Sensitivity Analysis & Sampling

Epid 814

Sensitivity analysis

- Uncertainty quantification: examines the variation in model outputs & behaviors
- **Sensitivity analysis**: examines which inputs/parameters drive that variation
    - If you change parameter p1, how much change in our output y (or other quantity of interest) do you see?

Goals

- Capture the frequency/distribution of different outputs/behaviors observed across parameter space as a function of the parameters

- Search for extremes/oddities, i.e. potentially uncommon behaviors that match a criteria (e.g. costly, interesting), illustrate the extreme range of behaviors

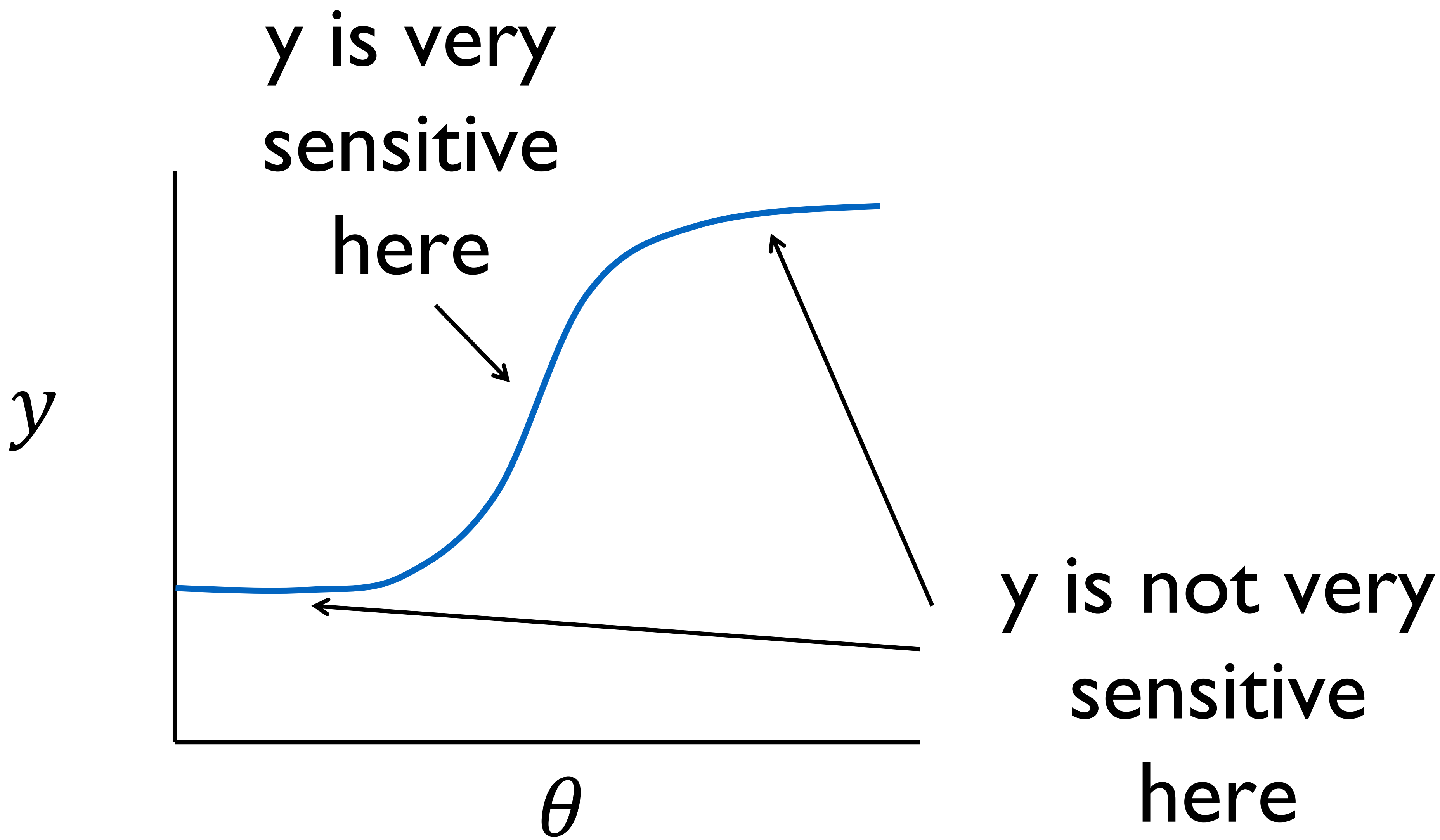
Goals

- Find sensitive/insensitive parameters or parameter combinations
  - Use these to decide what parameters to adjust/tune/intervene on
  - Reduce model complexity by fixing insensitive parameters

Basic setup

- Adjust a parameter or multiple parameters of interest
- Evaluate the model behavior/output
- Look for trends in how output changes as a function of parameters

Sensitivity is inherently a local attribute



Sensitivity/robustness/identifiability tradeoff

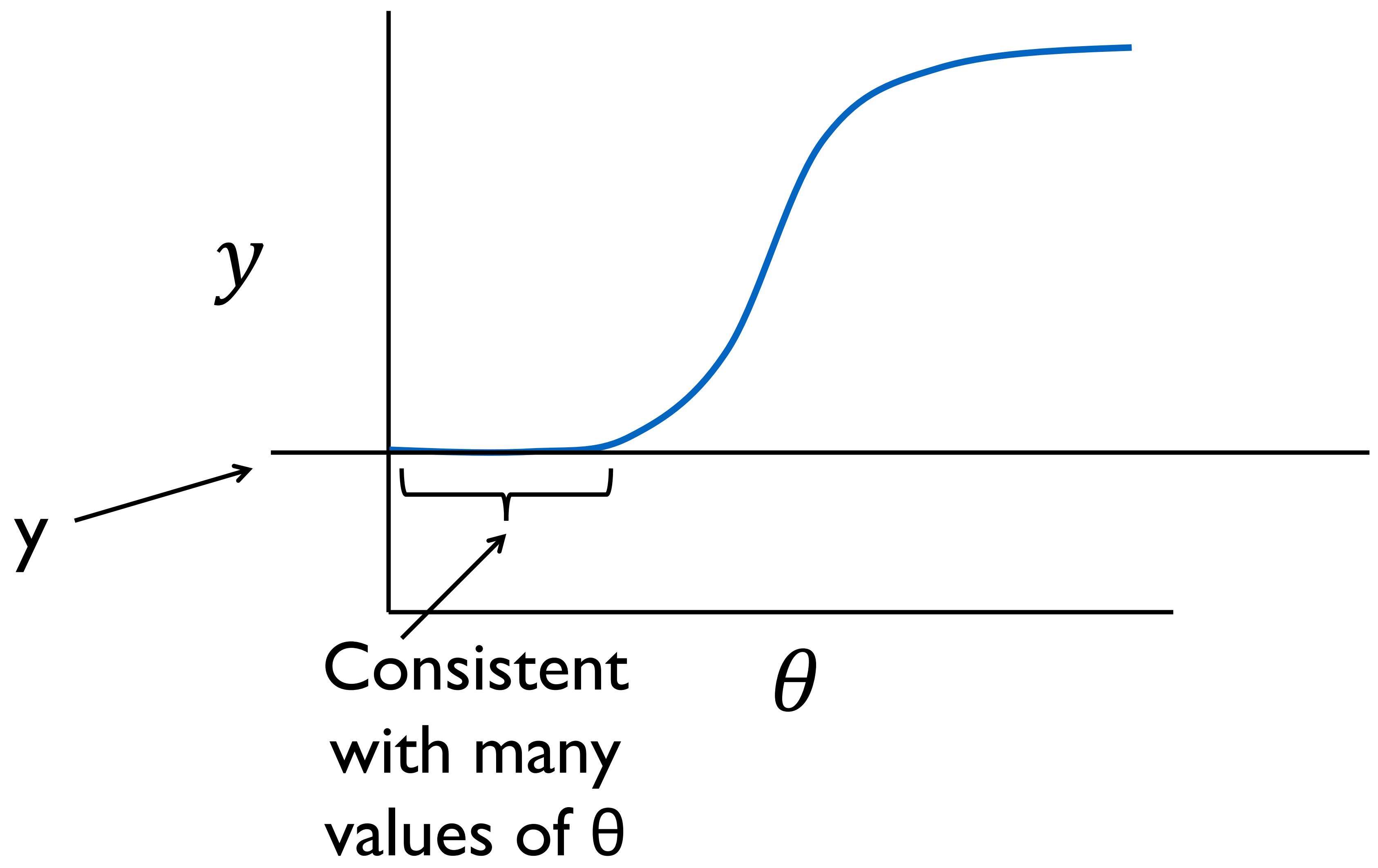
- Sensitivity - how much the model output changes as a function of the parameter(s)

- Robustness - is the model able to reproduce similar behavior across a range of parameter space?

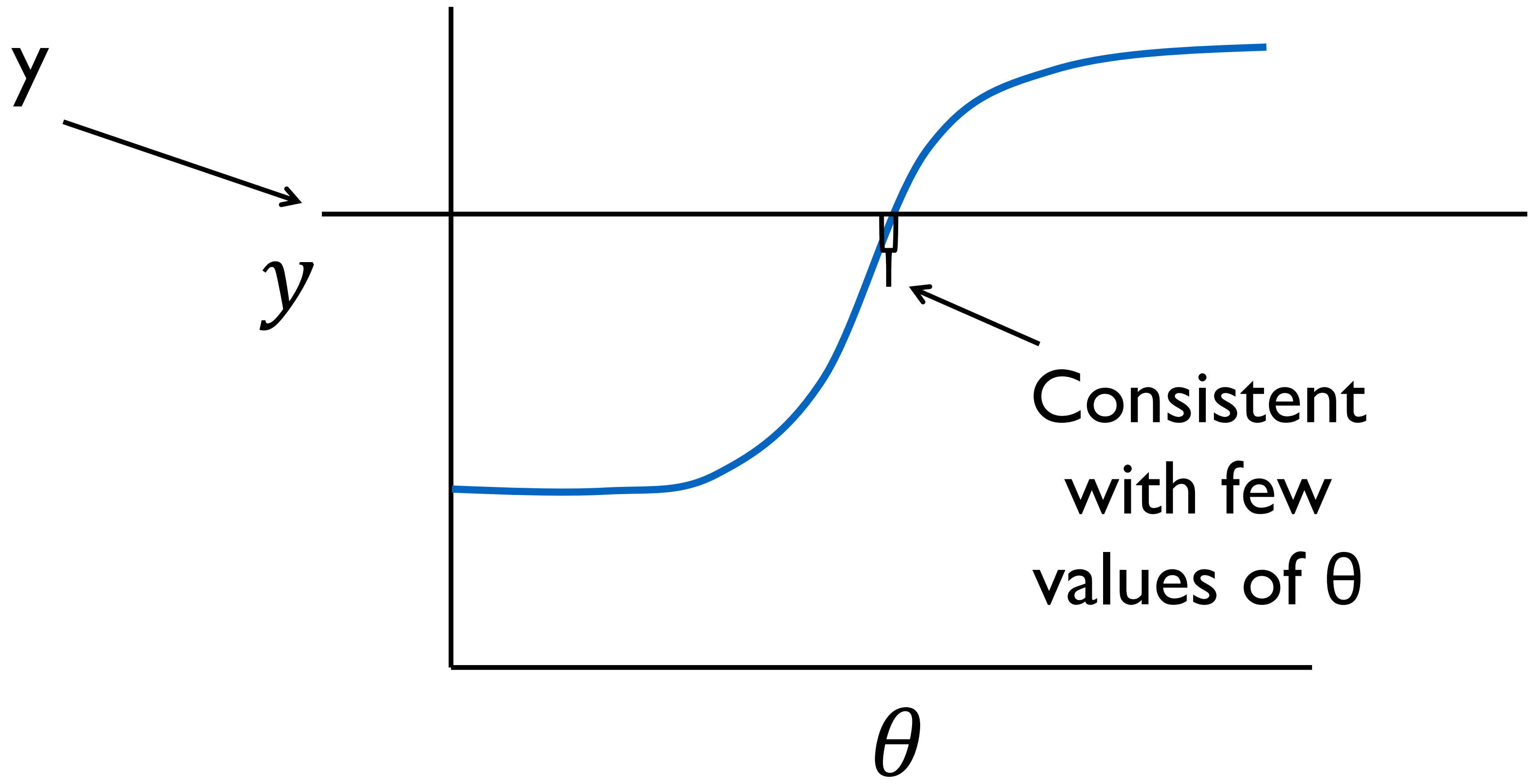
Sensitivity/robustness/identifiability tradeoff

- When a behavior is robust, we may have more confidence in it—but, this means we cannot be sure of what parameters generated the behavior
    - Unidentifiability
- Similarly, when the output is highly sensitive, we may be better able to infer what parameter conditions must be

Sensitivity/robustness tradeoff



Sensitivity/robustness tradeoff

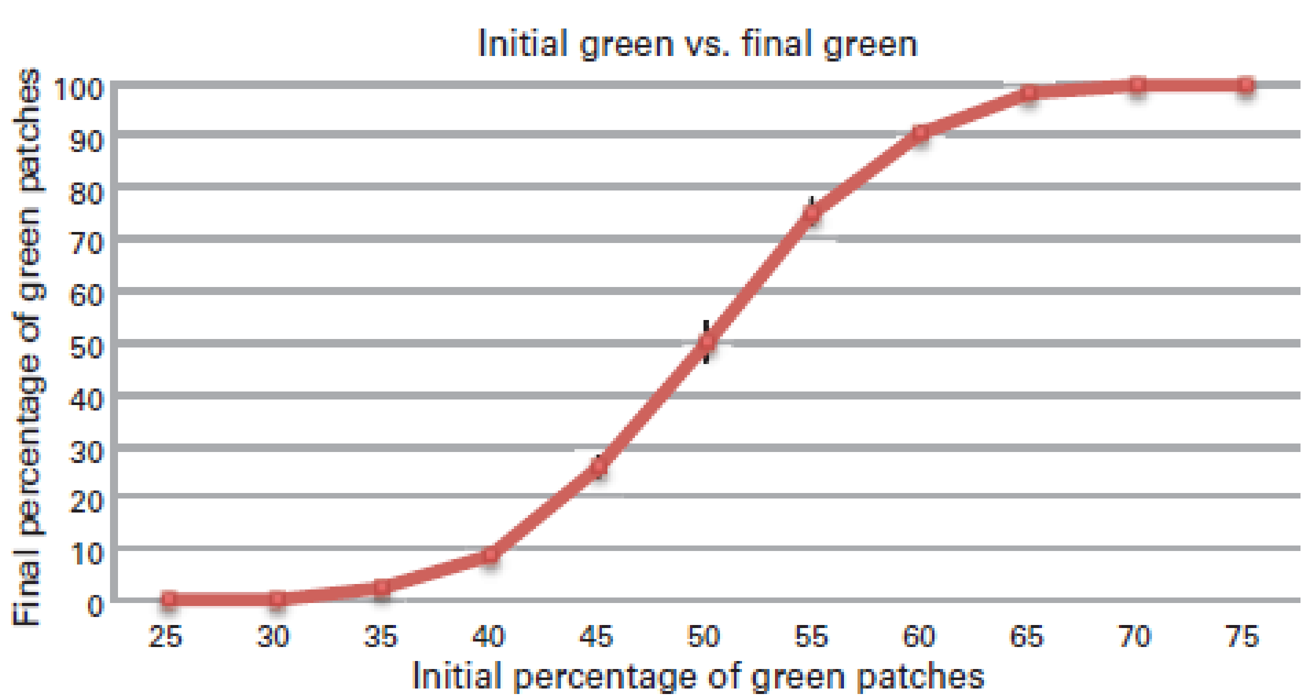


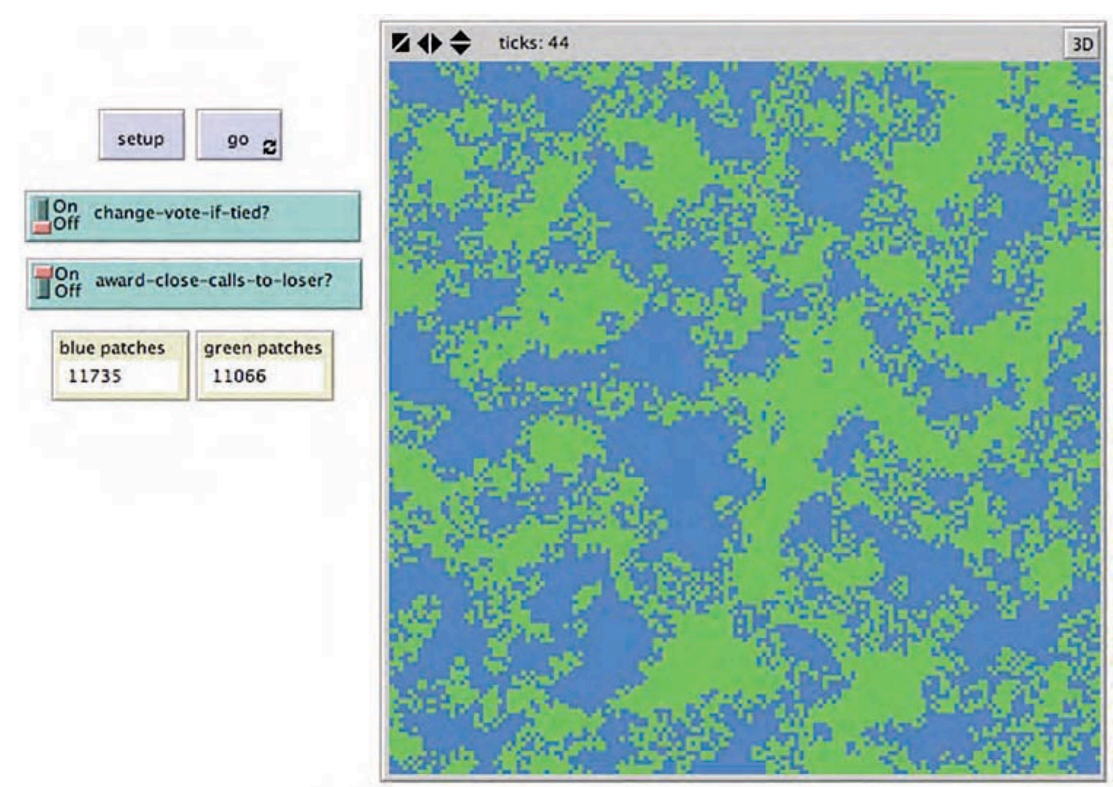
Local Methods

- One-at-a-time approaches
- Derivative methods for local sensitivity

One-at-a-time approach

- Adjust one parameter at a time, fixing the rest to pre-specified values
- Example: voting model initial fraction yes/no





Derivative-based local sensitivity

- Model $\dot{x} = f\left(x, t, p\right)$
- Output $y = g(x, t, p)$

- Output **sensitivity** to parameter variations

$$dy / dp$$

- Meaning depends on magnitude of y and p—often more useful to look at **relative sensitivity**

$$\frac{dy / y}{dp / p} = \frac{dy}{dp}\frac{p}{y}$$

How to calculate local sensitivity?

- Many methods—practically speaking, often done simply by testing small perturbations (e.g. 5% change) of the parameters and seeing how the output changes

- Relative sensitivity:

$$\frac{dy/y}{dp/p} \approx \frac{\Delta y/y}{\Delta p/p} = \frac{\%\text{ change in } y}{\%\text{ change in } p}$$

Forward sensitivity equations for ODEs

- Extended ODE system that allows simulation of the model and the sensitivity functions at the same time

$$
\begin{aligned}
\frac{dx}{dt} &= g(t, x(t,\theta), \theta) \\
\frac{d}{dt}\frac{\partial x}{\partial \theta} &= \frac{\partial g}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial g}{\partial \theta},
\end{aligned}
$$

- Take your ODE system and apply $\partial/\partial\theta$ (with chain rule and assuming you can switch derivative order)

SIR model

$$\frac{dS}{dt} = -\beta S \frac{I}{N}$$
$$\frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

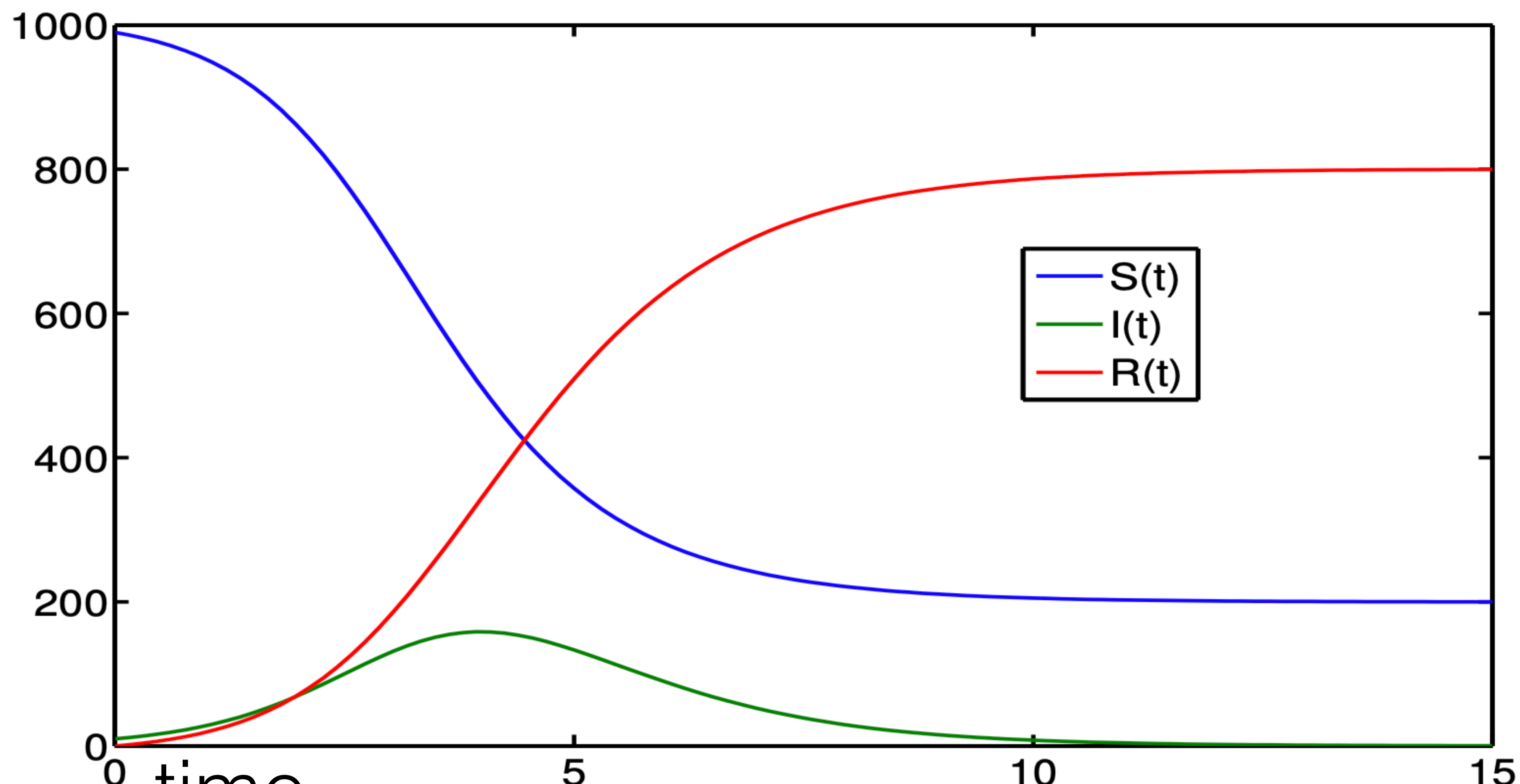
y = I(t)

- We assume the initial conditions $(S(0), I(0), R(0))$ and the population size N are known.
- The vector of model parameters is $\theta = (\beta, \gamma)$.
- The basic reproductive number is $\mathcal{R}_0 = \beta/\gamma$. Whenever $\beta/\gamma > 1$ then an outbreak occurs.

Borrowed from Ariel Cintron-Arias NIMBioS 2014 Parameter Estimation Tutorial

$$\frac{dS}{dt} = -\beta S \frac{I}{N}$$
$$\frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

y = I(t)



time

Borrowed from Ariel Cintron-Arias NIMBioS 2014 Parameter Estimation Tutorial

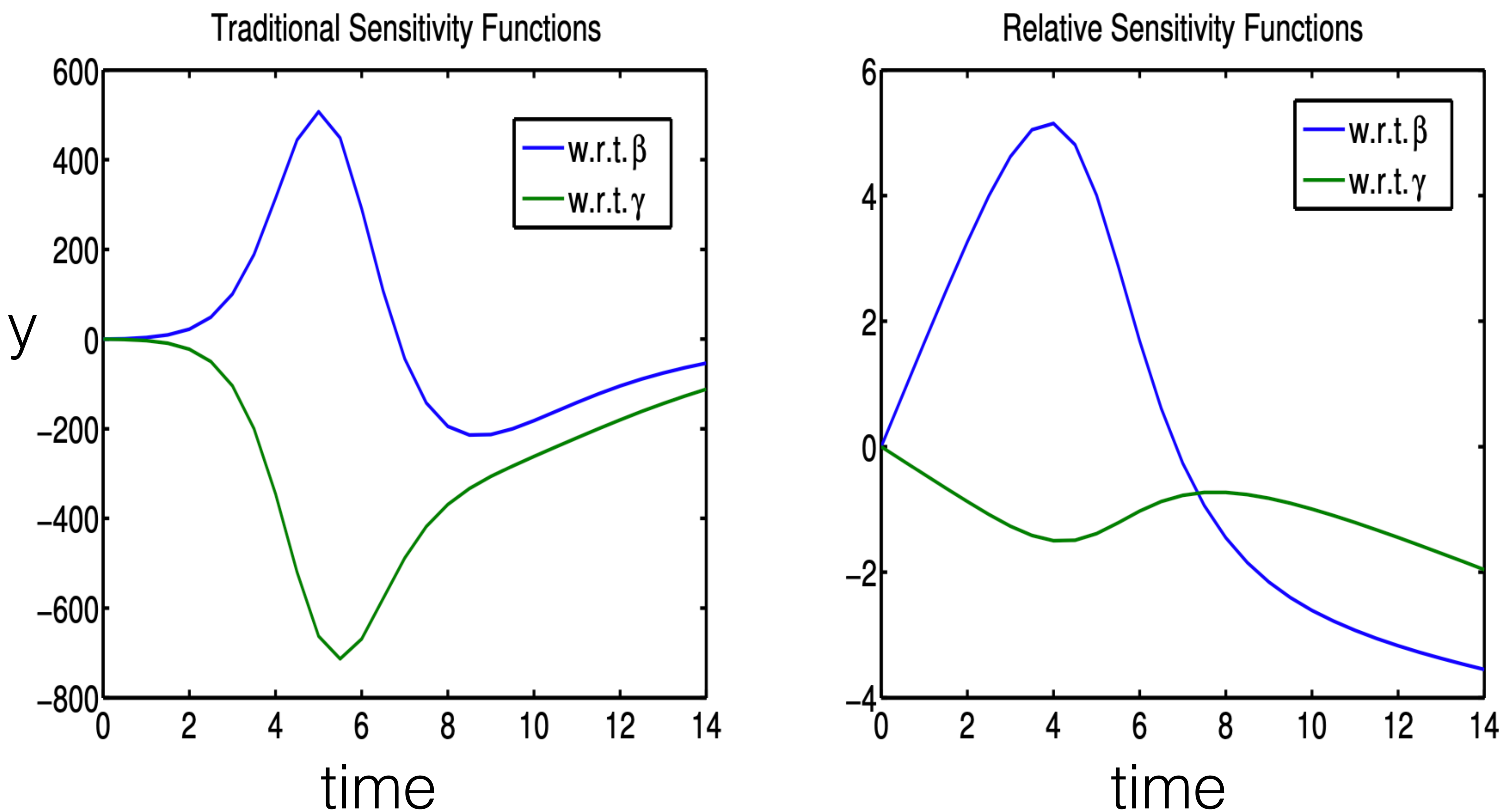
$$
\begin{aligned}
\frac{d}{dt}x(t) &= g(x(t,\hat{\theta}),\hat{\theta}), \\
\frac{d}{dt}\phi(t) &= \frac{\partial g}{\partial x}\phi(t) + \frac{\partial g}{\partial \theta},
\end{aligned}
$$

where $\phi(t) = \frac{\partial x}{\partial \theta}(t,\theta)$, $\frac{\partial g}{\partial x} = \begin{bmatrix} -\hat{\beta}\frac{I}{N} & -\hat{\beta}\frac{S}{N} & 0 \\ \hat{\beta}\frac{I}{N} & \hat{\beta}\frac{S}{N} - \hat{\gamma} & 0 \\ 0 & \hat{\gamma} & 0 \end{bmatrix}$ and

$$
\frac{\partial g}{\partial \theta} = \begin{bmatrix} -S\frac{I}{N} & 0 \\ S\frac{I}{N} & -I \\ 0 & I \end{bmatrix}.
$$

Borrowed from Ariel Cintron-Arias NIMBioS 2014 Parameter Estimation Tutorial

Sensitivity functions



Borrowed from Ariel Cintron-Arias NIMBioS 2014 Parameter Estimation Tutorial

Sensitivity Analysis

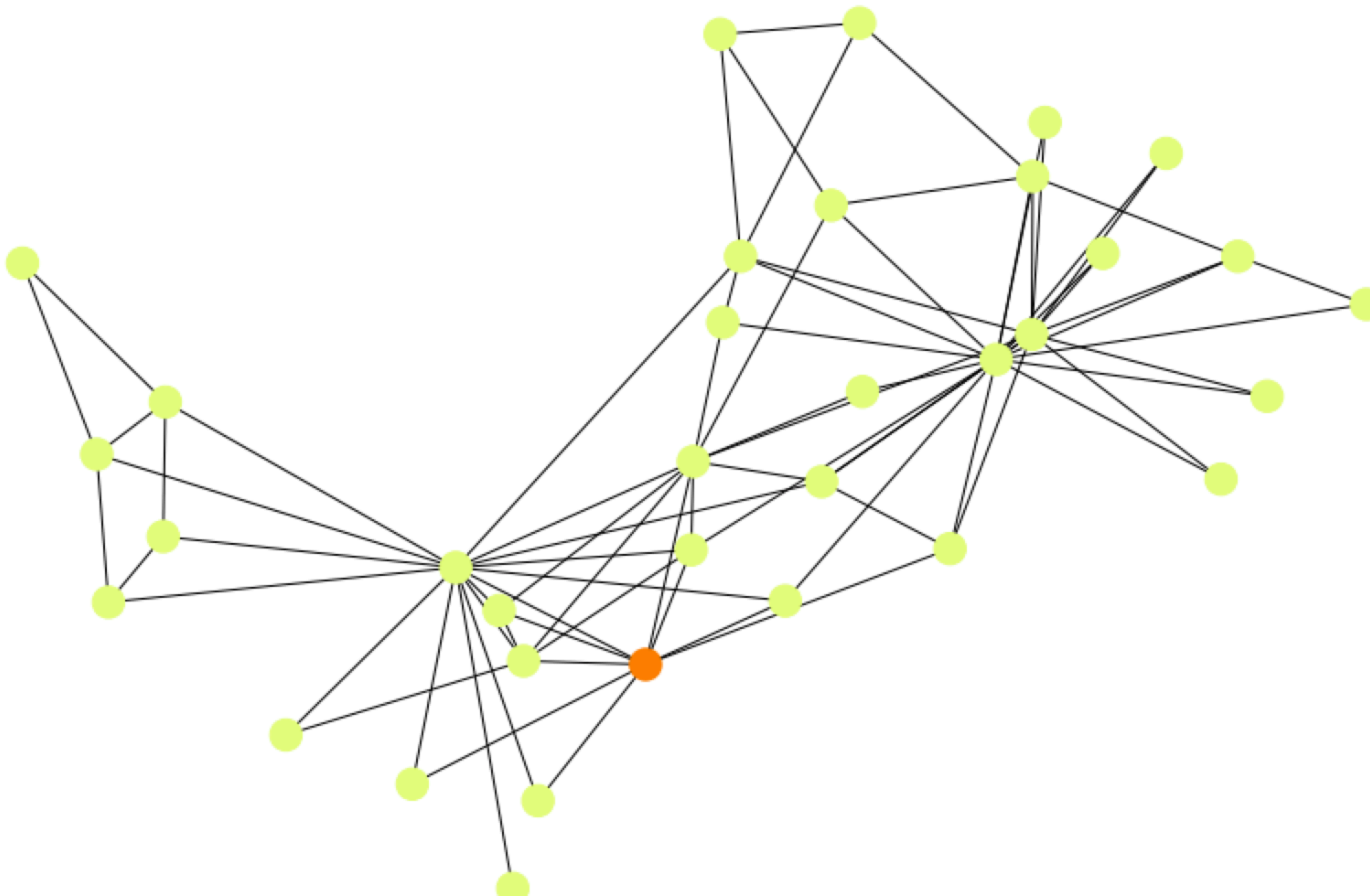
- Sensitivity is inherently a local attribute
- But often we want to know about **global sensitivity** over a wide range of values of θ
- Helps to know where to allocate resources in general for a variety of scenarios

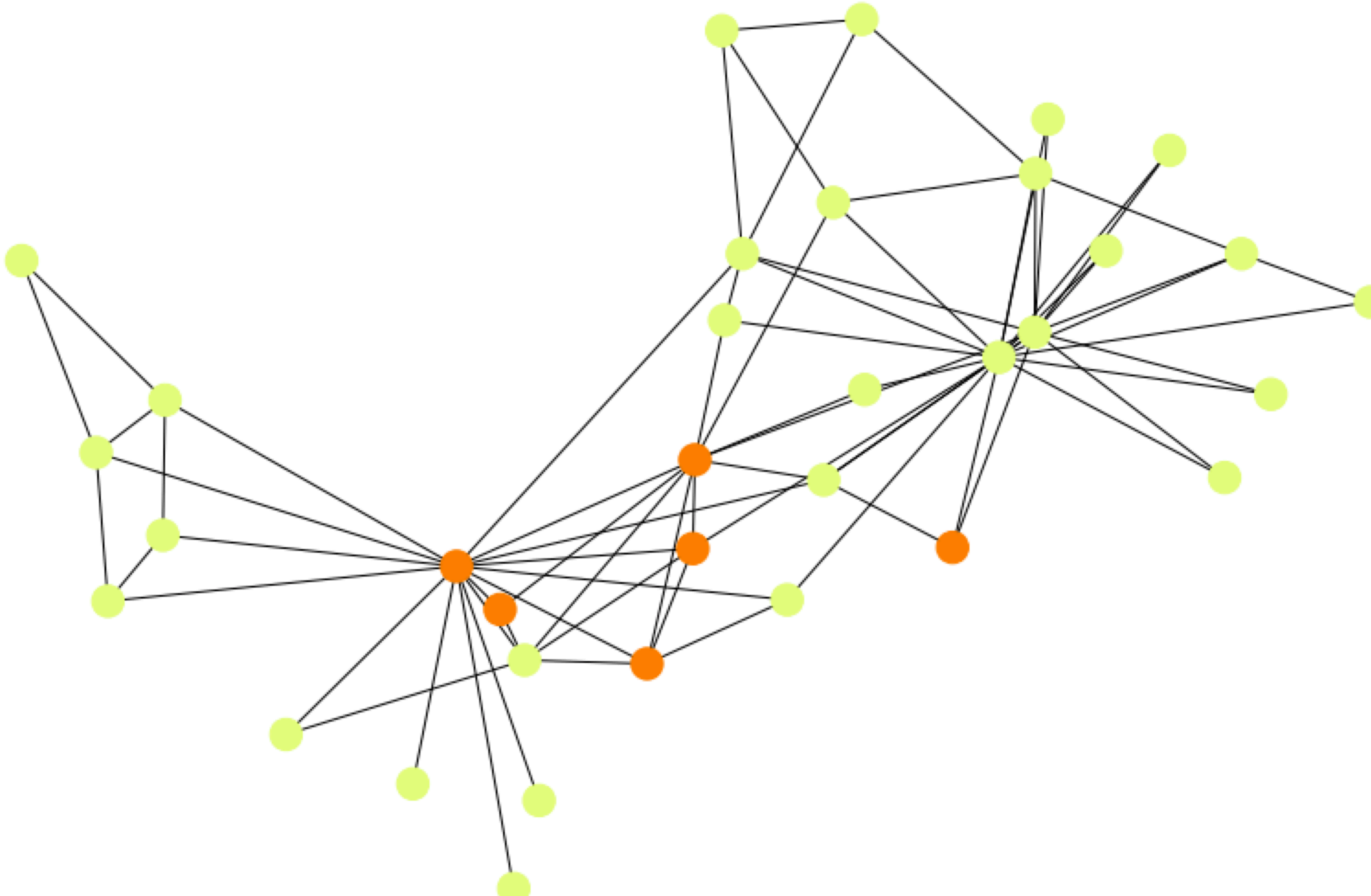
Global Methods

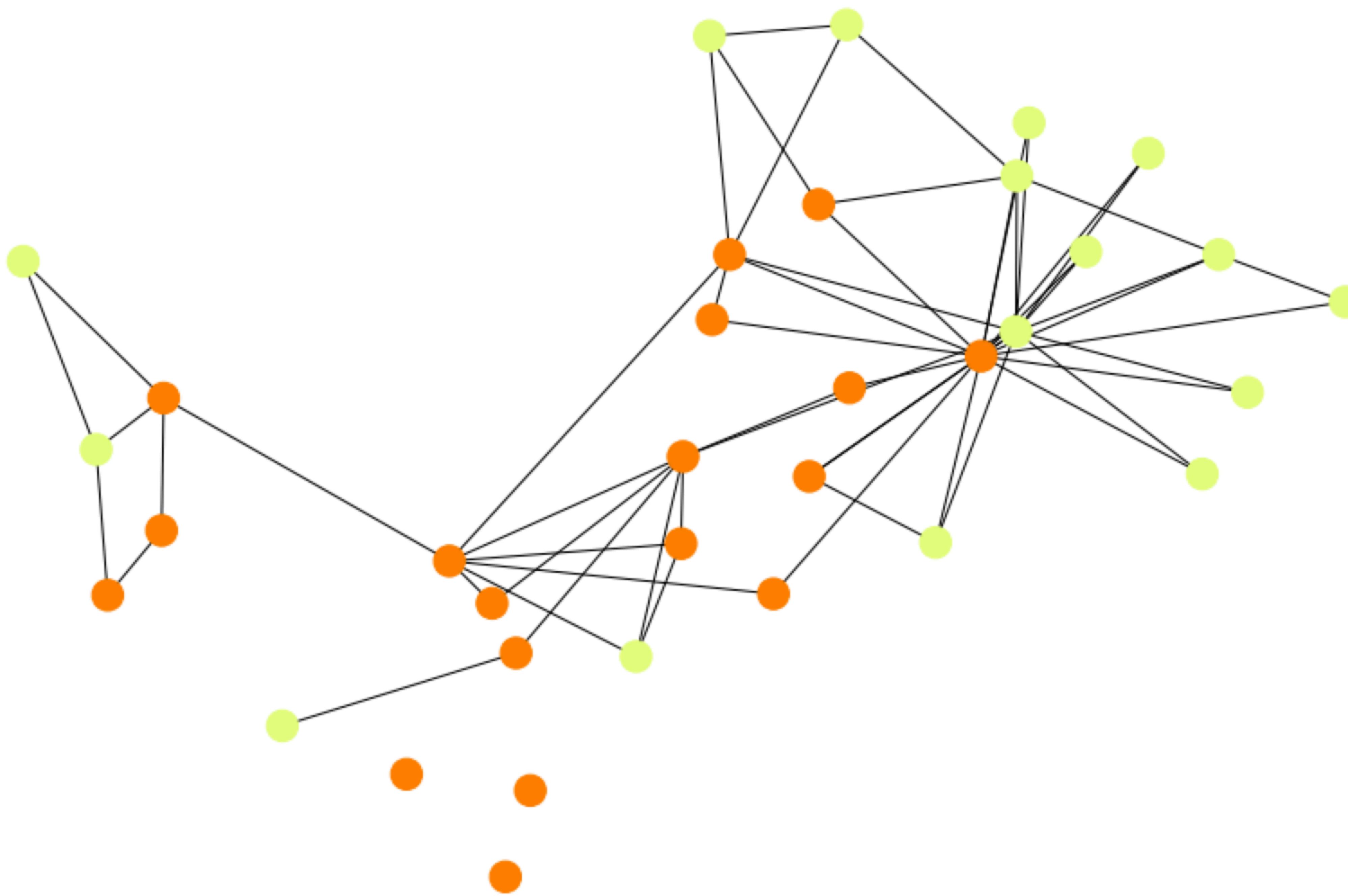
- Sampling-based methods
  - Visual approaches! (sample and look at scatterplots, etc.)
  - Regression-based methods
  - Variance-based methods

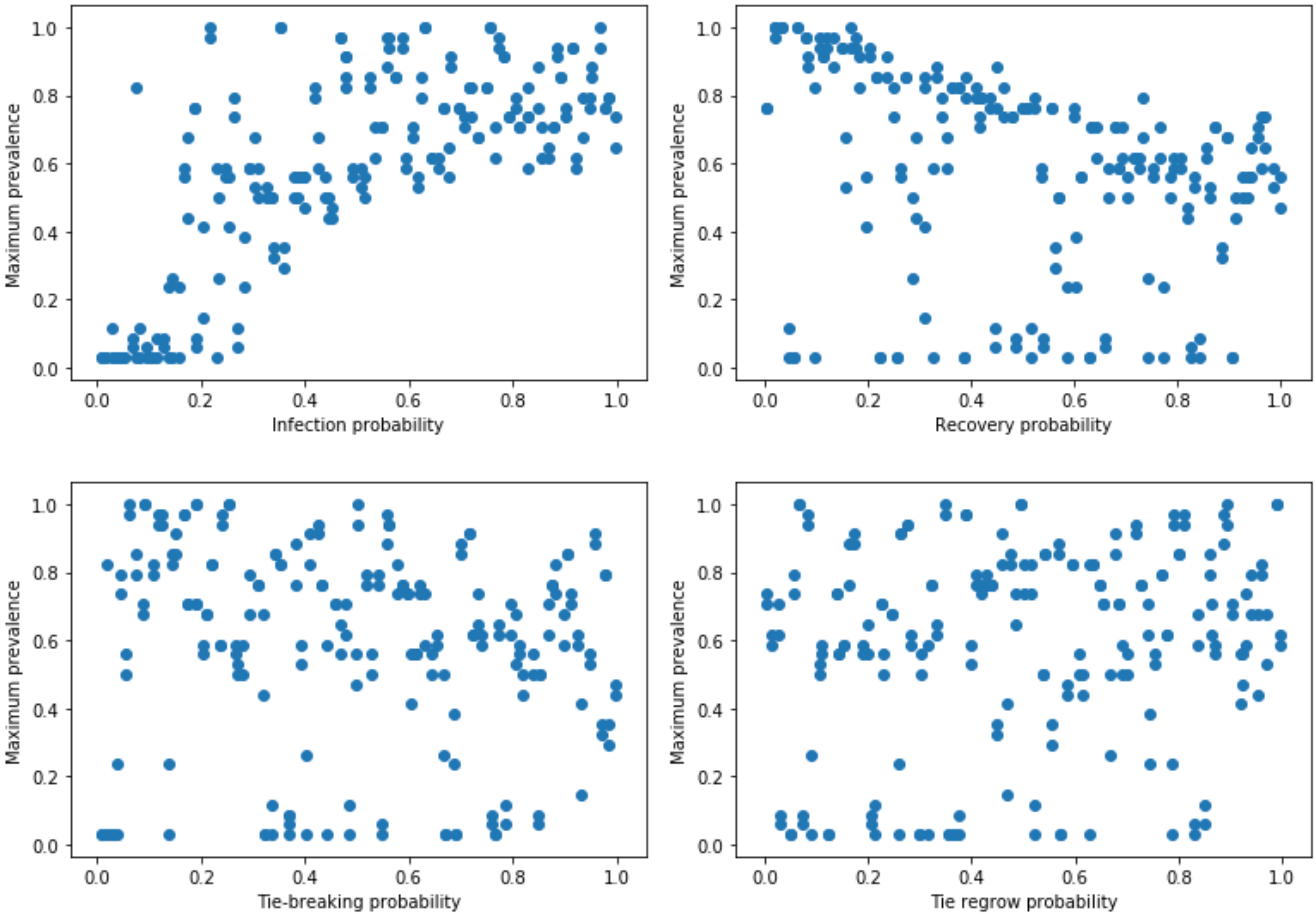
Sensitivity analysis: sampling & visualization

- Sample parameter space, then plot relationships between parameters and model output(s)
- Example: SIR model on the karate club network with tie breaking and regrowth









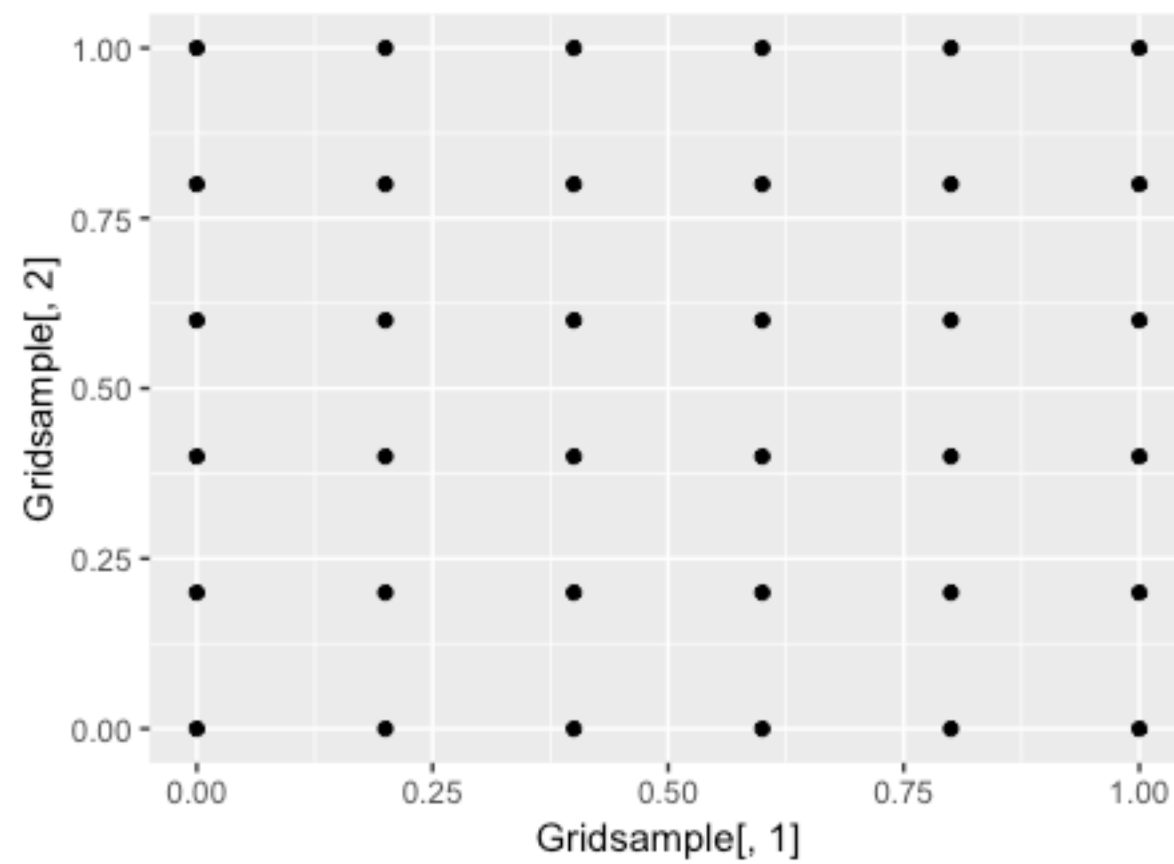
Maximum prevalence
Infection probability
Maximum prevalence
Recovery probability
Maximum prevalence
Tie-breaking probability
Maximum prevalence
Tie regrow probability

How to sample?

- Global sensitivity relies on sampling a sometimes high-dimensional parameter space

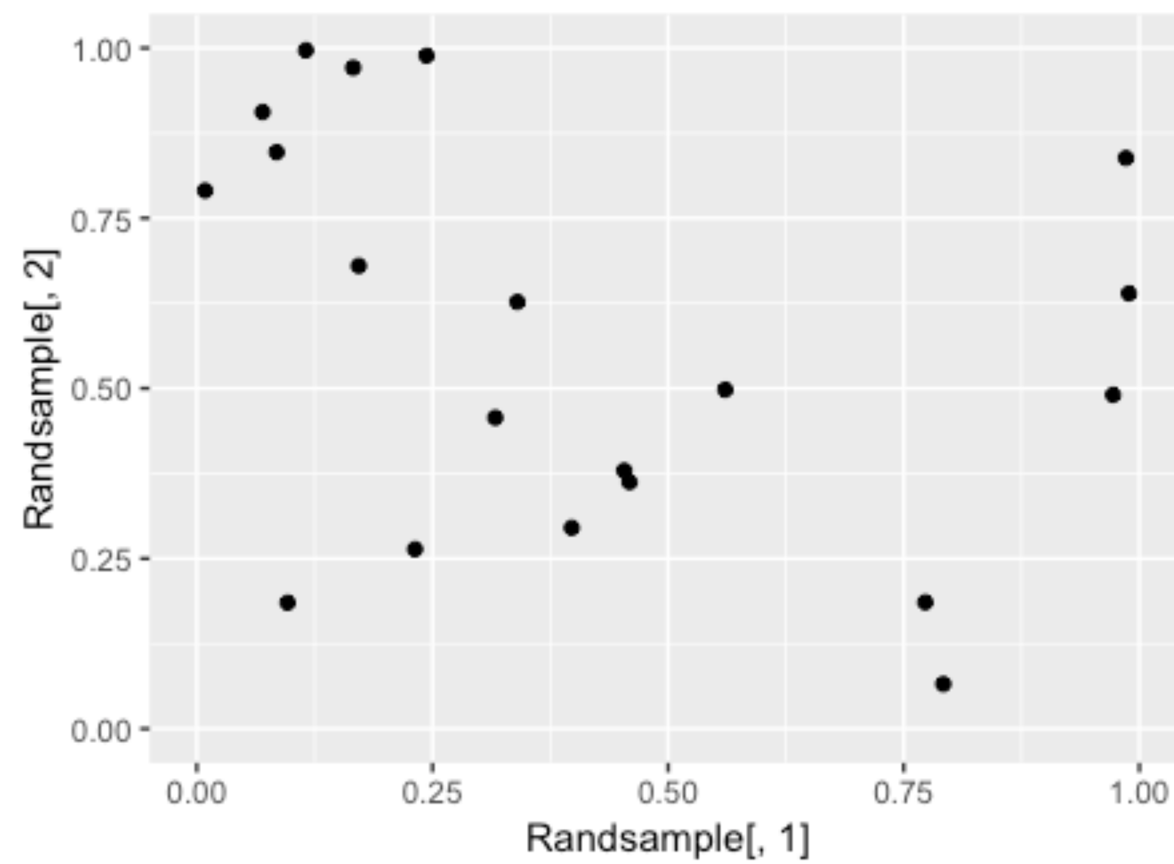
Sampling parameter space

- **Grid sampling**
- Typically choose uniform distribution of points
- Good coverage of space
- Computationally expensive! Becomes infeasible as dimension increases



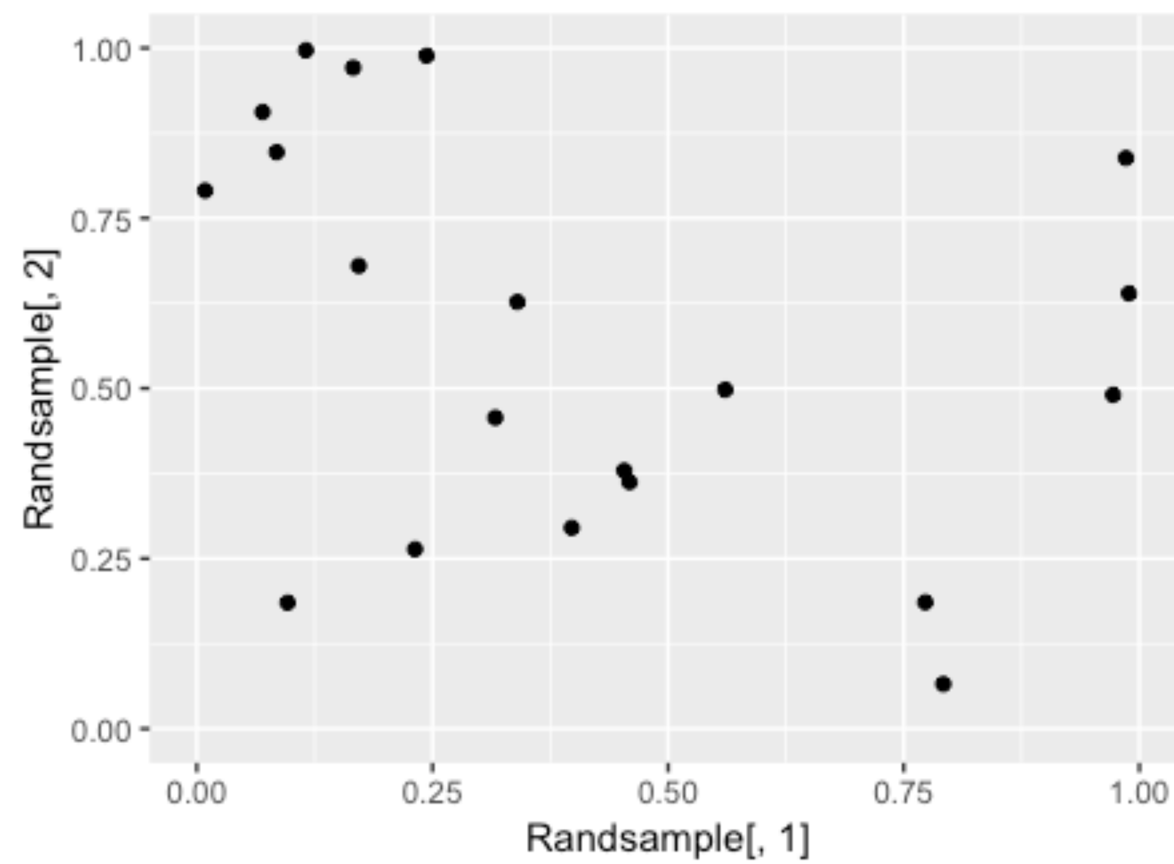
Sampling parameter space

- **Random sampling**
- Often done with uniform distribution, but can choose any distribution
- However, may leave big blank spots, require many samples to fully explore the space



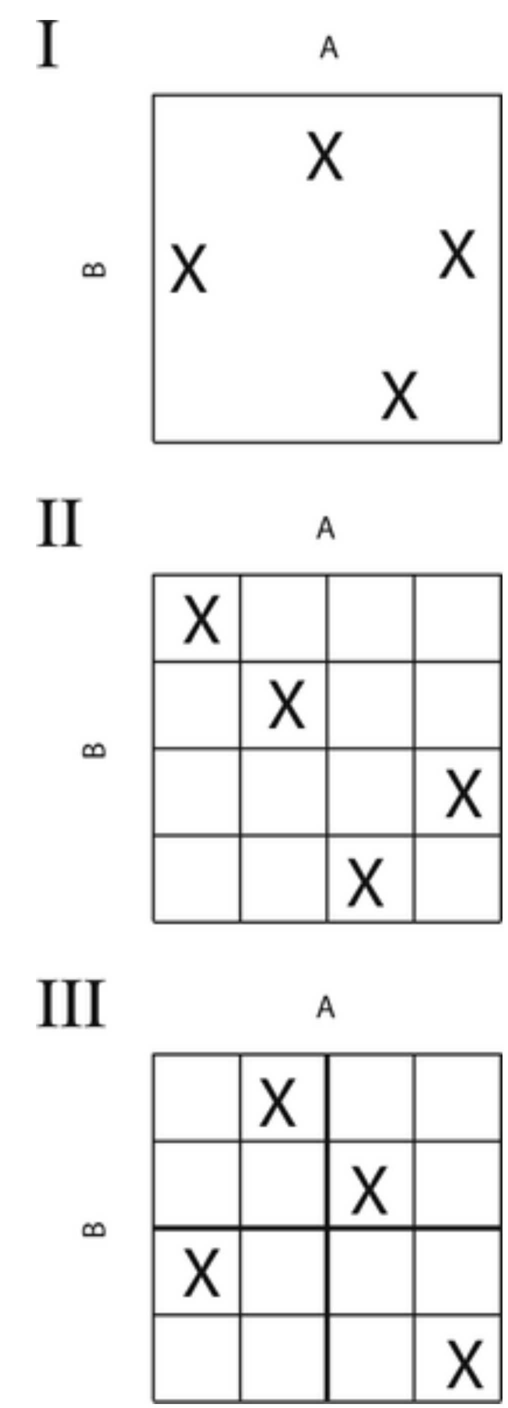
Sampling parameter space

- More efficient ways to explore the space?
  - Latin hypercube sampling (& variants, orthogonal, etc.)
  - Sobol sampling (& other low-discrepancy sequences)



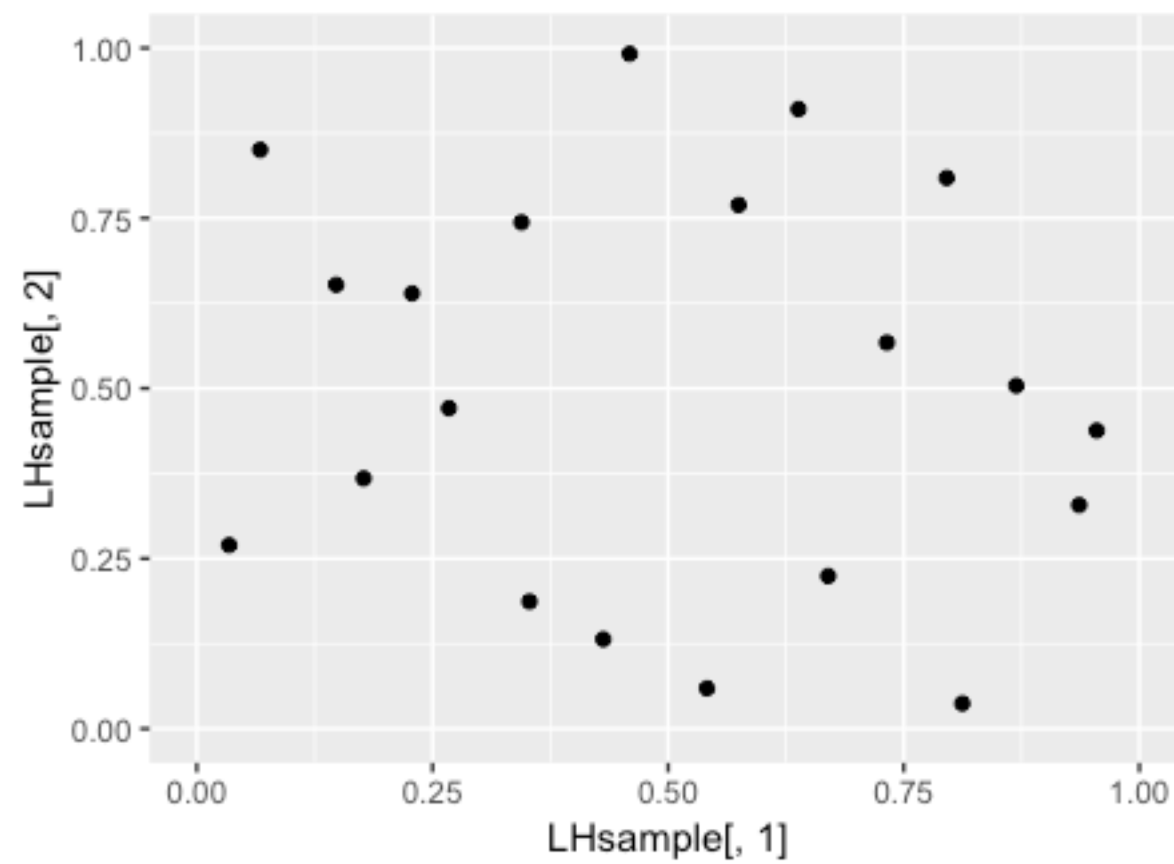
Latin hypercube sampling

- Kind of like sudoku
- Divide space into a grid of rows & columns
- Choose one square in each row and each column
- Choose a random point within that square



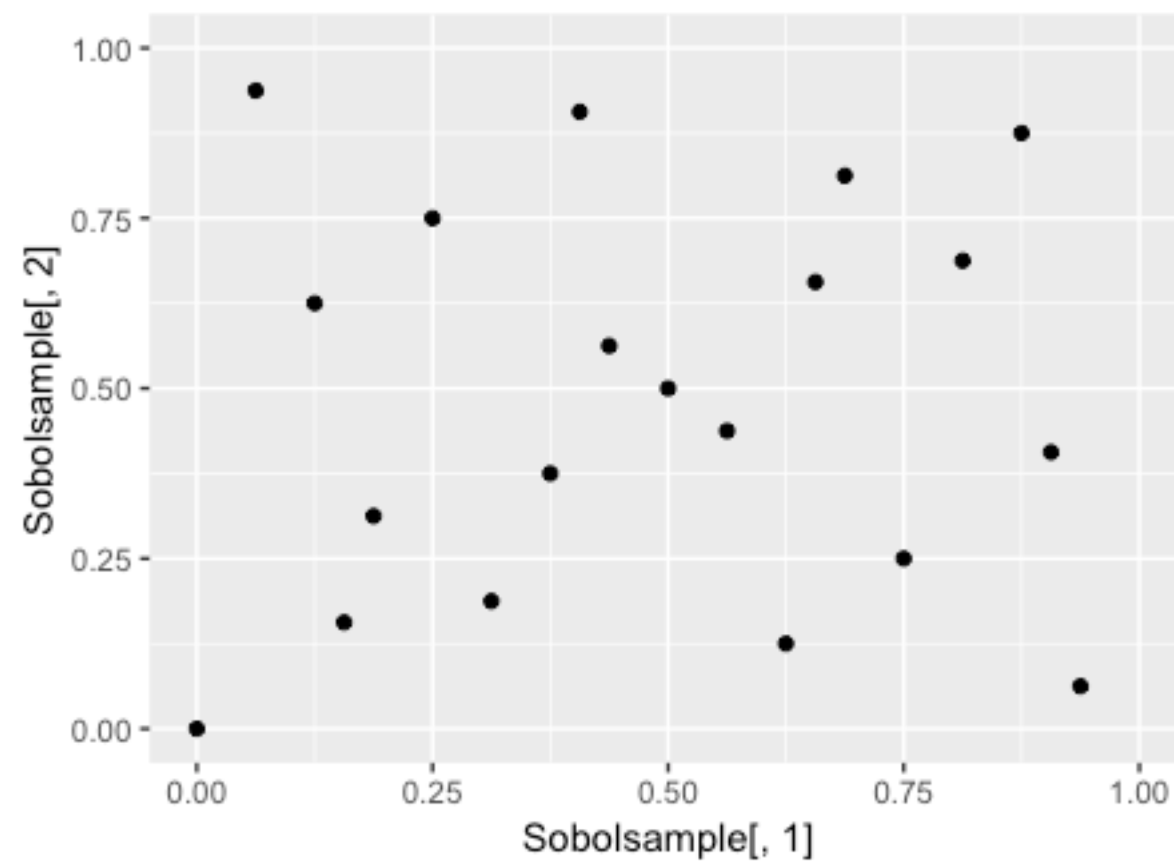
Latin hypercube sampling

- Still an element of randomness
- Ensures better coverage of the space/faster convergence to the sampled distribution



Sobol sampling

- Low-discrepancy sequence (see also Halton, Faure)
- Generates a sequence that samples the space evenly but requires few points
- Convergence can be better than LHS



Kucherenko et al., Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques. https://arxiv.org/abs/1505.02350

How many samples to take?

- Tough to say! Balance computational intensiveness with good coverage (often >>100, e.g. in the 1K to 10K range depending on number of parameters)

- May need to run more than one sample for a given point due to stochasticity (since different runs may give different behaviors)

- Some methods have rules of thumb, e.g. for LHS, $N_S > (4/3) \times N_P$ has been proposed, but you will often want much more than this bound

Blower and Dowlatabadi (1994), McKay et al. (1979)

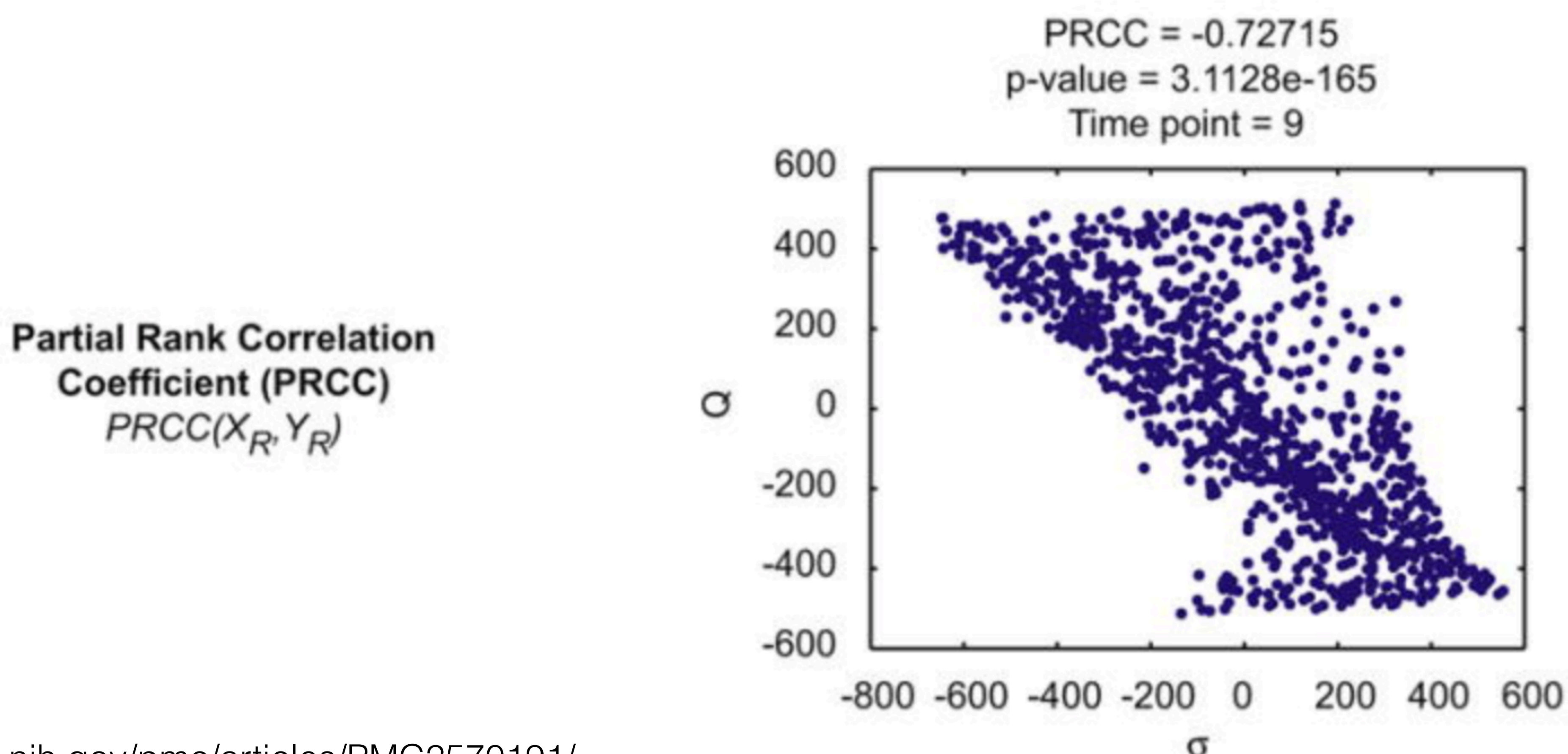
Regression based methods

- Fit linear trend to the data
- Pearson correlation coefficient - correlate parameter and output
  - However, only works for linear relationships
- For nonlinear but monotonic relationships, rank-based correlation coefficients are often useful (draw example)

Partial rank correlation coefficient

- A rank correlation coefficient that accounts for the effects of the other parameters

- Requires monotonic relationship with the output



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2570191/

Variance-based methods

- Decomposition of variance (also called the Sobol method)
- Determines how much of the variance in output is due to each parameter
- Analogous to an ANOVA
- Direct calculation
- Faster options, e.g. eFAST

Word of caution: interpreting statistical results on model outputs

- What does the p-value on one of these regression or variance-based statistics mean?

- What is the source of the uncertainty in these estimates?

Dimension reduction & parameter selection

- Fixing insensitive parameters
- Parameter subset selection methods
  - Find subsets of parameters or potentially new parameter combinations that explain most of the behavior (by fixing parameters or parameter combinations that are insensitive)

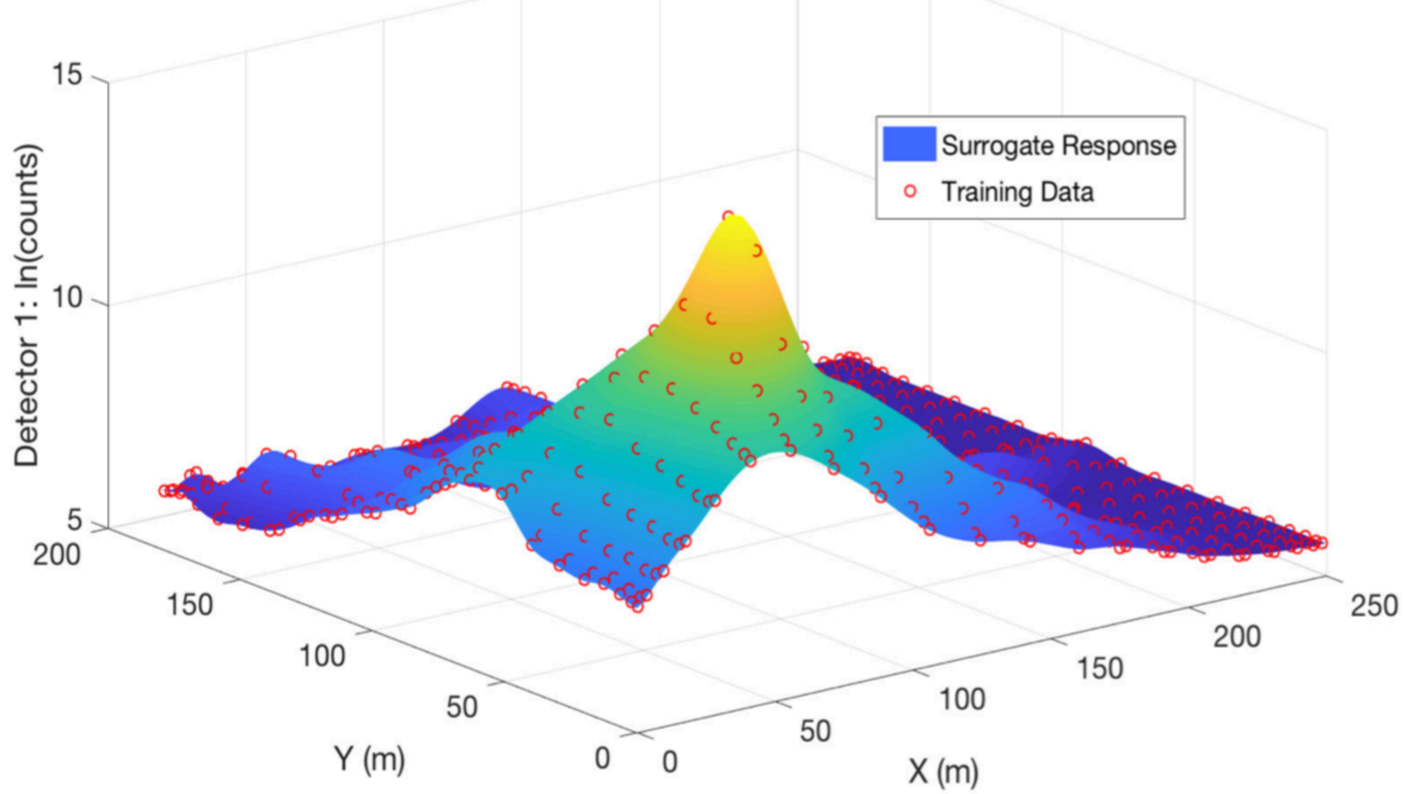
Surrogate models

- Global sensitivity methods can be highly computationally expensive—many ABMs take too long to run to be feasible with the number of samples needed to explore space

- Surrogate models (also called emulators, response surfaces) provide another option

Surrogate models

- Idea is to fit a surface or function to the model output(s) as a function of the parameters
  - Choose a functional form that is cheap to evaluate many times (e.g. polynomial, linear)



Cook JA, Smith RC, Hite JM, Stefanescu R, Mattingly J. Application and Evaluation of Surrogate Models for Radiation Source Search. Algorithms. 2019 Dec;12(12):269.

Surrogate models

- Use a smaller number of points to fit the surface, then sample a large number of points to run sensitivity analysis

- Re-run using the true model on regions of interest

Next time - Journal Club

- Marino, Simeone, et al. "A methodology for performing global uncertainty and sensitivity analysis in systems biology." Journal of theoretical biology 254.1 (2008): 178-196.