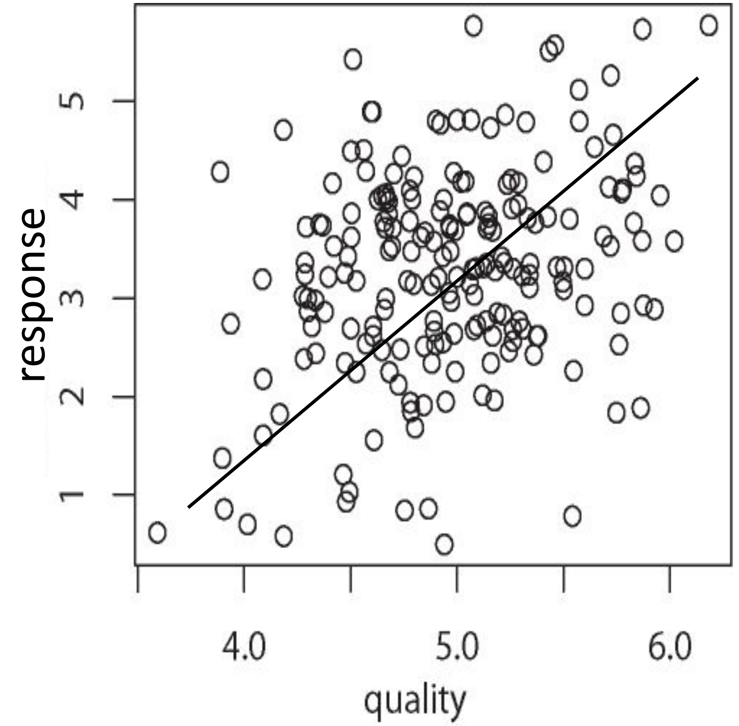
Characterizing Uncertainty

Linear Model

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$

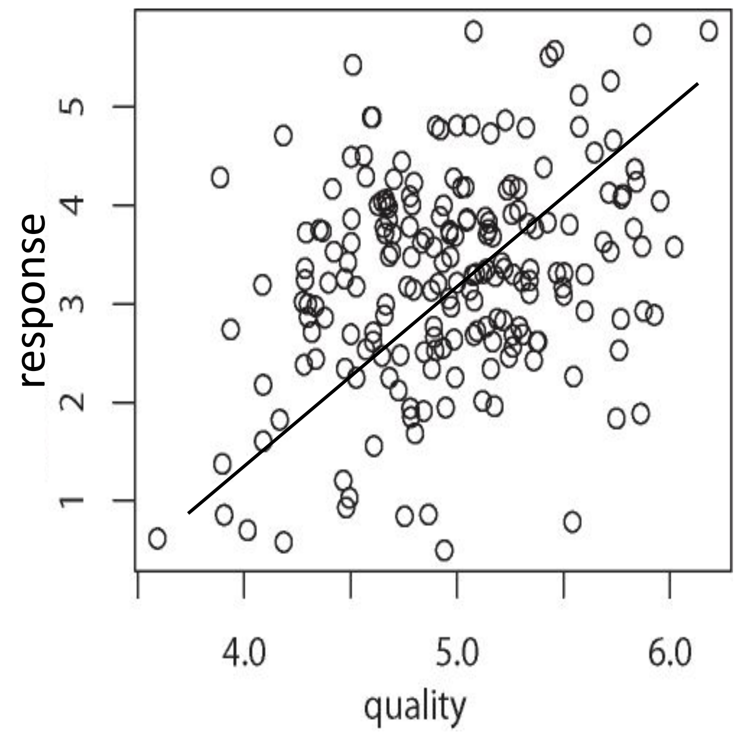


Linear Model

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent

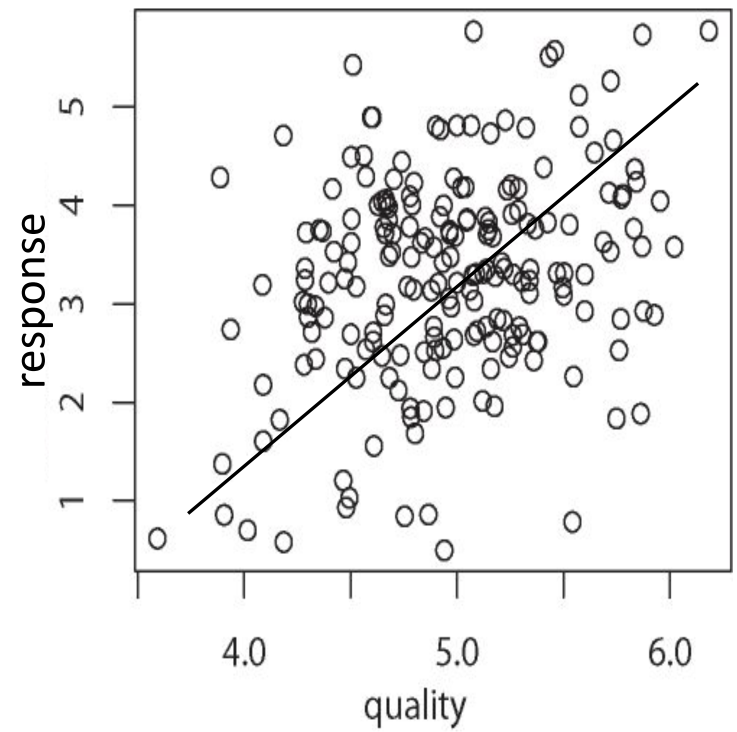


Linear Model

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent
- Homoskedasticity

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$

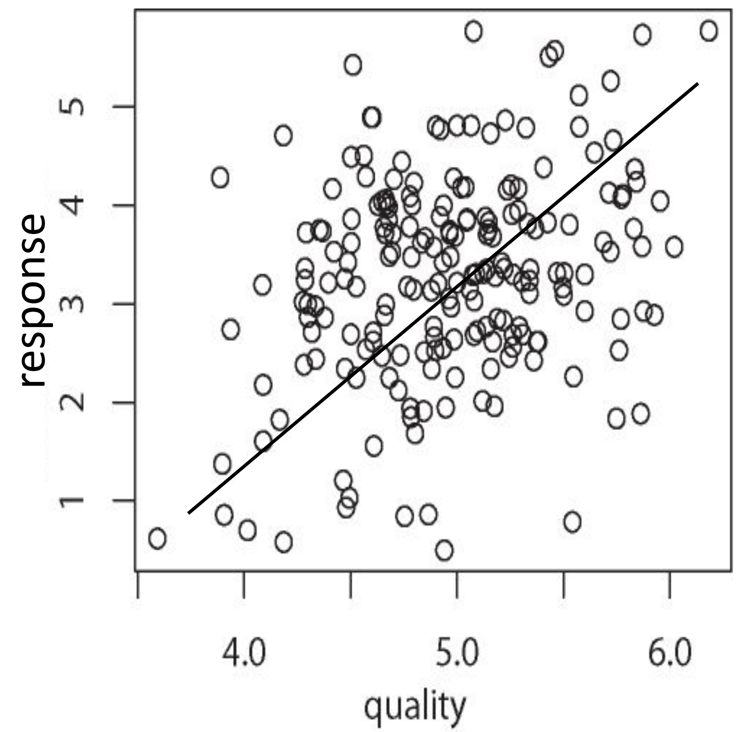


Linear Model

Classic Assumptions

- Error in Y is measurement error
- Normally distributed error
- Observations are independent
- Homoskedasticity
- No error in X variables

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$



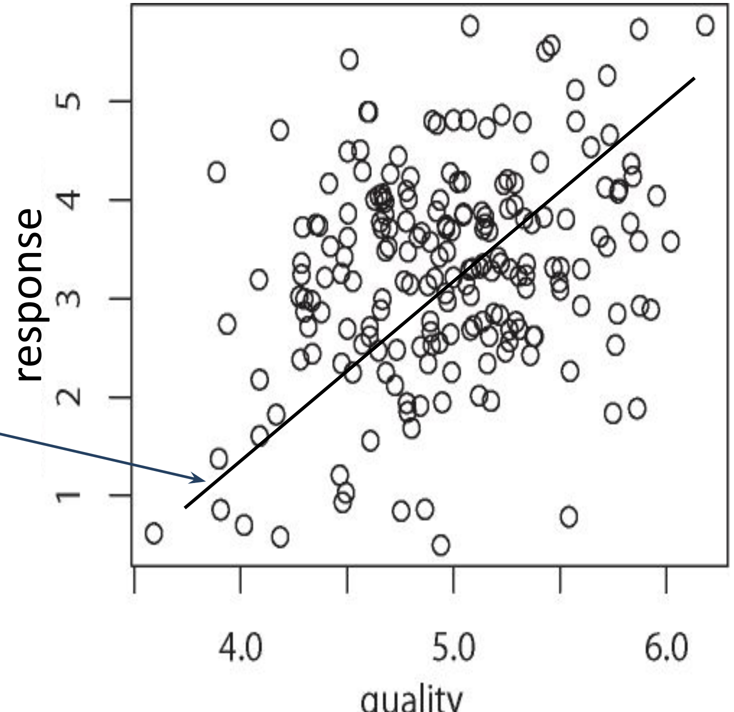
Linear Model

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$

Data Model Y

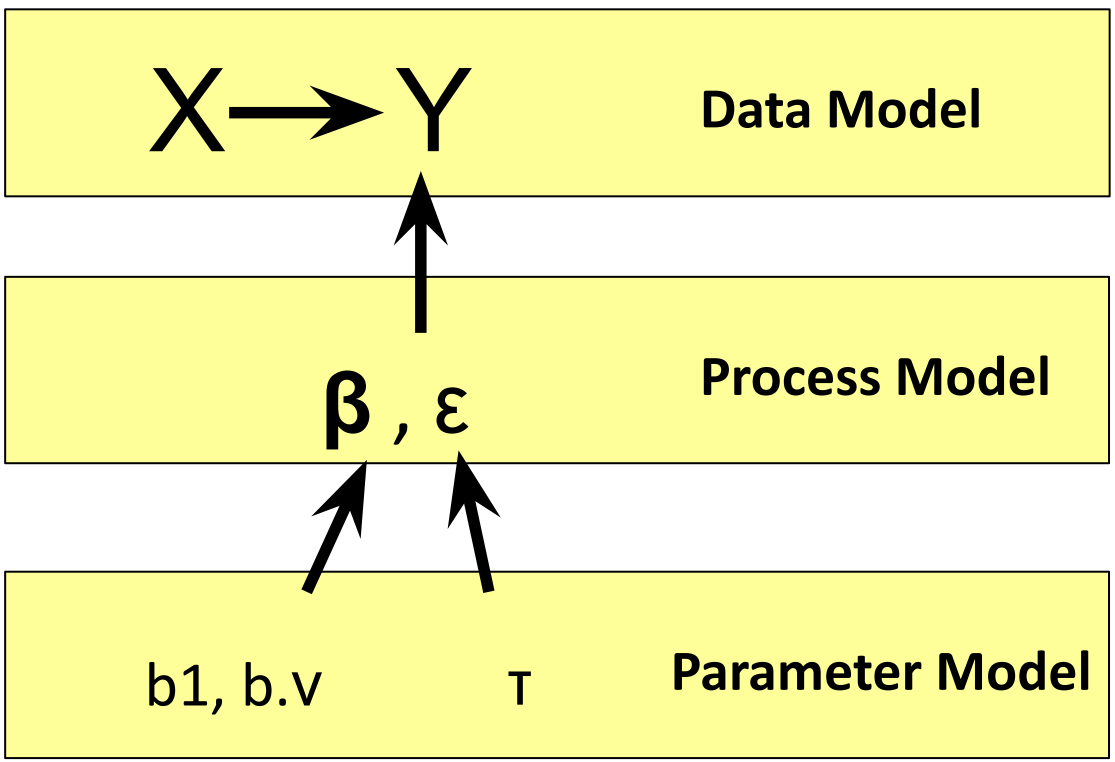
Process Model $\beta_0 + \beta_1(x_i)$

Parameters $\boldsymbol{\beta}$

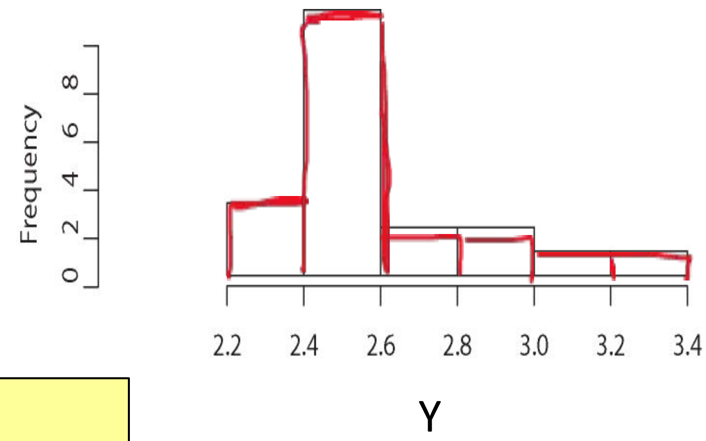


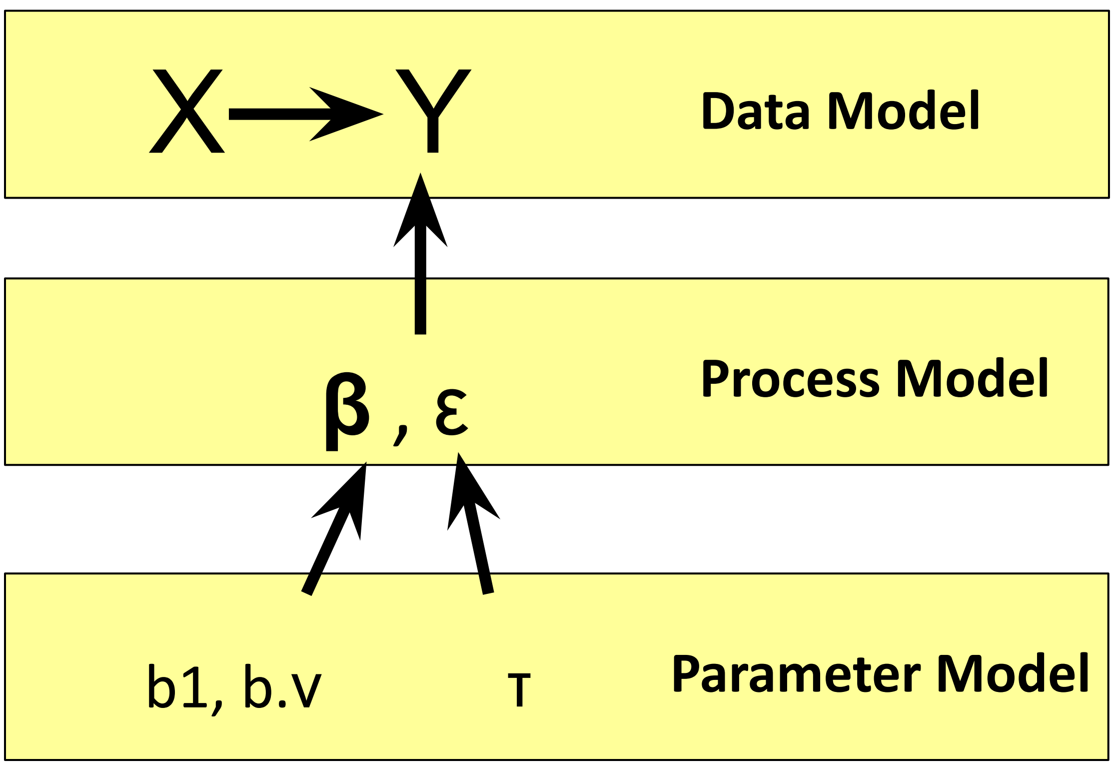
Linear Model – Graph Notation

$$y_i \sim \beta_0 + \beta(x_i) + \varepsilon_i$$



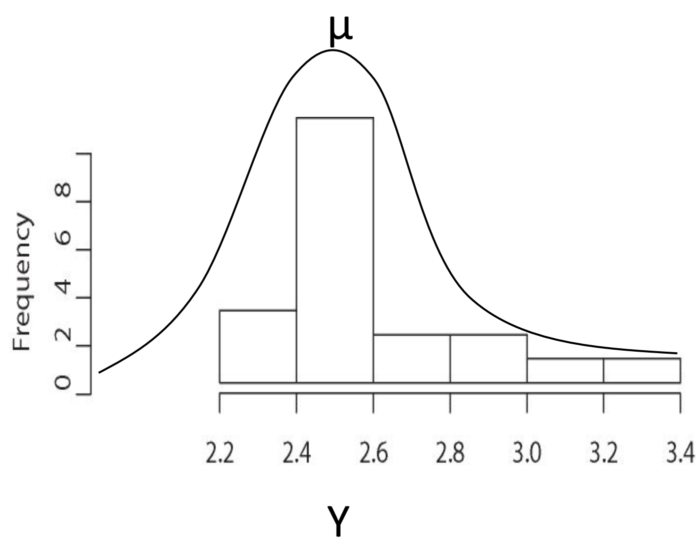
Linear Model –





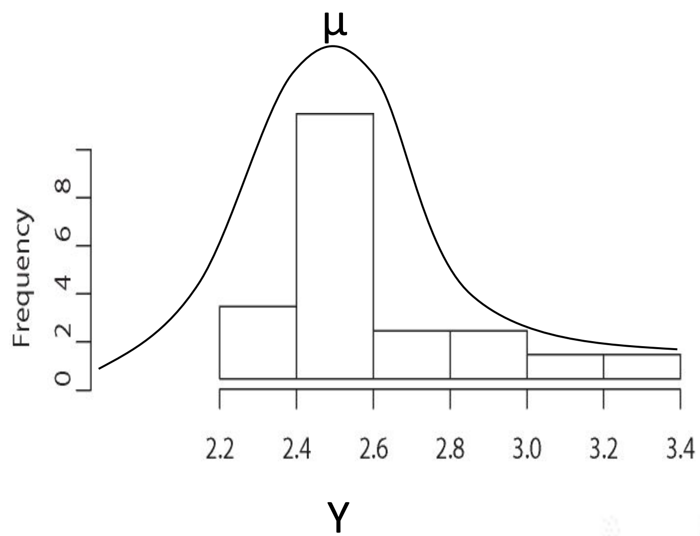
Data & Distributions

$y \sim N(\mu, \varepsilon)$



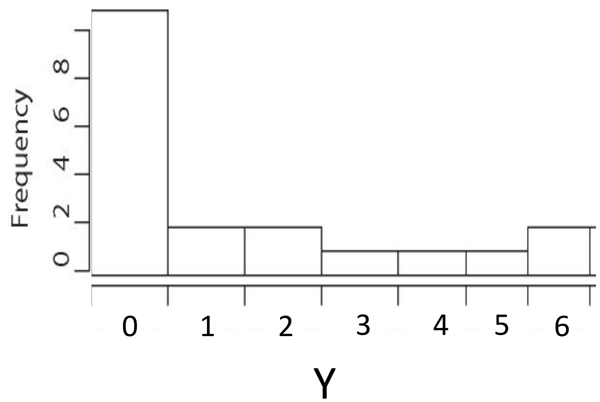
Expected relationship between data samples and the range of all possible data is described by a probability distribution.

$y \sim N(\mu, \varepsilon)$

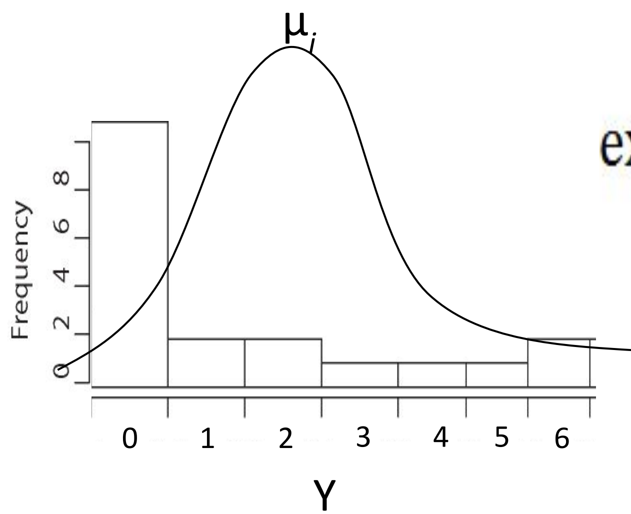


$$p(y|\mu) \propto \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$

Data & Distributions

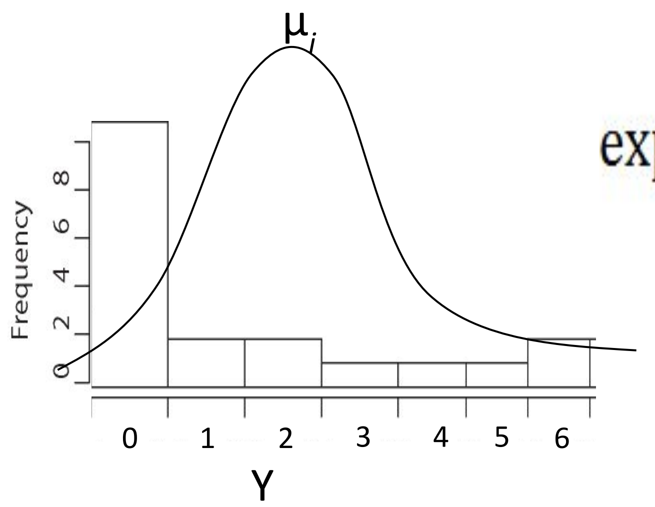


Data & Distributions

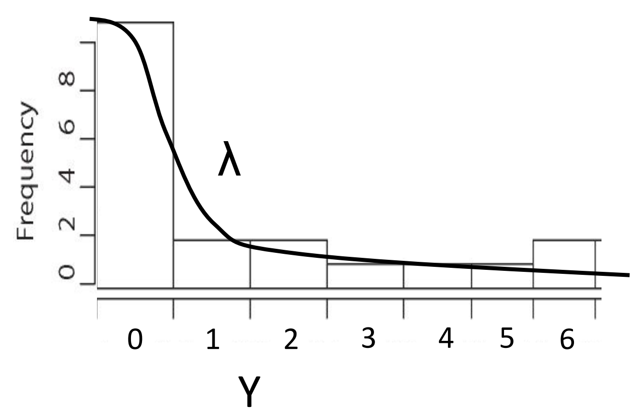


$$\exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$

Data & Distributions

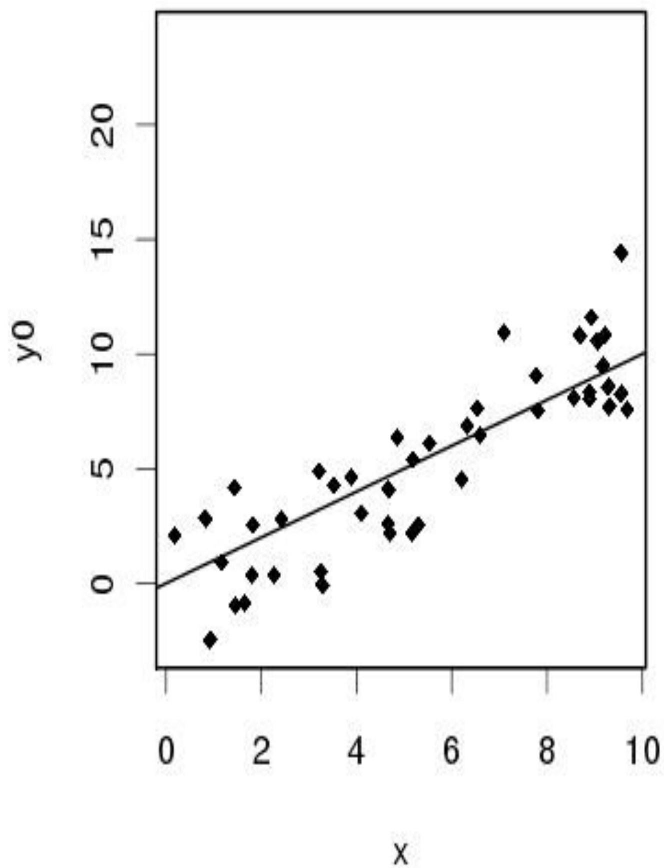


$$\exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$

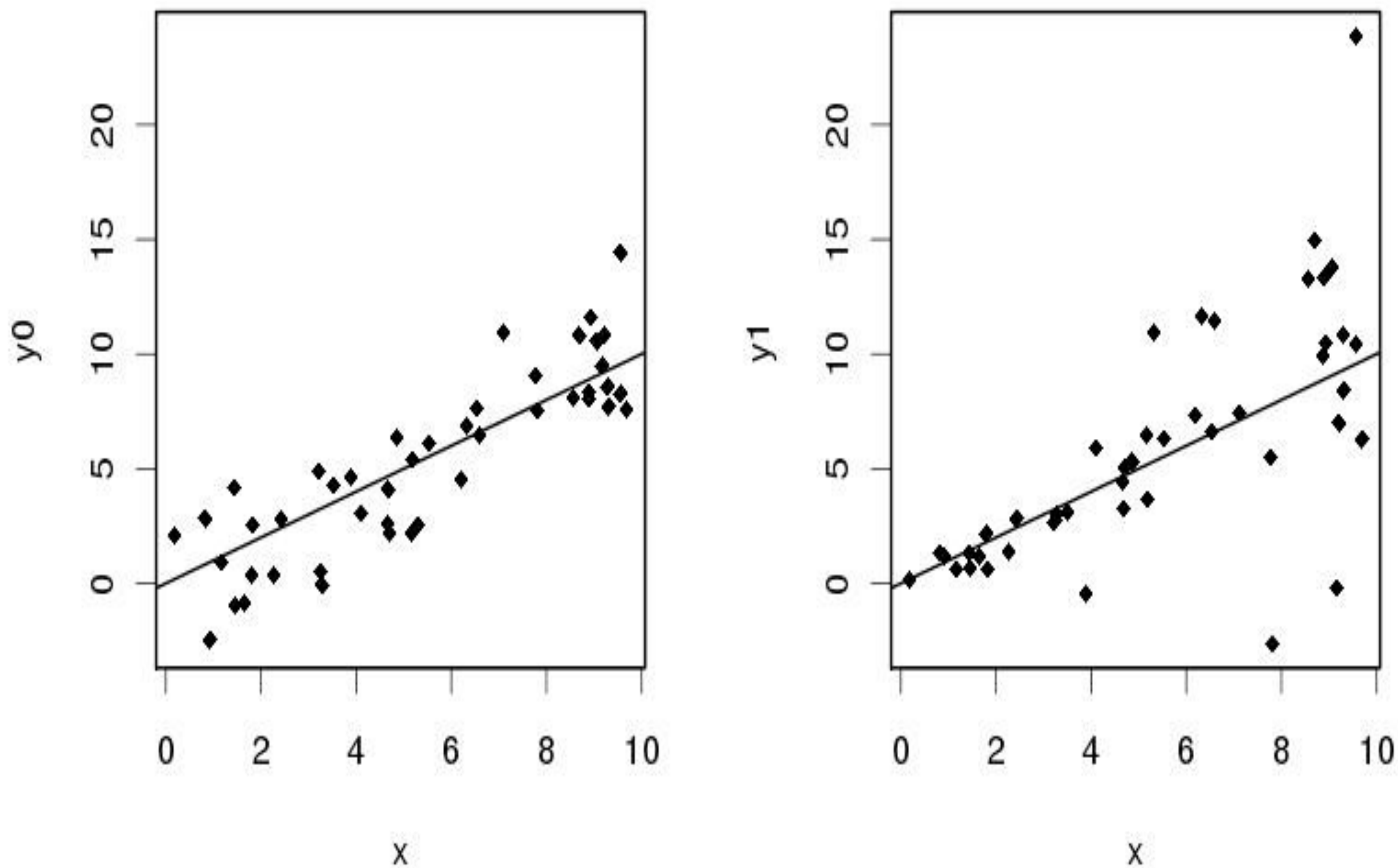


$$\frac{\lambda^y}{y!}e^{-\lambda}$$

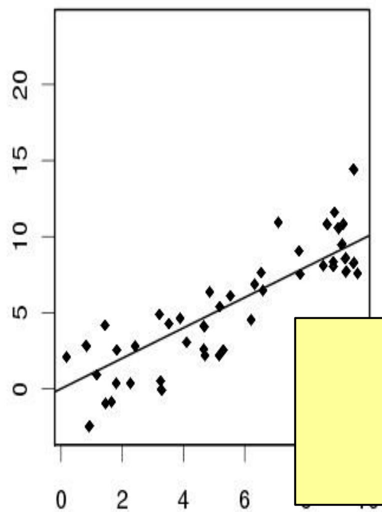
Uncertainty & Variance



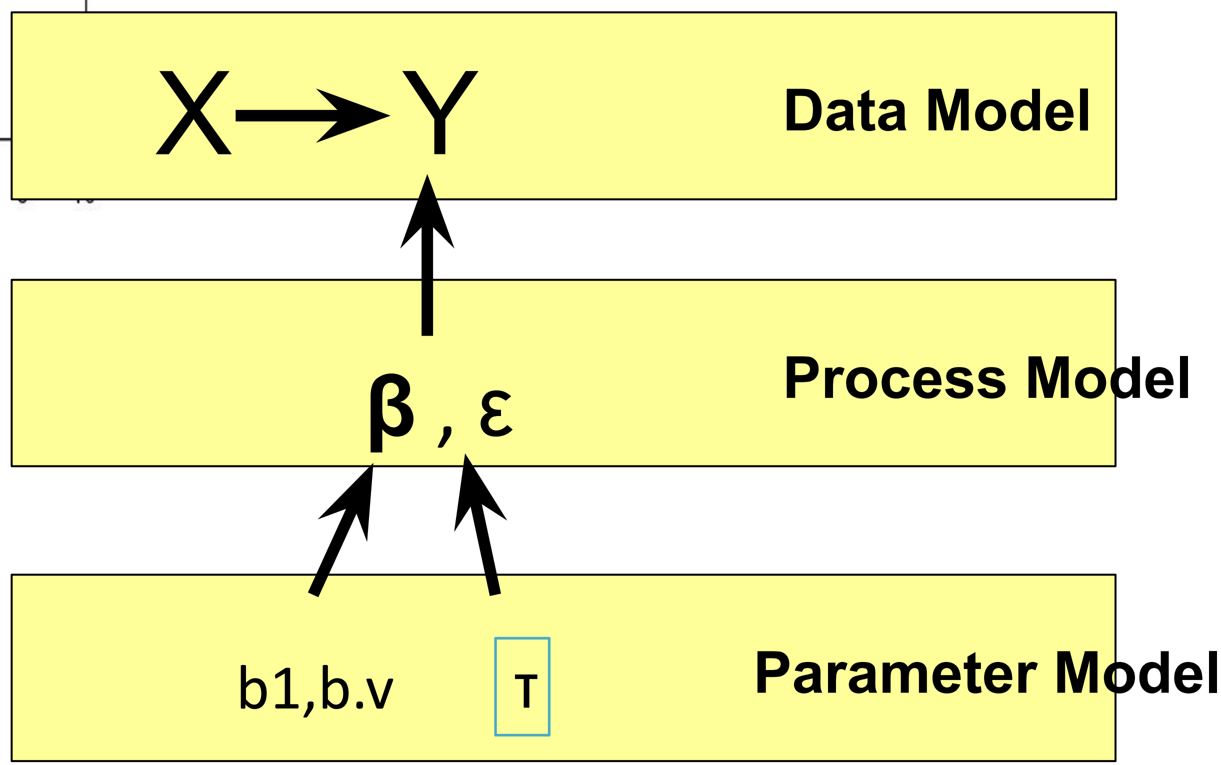
Uncertainty & Variance



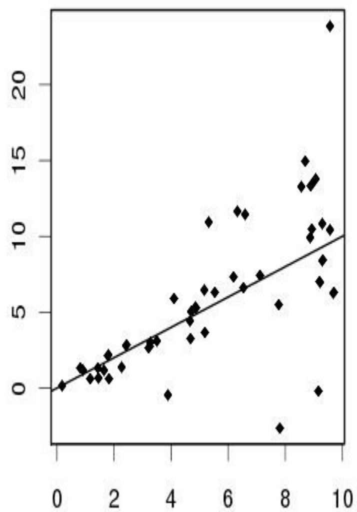
Uncertainty & Variance



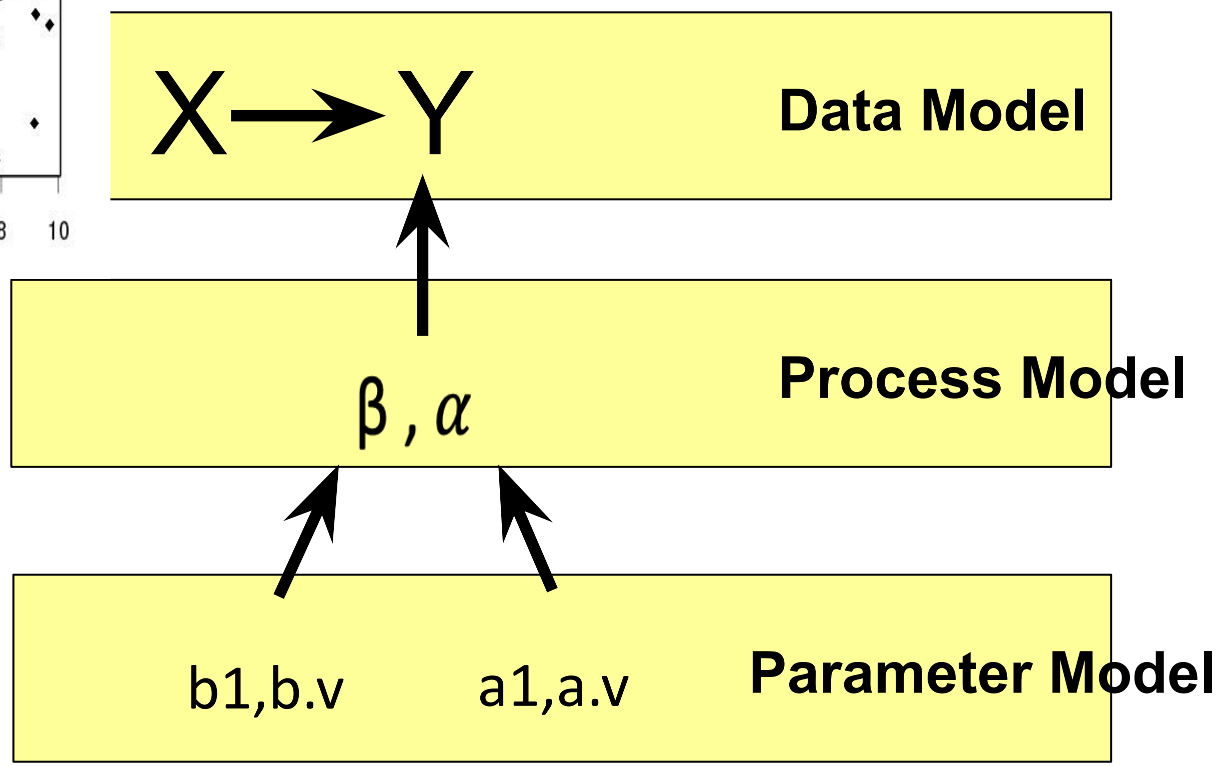
$\varepsilon_i \sim N(0,\tau)$



Uncertainty & Variance: Heteroskedasticity



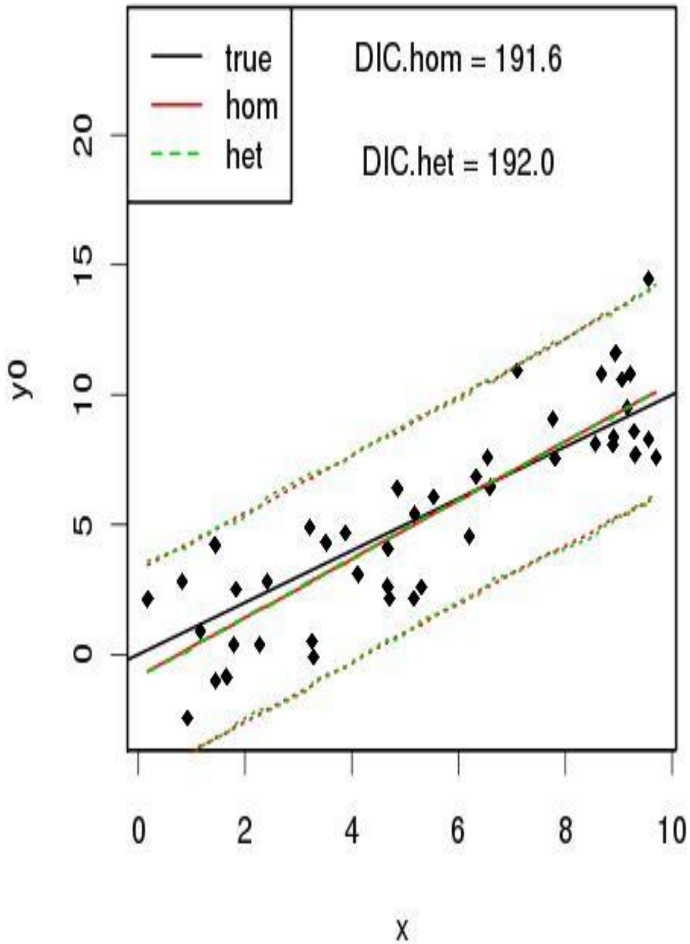
$$y \sim N(\beta_1 + \beta_2 x, (\alpha_1 + \alpha_2 x)^2)$$

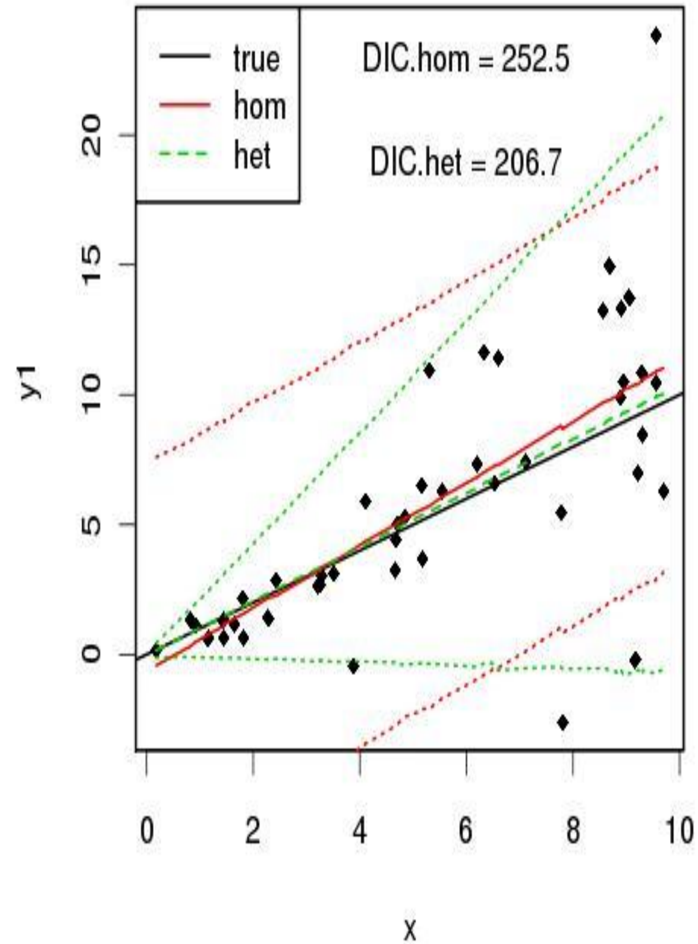


Uncertainty & Variance: Heteroskedasticity

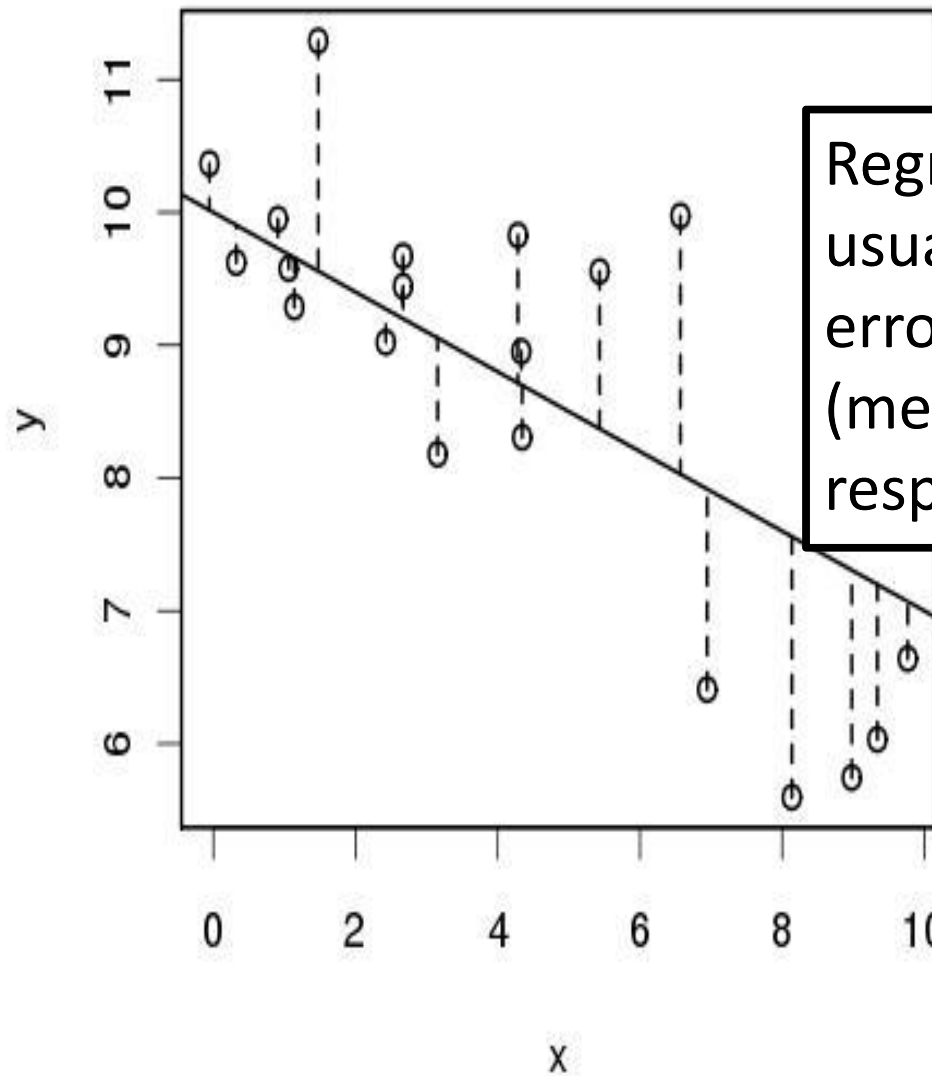
$y \sim N(\beta_1 + \beta_2 x, s^2)$

$y \sim N\left(\beta_1 + \beta_2 x, (\alpha_1 + \alpha_2 x)^2\right)$



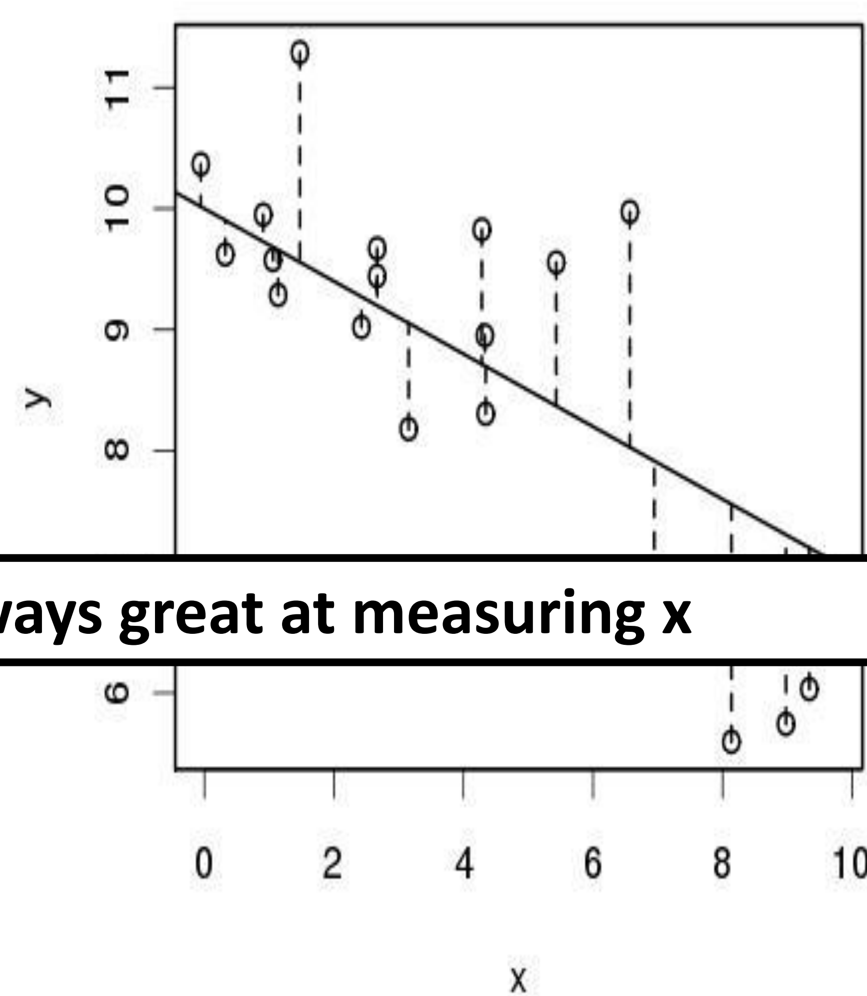


Uncertainty & Variance



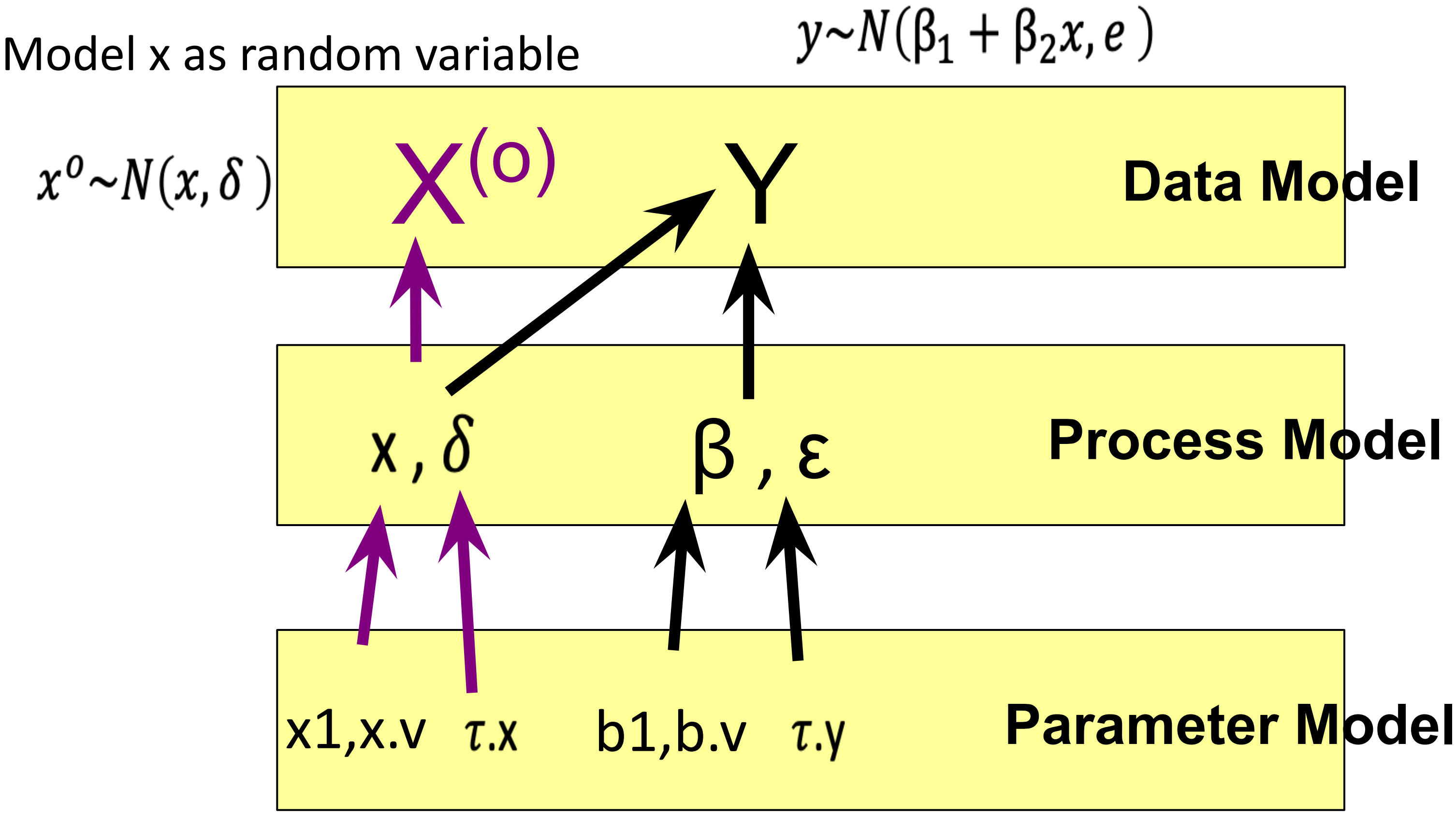
Regression model usually assumes all error is in Y (measurement of response)

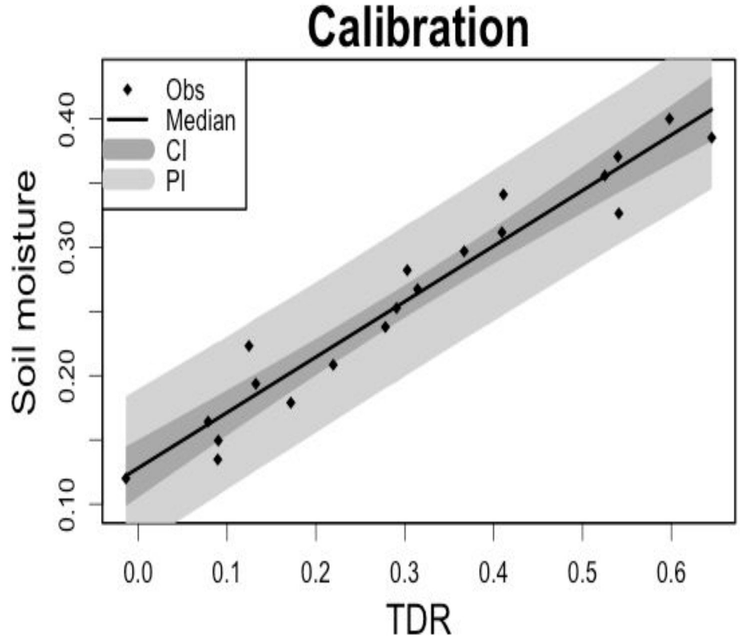
Uncertainty & Variance

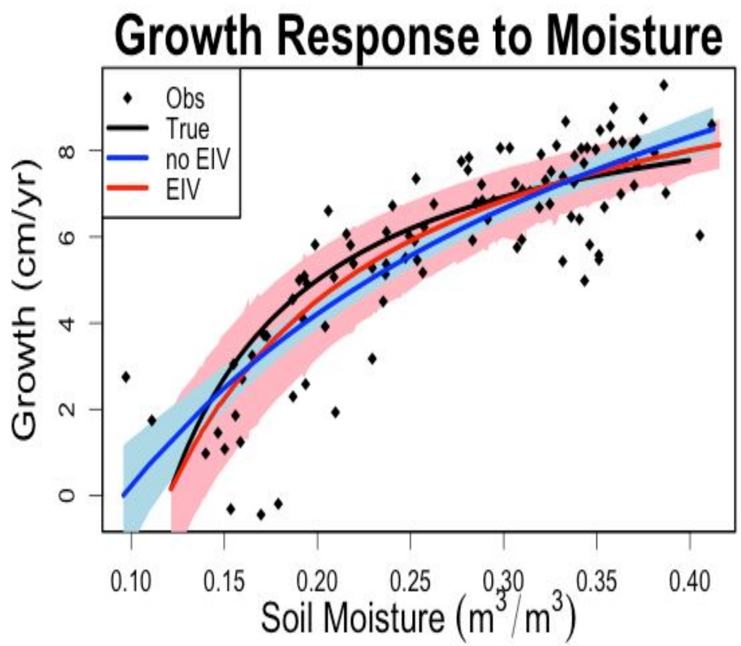


But we're not always great at measuring x either…

Uncertainty & Variance: Errors in Variables




Calibration
Obs
Median
CI
PI
Soil moisture
0.10
0.20
0.30
0.40
0.0
0.1
0.2
0.3
0.4
0.5
0.6
TDR

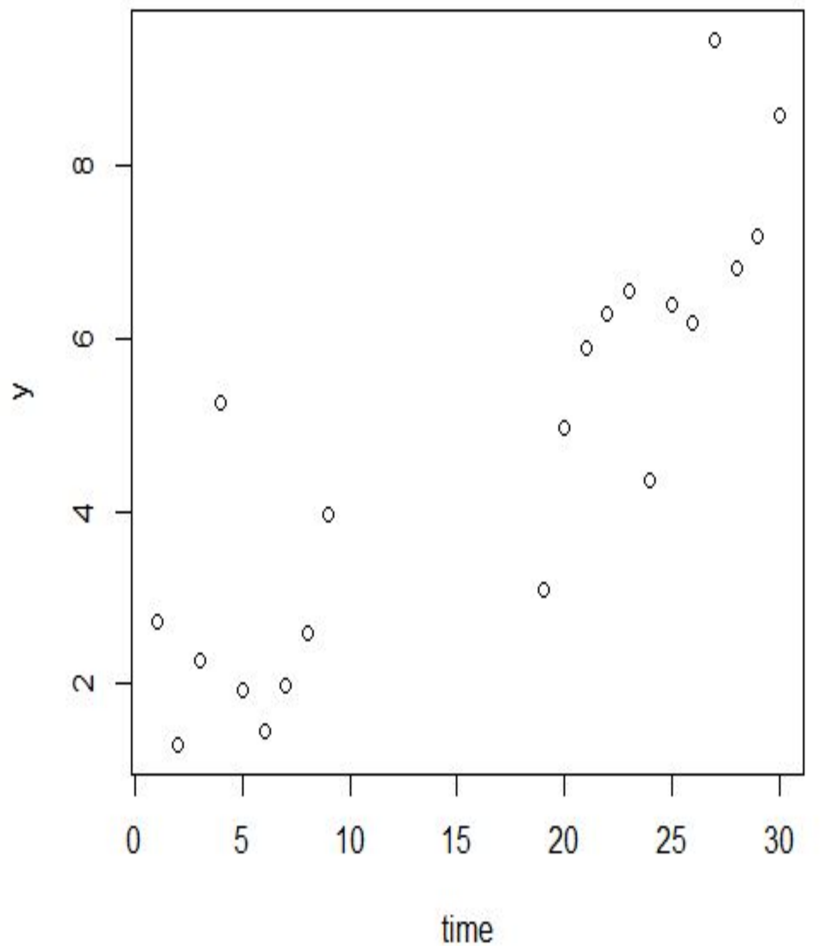

Growth Response to Moisture
Obs
True
no EIV
EIV
Growth (cm/yr)
0
2
4
6
8
0.10
0.15
0.20
0.25
0.30
0.35
0.40
Soil Moisture (m^3/m^3)

Latent Variables

- State not directly observed
  - Missing data

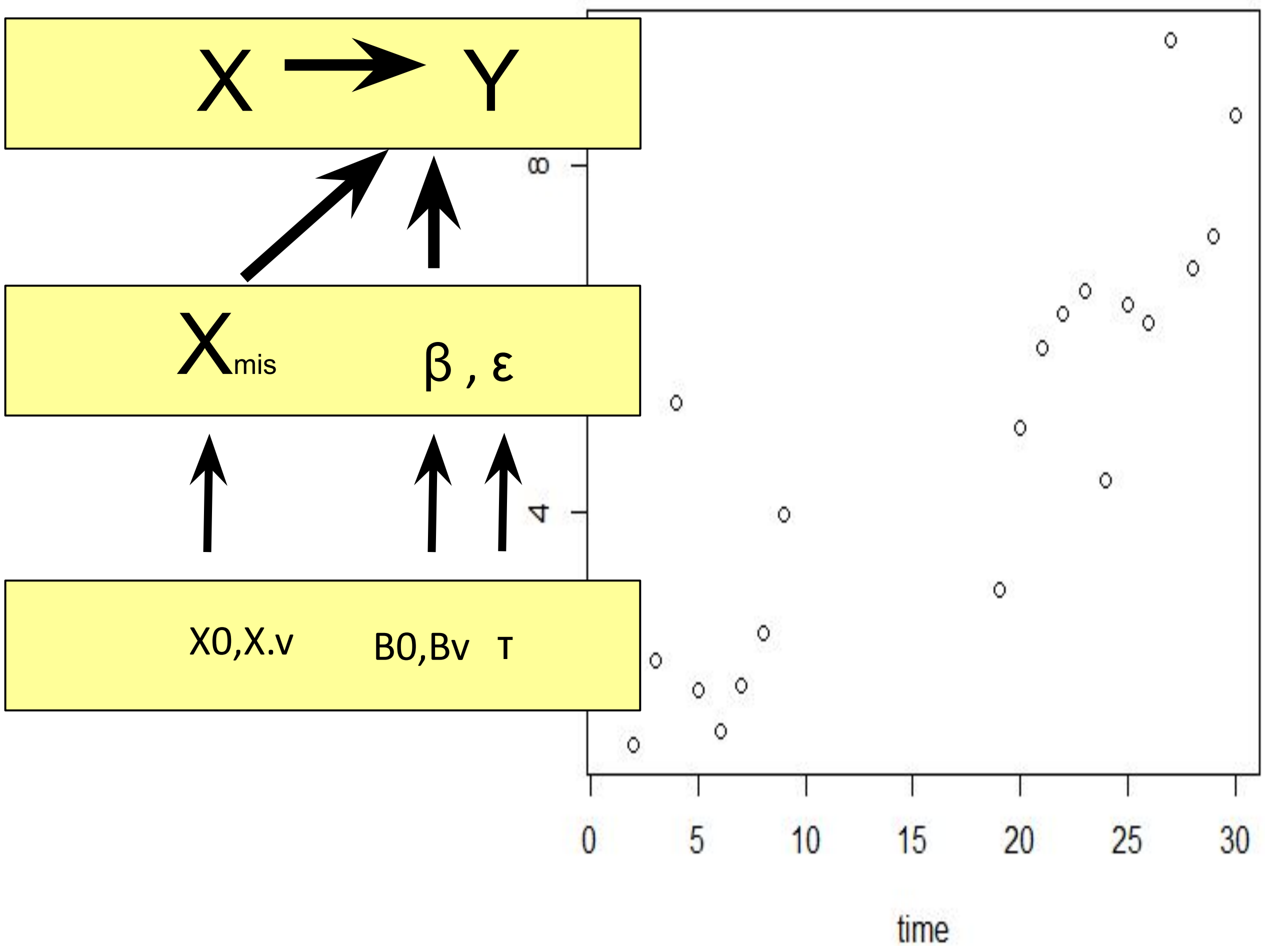
Latent Variables

Missing Data



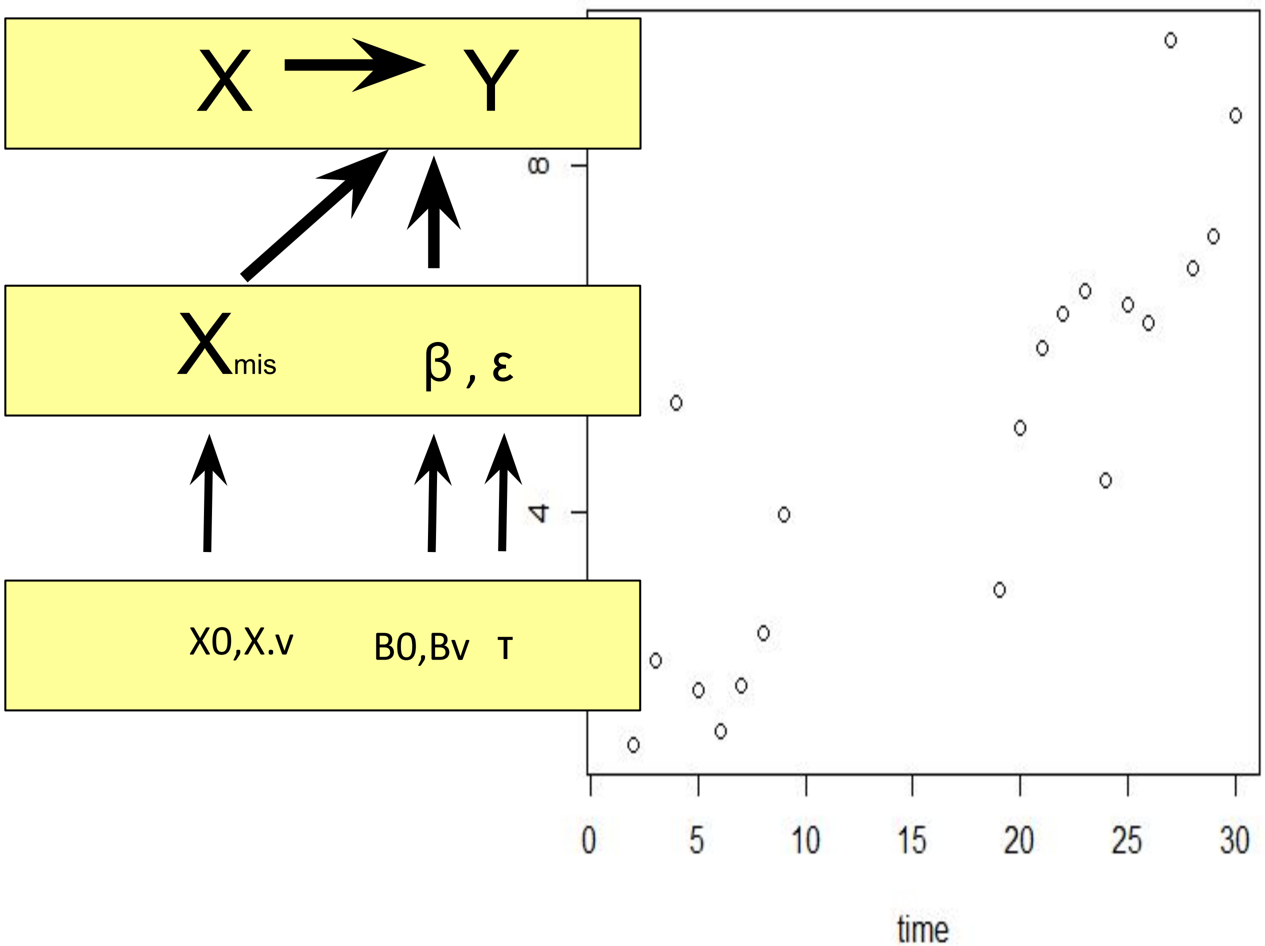
Latent Variables

Missing Data



Latent Variables

Missing Data

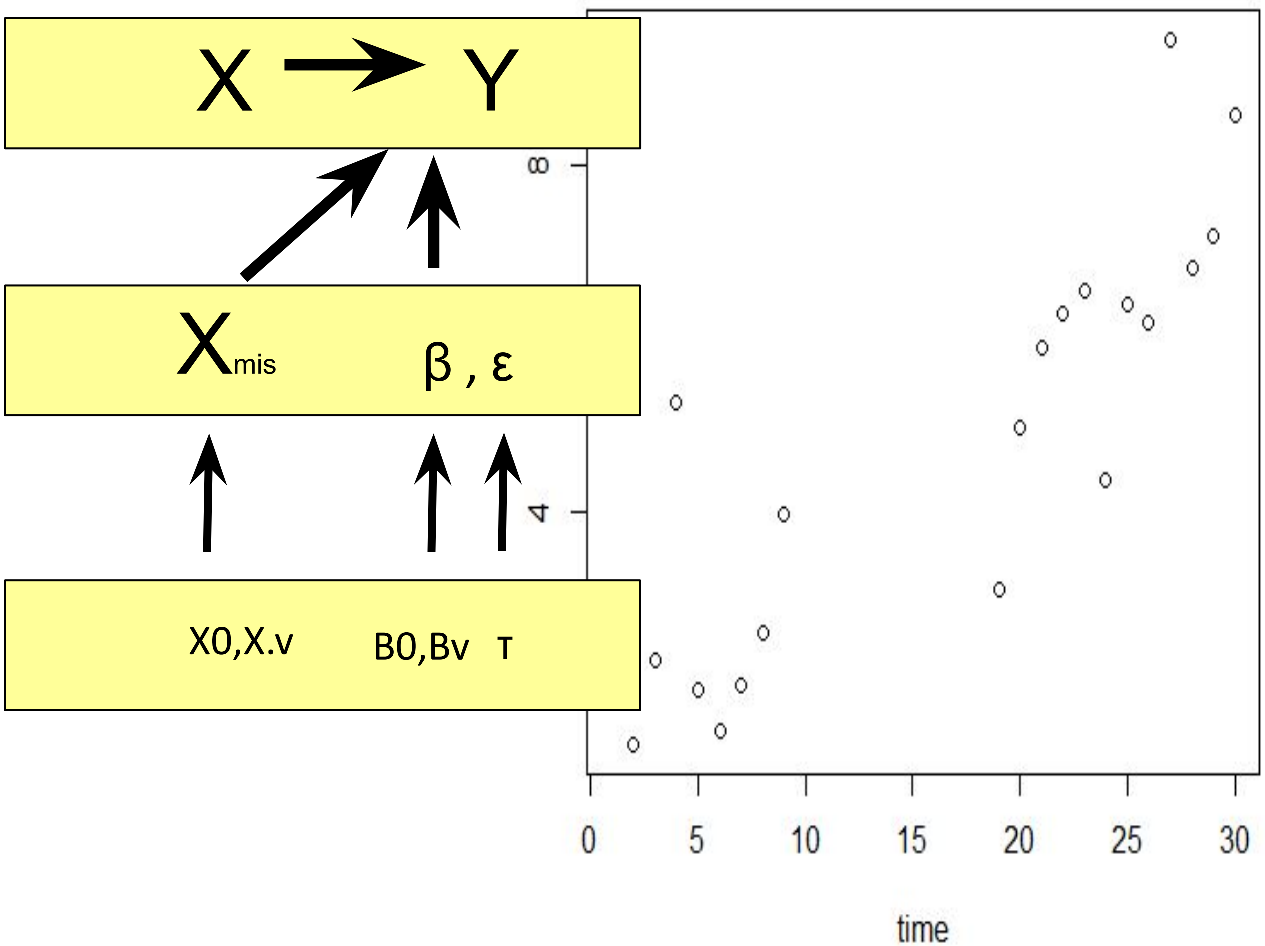


- Update the regression model conditioned on the current values of the missing data

- Update the missing data based on the current regression model and covariate values

Latent Variables

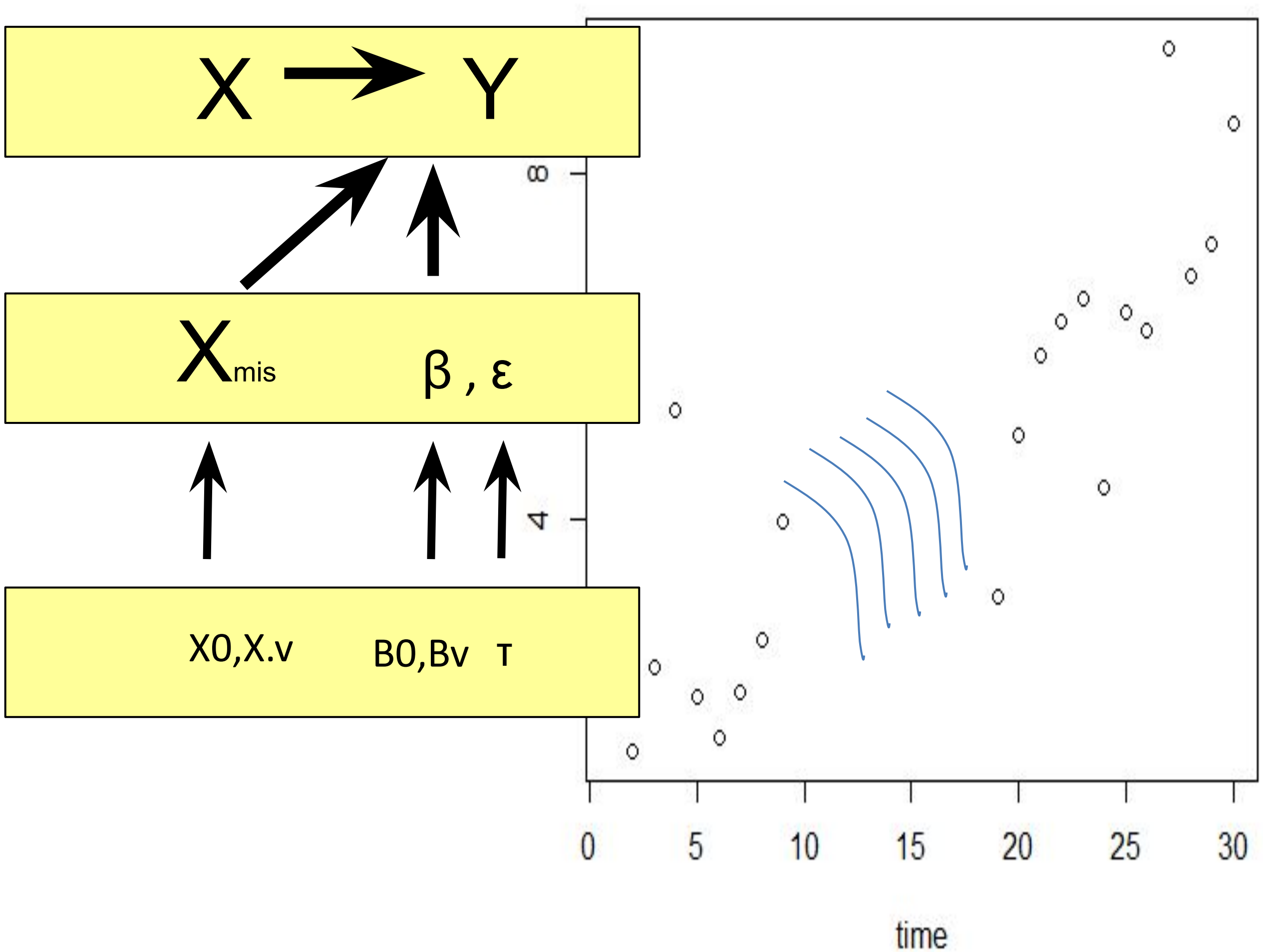
Missing Data



data is ***missing at random***

Latent Variables

Missing Data



data is *__missing at random__*

Latent Variables

- State not directly observed
  - Missing data
  - Proxy measures

Latent Variables

- State not directly observed
  - Missing data
  - Proxy measures

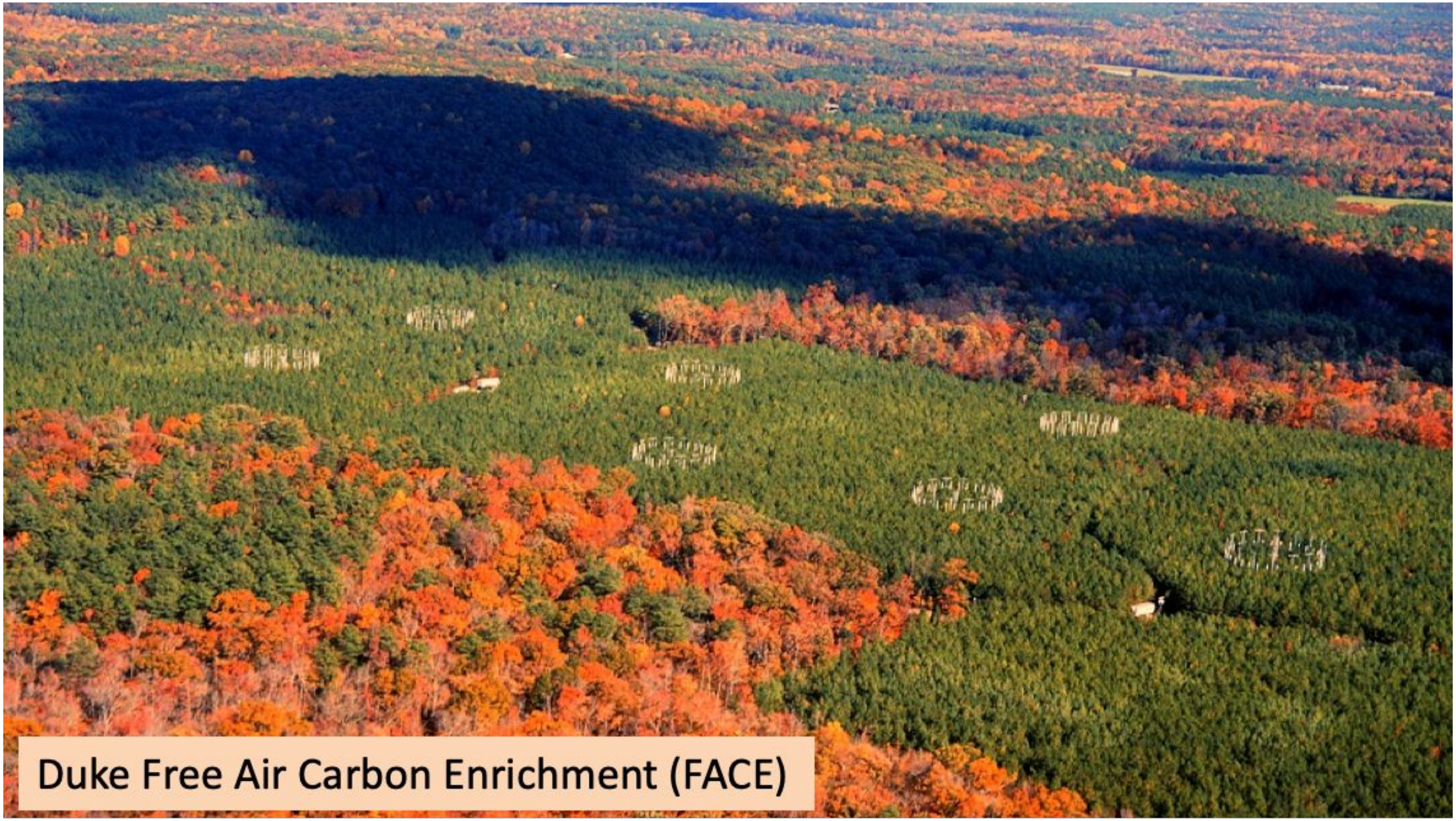
Will rising atmospheric CO_2 increase tree fecundity?

Latent Variables

Ignoring variable latency can lead to incorrect or falsely <u>overconfident</u> conclusions (and wonky bad forecasts)

Latent Variables

Will rising atmospheric CO_2 increase tree fecundity?



Latent Variables

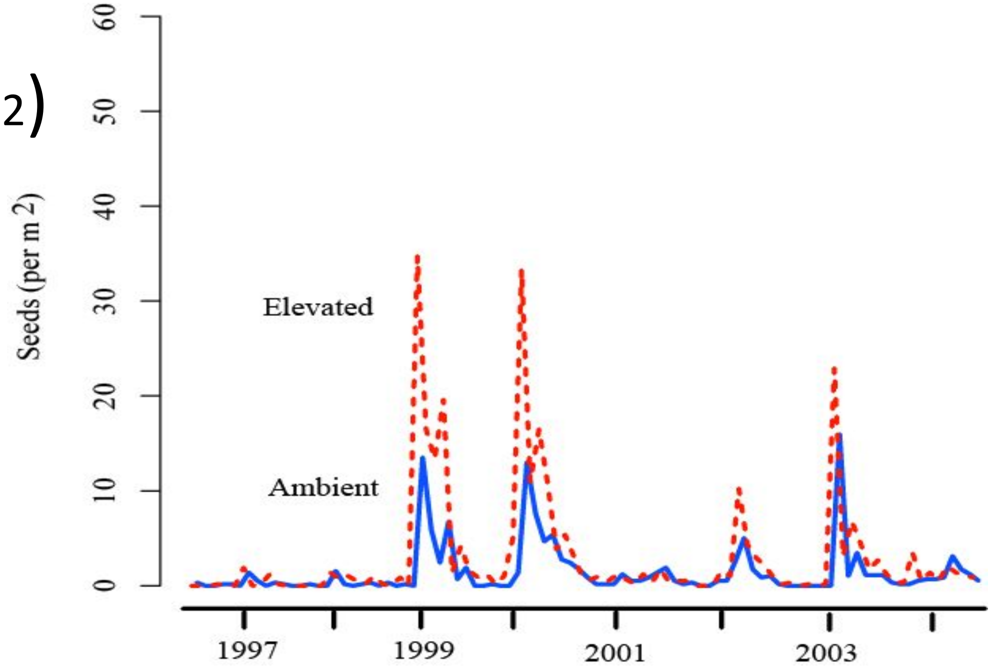
Will rising atmospheric CO_2 increase tree fecundity?

Fecundity ~ f(CO_2)

Latent Variables

Will rising atmospheric CO_2 increase tree fecundity?

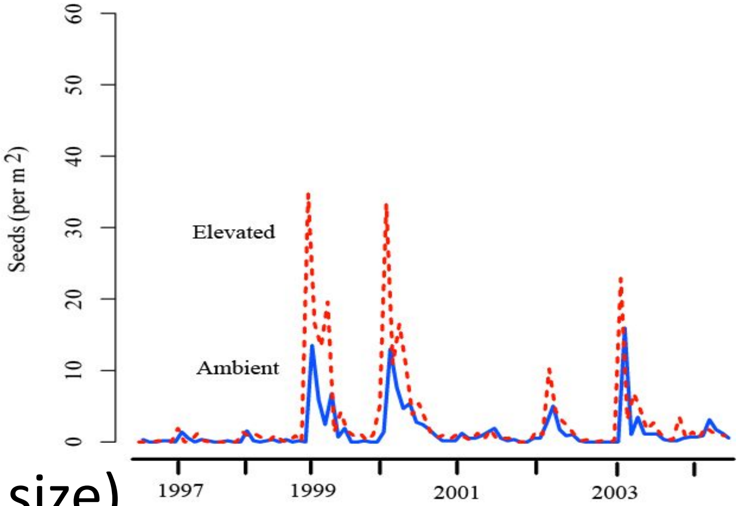
Seeds/m^2~ $f(CO_2)$



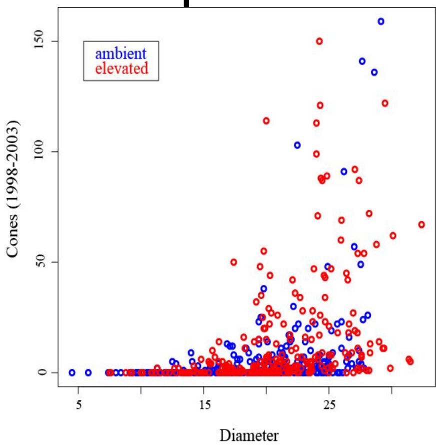
Latent Variables

Will rising atmospheric CO_2 increase tree fecundity?

Seeds/m^2~ f(CO_2)



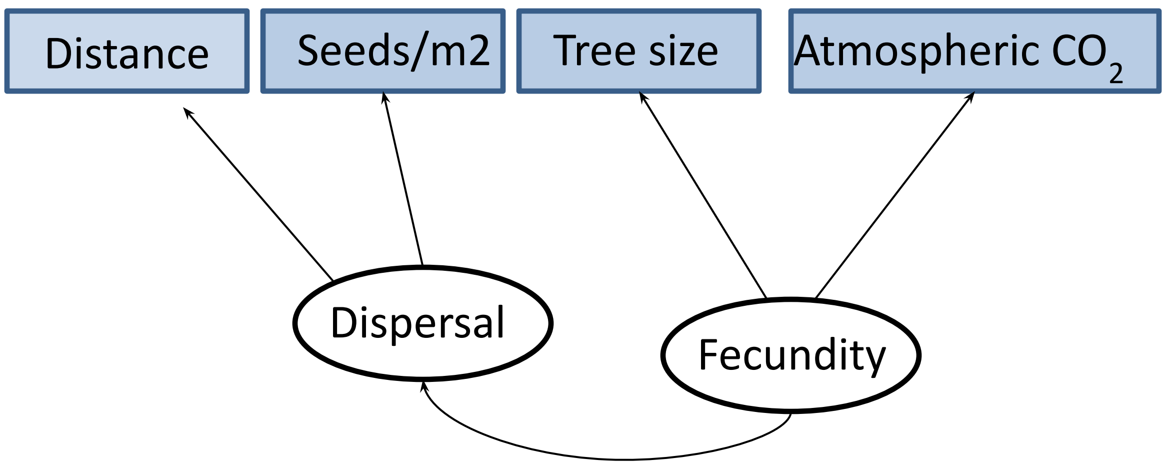
Seed production ~ f (tree size)



Seeds observed~ f(trees/m^2, dispersal)

Latent Variables

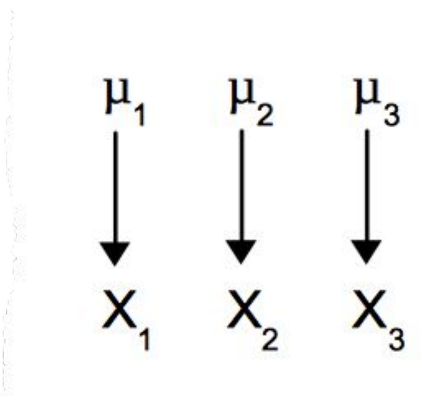
Will rising atmospheric CO_2 increase tree fecundity?



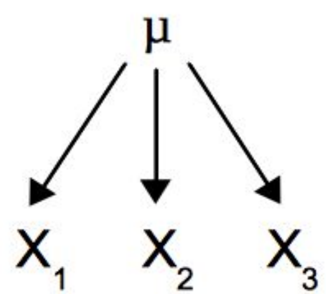
Fecundity ~ f(CO_2 + tree size) & f(Obs)

Hierarchical Bayes

Hierarchical Bayes

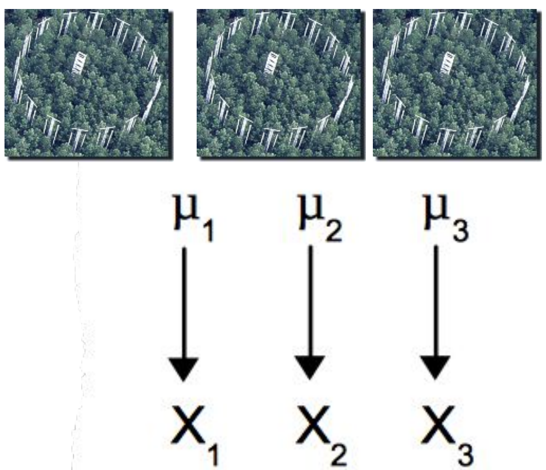


Independent

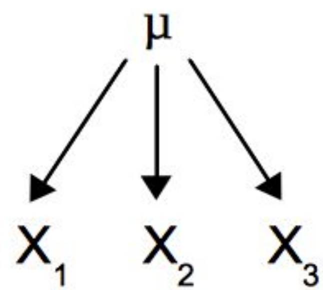


Shared

Hierarchical Bayes

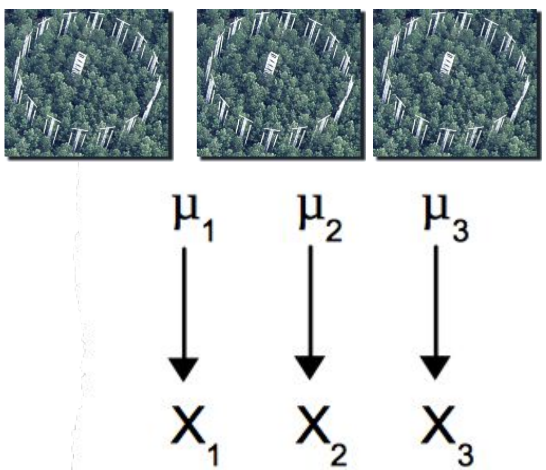


Independent

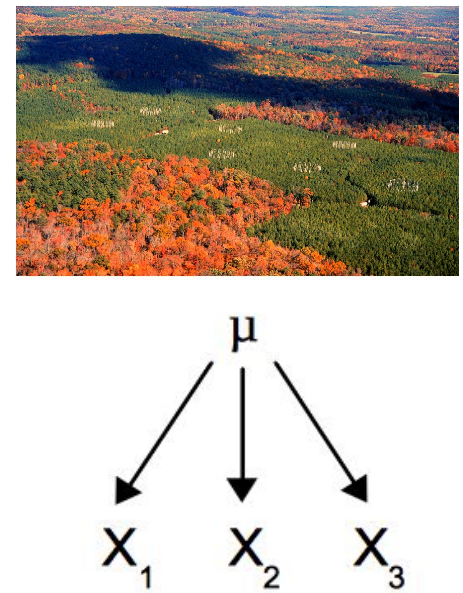


Shared

Hierarchical Bayes

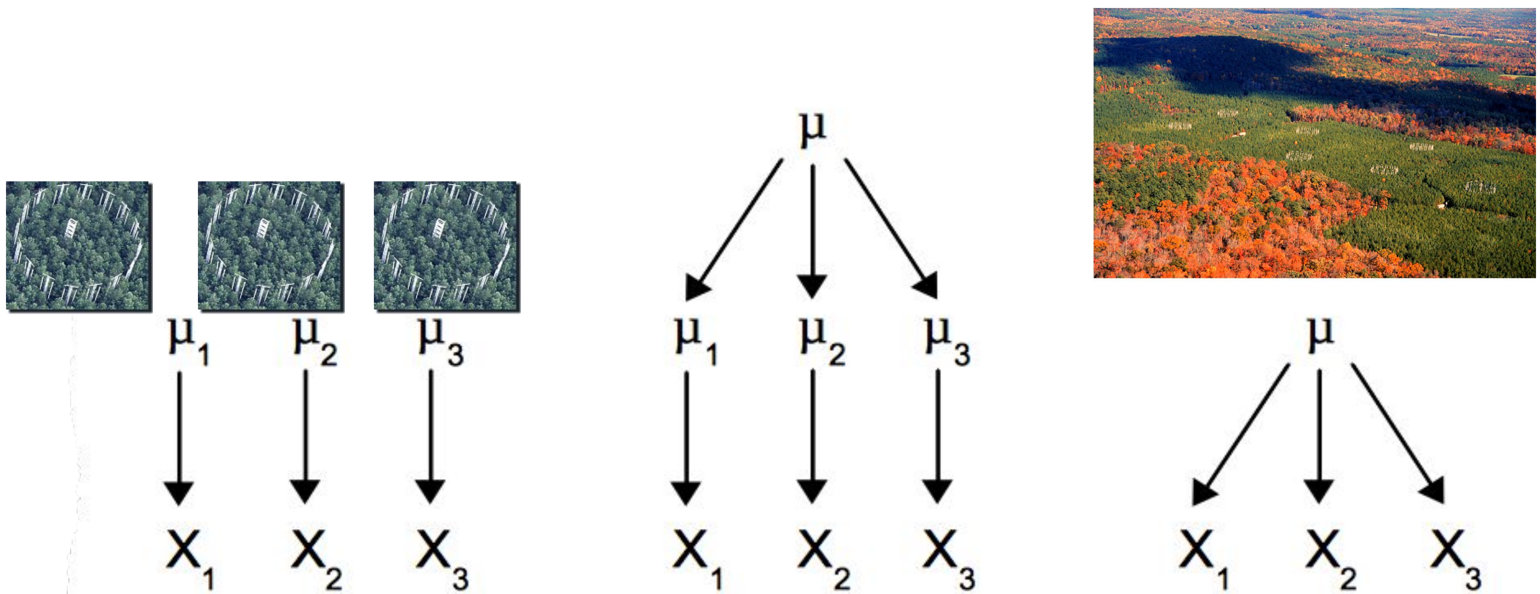


Independent



Shared

Hierarchical Bayes



Independent Hierarchical Shared

Hierarchical Bayes

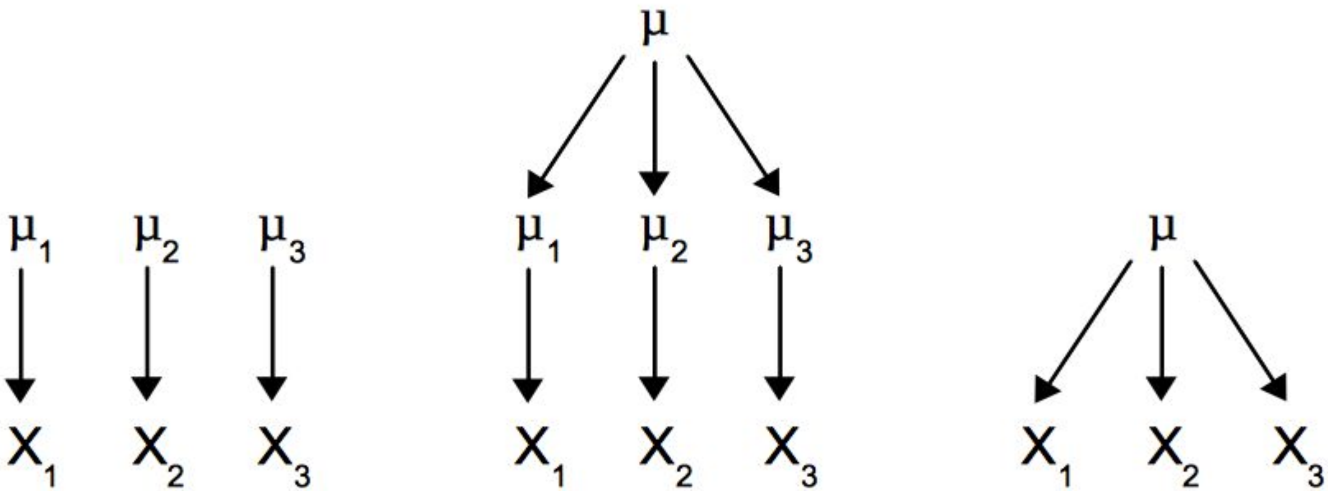
Model variability in the parameters

Partition variability explicitly into multiple terms

Borrow strength across data sets

Hierarchical Bayes

Predicting X_4?



Independent Hierarchical Shared

Hierarchical Bayes

Model variability in the parameters

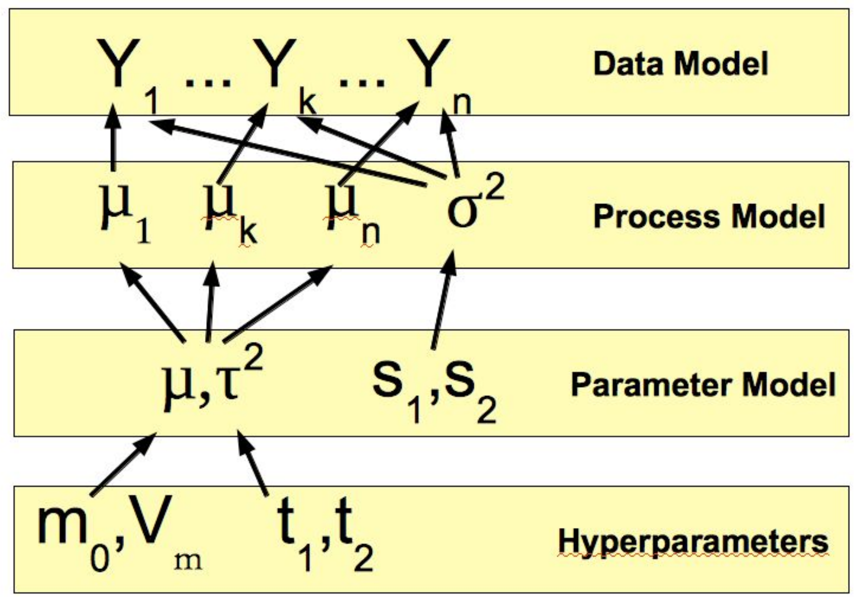
Partition variability explicitly into multiple terms

Borrow strength across data sets

Hierarchical with respect to parameters

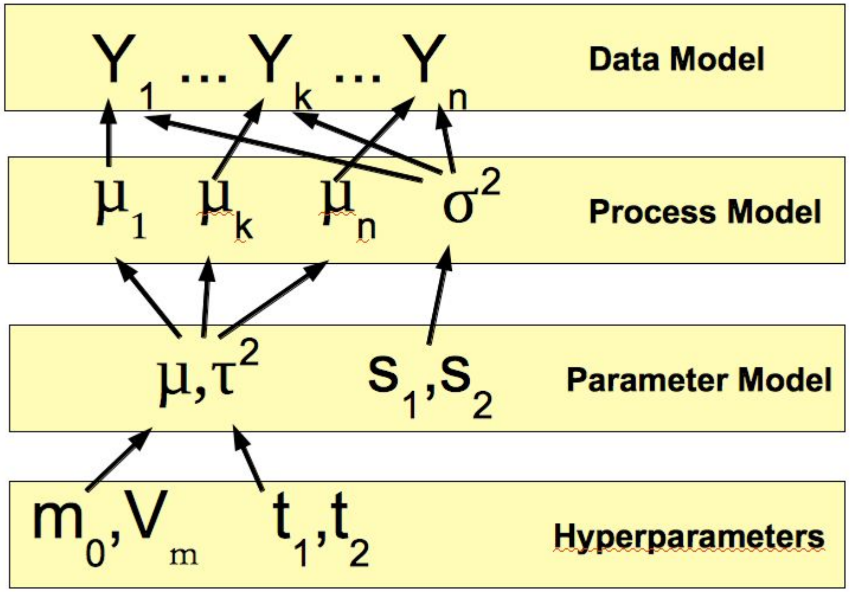
Pay attention to subscripts!

Hierarchical Bayes



Hierarchical Bayes

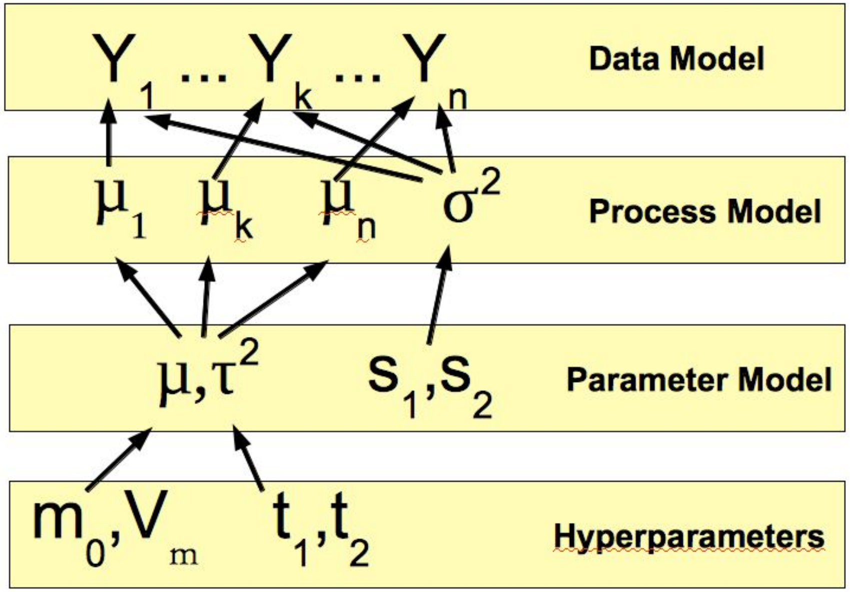
$Y_k \sim N(u_k, \sigma^2)$



Hierarchical Bayes

$Y_k \sim N(u_k, \sigma^2)$

$u_k \sim N(u, \tau^2)$
$\sigma^2 \sim IG(s_1, s_2)$

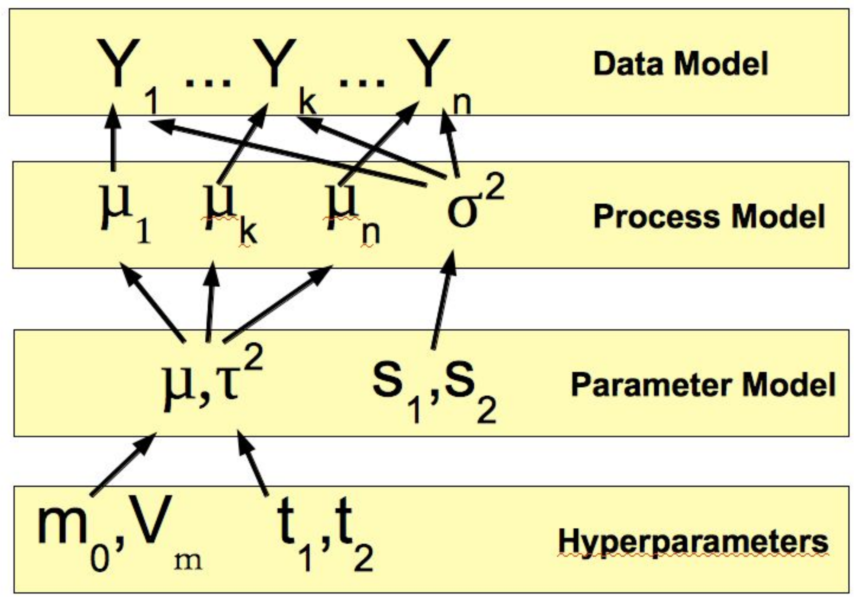


Hierarchical Bayes

$Y_k \sim N(u_k,\sigma^2)$

$u_k \sim N(u,\tau^2)$
$\sigma^2 \sim IG(s_1,s_2)$

$u \sim N(u0,v_m)$
$\tau^2 \sim IG(t1,t2)$



Hierarchical Bayes – random effects model

$Y_k \sim N(u_k, \sigma^2)$

$Y_k \sim N(u_g + \alpha_k , \sigma^2)$

$u_k \sim N(u, \tau^2)$
$\sigma^2 \sim IG(s_1, s_2)$

$u \sim N(u0, v_m)$
$\tau^2 \sim IG(t1, t2)$

Hierarchical Bayes – random effects

Random effects attempt to account for the unexplained variance associated with some group (plot, year, etc.) due to all the things that were not measured

May point to scales that need additional explanation

Adding covariates may explain some portion of this variance, *but there's always something you didn't measure*

Hierarchical Bayes – random effects model

$Y_k \sim N(u_k, \sigma^2)$

$u_k \sim N(u, \tau^2)$

$\sigma^2 \sim IG(s_1, s_2)$

$u \sim N(u0, v_m)$

$\tau^2 \sim IG(t1, t2)$

$Y_k \sim N(u_g + \alpha_k\ , \sigma^2)$

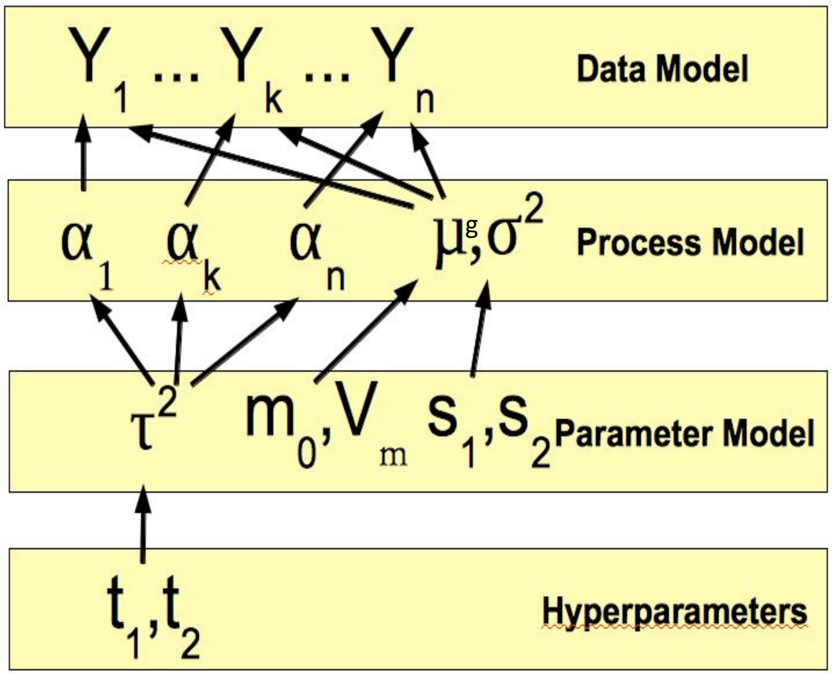
$\boldsymbol{\alpha_k \sim N(0, \tau^2)}$

$\sigma^2 \sim IG(s_1, s_2)$

$u_g \sim N(u0, v_m)$

$\tau^2 \sim IG(t1, t2)$

Hierarchical Bayes – random effects model



$Y_k \sim N(u_g + \alpha_k, \sigma^2)$

$\boldsymbol{\alpha_k \sim N(0,\tau^2)}$

$\sigma^2 \sim IG(s_1, s_2)$

$u_g \sim N(u0, v_m)$

$\tau^2 \sim IG(t1, t2)$

Hierarchical Bayes – random effects model

$Y_k \sim N(u_g + \alpha_k , \sigma^2)$

REEs have mean 0

$\boldsymbol{\alpha_k \sim N(0,\tau^2)}$

$\sigma^2 \sim IG(s_1,s_2)$

$u_g \sim N(u0,v_m)$

$\tau^2 \sim IG(t1,t2)$

Hierarchical Bayes – random effects model

$Y_k \sim N(u_g + \alpha_k , \sigma^2)$

REEs have mean 0

$\boldsymbol{\alpha_k \sim N(0,\tau^2)}$

$\sigma^2 \sim IG(s_1,s_2)$

$u_g \sim N(u0,v_m)$

REs variance attributes portion of uncertainty to specific source

$\tau^2 \sim IG(t1,t2)$

Hierarchical Bayes – random effects model

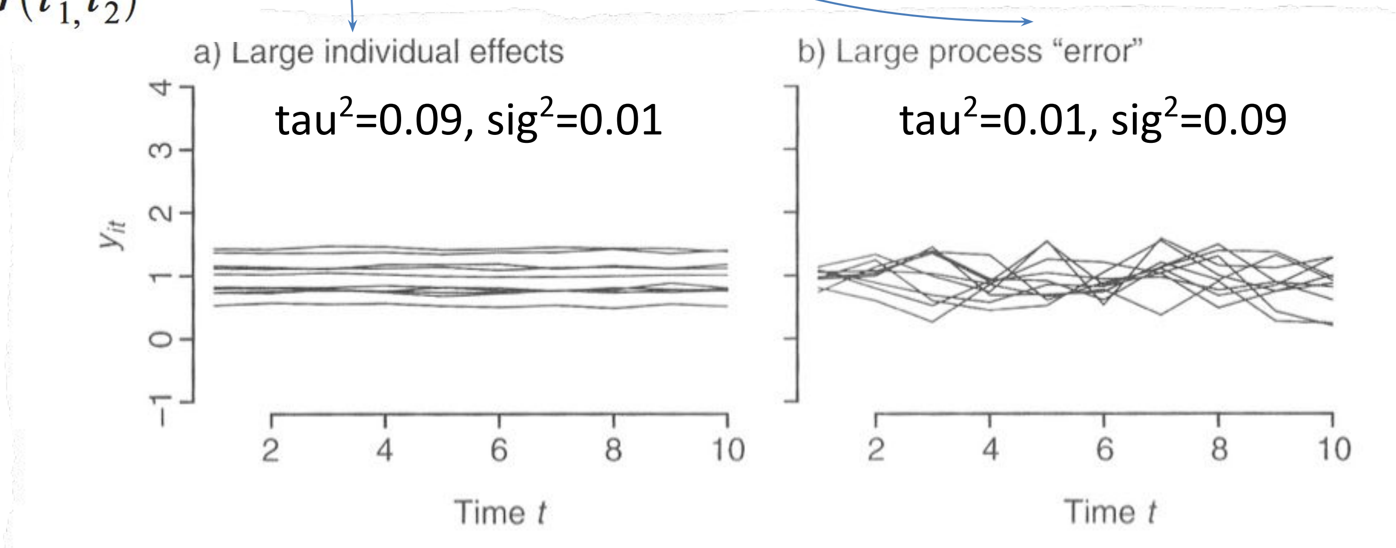
$$Y_k \sim N(\mu_g + \alpha_k, \sigma^2)$$

$$\alpha_k \sim N(0, \tau^2)$$

$$\sigma^2 \sim IG(s_1, s_2)$$

$$\mu_g \sim N(\mu_{0,} V_\mu)$$

$$\tau^2 \sim IG(t_{1,} t_2)$$



Hierarchical Bayes – random effects model

RE's traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

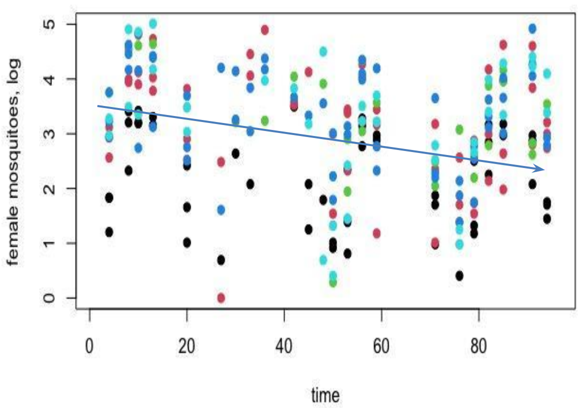
- *e.g., Plots, Years*
- used to account for lack of independence among reps

Hierarchical Bayes – random effects model

RE's traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

- *e.g., Plots, Years*
- used to account for lack of independence among reps

Replication is important for identifiability (i.e., partitioning variance between process error and REs)



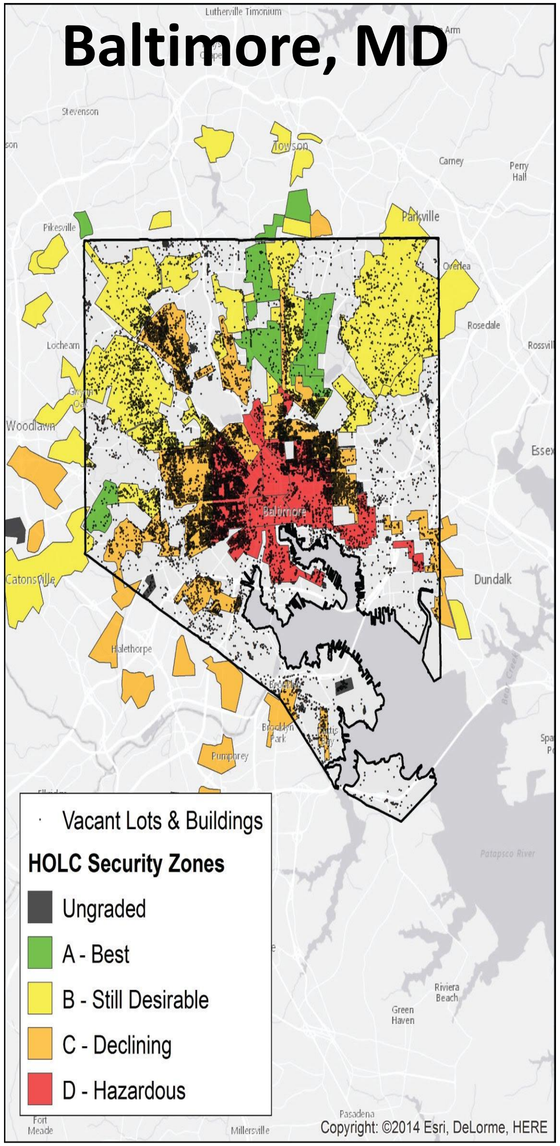
Hierarchical Bayes – random effects model

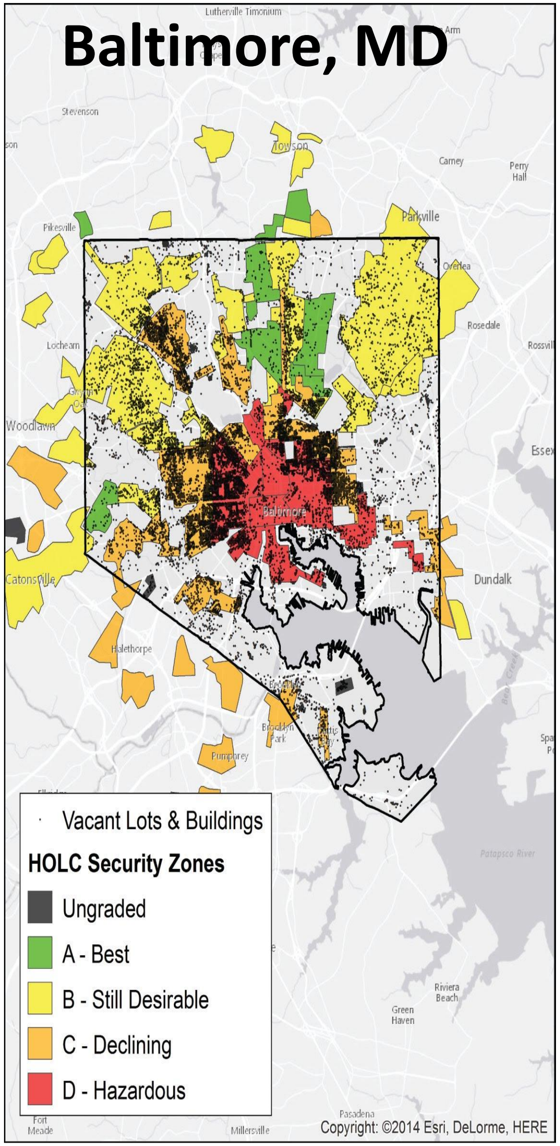
RE’s traditionally aspects of the study that would likely change if experiment/observation protocol was replicated

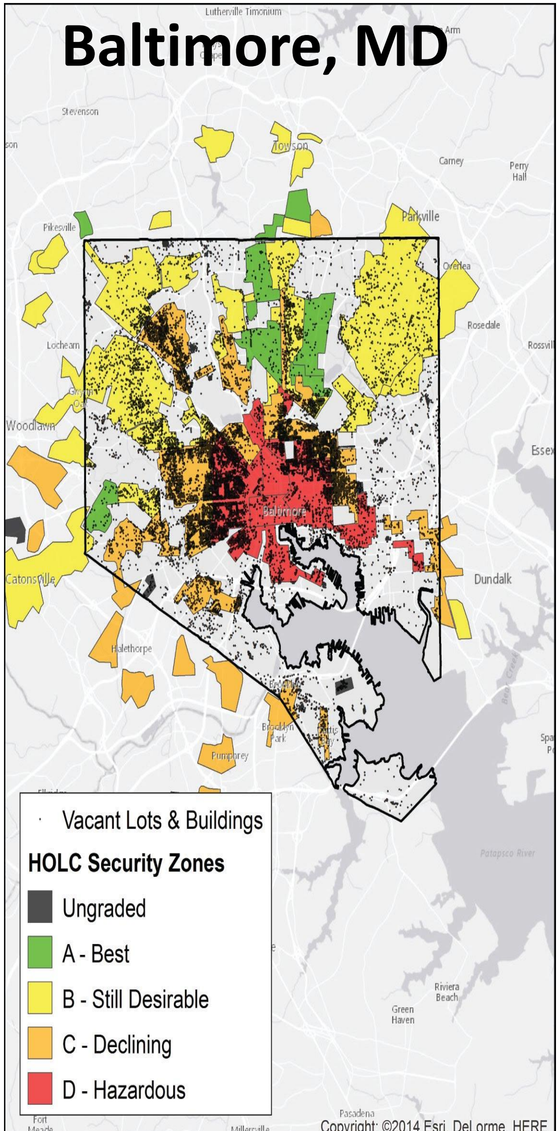
- *e.g., Plots, Years*
- used to account for lack of independence among reps

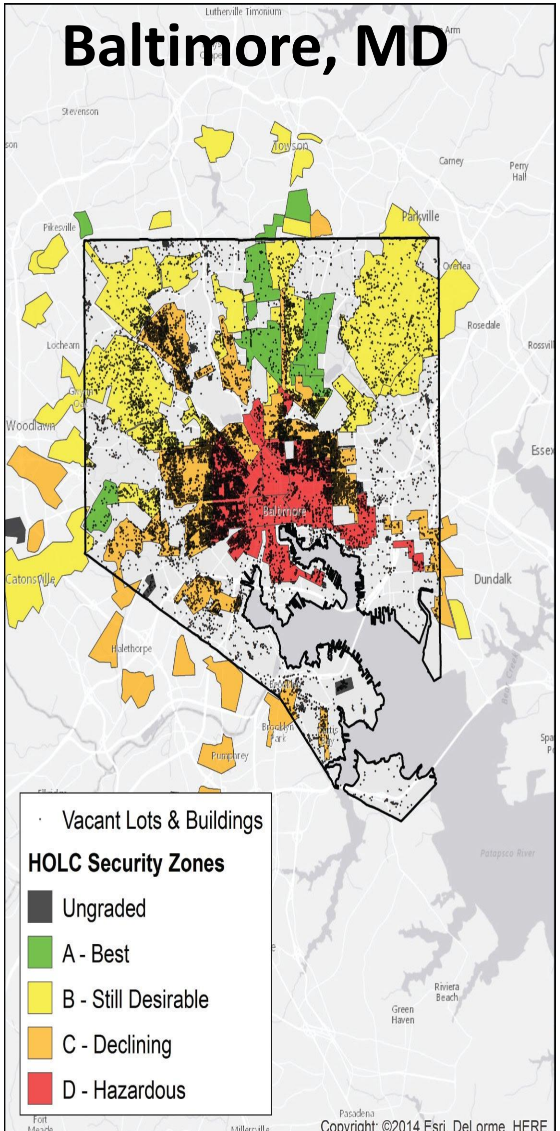
Replication is important for identifiability (i.e., partitioning variance between process error and REs)

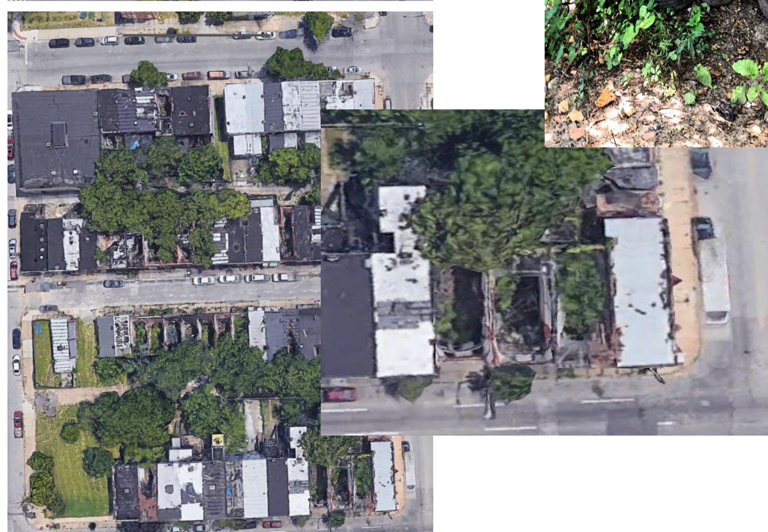
Treatments/covariates of interest are treated as fixed effects (mixed effects model)

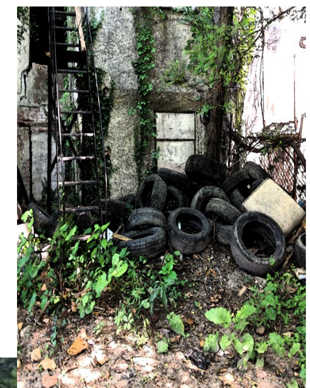

Baltimore, MD
Vacant Lots & Buildings
HOLC Security Zones
Ungraded
A - Best
B - Still Desirable
C - Declining
D - Hazardous
Copyright: ©2014 Esri, DeLorme, HERE

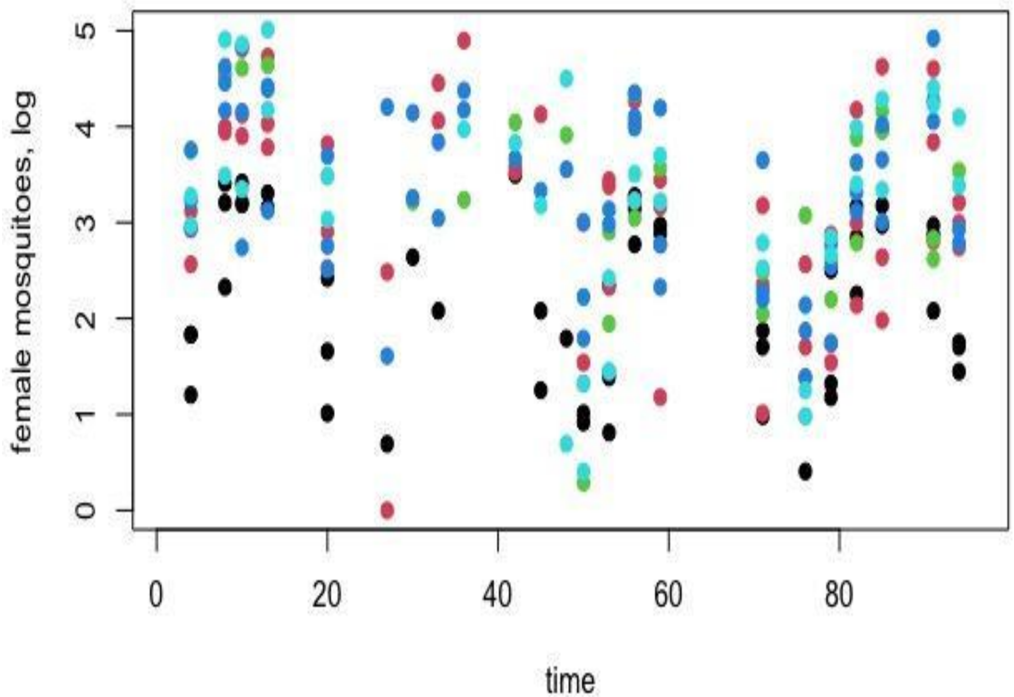

Baltimore, MD
Vacant Lots & Buildings
HOLC Security Zones
Ungraded
A - Best
B - Still Desirable
C - Declining
D - Hazardous
Copyright: ©2014 Esri, DeLorme, HERE


Baltimore, MD
Lutherville Timonium
Towson
Parkville
Pikesville
Overlea
Rosedale
Woodlawn
Essex
Baltimore
Catonsville
Dundalk
Halethorpe
Vacant Lots & Buildings
HOLC Security Zones
Ungraded
A - Best
B - Still Desirable
C - Declining
D - Hazardous
Copyright: ©2014 Esri, DeLorme, HERE


Baltimore, MD
Vacant Lots & Buildings
HOLC Security Zones
Ungraded
A - Best
B - Still Desirable
C - Declining
D - Hazardous
Copyright: ©2014 Esri, DeLorme, HERE

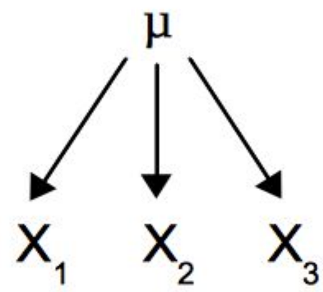






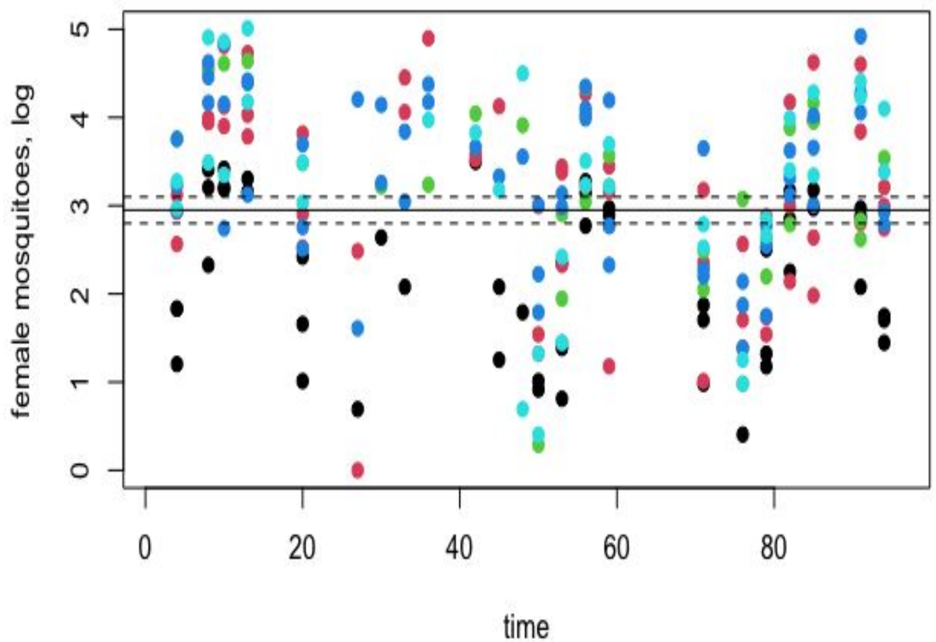
female mosquitoes, log
0
1
2
3
4
5
0
20
40
60
80
time

```
Global Mean= "
model{
  mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)

for(i in 1:nblocks){
  for(t in 1:ndates){
    X[i,t] ~ dnorm(mu, tau)
  }}}
```

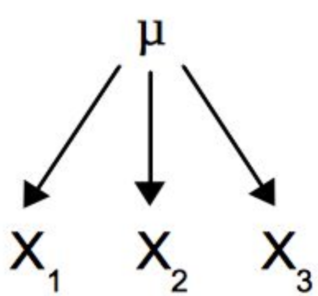


Shared

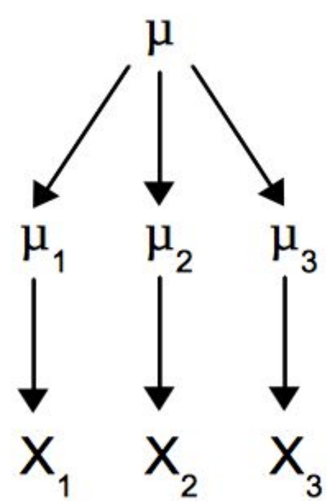
```
Global Mean= "
model{
  mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)

for(i in 1:nblocks){
  for(t in 1:ndates){
    X[i,t] ~ dnorm(mu, tau)
  }}}
```



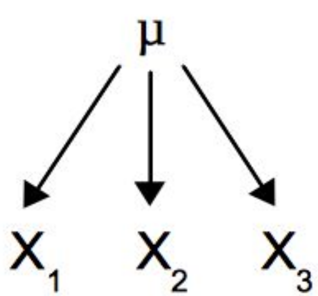
```
Global Mean= "
model{
  mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)

for(i in 1:nblocks){
  for(t in 1:ndates){
    X[i,t] ~ dnorm(mu, tau)
  }}}
```




```
Space= "
model{
  mu ~ dnorm(0,0.001)
  tau ~ dgamma(0.001,0.001)
  tau.sp~dgamma(0.001,0.001)

for(i in 1:nblocks){
  alpha.sp[i]~dnorm(0,tau.sp)
  Emu[i]=mu + alpha.sp[i]

for(t in 1:ndates){
    X[i,t] ~ dnorm(Ex[i], tau)
  }}}
"
```



```
Global Mean= "
model{
  mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)

for(i in 1:nblocks){
  for(t in 1:ndates){
    X[i,t] ~ dnorm(mu, tau)
  }}}
```

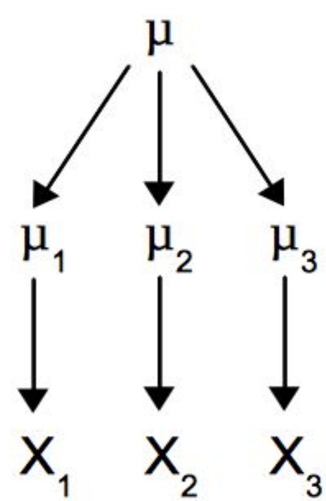



```
Space= "
model{
  mu ~ dnorm(0,0.001)
  tau ~ dgamma(0.001,0.001)
  tau.sp~dgamma(0.001,0.001)

for(i in 1:nblocks){
  alpha.sp[i]~dnorm(0,tau.sp)
  Emu[i]=mu + alpha.sp[i]

for(t in 1:ndates){
    X[i,t] ~ dnorm(Ex[i], tau)
  }}}
"
```

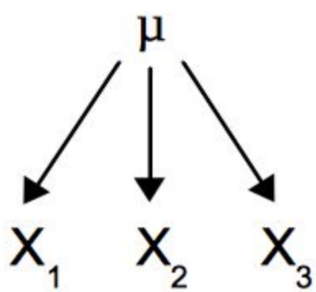
Parameter Model (Priors)

Process Model

Data Model



```
Global Mean= "
model{
 mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)

for(i in 1:nblocks){
 for(t in 1:ndates){
   X[i,t] ~ dnorm(mu, tau)
 }}}
```



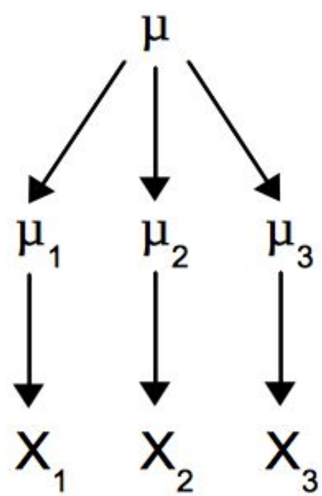

```
Space= "
model{
 mu ~ dnorm(0,0.001)
 tau ~ dgamma(0.001,0.001)
 tau.sp~dgamma(0.001,0.001)

for(i in 1:nblocks){
 alpha.sp[i]~dnorm(0,tau.sp)
 Emu[i]=mu + alpha.sp[i]

for(t in 1:ndates){
   X[i,t] ~ dnorm(Ex[i], tau)
 }}}
"
```

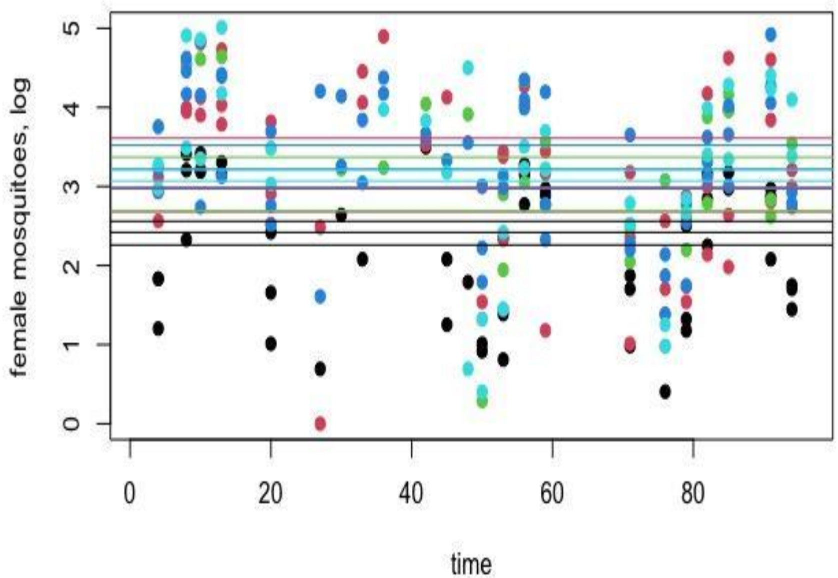
Parameter Model (Priors)

Hyperprior

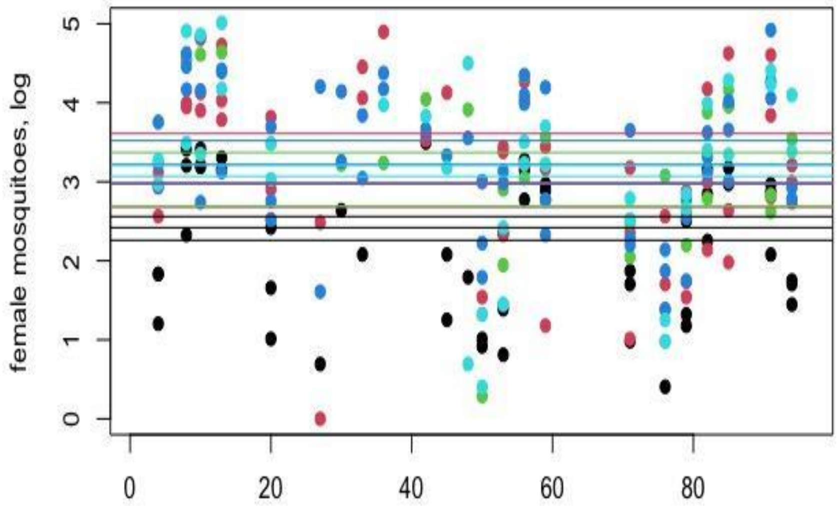
Process Model

Data Model

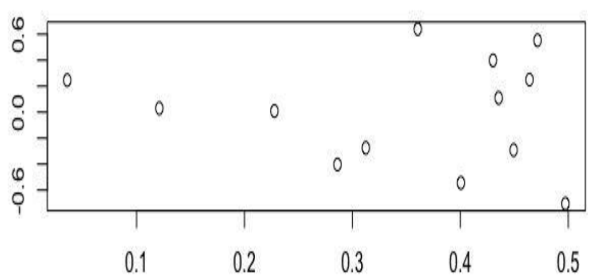





```
Space= "
model{
  mu ~ dnorm(0,0.001)
  tau ~ dgamma(0.001,0.001)
  tau.sp~dgamma(0.001,0.001)

for(i in 1:nblocks){
  alpha.sp[i]~dnorm(0,tau.sp)
  Emu[i]=mu + alpha.sp[i]

for(t in 1:ndates){
   X[i,t] ~ dnorm(Ex[i], tau)
  }}}
"
```

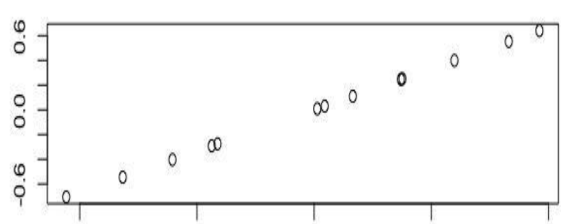


alpha.sp



Vegetation Cover

alpha.sp

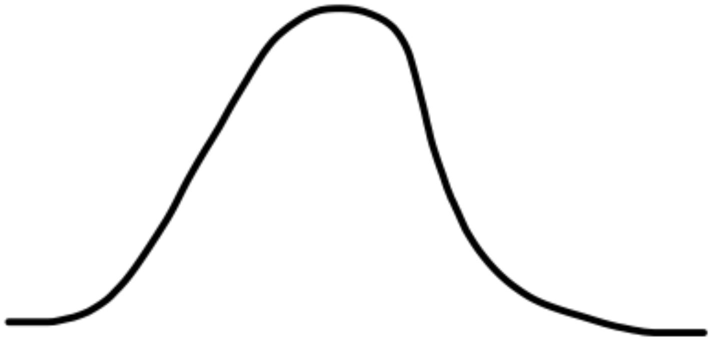


Temperature

```
Space= "
model{
  mu ~ dnorm(0,0.001)
  tau ~ dgamma(0.001,0.001)
  tau.sp~dgamma(0.001,0.001)

for(i in 1:nblocks){
  alpha.sp[i]~dnorm(0,tau.sp)
  Emu[i]=mu + alpha.sp[i]

for(t in 1:ndates){
   X[i,t] ~ dnorm(Ex[i], tau)
  }}}
"
```

Model	**mu (95% ci)**	**tau (sigma)**	**tau.sp (sigma)**	**Deviance**
Global	2.99 (2.98, 3.13)	0.85 (1.18)	NA	698
Spatial RE	2.97 (2.66, 3.28)	1.04 (0.96)	4.98 (0.20)	654

Model	mu (95% ci)	tau (sigma)	tau.sp (sigma)	Deviance
Global	2.99 (2.98, 3.13)	0.85 (1.18)	NA	698
Spatial RE	2.97 (2.66, 3.28)	1.04 (0.96)	4.98 (0.20)	654

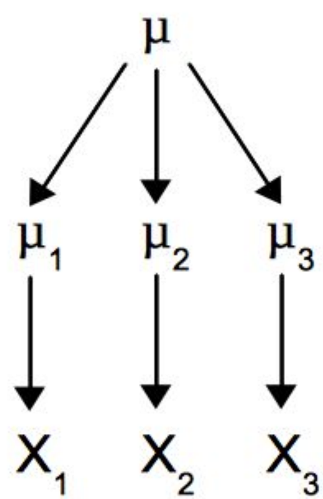


defines a distribution that can be used to infer mean at out-of-sample site.
Tighter REs (smaller variance) = more precise prediction

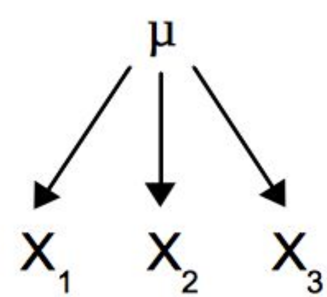
Hierarchical Bayes

Hierarchical model allows predictions about the unobserved

Out-of-sample predictions integrate over the random effects variance in known sites/years – will include more uncertainty than in-sample estimates.



Hierarchical



Shared

Hierarchical Bayes

Ecology is complex. You cannot measure everything.

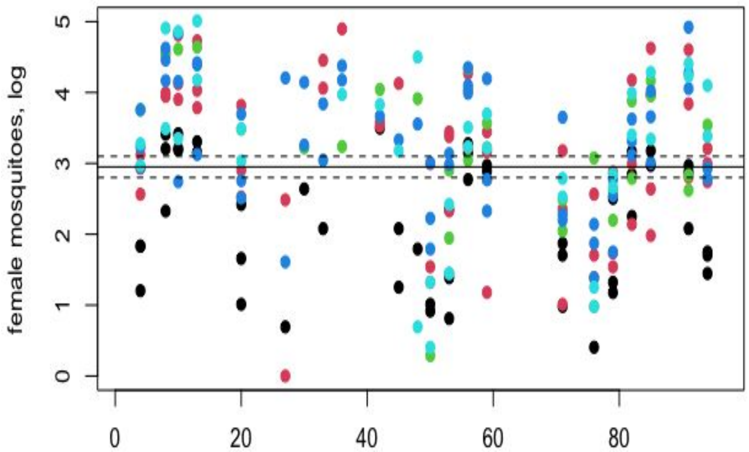
Hierarchical models can help identify important drivers and account for non-independence in sampling scheme

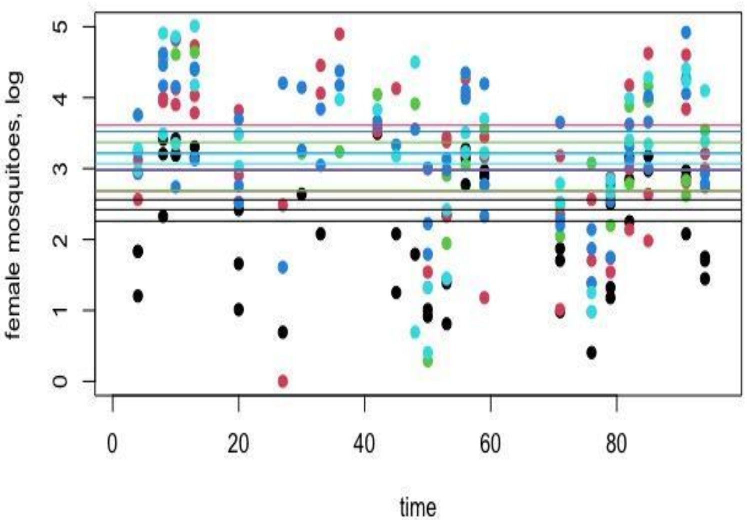
Hierarchical model allows for borrowed strength from data-rich to data-poor

- predictions about the unobserved
- integrate over the random effects variance = more real reflection of uncertainty in out-of-sample predictions

Ecological Forecasting is about characterizing uncertainty

Ecological Forecasting is about characterizing uncertainty





What will mosquito abundance be next month?

When and where should city invest in control?