A patch SIR model for predicting flu outbreaks using airport usage statistics

Michelle Chen

Mentor: Dorothy Wallace

Dartmouth College

Hanover, NH 03755

Introduction

- March/April 2009: H1N1 influenza was detected in Southern California and Texas
- Outbreaks may be linked to foreign travel to/from Mexico
- Hypothesis: H1N1 may have been transmitted from within the U.S. via large regional airline hubs.

Why use airports?

- One study indicates that properties of the statistical distribution of travel may influence rates of disease propagation.

Regions

- New York (NY): New York City and Newark
- Colorado (CO): Denver
- California (CA): Los Angeles, San Francisco, Oakland, and Santa Ana
- Texas (TX): Dallas and Houston
- Georgia (GA): Atlanta
- Illinois (IL): Chicago

Classic Model

- SIR Model for large homogeneous well-mixed populations
- Regions act as patches and are connected by migration.
- Accuracy is dependent on rates.

Generic Model

- Growth Rate: $(a-b)(1-S_i/k_i)-dS_i$

Parameter	Description
a	Natural birth rate
b	Natural death rate
d	Disease death rate
k_i	Equilibrium population at the disease free equilibrium corresponding to the actual population of city i.

General Equations

$$S_i^{'} = (a-b)(S_i)(1 - S_i / k_i) - BS_iI_i /(S_i + I_i + R_i) - \sum_{j \sim i} m_{(i,j)}S_i + \sum_{j \sim i} m_{(j,i)}S_j$$

$$I_i^{'} = BS_iI_i /(S_i + I_i + R_i) - (v + b + d)I_i - \sum_{j \sim i} m_{(i,j)}I_i + \sum_{j \sim i} m_{(j,i)}I_j$$

$$R_i^{'} = vI_i - bR_i - \sum_{j \sim i} m_{(i,j)}R_i + \sum_{j \sim i} m_{(j,i)}R_j$$

Parameter	Description
B	Transmission rate
v	Recovery rate

Sample Equations

- dNY = (a-b)*NY*(1-NY/19151072) + m_CO_NY*CO + m_CA_NY*CA + m_TX_NY*TX + m_GA_NY*GA + m_IL_NY*IL - m_NY*NY - B*NY*I_NY/(NY+I_NY+R_NY)

- dI_NY = B*NY*I_NY/(NY+I_NY+R_NY) - (d+b+v)*I_NY + m_CO_NY*I_CO + m_CA_NY*I_CA + m_TX_NY*I_TX + m_GA_NY*I_GA + m_IL_NY*I_IL - m_NY*I_NY

- dR_NY = v*I_NY - b*R_NY + m_CO_NY*R_CO + m_CA_NY*R_CA + m_TX_NY*R_TX + m_GA_NY*R_GA + m_IL_NY*R_IL - m_NY*R_NY

Migration Rates

- Assumption: traffic to and from destinations other than the six destinations modeled was negligible
- Equation: $m_{state1\_state2} = q_{state1} * q_{state2} / \sum_{state1}$

| Parameter	Description
m_{state1_state2}	Migration from state 1 to state 2
q_{state}	Total average daily traffic through the state's airports
$\sum_{state1}$	Average daily traffic through the five remaining airports in the model (exclude state 1)

Migration Rates

- Sample Equation:

$$m_{NY} = m_{NY\_CO} + m_{NY\_CA} + m_{NY\_TX} + m_{NY\_GA} + m_{NY\_IL}$$

Parameter	Description
m_state	Migration away from the state

Graphical Depiction
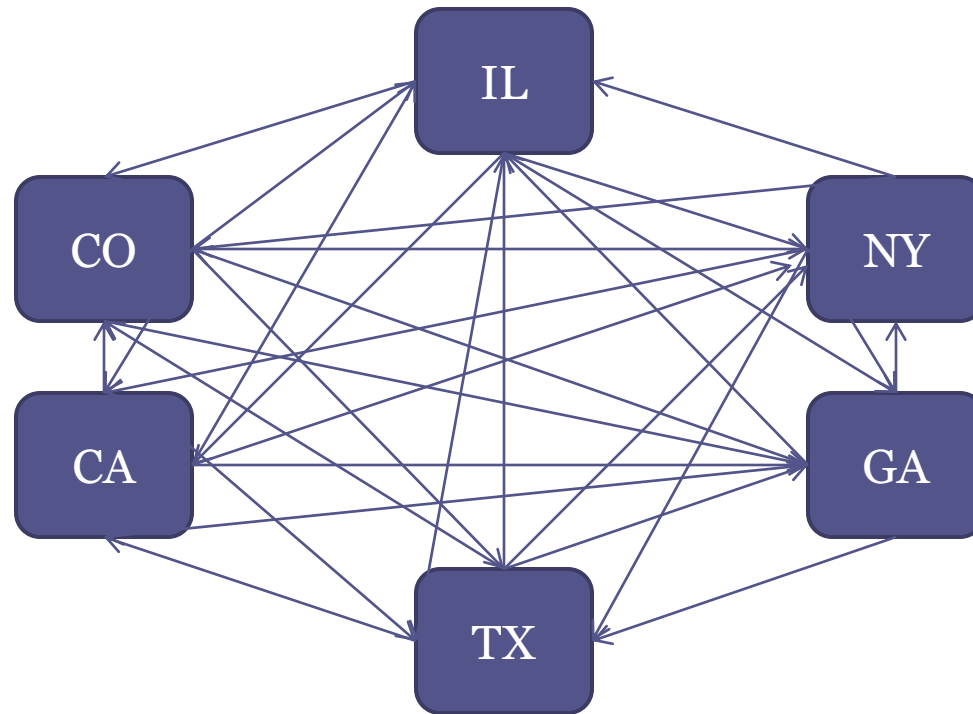
Other Rates

- Transmission Rate: 1.5 new cases per three days
- Recovery Rate: estimated from CDC data to be 14 days
- Birth Rate & Death Rate: CDC data from 2010

Table of Values

Parameter	Value	Parameter	Value
NY	19151072	m_CO_IL	4.23E-05
CO	2599235	m_CA_NY	0.000137123
CA	17293449	m_CA_CO	6.72E-05
TX	12564909	m_CA_TX	0.000141255
GA	5540092	m_CA_GA	0.000117897
IL	9622245	m_CA_IL	0.000109452
m_NY_CO	5.93E-05	m_TX_NY	0.000125624
m_NY_CA	0.000133771	m_TX_CO	6.16E-05
m_NY_TX	0.00012474	m_TX_CA	0.000138777
m_NY_GA	0.000104113	m_TX_GA	0.00010801
m_NY_IL	9.67E-05	m_TX_IL	0.000100273
m_CO_NY	5.30E-05	m_GA_NY	0.000100813
m_CO_CA	5.86E-05	m_GA_CO	4.94E-05
m_CO_TX	5.46E-05	m_GA_CA	0.000111369
m_CO_GA	4.56E-05	m_GA_TX	0.00010385

Table of Values

Parameter	Value	Parameter	Value
m_GA_IL	8.05E-05	I_NY	0
m_IL_NY	9.23E-05	I_CO	0
m_IL_CO	4.52E-05	**I_CA**	**2**
m_IL_CA	0.000101972	**I_TX**	**1**
m_IL_TX	9.51E-05	I_GA	0
m_IL_GA	7.94E-05	I_IL	0
m_NY	0.000518612	R_NY	0
m_CO	0.000254107	R_CO	0
m_CA	0.000572914	R_CA	0
m_TX	0.000534237	R_TX	0
m_GA	0.000445897	R_GA	0
m_IL	0.000413958	R_IL	0
a	0.0135	d	0.01043575
b	0.008036	v	0.07142857
B	**0.5**		

pop_test_6

- Used to test the migration rates with no disease present to verify that the populations of the regions remained at equilibrium.
- Key Difference: the model included only the susceptible population.
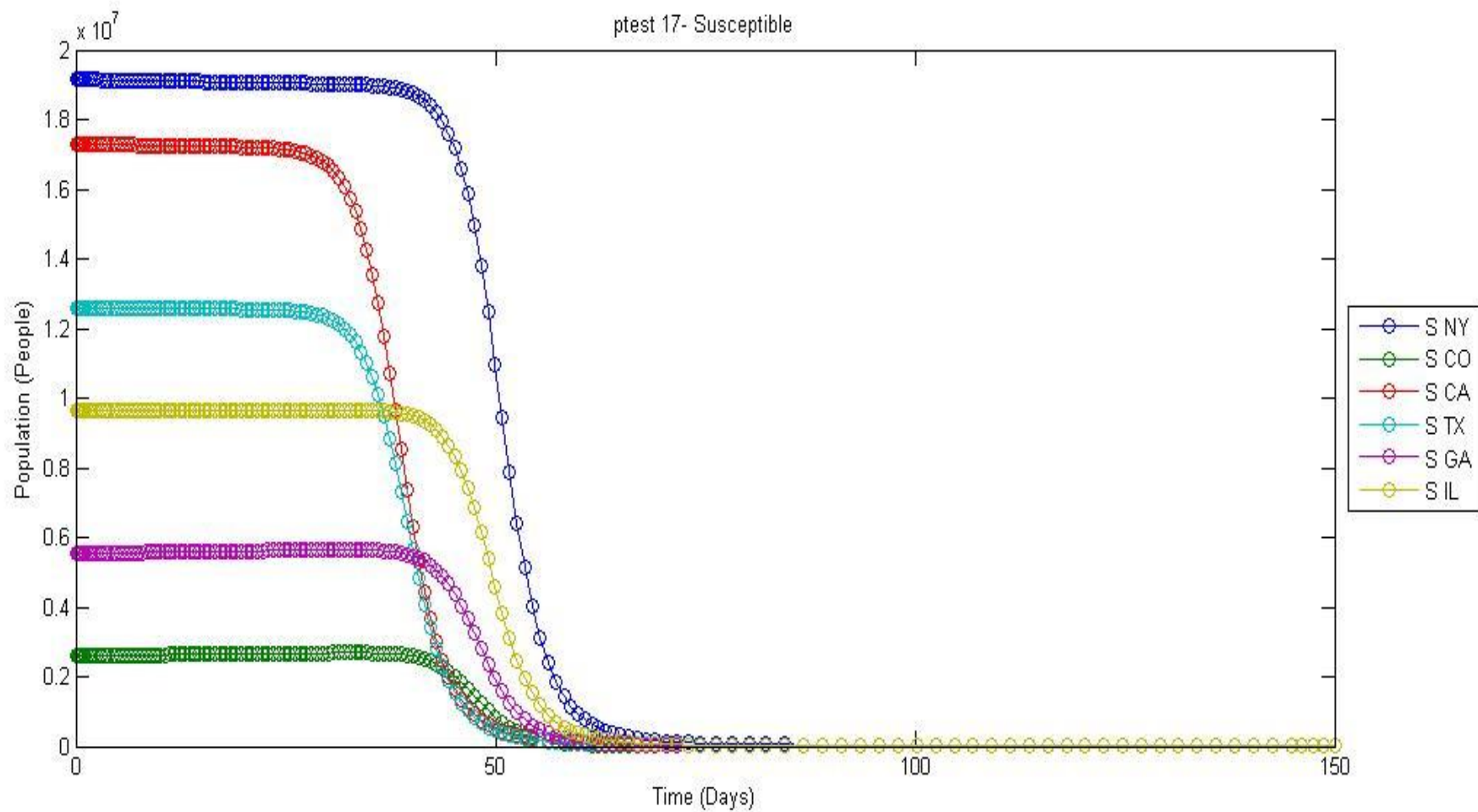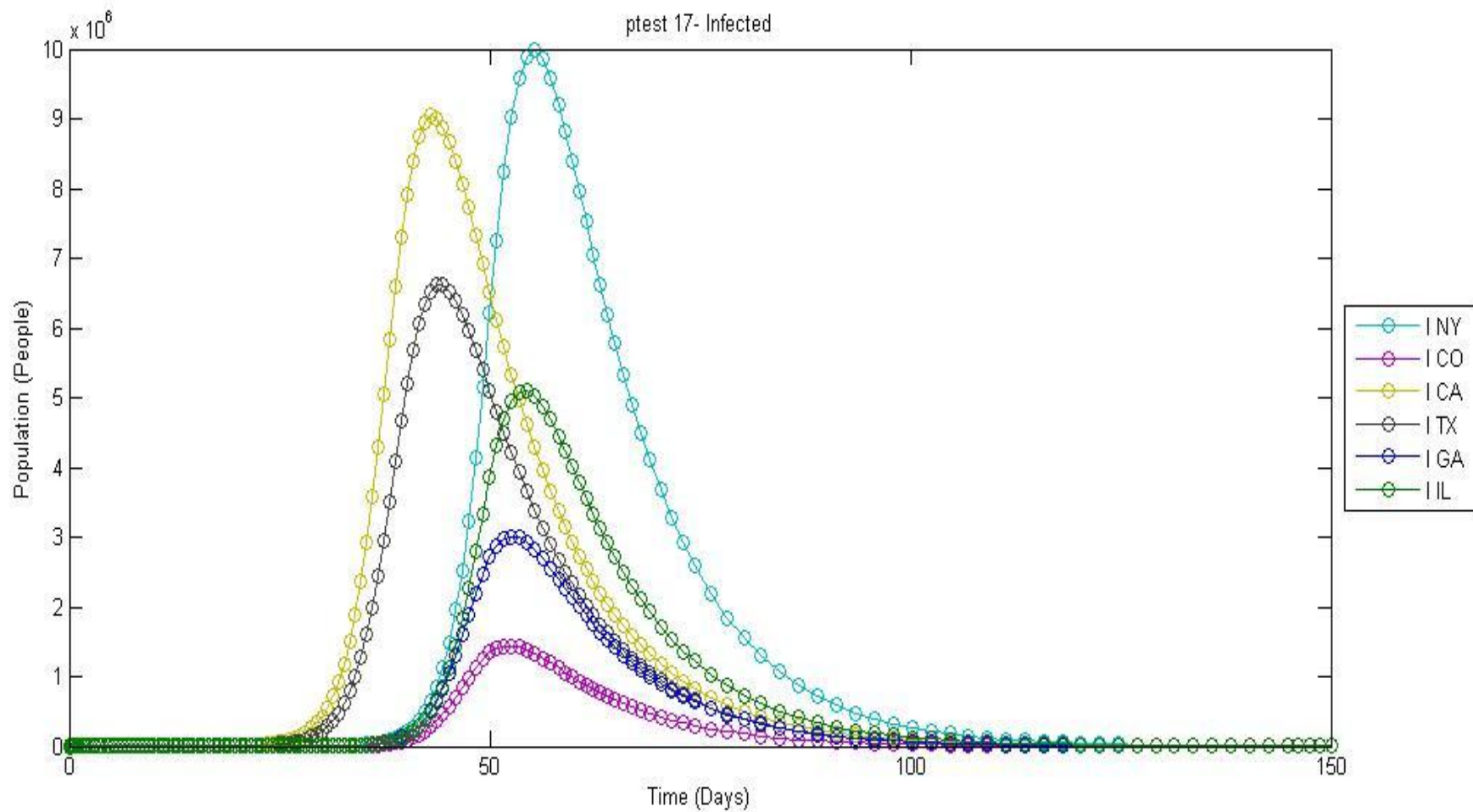
pop_test_6
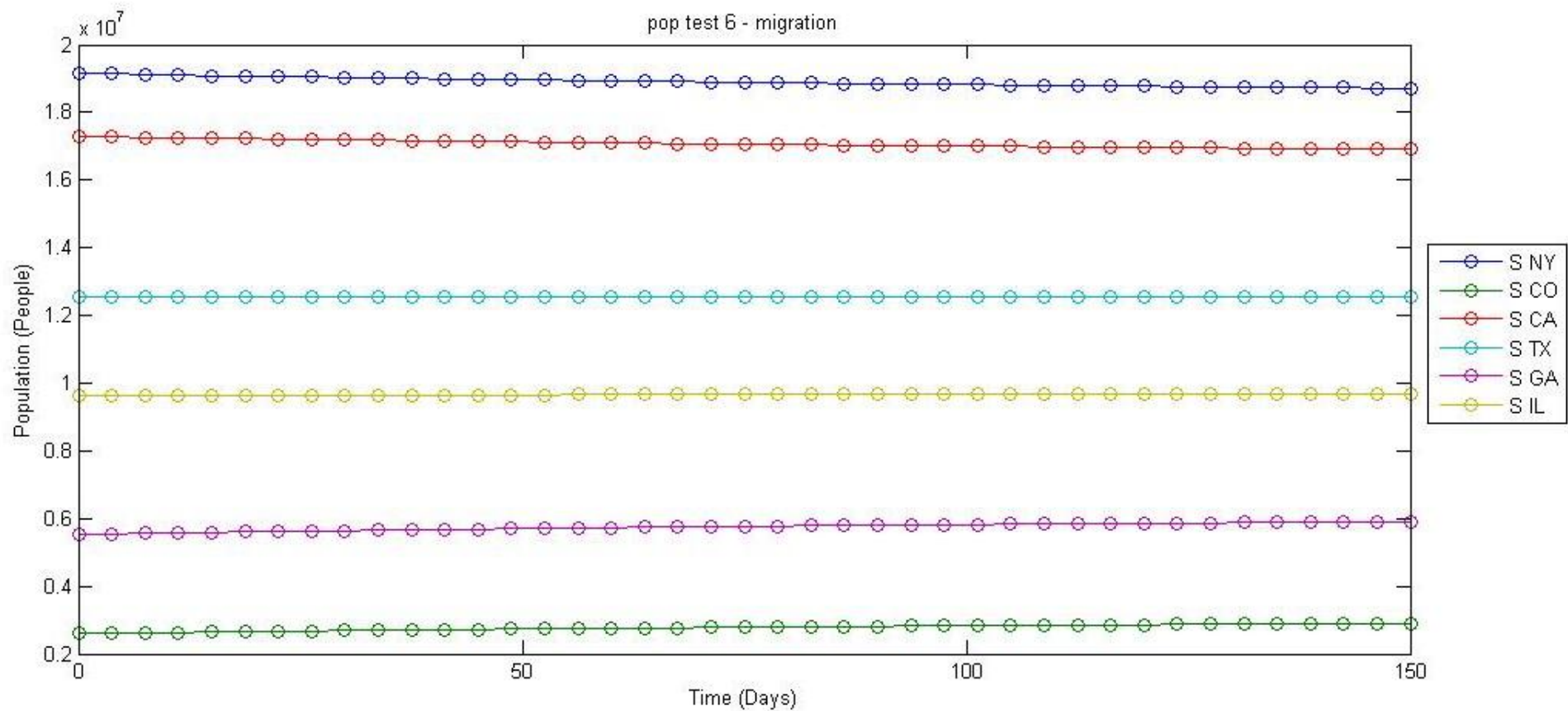
Two Versions of the Model

- Both models were run with disease introduced in California and Texas
- Difference: In the second model, birth rates depended on the entire population of the region rather than on just the susceptible individuals.
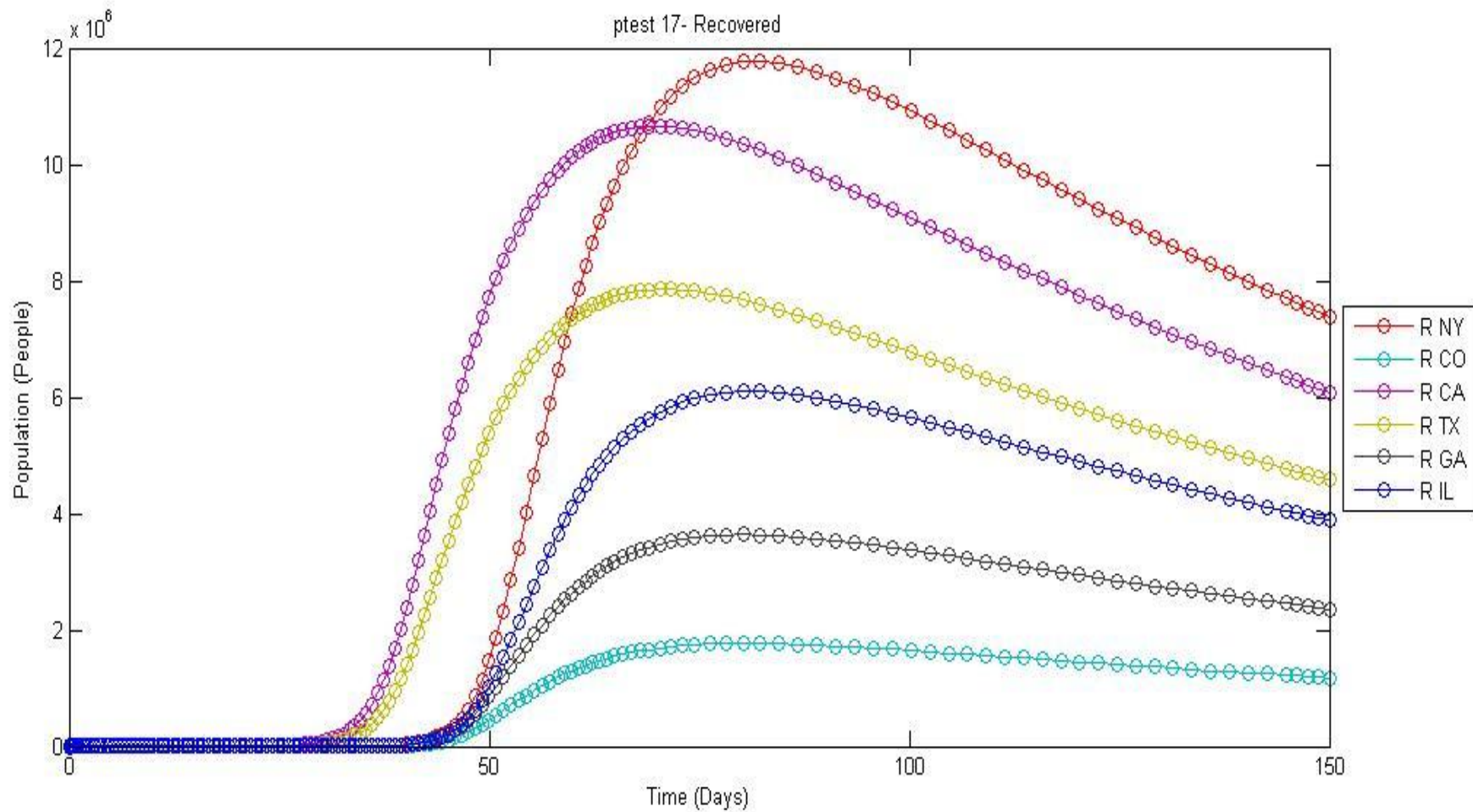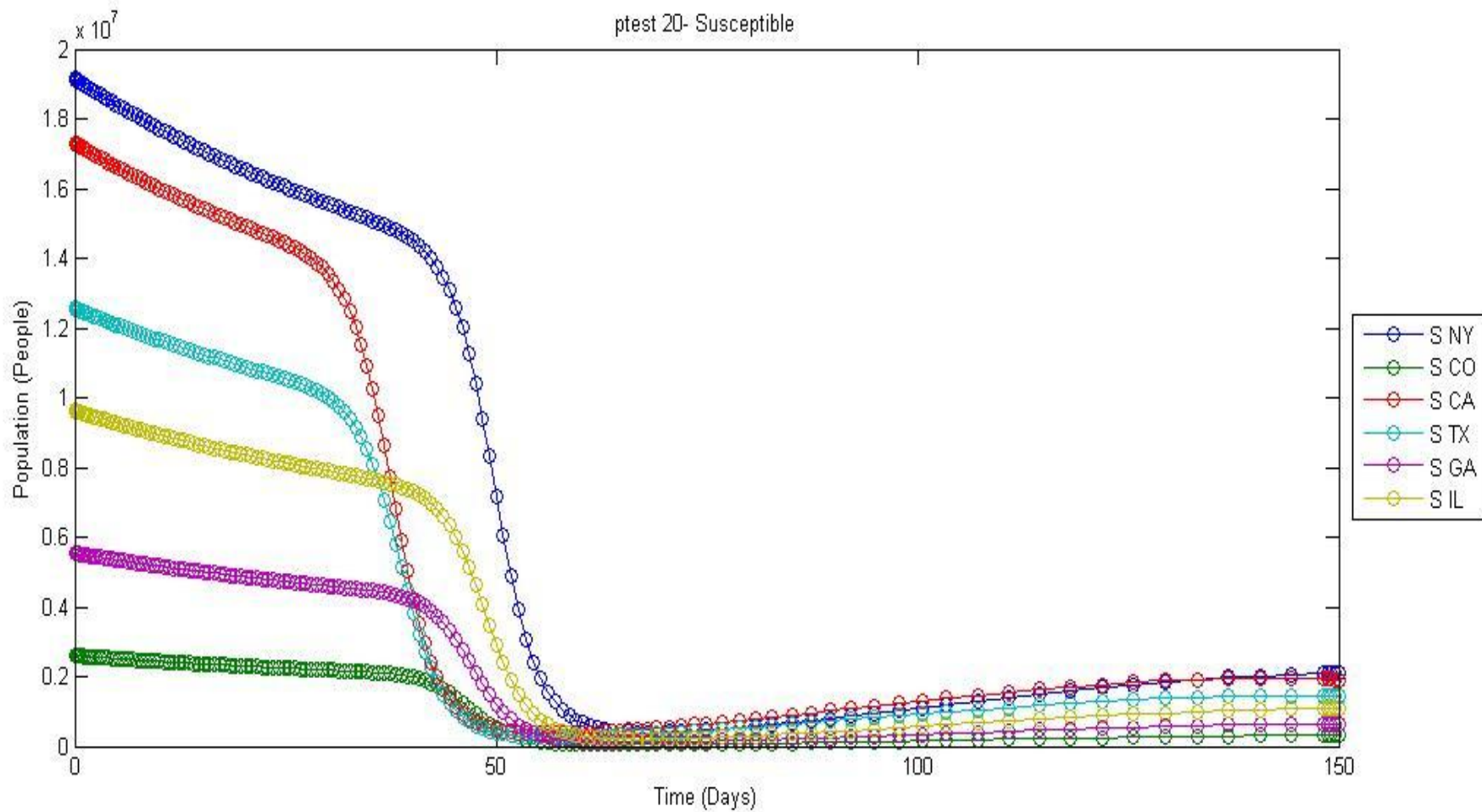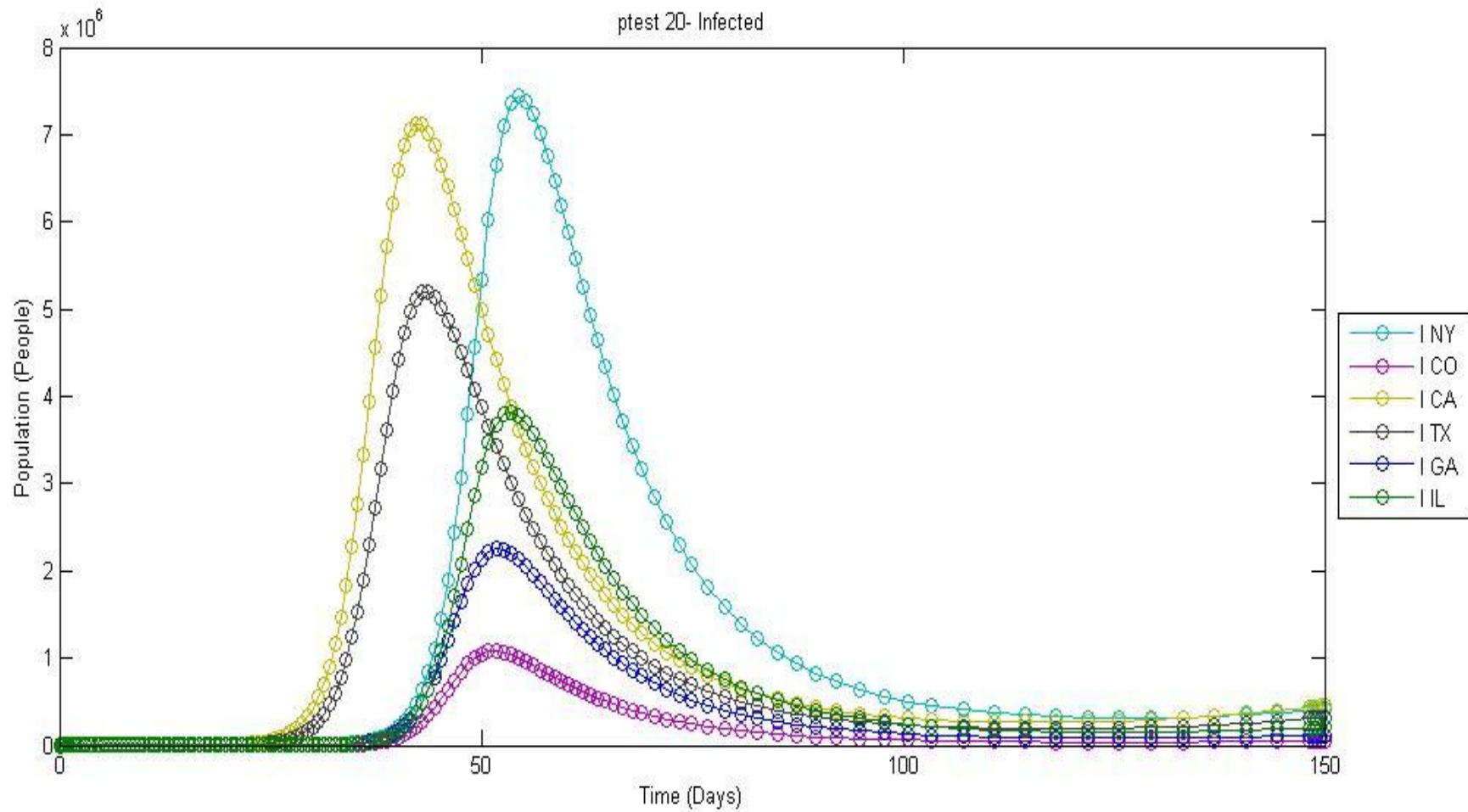
ptest_17 Susceptible

ptest_17 Infected
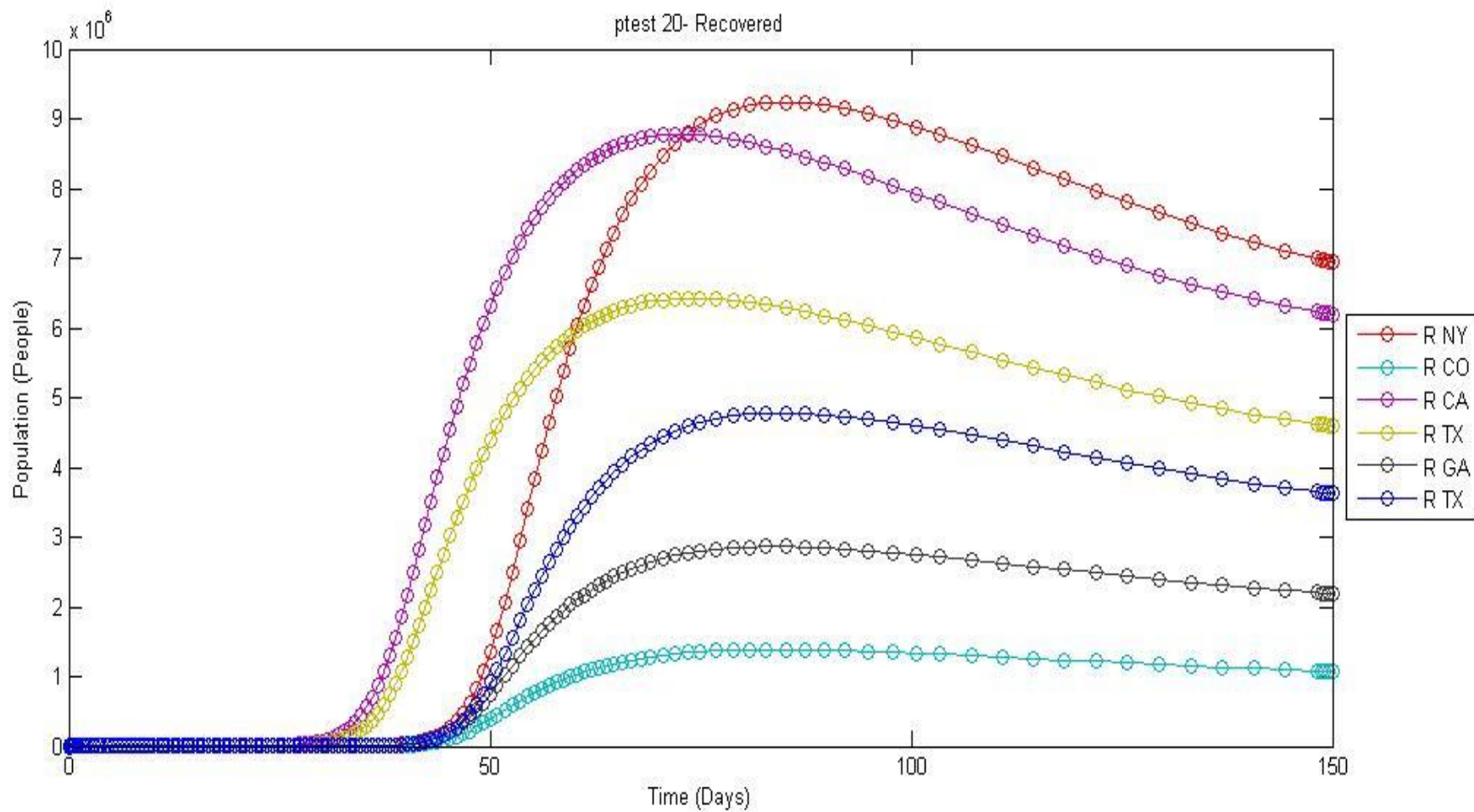
ptest_17 Recovered

ptest_20 Susceptible

ptest 20- Susceptible

ptest_20 Infected

ptest_20 Recovered

Comparison to Known Data

- Models were compared with patient data in Regions 2, 4, 5, 6, 8, and 9 that report to the Center for Disease Control (CDC)

Region	States
2	**New York**, New Jersey
4	Kentucky, Tennessee, North Carolina, South Carolina, **Georgia**, Florida, Alabama, Mississippi
5	Minnesota, Wisconsin, Michigan, **Illinois**, Indiana, Ohio
6	New Mexico, **Texas**, Olkahoma, Louisiana, Arkansas
8	Utah, **Colorado**, Wyoming, Montana, South Dakota, North Dakota
9	**California**, Nevada, Arizona

Disease Prevalence

- The CDC reported 43,677 laboratory confirmed cases of H1N1 through July 2009.
- The estimated under reporting is 79 actual infections per reported case.
- The CDC's statistical model estimated 1.8 million to 5.7 million cases occurred

Peak Prevalence Results

Region	Peak Value (Model)	Peak Value (Data)	79 * Peak Value (Data)
NY	9,981,985	275	21,725
CO	1,442,394	2,134	168,586
CA	9,042,460	1,192	94,168
TX	6,614,867	652	51,508
GA	3,010,005	815	64,385
IL	5,106,028	1,763	139,277
National	28,000,000	10,050	793,950

Comparison Using an Integration Method

- Assumption: 14 day duration of disease
- First 15 weeks of data for each region.
- Total cases are multiplied by 79 as recommended.
- Model predicts values about 15 times as large as the statistically adjusted data.

Peak Prevalence Results Using Integration

| Region	79 * Total Cases (Data)	Integrated Value (Model)/14
NY	236,052	14,975,307
CO	103,964	2,192,060
CA	474,869	13,758,223
TX	278,870	10,087,391
GA	514,053	4,551,770
IL	643,455	7,691,961
National	3,077,919	53,256,713

Timing of Peak Prevalence

- Errors for the default transmission rate were within a couple of weeks.
- Exceptionally Poor Matches: CA and TX
- If you adjust the transmission rate for a single state could often bring the model and data into agreement.
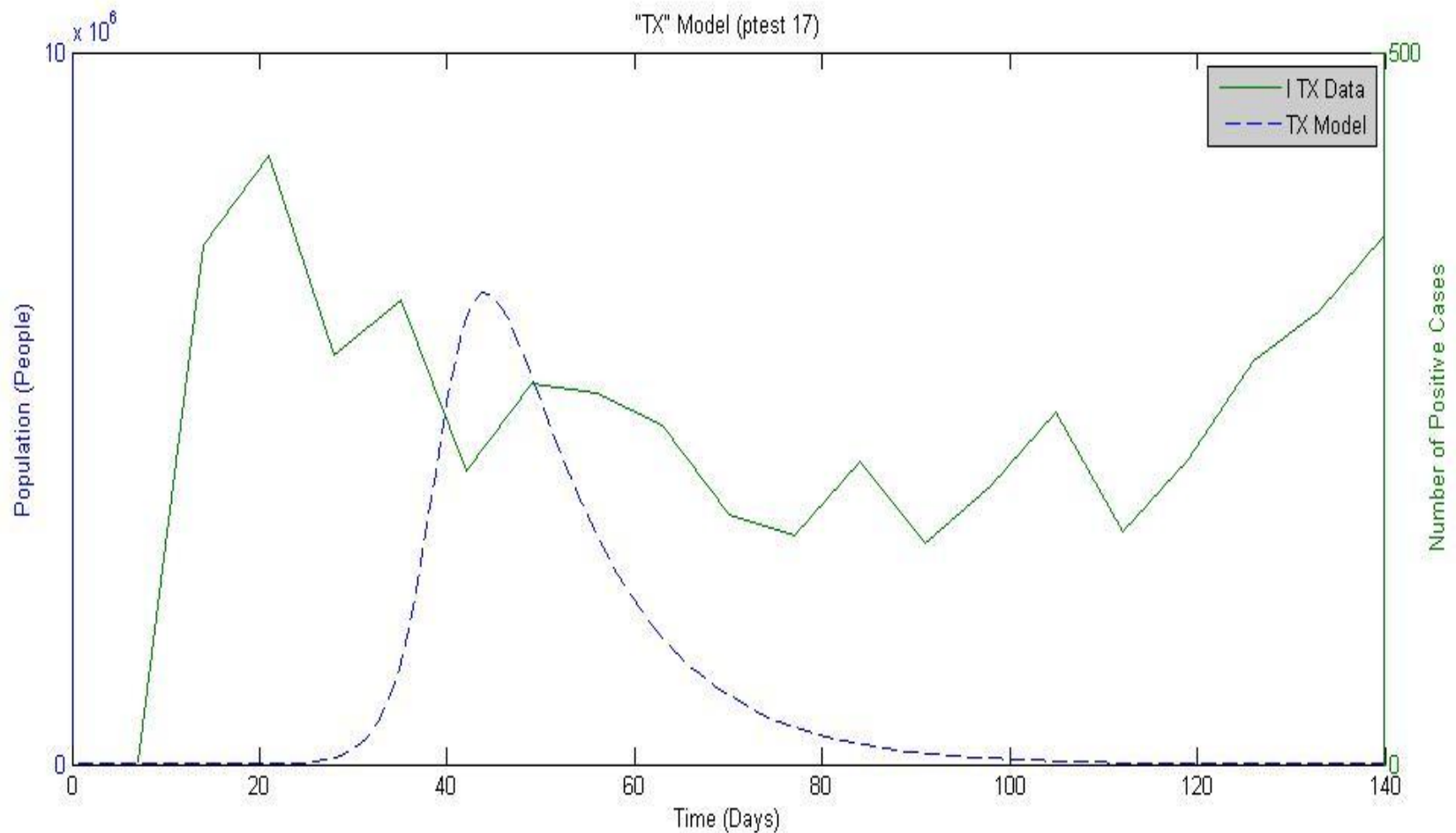
Example of a Good Match
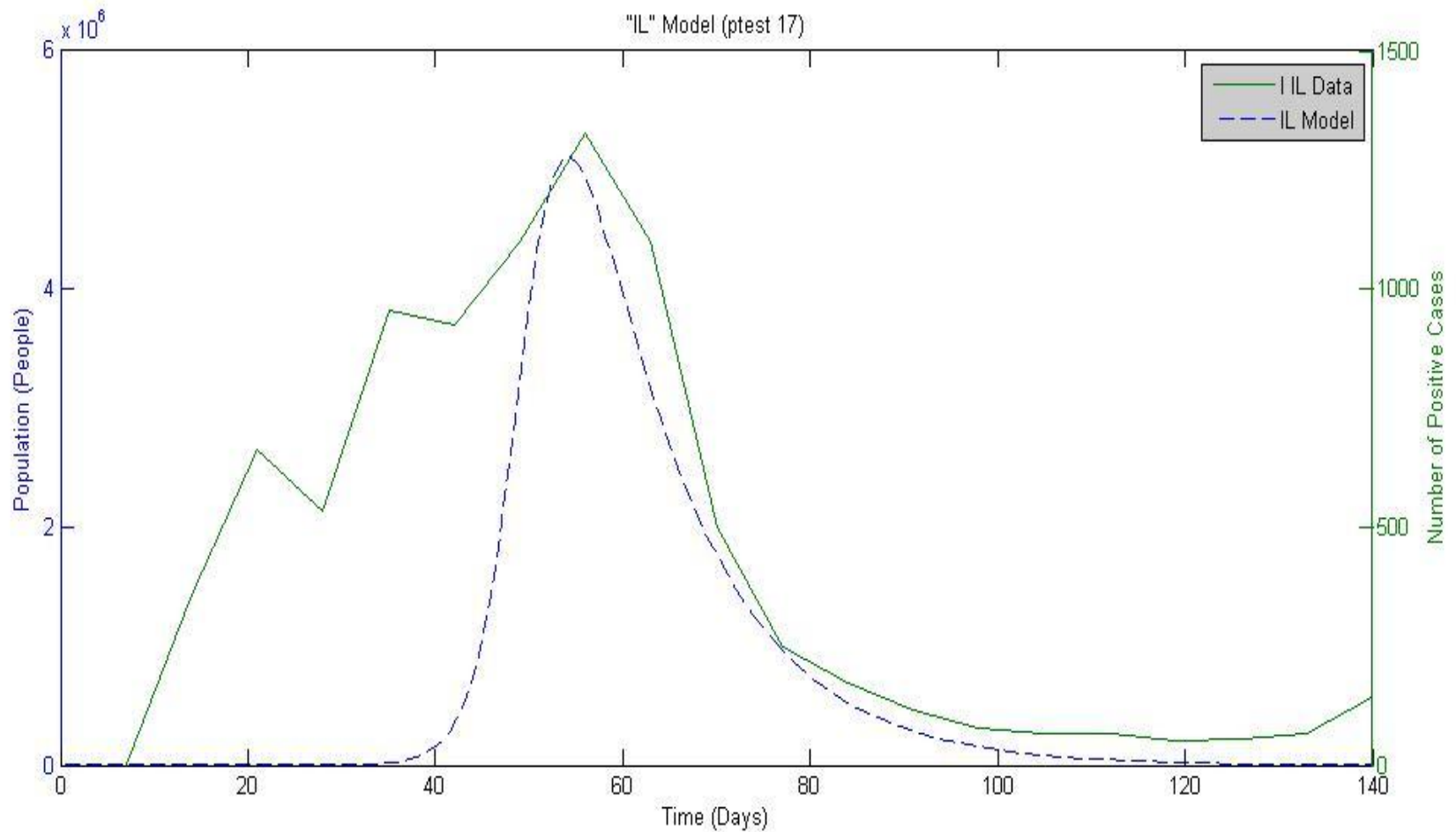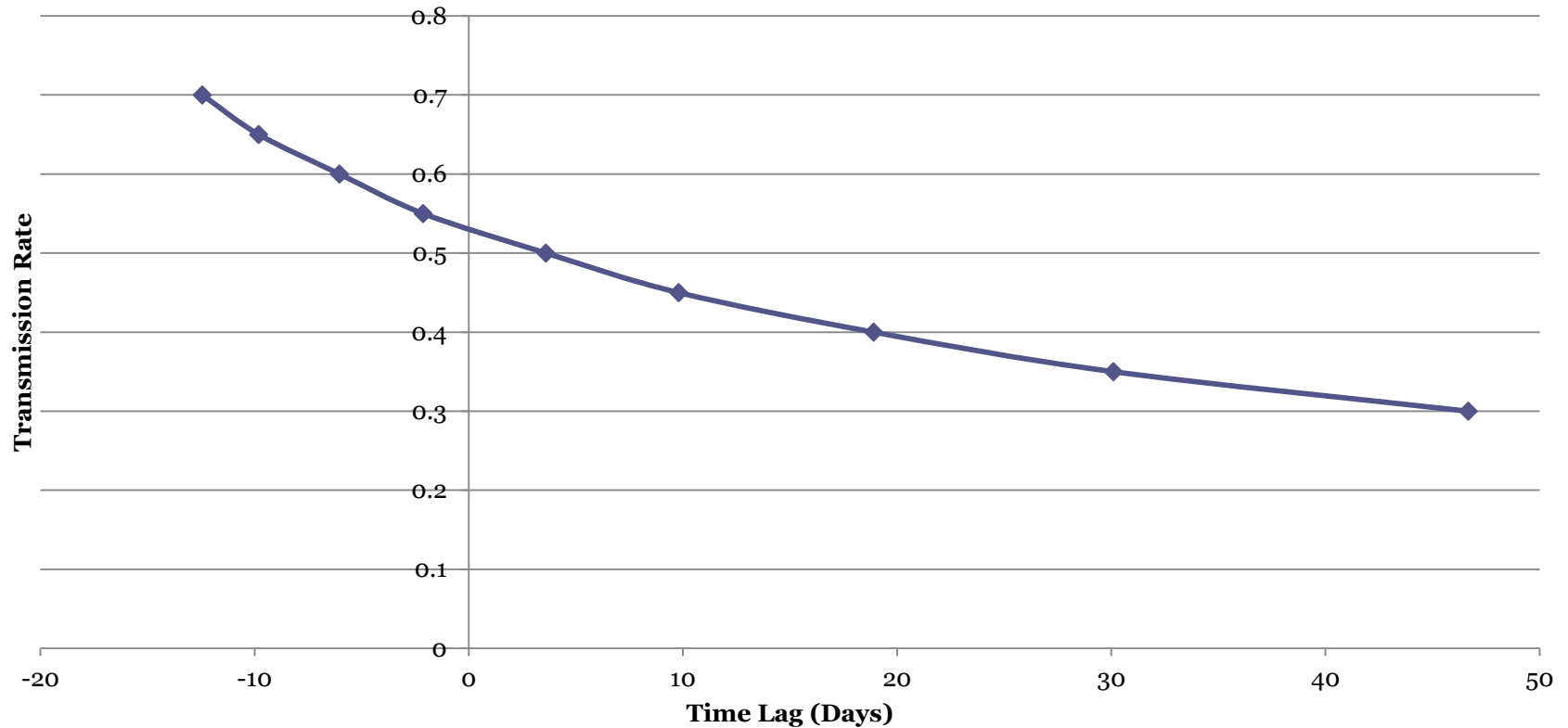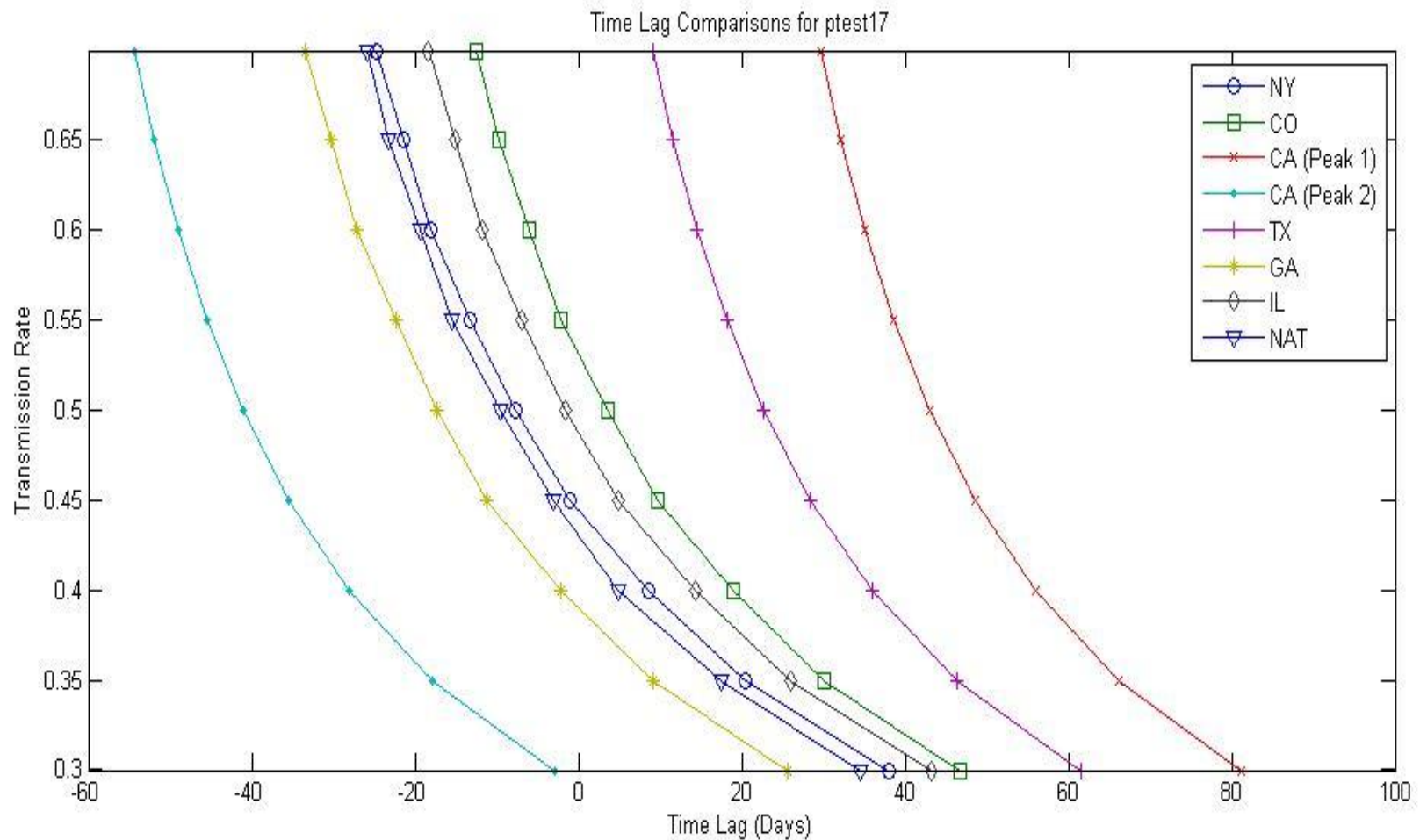
California's Poor Match
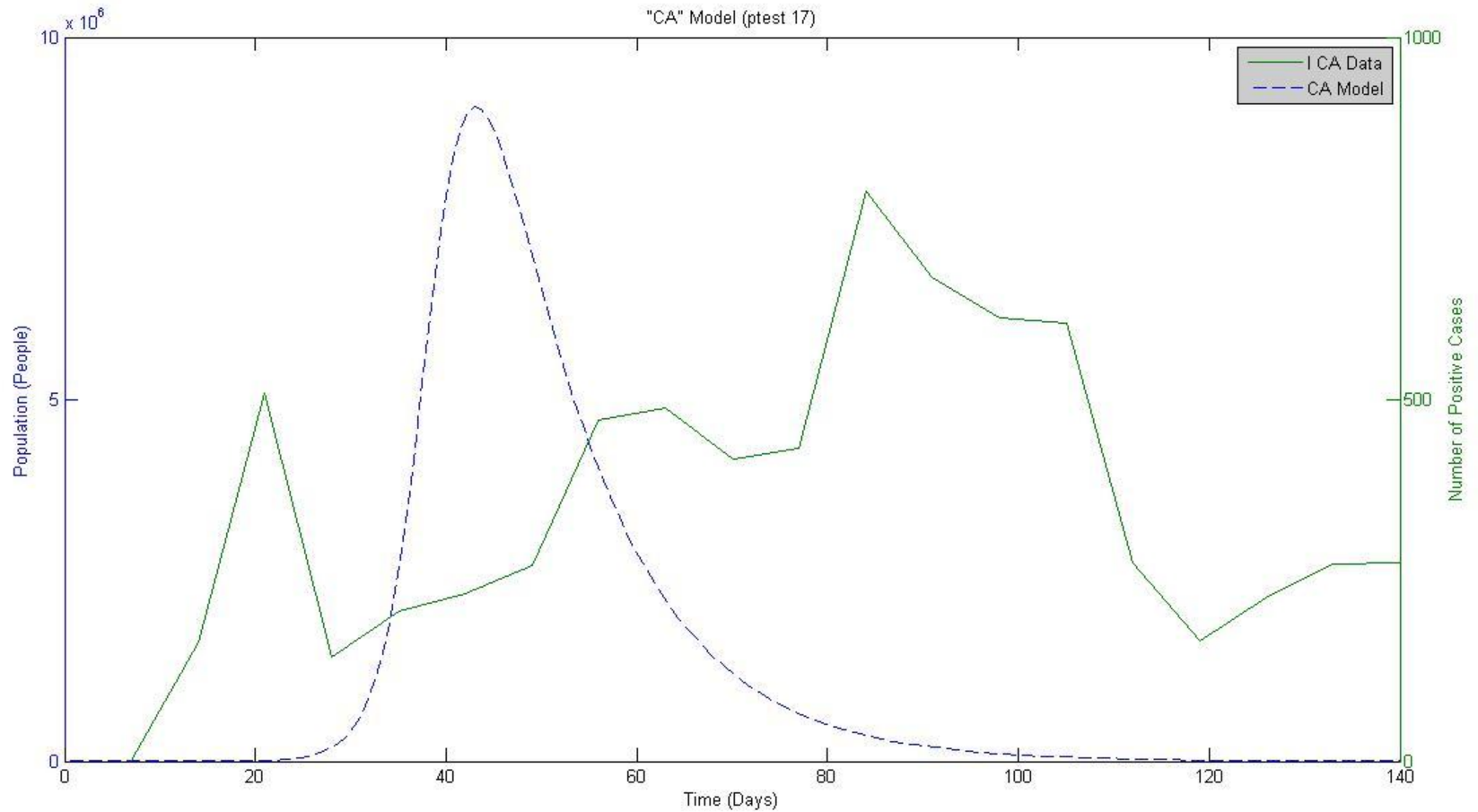
Texas's Poor Match

Time Lag v. Transmission Rates

Time Lag v. Transmission Rates for CO (ptest_17)

Time Lag v. Transmission Rates

Time Lag Comparisons for ptest17
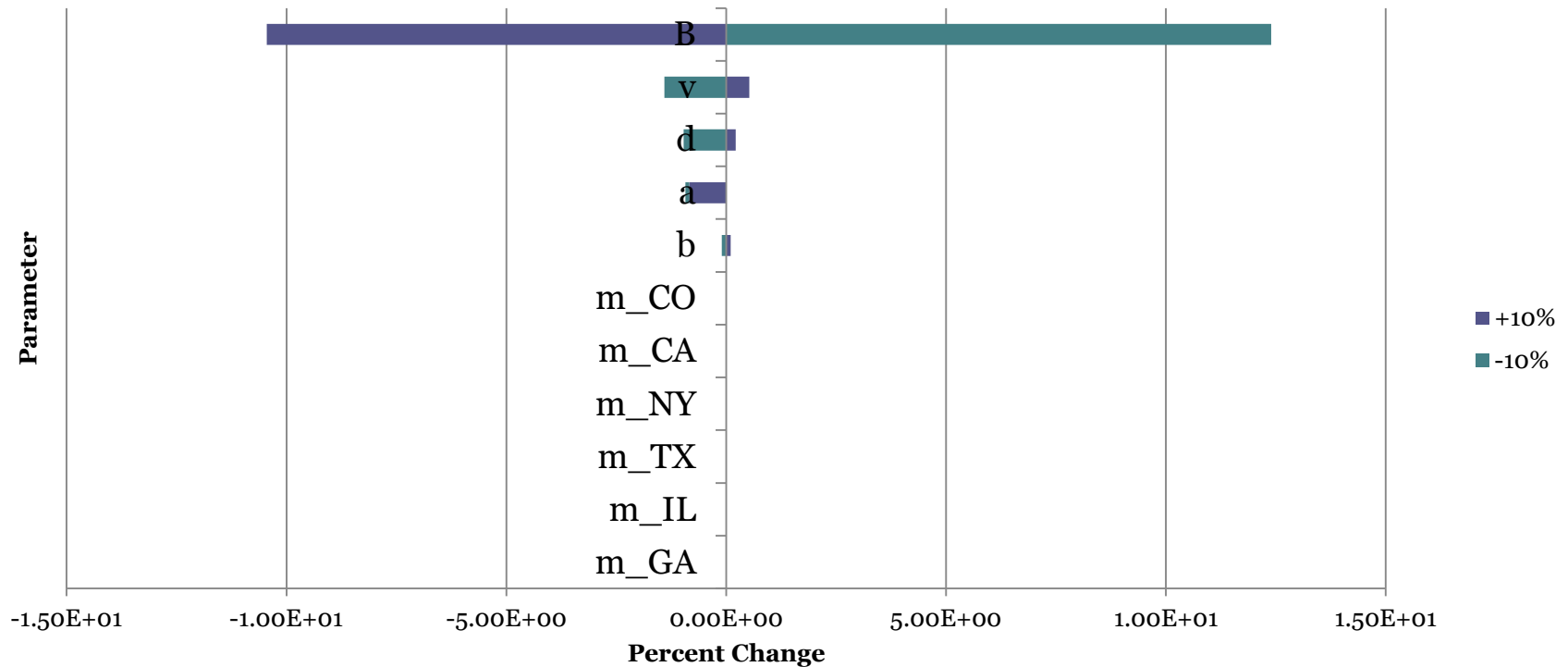
Sensitivity Analysis

- This was conducted on parameters and initial conditions with respect to the error observed between peak case load in the model versus data.
- Each parameter was varied up and down 10%.
- In general, disease related parameters had a much larger effect than migration rates.

Sensitivity Analysis for Peak Time
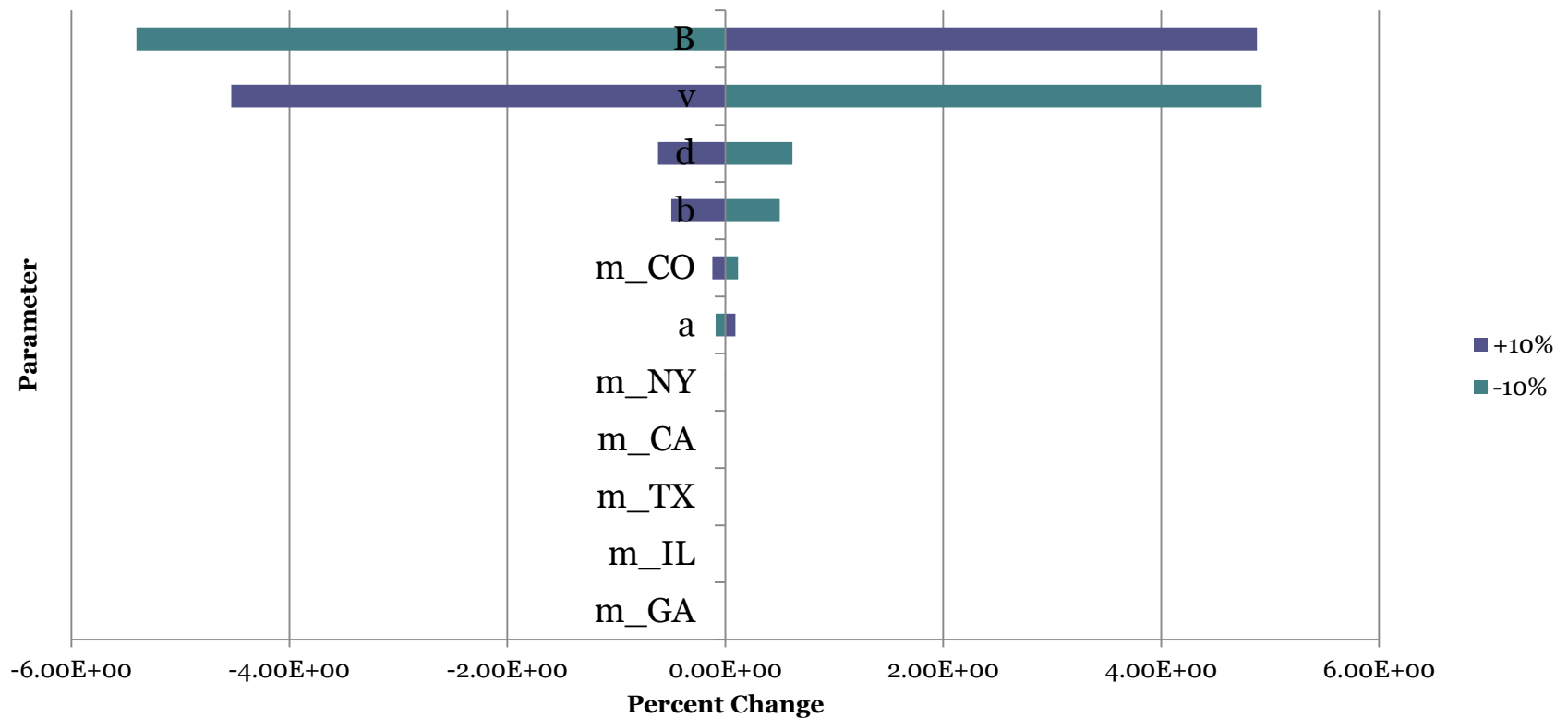

Sensitivity Analysis Comparing CO Peak Days (ptest_17)

Sensitivity Analysis for Peak Time

Region	Top 5	Negligible Effect
NY	B, v, d, a, b	$m\_CO, m\_CA, m\_TX, m\_NY, m\_IL, m\_GA$
CO	B, v, d, a, b	$m\_CO, m\_CA, m\_NY, m\_TX, m\_IL, m\_GA$
CA	B, v, d, b, m_CA	$m\_NY,\ m\_CO, m\_TX, m\_IL, a, m\_GA$
TX	B, v, d, b, m_CA	$m\_NY,\ m\_CO, m\_TX, m\_IL., a, m\_GA$
GA	B, v, d, b, a	$m\_NY, m\_CA, m\_CO, m\_TX, m\_IL, m\_GA$
IL	B, v, d, b, a	$m\_NY,\ m\_CA, m\_CO, m\_IL, m\_TX, m\_GA$
NAT	B, v, m_NY, a, d	$b, m\_CO, m\_CA, m\_IL, m\_TX, m\_GA$

Sensitivity Analysis for Peak Value

Sensitivity Analysis Comparing CO Values (ptest_17)

Sensitivity Analysis for Peak Values

Region	Top 5	Negligible Effect
NY	B, v, d, a, b	$m\_CO, m\_CA, m\_TX, m\_NY, m\_IL, m\_GA$
CO	B, v, d, b, m_CO	$a, m\_NY, m\_CA, m\_TX, m\_IL, m\_GA$
CA	B, v, d, b, a	$m\_CA, \ m\_CO, m\_NY, m\_TX, m\_GA, m\_IL$
TX	B, v, d, b, a	$m\_TX, \ m\_CO, m\_CA, m\_NY. \ m\_GA, m\_IL$
GA	B, v, d, b, m_GA	$a, m\_CA, m\_NY, m\_TX, m\_IL, m\_CO$
IL	B, v, d, b, a	$m\_IL, \ m\_NY, m\_CA, m\_CO, m\_GA, m\_TX$
NAT	B, v, d, b, a	$m\_NY, m\_CA, m\_TX, m\_IL, m\_GA, m\_CO$

Conclusion: Estimates of Disease Prevalence

- FluView Disclaimer: Data is from the U.S. World Health Organization (WHO) Collaborating Laboratories and the National Respiratory and Enteric Virus Surveillance System (NREVSS)

- The model does not take into account behavioral changes of individuals.

- To utilize our model, it would have to be adjusted in order to predict the disease burden in advance.

Conclusion: Timing of Peak Prevalence

- Transmission rates within a range of .4 to .55 lead to good agreement on timing with data from GA, NY, CO, IL, and NAT.
- Texas: There is an early peak followed by a period of low case reports and then a second peak.
- California: There were two peaks.
- The model may suggest that only a few cases arriving from abroad have the potential to spread through migration within the country.

Conclusion: General Utility of this Method

- Sensitivity analysis shows that small variations in migration do not have a noticeable effect on timing of peak prevalence.
- How well a patch SIR model will predict the timing of an epidemic is dependent on getting accurate estimates of transmission rates and other disease-related parameters.
- Our method has the advantage of using publicly available data and epidemiological parameters that may be estimated from the first few cases.

References

1. Novel Swine-Origin Inuenza A (H1N1) Virus Investigation Team, *N Engl J Med* 360 2605-2615 (2009)
2. C. Fraser et al, *Science* 324 1557 (2009)
3. R. M. Anderson and R. M. May, *Nature* 280, 361 (1979).
4. M. J. Keeling and L. Danon, Mathematical modelling of infectious diseases *British Medical Bulletin* 92 (2009)
5. V. Colizza, A. Barrat, M. Barthelemy, and A. Vespignani *PNAS* 103 7 2015-2020 (2006)
6. Wallace, Hu-Wang, Chen *Symmetries and Groups in Contemporary Physics (Nankai Series in Pure, Applied Mathematics and Theoretical Physics)*, World Scientific, in press (2013)
7. Center for Disease Control and Prevention (Atlanta 2009), [Data File]. Retrieved May 3, 2012 fromhttp://www.cdc.gov/nchs/births.htm

References

8. Center for Disease Control and Prevention (Atlanta 2009), [Data File]. Retrieved May 3, 2012 from http://www.cdc.gov/nchs/deaths.htm
9. Center for Disease Control and Prevention (Atlanta 2010). Retrieved May 3, 2012 from http://www.cdc.gov/tb/statistics/reports/2010/pdf/report2010.pdf
10. A. Gardner (ABC News, New York 2009). Retrieved May 3, 2012 from http://abcnews.go.com/Health/Healthday/story?id=8439860
11. S. Cauchemez, et al, *N Engl J Med* 361 2619-2627 (2009)
12. B. J. Coburn, B. G. Wagner and S. Blower, *BMC Medicine* 7 30 (2009)

References

13. Center for Disease Control and Prevention (Atlanta 2010) Retrieved May 3, 2012 from http://www.cdc.gov/h1n1flu/qa.htm
14. Airports Council International, North America (Montreal 2009), [Data File] Retrieved May 3, 2012 from http://aci-na.org/content/airport-traffic-reports
15. Center for Disease Control and Prevention (Atlanta, 2010). [Data Set]. Retrieved Oct. 3, 2012 from http://gis.cdc.gov/grasp/fluview/fluportal dashboard.html
16. Center for Disease Control and Prevention (Atlanta 2010).[Data Set]. Retrieved Oct. 3, 2012 from the Division of Tuberculosis Eliminations online ordering system at http://www.cdc.gov/tb/.

References

17. C. Reed, F. J. Angulo, D. L. Swerdlow, M. Lipsitch, M. I. Meltzer, D. Jernigan, et al, *Emerg Infect Dis* (2009) Retrieved Jul. 31, 2013 from http://wwwnc.cdc.gov/ eid/ article/15/12/09-1413.htm
18. Center for Disease Control and Prevention (Atlanta 2009). *MMWR* 58 15, 400-402. Retrieved Oct. 10, 2012 from http://www.cdc.gov/mmwr/preview/ mmwrhtml/mm5815a5.htm
19. S. Funk, Marcel Salathe and V. A. A. Jansen, *J. R. Soc. Interface* 7 12471256 (2010)