

# Characterizing regional differences in flu transmission in the US

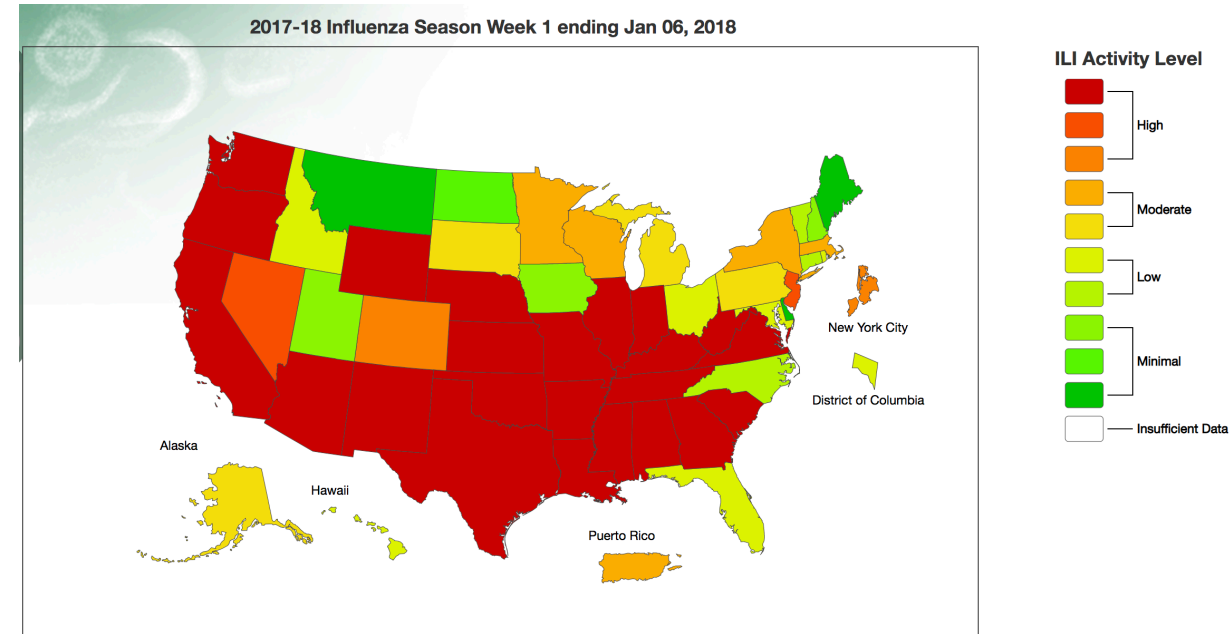
Zihao Wang, Pascual lab

# Outline

- Flu Background
- An epidemiology model incorporating viral evolution [1]
- Regional Heterogeneity in influenza incidence
- Methods
- Preliminary results

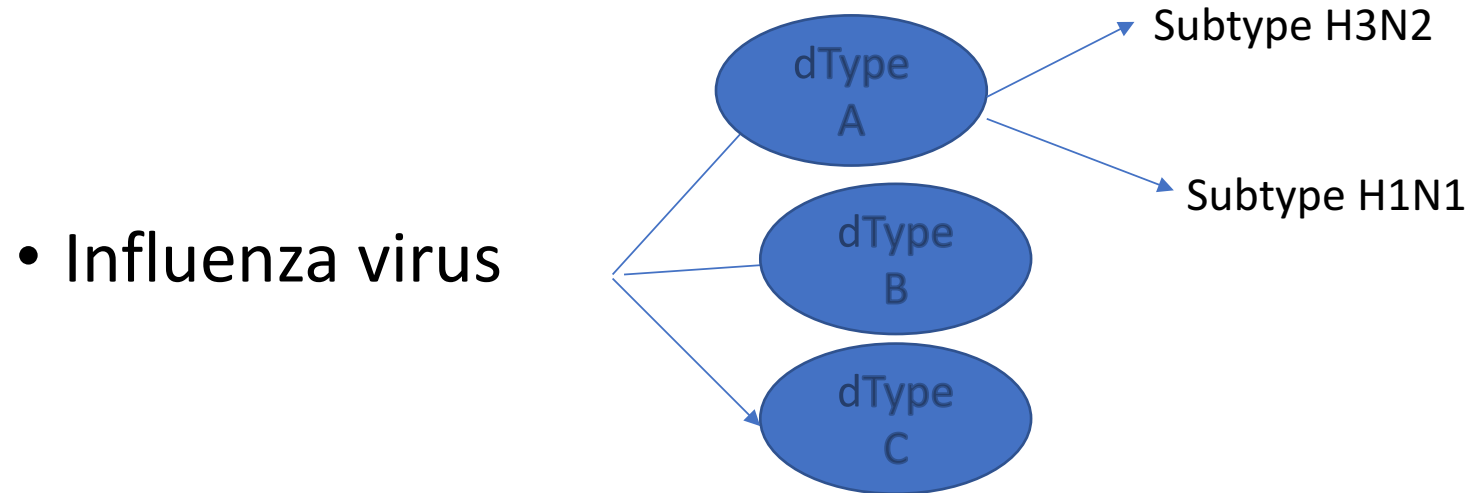
# Motivation:

- Causing 3,000,000 ~ 5,000,000 cases of severe illness and 250,000 ~ 500,000 deaths annually
- Difficult to predict



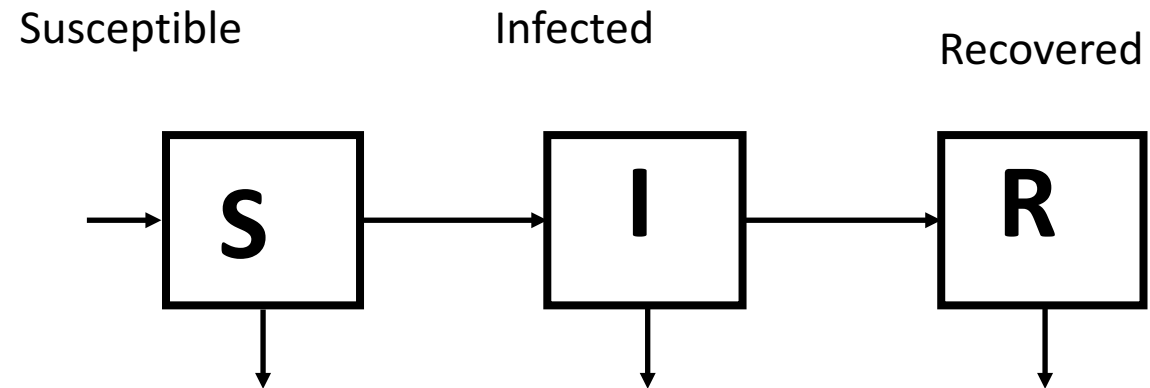
Source: Center for Disease Control and Prevention

# Flu Background



- Antigens: molecular structures on the surface of viruses that are recognized by the immune system and are capable of triggering an immune response (antibody production)
- Within a subtype, the virus can have different antigenic characterizations

# Model: SIR model



# Antigenic evolution causes loss of immunity

Infected or not

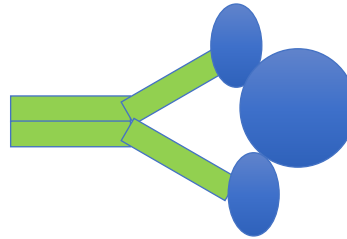
Immune or not

sequence

Season a

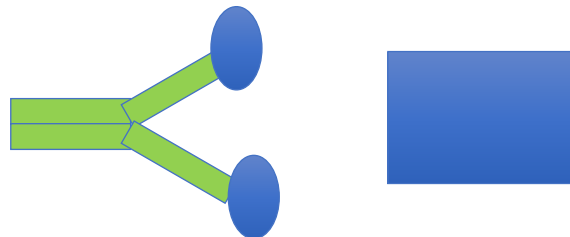


Season b

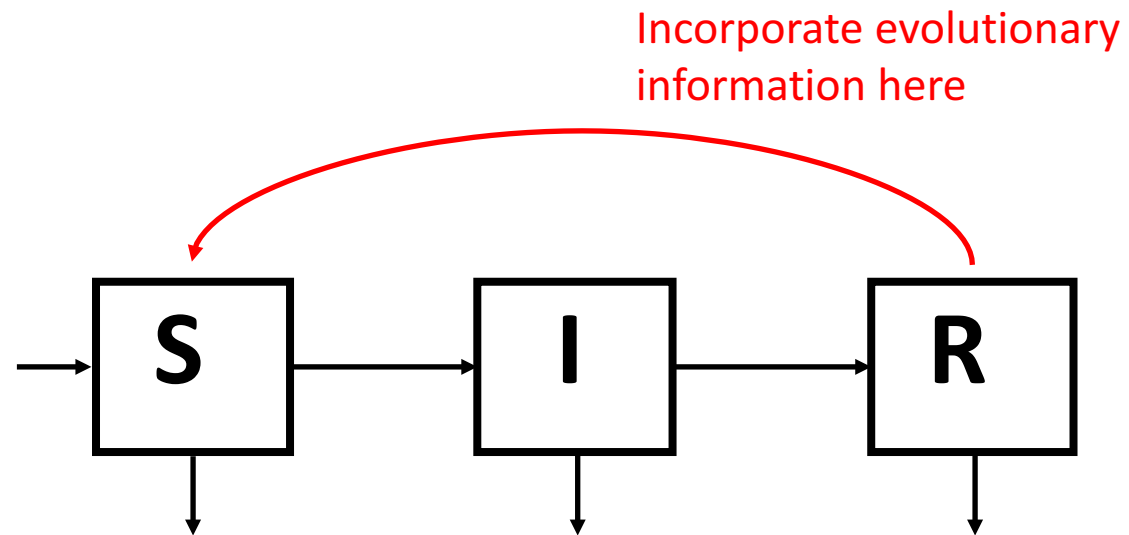


Can lose immunity when  
antigen evolves

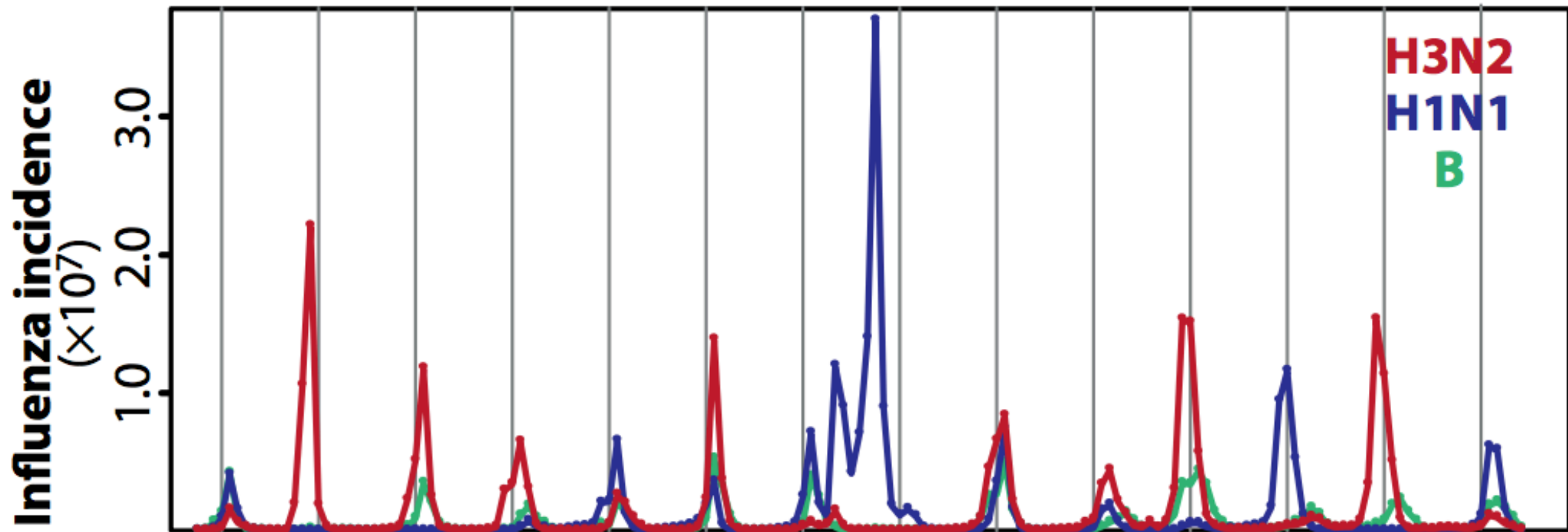
Season c



# SIRS model

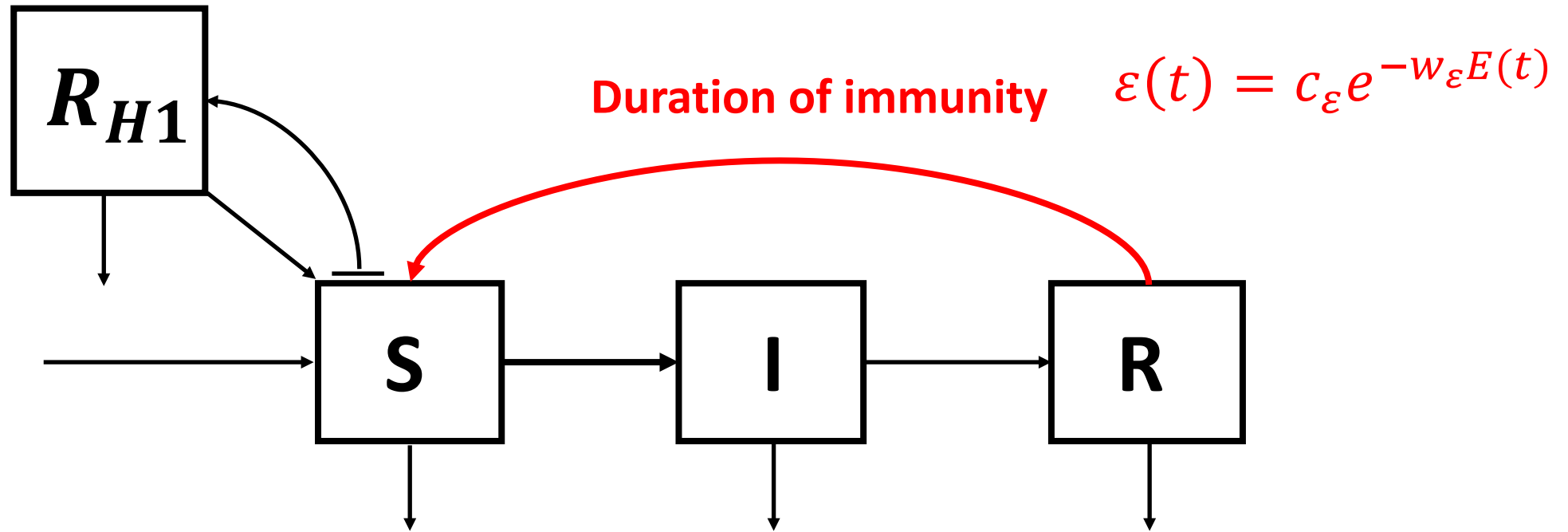


# Cross-immunity of H3N2 and H1N1

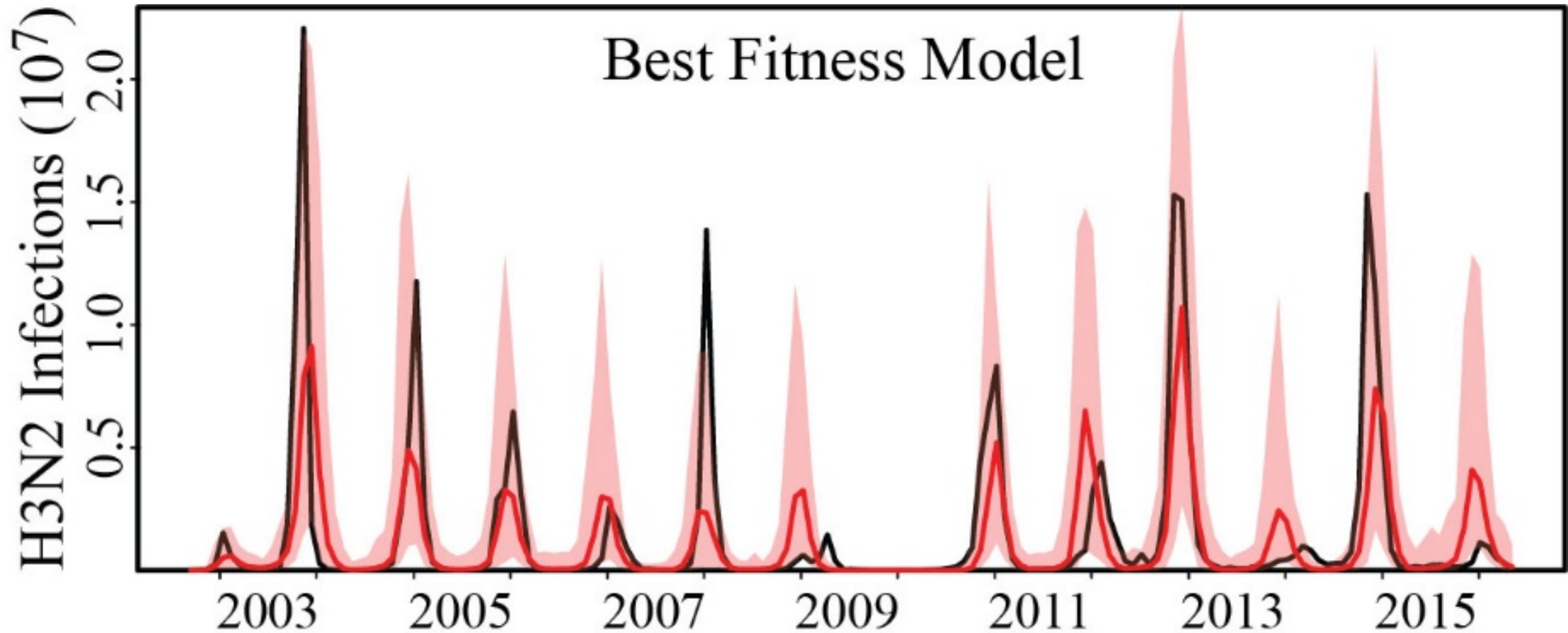




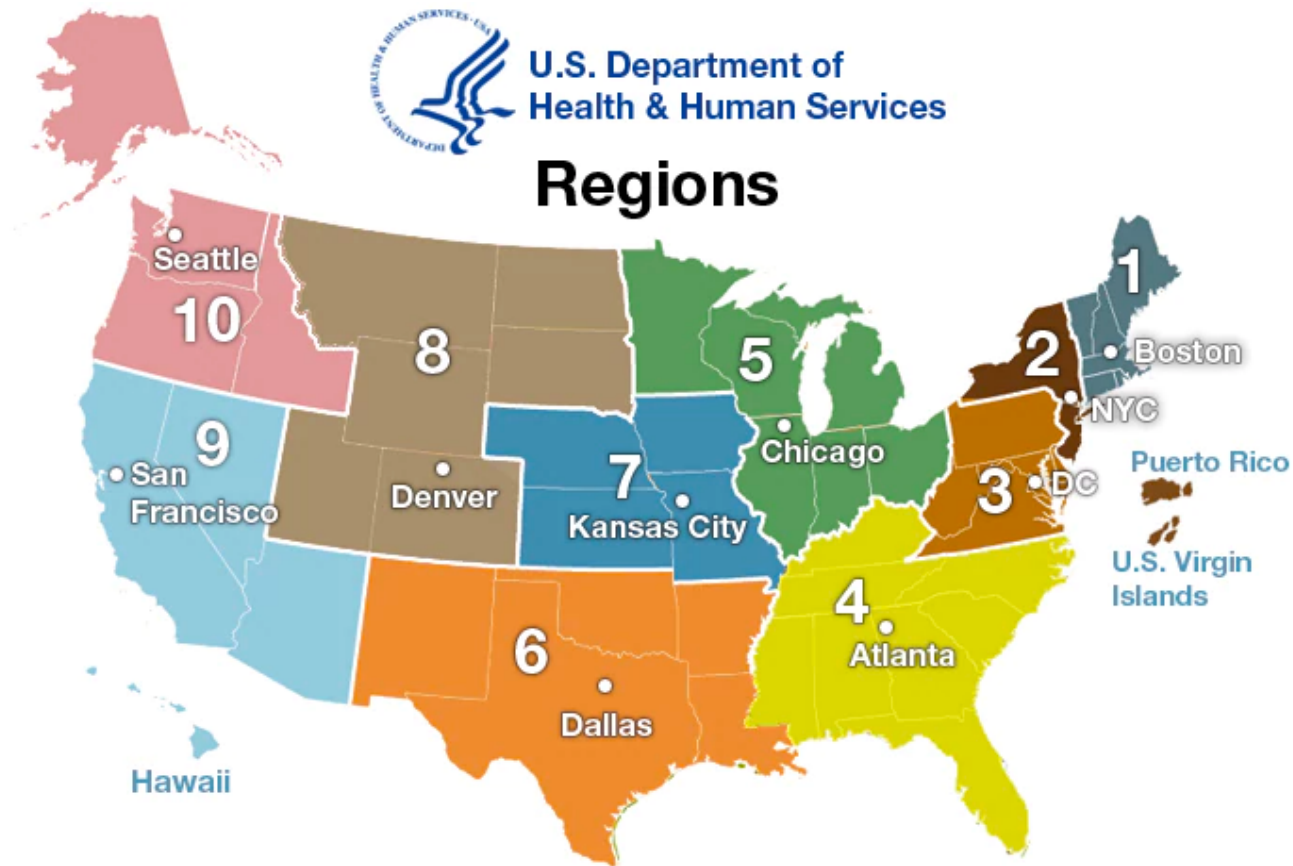
# Model informed by Evolution and H1N1 levels



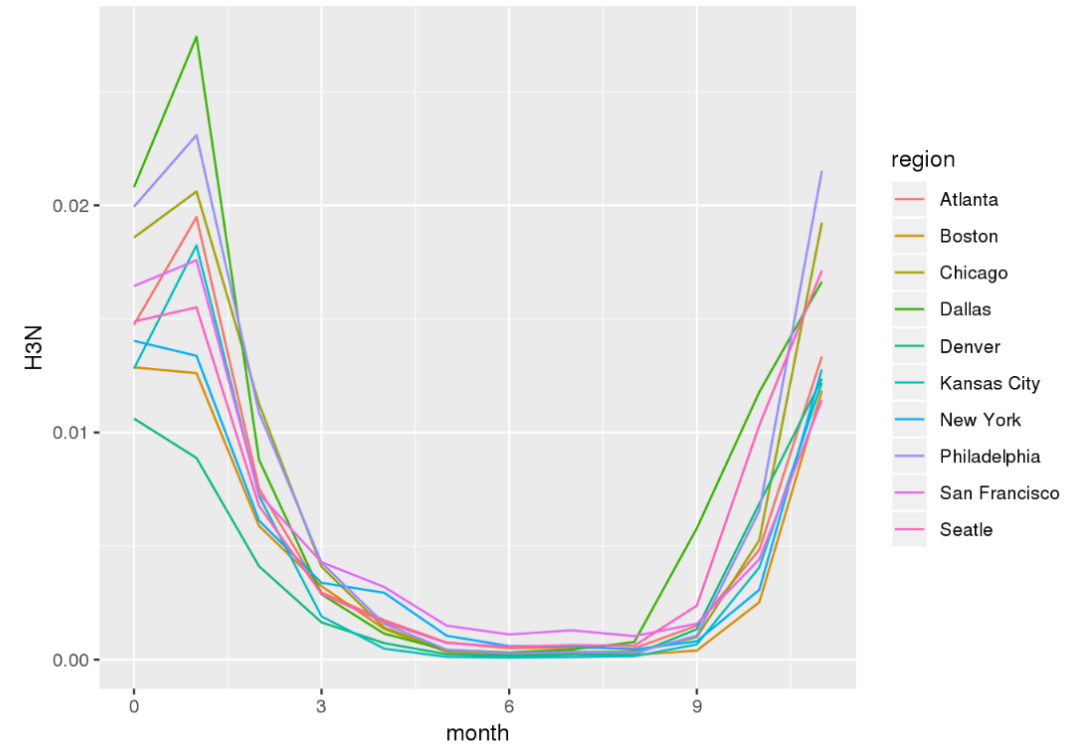
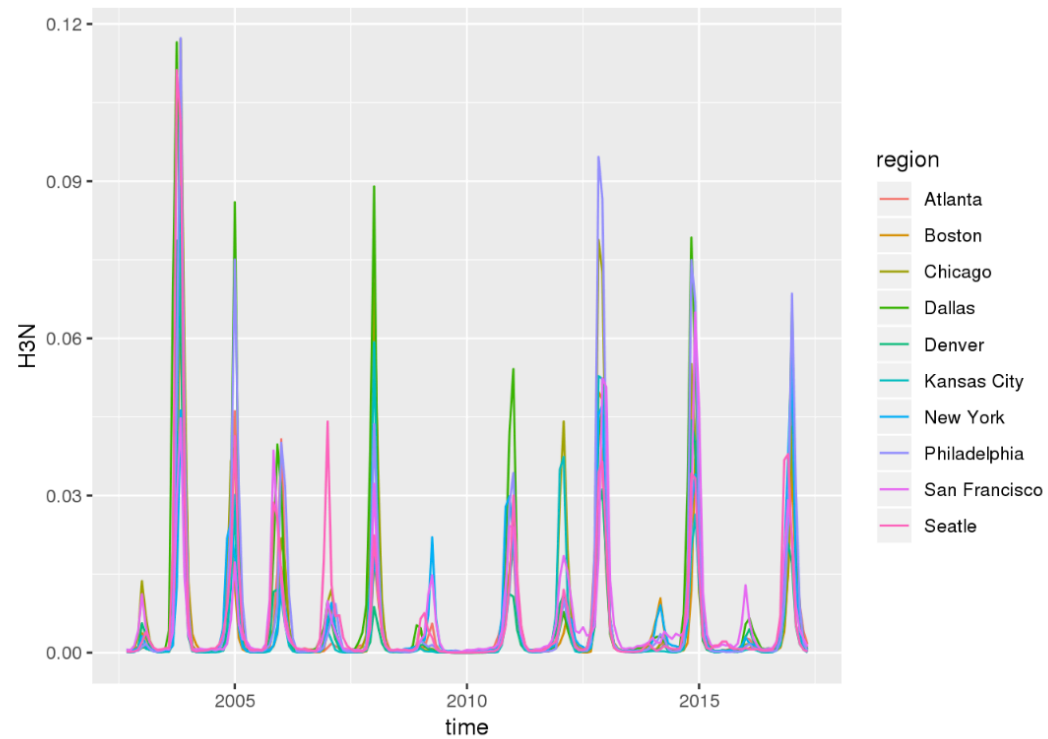
# Results from the original model



# Regional Heterogeneity in influenza incidence



# Regional Heterogeneity in influenza incidence



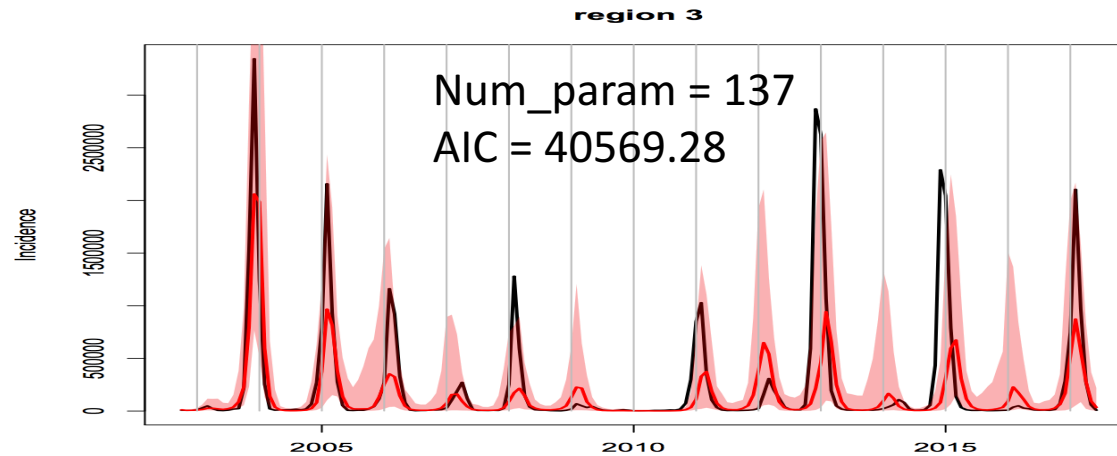
# Research question

- Can we explain the heterogeneity in flu incidence among regions by differences in underlying transmission parameters?
- If so , then can we make better predictions using regional data?

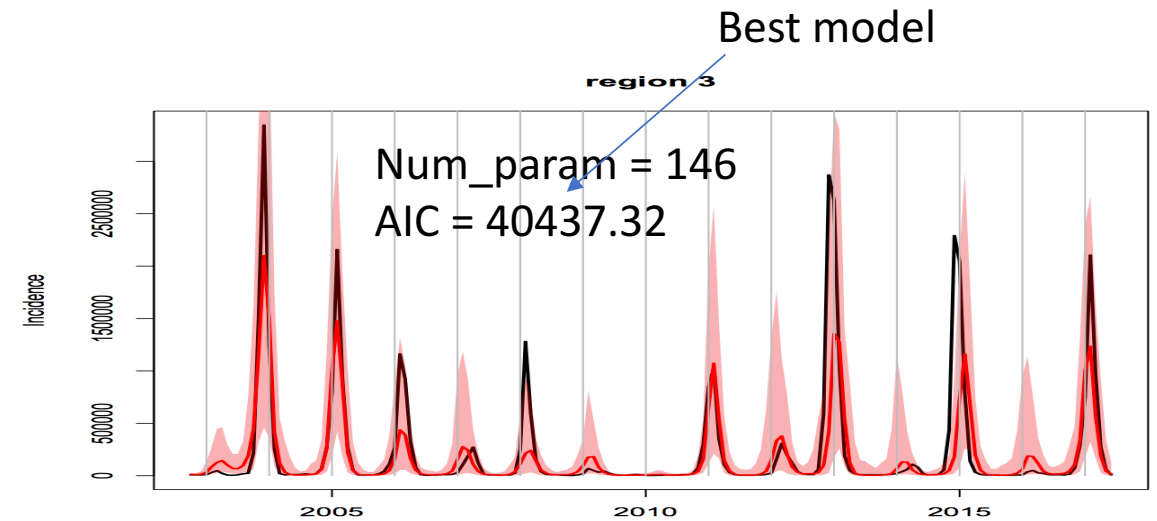
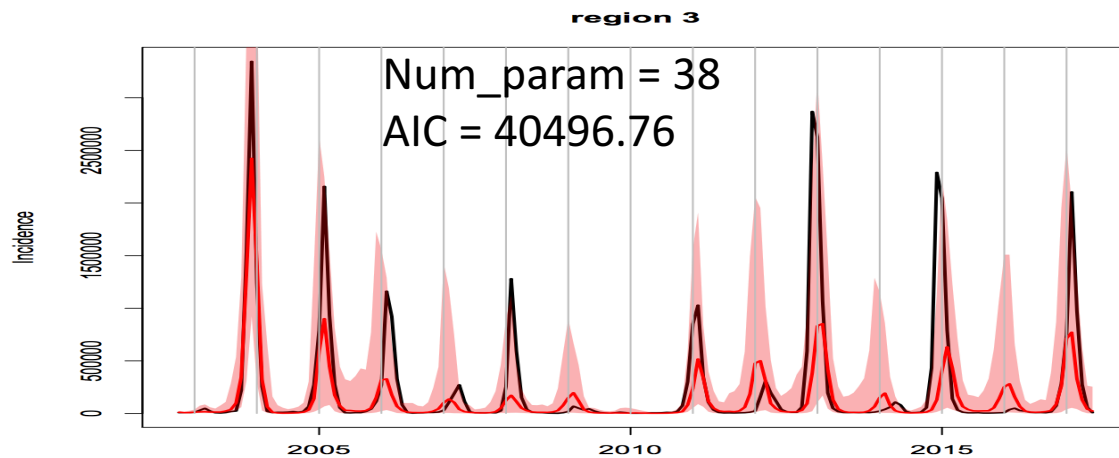
# Methods

- Build 10 models for the 10 regions
- Make some parameters shared across regions and some specific
- Fit the model to the incidence data

# Model Selection



- These are simulation plot for region 3 data
- The AIC needs to be taken with a grain of salt



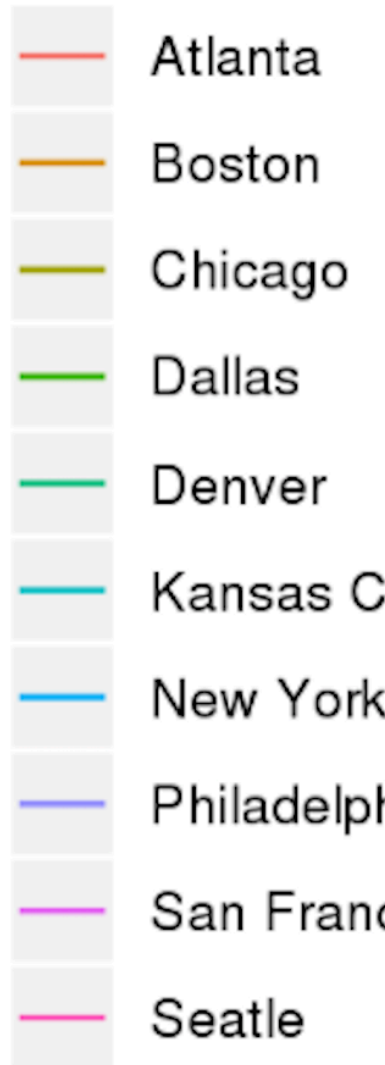
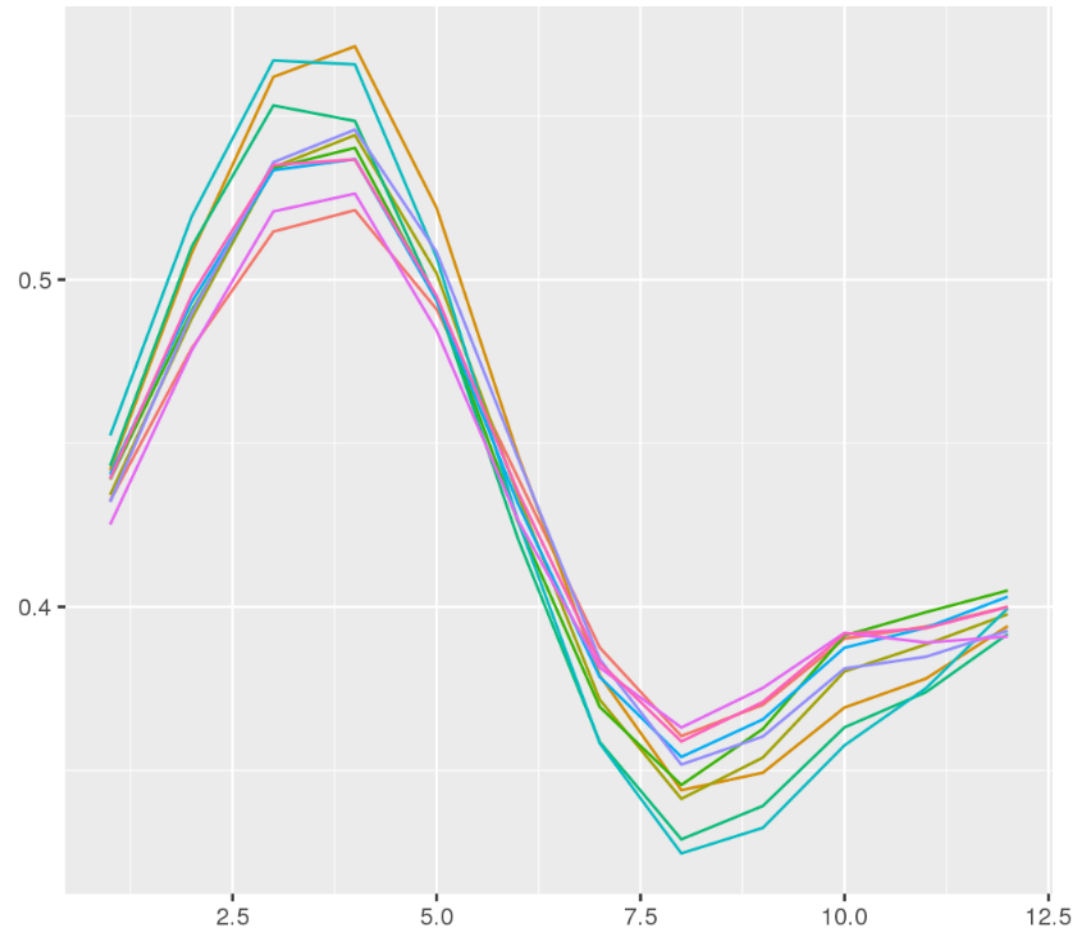
# Differences in seasonality

$$\beta(t) = \exp \left[ \sum_{i=1}^6 w_i s_i \right] \frac{d\Gamma}{dt}$$

Y is the transmission rate  
(person/(person\*day) without process  
noise;  
it shows seasonality of transmission

X is the 12 months

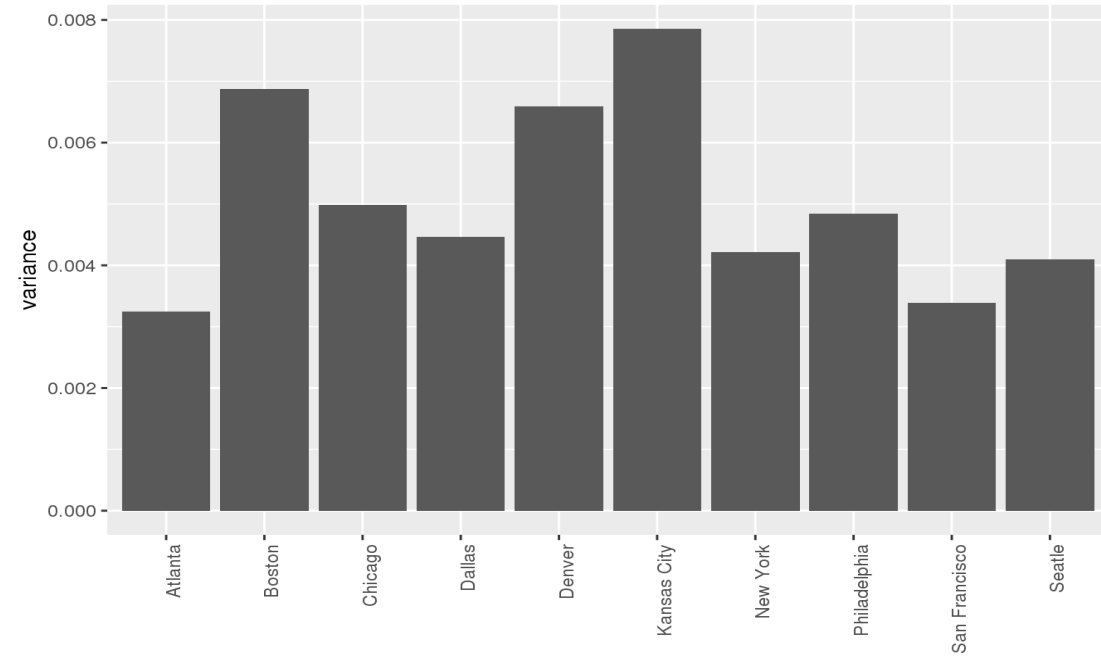
The regions are shown  
as the names of one of  
the cities.





# Seasonality and temperature

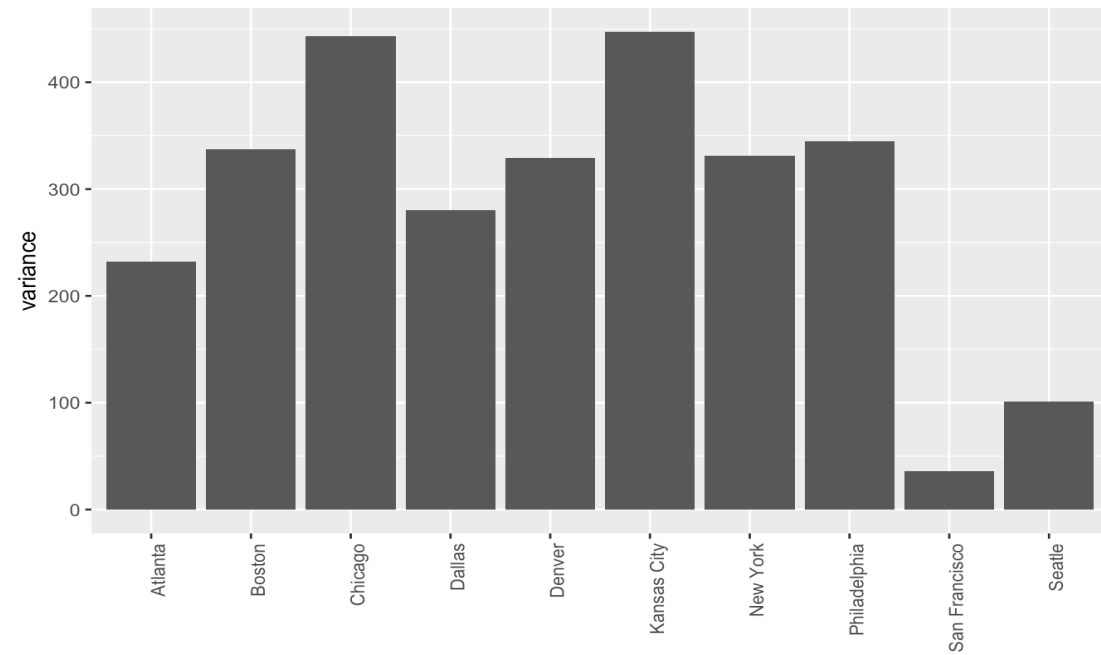
Variance of  
transmission rate



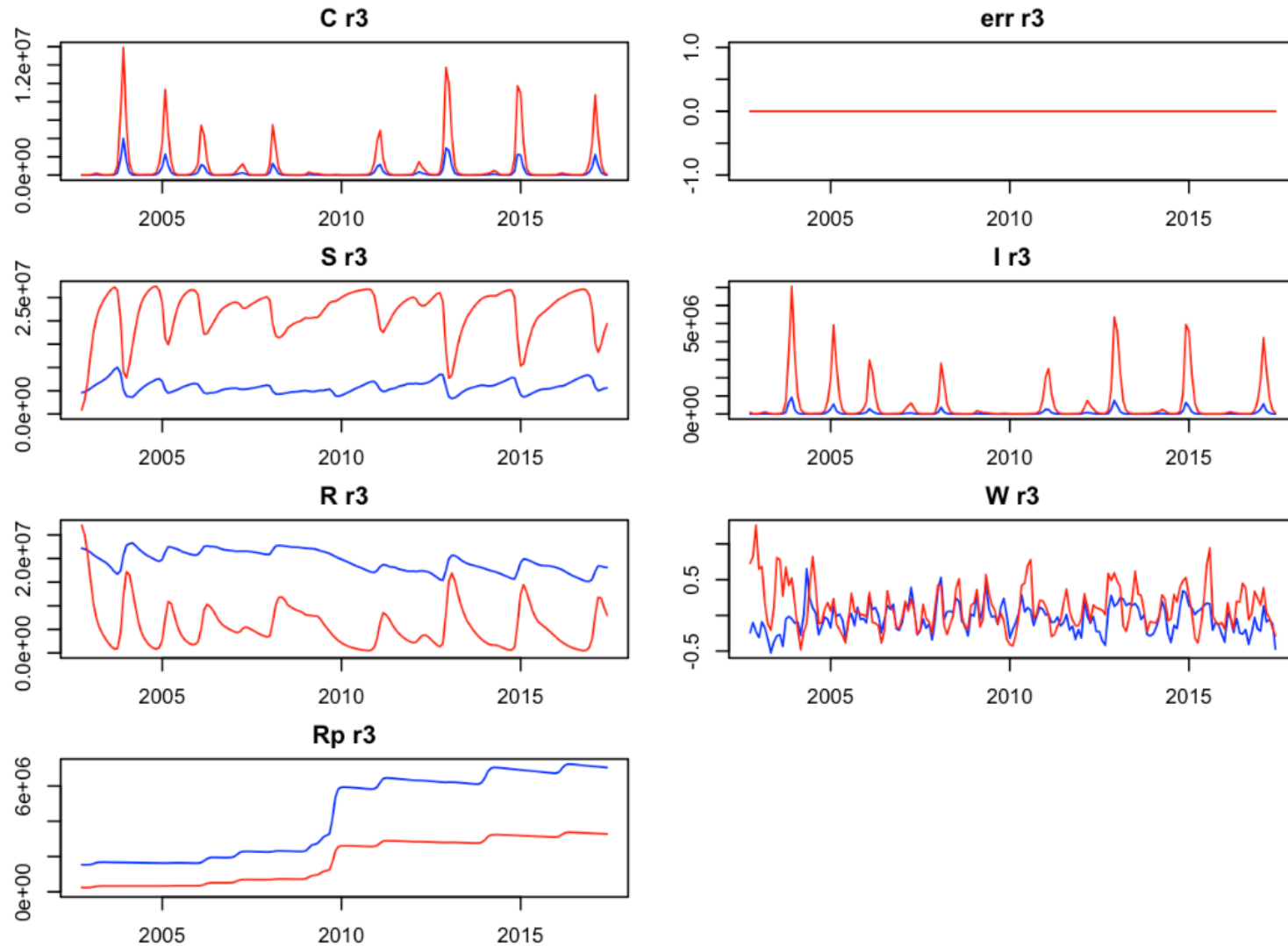
Correlation = 0.6784565

P-value = 0.03103

Variance of  
temperature



# Local optimum and identifiability of parameters



Two parameter combinations that yield similar likelihood may have very different biological meanings.

The red and blue line corresponds to the hidden states of such parameter combinations. Both of them might be in the local optimum of the likelihood surface.

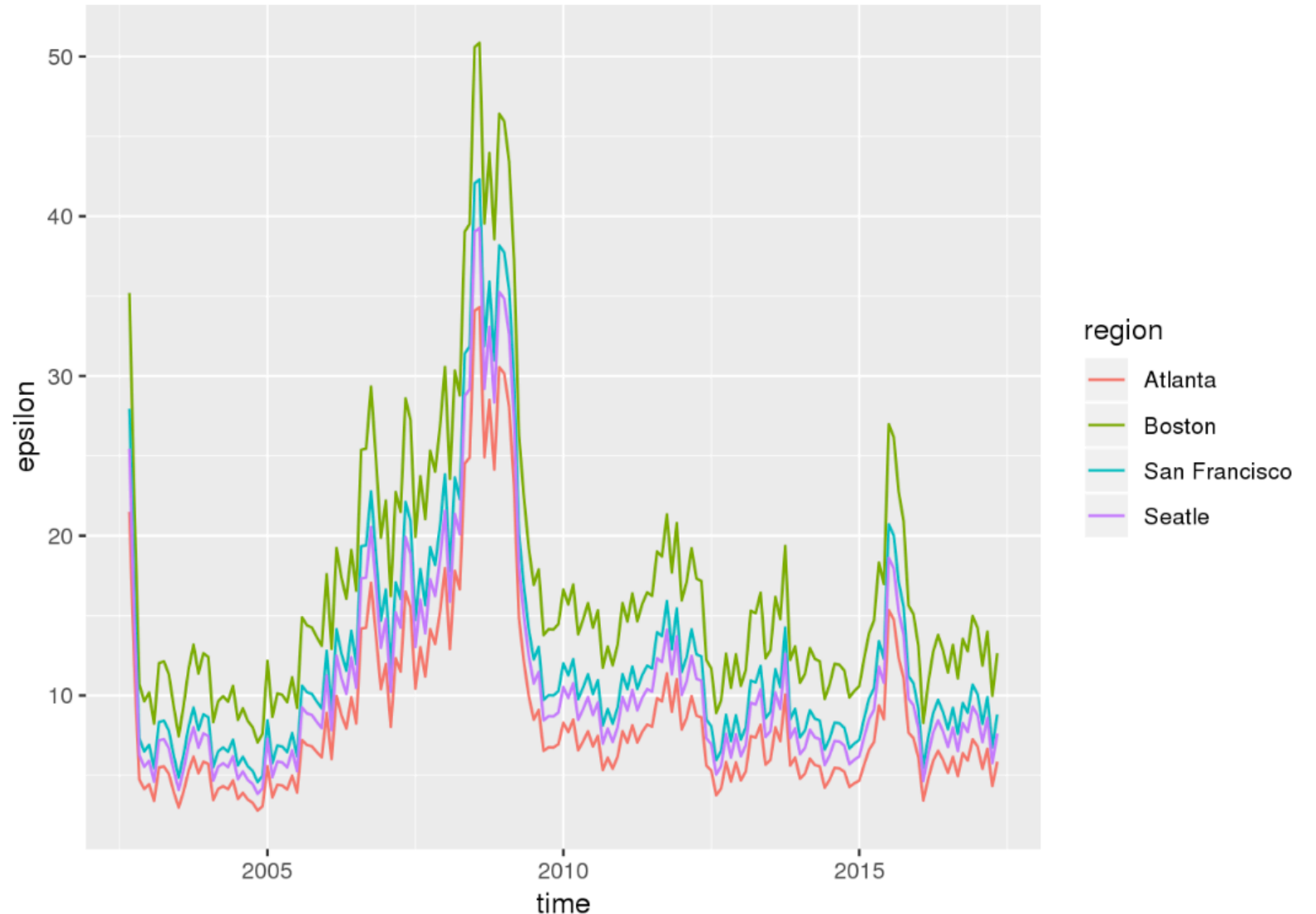
It also causes problems when I do global searches of the parameters.

# Local optimum and identifiability of parameters

Epsilon: duration of immunity (years)

The best model tells me the response to evolution in duration of immunity should be different across regions.

However, it is possible that this is a tradeoff between transmission rate and duration of immunity



# Summary

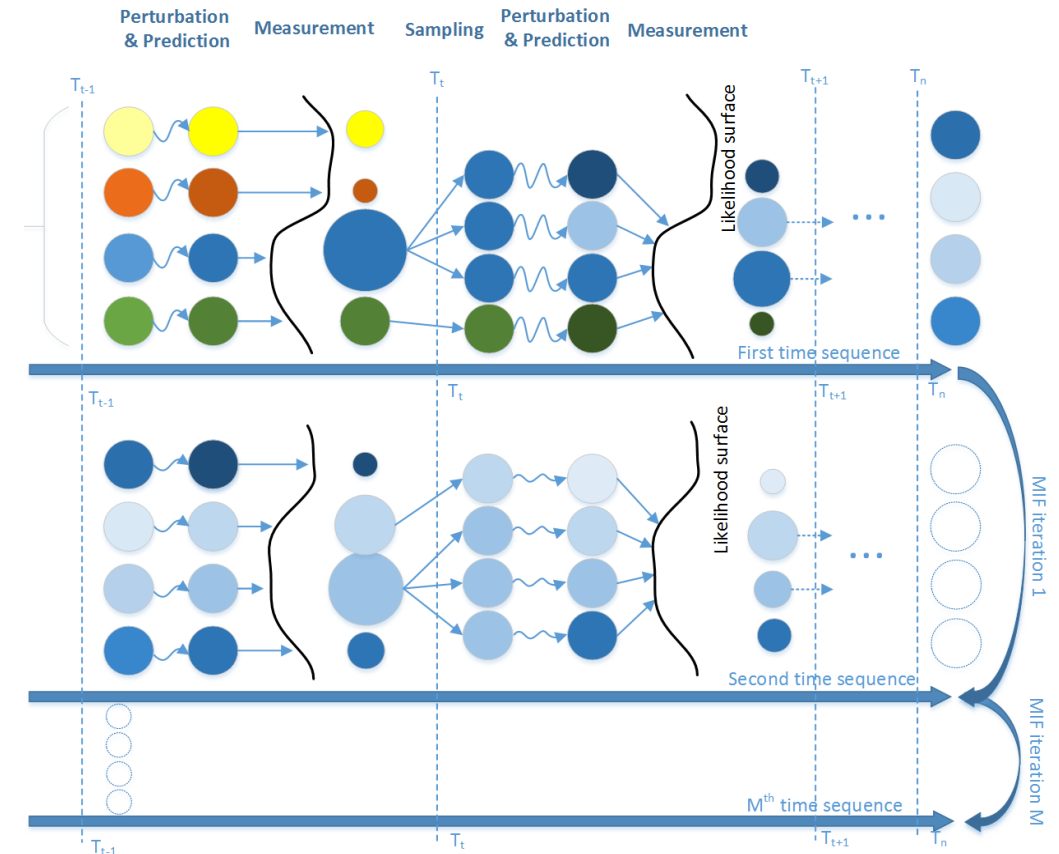
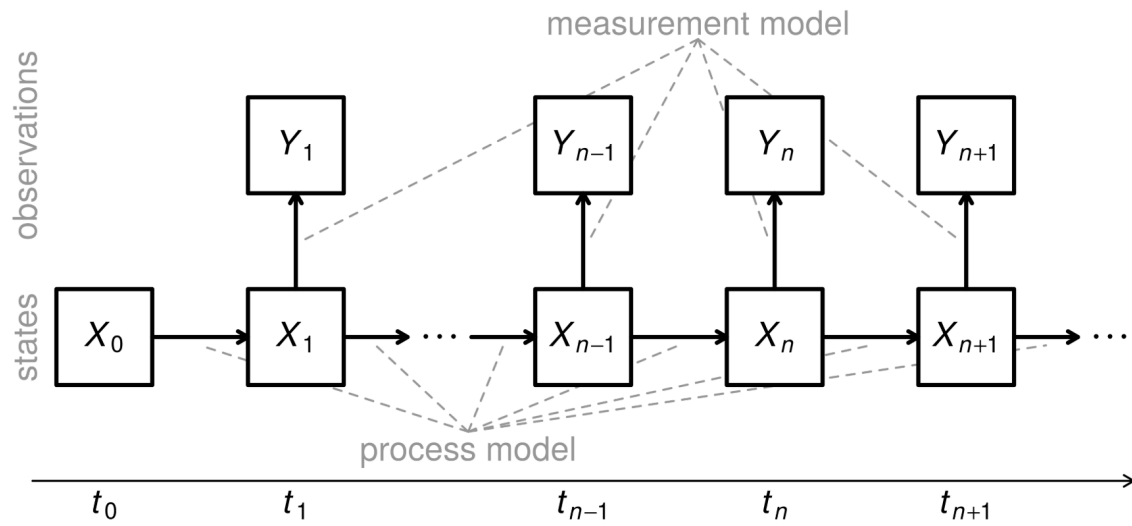
I have done:

- Data collection and preprocessing for regional incidence data
- Build models for the 10 regions with different combinations of shared and specific parameters
- Parameter searching
- Model selection and parameter interpretation

# POMP process and MIF method

POMP: Partially Observed Process Model

MIF: Sequential Monte Carlo method  
(models with noise; hidden variables)



E. Ionides and A. King , UM  
R Package pomp

# Equations

Transmission rate

seasonality

$$\beta(t) = \exp \left[ \sum_{i=1}^6 w_i s_i + w_\beta E(t) \right] \frac{d\Gamma}{dt}$$

process noise

evol. index

$$\frac{dS}{dt} = \left( \mu N(t) + \frac{dN(t)}{dt} \right) - \beta(t) S \left( \frac{I}{N(t)} \right)^\alpha + \frac{R}{\epsilon(t)} - \tau - \mu S - \Lambda_{H1}(t) + \frac{R_{H1}}{\epsilon_{H1}} \quad (1)$$

Process model

Measurement noise  
(normal error)  
>> Likelihood

$$\frac{dI}{dt} = \beta(t) S \left( \frac{I}{N(t)} \right)^\alpha + \tau - \frac{I}{\gamma} - \mu I \quad (2)$$

Duration of immunity

$$\frac{dR}{dt} = \frac{I}{\gamma} - \frac{R}{\epsilon(t)} - \mu R \quad (3)$$

$$\epsilon(t) = c_\epsilon e^{-w_\epsilon E(t)}$$

# Incidence data

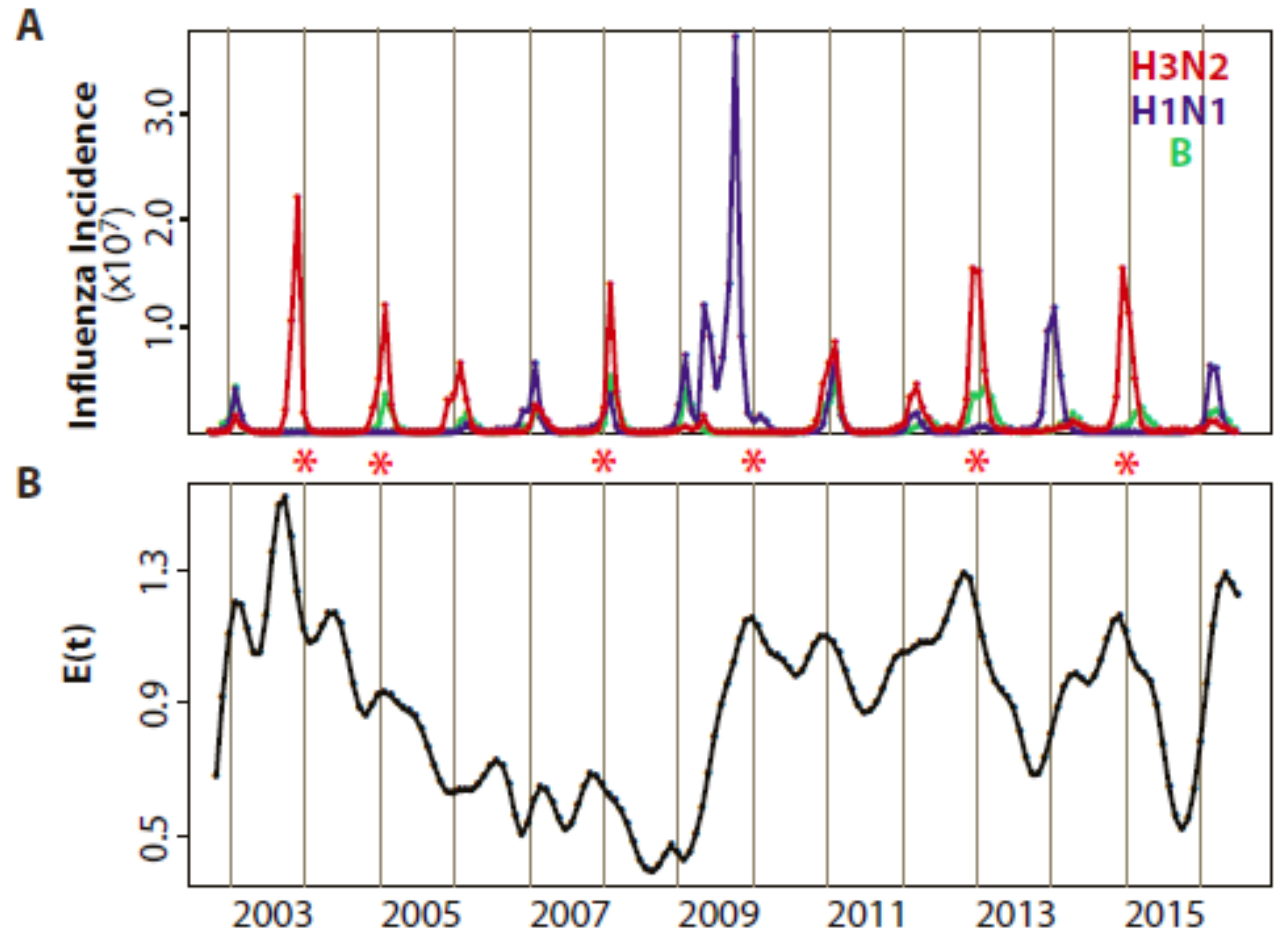
Outpatient illness surveillance data and viral surveillance data from FluView of the US Centers for Disease Control and Prevention (CDC)

**Incidence** = ILI (influenza like illness) percentage

**X** proportion of ILI samples that test positive for influenza

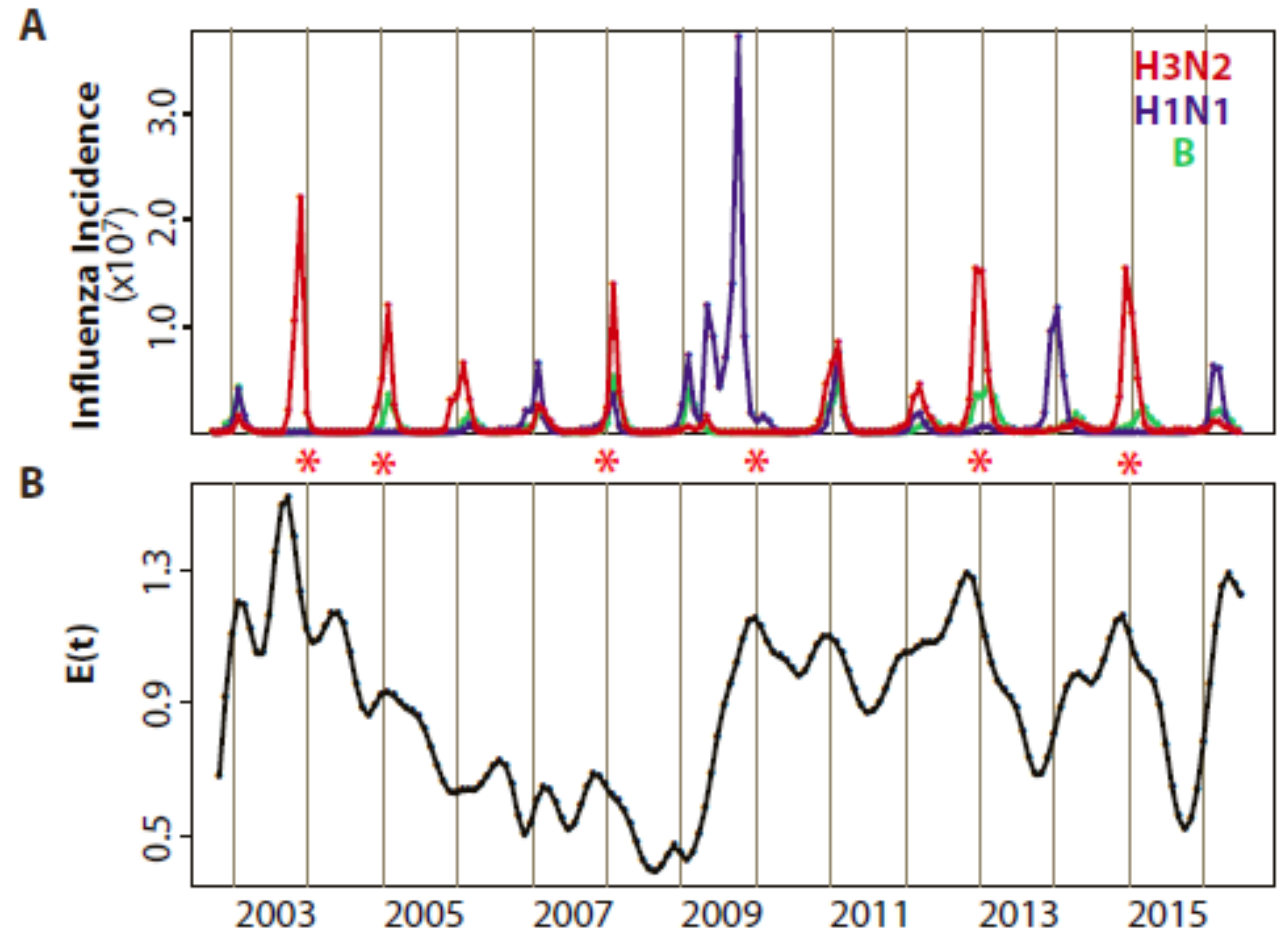
**X** subtype specific proportion

**X** population size



# Evolutionary Index

$$E(t) = \sum_{s=0}^{19} d(t,s) e^{-\frac{s+1}{\theta}}$$





# Acknowledgement

I would like to thank Professor Pascual and Xiangjun Du for their general guidance on their project, Rahul for discussing with me many times on the problem of parameter estimation, everyone in Pascual Lab for helping me with preparing for this presentation, and everyone in this fellowship program for creating this wonderful environment for learning and working.

This work was supported in part by the BSCD Summer Fellowship in Quantitative Biology, funded by BSD/BSCD Curriculum Innovation Fund at the University of Chicago.”

# Citation

- [1]"Climate of 100 Selected U.S. Cities." Infoplease.  
<<https://www.infoplease.com/science-health/weather/climate-100-selected-us-cities/>>.
- [2] Xiangjun Du, Aaron A. King, Robert J. Woods and Mercedes Pascual, Evolution-informed forecasting of seasonal influenza A (H3N2)
- [3] Slides from Professor Pascual and Rahul Subramanian
- [4] Aaron King and Dao Nguyen and Edward Ionides, Statistical Inference for Partially Observed Markov Processes via the R Package pomp, Journal of Statistical Software, volume 69, number 12, 2016