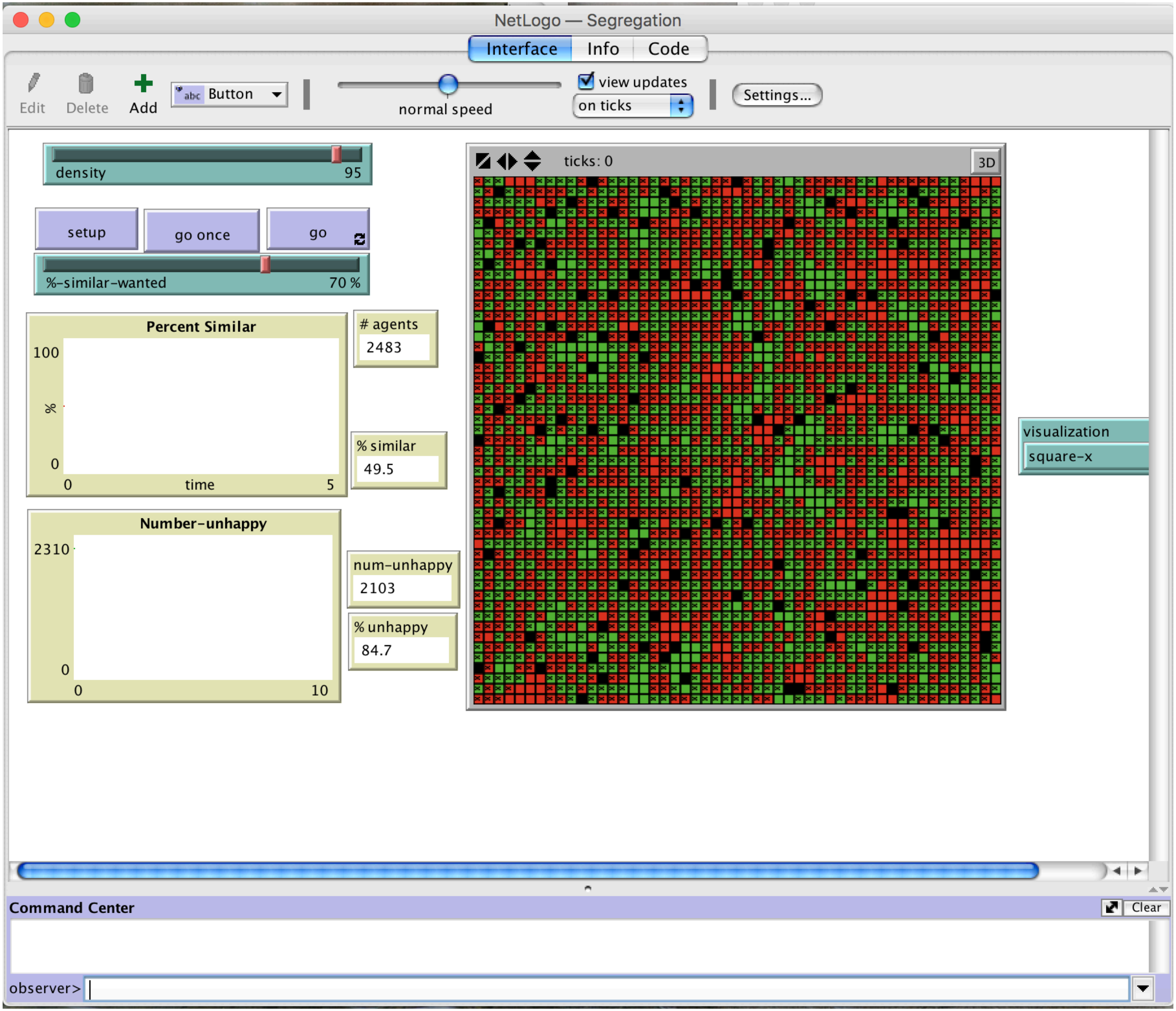
Agent-based Models: Public Health Practice

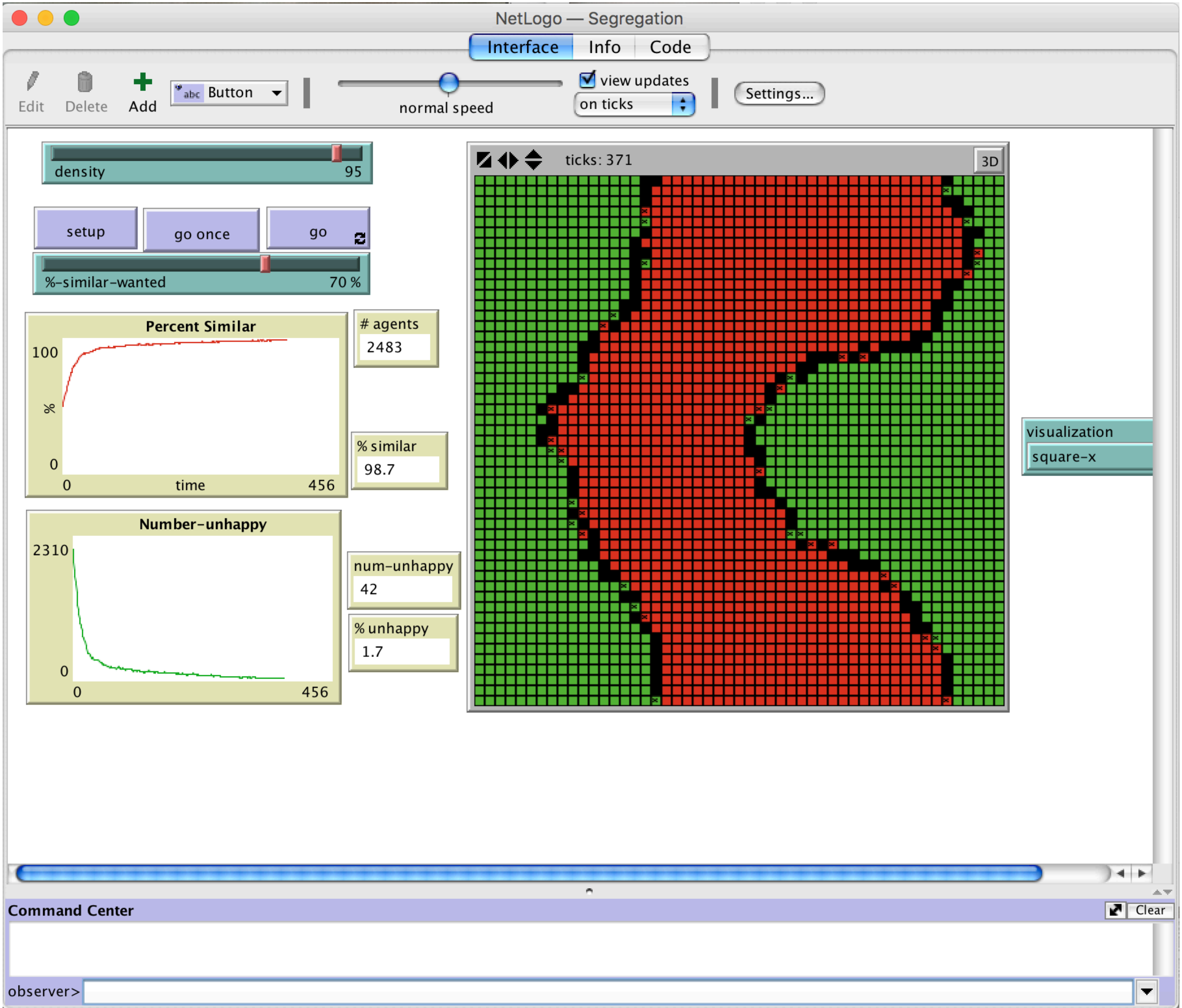
Objectives

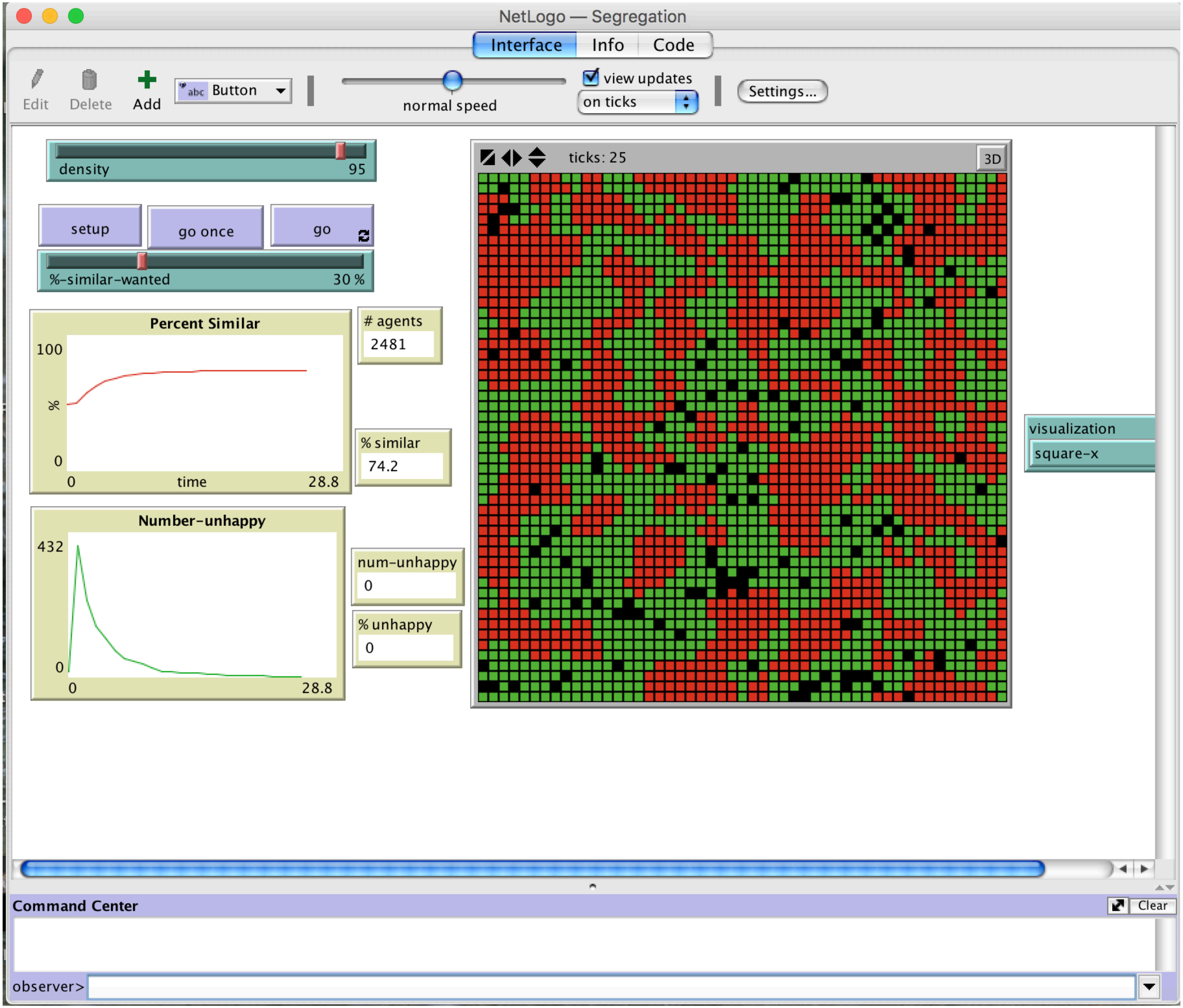
- Some examples of findings from agent-based models
- A worked example of an agent-based modeling project from start to finish(?)
- Please do interrupt, pose questions, etc.

Schelling Segregation Model

- Remarkably simple model of preference for “like living with like”
- Two types of agent
  - Satisfied if at least *p* percent of neighbors are the same as them
  - Dissatisfied if lower, move to a new vacant location


NetLogo — Segregation
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
view updates
on ticks
Settings...
density
95
setup
go once
go
%-similar-wanted
70 %
Percent Similar
100
%
0
0
time
5
agents
2483
% similar
49.5
Number-unhappy
2310
0
0
10
num-unhappy
2103
% unhappy
84.7
ticks: 0
3D
visualization
square-x
Command Center
Clear
observer>


NetLogo — Segregation
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
view updates
on ticks
Settings...
density
95
setup
go once
go
%-similar-wanted
70 %
ticks: 371
3D
Percent Similar
100
%
0
time
456
agents
2483
% similar
98.7
Number-unhappy
2310
0
456
num-unhappy
42
% unhappy
1.7
visualization
square-x
Command Center
Clear
observer>



NetLogo — Segregation
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
view updates
on ticks
Settings...
density
95
setup
go once
go
%-similar-wanted
30 %
Percent Similar
100
%
0
0
time
28.8
agents
2481
% similar
74.2
Number-unhappy
432
0
0
28.8
num-unhappy
0
% unhappy
0
ticks: 25
3D
visualization
square-x
Command Center
Clear
observer>

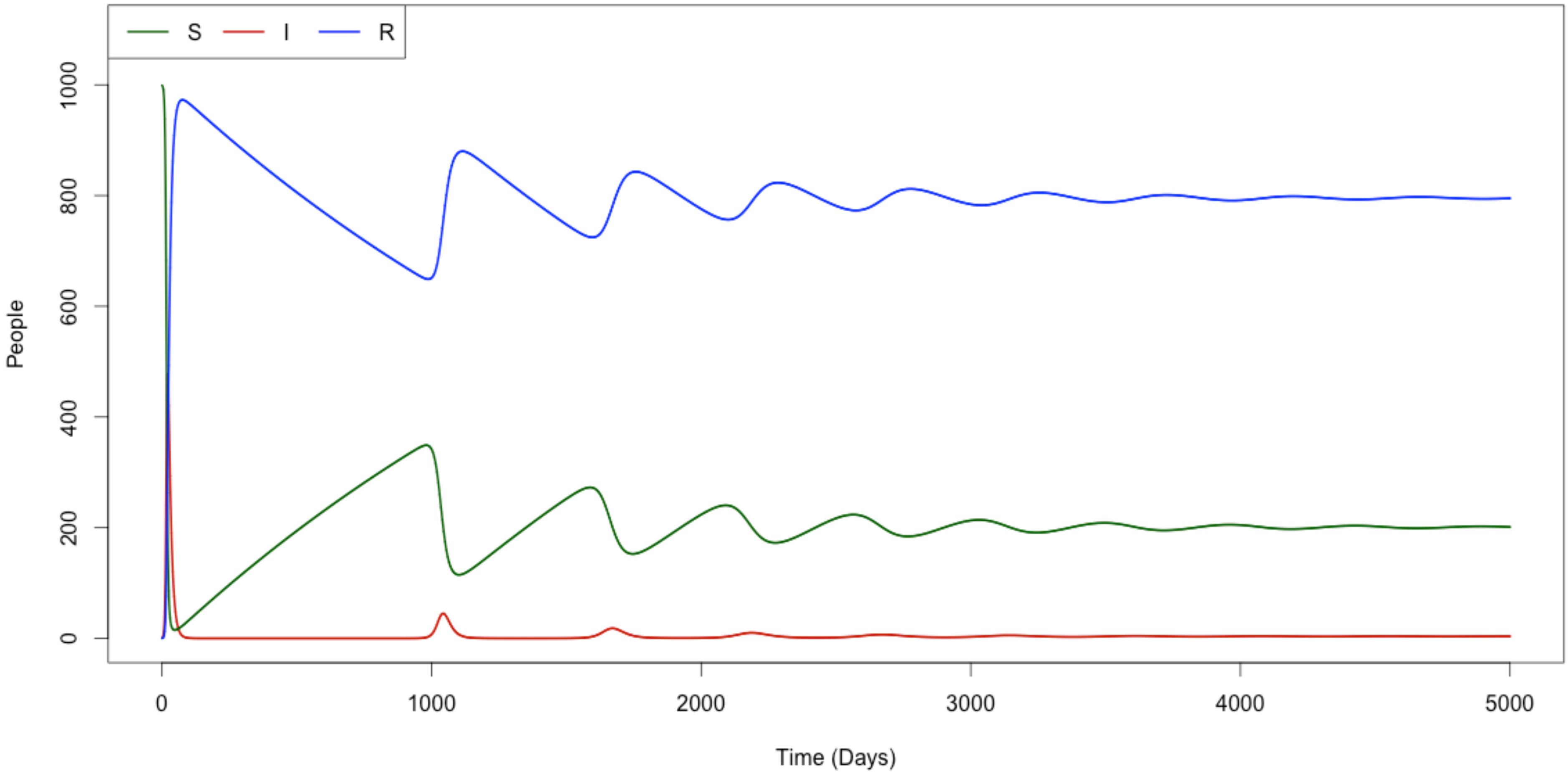
What Do We Get From This?

- Mild amounts of assortative preference can result in a highly segregated population
- “I’m willing to live in a neighborhood where 70% of my neighbors look not like me” seems like it would result in lots of diversity

Simple SIR Model

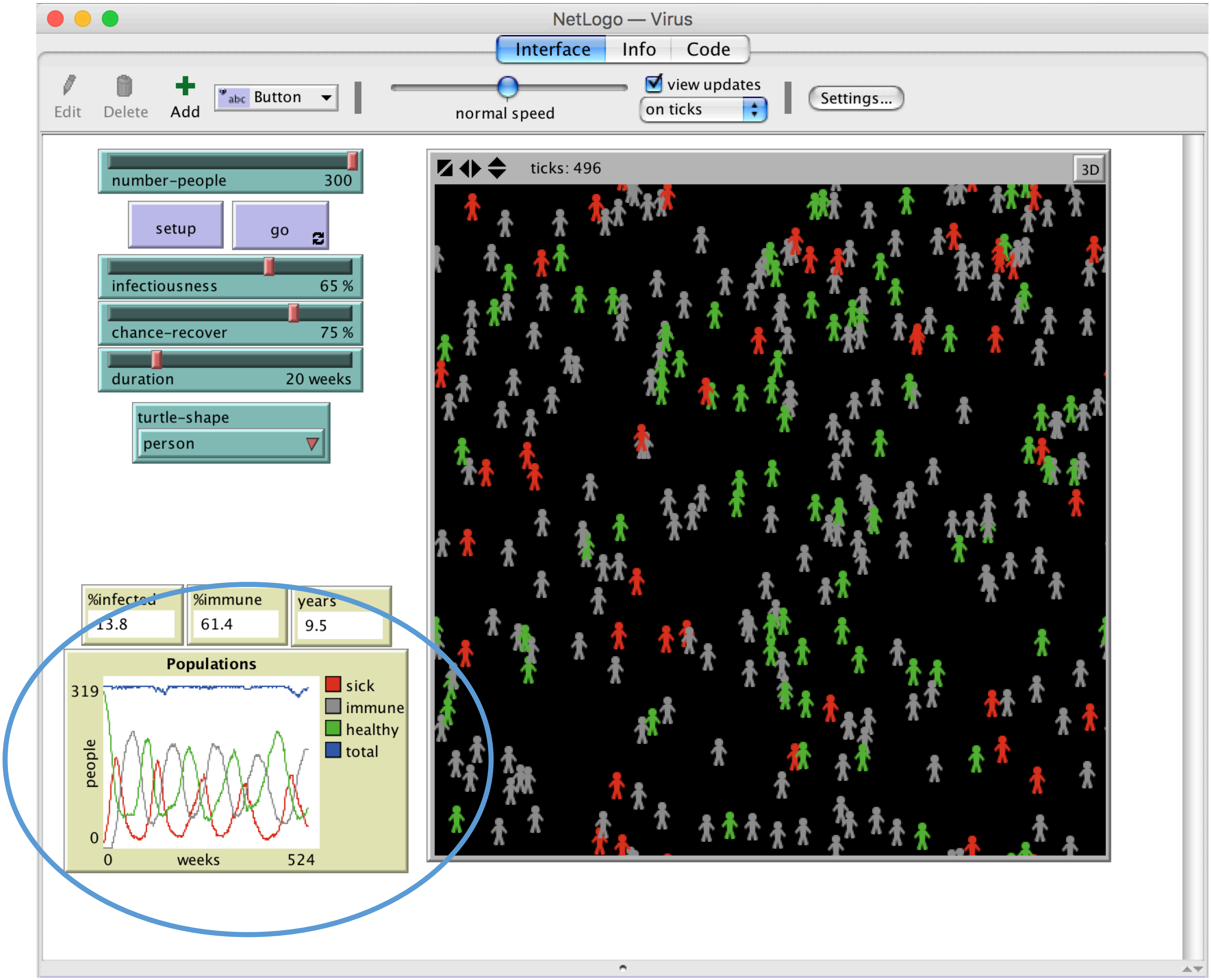
- Lets consider the standard “Susceptible – Infected – Recovered” model in mathematical epidemiology
- Add birth and death
- What does an agent-based form of this model look like?
- What does it do?

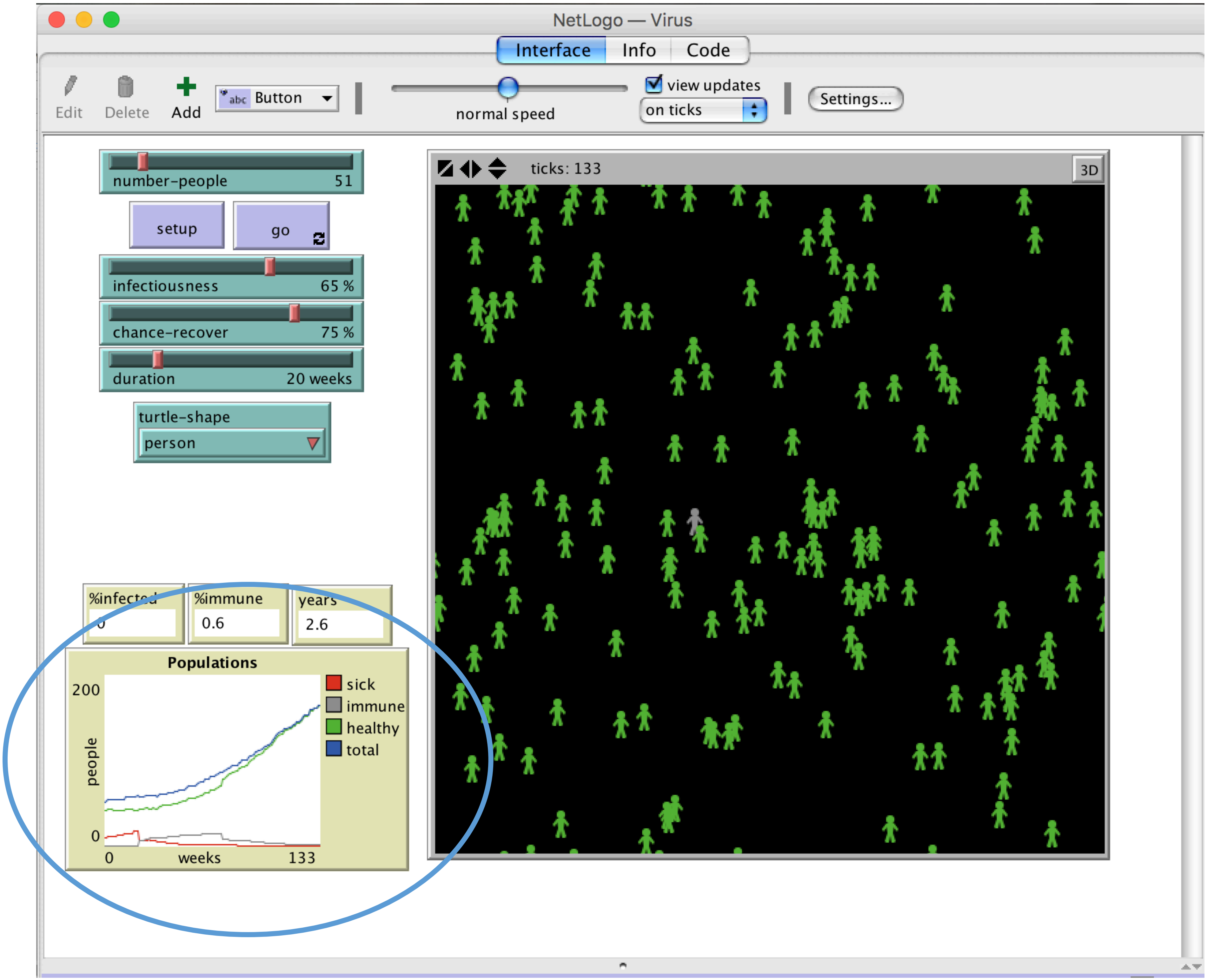
Deterministic Form



Damped oscillations

Hugely exaggerated birth and death rate in this example


NetLogo — Virus
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
view updates
on ticks
Settings...
number-people
300
setup
go
infectiousness
65 %
chance-recover
75 %
duration
20 weeks
turtle-shape
person
ticks: 496
3D
%infected
13.8
%immune
61.4
years
9.5
Populations
319
0
people
0
weeks
524
sick
immune
healthy
total


NetLogo — Virus
Interface
Info
Code
Edit
Delete
Add
Button
normal speed
view updates
on ticks
Settings...
number-people
51
setup
go
infectiousness
65 %
chance-recover
75 %
duration
20 weeks
turtle-shape
person
ticks: 133
3D
%infected
%immune
0.6
years
2.6
Populations
200
0
people
0
weeks
133
sick
immune
healthy
total

What Happened!?

- ”Stochastic extinction”
- Many deterministic models have populations that asymptotically approach zero, but are never actually zero
- There’s always some small fraction of infected individuals
  - “Atto-fox Problem” – 10^{-18} foxes
- For a large population, this might not be a big deal
- For a small population?
- This may have a huge impact on disease dynamics
- Elucidating if stochastic extinction if common or rare requires many simulations

Ebola Modeling

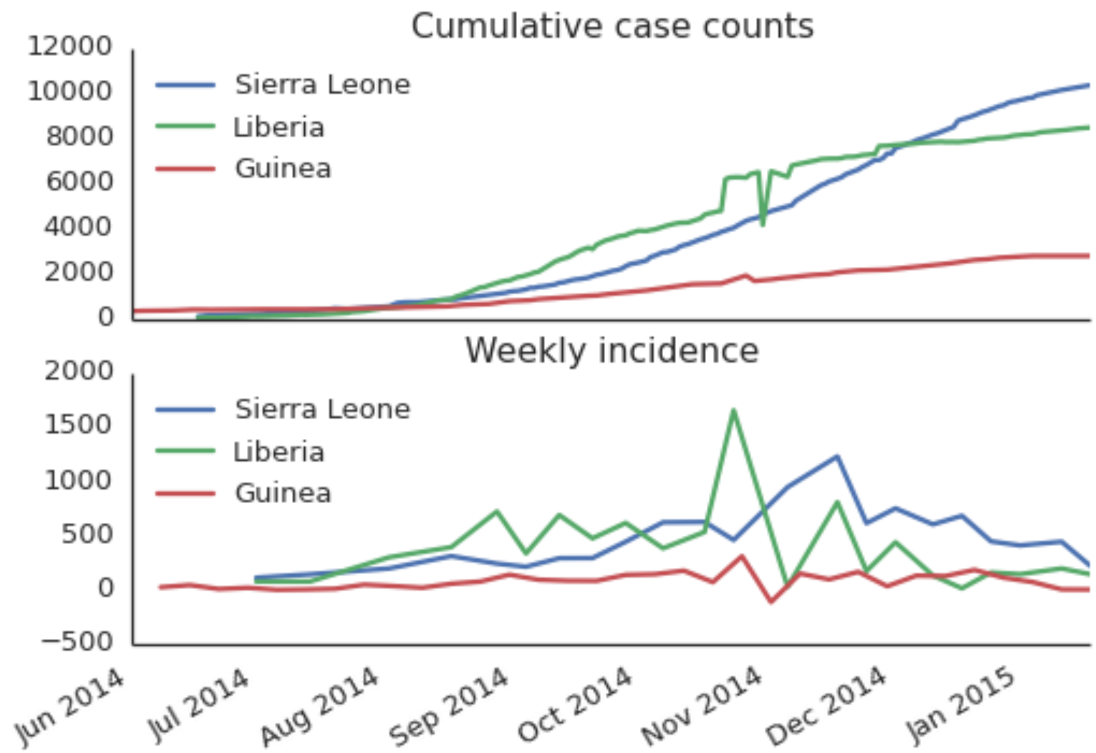
- West African Ebola outbreak the largest one in history by far
- NDSSL tasked by several groups with providing forecasts and intervention evaluation
- Using this experience as a “worked example” of the agent-based modeling enterprise
- “The Story of my Postdoc”

NDSSL Ebola Modeling Team

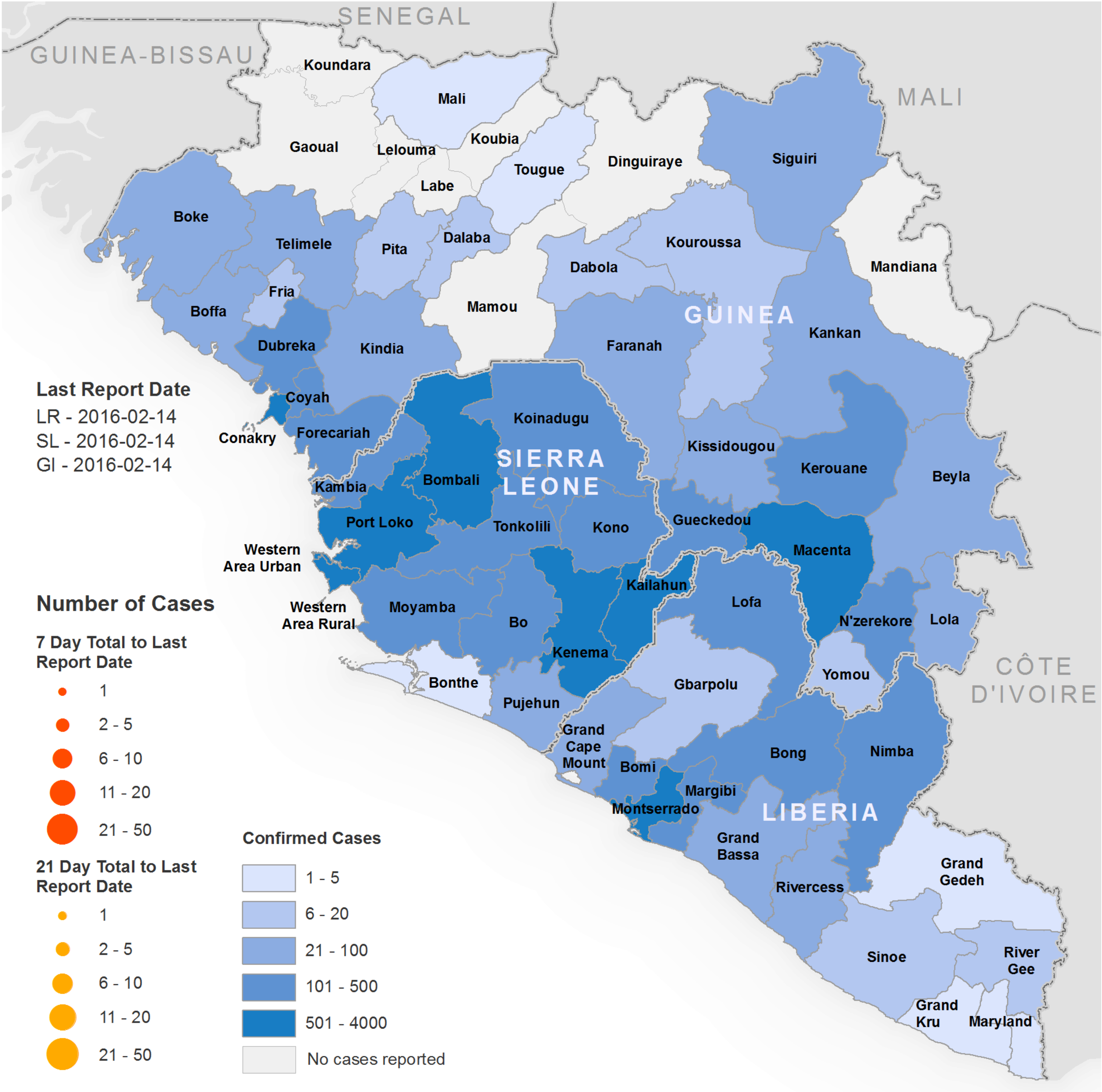
Staff: Abhijin Adiga, Kathy Alexander, Chris Barrett, Richard Beckman, Keith Bisset, Jiangzhuo Chen, Youngyoun Chungbaek, Stephen Eubank, Sandeep Gupta, Maleq Khan, Chris Kuhlman, Eric Lofgren, Bryan Lewis, Achla Marathe, Madhav Marathe, Henning Mortveit, Eric Nordberg, Paula Stretz, Samarth Swarup, Meredith Wilson,Mandy Wilson, and Dawen Xie, with support from Ginger Stewart, Maureen Lawrence-Kuether, Kayla Tyler, Kathy Laskowski, Bill Marmagas

Students: S.M. Arifuzzaman, Aditya Agashe, Vivek Akupatni, Caitlin Rivers, Pyrros Telionis, Jessie Gunter, Elisabeth Musser, James Schlitt, Youssef Jemia, Margaret Carolan, Bryan Kaperick, Warner Rose, Kara Harrison

The Ebola Epidemic



	Cases	**Deaths**
Guinea	3804	2536
Liberia	10666	4806
Sierra Leone	14122	3955
Worldwide	**28603**	**11301**



SENEGAL
GUINEA-BISSAU
MALI
GUINEA
SIERRA LEONE
LIBERIA
CÔTE D'IVOIRE
Koundara
Mali
Gaoual
Lelouma
Koubia
Tougue
Dinguiraye
Siguiri
Labe
Boke
Telimele
Pita
Dalaba
Kouroussa
Mandiana
Dabola
Fria
Boffa
Mamou
Kankan
Dubreka
Kindia
Faranah
Coyah
Conakry
Forecariah
Koinadugu
Kissidougou
Kerouane
Beyla
Kambia
Bombali
Port Loko
Tonkolili
Kono
Gueckedou
Macenta
Western Area Urban
Western Area Rural
Kailahun
Lofa
Moyamba
Bo
N'zerekore
Lola
Kenema
Yomou
Bonthe
Gbarpolu
Pujehun
Grand Cape Mount
Bomi
Bong
Nimba
Margibi
Montserrado
Grand Bassa
Rivercess
Grand Gedeh
Sinoe
River Gee
Grand Kru
Maryland
Last Report Date
LR - 2016-02-14
SL - 2016-02-14
GI - 2016-02-14
Number of Cases
7 Day Total to Last Report Date
1
2 - 5
6 - 10
11 - 20
21 - 50
21 Day Total to Last Report Date
1
2 - 5
6 - 10
11 - 20
21 - 50
Confirmed Cases
1 - 5
6 - 20
21 - 100
101 - 500
501 - 4000
No cases reported

Two Major Questions

- What is going to happen?
- Why is this happening?

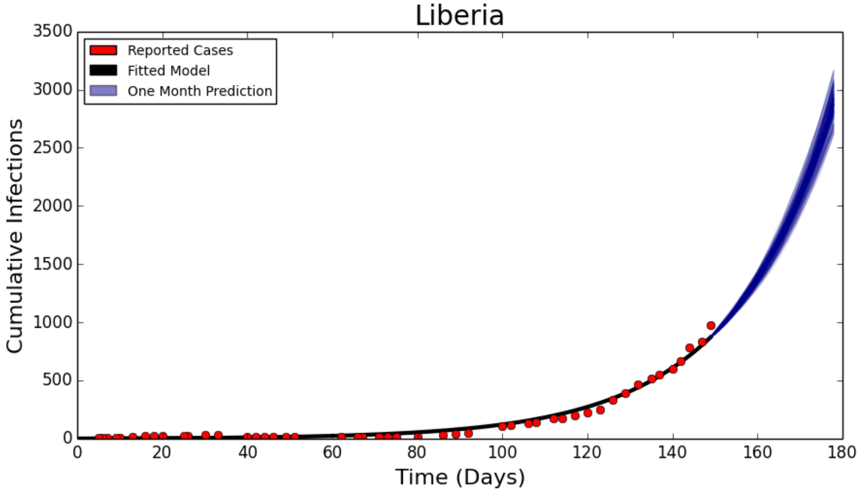
What Is Going to Happen?

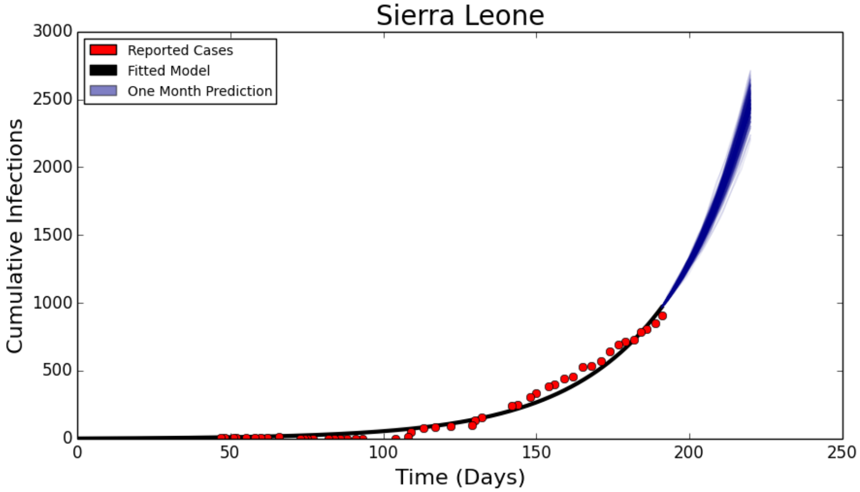
- What is the question or system you want to model?
  - We need a forecast for Ebola transmission in West Africa to determine where the epidemic is headed in both space and time
- Why does it need to be modeled?
  - Urgent public health need, statistical forecasts might not be able to capture the full dynamics of the infection
- What kind of model does it need?
  - A spatial model would be nice, and detail would be nice, but isn't essential
- How fast do you need an answer?
  - **Forecast turnaround times on the order of days to "a week".**


STOP
FreeFoto.com

Fast Prototype Models for Forecasting

- Agent-based model was not yet ready, would take a fair amount of time to build a working population of West Africa and get it to fit to data
- Already under fairly high demand doing briefings, conference calls, etc.
- A stochastic compartmental model was used instead to give rapid answers to questions while the agent-based model was developed.





Why Is This Happening?

- The situation in West Africa is considerably worse than any other historical Ebola outbreak
  - Ugandan outbreak had 425 human cases
  - 67X difference in magnitude
  - Widespread dissemination of disease
- "Something must be different"

Viral Evolution

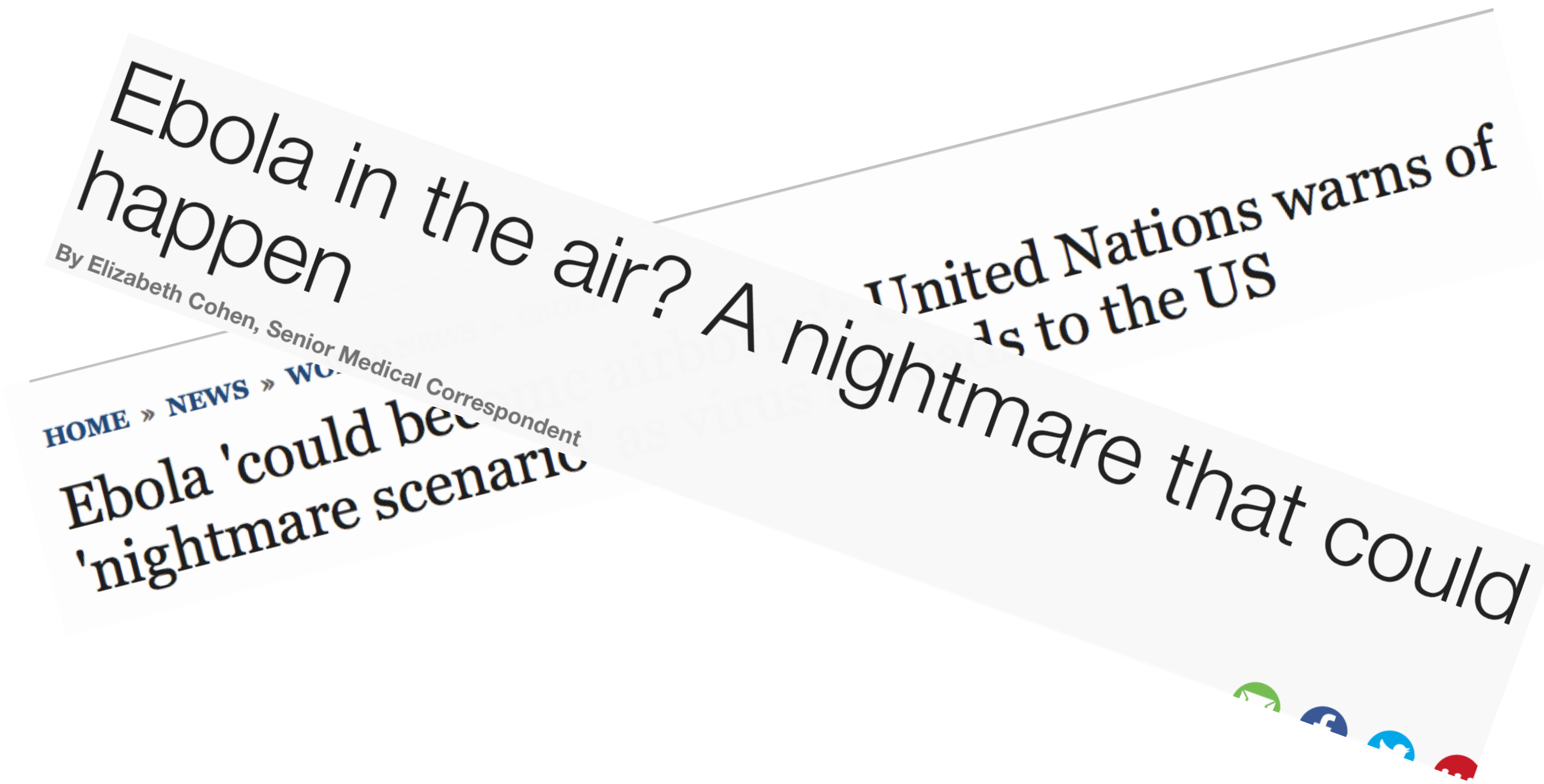
- Is the virus more transmissible? More virulent? Is Ebola airborne?

HOME » NEWS » WORLD NEWS » EBOLA

Ebola 'could become airborne': United Nations warns of 'nightmare scenario' as virus spreads to the US

Viral Evolution

- Is the virus more transmissible? More virulent? Is Ebola airborne?



Viral Evolution

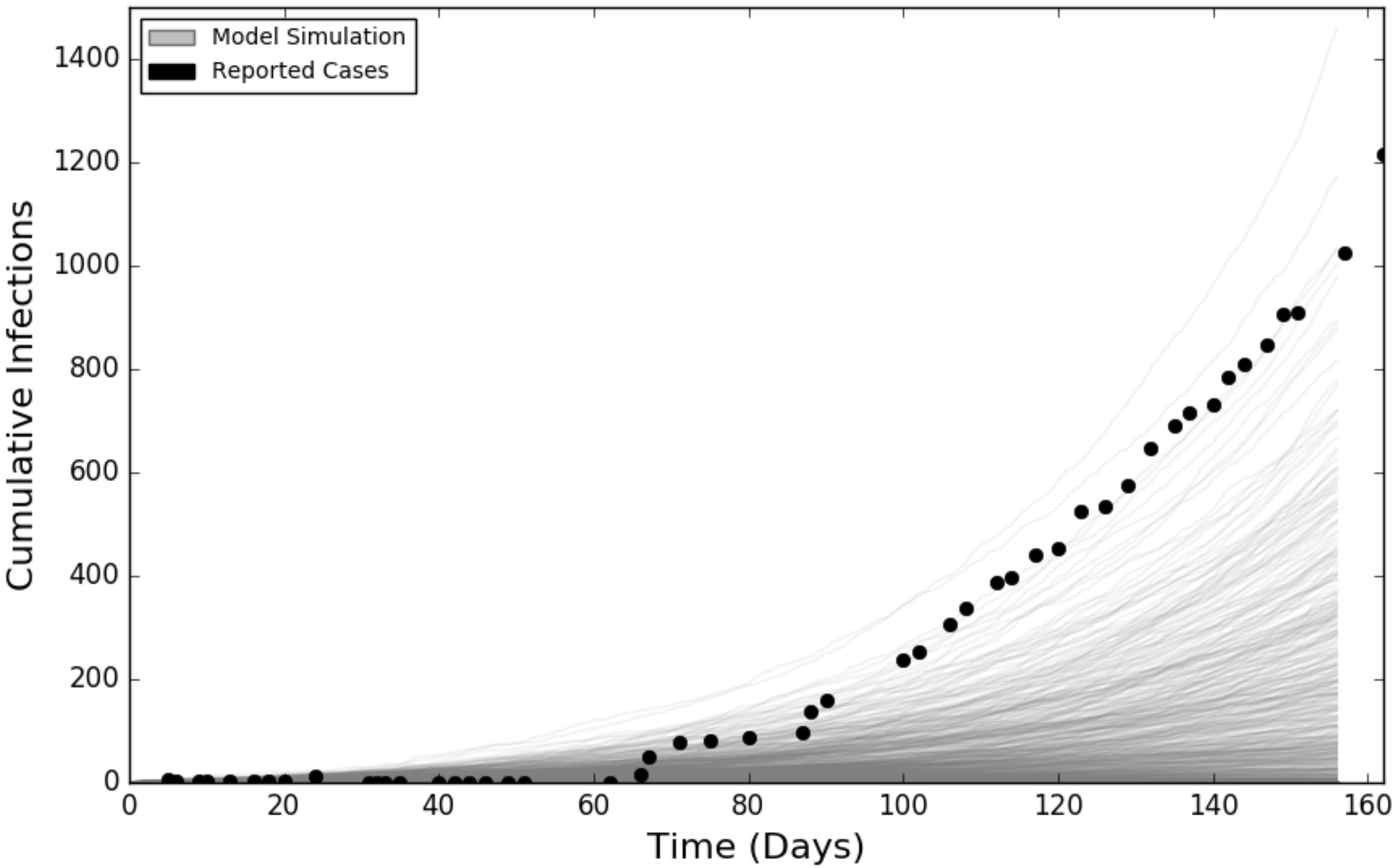
- Is the virus more transmissible? More virulent? Is Ebola airborne?

Population Mobility

- Little is known about the mobility patterns in West Africa
- Believed to be very high – displaced persons, porous borders, informal transport networks

H_0: Maybe This Is Just Bad Luck

- Unstated alternative to the two previous hypotheses, as well as a number of others
- Stochastic systems can have high variability – this is possibly just bad luck
- Result of an idle thought experiment in an early forecast model



Model Simulation
Reported Cases
Cumulative Infections
Time (Days)
0
200
400
600
800
1000
1200
1400
0
20
40
60
80
100
120
140
160

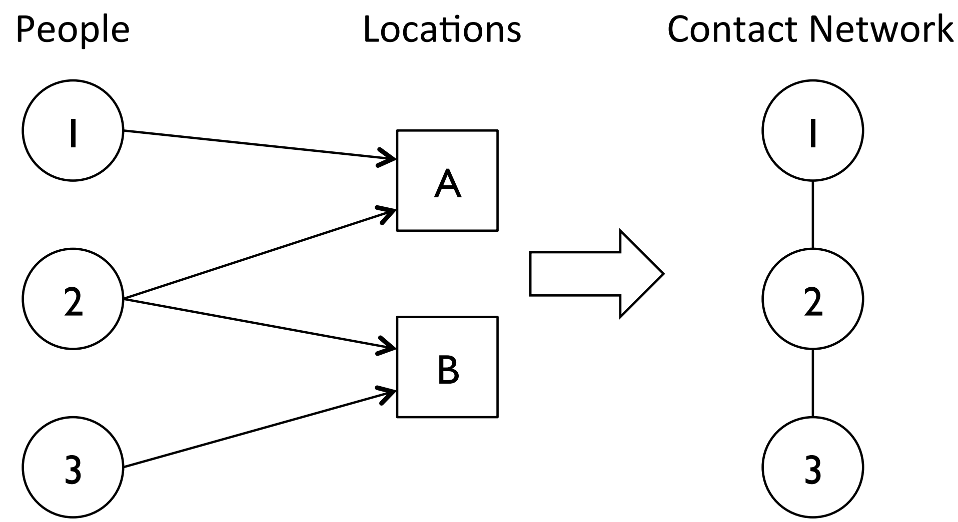
Role of Change

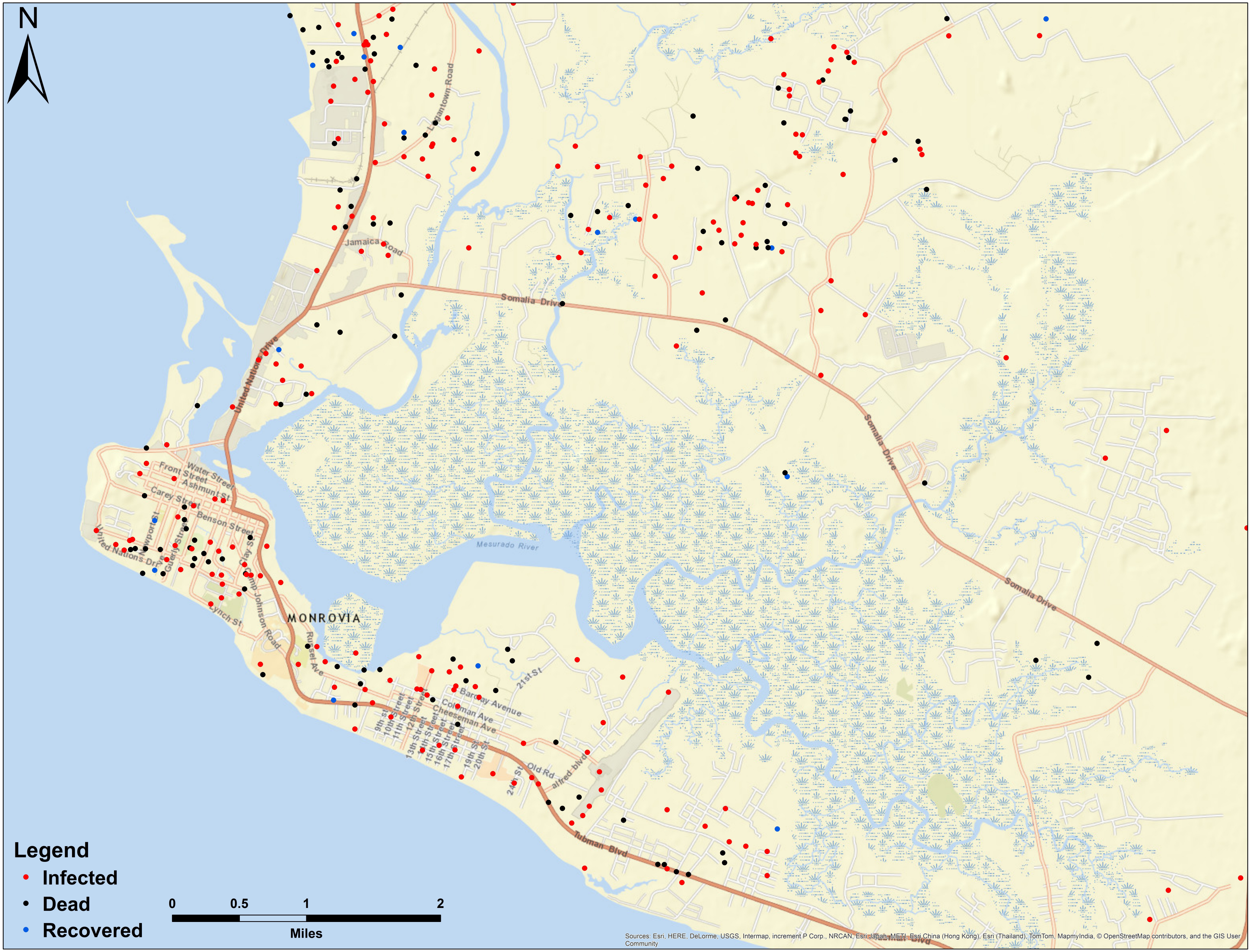
- Stochastic forecast model appears to show the current epidemic as a “bad case”, but not the worst case.
- The parameters that give us the model in the forecast also appear to yield much less severe epidemics
- Can a parameter set that explains *this* epidemic also explain *all other* epidemics?

- What is the question or system you want to model?
  - We need to assess the stochastic variability of the current Ebola epidemic
- Why does it need to be modeled?
  - It's impossible to collect real data on thousands of Ebola epidemics in West Africa
- What kind of model does it need?
  - If we need to test movement patterns, an agent-based model is the most intuitive way to do that
- How fast do you need an answer?
  - This is a research question, and can be done at the standard pace of research

Creating a Synthetic Population

- NDSSL uses agent-based models to build “synthetic populations” that use large amounts of data to build an artificial but realistic population
- This population then goes about their lives, and the simulation records what individuals are in the same place at the same time
- This yields a weighted contact network that disease can be simulated spreading across




N
Legend
Infected
Dead
Recovered
0
0.5
1
2
Miles
MONROVIA
Mesurado River
Somalia Drive
United Nations Drive
Jamaica Road
Tubman Blvd
Barclay Avenue
Coleman Ave
Cheeseman Ave
Water Street
Front Street
Ashmun St
Carey Street
Benson Street
Lynch St
Camp Johnson Road
Russel Ave
21st St
Old Rd
Sources: Esri, HERE, DeLorme, USGS, Intermap, increment P Corp., NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), TomTom, MapmyIndia, © OpenStreetMap contributors, and the GIS User Community

Synthetic Populations

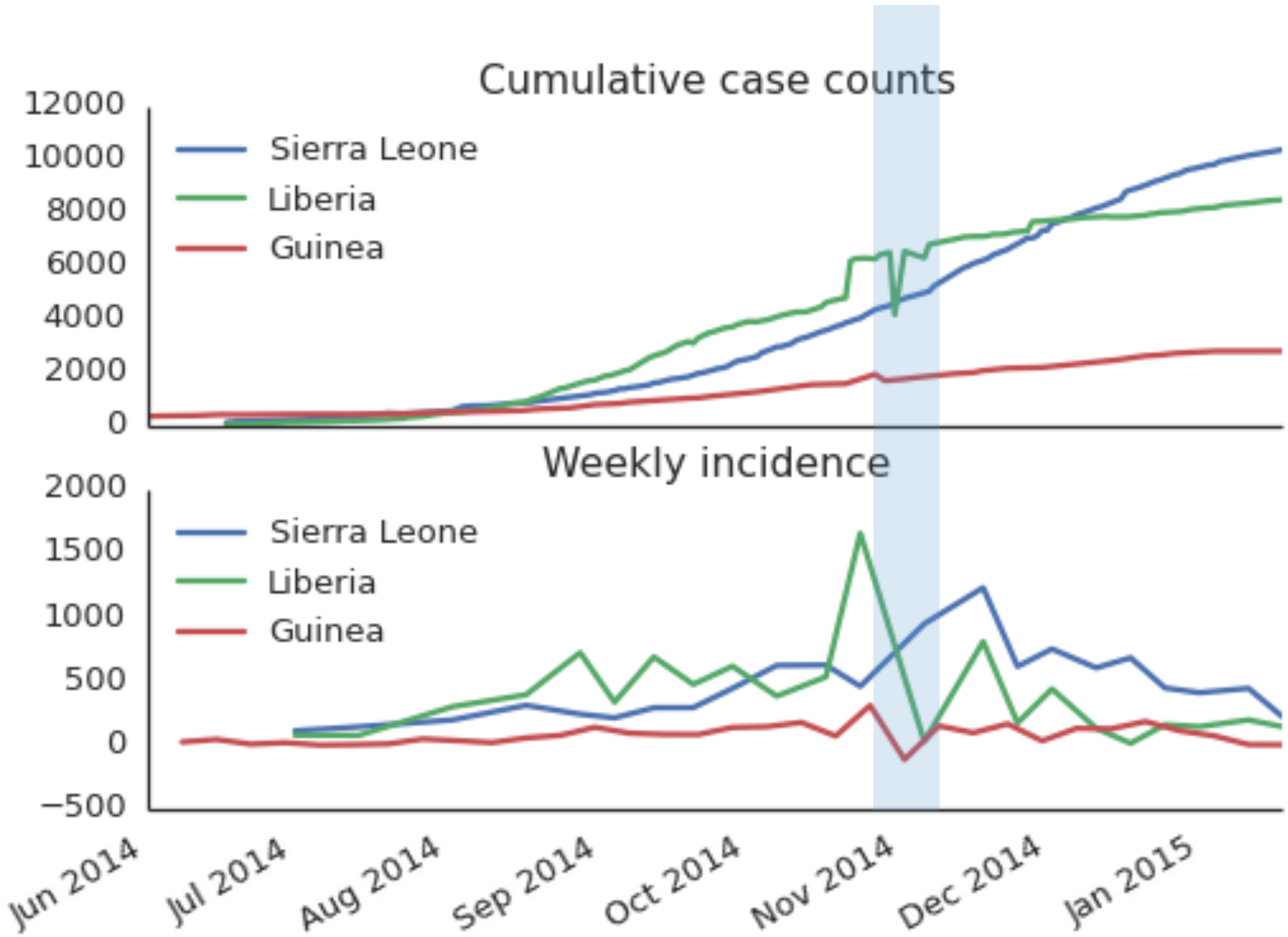
- All of the synthetic populations made for West Africa are available at:
  - https://www.bi.vt.edu/ndssl/featured-projects/ebola
  - Visualization tools available from the same site

Why Do It This Way

- Decouples the disease transmission process from the simulation of the population
  - You can overlay multiple diseases on the same network
  - You can exploit faster network simulation algorithms to speed up computation
  - Applications beyond disease modeling
    - e.g. Environmental epidemiology studies – we have latitude and longitude to combine with exposure data for an entire synthetic population
  - Can be done *before* a specific disease model is needed

The Problem of Data

- West Africa is not the ideal place to conduct research
- Ebola data is past “noisy” and into very, very messy
  - Delays in reporting
  - Temporal effects in reporting – large surges of cases followed by lulls
  - Revision of previous errors occasionally results in negative incidence
  - Data from multiple sources, many people keeping detailed line listings in-house
- Very labor intensive collection of reported cases from WHO, Ministries of Health, etc.
  - https://github.com/cmrivers/ebola



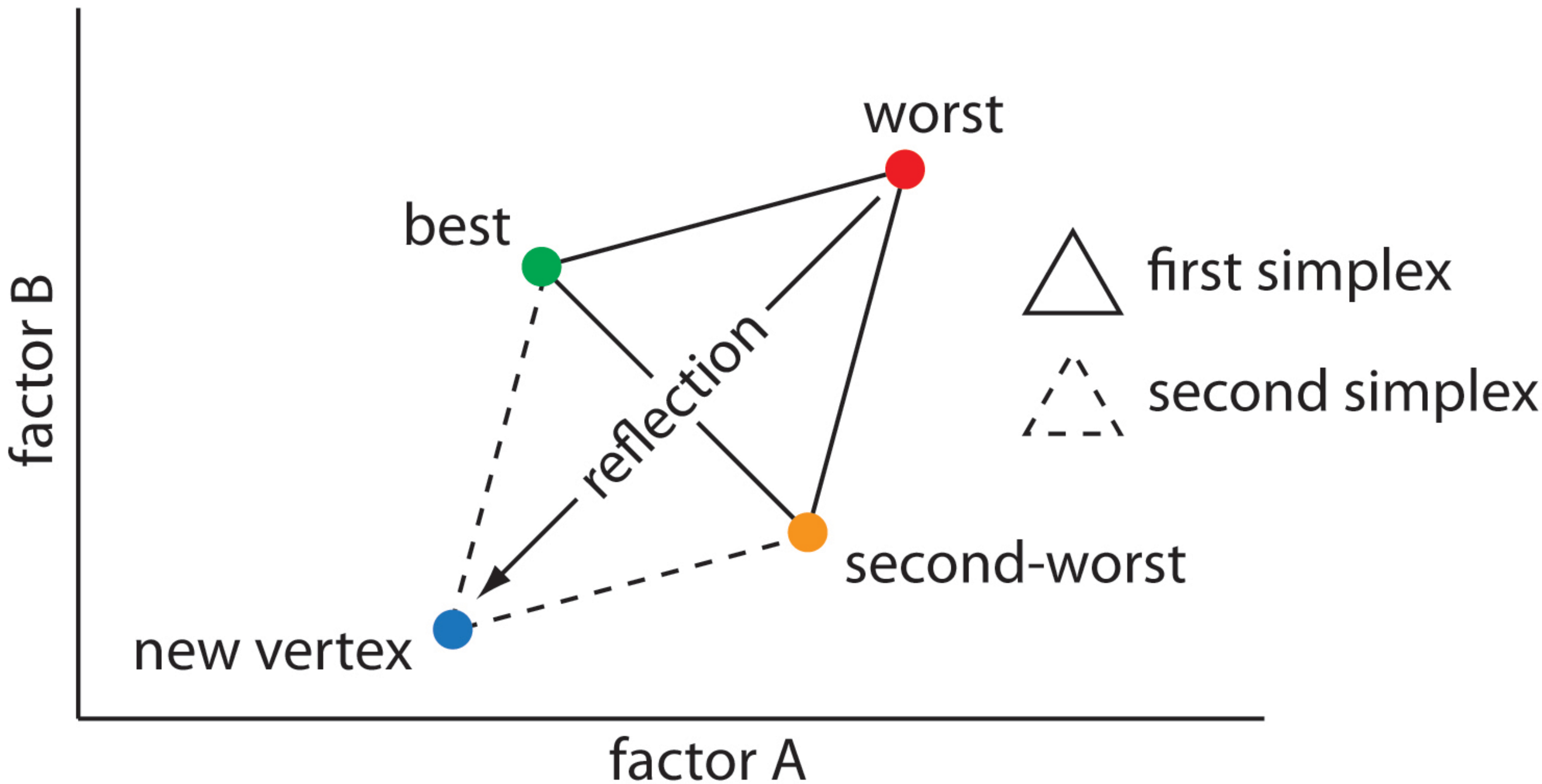
Cumulative case counts
12000
10000
8000
6000
4000
2000
0
Sierra Leone
Liberia
Guinea
Weekly incidence
2000
1500
1000
500
0
−500
Sierra Leone
Liberia
Guinea
Jun 2014
Jul 2014
Aug 2014
Sep 2014
Oct 2014
Nov 2014
Dec 2014
Jan 2015

Fitting Stochastic Models

- As mentioned previously, this is a non-trivial problem
- Inherent conflict between forecasting and this research question, with methods being developed simultaneously
  - Forecast: Need to get as close to *this* epidemic as possible. Potential outcomes that didn't happen aren't of interest
  - Stochastic Variability: Need to capture all the potential outcomes. An epidemic that dies out on day 1 is just as important
- How do you define a "miss", let alone write an objective function to penalize it?

Fitting Stochastic Models

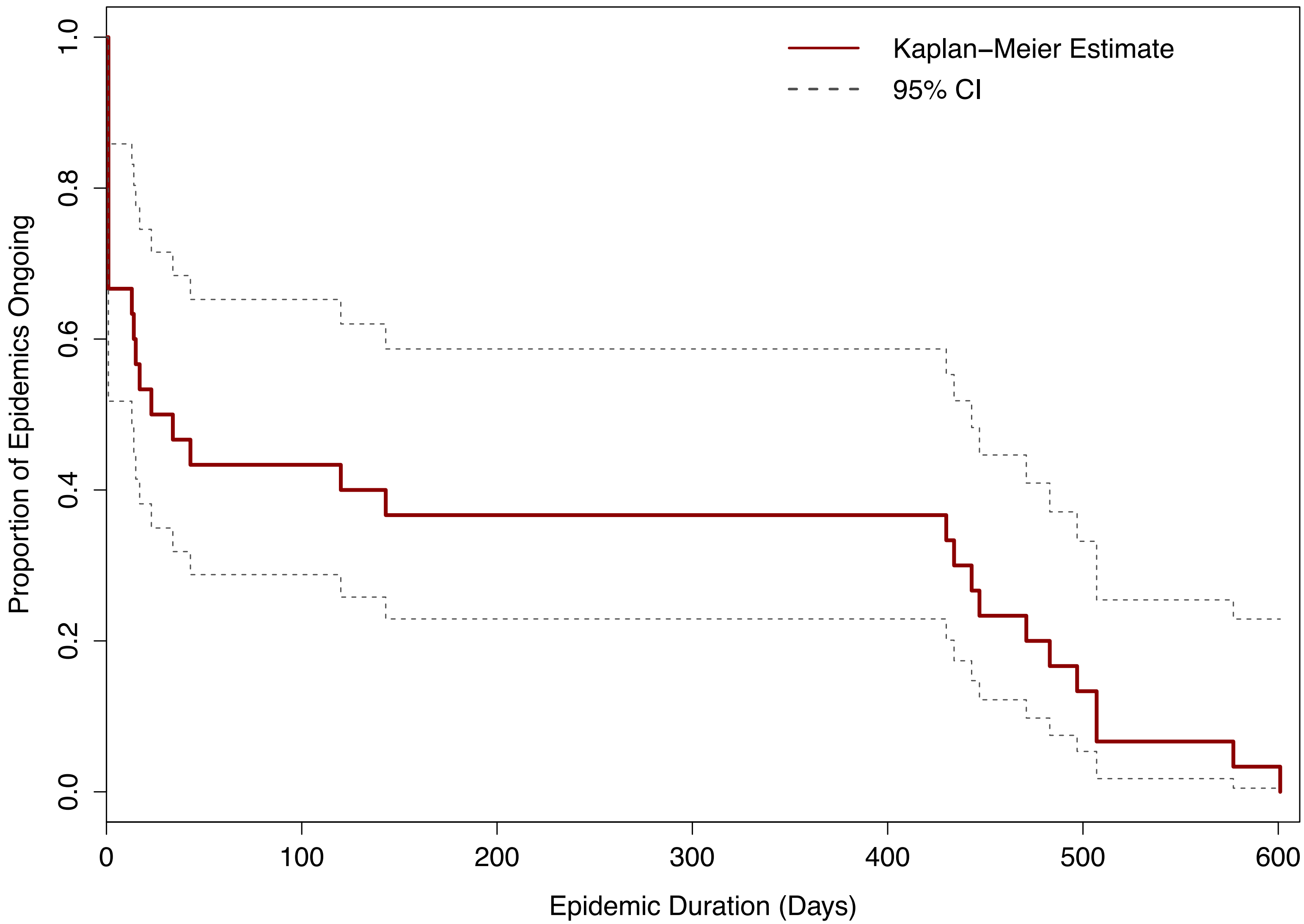
- Parameters Fit:
  - Transmission probability
  - Two ”pivot points” that would allow the epidemic to change
    - Both timing and change in transmissibility are free parameters
  - Fit to the minimum distance of an ensemble of 25 model runs compared to the time series of reported cases
    - Obviously fairly time consuming
  - Used Nelder-Mead method to optimize the parameter fit
    - Strong assumptions behind this method – unimodal objective function among them


worst
best
first simplex
second simplex
reflection
second-worst
new vertex
factor B
factor A

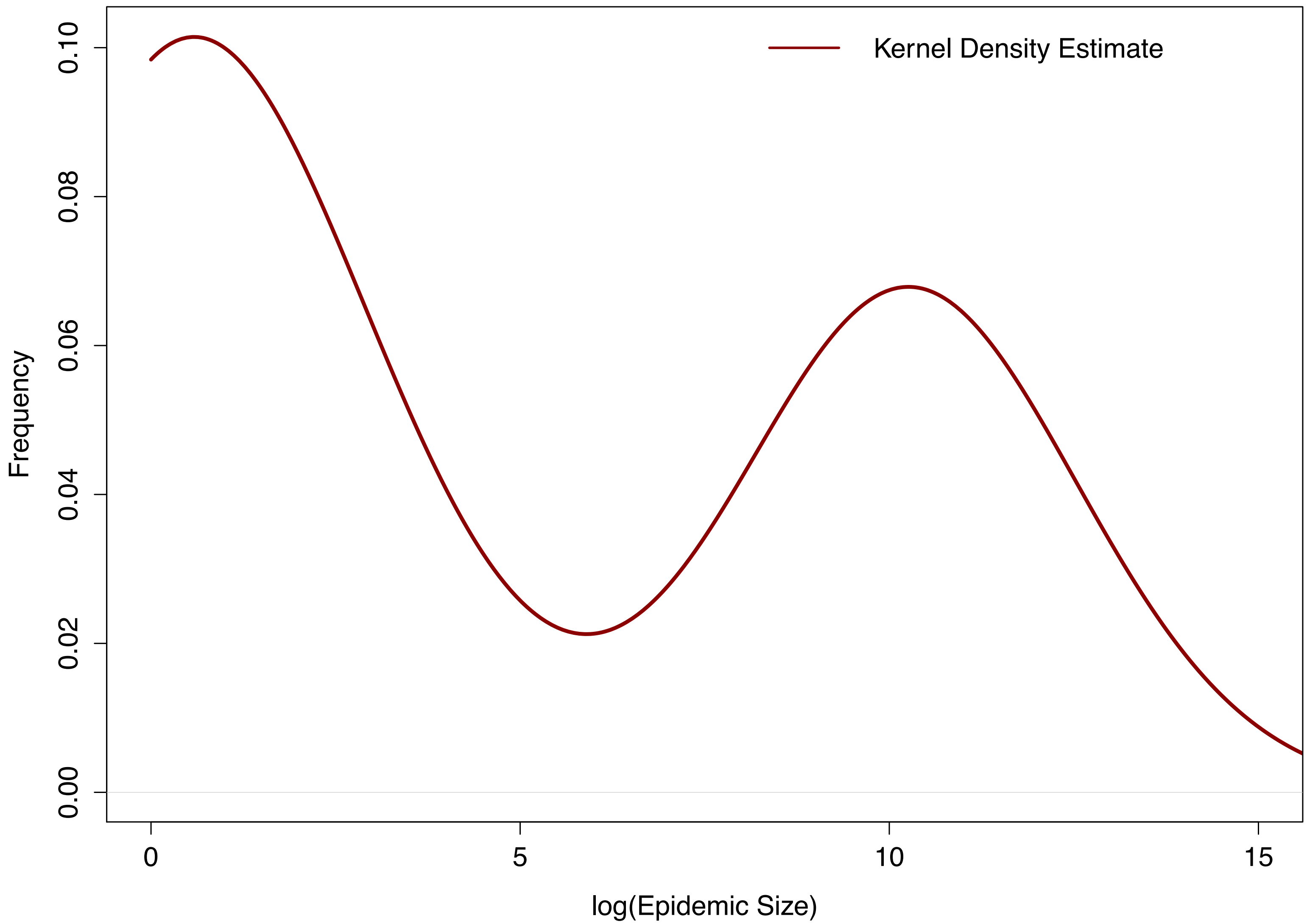
What Results to Measure

- Where this epidemic relates to other epidemics in terms of **time**
  - This is also implicitly a forecast of how long the epidemic is going to last, though not conditioned on what has already happened
- Where this epidemic relates to other epidemics in terms of **size**

Time


Kaplan–Meier Estimate
95% CI
Proportion of Epidemics Ongoing
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
500
600
Epidemic Duration (Days)

Size



Kernel Density Estimate
Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
5
10
15
log(Epidemic Size)

Can This Ebola Outbreak Explain the Past History of Ebola?

- Yes and No
- While the historical outbreaks are accounted for in this model, they fall in a “trough” in the density of model outputs
  - The model struggles with “medium sized outbreaks”
- This model *does* suggest that small, short lived micro-epidemics of Ebola are possible with the same transmission parameters as the current outbreak
  - This may not be the first Ebola outbreak in West Africa
  - Small outbreaks may be hard to distinguish from the noisy signal of Lassa, Malaria, etc.

Next Steps and Other Perspectives

- Change the way the synthetic contact network is built to forbid long distance travel
  - This addresses one of the two main hypotheses for why this Ebola outbreak was big
  - Mobility patterns may be different than in previous outbreaks
- "Can this Ebola outbreak give us the previous outbreaks as well?" is different from "Can previous outbreaks have predicted this outbreak?"
  - It's possible that the "true" parameter distribution has the current outbreak as an outlier – we're attempting to fit to it

Social Epidemiology

- Switching gears from outbreak response to social epidemiology (with both infectious and non-infectious disease examples)
- Social epidemiology is another good place for agent-based models
  - Wide-range policy questions
  - Difficult to conduct randomized trials
  - Pulls data from many different sources
- **Tons** of social epidemiology is trying to get at the emergent outcomes of the decisions lots of individuals make
  - **This should sound familiar**
- We're also moving past my work

Population Mixing and Public Health Emergencies

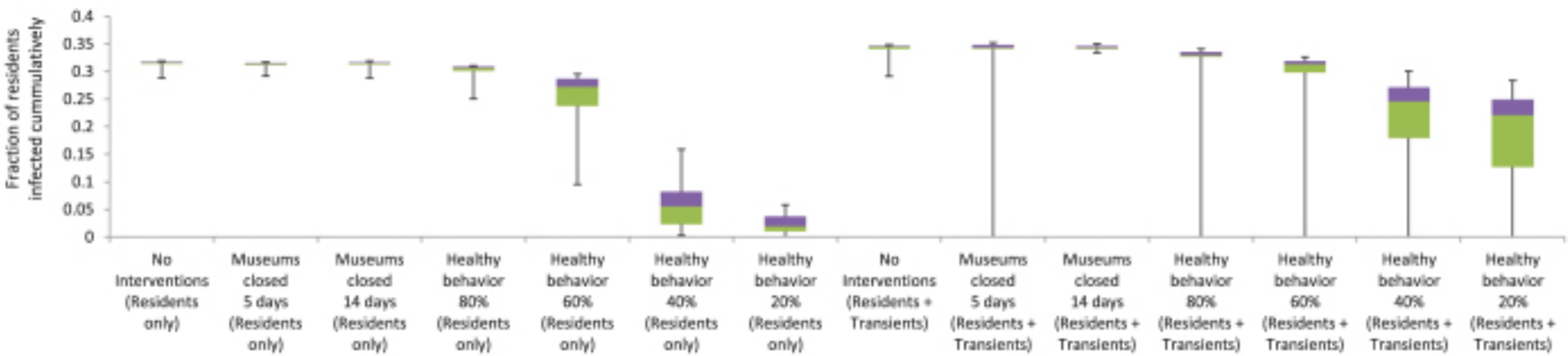
- NDSSL has a working, high-resolution model of Washington, D.C.
- Strangely structured population – lots of tourists, lots of residents, but a distinct division between them
- Naturally some interest in modeling response to emergencies

Pandemic Flu

- Parikh N, Youssef M, Swarup S, Eubank S. Modeling the effect of transient populations on epidemics in Washington DC. *Scientific Reports*. 2013;3:3152. doi:10.1038/srep03152.
- Transient populations (tourists) had a large impact on the amount if disease during a simulated epidemic
- These took place primarily at tourist sites
- Should we close these sites?
- Are there alternative strategies?

Museum Closure

- Closing museums had little impact on the epidemic
- Drives transient populations to other parts of the city where they come in contact with residents
- Promoting health behavior (hand sanitization etc.) at tourist destinations had a bigger impact



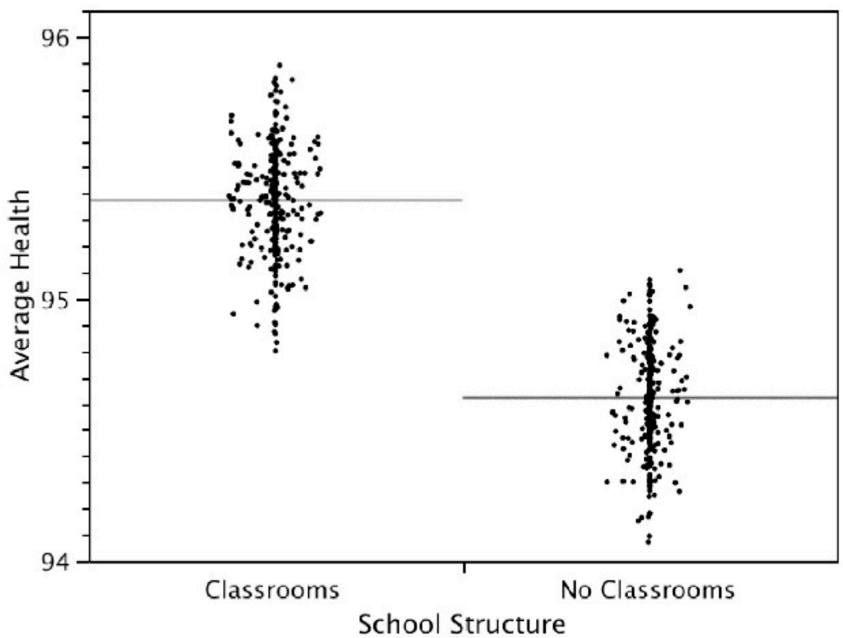
School Closure

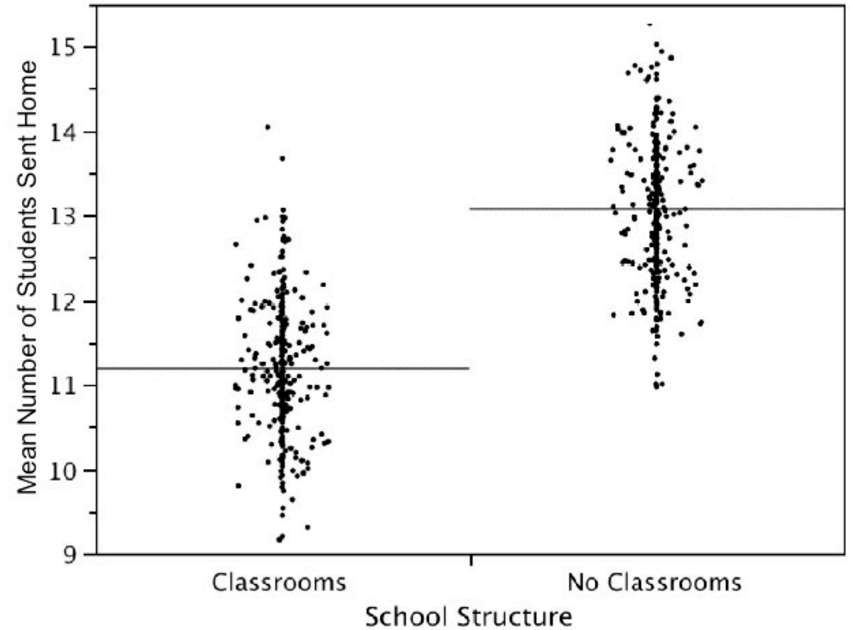
- Similar question in Lofgren, E. T., Rogers, J., Senese, M., & Fefferman, N. H.. (2008). Pandemic preparedness strategies for school systems: is closure really the only way?. *Annales Zoologici Fennici*, *45*(5), 449–458
- Is there a way to mitigate the impact of school closure (lost school days, social justice impacts, etc.) for pandemic preparedness

Simple School Model

- 10 classrooms
  - Each class room an 8x8 grid housing 20 students
  - Larger 20x20 common area (cafeteria, gymnasium)
  - Students inhabit one cell, and can interact with their adjacent cells
  - Can move one cell/unit time
    - 3x speed in common areas
  - In classrooms from 0900 to 1100 and 1300 to 1500
  - Other times in common area
  - Students shed infectious material into the environment, where other students can be infected by contact with them or fomites

Impact of School Structure

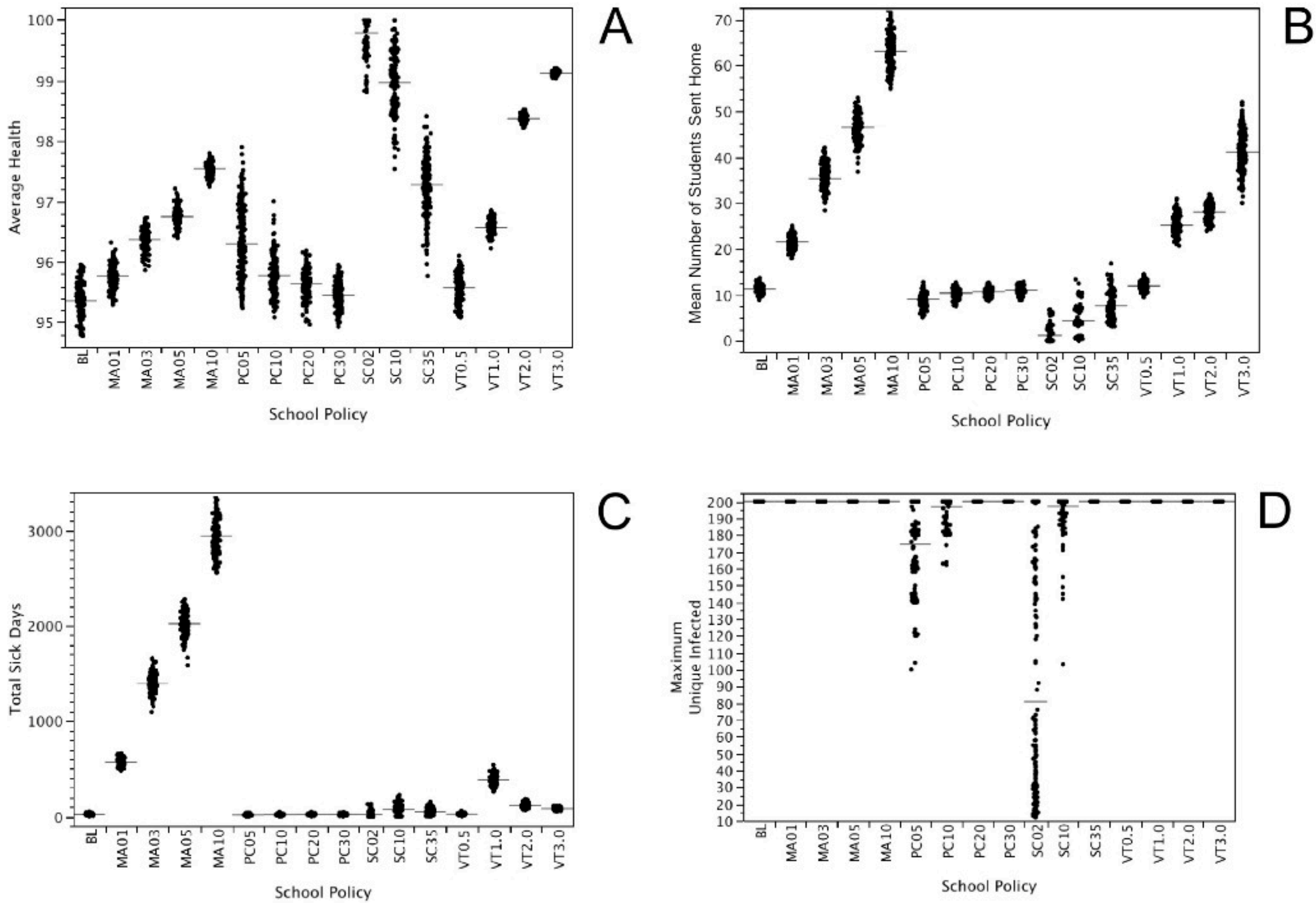




Clearly, large-scale mixing invites larger epidemics

Policy Interventions

- The “standard” policy for addressing a pandemic was to close schools down
  - Negative outcomes for this
    - Lost school days for all kids, sick and well
    - Differential economic impact of kids being sent home
    - Unknown hidden impacts of closing school
      - “They just go stay with someone’s grandma”
- Are there alternative interventions?
  - Closing common areas
    - Less impactful
    - Likely preferred by everyone except well students
  - Mandatory absence for sick kids based on time rather than health


A
Average Health
School Policy
B
Mean Number of Students Sent Home
School Policy
C
Total Sick Days
School Policy
D
Maximum
Unique Infected
School Policy
BL
MA01
MA03
MA05
MA10
PC05
PC10
PC20
PC30
SC02
SC10
SC35
VT0.5
VT1.0
VT2.0
VT3.0

Alternative Strategies

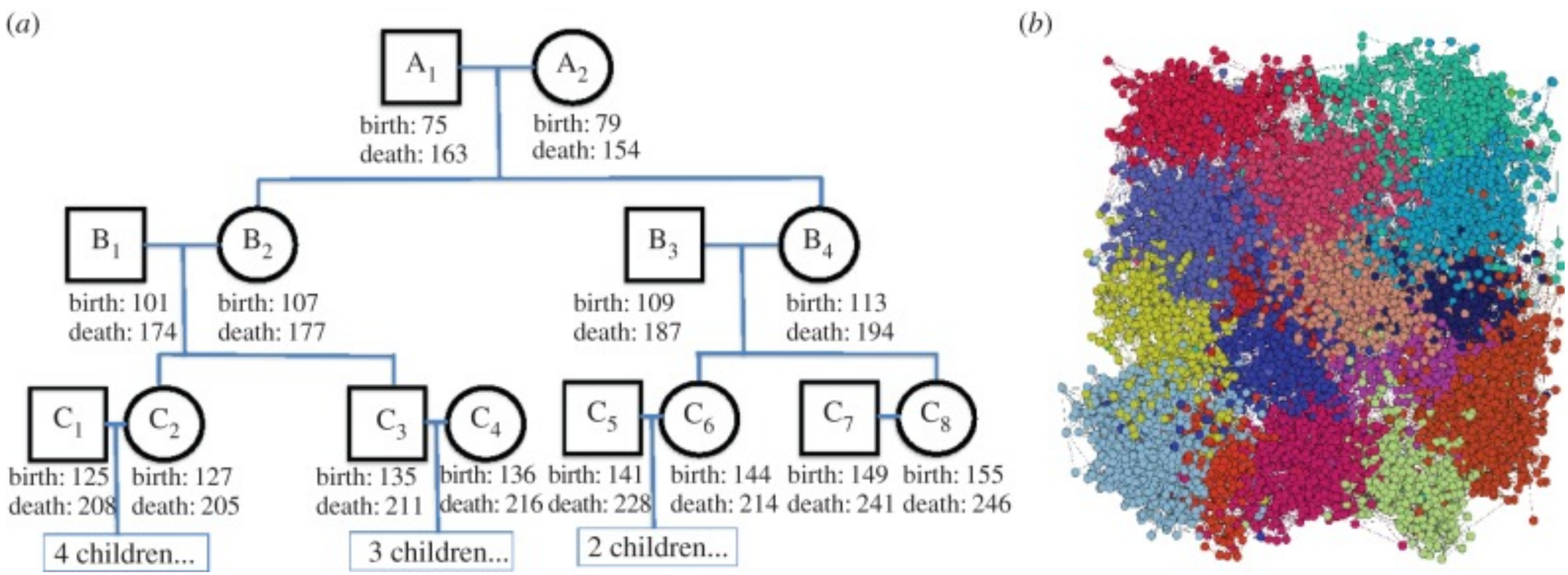
- “School Open” strategies may help mitigate an epidemic while lessening the burden of response
- Is this an “acceptable compromise” between more effective strategies with considerably higher costs to them
- Much less consequence for a “false alarm”

Incarceration

- Lum K, Swarup S, Eubank S, Hawdon J. 2014 The contagious nature of imprisonment: an agent-based model to explain racial disparities in incarceration rates. J. R. Soc. Interface 11: 20140409.
- There are observed racial disparities in both incarcerations between black and white Americans in the United States **and** in sentence length
  - Can the latter explain the former
- Assume incarceration is “contagious” – being placed in jail makes your family/close associates etc. more likely to be placed in jail

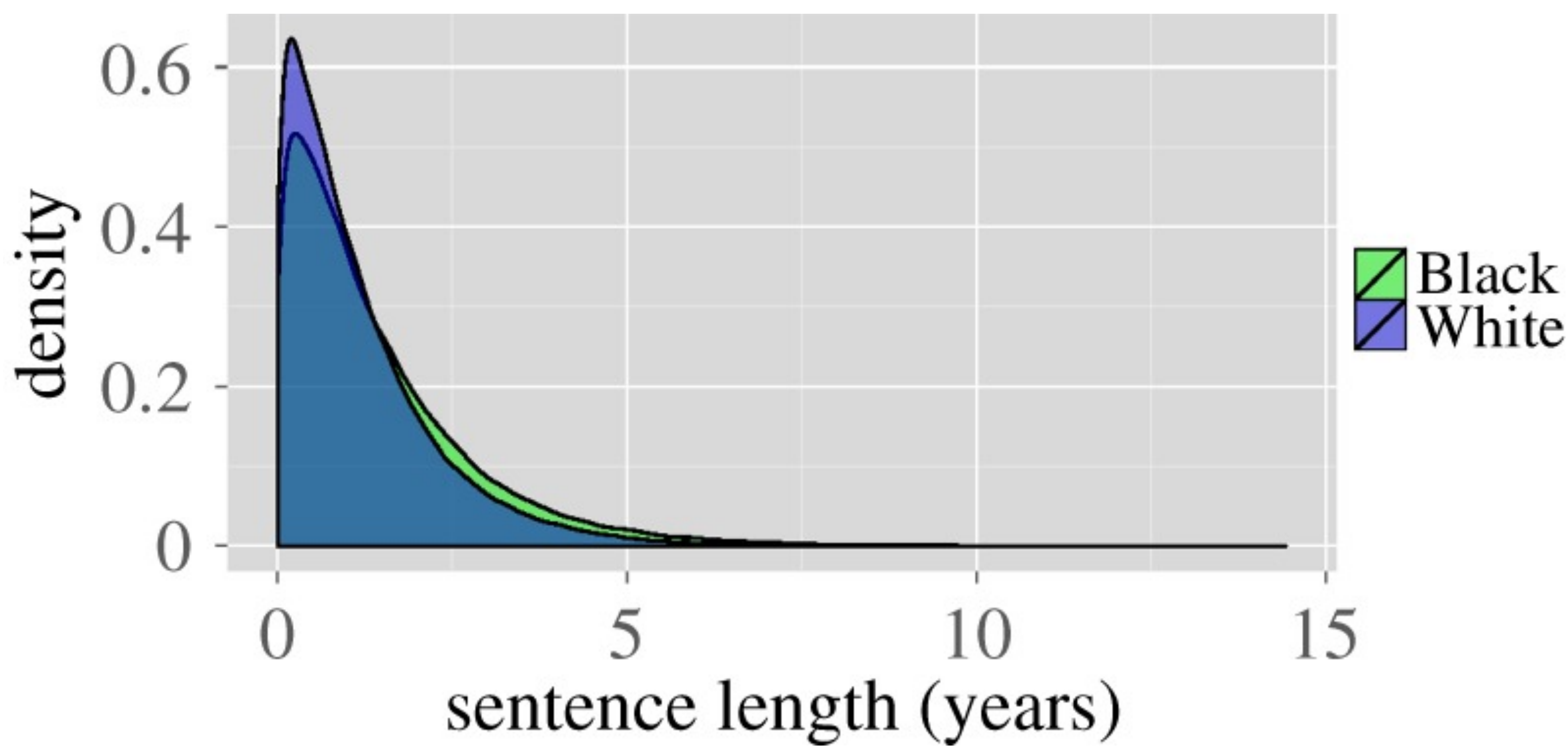
Modeling Families

- Complex formation of population
- People are born, assigned a geographic space, friends and spouses are selected, babies are born and people die
- Relationships represented by network edges
- 8856 agents with 61376 ties between them



Sentence Generation

- Drawn from different distributions for blacks vs. whites based on sentencing data from the Bureau of Justice Statistics



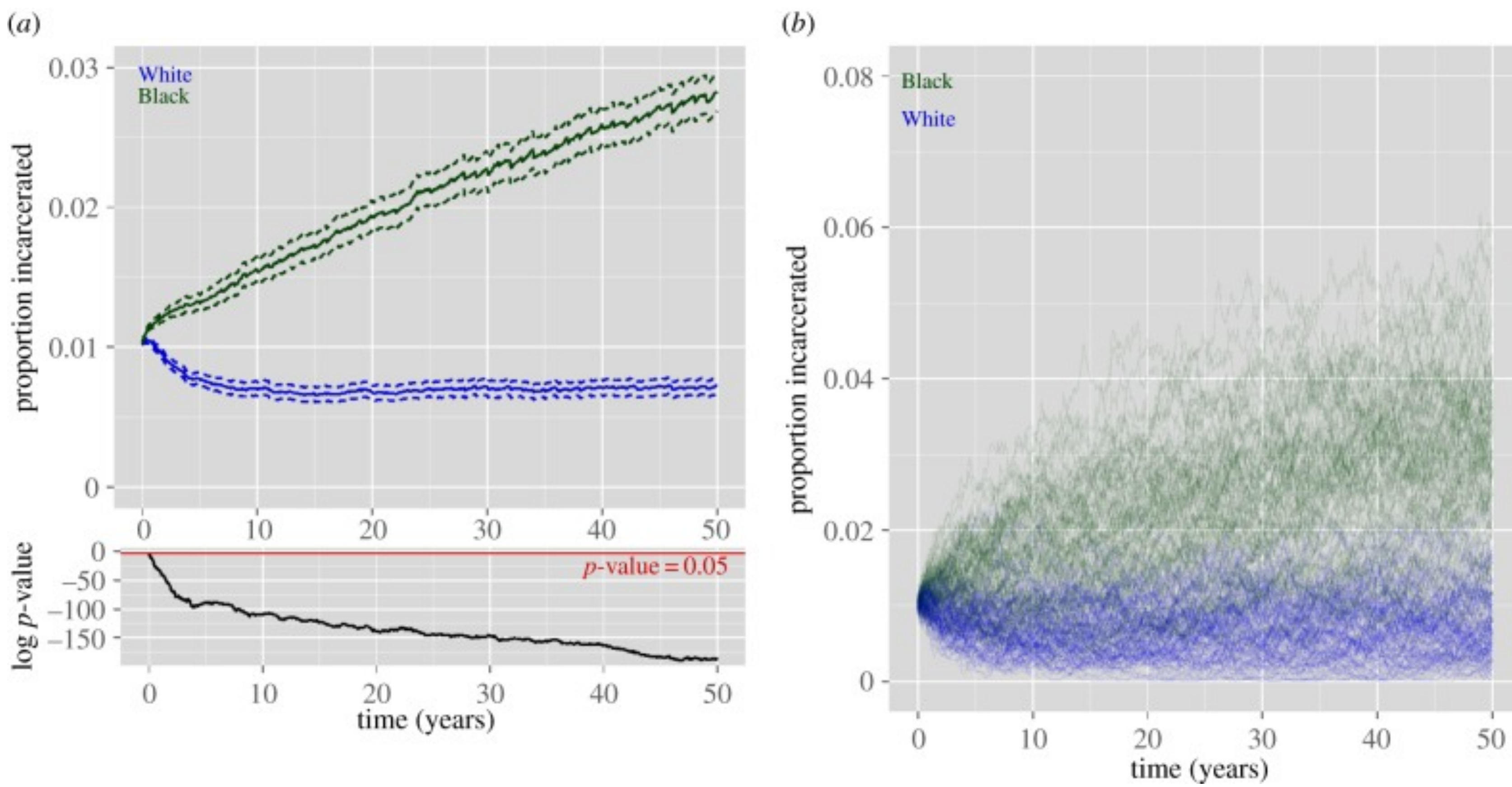
Incarceration Transmission

- Incarceration is passed from an inmate along their relationships
- Cumulative probability will vary by sentence length

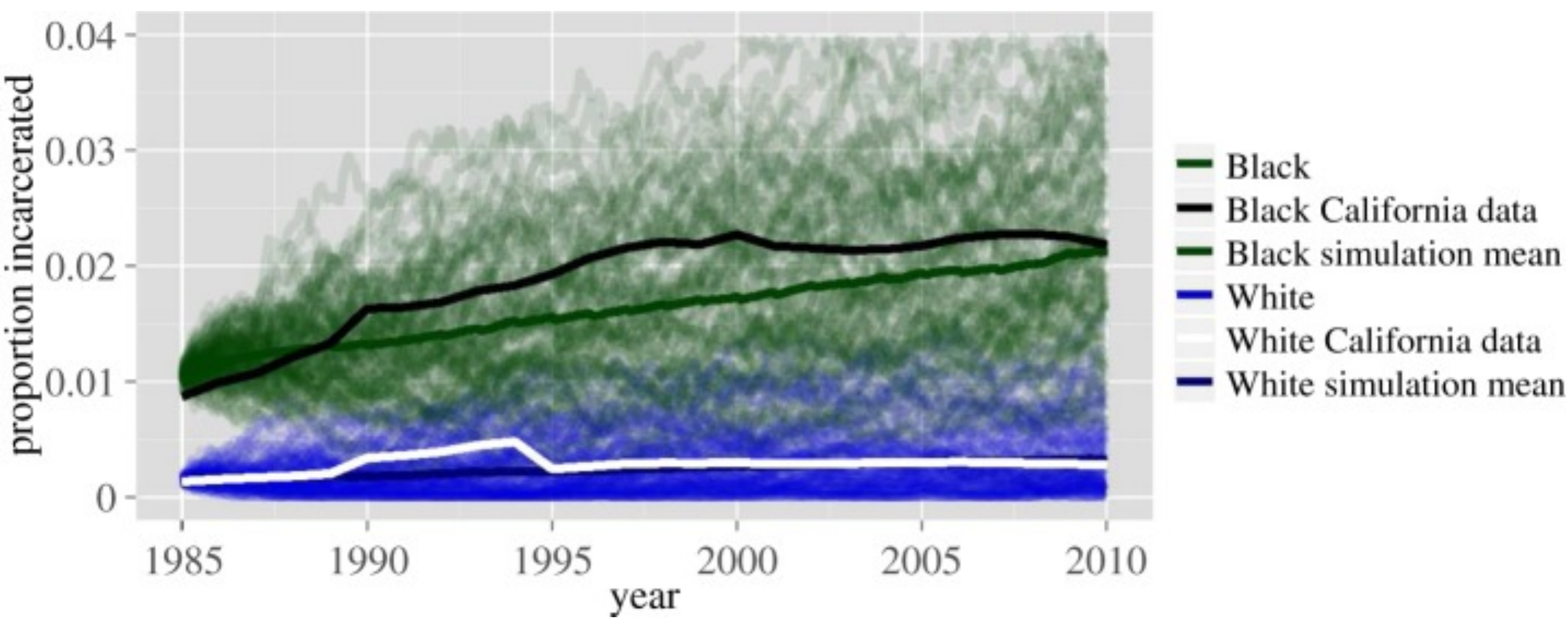
	survey		White		Black	
	women	men	women	men	women	men
mother	0.012	0.048	0.012	0.046	0.014	0.056
father	0.147	0.148	0.138	0.138	0.163	0.163
sister	0.107	0.059	0.101	0.058	0.121	0.069
brother	0.377	0.349	0.324	0.303	0.370	0.347
spouse	0.059	0.011	0.057	0.011	0.069	0.013
adult child	0.213	0.085	0.194	0.082	0.227	0.098

Simulate these Epidemics

- 500 simulations, 250 each for the “Black” and “White” sentence distribution



Comparing to Data



Discussion

- We’ve now touched on **many** agent-based modeling examples, from summaries of their results to a pretty deep look at one project
- Questions?
- Discussions?
- Particular problems you have that you think might be amenable to agent-based modeling?