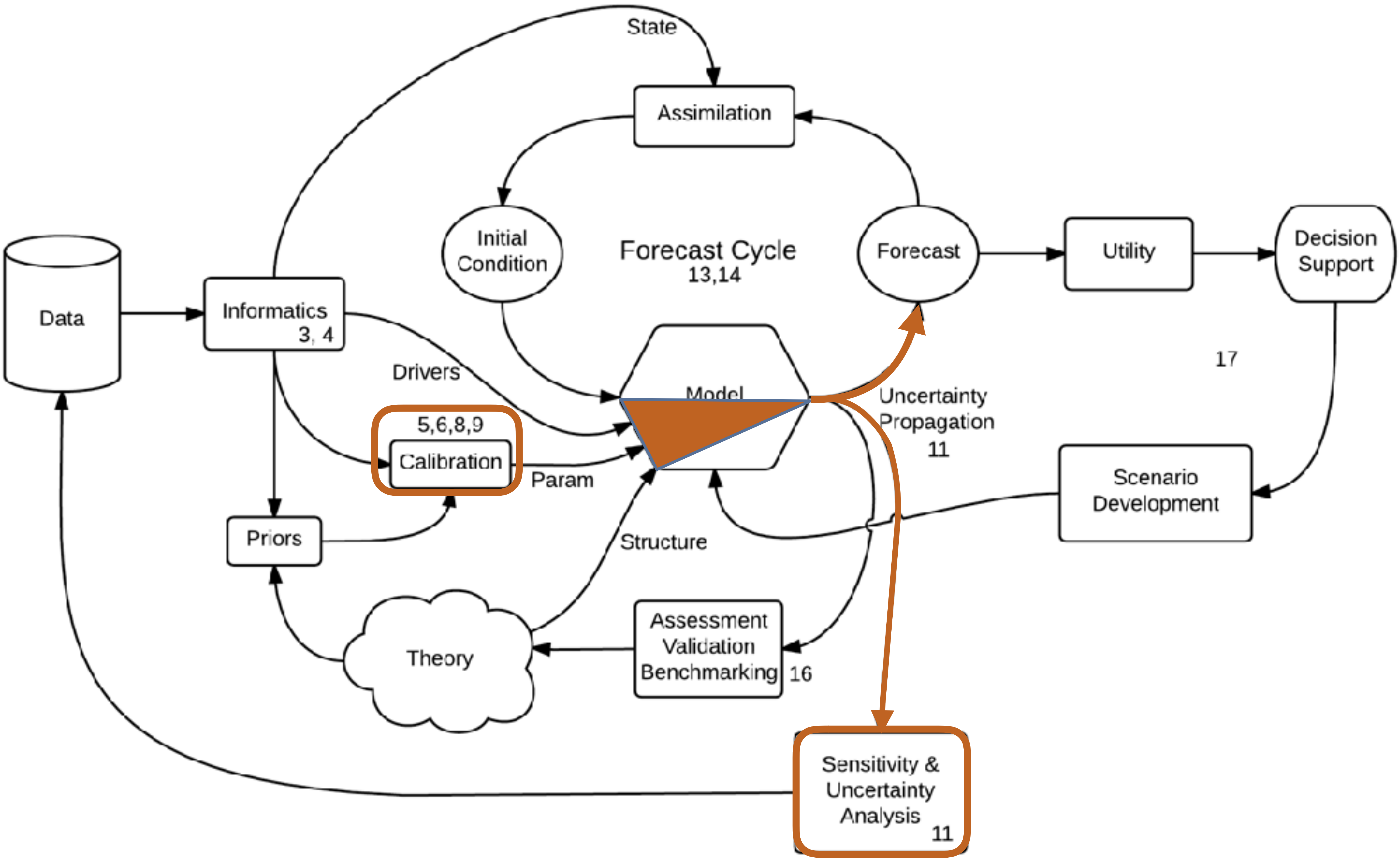
PROPAGATING, ANALYZING, AND REDUCING UNCERTAINTY


State
Assimilation
Initial Condition
Forecast Cycle
13,14
Forecast
Utility
Decision Support
Data
Informatics
3, 4
Drivers
Model
Uncertainty Propagation
11
17
5,6,8,9
Calibration
Param
Scenario Development
Priors
Structure
Assessment Validation Benchmarking
16
Theory
Sensitivity & Uncertainty Analysis
11

Concepts

- Sensitivity Analysis $\partial x \rightarrow \partial y$
  How does a change in X translate into a change in Y?
- Uncertainty Propagation $Var[x] \rightarrow Var[y]$
  How does the uncertainty in X affect the uncertainty in Y?
  How do we forecast Y with uncertainty?
- Uncertainty Analysis
  which sources of uncertainty are most important?
- Optimal Design
  How do we best reduce the uncertainty in our forecast?

Sensitivity Methods

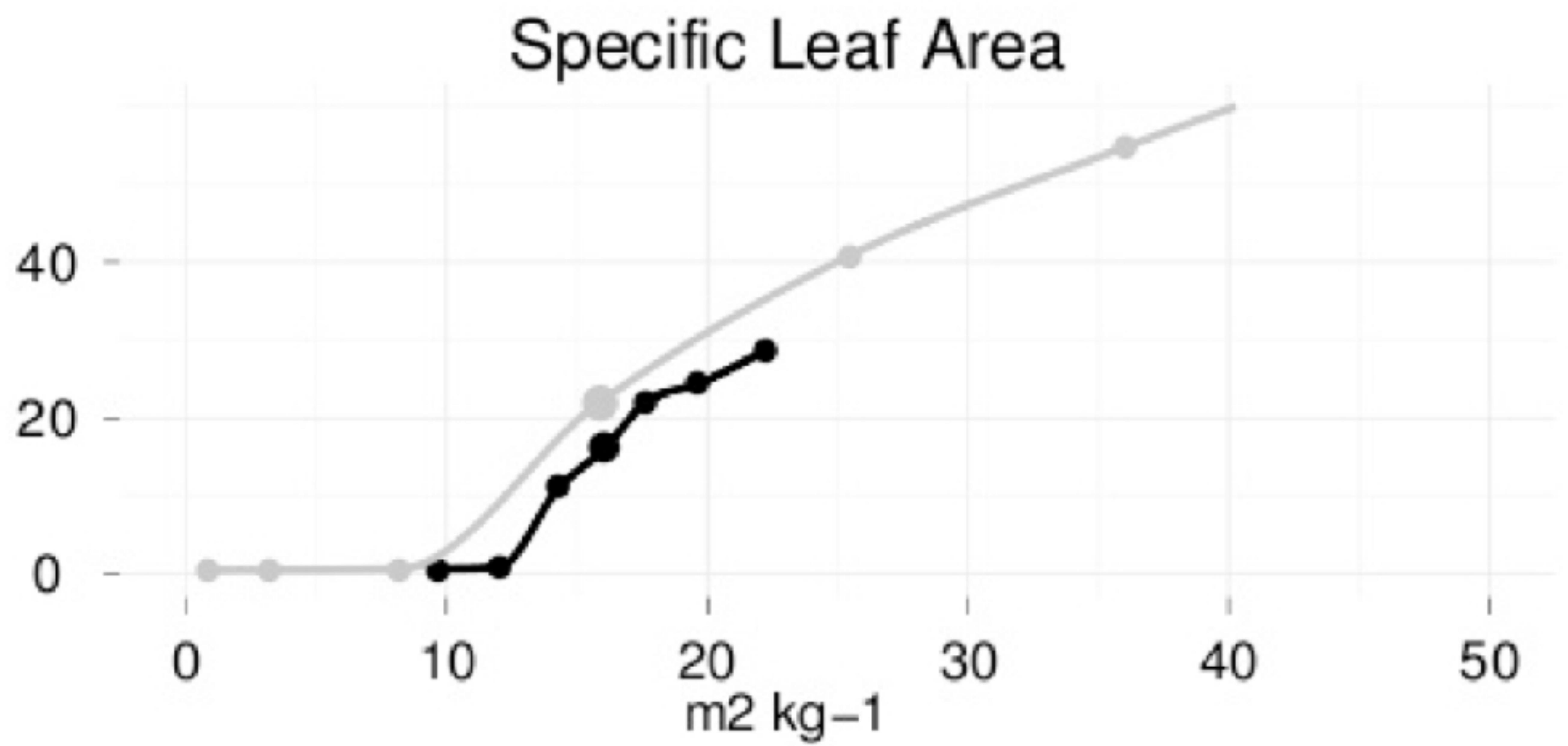
- Local
  - Analytical: $df/d\Theta$
  - One-at-a-time perturbations
- Global
  - Monte Carlo
  - Sobol
  - Emulators
  - Elementary Effects
  - Group Sampling

Extensive but Costly

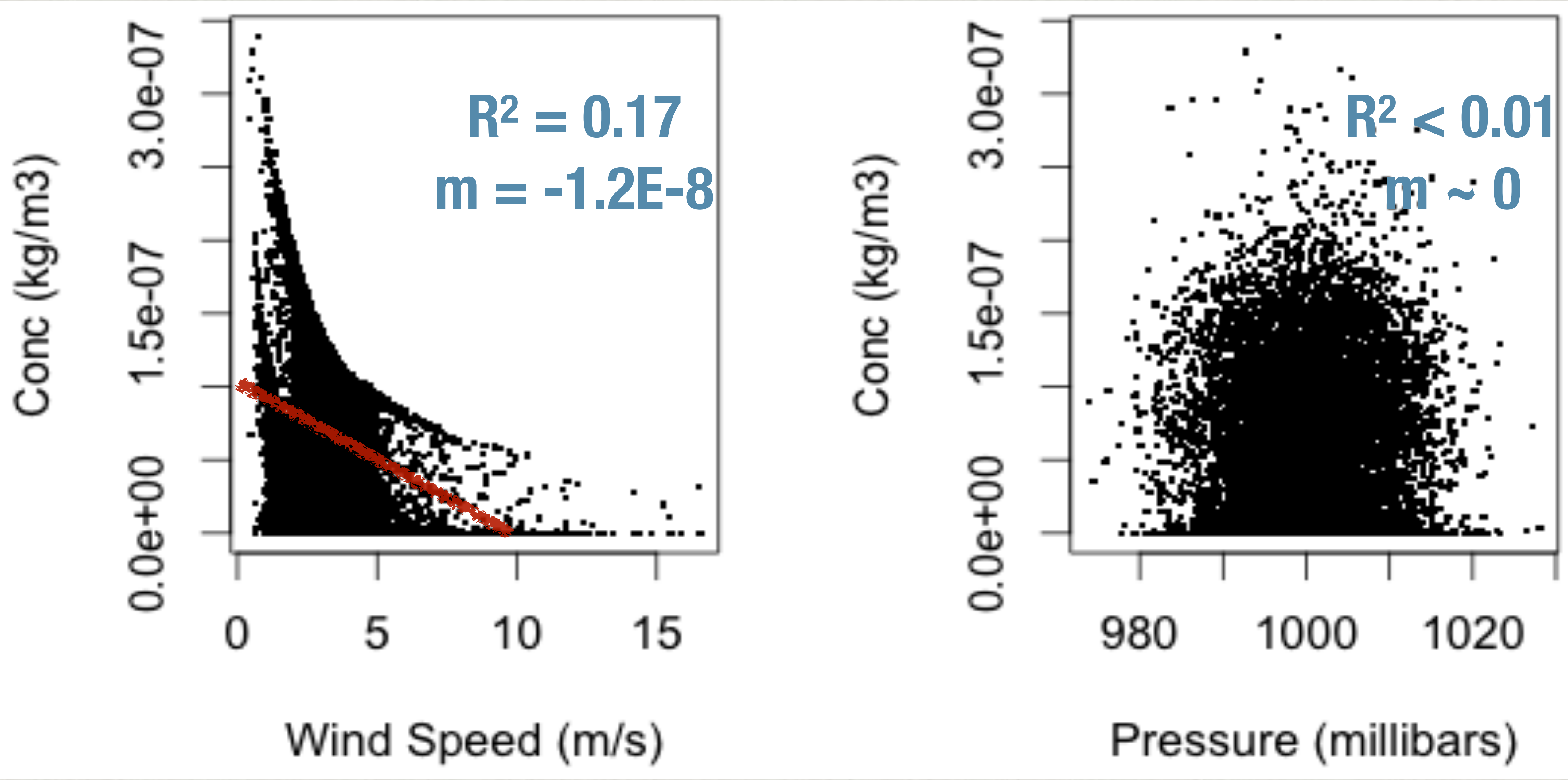
Sparse but Cheap

Saltelli et al. 2008. Global Sensitivity Analysis

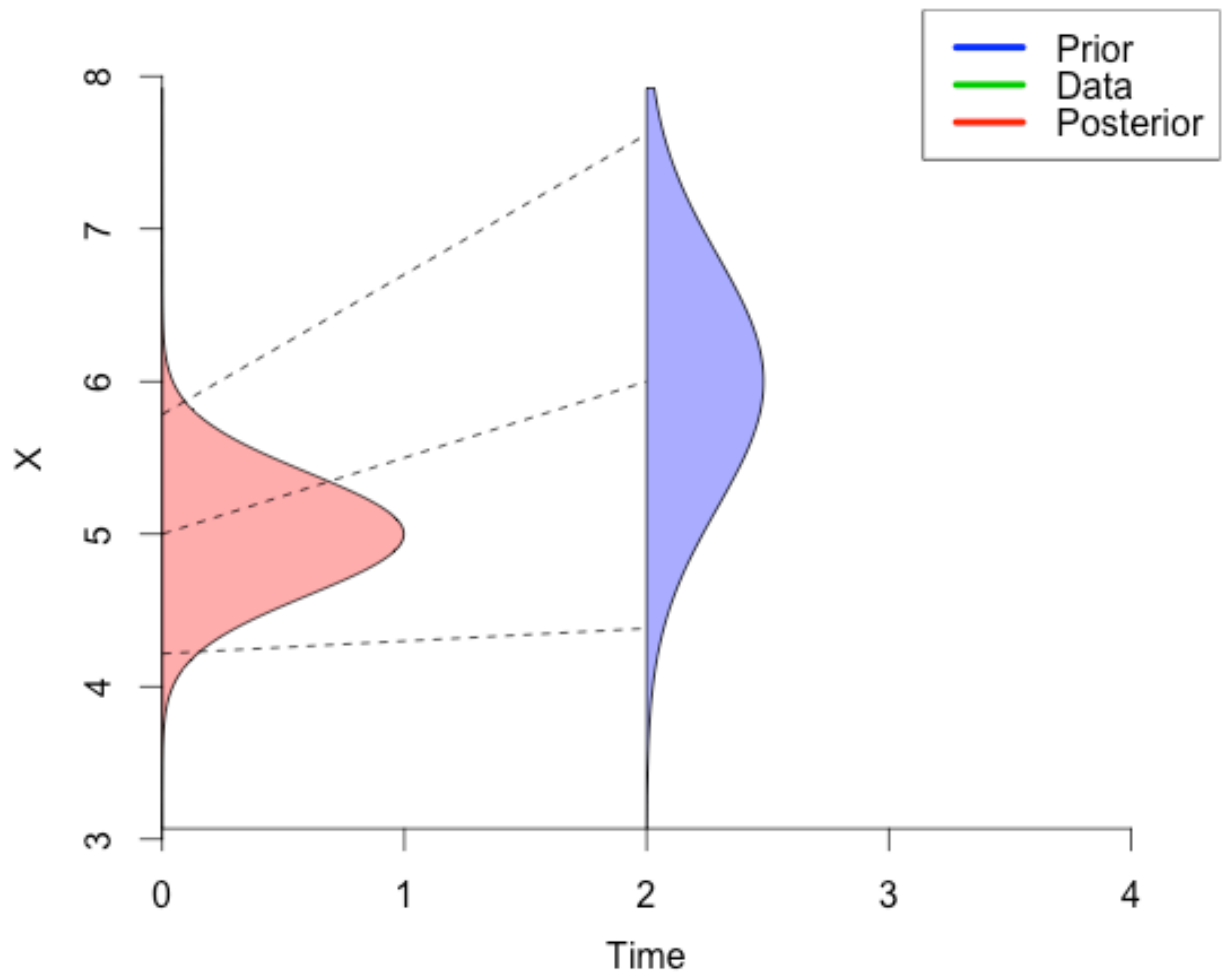
Sensitivity Analysis



Monte Carlo Sensitivity

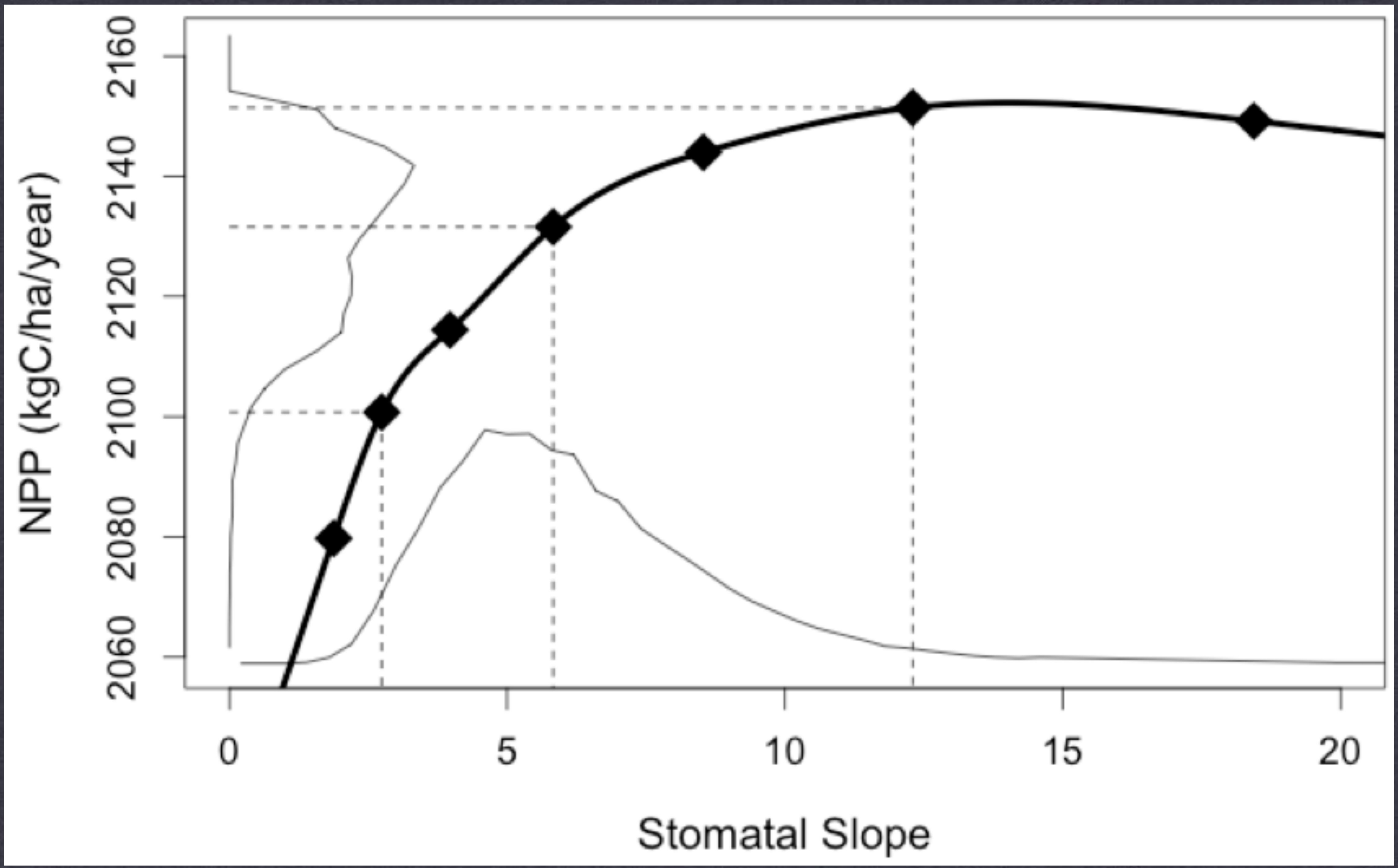


Free if you do MC uncertainty propagation or MCMC



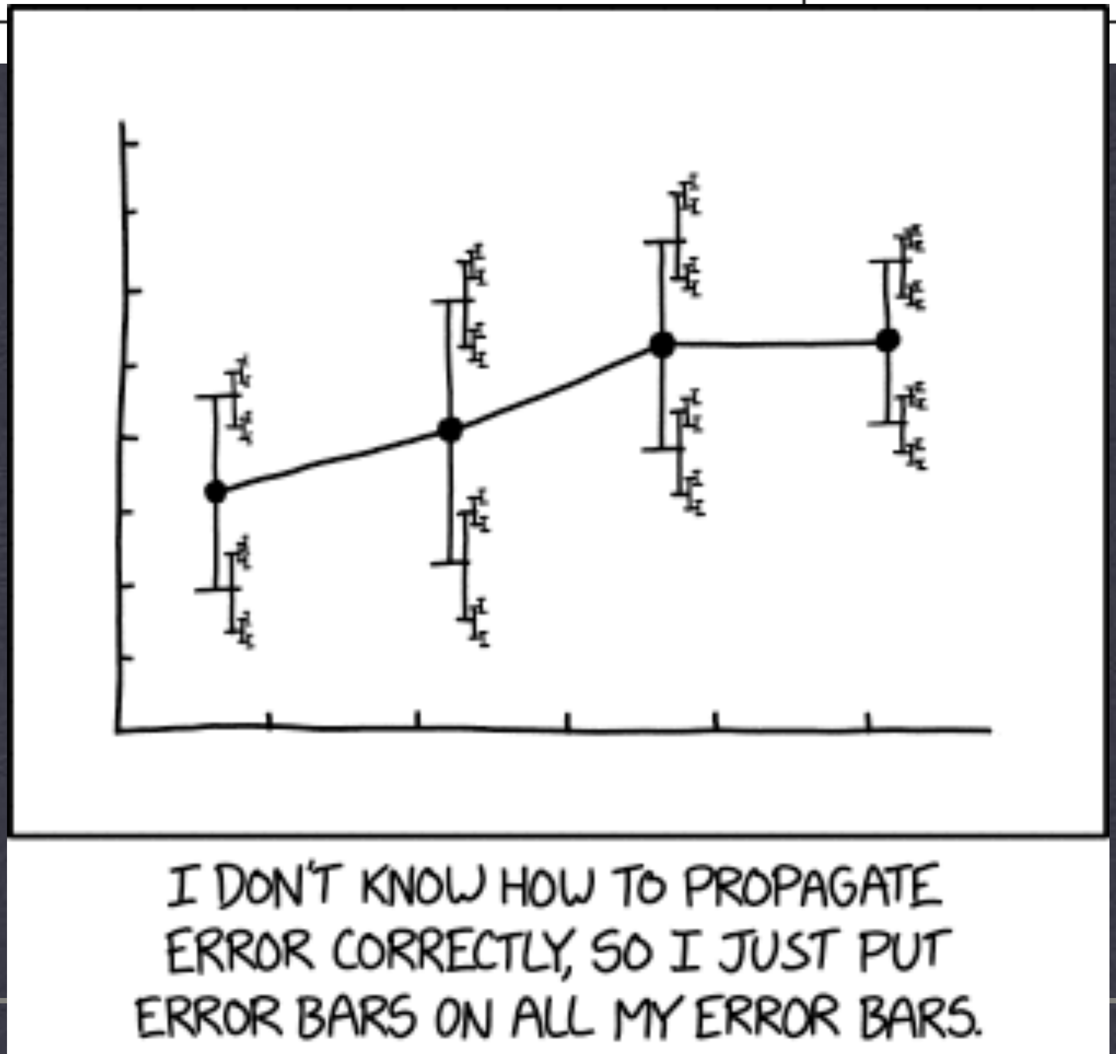
Prior
Data
Posterior
X
8
7
6
5
4
3
0
1
2
3
4
Time

UNCERTAINTY PROPAGATION



UNCERTAINTY PROPAGATION

		Output	
		Distribution	Moments
Approach	Analytic	Variable Transform	Analytical Moments Taylor Series
	Numeric	Monte Carlo	Ensemble



Typo in book
(puts Taylor in with Ensemble)

VARIABLE TRANSFORM

$$P_Y[y] = P_\theta[f^{-1}(y)] \cdot \left| \frac{d\, f^{-1}(y)}{dy} \right|$$

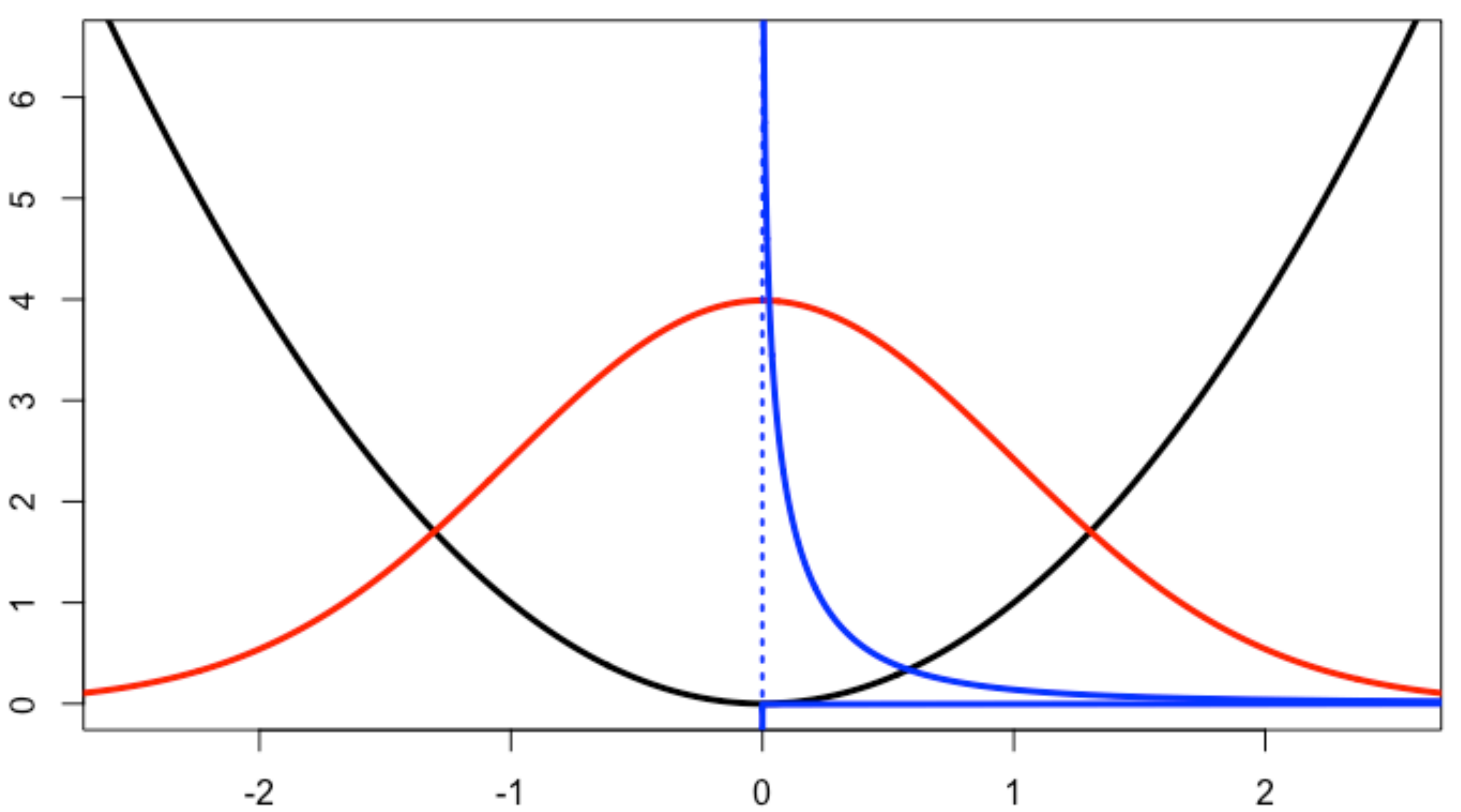
$X \sim N(0,1)$

$Y = X^2$

$Y \sim \chi^2$

$E[f(\bar{X})] = 0$

$E[\overline{f(X)}] = 1$



Analytical Moments

$$Var(aX) = a^2 Var(X)$$

$$Var(X+b) = Var(X)$$

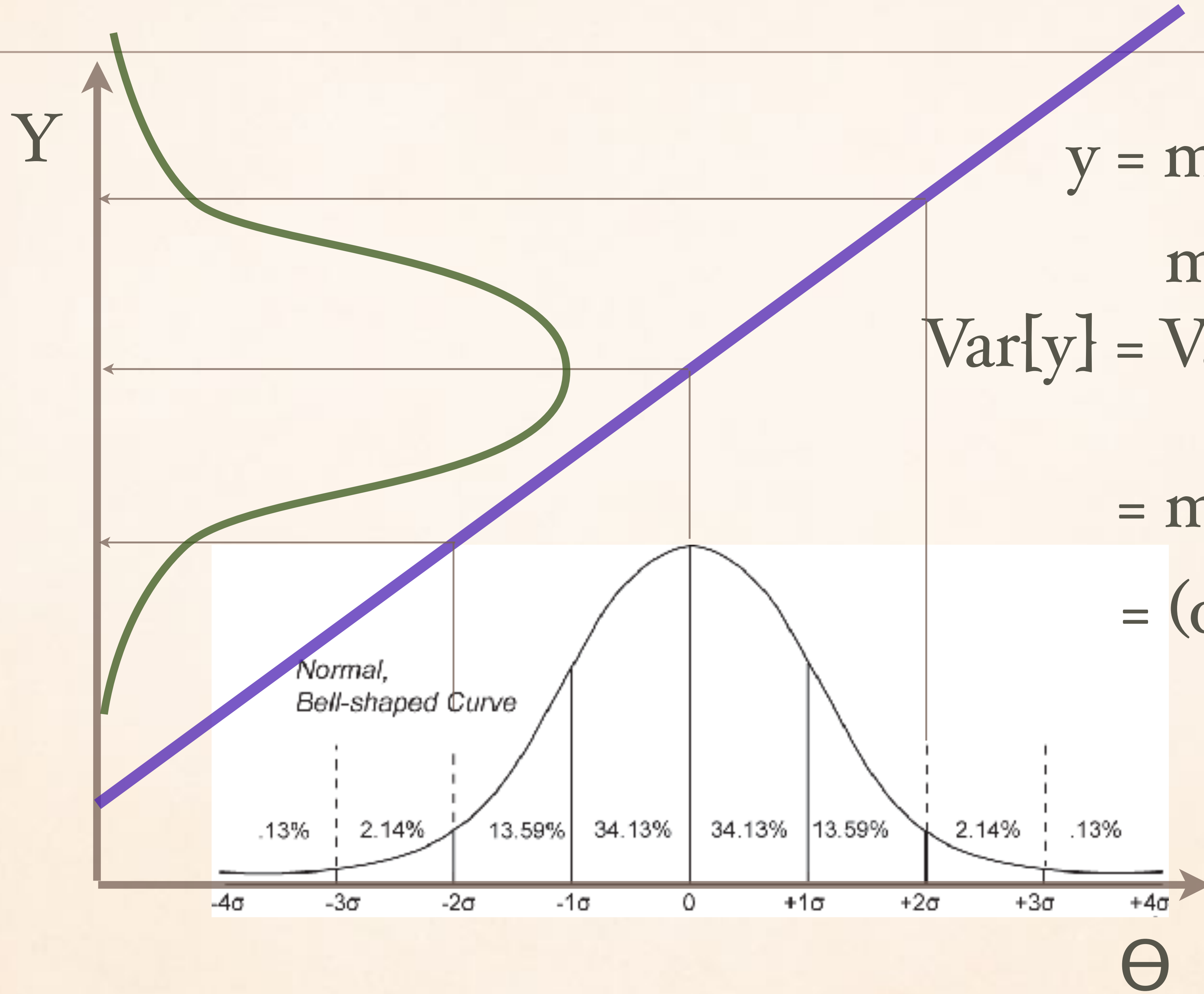
$$Var(X+Y) = Var(X) + Var(Y) + 2\mathrm{Cov}(X,Y)$$

$$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2\mathrm{abCov}(X,Y)$$

$$Var\left(\sum X\right) = \sum Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

$$Var(X) = Var(E[X|Y]) + E[Var(X|Y)]$$

REL'N TO SENSITIVITY



$$y = m\Theta + b$$

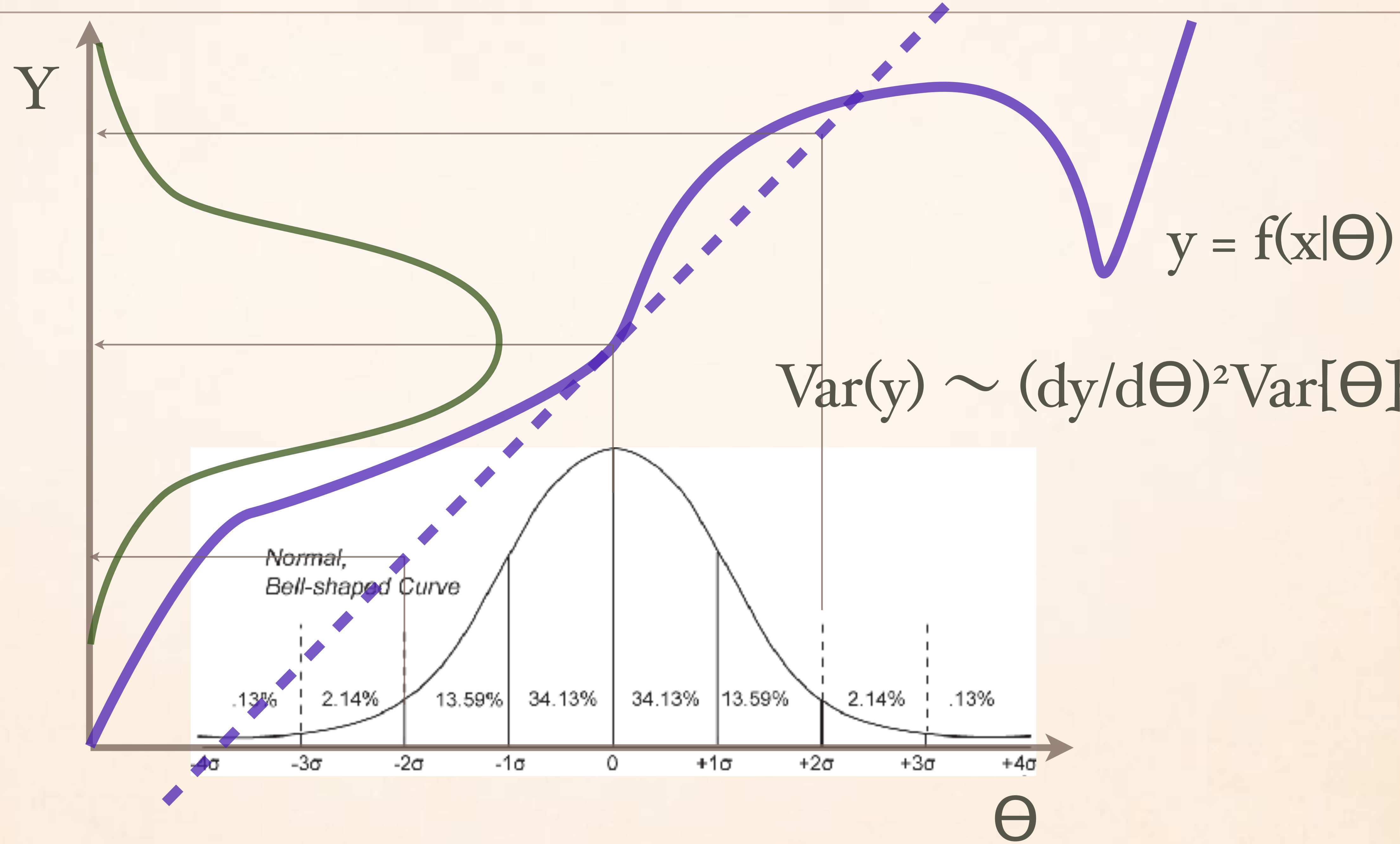
$$m = dy/d\Theta$$

$$\mathrm{Var}\{y\} = \mathrm{Var}[m\Theta + b\}$$

$$= m^2\mathrm{Var}[\Theta]$$

$$= (dy/d\Theta)^2\mathrm{Var}[\Theta\}$$

LINEAR APPROX



TAYLOR SERIES

$$Var[f(x|\theta)] \approx Var\left[f(x|\bar{\theta}) + \frac{\frac{df}{d\theta}(x|\bar{\theta})}{1!}(\theta - \bar{\theta}) + \ldots \right]$$

$$var[f(x)] \approx \left(\frac{\partial f}{\partial \theta_i} \right)^2 var[\theta]$$

TAYLOR SERIES

$$Var[f(x|\theta)] \approx Var\left[f(x|\bar{\theta}) + \frac{\frac{df}{d\theta}(x|\bar{\theta})}{1!}(\theta - \bar{\theta}) + \ldots \right]$$

$$var[f(x)] \approx \sum \left(\frac{\partial f}{\partial \theta_i} \right)^2 var[\theta_i] +$$

$$\sum_{i \neq j} \left(\frac{\partial f}{\partial \theta_i} \right) \left(\frac{\partial f}{\partial \theta_j} \right) cov[\theta_i, \theta_j]$$

$$Y_{t+1} = f(Y_t, X_t|\theta) + \varepsilon$$

$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{df}{dY}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC \\ uncert}} + \underbrace{\left(\frac{df}{dX}\right)^2}_{\substack{driver \\ sens}} \underbrace{Var[X]}_{\substack{driver \\ uncert}} + \underbrace{\left(\frac{df}{d\theta}\right)^2}_{\substack{param \\ sens}} \underbrace{Var[\theta]}_{\substack{param \\ uncert}} + \underbrace{Var[\varepsilon]}_{\substack{process \\ error}}$$

COV & SCALING

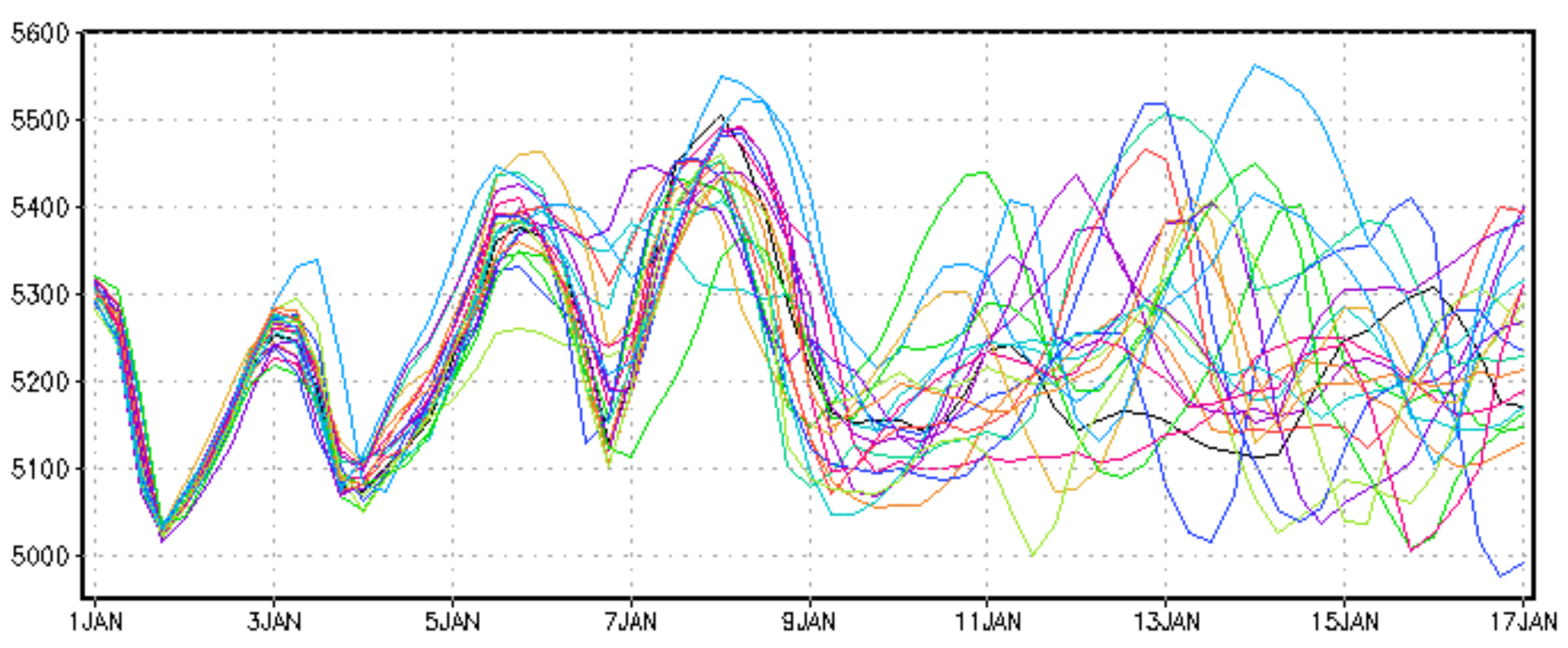
- **Scaling very dependent on spatial and temporal auto- & cross-correlation**

$$\sum\sum\frac{\partial f}{\partial X_i}\frac{\partial f}{\partial X_j}COV[X_i,X_j]$$

UNCERTAINTY PROPAGATION

		Output	
		Distribution	Moments
Approach	Analytic	Variable Transform	Analytical Moments Taylor Series
	Numeric	Monte Carlo	Ensemble

Numerical Approximation

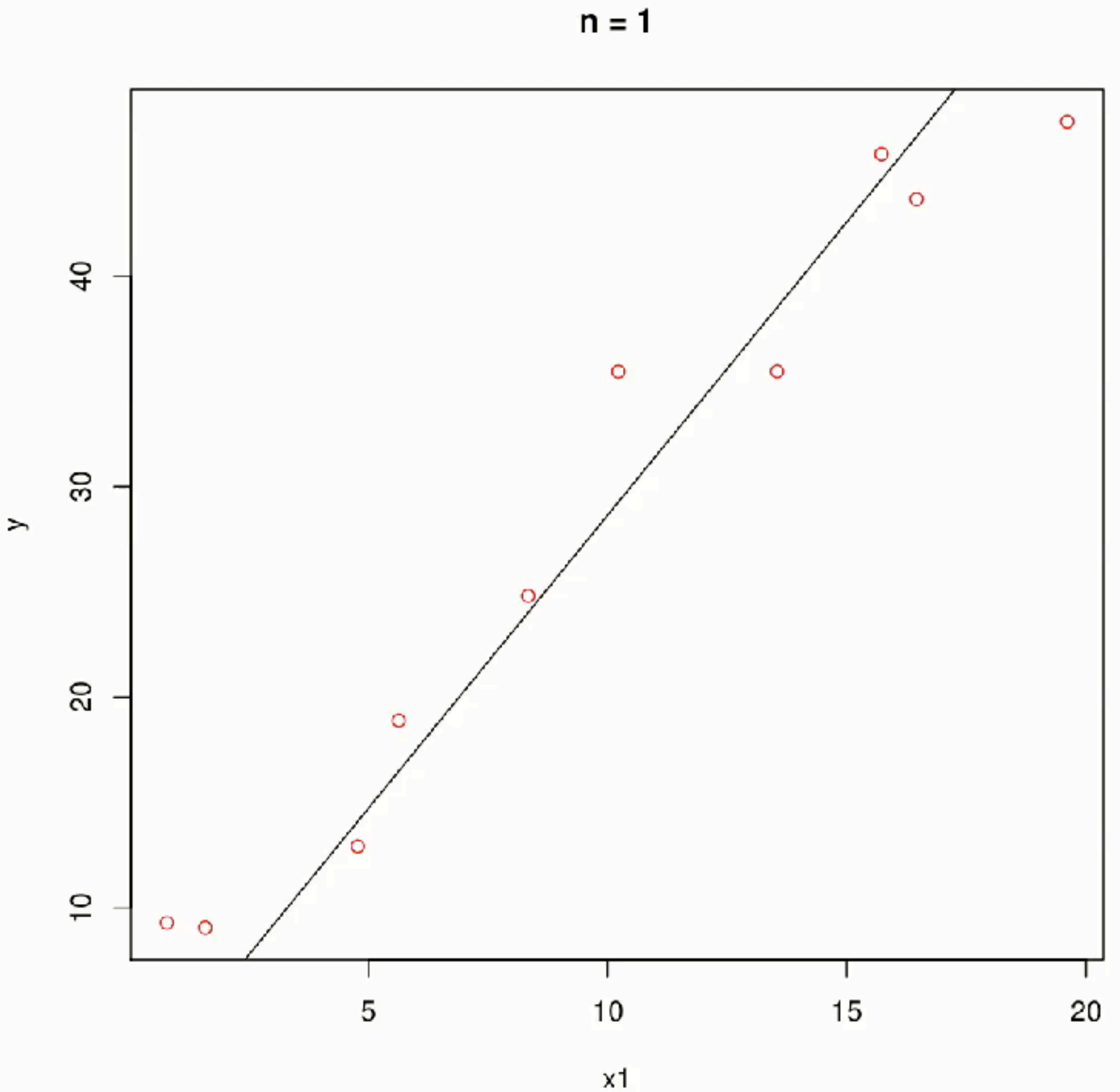


- Monte Carlo Simulation --> Distribution
- Ensemble Analysis --> Moments

MONTE CARLO UNCERTAINTY

- for (i in 1:n)
  - draw random values from distributions
  - run model
  - save results
- summarize distributions

Already have this from MCMC!


n = 1
y
x1
10
20
30
40
5
10
15
20

ENSEMBLE UNCERTAINTY

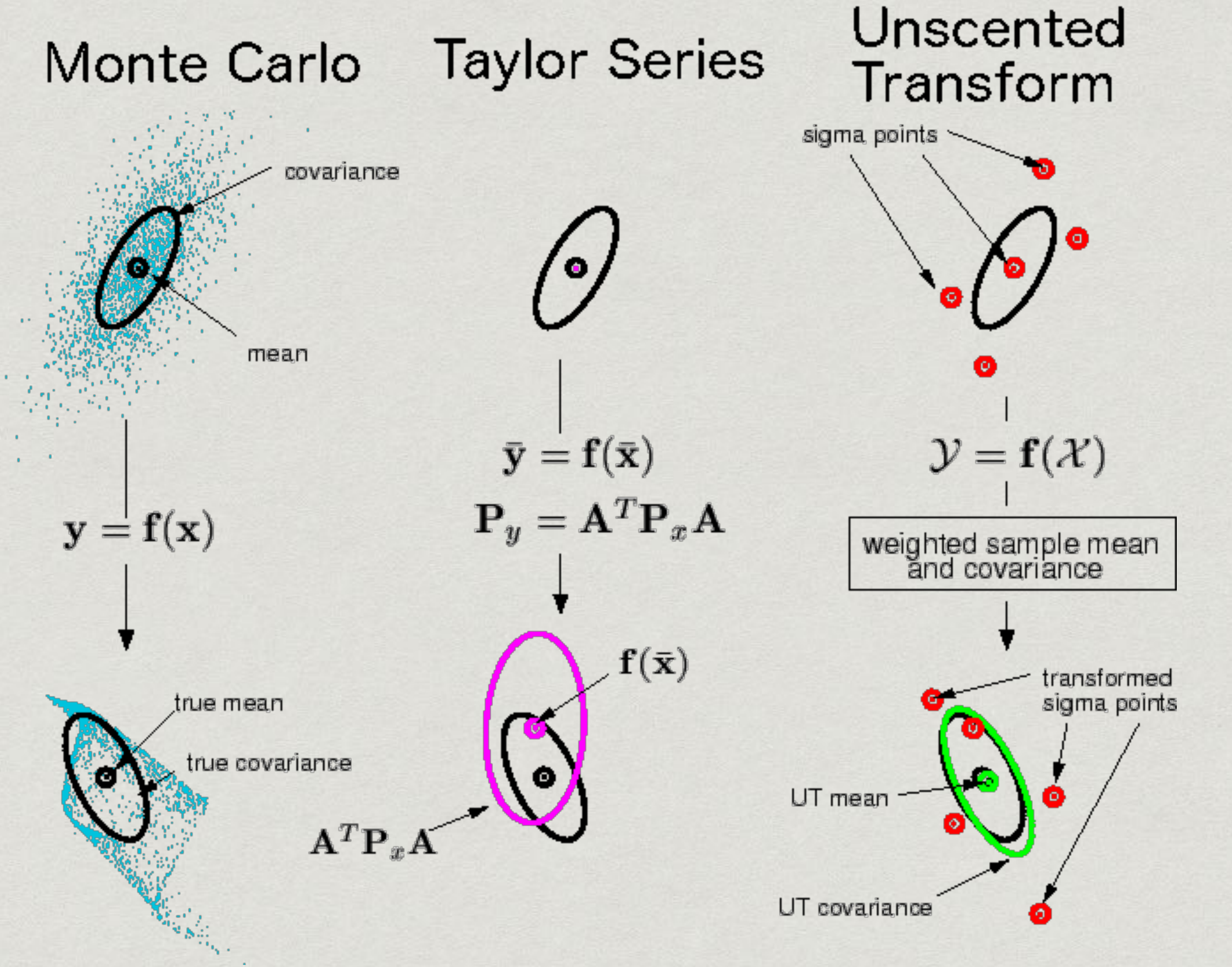
- for (i in 1:n) ← Requires smaller N to estimate moments than to approximate full PDF
  - draw random values from distributions
  - run model
  - save results

Already have this from MCMC!

- **Fit PDF to results**
- **Use PDF for intervals, etc.**

INPUT DISTRIBUTION

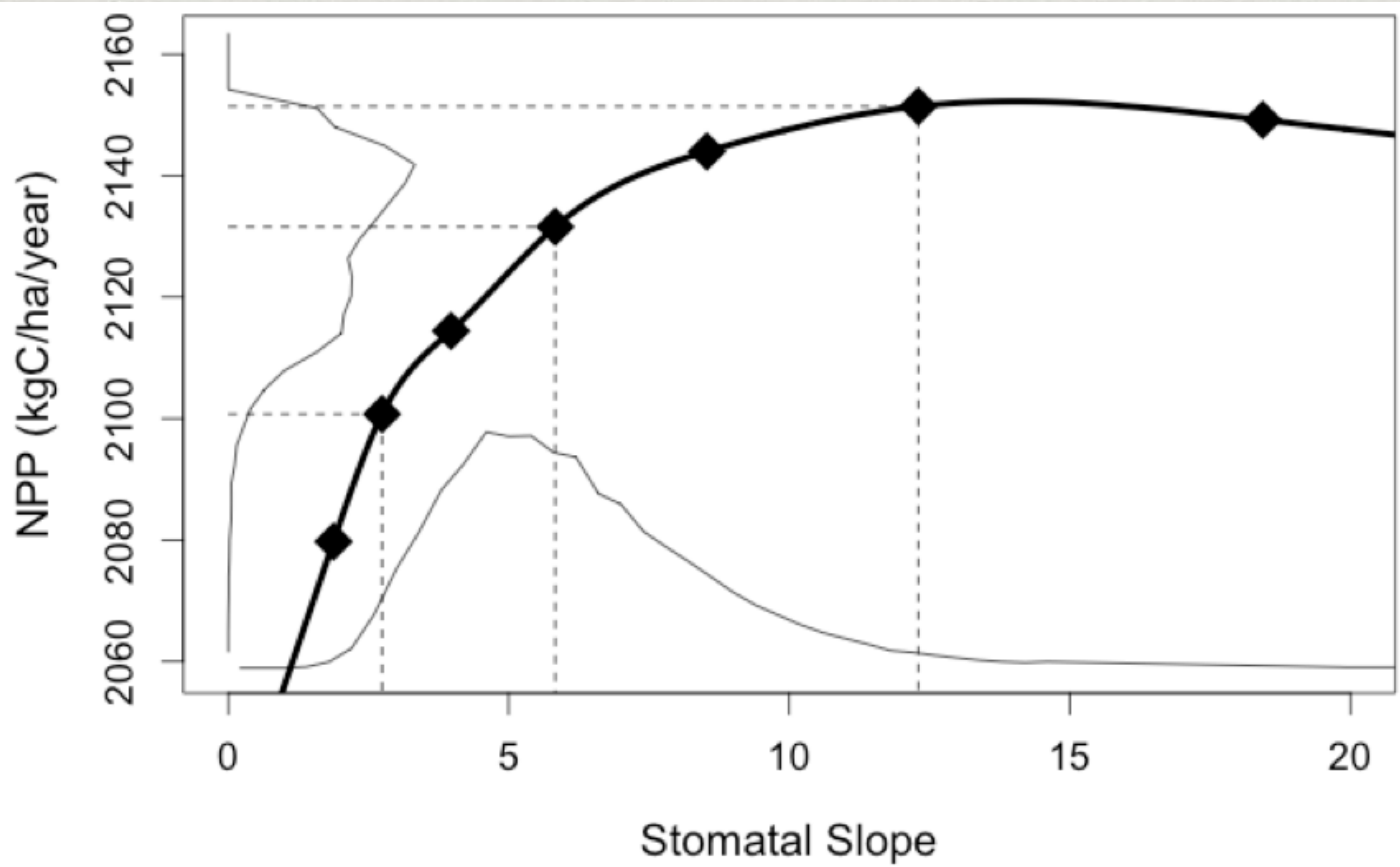
OUTPUT DIST.

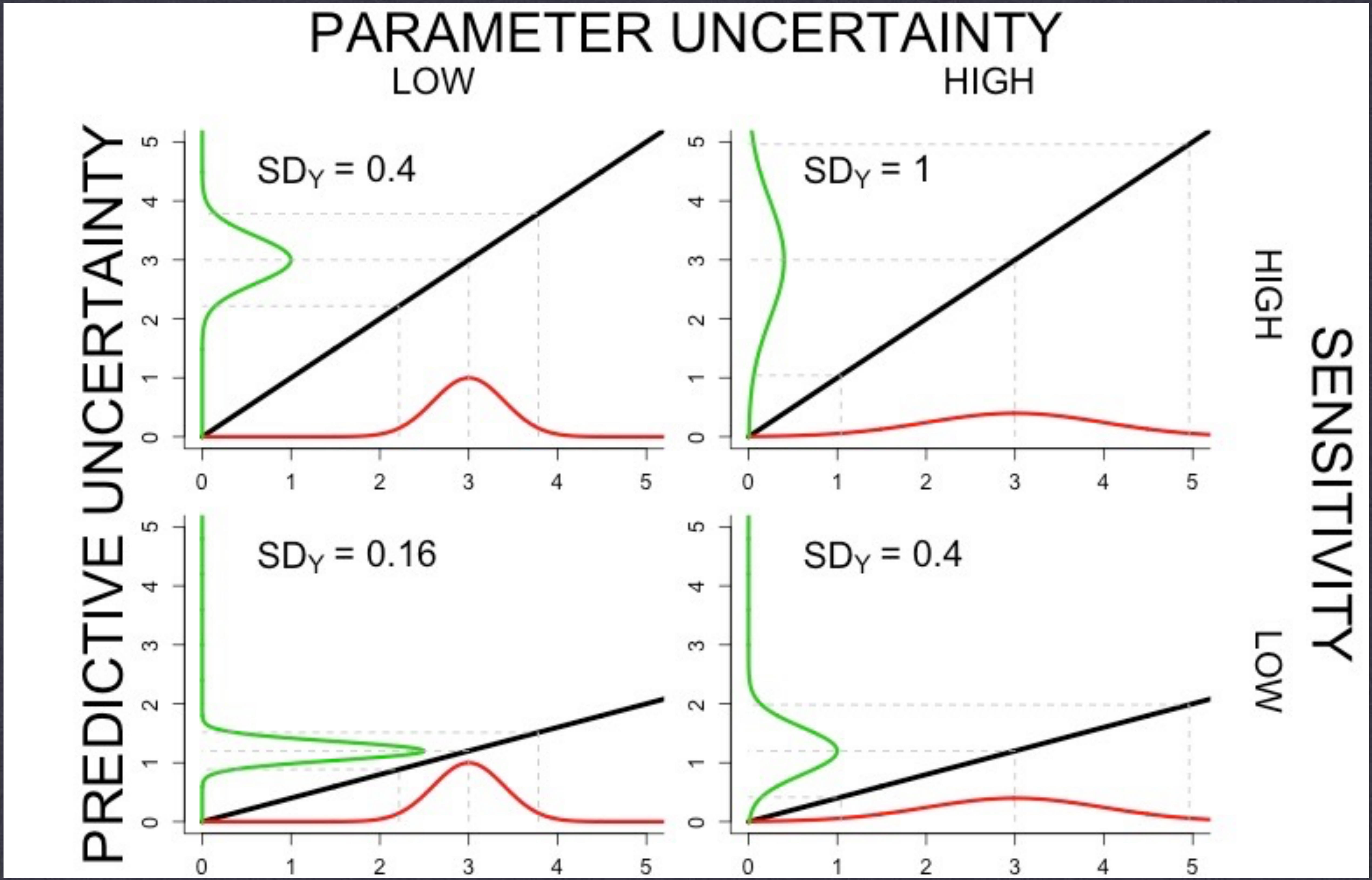


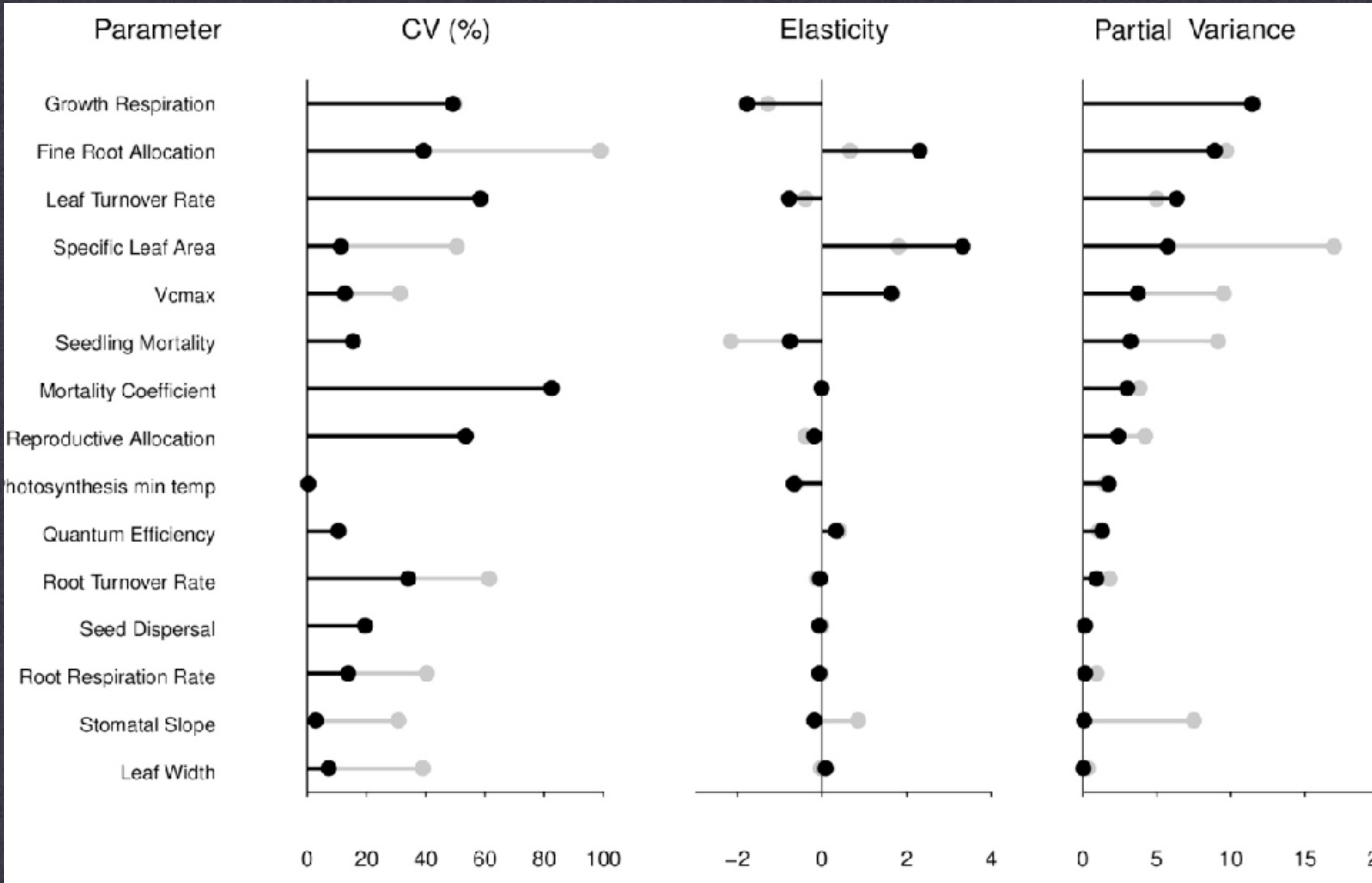
UNCERTAINTY PROPAGATION

Approach		Output	
		Distribution	Moments
	Analytic	Variable Transform	Analytical Moments Taylor Series
	Numeric	Monte Carlo	Ensemble

Uncertainty Analysis



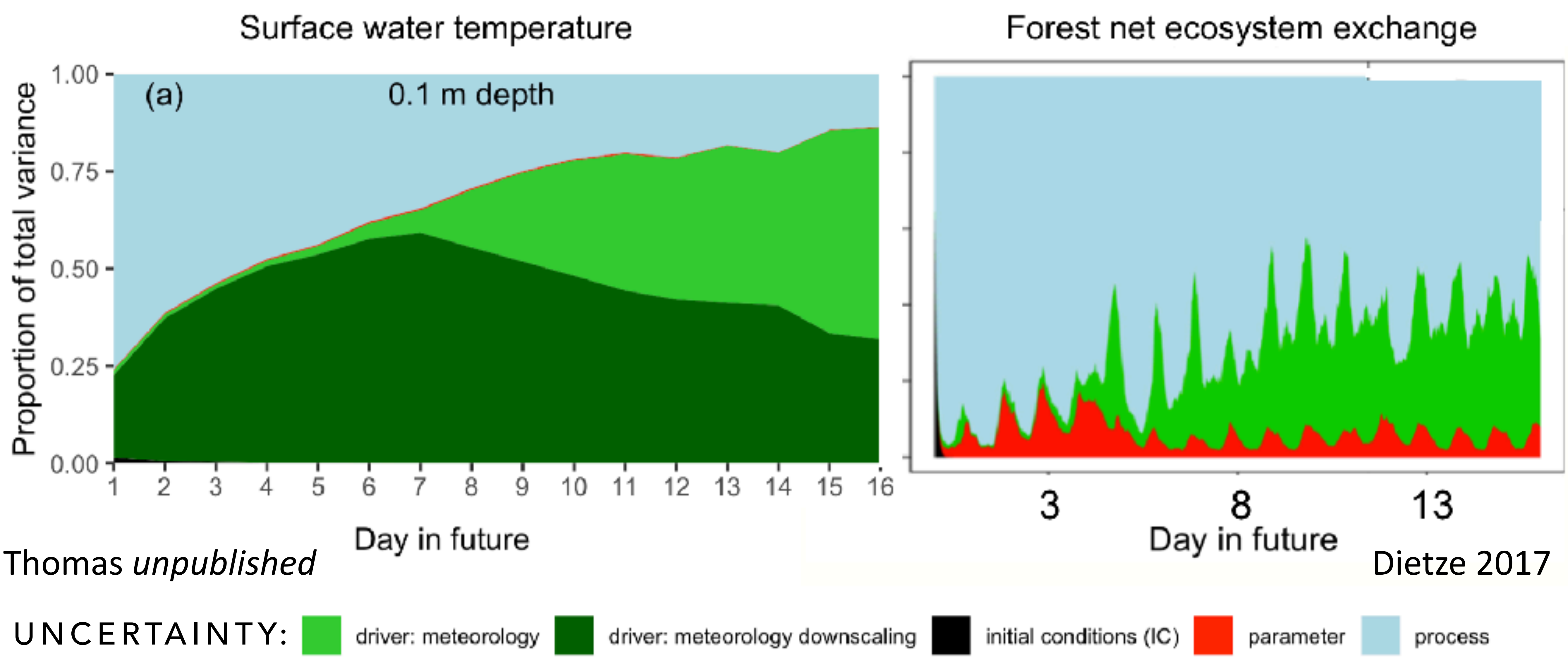

PARAMETER UNCERTAINTY
LOW
HIGH
PREDICTIVE UNCERTAINTY
SENSITIVITY
HIGH
LOW
$SD_Y = 0.4$
$SD_Y = 1$
$SD_Y = 0.16$
$SD_Y = 0.4$



VARIANCE DECOMPOSITION

SWITCHGRASS YIELD, CENTRAL ILLINOIS

How do the drivers of forecast uncertainty vary across ecological system?



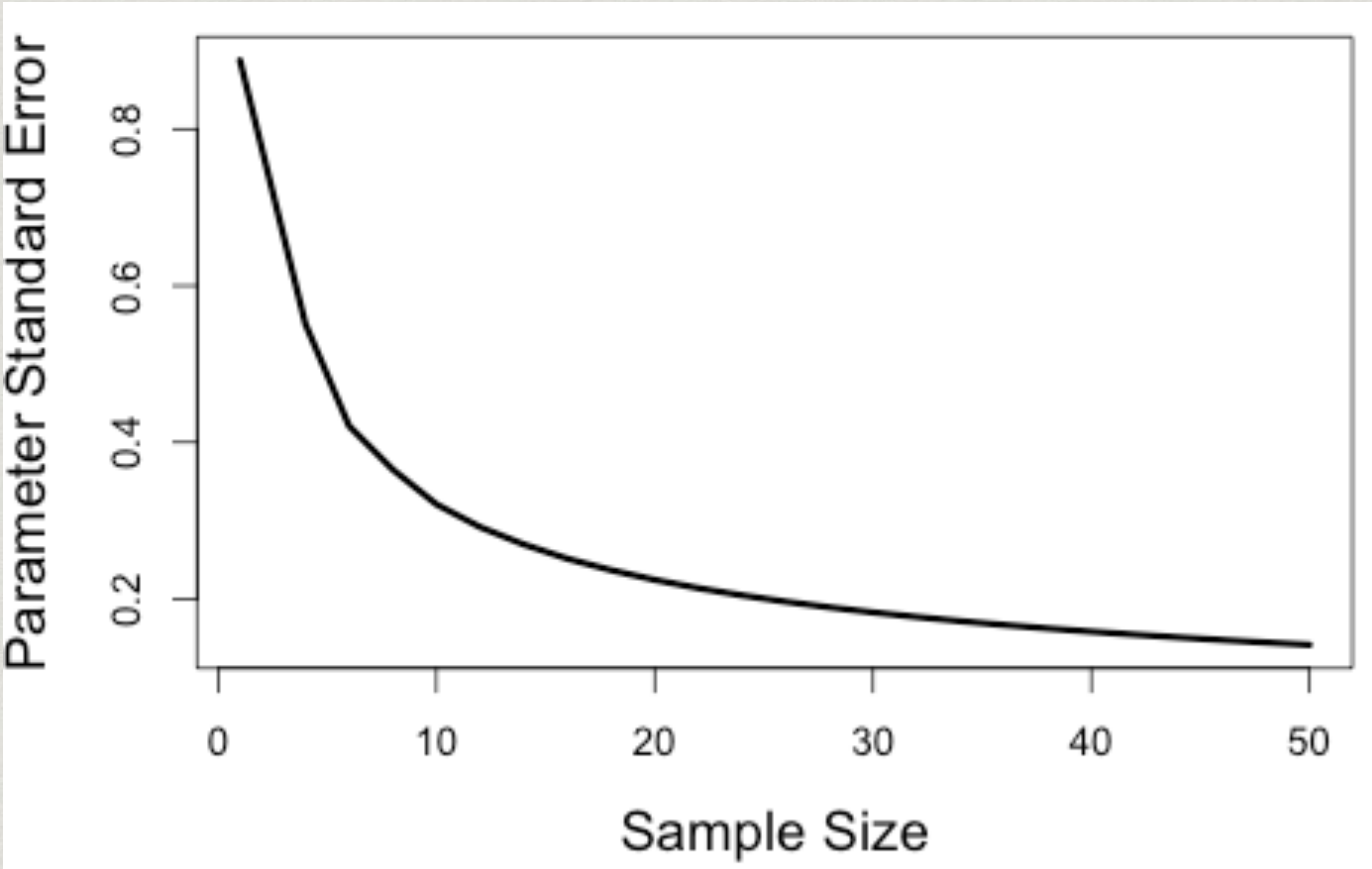
Thomas *unpublished*

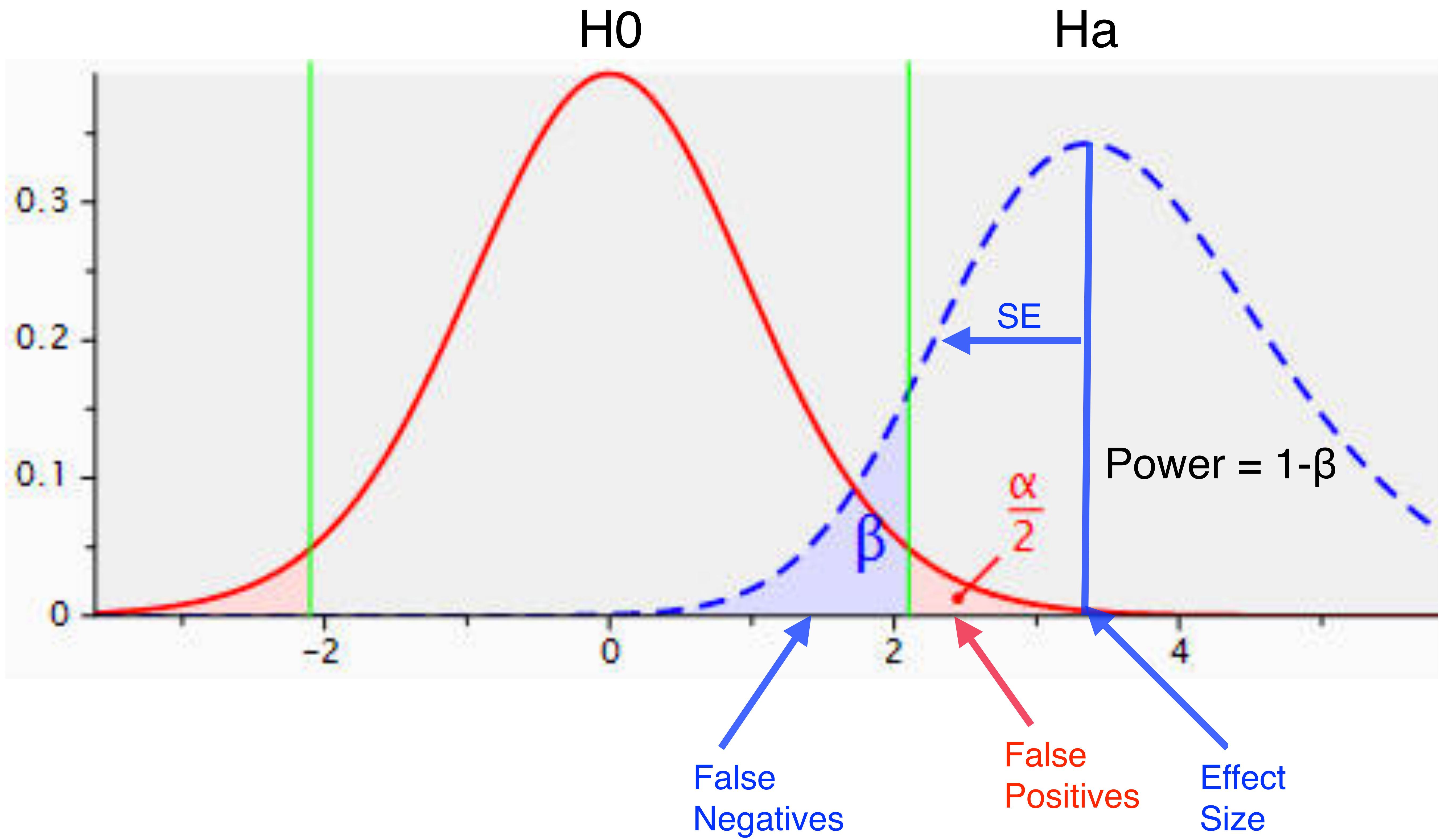
Dietze 2017

Tools for model-data feedbacks

- **Power analysis**
  - Sample size needed to detect an effect size
  - Minimum effect size detectable given a size
- **Observational design**
  - What do I need to measure?
  - Where should I collect new data?
  - How do I gain new info most efficiently?

SE ∝ 1/sqrt(n)


Parameter Standard Error
0.2
0.4
0.6
0.8
0
10
20
30
40
50
Sample Size

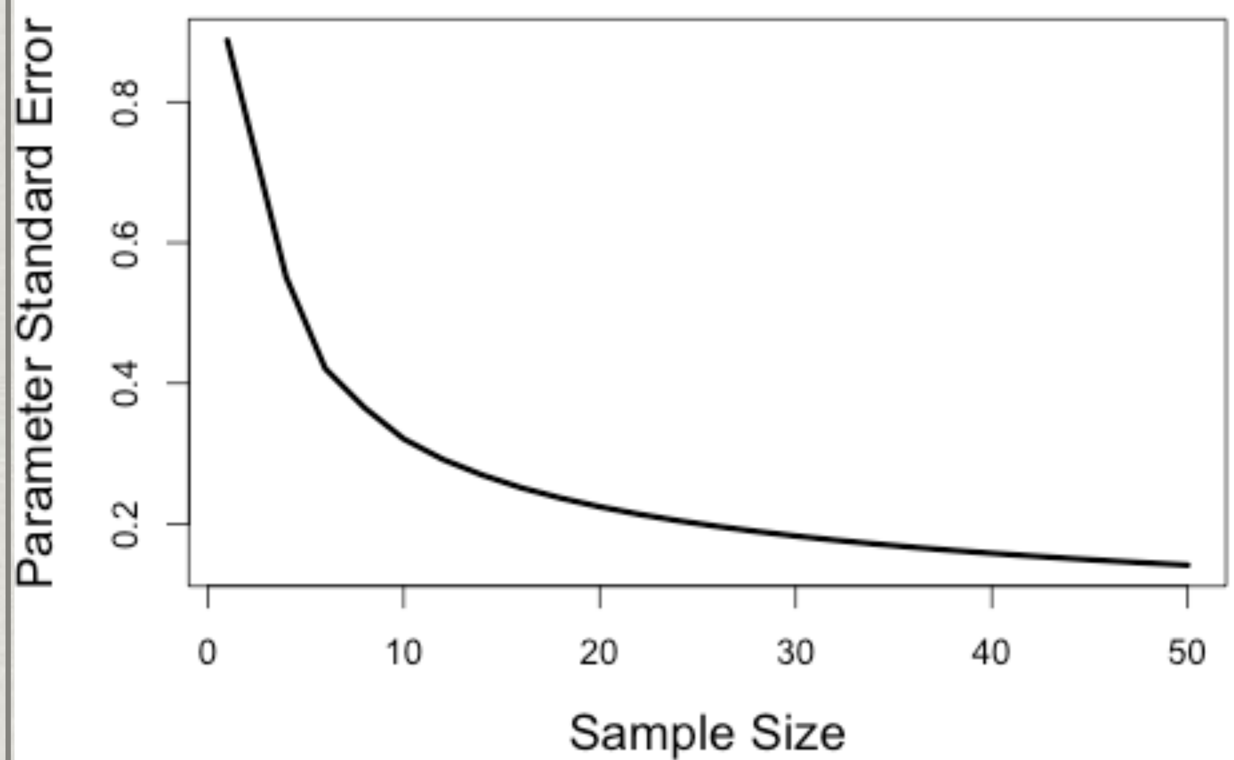


Power = f(effect size, SE)

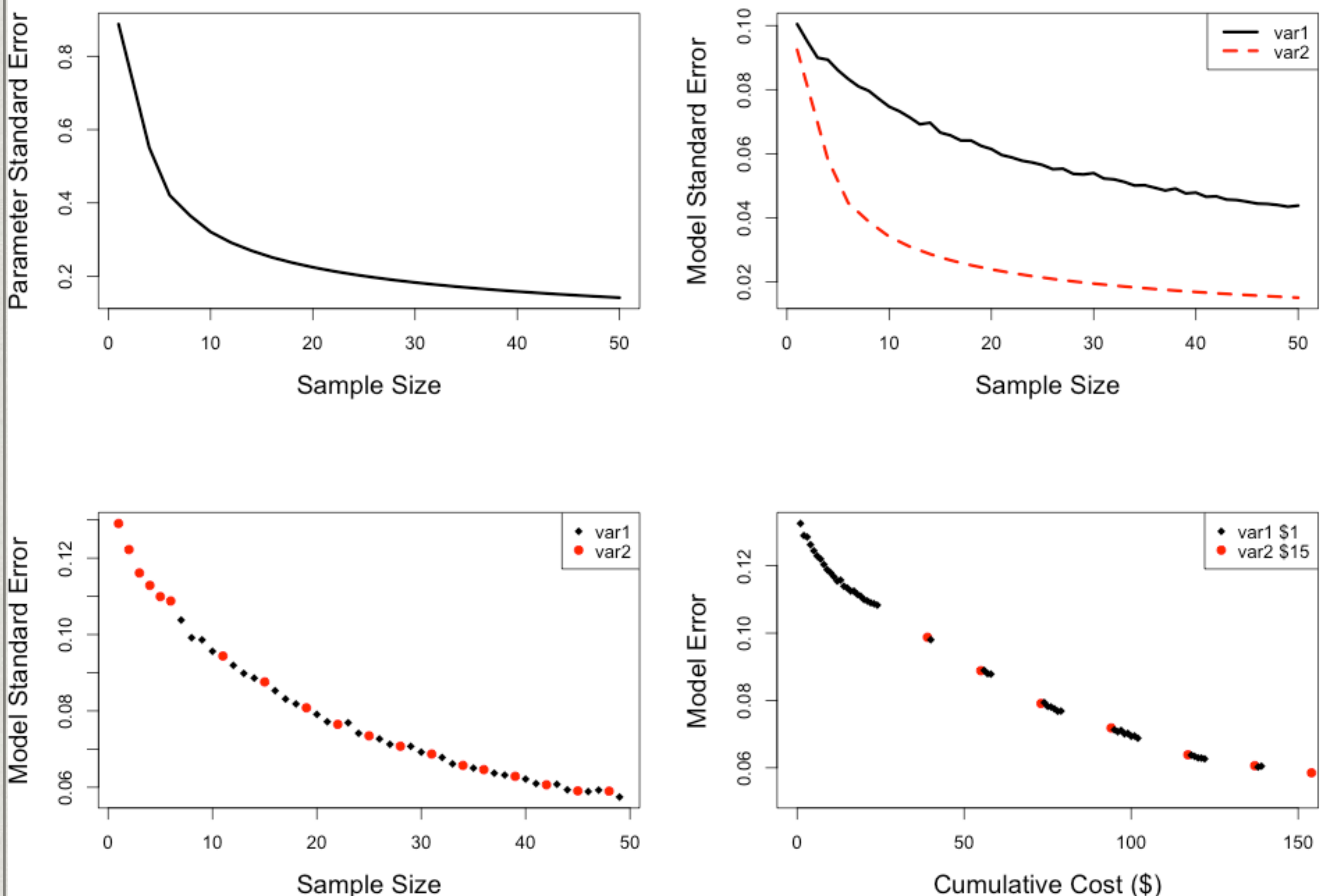
Pseudo-data simulation

```
for(k in 1:M)
    Draw random data of size N
    Fit model
    Save Parameter(S) of interest
```

- Nonparameteric bootstrap: resample data
- Parameteric bootstrap: assume param, sim data
- Embed in overall loop over N or different effect sizes
- Summarize distribution



Parameter Standard Error
0.2
0.4
0.6
0.8
0
10
20
30
40
50
Sample Size



Parameter Standard Error
Sample Size
Model Standard Error
var1
var2
Sample Size
Model Standard Error
var1
var2
Sample Size
Model Error
var1 $1
var2 $15
Cumulative Cost ($)

Observing System Simulation Experiments

- Simulate "true" system
- Simulate pseudo-observations
- Assimilate pseudo-observations
- Assess impact on estimates

- **Augment an existing network**
  - **Additional locations**
  - **New Sensors**
- **Common in Weather, Remote Sensing, Oceanography**