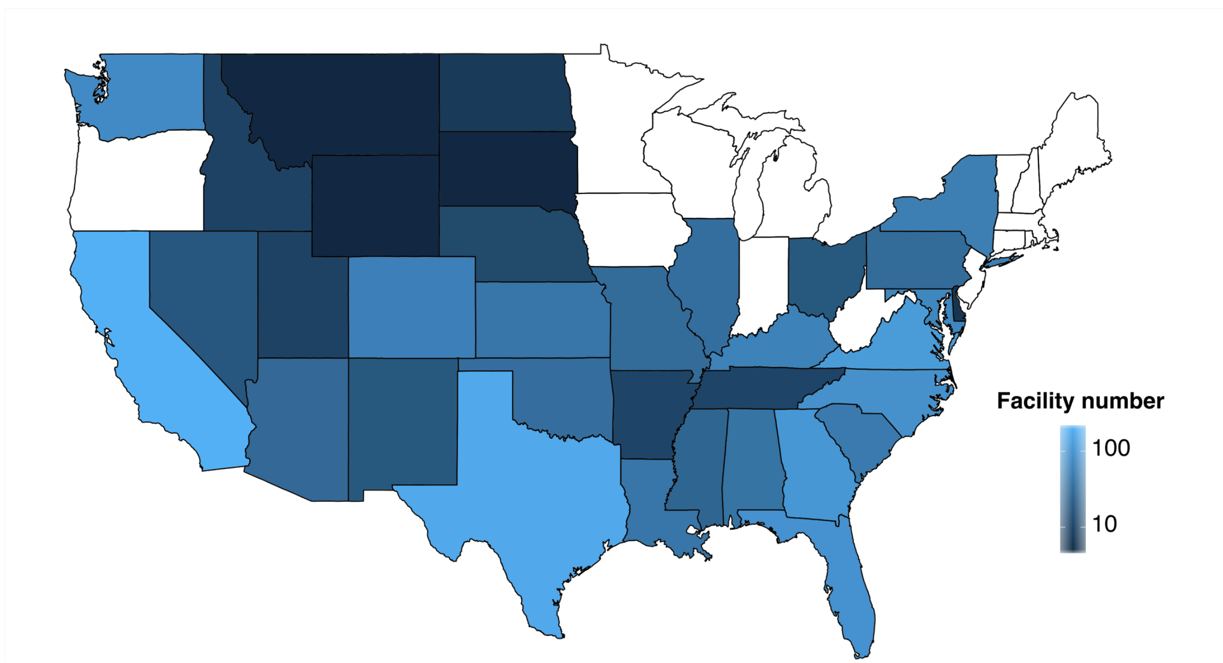
Operational Dissemination:
A Web Portal (cpid.iri.columbia.edu)

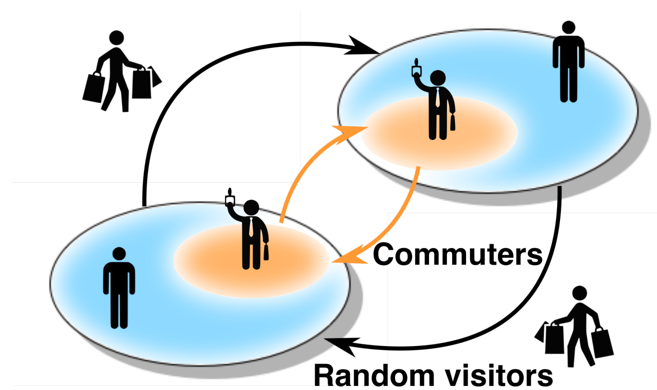


Large-Scale Spatial-Temporal Forecast of Influenza







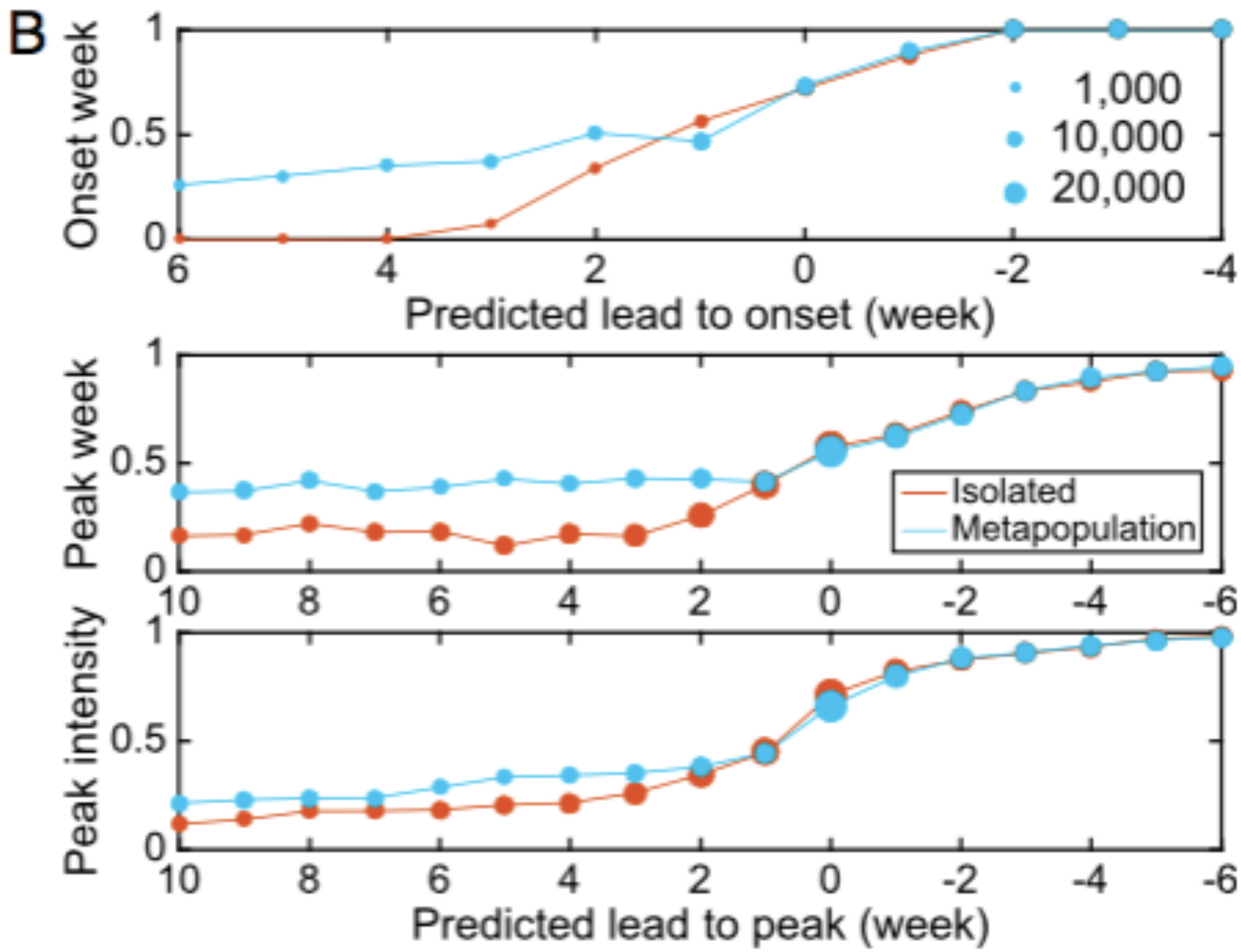


$$\frac{dI_n^k}{dt} = \frac{\beta_n(t) S_n^k I_n}{N_n} - \frac{I_n^k}{D} - \frac{\theta I_n^k}{N_n} \sum_{m \neq n} \bar{N}_m^n + \frac{\theta N_n^k}{N_n} \sum_{m \neq n} \bar{N}_n^m \frac{I_m}{N_m}$$

$$\frac{dS_n^k}{dt} = \frac{N_n^k - S_n^k - I_n^k}{L} - \frac{\beta_n(t) S_n^k I_n}{N_n} - \frac{\theta S_n^k}{N_n} \sum_{m \neq n} \bar{N}_m^n + \frac{\theta N_n^k}{N_n} \sum_{m \neq n} \bar{N}_n^m \frac{S_m}{N_m}$$

Pei et al., 2018

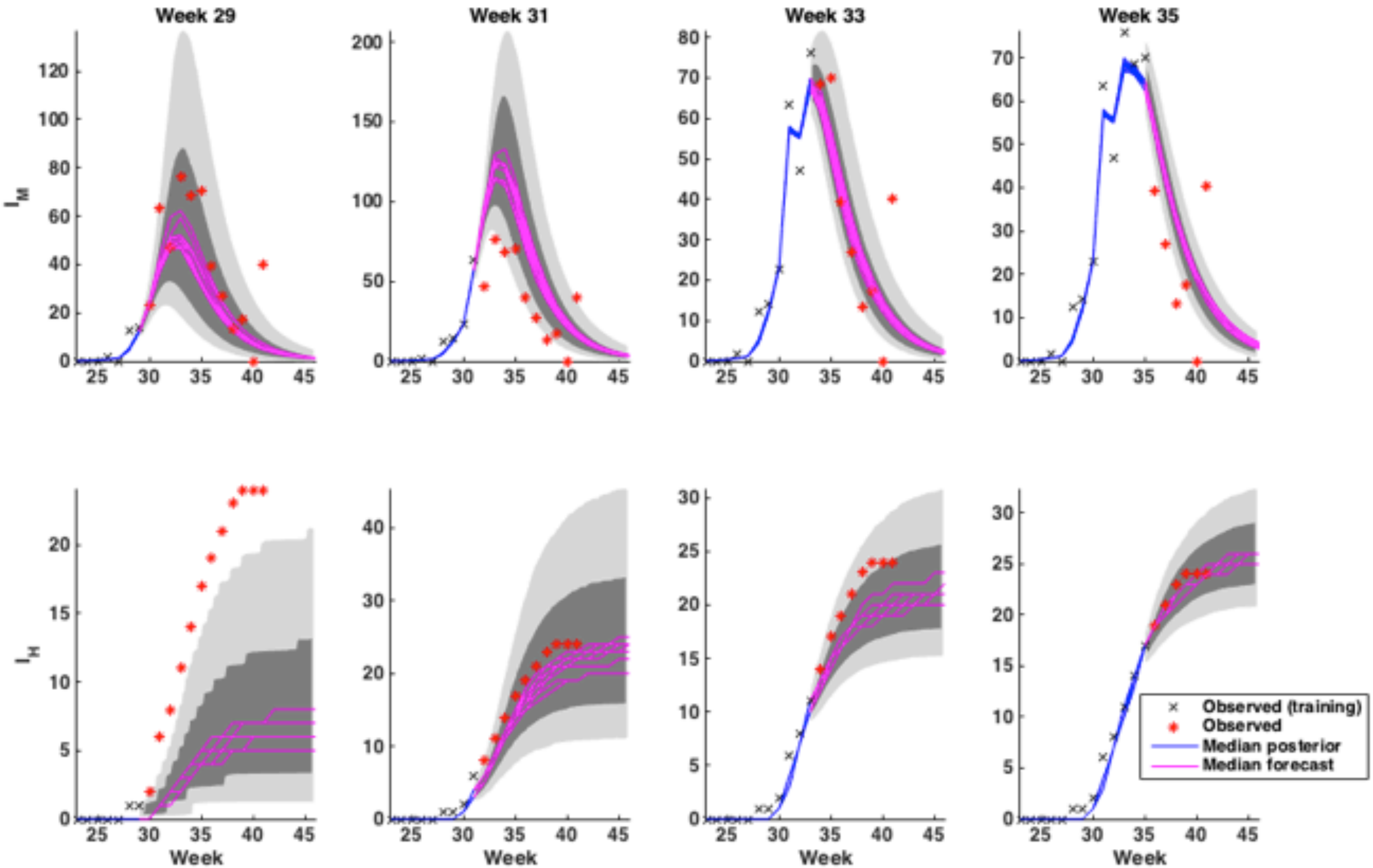
Large-Scale Spatial-Temporal Forecast of Influenza



Accuracy considerably improved, particularly for prediction of outbreak onset (compared to forecasts run in isolation at each site)

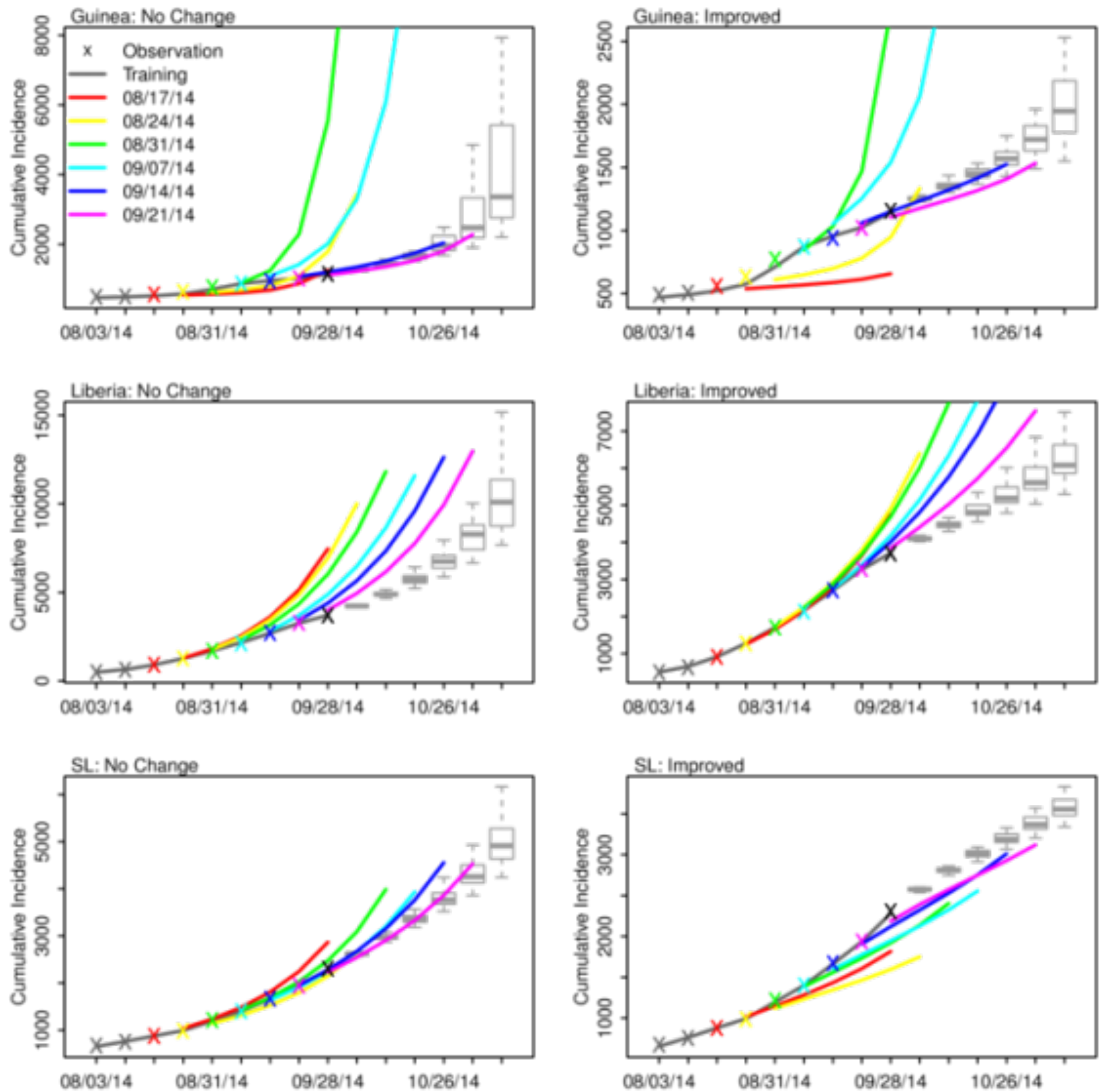
Pei et al., 2018

West Nile Virus



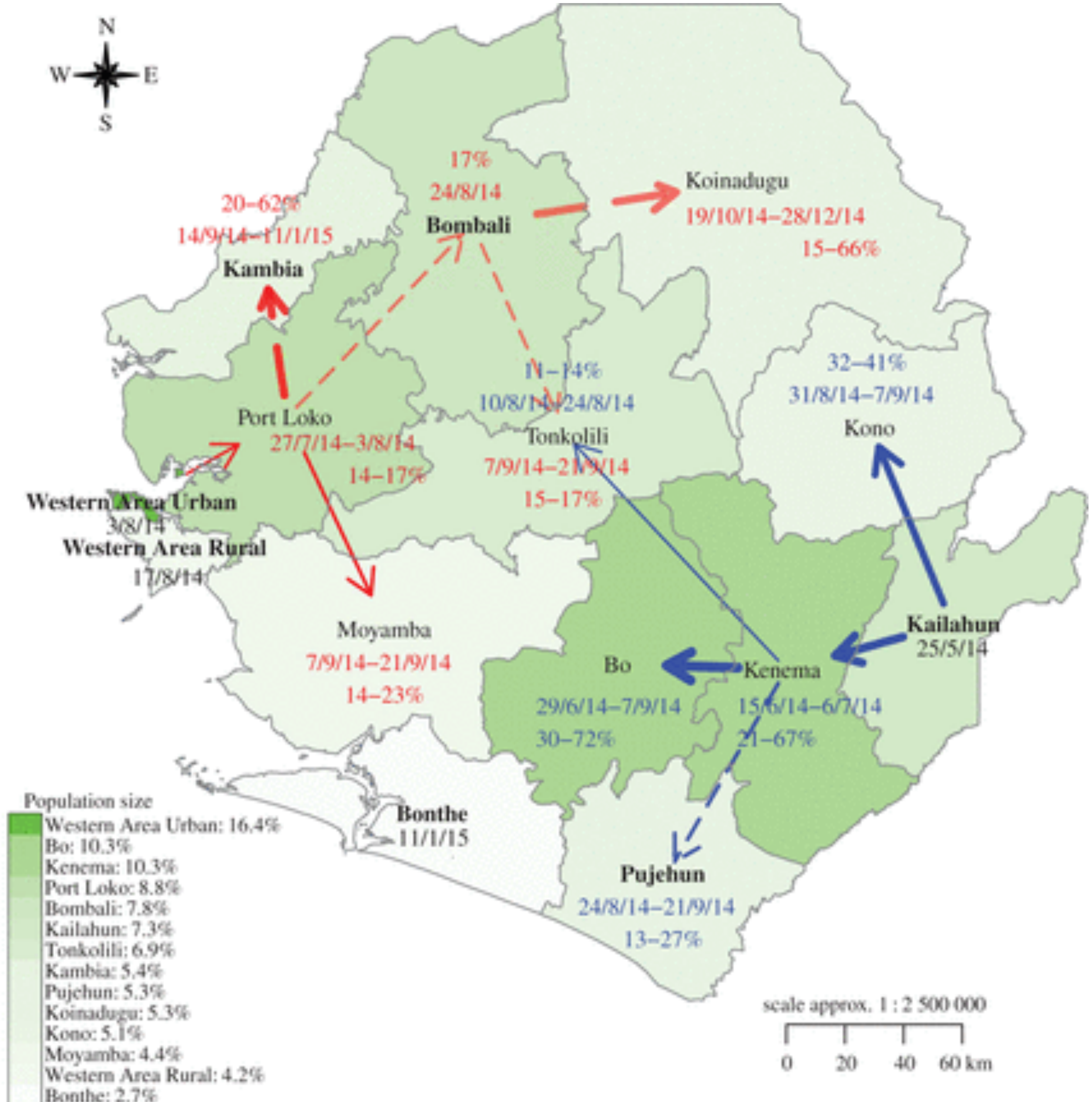
DeFelice et al., 2017

West Africa Real-Time Ebola Forecasts



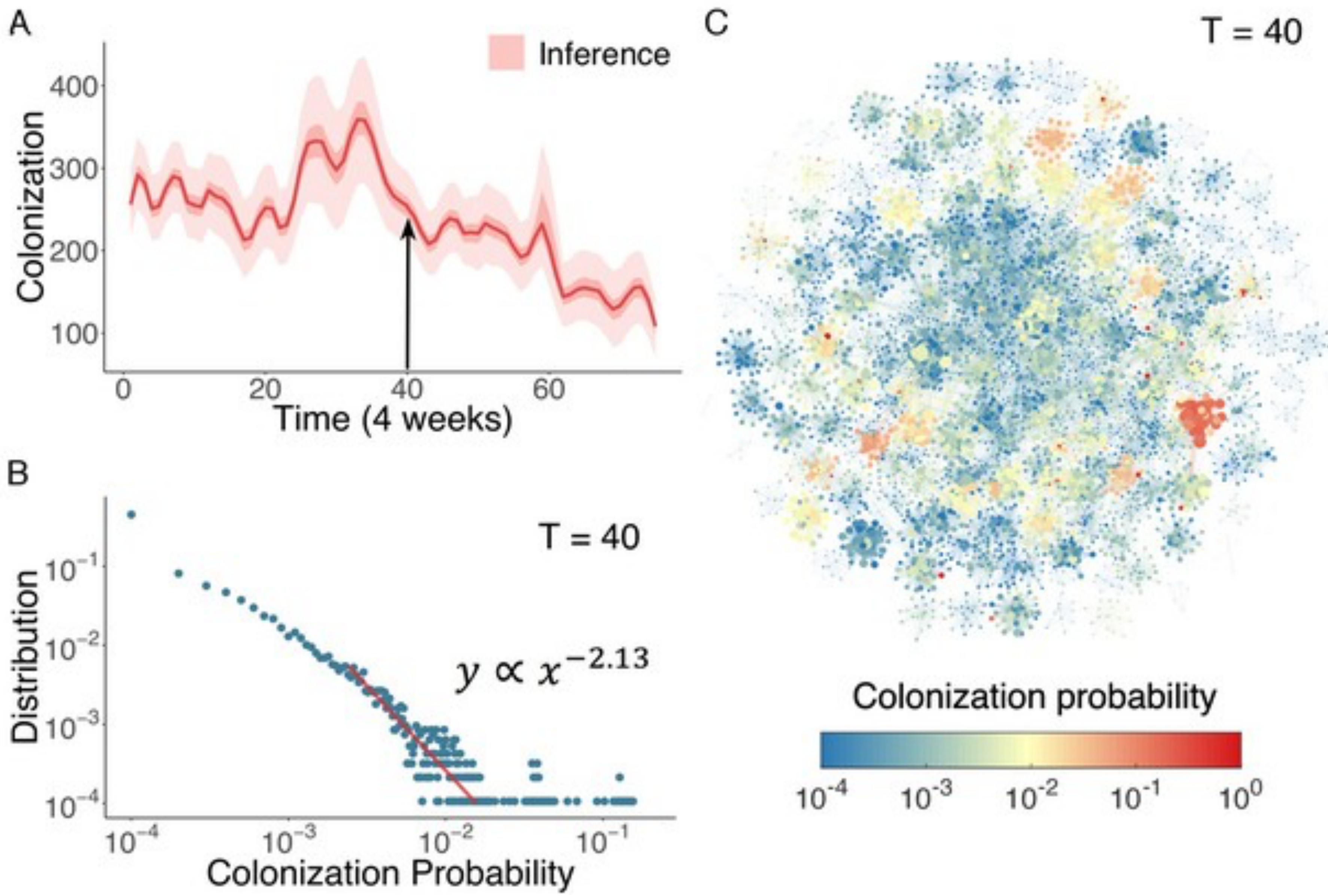
Shaman et al., 2014

Inference of the Spatial Spread of Ebola



Yang et al., 2015

Inference of Asymptomatic Colonization of MRSA



Pei et al., 2018

Understanding Prevalence and Transmission of Respiratory Viruses

- ***Cohort*** — 214 individuals from October 2016 to April 2018.

  (two daycares, CUMC, pediatric and adult ED, high school). Weekly swabs + daily symptoms .

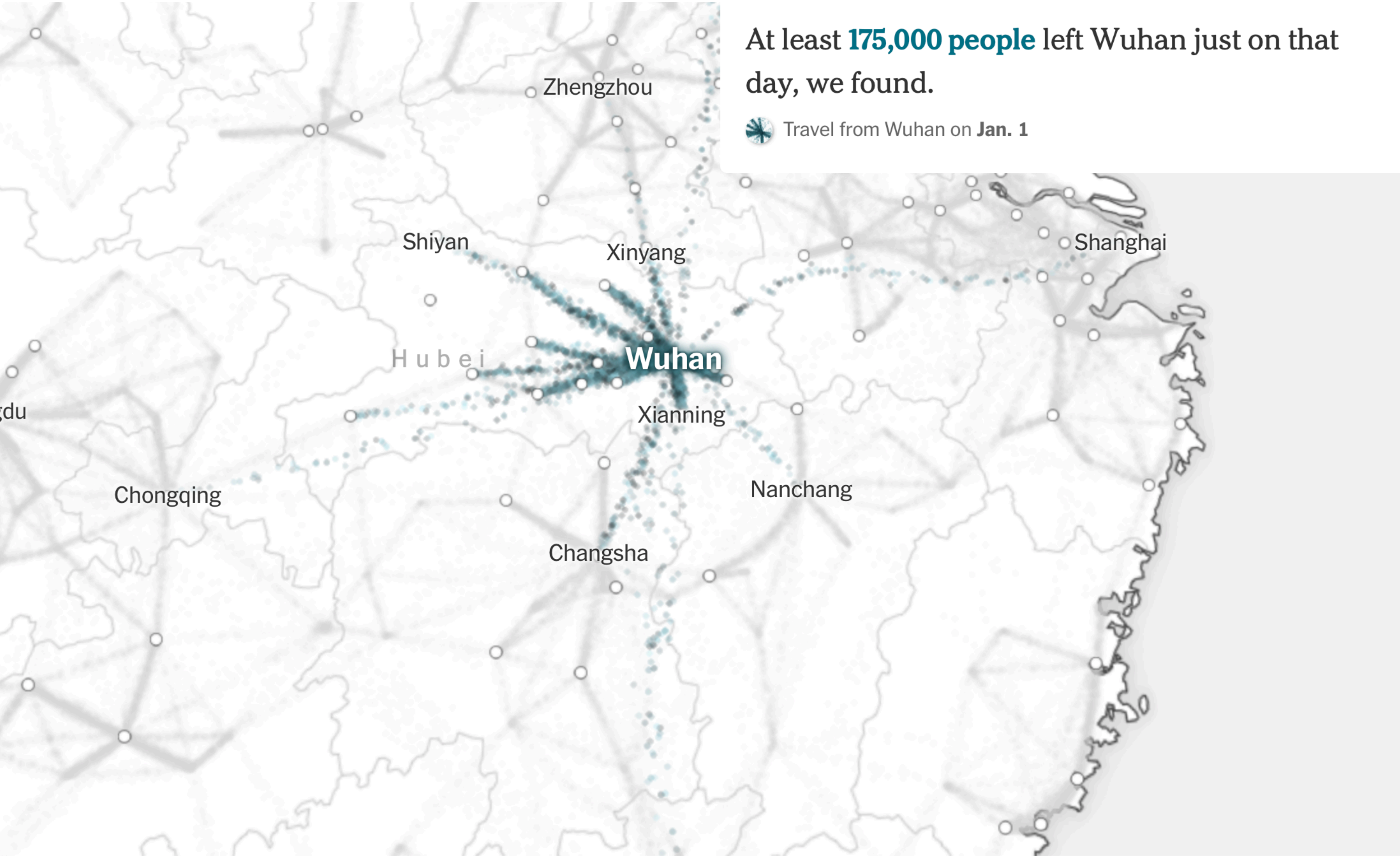
Virome of Manhattan

Most Infections Undocumented

VIRUS	EPISODES*	MA	P(MA|v_i)	HOME	P(HOME|v_i)	MEDS	P(MEDS|v_i)
Influenza	32	7	0.22	14	0.44	18	0.56
RSV	30	2	0.07	6	0.20	12	0.40
PIV	30	3	0.10	4	0.15	9	0.30
HMPV	20	4	0.20	7	0.35	10	0.50
HRV	275	24	0.09	31	0.11	70	0.25
Adenovirus	63	9	0.14	10	0.16	14	0.22
Coronavirus	137	6	0.04	13	0.09	36	0.25

*group of consecutive weekly specimens from a given individual that were positive for the same v
(allowing for a one-week gap to account for false negatives and temporary low shedding).

COVID-19 Rapid Spread



New York Times, March 22, 2018

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?

$$\frac{dS_i}{dt} = -\frac{\beta S_i I_i^r}{N_i} - \frac{\mu\beta S_i I_i^u}{N_i} + \theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} S_i}{N_i - I_i^r}$$

$$\frac{dE_i}{dt} = \frac{\beta S_i I_i^r}{N_i} + \frac{\mu\beta S_i I_i^u}{N_i} - \frac{E_i}{Z} + \theta \sum_j \frac{M_{ij} E_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} E_i}{N_i - I_i^r}$$

$$\frac{dI_i^r}{dt} = \alpha \frac{E_i}{Z} - \frac{I_i^r}{D}$$

$$\frac{dI_i^u}{dt} = (1-\alpha)\frac{E_i}{Z} - \frac{I_i^u}{D} + \theta \sum_j \frac{M_{ij} I_j^u}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} I_i^u}{N_i - I_i^u}$$

$$N_i = N_i + \theta \sum_j M_{ij} - \theta \sum_j M_{ji}$$

- Metapopulation network model representing 375 cities in China
- Use Tencent travel records during the Chunyun spring festival
- Coupled with data assimilation methods
- Use daily observations from all 375 cities
- Simulate January 10-23

Li et al., 2020

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?

$$\frac{dS_i}{dt} = -\frac{\beta S_i I_i^r}{N_i} - \frac{\mu\beta S_i I_i^u}{N_i} + \theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} S_i}{N_i - I_i^r}$$

$$\frac{dE_i}{dt} = \frac{\beta S_i I_i^r}{N_i} + \frac{\mu\beta S_i I_i^u}{N_i} - \frac{E_i}{Z} + \theta \sum_j \frac{M_{ij} E_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} E_i}{N_i - I_i^r}$$

$$\frac{dI_i^r}{dt} = \alpha \frac{E_i}{Z} - \frac{I_i^r}{D}$$

$$\frac{dI_i^u}{dt} = (1-\alpha)\frac{E_i}{Z} - \frac{I_i^u}{D} + \theta \sum_j \frac{M_{ij} I_j^u}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} I_i^u}{N_i - I_i^u}$$

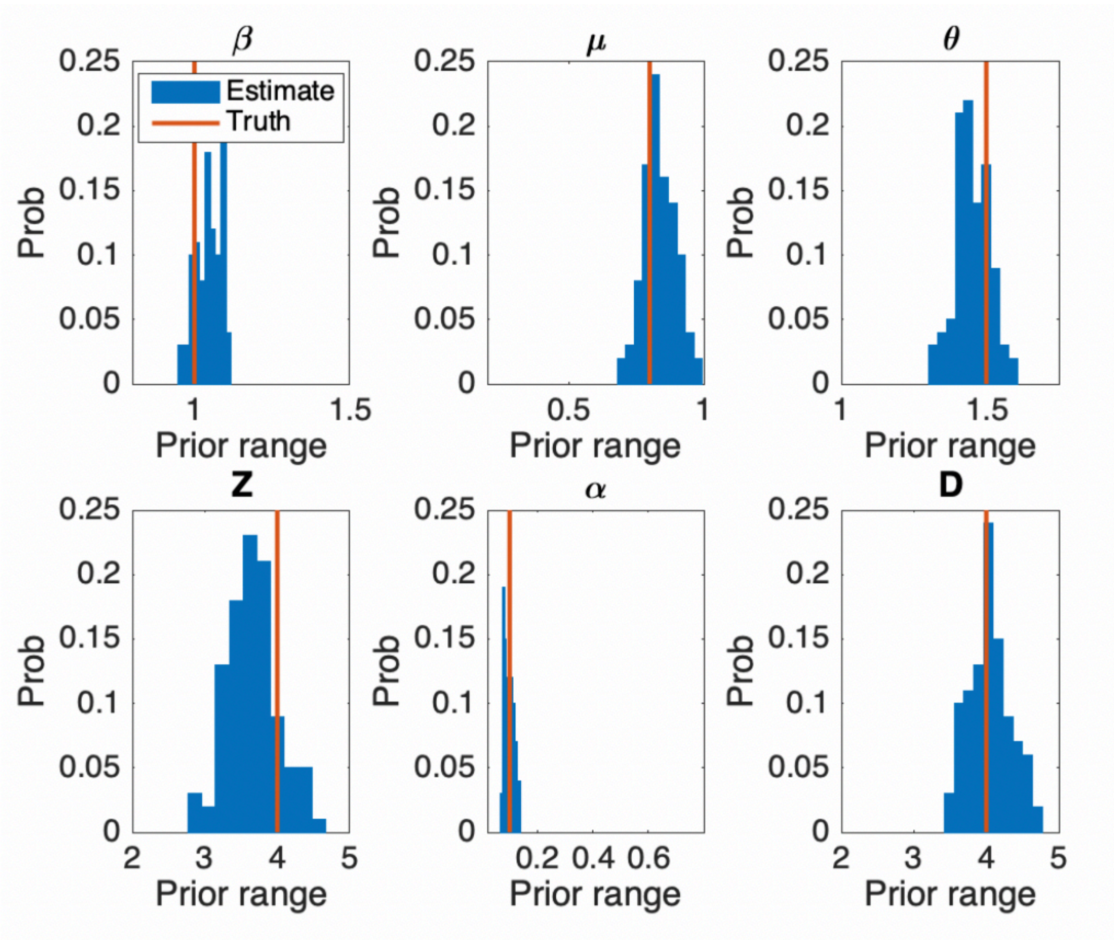
$$N_i = N_i + \theta \sum_j M_{ij} - \theta \sum_j M_{ji}$$

Li et al., 2020

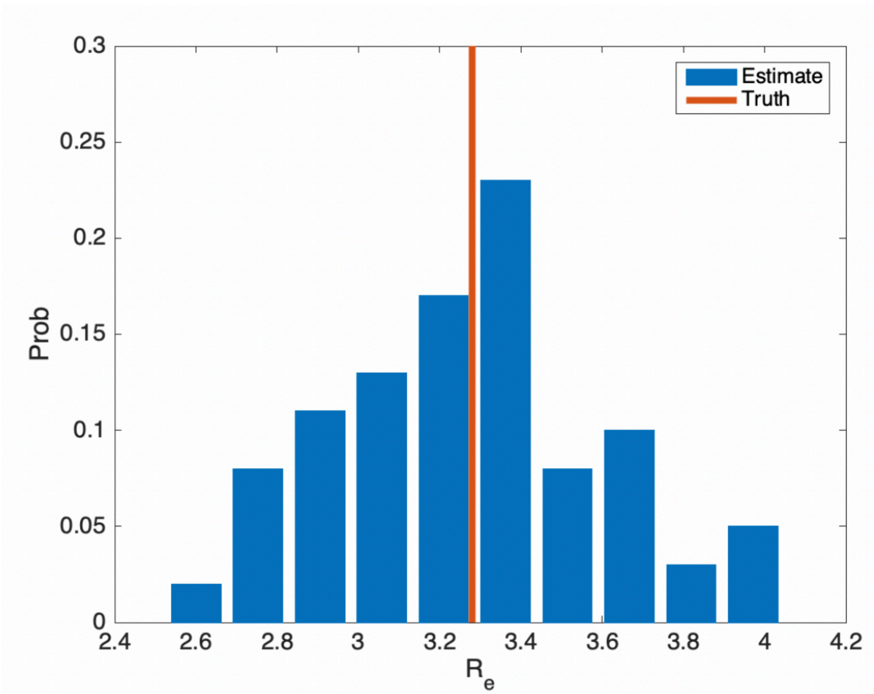
- Simulate January 10-23
- Prior to travel restrictions
- The model separately represents documented and undocumented infections
- The model has a separate contagiousness for documented/ undocumented infections

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?



- Synthetic test of model-inference parameter estimation using model-generated observations



Li et al., 2020

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?

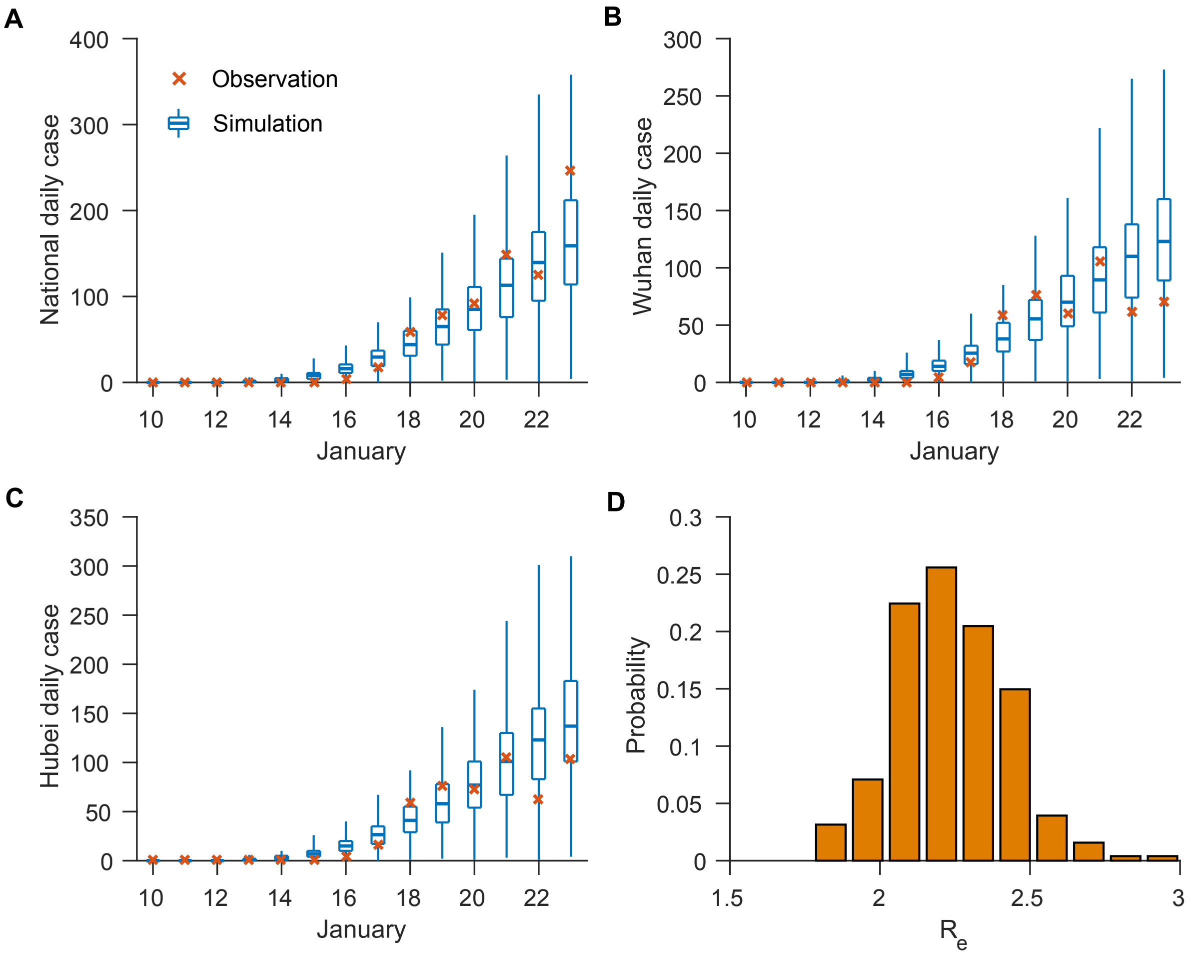
Parameter	Median (95% CIs)
Transmission rate (β, days^{-1})	1.12 (1.04, 1.18)
Relative transmission rate (μ)	0.55 (0.46, 0.62)
Latency period (Z, days)	3.69 (3.28, 4.03)
Infectious period (D, days)	3.48 (3.18, 3.74)
Reporting rate (α)	0.14 (0.10, 0.18)
Basic reproductive number (R_e)	2.38 (2.04, 2.77)
Mobility factor (θ)	1.36 (1.28, 1.43)

- Estimate that 14% of infections are documented
- 86% are undocumented
- Per person, undocumented infections are on average half as contagious (55%) as documented infections
- 2.38 reproductive number

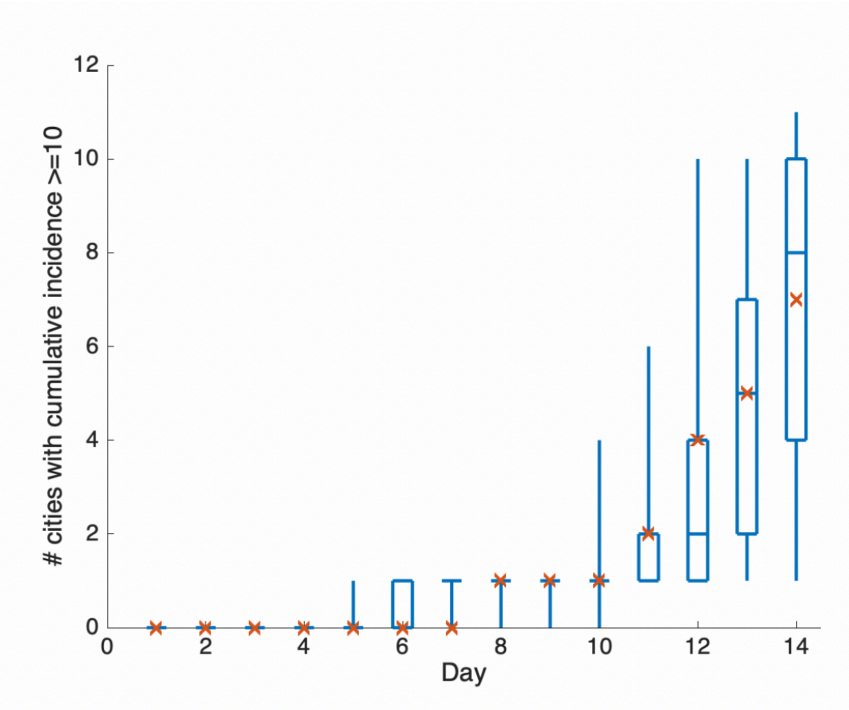
Li et al., 2020

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?



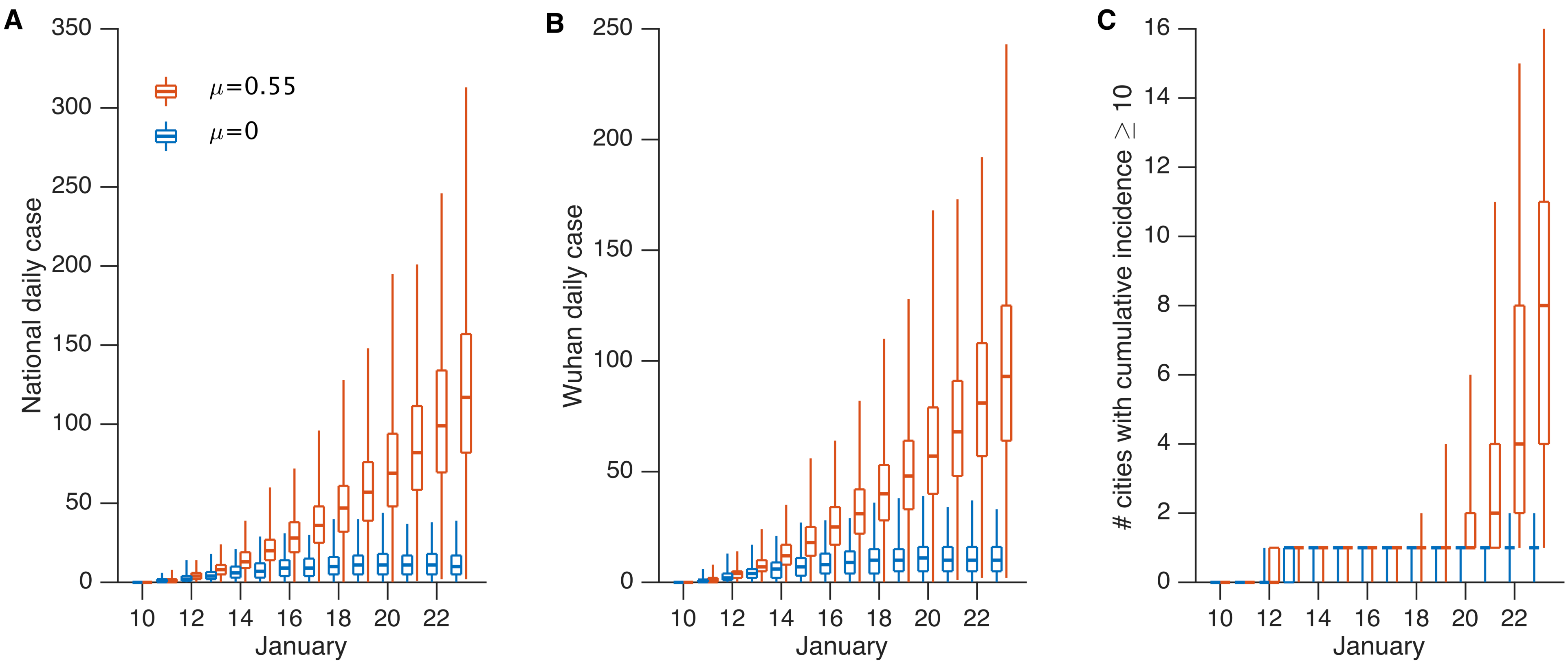
- Simulations with the parameter estimates match the observed outbreak



Li et al., 2020

Inference of Undocumented COVID-19 Infections

Are contagious, undocumented infections supporting the rapid spread of disease?



- Simulations show without transmission from undocumented cases, confirmed cases decrease 79%

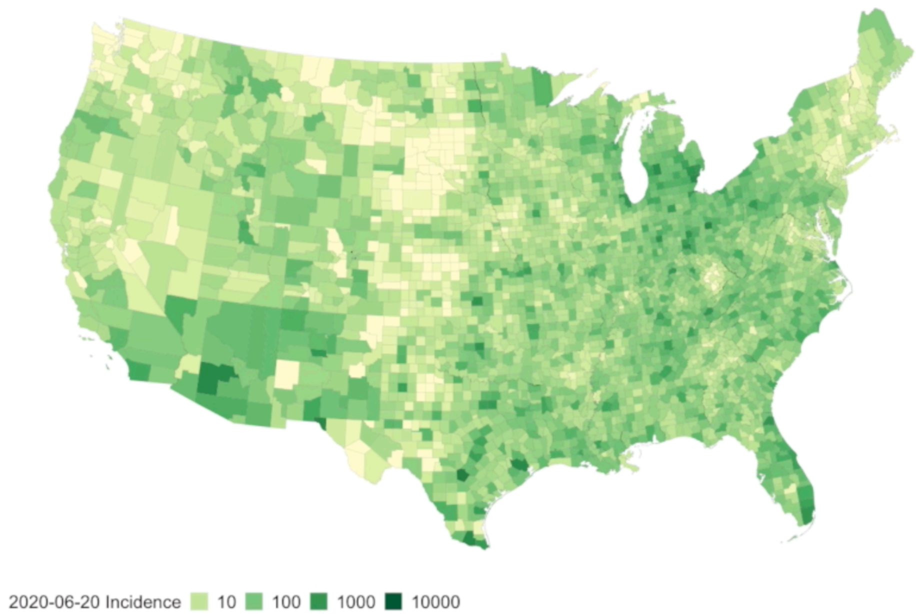
Li et al., 2020

Documentation History of CoV

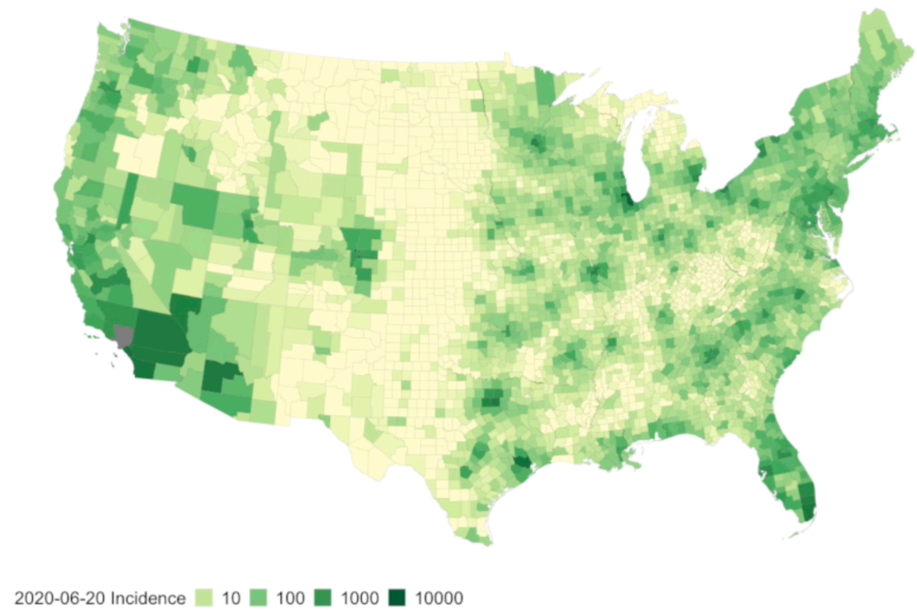
- SARS: sub-clinical infection rates believed to be low (WHO, 2003)
- MERS: 21% of laboratory identified cases were mild or asymptomatic (WHO, 2018)
- Seasonal Coronaviruses (229E, OC43, NL63, HKU1)
  - 135 infection events
  - >60% mild or asymptomatic
  - 4% sought medical care (all had either OC43 or HKU1—the two seasonal betacoronaviruses) (Shaman and Galanti, 2020)
- Our model-inference approach identifies a 14% documentation rate prior to travel restrictions (Li et al. 2020) and indicates that undocumented infections contribute substantially to COVID-19 transmission.

Projections for the US

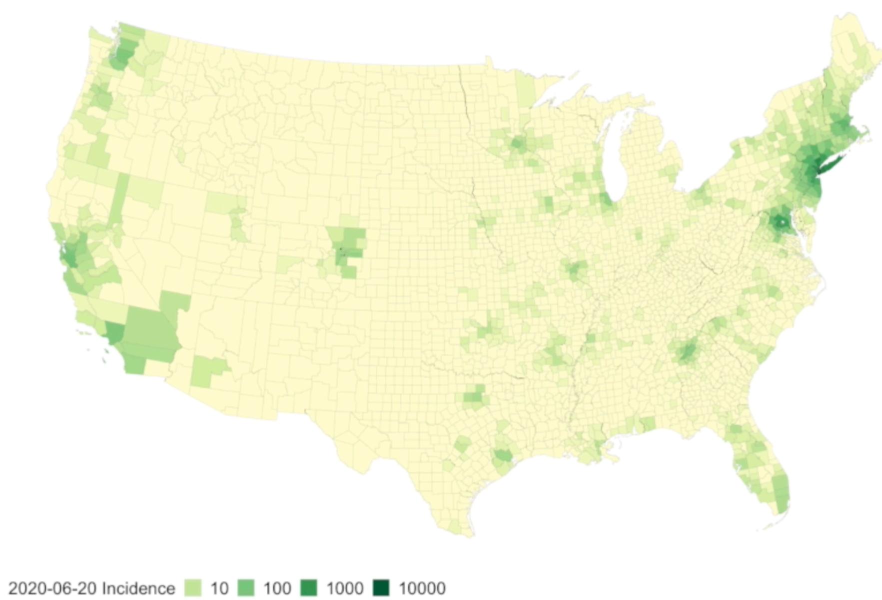
No Control Simulation - June 20, 2020



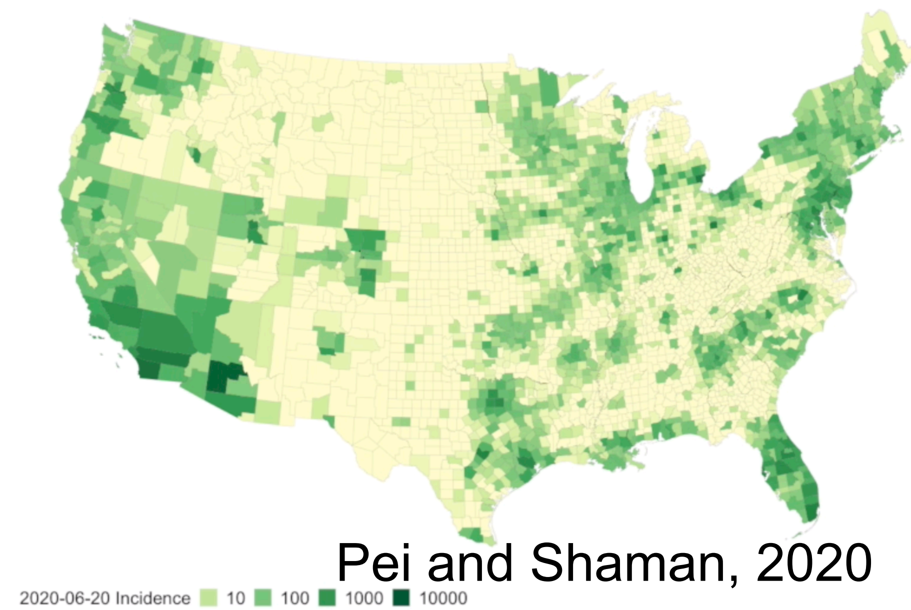
25% Transmission Reduction Simulation - June 20, 2020



50% Transmission Reduction Simulation - June 20, 2020



95% Movement Reduction Simulation - June 20, 2020



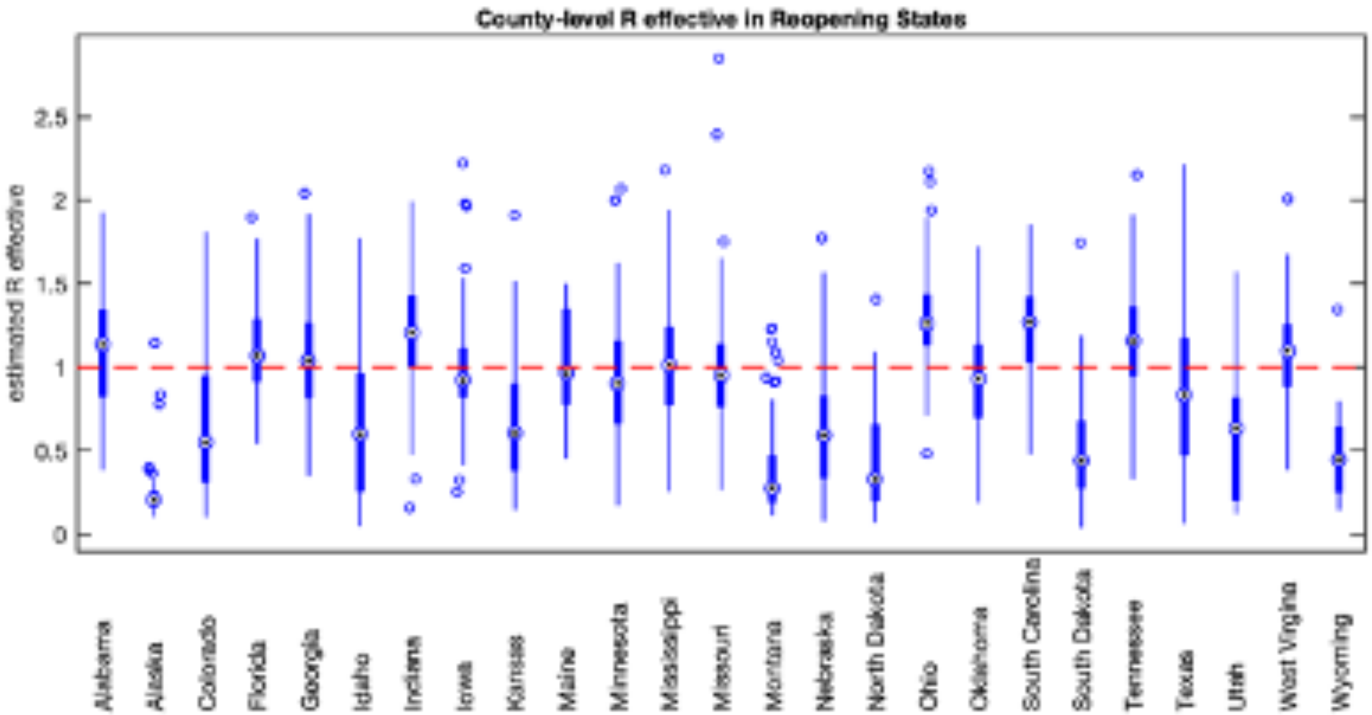
Pei and Shaman, 2020

Initial Estimates for the US (through March 13, 2020)

Parameter	Median (95% CIs)
Transmission rate (β, days^{-1})	0.95 (0.84, 1.06)
Relative transmission rate (μ)	0.64 (0.56, 0.70)
Latency period (Z, days)	3.59 (3.28, 3.99)
Infectious period (D, days)	3.56 (3.21, 3.83)
Reporting rate (α)	0.080 (0.069, 0.093)
Basic reproductive number (R_e)	2.27 (1.87, 2.55)
Mobility factor (θ)	0.15 (0.12,0.17)

Pei and Shaman, 2020

Re-Opening



Effective reproduction number for counties in reopening states as of May 2, 2020.

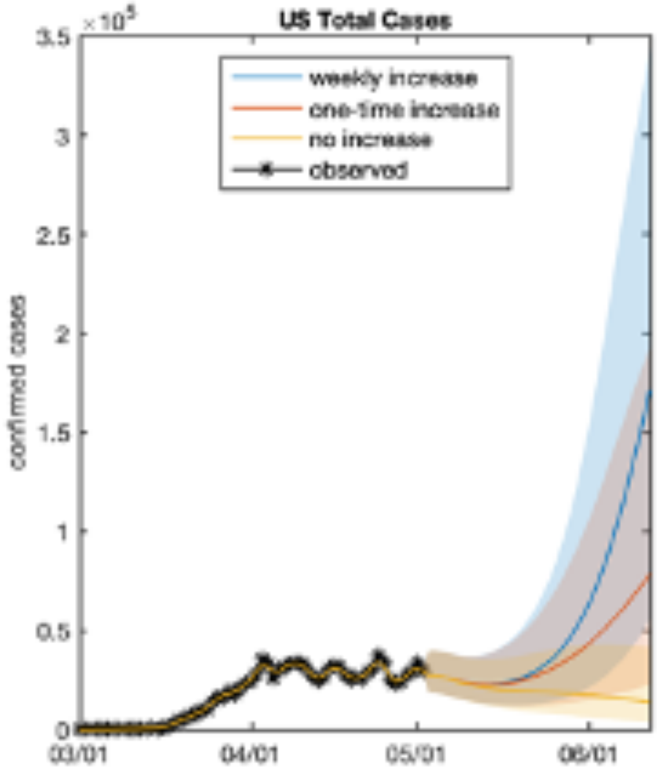
1. Weekly 20% decrease in places with growing weekly cases and a one-time 10% increase in places with return to work (latter supersedes the former)

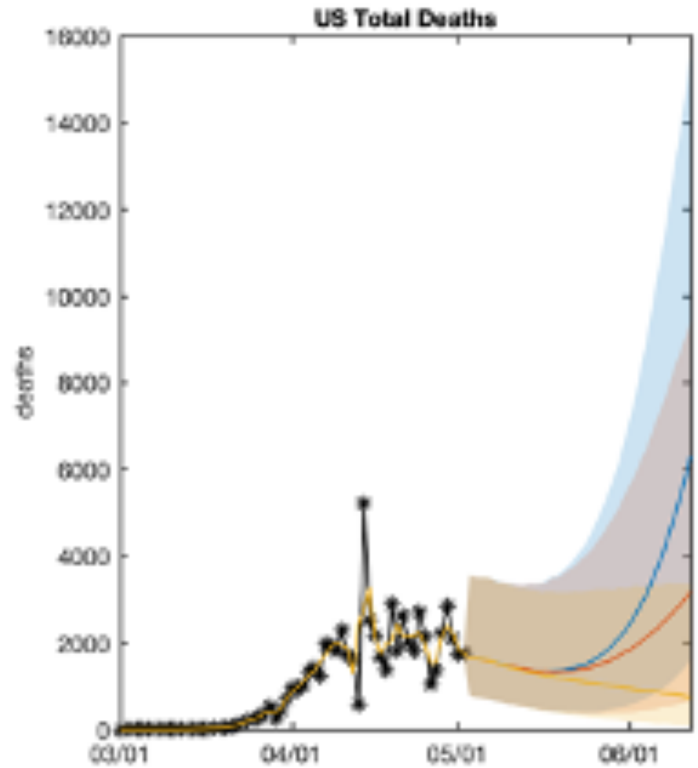
2. Weekly 20% decrease in places with growing weekly cases and a weekly 10% increase in places with return to work (latter supersedes the former)

3. Weekly 20% decrease in places with growing weekly cases

Yamana et al., 2020

Re-Opening





Projections of Re-Opening

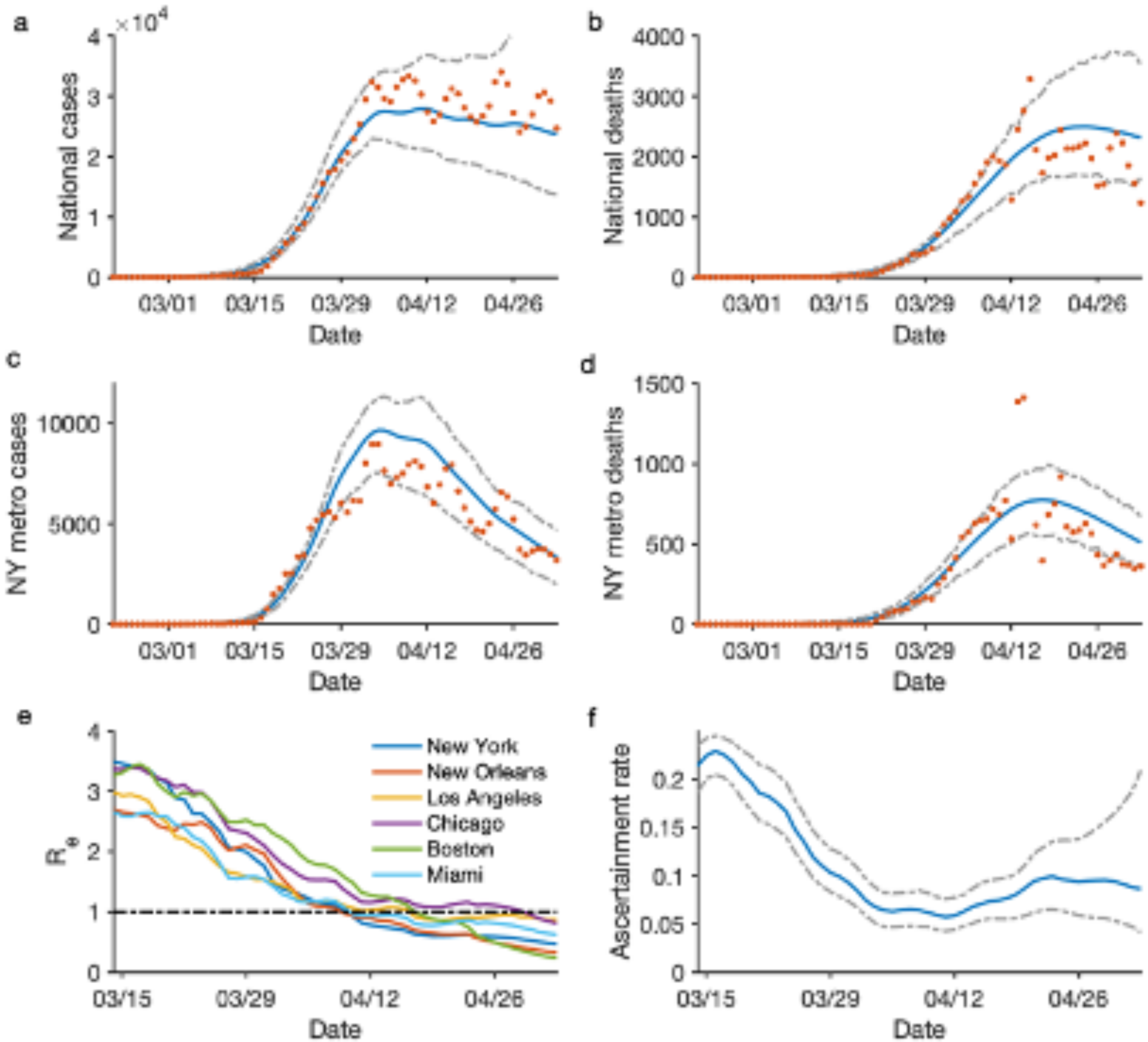
1. Weekly 20% decrease in places with growing weekly cases and a one-time 10% increase in places with return to work (latter supersedes the former)

2. Weekly 20% decrease in places with growing weekly cases and a weekly 10% increase in places with return to work (latter supersedes the former)

3. Weekly 20% decrease in places with growing weekly cases

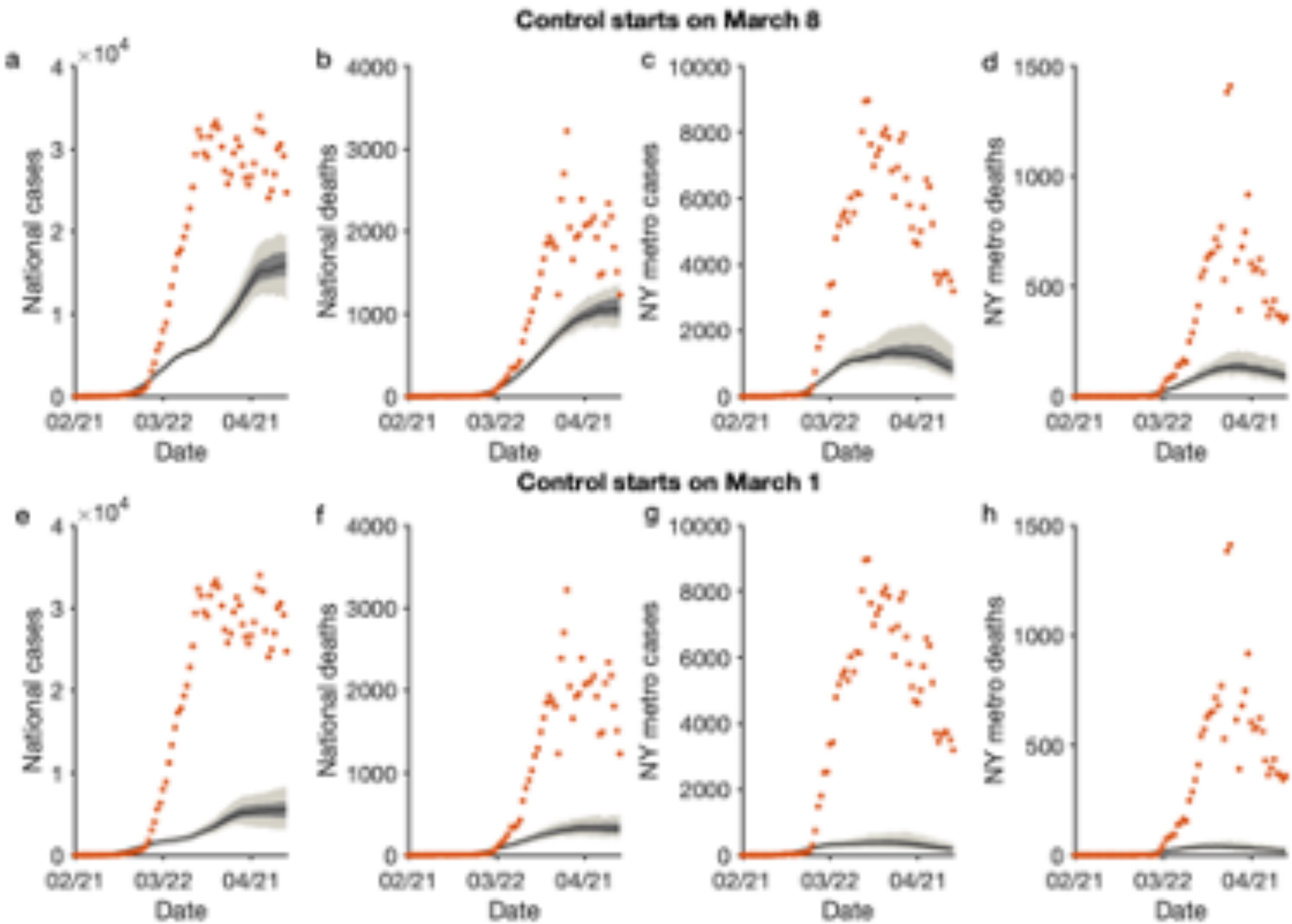
Yamana et al., 2020

Inference, Fitting and Projection



Pei et al., 2020

Counterfactuals



Pei et al., 2020