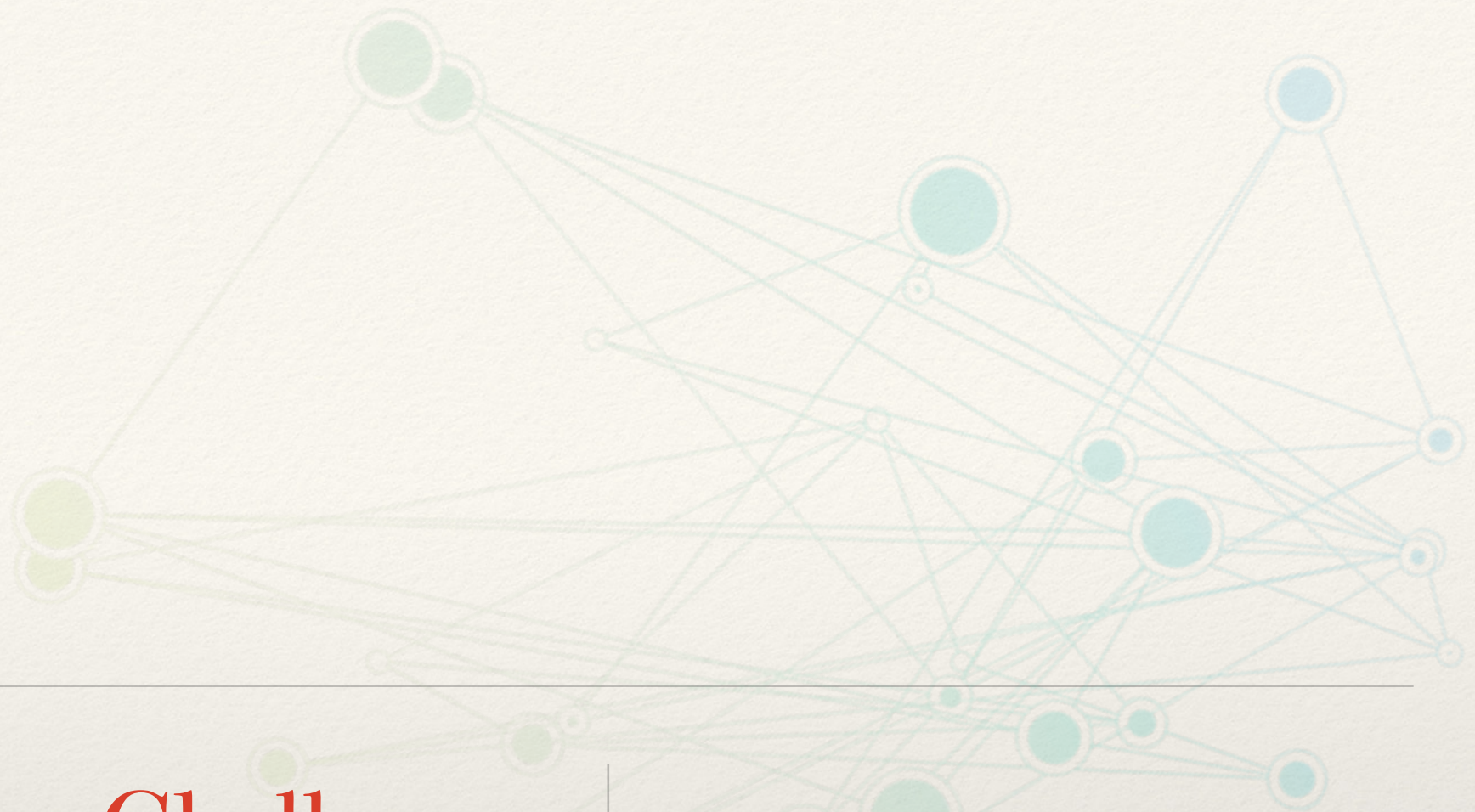




Ebola Forecasting Challenge

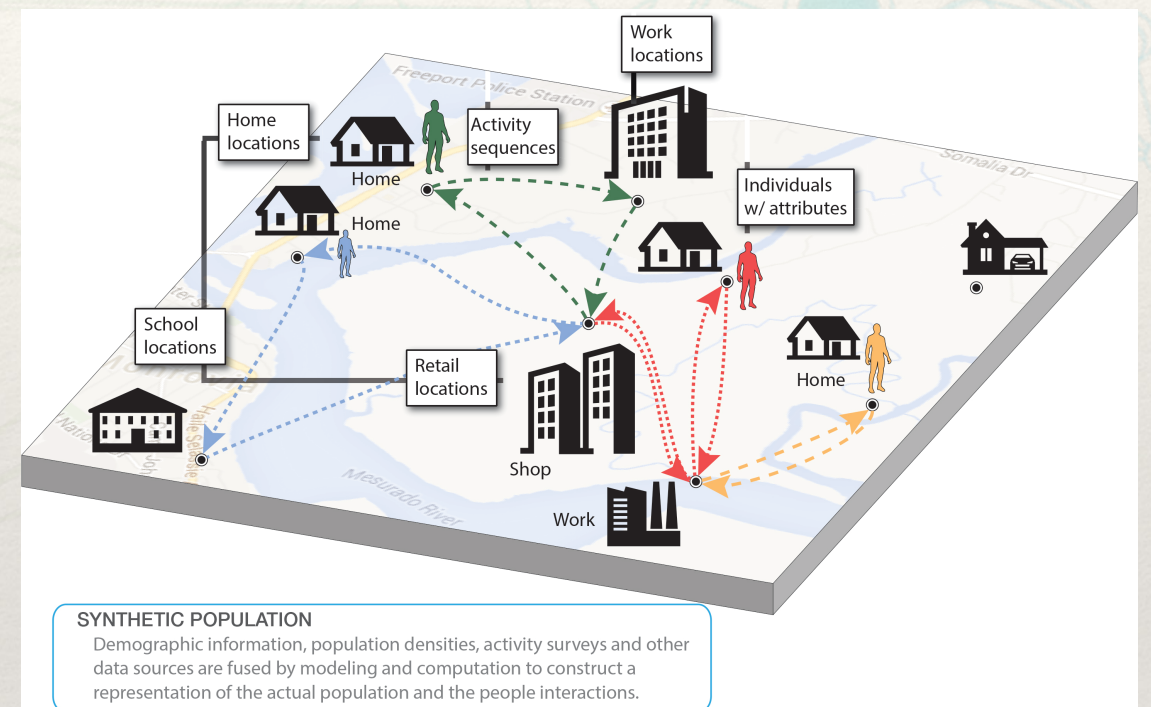
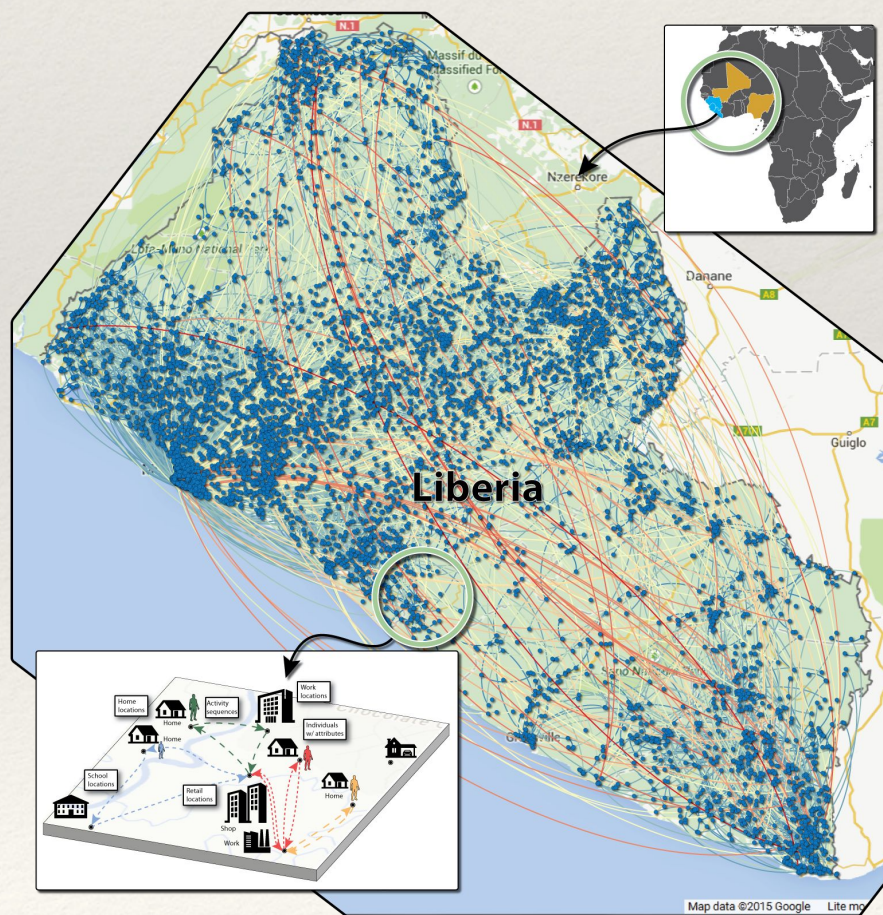
Team: Virginia Tech

- Synthetic population
- Model description
- Calibration methodology

Srini Venkatramanan et al.

Synthetic Population

- ❖ Who, where, when and what?
 - ❖ individuals with household structure
 - ❖ home and other activity locations
 - ❖ activity sequences

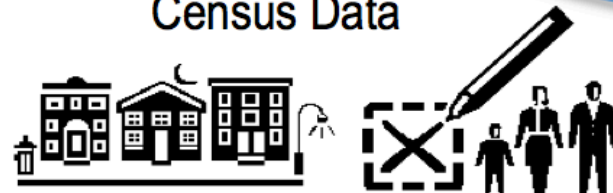


Liberia

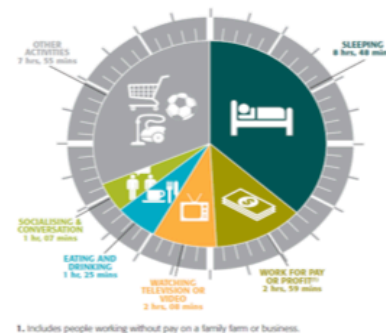
LandScan Population Counts



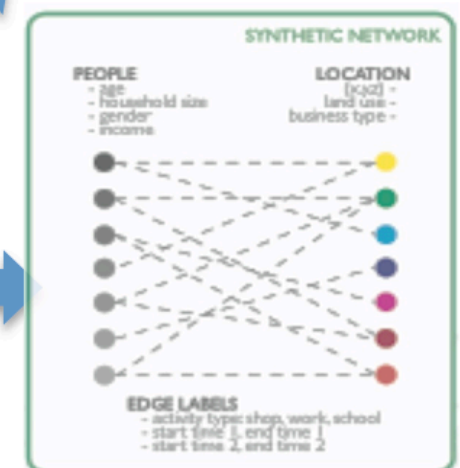
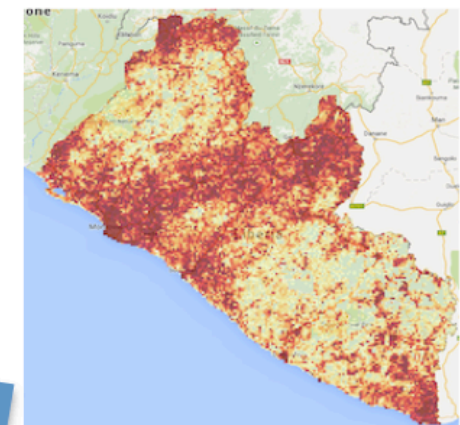
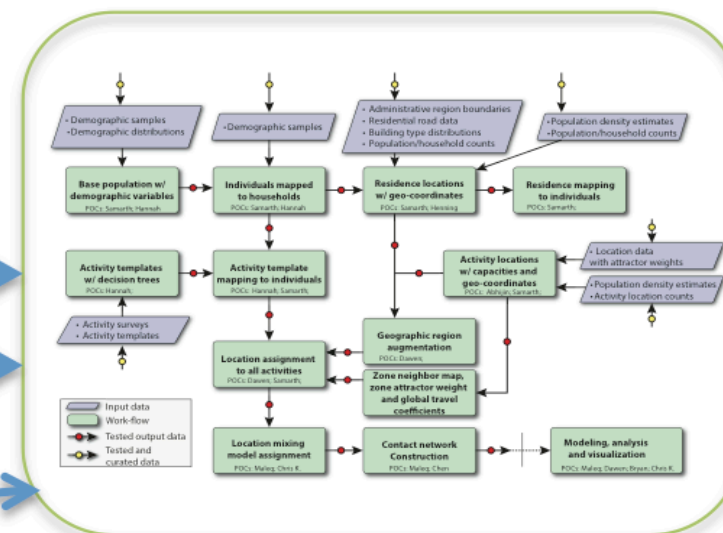
Census Data



Time Use Surveys



Network Synthesis



approx. 4M nodes, 80M edges*

Synthetic Population

* Data available at: <https://www.bi.vt.edu/ndssl/featured-projects/ebola/>

Simulation Environment

❖ *EpiFast:*

- ❖ Parallel algorithm and implementation for simulating standard SEIR model
- ❖ Scalable, computationally fast and allows for various dynamic interventions
 - ❖ Used for modeling hospitalized and funeral states, behavioral interventions
- ❖ Single stochastic replicate of Ebola-Liberia (400days) takes approximately 30 seconds on two 12-core 2.93 GHz Dual Processor Nodes (48GB Memory and 1 TB Internal Hard Drive).
- ❖ *Bisset et al., “EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems”, ACM Proceedings of the 23rd international conference on Supercomputing, pg. 430–439, 2009.*

Model Parameters

- ❖ *Initial infected*

- ❖ Source county from situation report per scenario
- ❖ Initial guess from the given epicurve data, **calibrated** in range [1,20] to correct for early reporting errors.

- ❖ *Infectious and Incubation Period Distribution*

- ❖ **Estimated** from patient database (sc-1 / test)
- ❖ When unavailable set to an earlier estimate from 2014 outbreak calibrations

- ❖ *Base transmissibility*

- ❖ probability of infection per minute of contact
- ❖ Initial guess 5.5E-5, **calibrated** in the range [3E-5, 8E-5]

- ❖ *Natural isolation*

- ❖ 70% reduction in transmissibility for non-household interactions (**fixed**)

Model Parameters

- ❖ *Funeral-induced transmissibility reduction*
 - ❖ Early time points set to zero
 - ❖ Once safe burials were introduced efficacy and compliance were **calibrated**
- ❖ *Hospitalization-induced transmissibility reduction*
 - ❖ Efficacy **calibrated** in the range $[0.1, 0.8]$
- ❖ *Travel reduction*
 - ❖ Long distance travel links were damped / removed to better mimic Merler et al.'s synthetic population (**fixed**)
- ❖ *Threshold Fraction*
 - ❖ Threshold on incidences per county before first behavioral intervention kicks in. Range $[5E-5, 1E-3]$. In addition to the threshold, efficacy of the intervention was also **calibrated**.
- ❖ *Action Delay*
 - ❖ Delay between the first and second behavioral intervention. Range $[45, 105]$. Initial guess 75. Efficacy of the intervention was also **calibrated**.

Model Calibration

- ❖ Parameter estimation:

*Given the ground truth (epidemic curve),
estimate model parameter values that best explain the data.*

- ❖ **Challenges:**

- ❖ Ground truth: single instance from a stochastic system; often noisy
- ❖ Model: High dimensional, highly non-linear, non-convex response surface

- ❖ **Our approaches:**

- ❖ **Optimization framework**

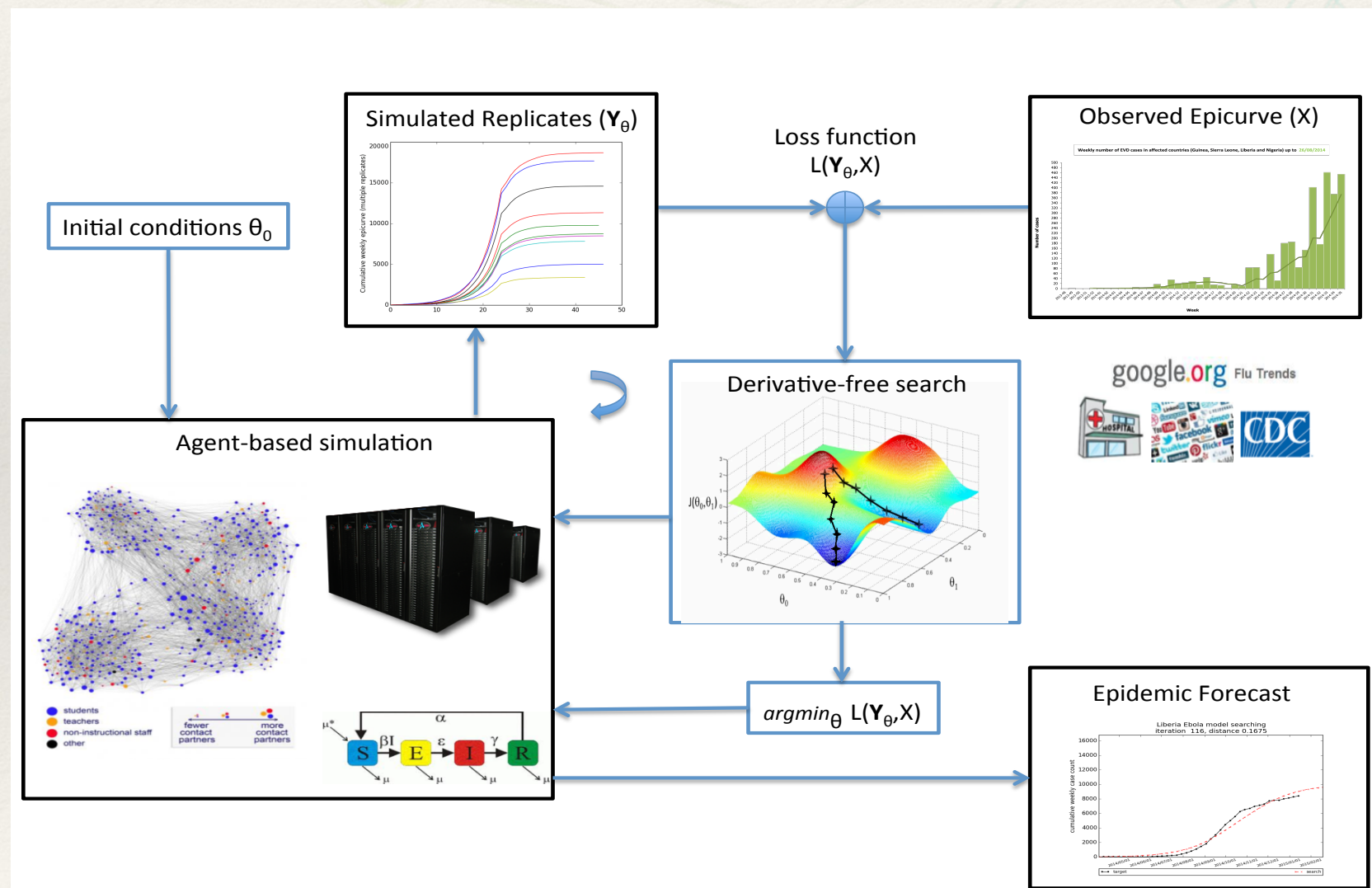
- ❖ Fewer simulation runs, quicker to estimate
 - ❖ Results are much easier to interpret for policy planners
 - ❖ **May converge to local minima**
 - ❖ **Difficult to capture parameter uncertainty**

- ❖ **Bayesian framework**

TP	Calibration method
1	Opt.
2	Bayes 1-phase
3	Bayes 2-phase
4	Opt.
5	Bayes 1-phase

Optimization Framework

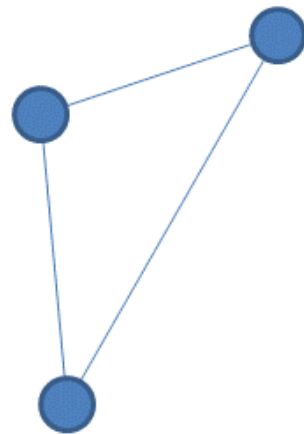
- ❖ *Search method*: Nelder-Mead simplex method (derivative-free)
- ❖ *Objective function*: average weighted-L1 distance between the simulated replicates and observed incremental epicurve $\min_{\theta} \frac{1}{R} \sum_{r=1}^R (\sum_{i=1}^N \alpha^{N-i} |Y_{\theta,r}(i) - X(i)|)$



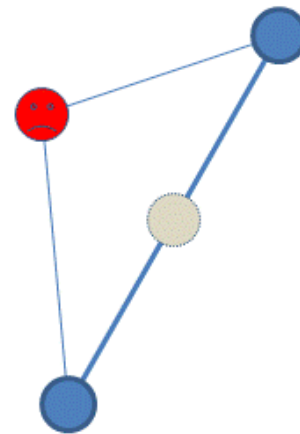
Nelder, John A.; R. Mead (1965). "A simplex method for function minimization". Computer Journal 7: 308–313.

Nelder-Mead Algorithm

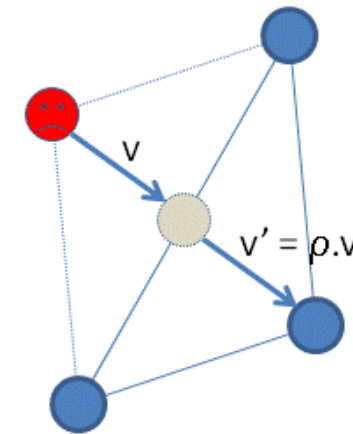
Main steps of the simplex algorithm



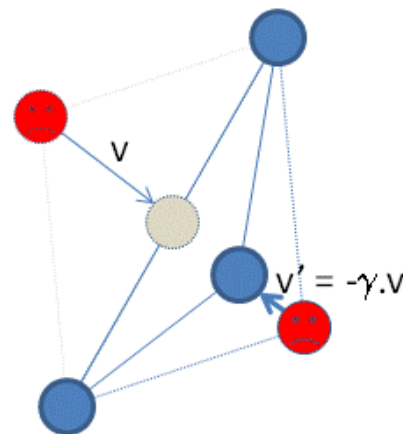
1: Initial Simplex



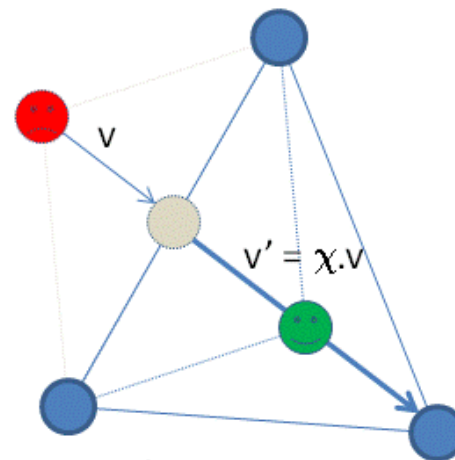
2: Center of gravity
(without the worst point)



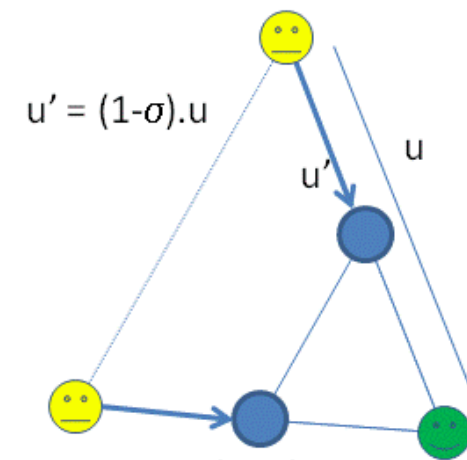
3: Reflection
(Default: $\rho = 1.00$)



4a: Contraction
(Default: $\gamma = 0.50$)



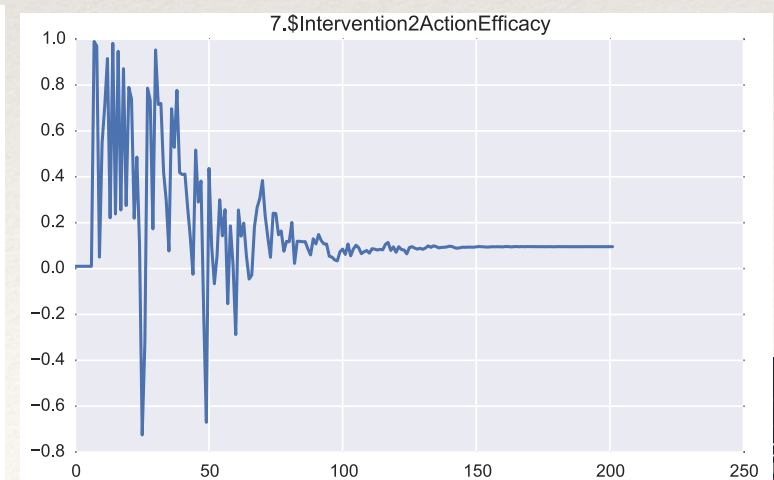
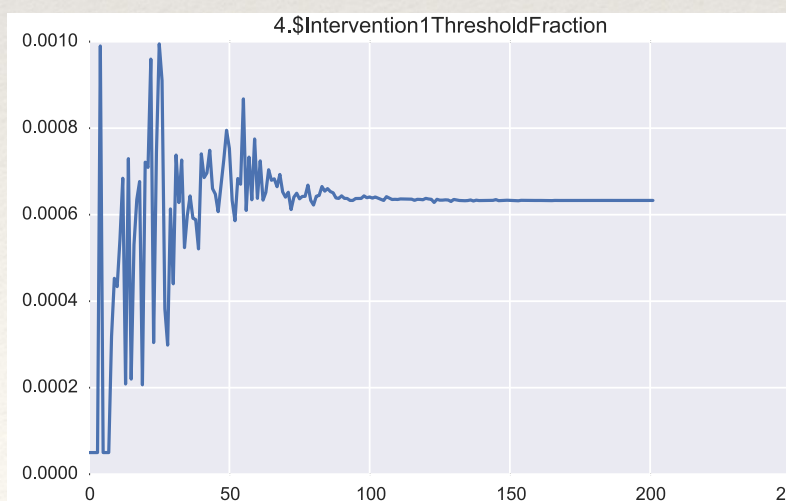
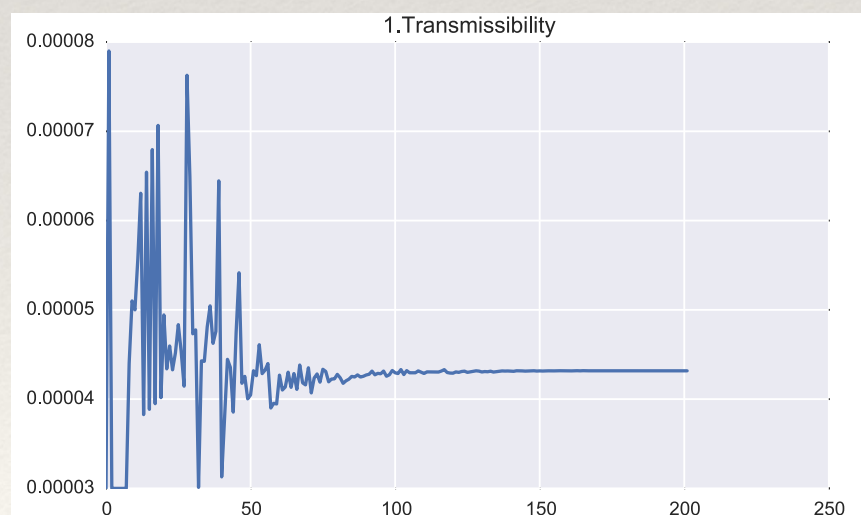
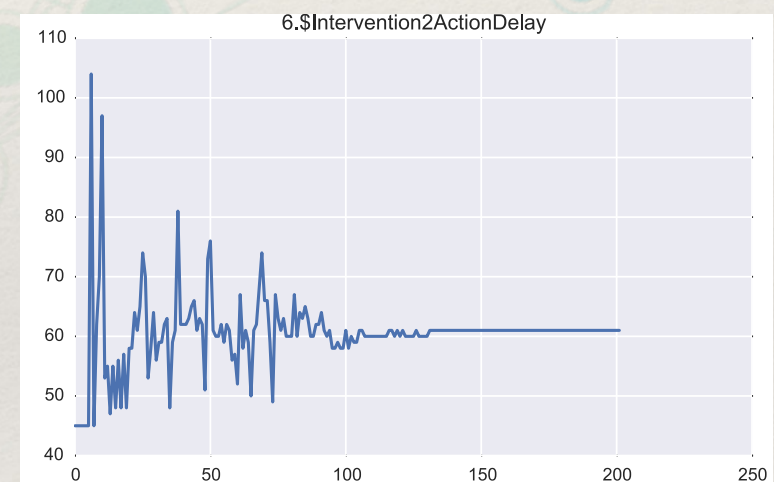
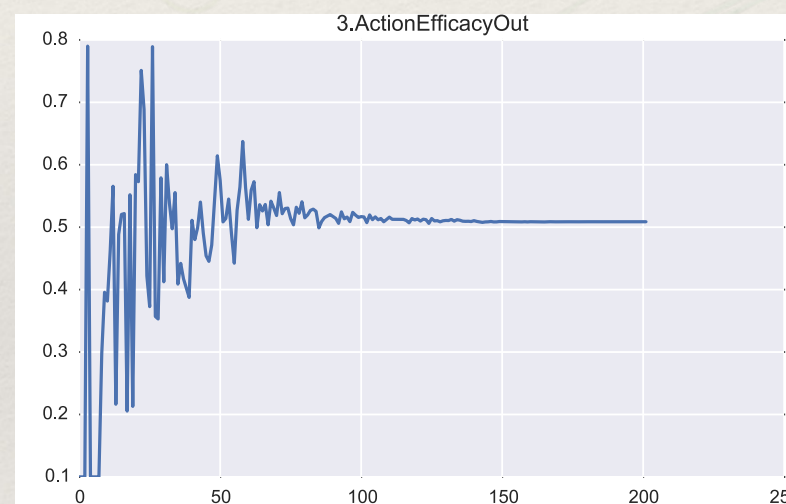
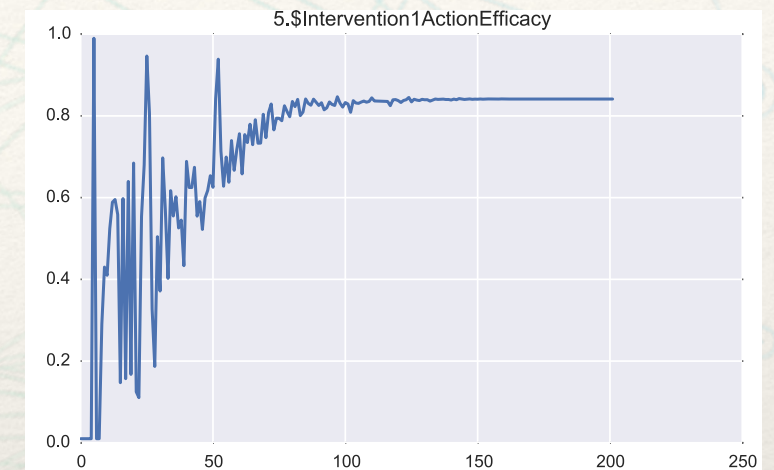
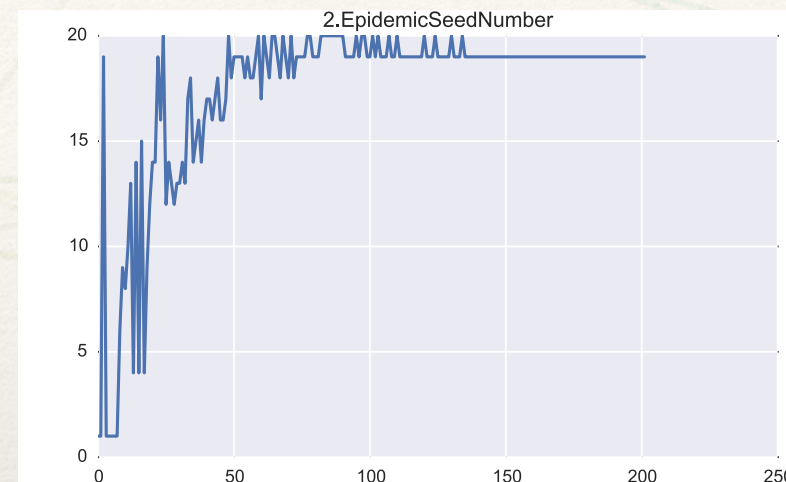
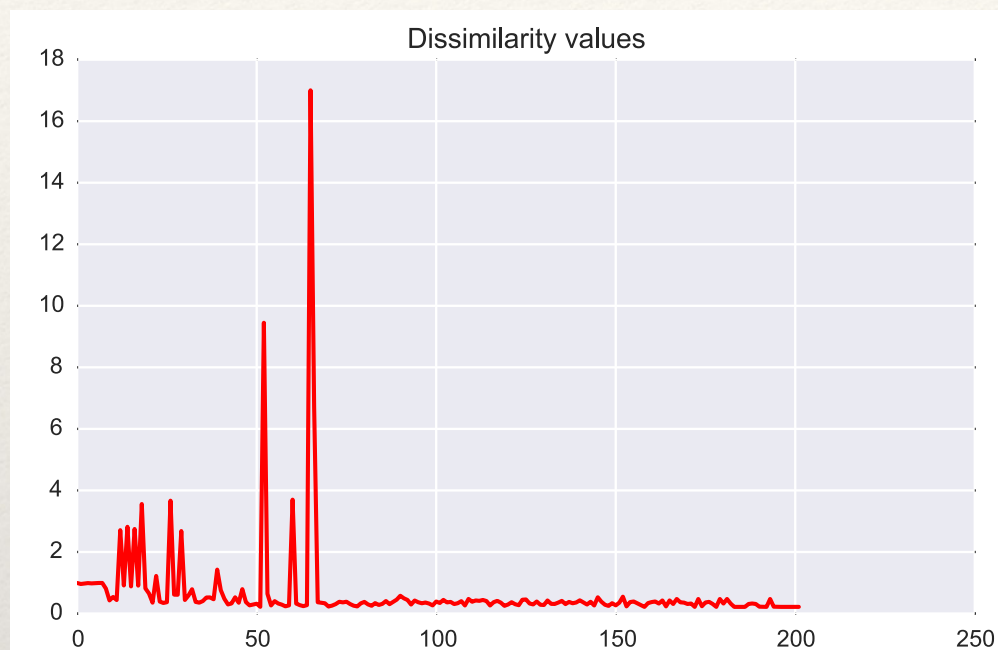
4b: Expansion
(Default: $\chi = 2.00$)



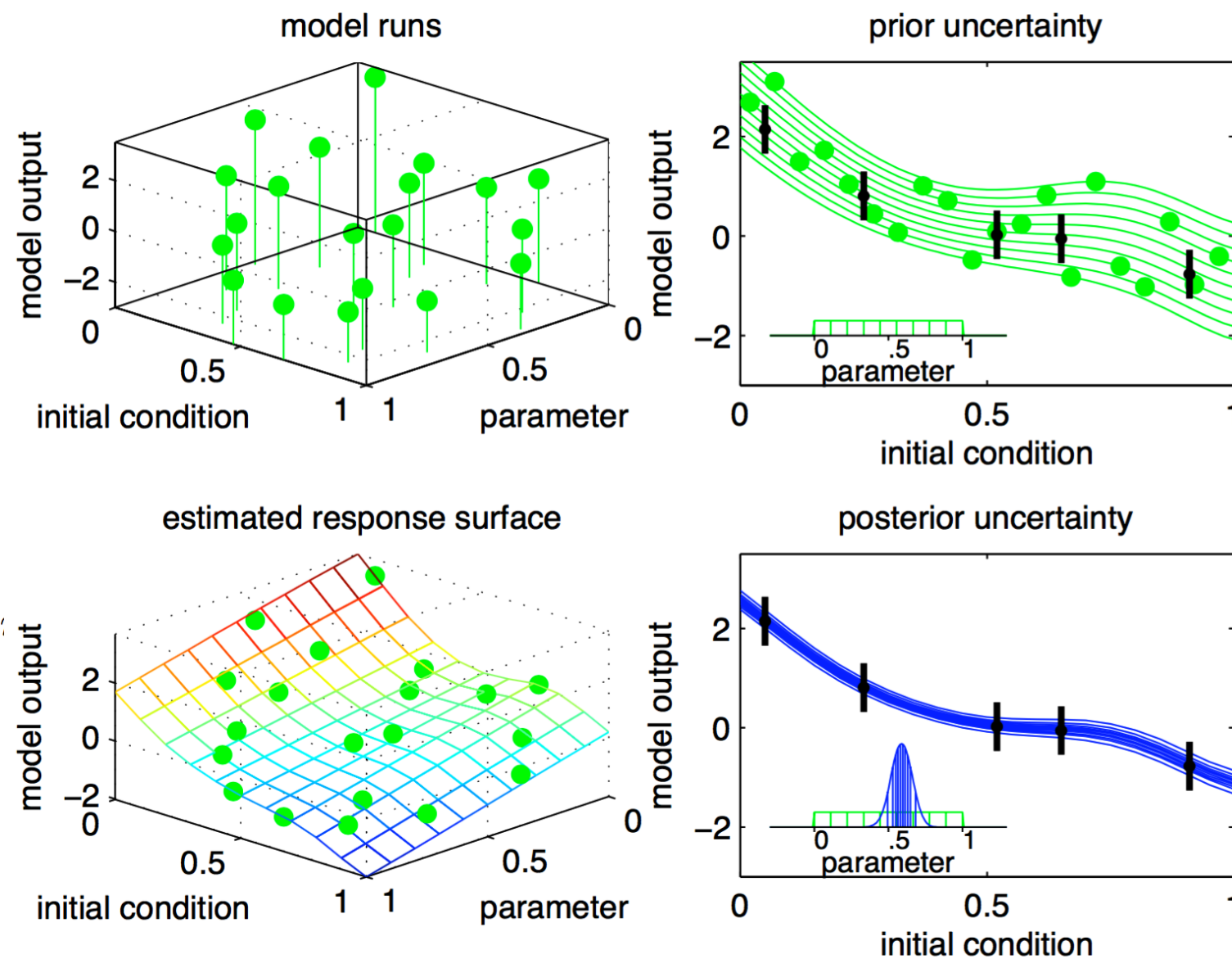
5: Shrinkage
(Default: $\sigma = 0.50$)

Nelder, John A.; R. Mead (1965). "A simplex method for function minimization". Computer Journal 7: 308–313.

Parameter Evolution

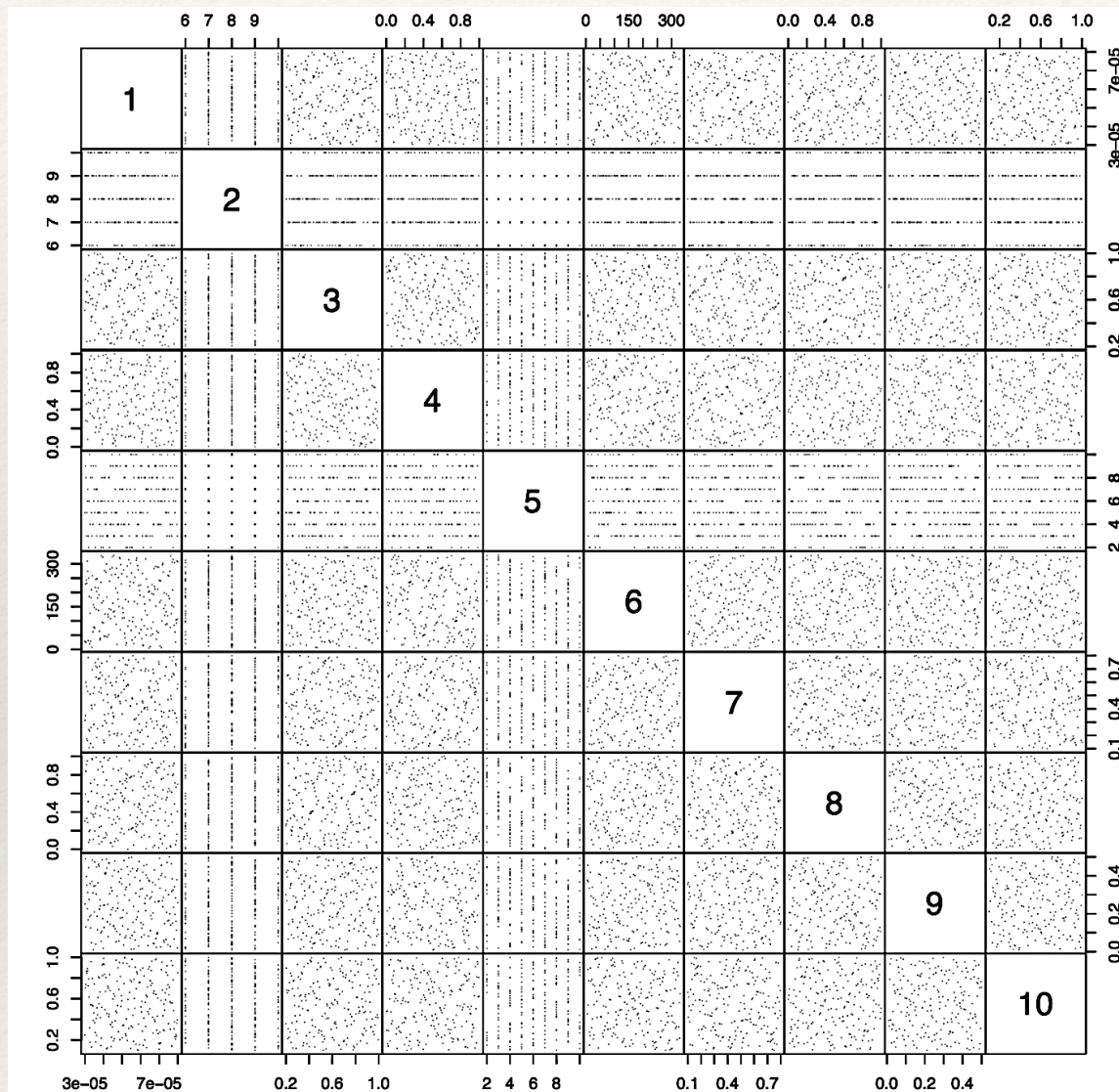


Bayesian Framework

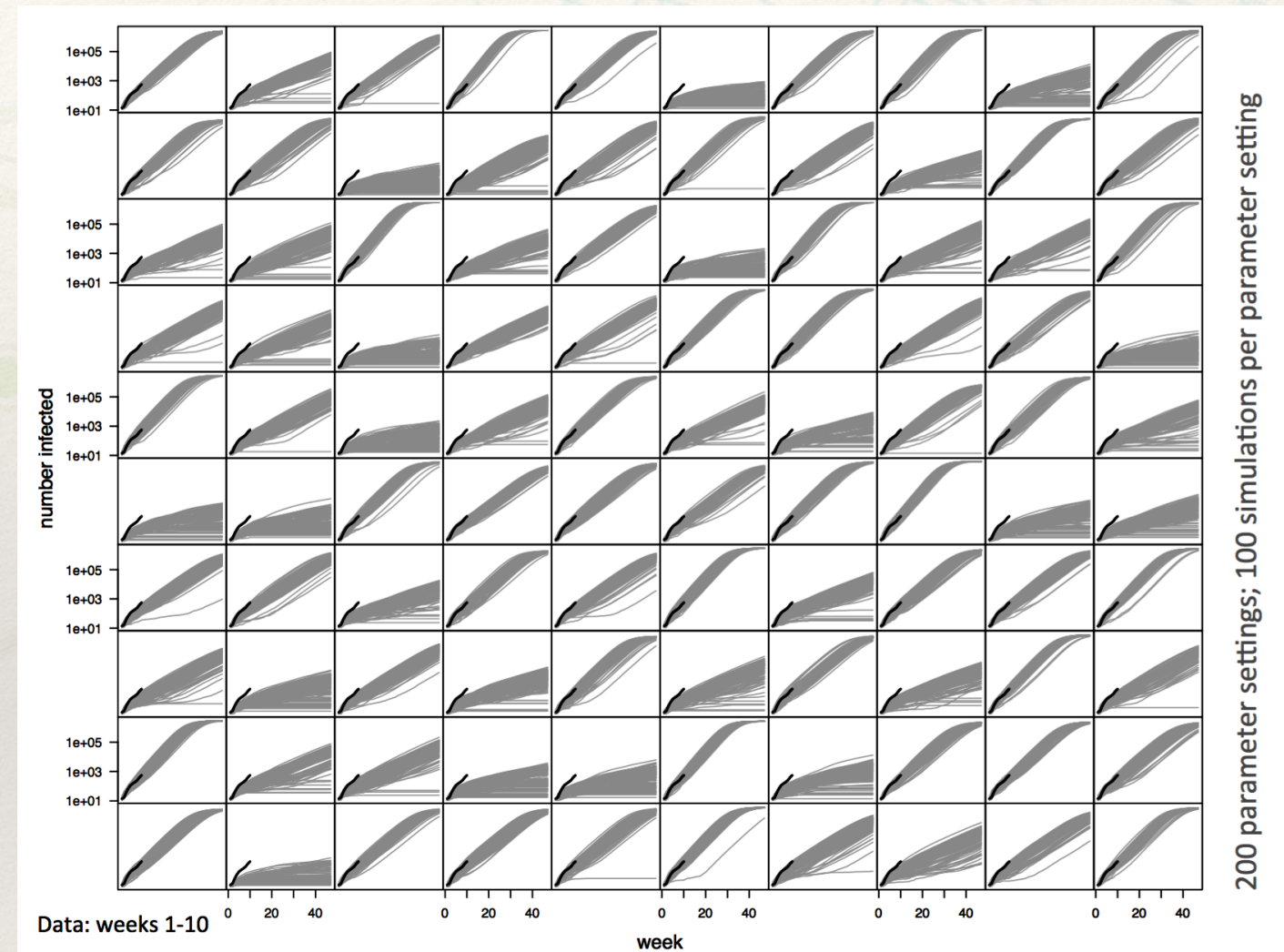


Higdon, Dave, et al. "Computer model calibration using high-dimensional output." Journal of the American Statistical Association (2012).

Digital Library



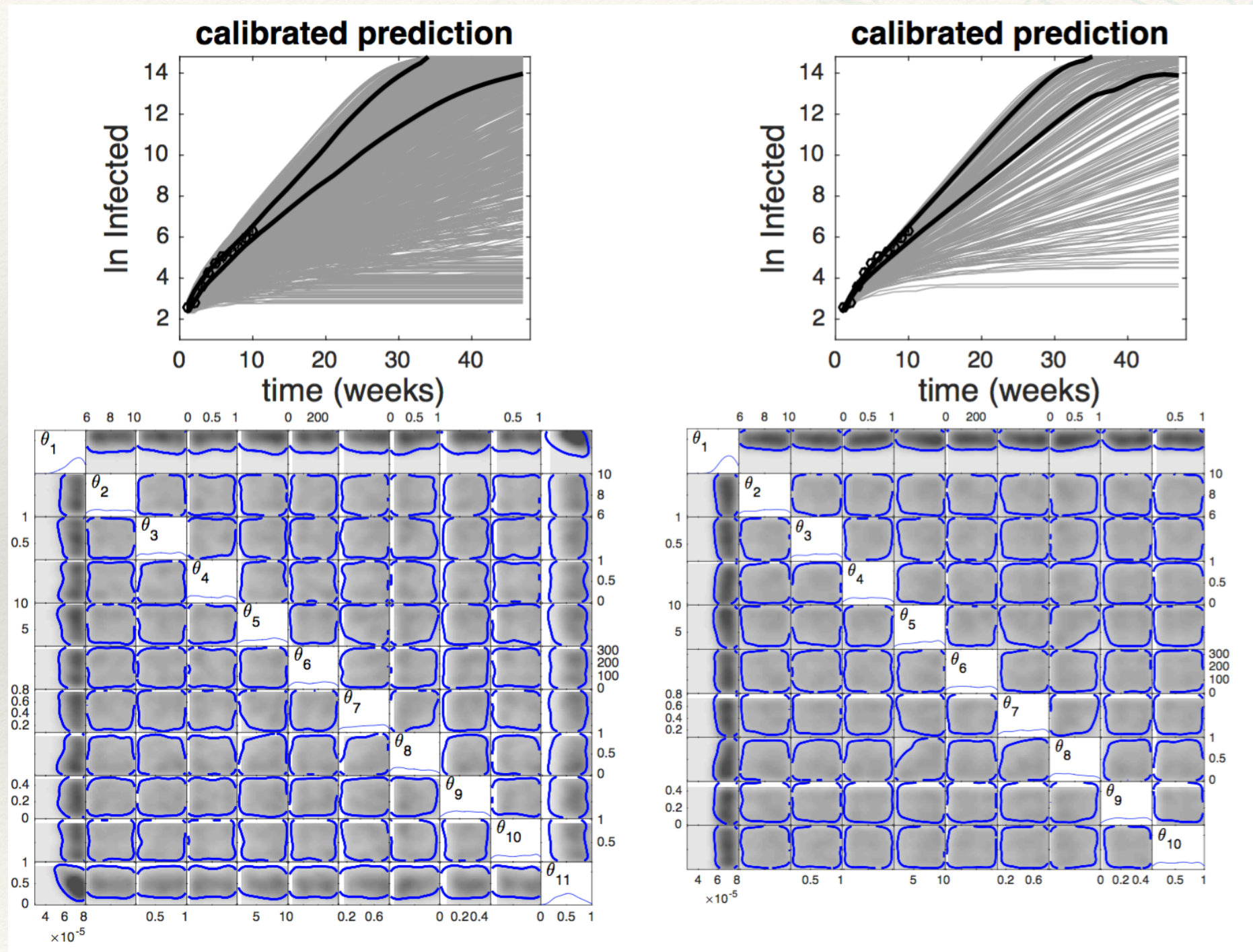
2-D space filling LHS design



Digital Library of epicurves

Posterior Distribution

Accounting for randomness



Using median simulation

Challenges Ahead

- ❖ *Forecasting a moving target*
 - ❖ Time varying disease and intervention parameters
 - ❖ Impact of information feedback into behavior
- ❖ *Sensitivity analysis and Uncertainty quantification*
 - ❖ Sensitivity to noise in data, model construction
 - ❖ Model reliability versus data utility
 - ❖ Model uncertainty, parameter uncertainty, inherent stochasticity
- ❖ *Multi-scale calibration*
 - ❖ Calibrating to multiple ground truth at various spatio-temporal scales
- ❖ *Fundamental limits to forecasting?*
 - ❖ How far ahead can we forecast reliably?
 - ❖ Computational complexity?

Acknowledgments

Sponsors



Ebola Challenge Team



Srimi Venkat
Computer Science



Bryan Lewis
Public Health



Jiangzhuo Chen
Software Systems



Dave Higdon
Statistics



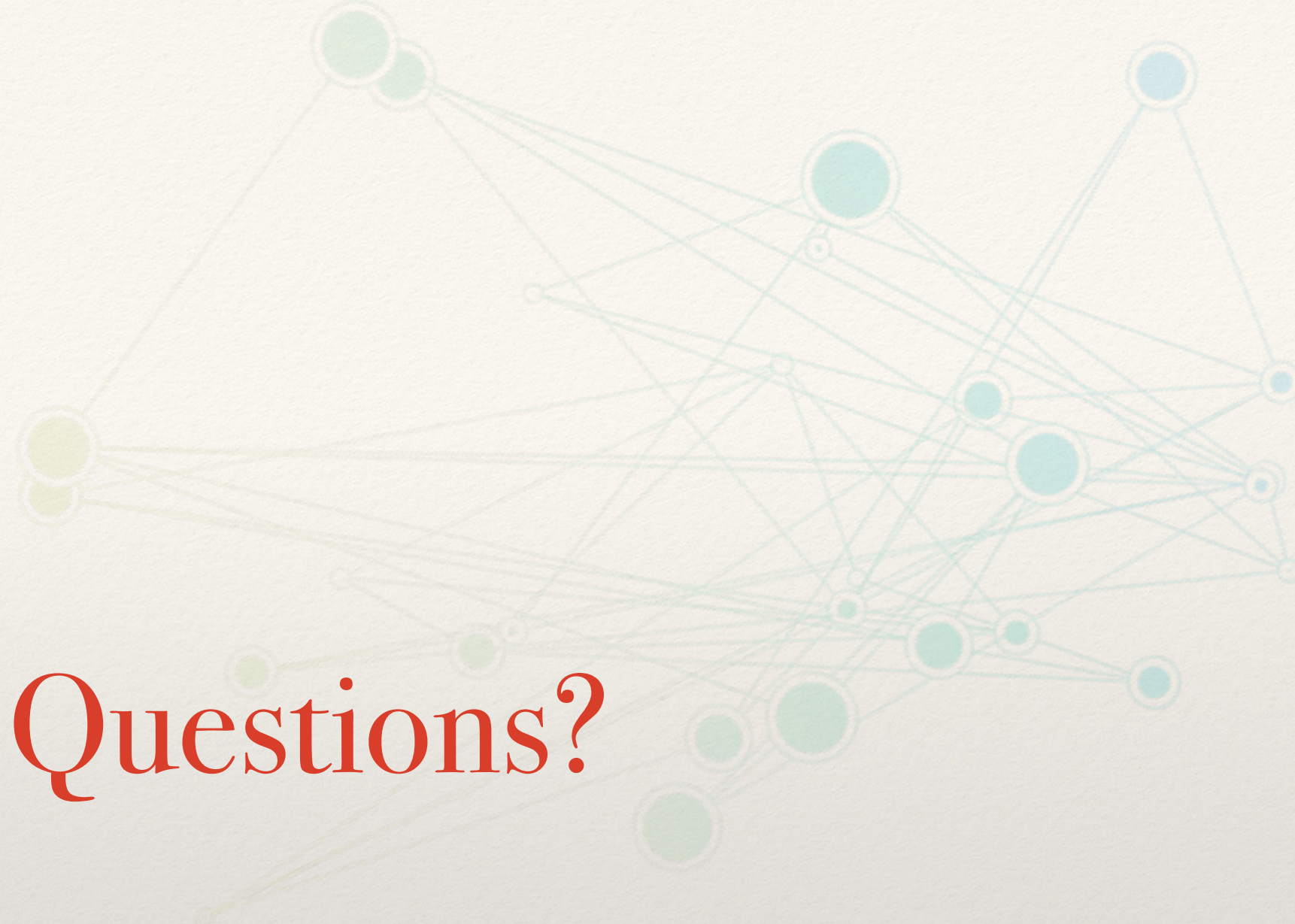
Anil Vullikanti
Computer Science



Madhav Marathe
Computer Science

NDSSL Staff



A network diagram with nodes and edges in the background. The nodes are represented by circles of varying sizes and colors (green, yellow, blue) and are interconnected by a web of thin lines, suggesting a complex network structure.

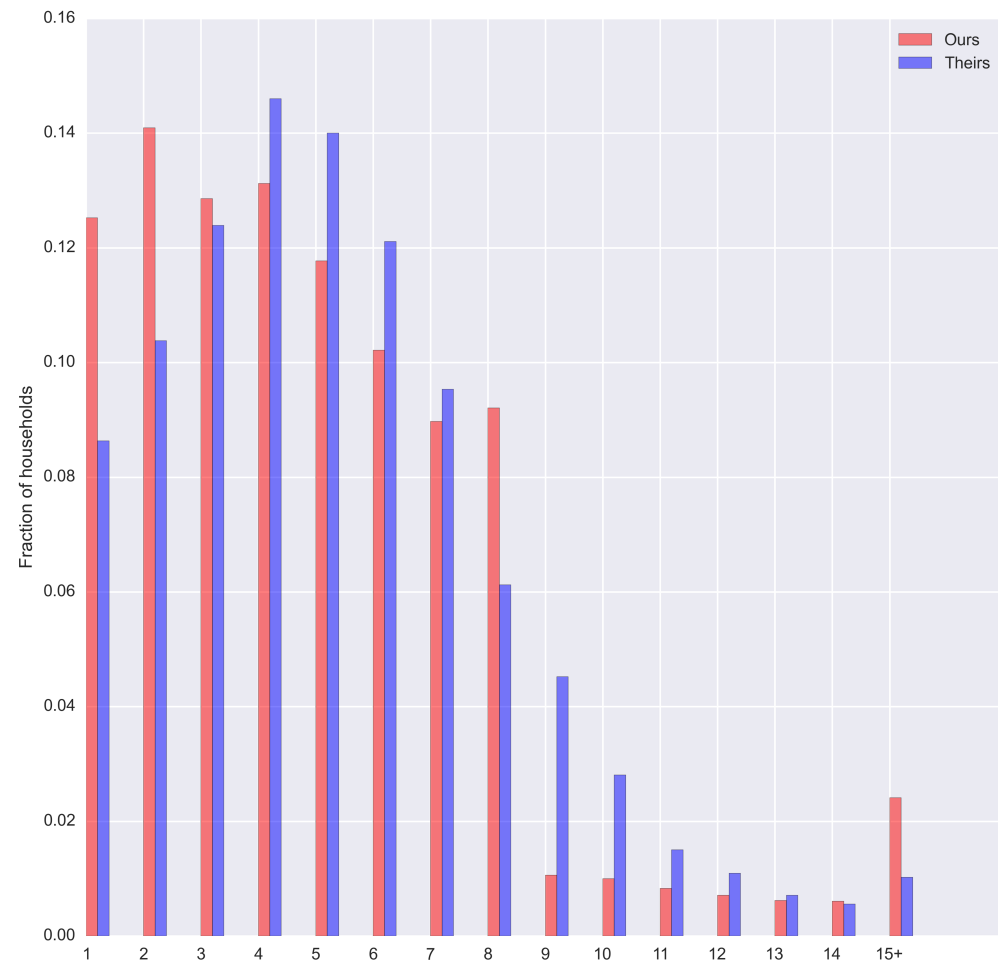
Questions?

vsriniv@vt.edu
bi.vt.edu/ndssl

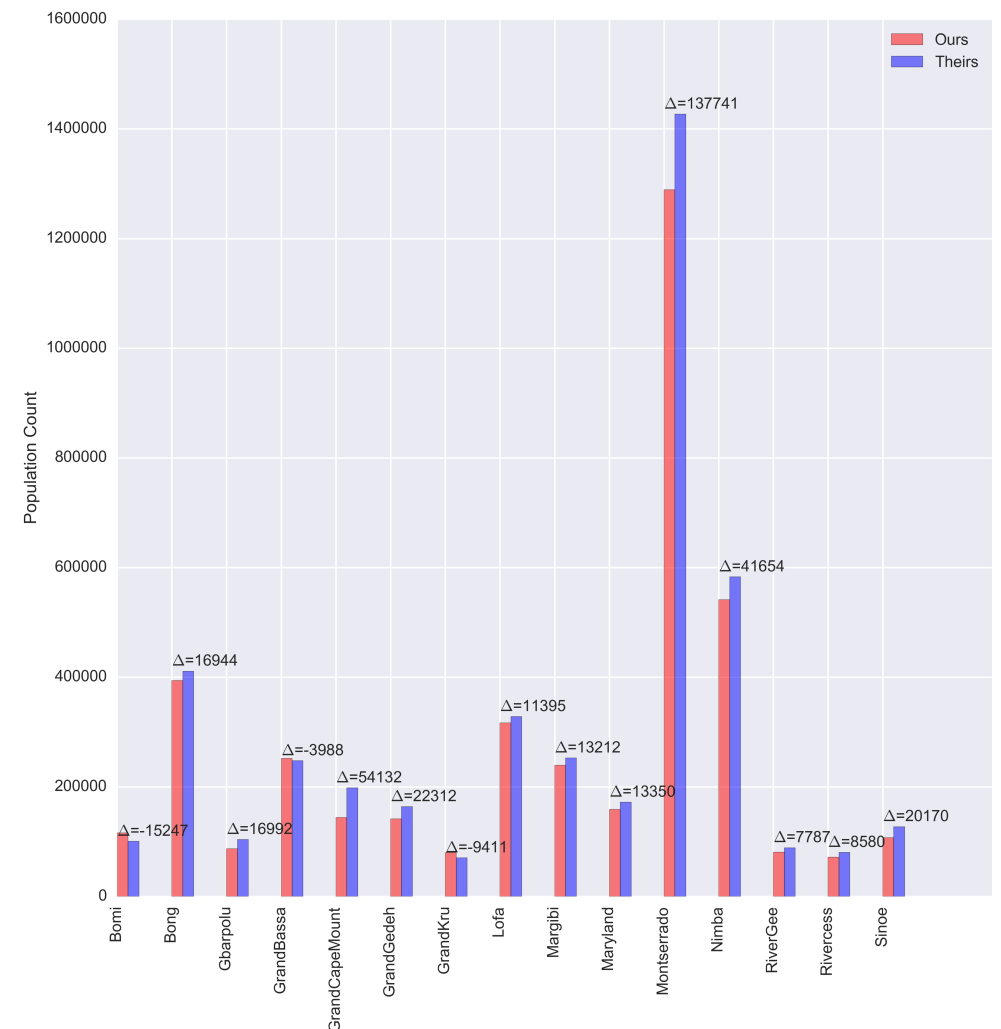


Backup slides

Demographics Comparison

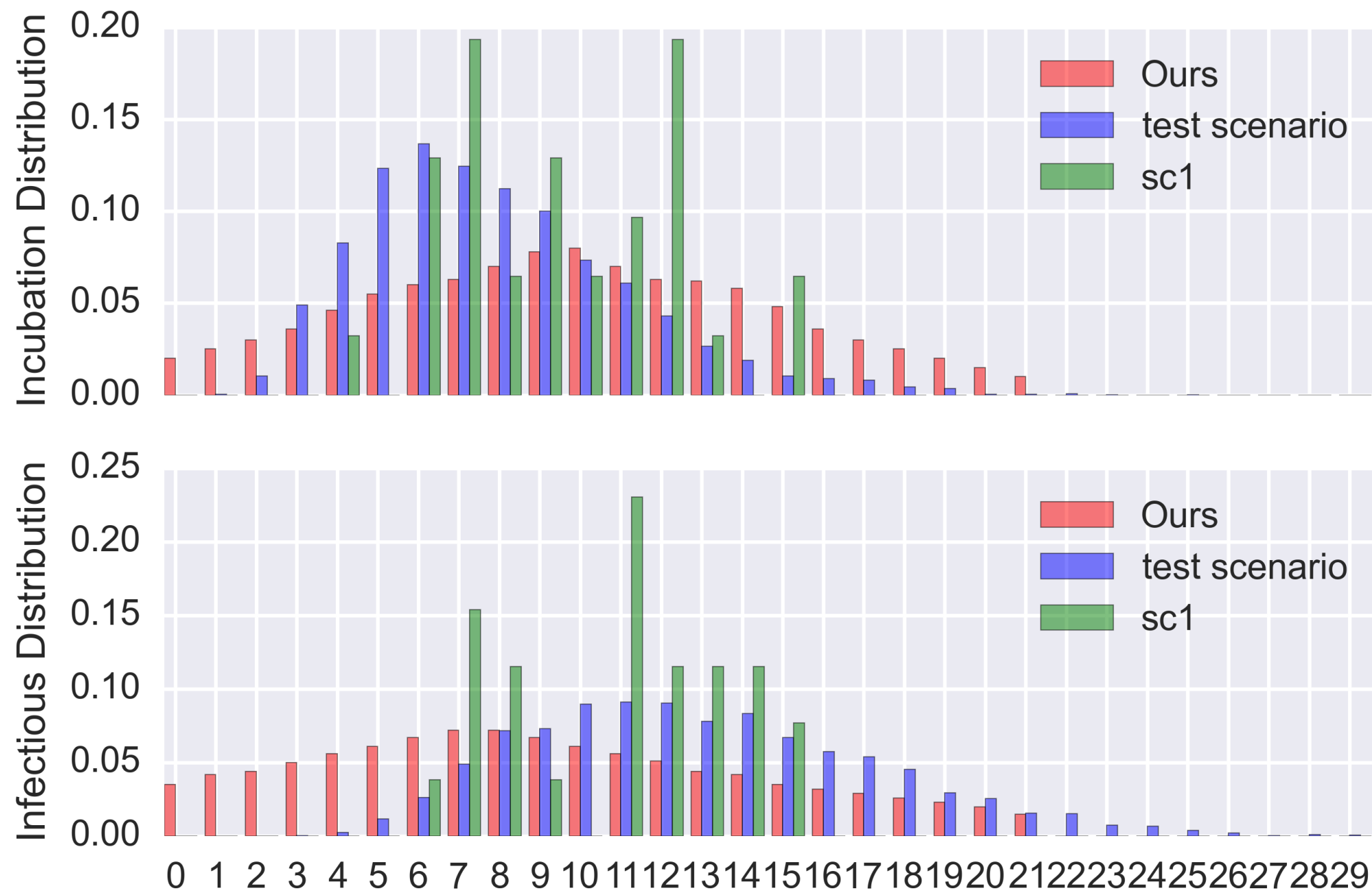


Household size distribution



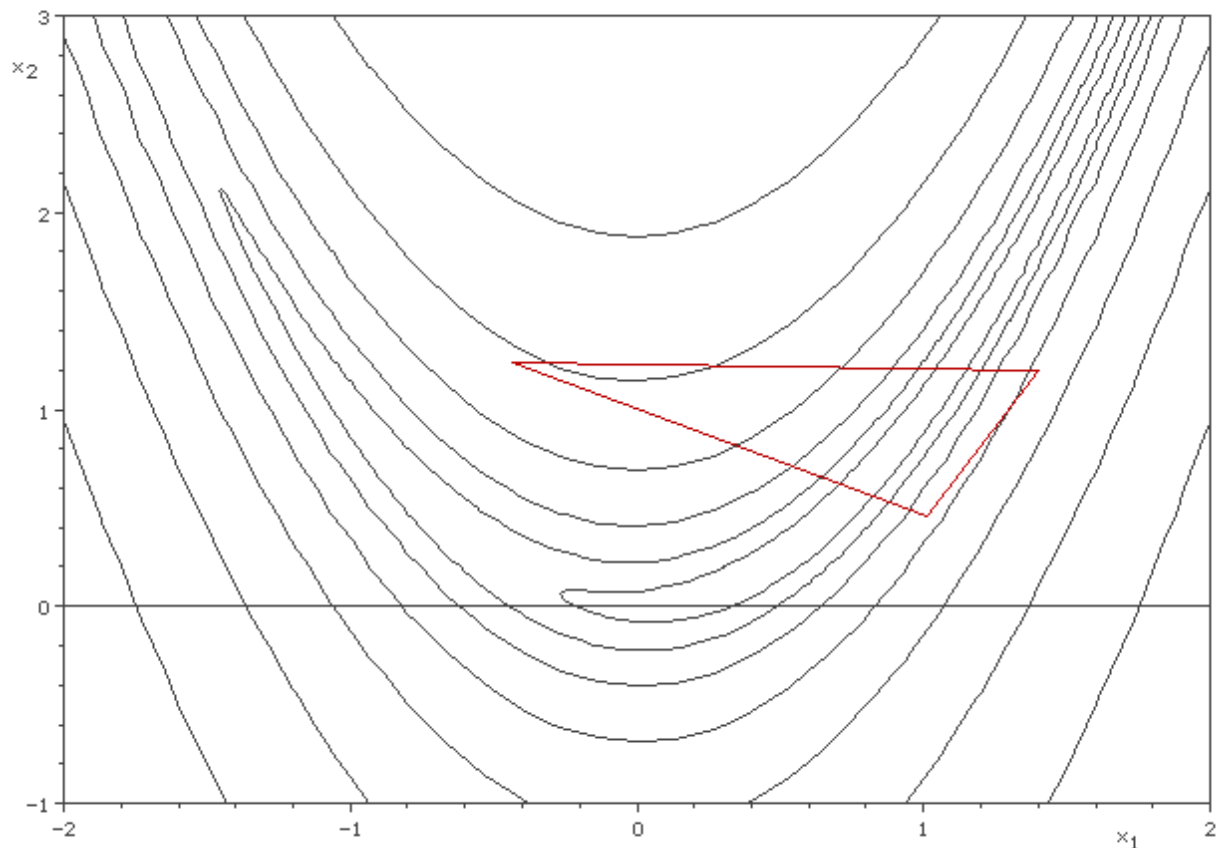
County populations

Distribution Comparison



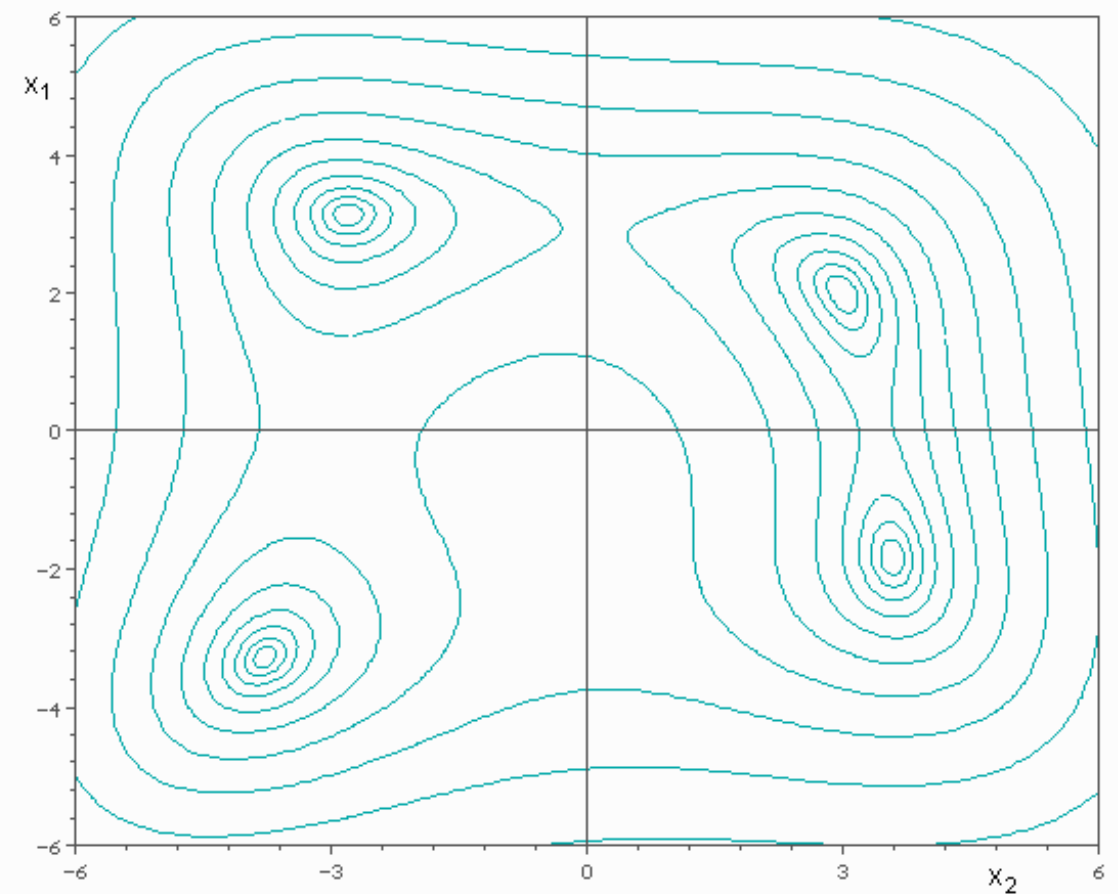
Nelder-Mead Animation

Nelder-Mead Simplex search over Banana Function



$$f(x, y) = (a - x)^2 + b(y - x^2)^2$$

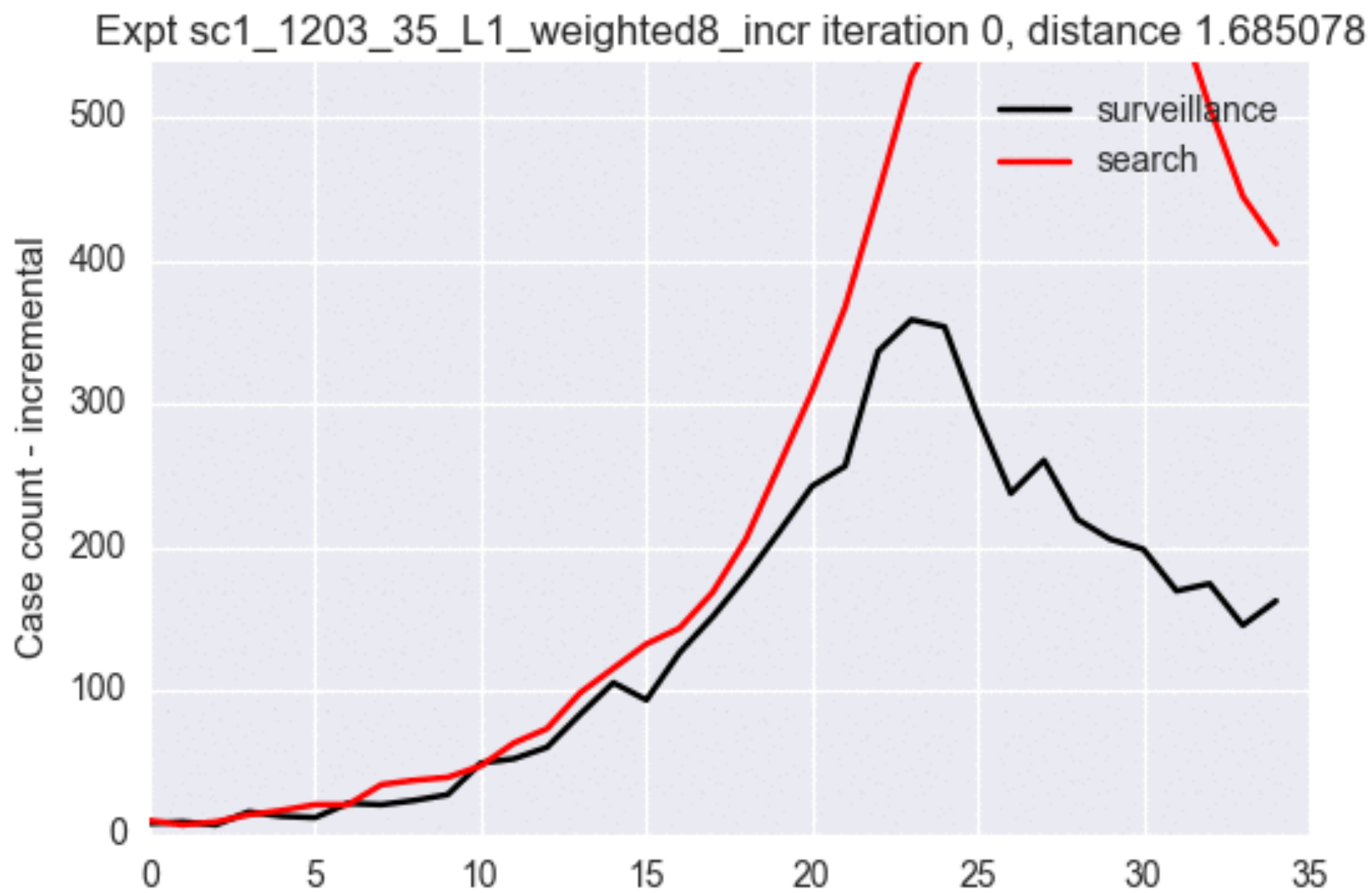
Nelder-Mead Simplex search over Himmelblau function



(c) P.A. Simionescu 2006

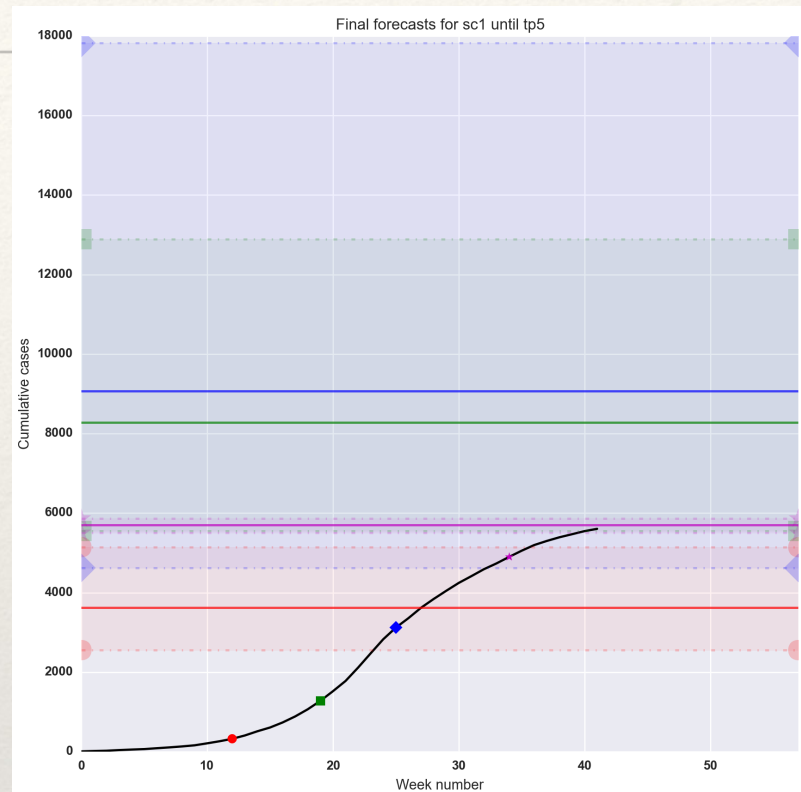
$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2.$$

Search Evolution

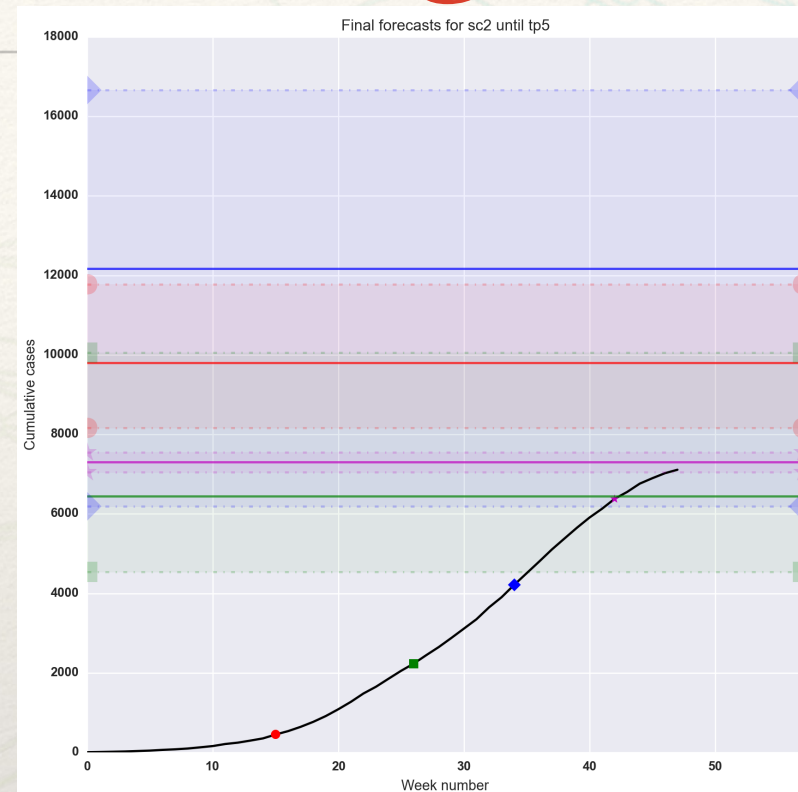


Our Forecasts - Long-term

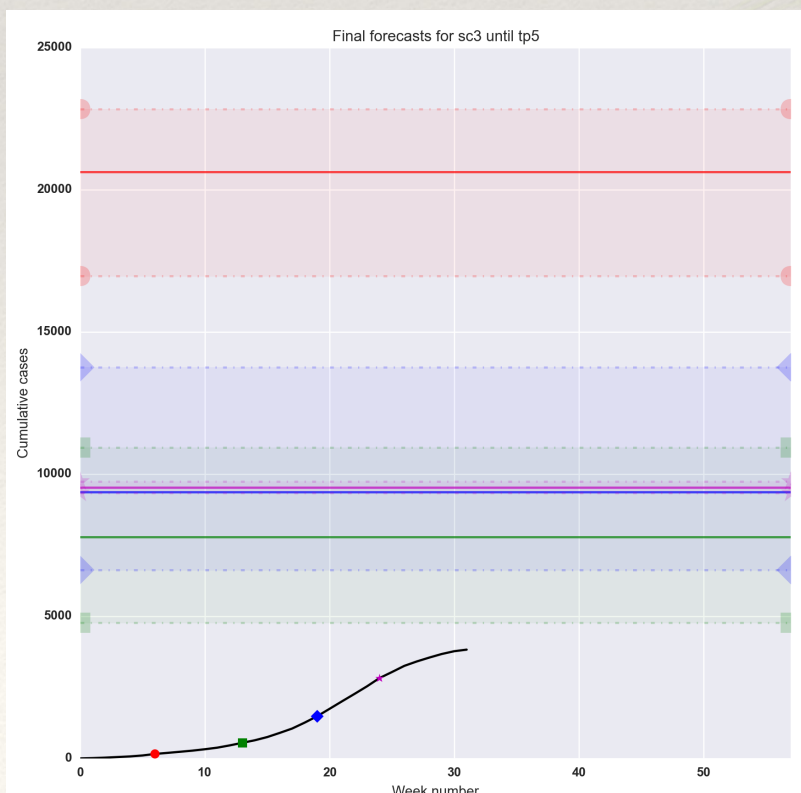
Sc. 1



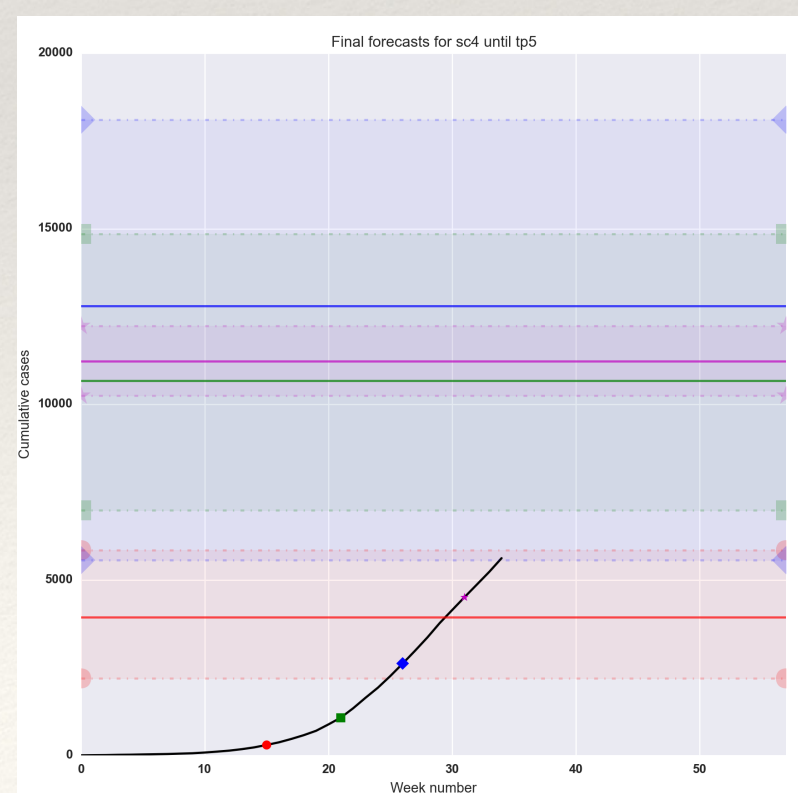
Sc. 2



Sc. 3

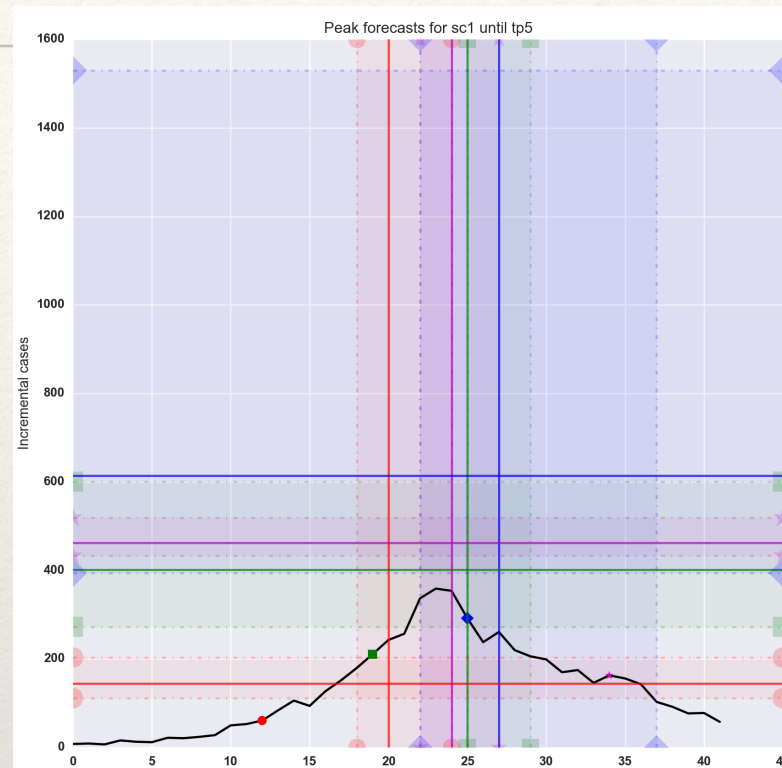


Sc. 4

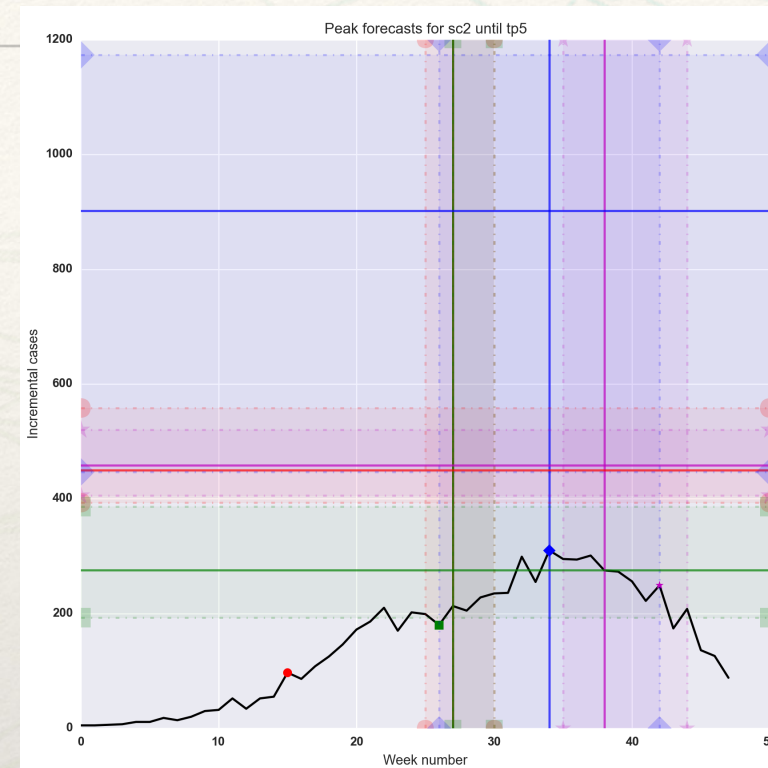


Our Forecasts - Peak

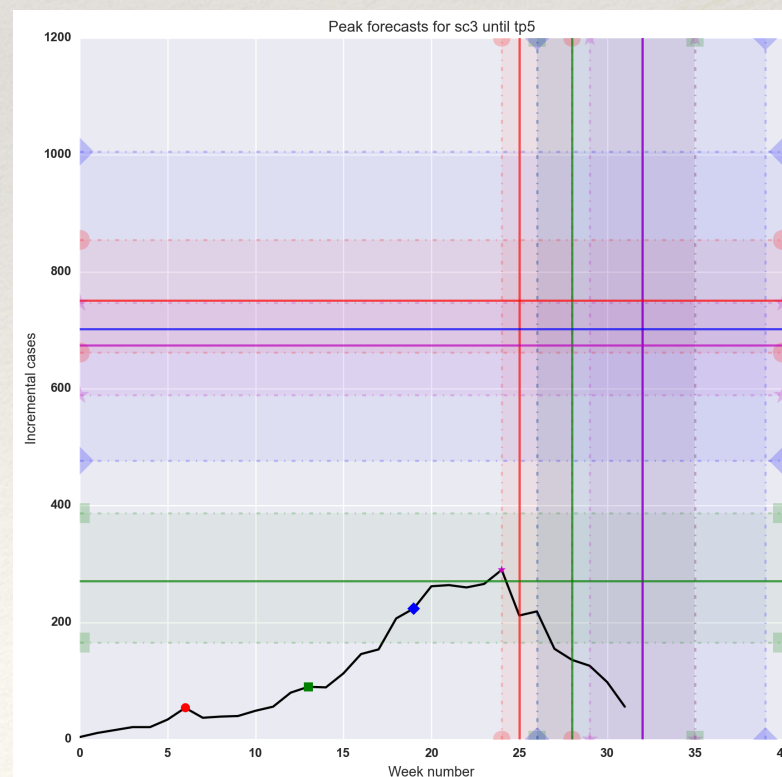
Sc. 1



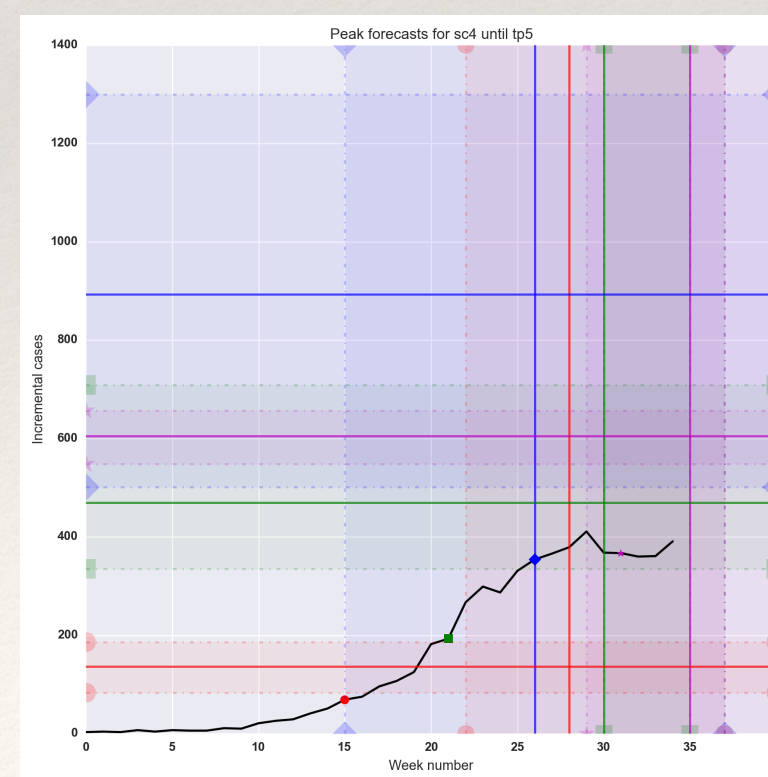
Sc. 2



Sc. 3

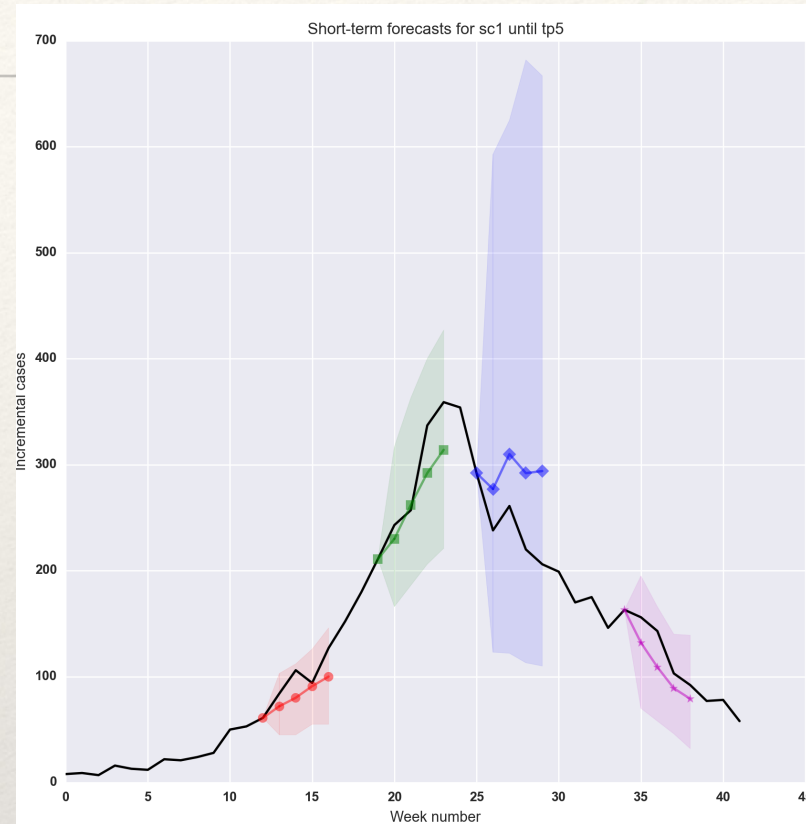


Sc. 4

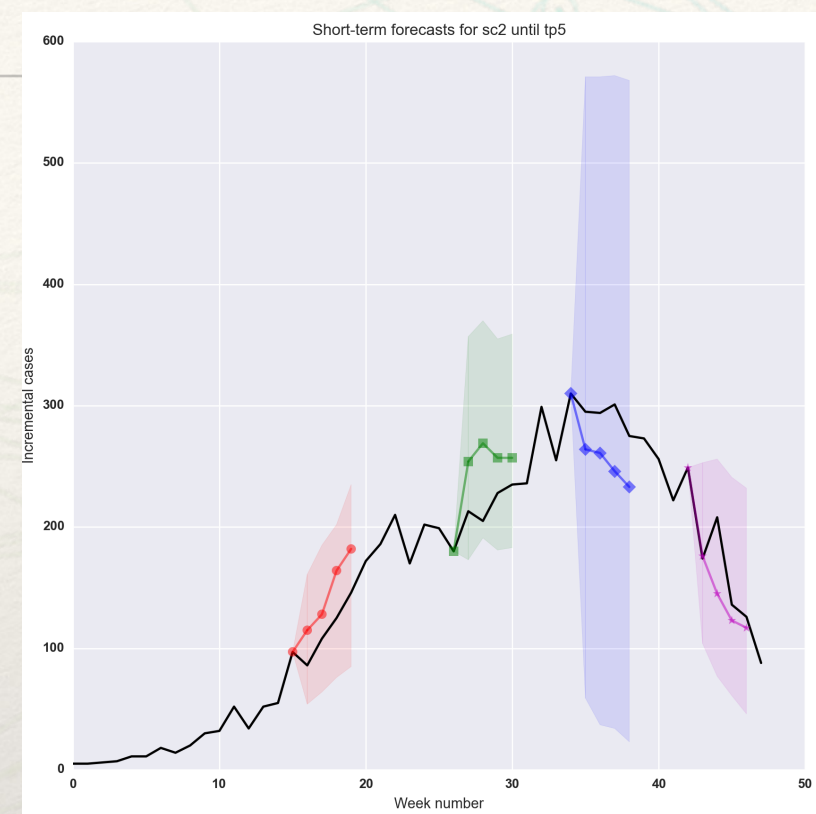


Our Forecasts - Short-term

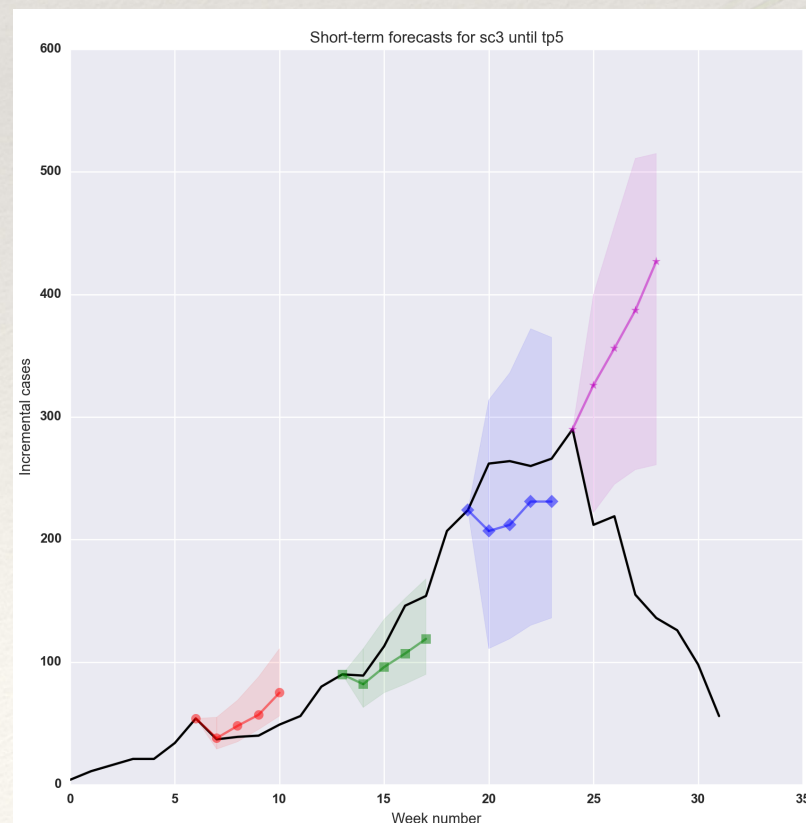
Sc. 1



Sc. 2



Sc. 3



Sc. 4

