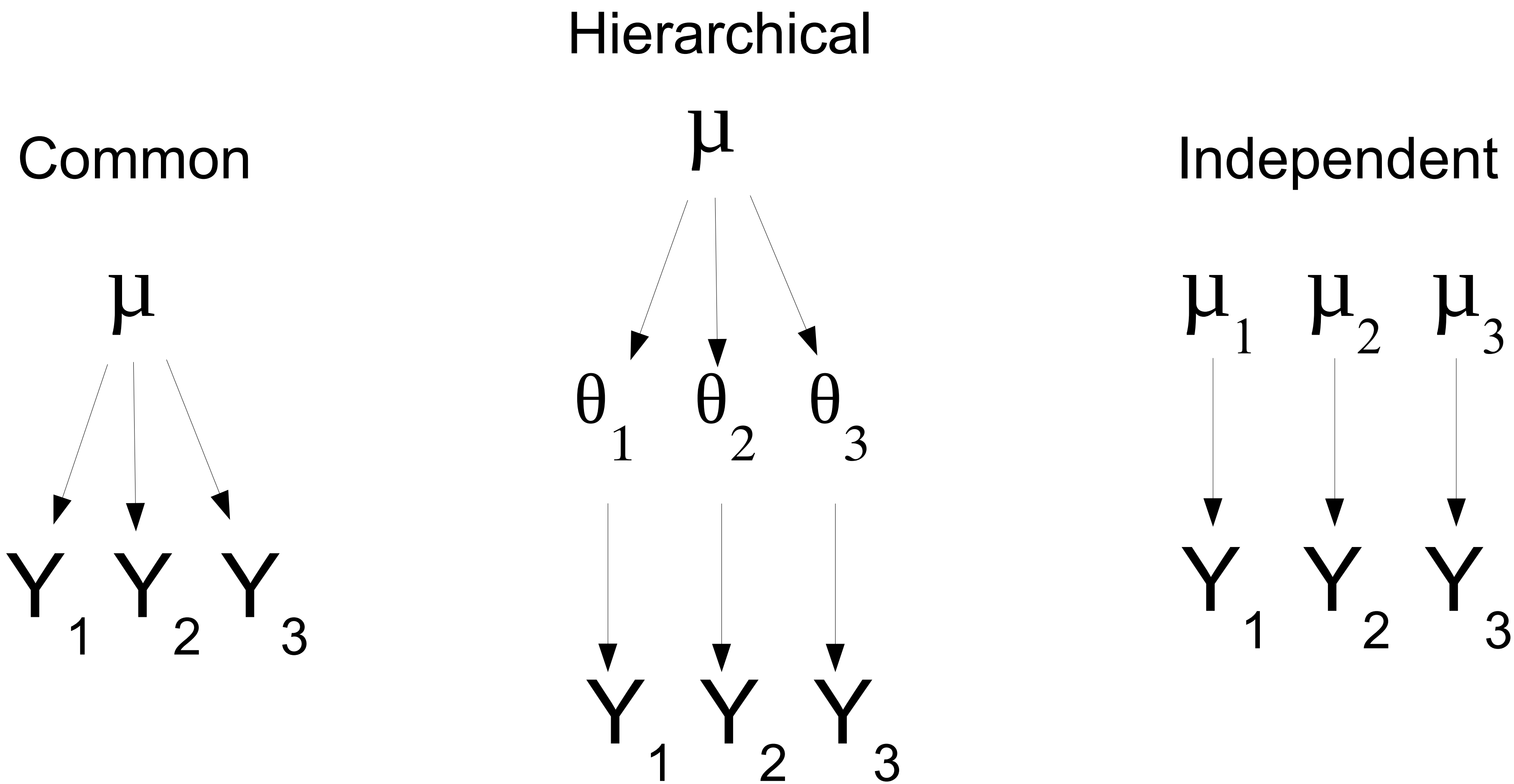
Hierarchical Bayes

Hierarchical Models



Independent Means, Common Variance

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

$$\mu_k \sim N(\mu, \tau^2)$$

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

At this point, this model is fitting each data set independently but assume the mean for each has the same prior

Independent Means, Common Variance

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

$$\mu_k \sim N(\mu, \tau^2)$$

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

For the hierarchical model, instead assume the prior contains unknown model parameters

Hierarchical Mean, Common Variance

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

$$\mu_k \sim N(\mu, \tau^2)$$

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

$$\left.\begin{aligned} &\mu \sim N(\mu_{0,} V_\mu) \\ &\tau^2 \sim IG(t_{1,} t_2) \end{aligned}\right\}$$

Then need to specify *hyperpriors* on our prior

Hierarchical Models

- Model variability in the parameters of a model
- Partition variability more explicitly into multiple terms
- Borrow strength across data sets

- Details usually in the SUBSCRIPTS
- Hierarchical with respect to parameters

Random Effects

- Common special case of Hierarchical models

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

$$\mu_k \sim N(\mu, \tau^2)$$

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

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

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

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

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

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

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

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

Random Effects

- Common special case of Hierarchical models

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

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

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

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

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

Random Effects

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

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

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

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

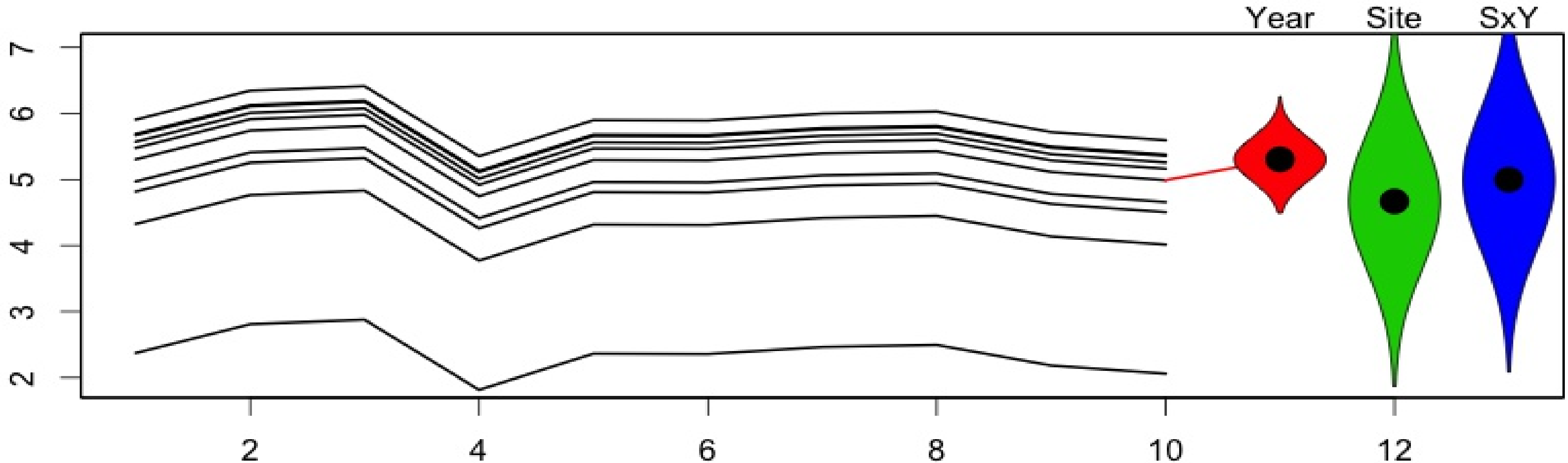
$$\tau^2 \sim IG(t_{1,} t_2)$$

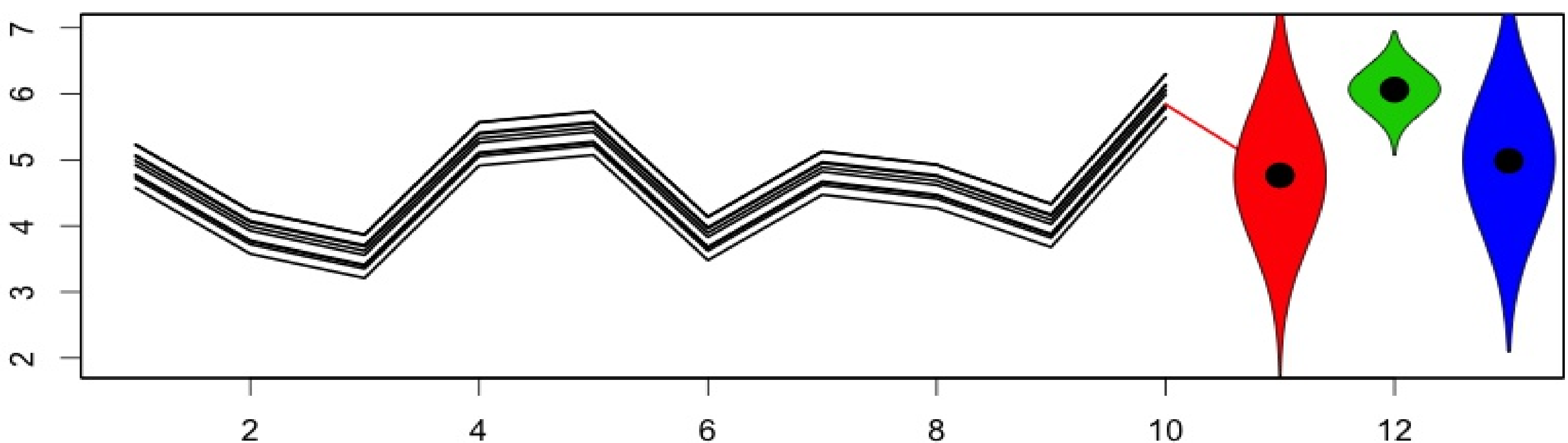
- Random effects always have mean 0
- Random effects variance attributes a portion of uncertainty to a specific source

What things can be random effects?

- Traditionally, random effects apply to aspects of the study that would not be the same if replicated
  - e.g. Plot, Block, Year, individual, etc.
  - Often used to account for a lack of independence
- Treatments and covariates of interest are usually treated as **fixed effects**
- Typically there is some degree of replication otherwise the random effect is not identifiably different from the residual “noise” term $J \sim N(0,\sigma^2)$

Why bother?
Impacts on inference...





Prediction

- Hierarchical model allows predictions about an unobserved species, sites, years, etc

- Out-of-sample predictions integrate over random effects, more uncertain than in-sample

- Posterior for new species/site/year could be updated with a relatively small number of observations

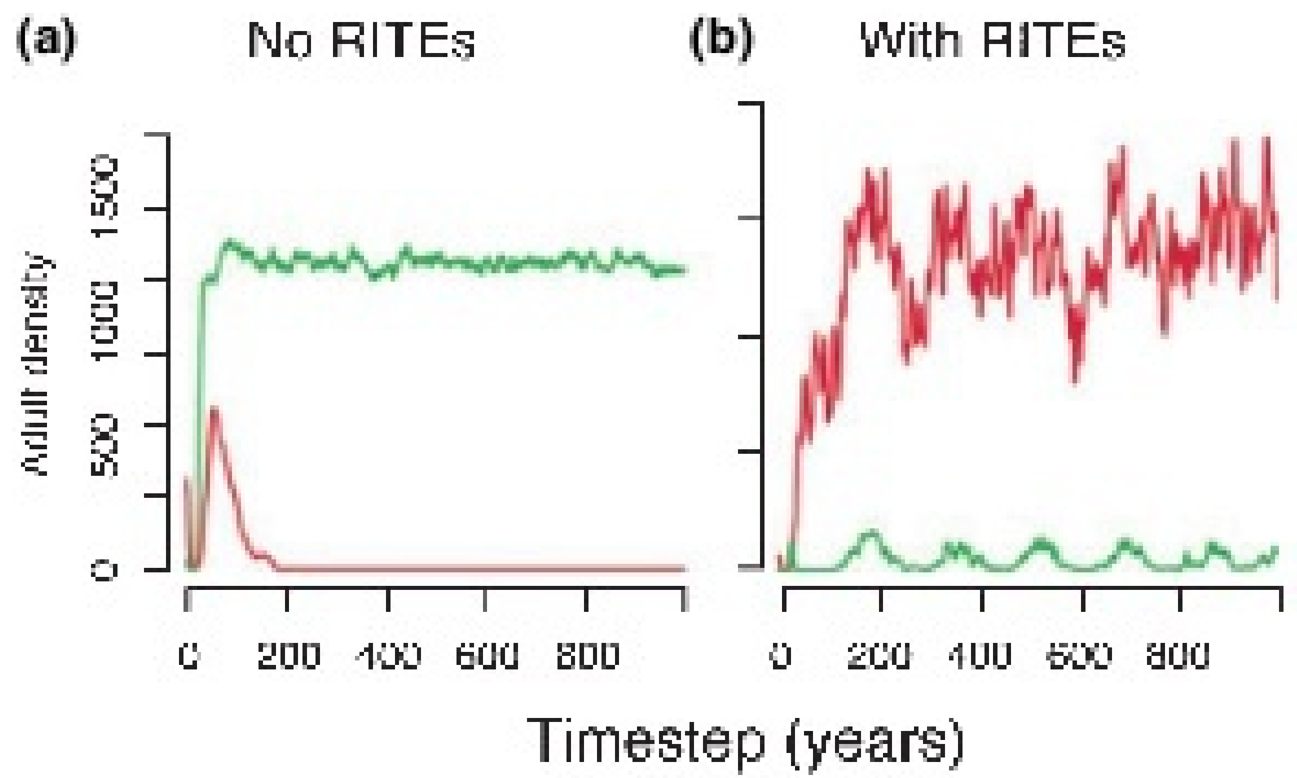
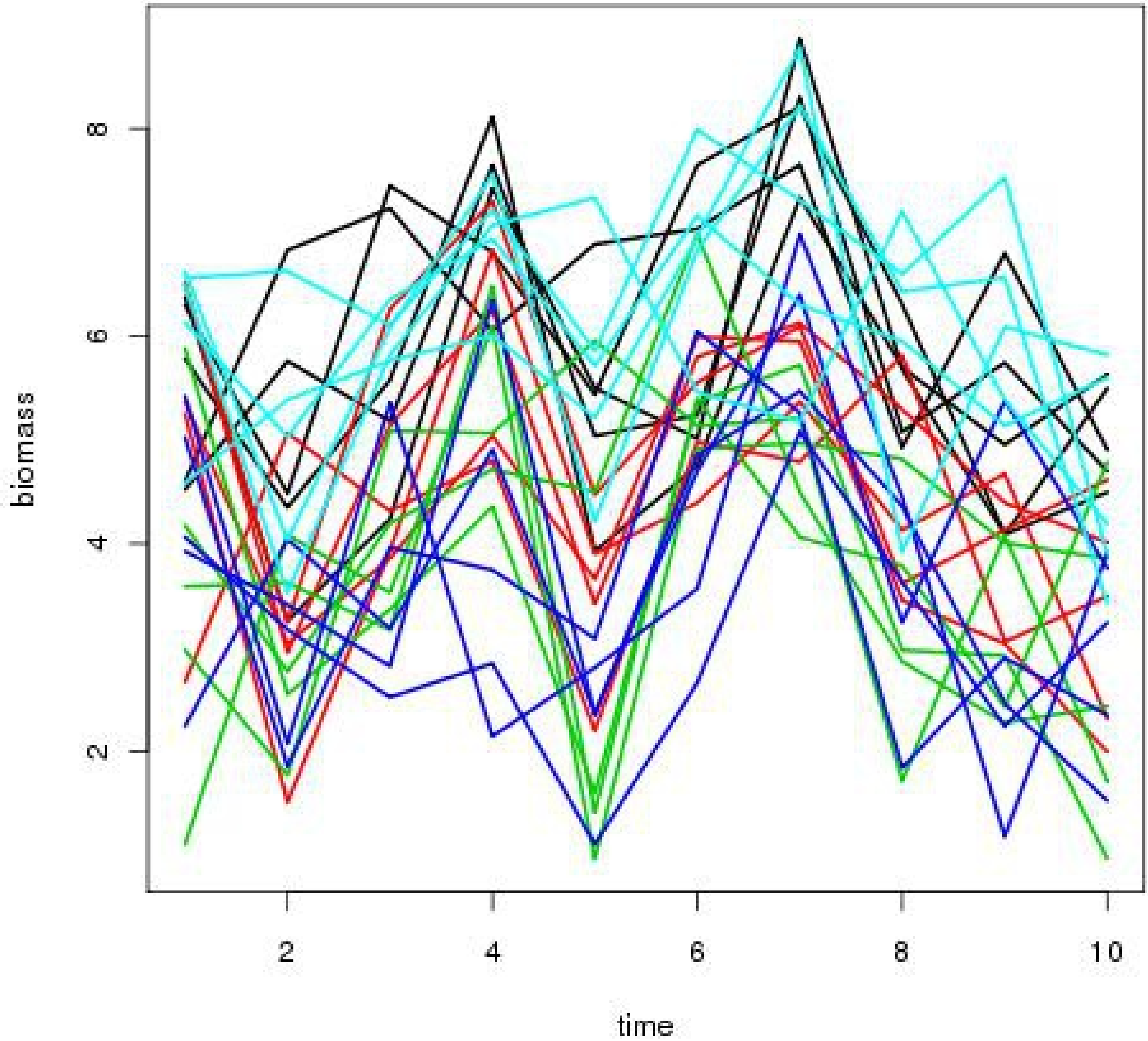


Figure 3 The impact of random individual effects (RITEs) on coexistence of two competing species. Two spatiotemporal and individual-based simulations were run using recruitment processes that are parameterized with data, summarized in Fig. 1. Panel (a) is the traditional approach having deterministic species differences and stochasticity in time, but no within-population heterogeneity, reflecting that fact the green species is the deterministic winner (Fig. 1a). Population heterogeneity in (b) means that green is not the deterministic winner, but rather both species win with some probability.

Start Simple

Progressively
Add Complexity

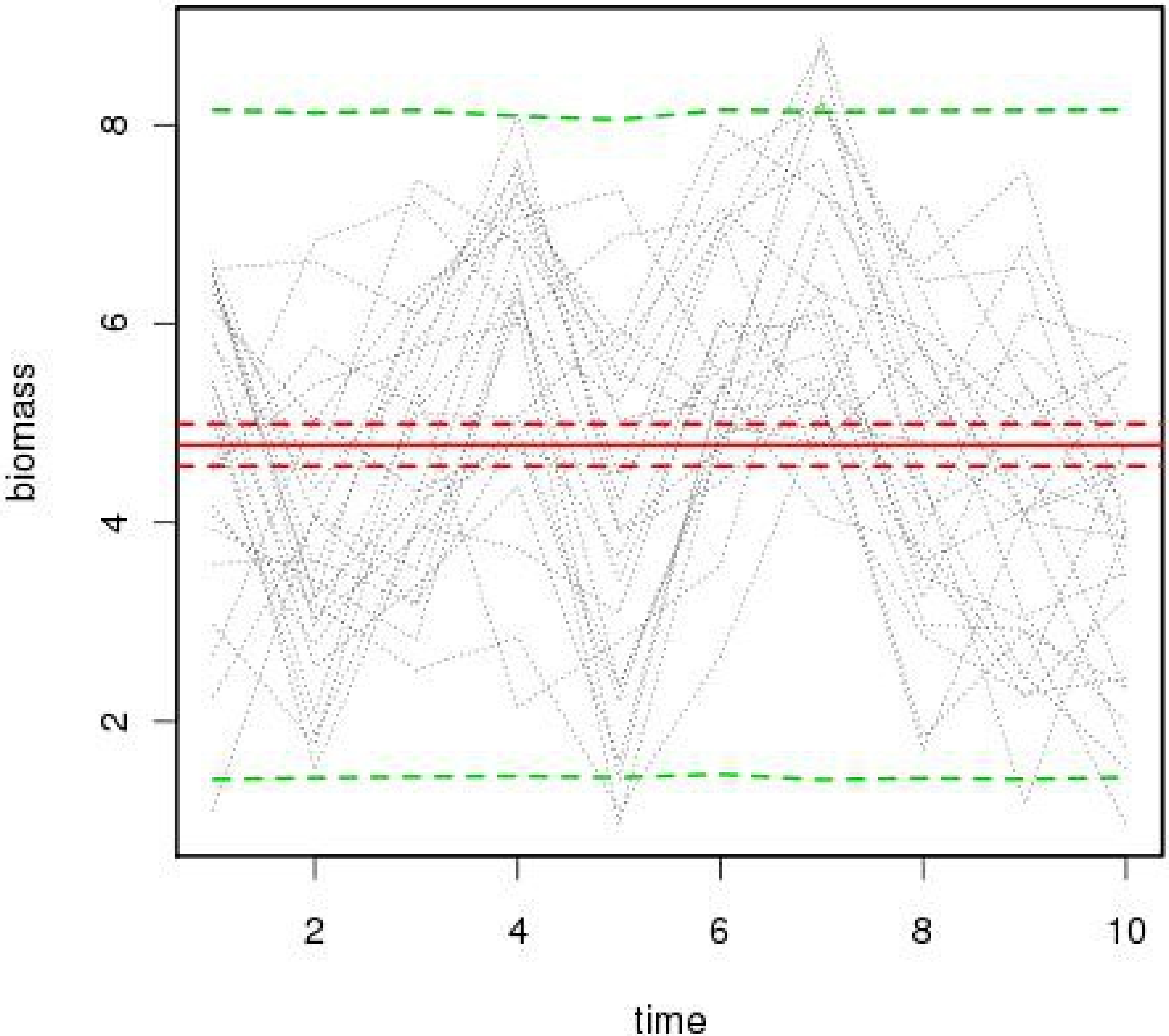
Example: Biomass by Block and Time



Model 1: Global Mean

```
model{
  mu ~ dnorm(0,0.001)                ## priors
  sigma ~ dgamma(0.001,0.001)

  for(t in 1:nt){          ## time
    for(b in 1:nb){        ## block
      for(i in 1:nrep){    ## individual
        x[t,b,i] ~ dnorm(mu,sigma)
      }
    }
  }
}
```

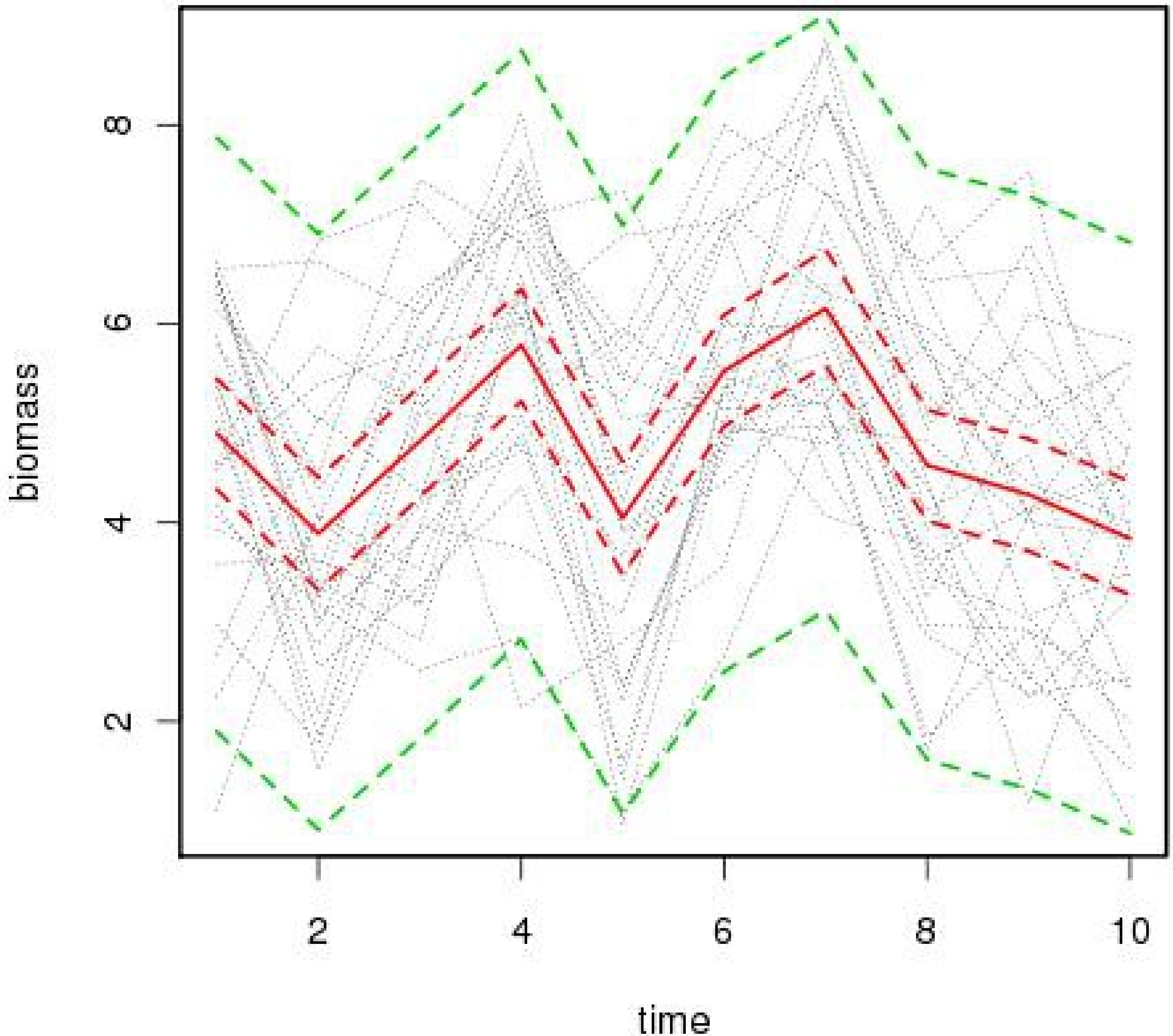


biomass
2
4
6
8
2
4
6
8
10
time

Model 2: Random Temporal Effect

```
model{
  mu ~ dnorm(0,0.001)                 ## priors
  sigma ~ dgamma(0.001,0.001)
  for(t in 1:nt){alpha.t[t] ~ dnorm(0,tau.t)}
  tau.t ~ dgamma(0.001,0.001)         ## hyperprior

  for(t in 1:nt){
    Ex[t] <- mu + alpha.t[t]          ## process model
    for(b in 1:nb){
      for(i in 1:nrep){
        x[t,b,i] ~ dnorm(Ex[t],sigma)  ## data model
      }
    }
  }
}
```

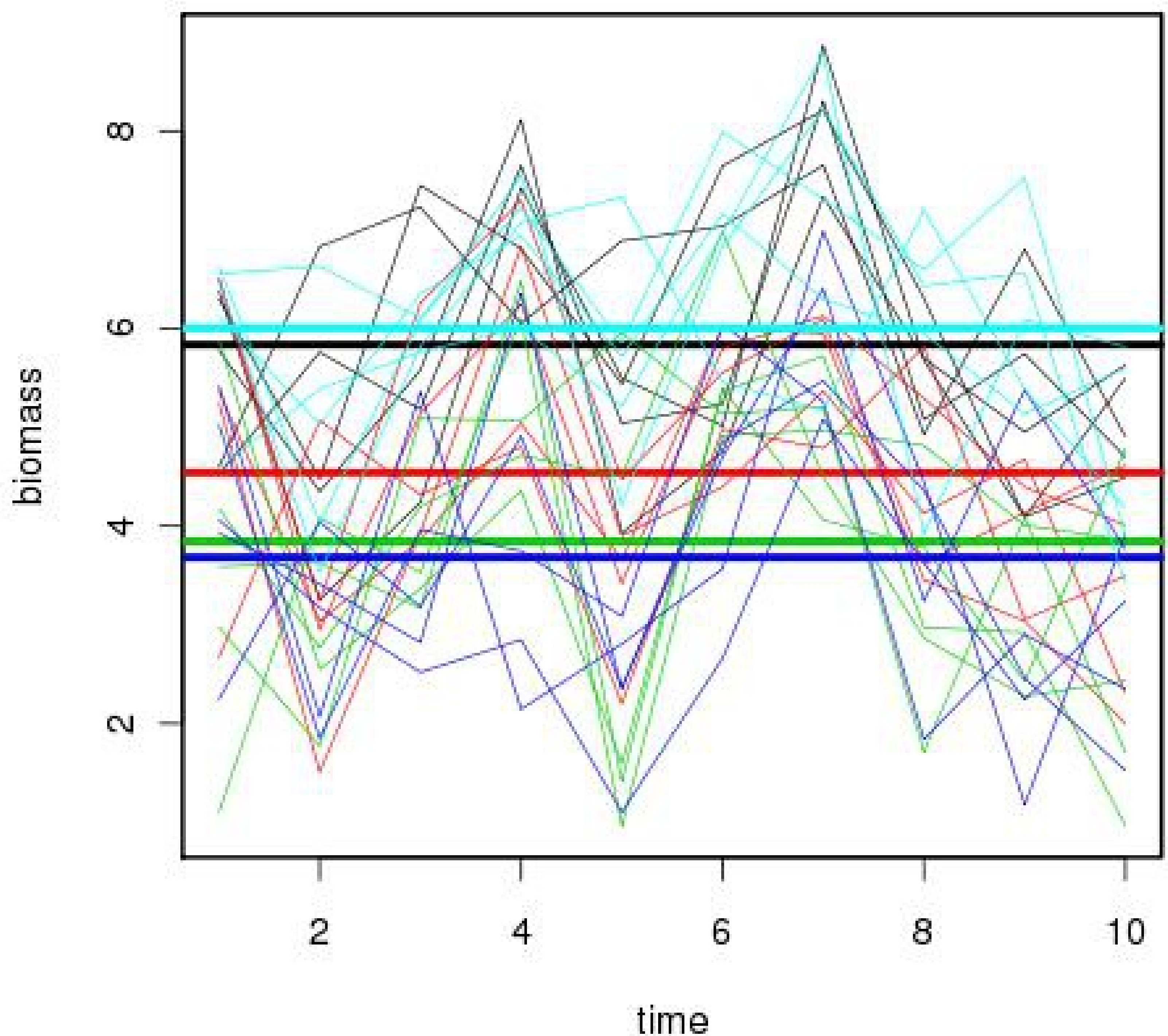


biomass
8
6
4
2
time
2
4
6
8
10

Model 3: Random Block Effect

```
model{
  mu ~ dnorm(0,0.001)            ## priors
  sigma ~ dgamma(0.001,0.001)
  tau.b ~ dgamma(0.001,0.001)
  for(b in 1:nb){ alpha.b[b] ~ dnorm(0,tau.b)}

  for(b in 1:nb){
    Ex[b] <- mu + alpha.b[b]
    for(t in 1:nt){
      for(i in 1:nrep){
        x[t,b,i] ~ dnorm(Ex[b],sigma)
      }
    }
  }
}
```



biomass
time
2
4
6
8
10

Model 4: Random Block & Time

```
model{

  mu ~ dnorm(0,0.001)              ## priors
  sigma ~ dgamma(0.001,0.001)
  tau.b ~ dgamma(0.001,0.001)
  tau.t ~ dgamma(0.001,0.001)
  for(t in 1:nt){alpha.t[t] ~ dnorm(0,tau.t) }
  for(b in 1:nb){alpha.b[b] ~ dnorm(0,tau.b) }

  for(t in 1:nt){
    for(b in 1:nb){
      Ex[t,b] <- mu + alpha.b[b] + alpha.t[t]
      for(i in 1:nrep){
        x[t,b,i] ~ dnorm(Ex[t,b],sigma)
      }
    }
  }
}
```

Summary Table

Model	mu	sigma	tau.t	tau.b	DIC
Global Mean	4.78 (0.11)	2.92 (0.27)			977.9
Random Time	4.75 (0.33)	2.23 (0.21)	0.97 (0.64)		919.8
Random Block	4.82 (0.69)	1.92 (0.18)		2.36 (3.62)	878.0
Random B x T	4.85 (0.75)	0.84 (0.08)	1.31 (0.67)	0.80 (0.60)	766.8

Random Effects Linear Model

$\mu_{i,k} = \underbrace{X_i \beta}_{\text{Fixed Effects}} + \underbrace{\alpha_k}_{\text{Random Effect}} + \underbrace{\epsilon_{i,k}}_{\text{Residual Error}}$	**Process model**
$\epsilon_{i,k} \sim N(0, \sigma^2)$	**Data model**
$\alpha_k \sim N(0, \tau^2)$	**Random effect**
$\sigma^2 \sim IG(s_1, s_2)$	**Error variance prior**
$\beta \sim N(B_{0,} V_\beta)$	**Fixed effects prior**
$\tau^2 \sim IG(t_{1,} t_2)$	**Random effects variance prior**

Explaining unexplained variance

- 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/process may explain some portion of this variance, but there's always something you didn't measure
- Sometimes additional fixed effects not justified (model selection)

Example: Year effects

- Consider the number of new young produced per adult female from population of birds
- Suppose adding a year effect shows significant year-to-year variability that is coherent through the whole population
- Based on the estimates of the year effects, could look for additional covariates that correlate with these values (e.g. different climate variables) without having to rerun the whole model
- Could refine the model to add additional drivers

Modeling Uncertainty

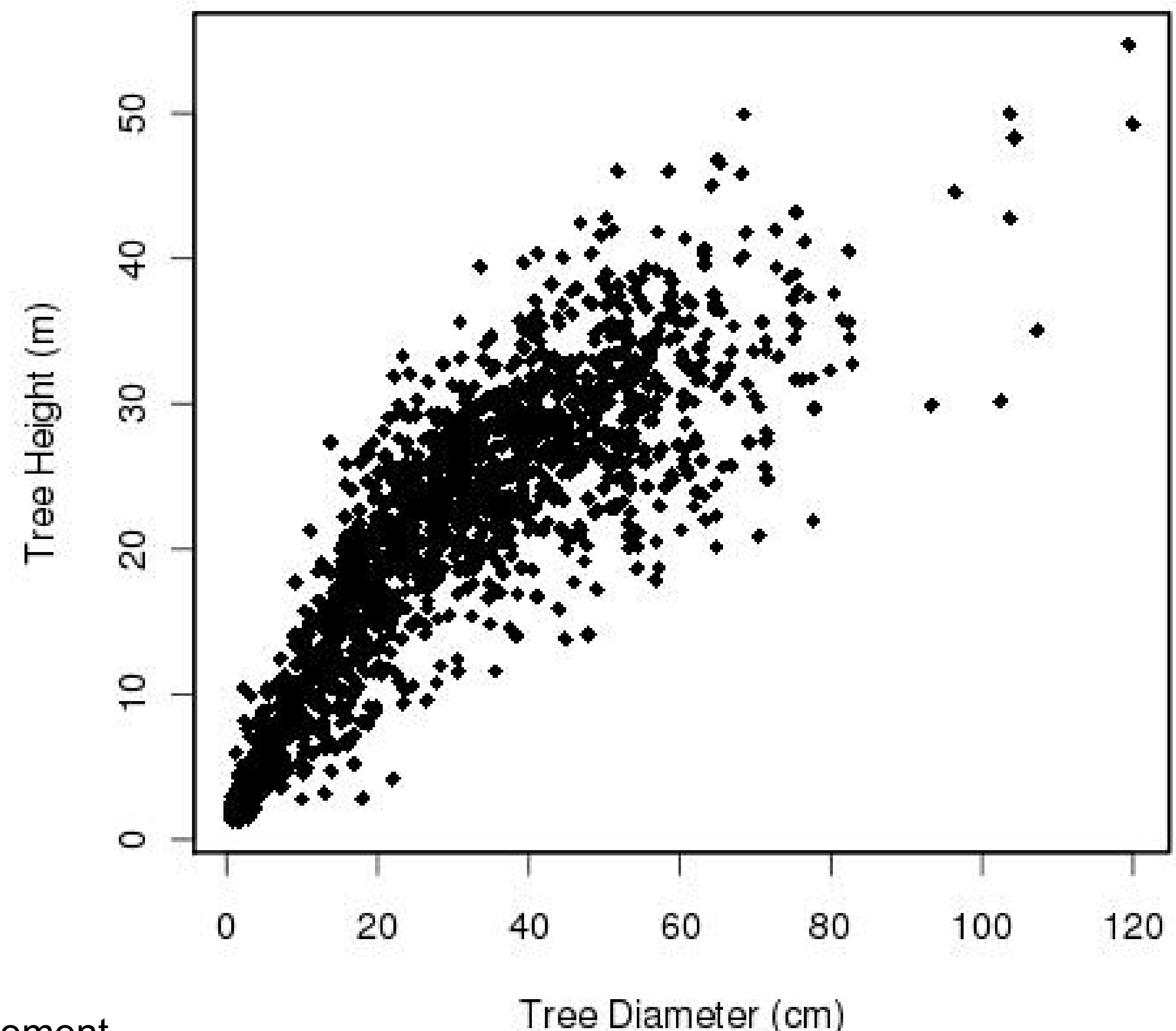
- Overall take home message:

  The proper accounting of uncertainty can be JUST AS IMPORTANT to making valid inference from your model as the process model and covariates

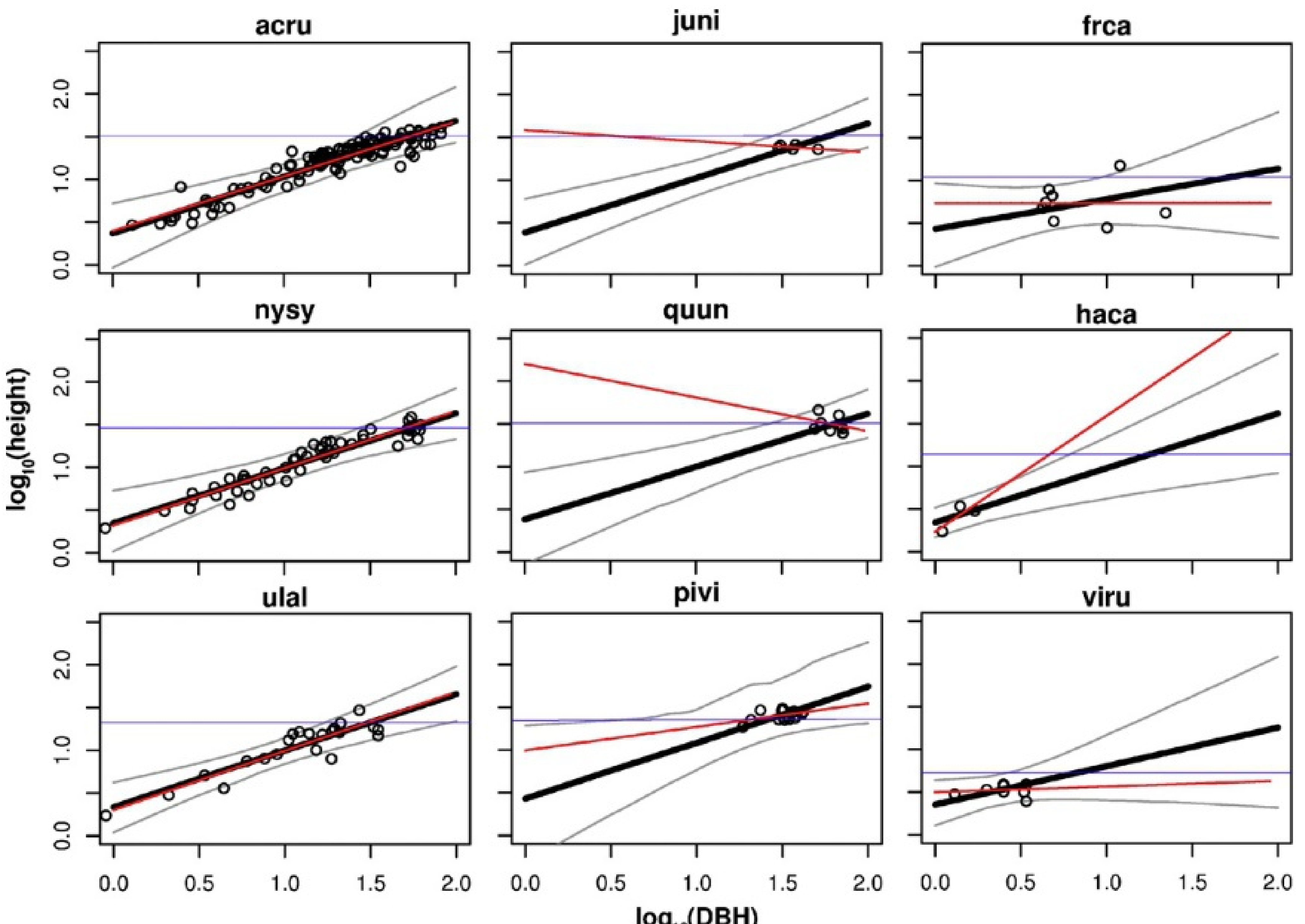
- Random effects are used to account for the impacts of unmeasured/unmeasurable covariates

Example: Tree Allometries

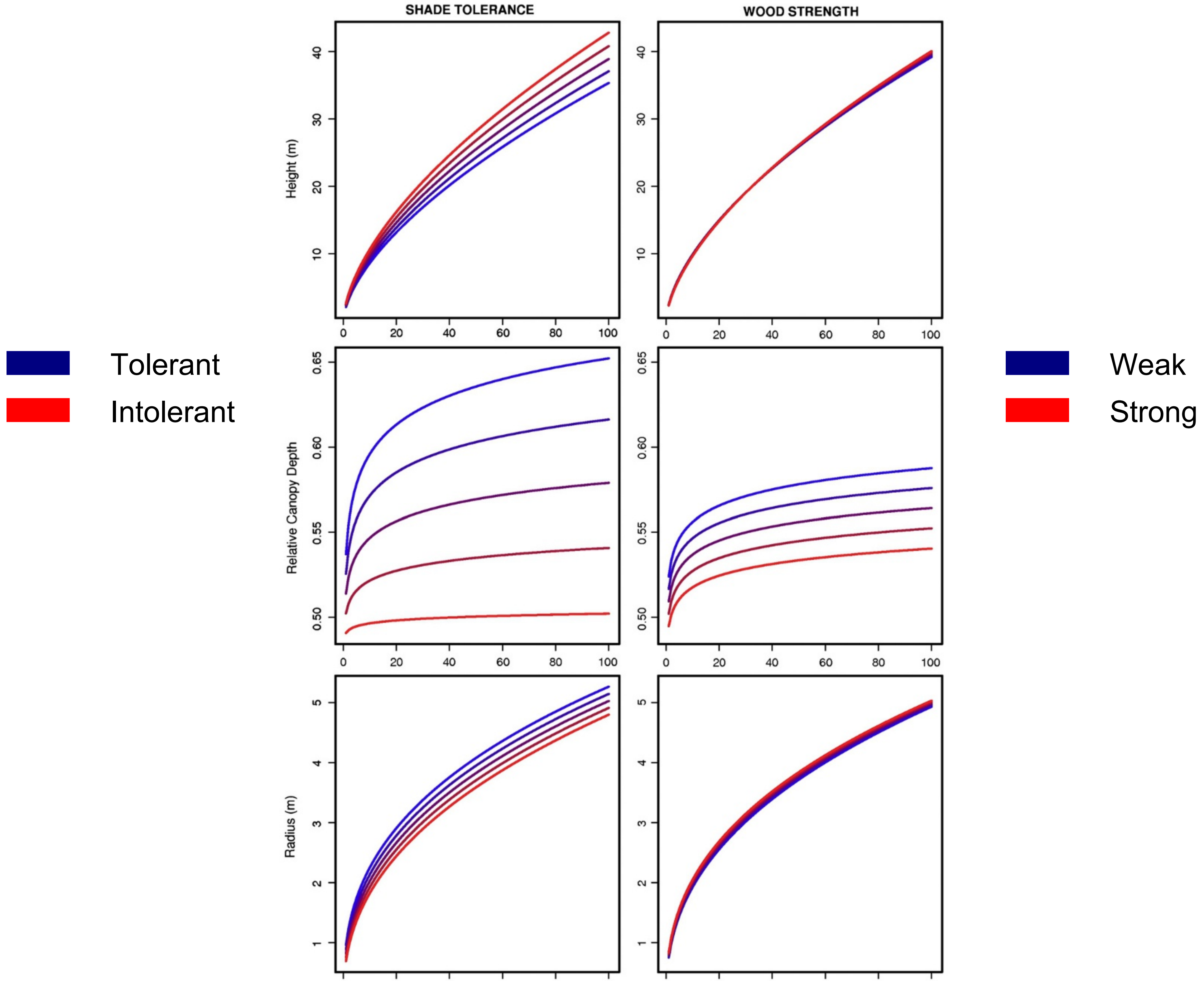
- 53 spp
- 1691 obs
- Mixed temperate
- North Carolina



Dietze et al 2008
Forest Ecology & Management



acru
juni
frca
nysy
quun
haca
ulal
pivi
viru
log10(height)
log10(DBH)
0.0
0.5
1.0
1.5
2.0



SHADE TOLERANCE
WOOD STRENGTH
Height (m)
Relative Canopy Depth
Radius (m)
Tolerant
Intolerant
Weak
Strong

Nonlinear Hierarchical Models

- Often takes more thought to decide which parameters you consider random and which are fixed
- Setting all parameters to random can often result in unidentifiablity
- Inclusion of covariates also challenging

Example: Coho salmon reproduction

- Beverton-Holt pop'n model with DD

$$r_t = \frac{s_t}{1/\alpha + s_t / r_m} e^{\epsilon_t}$$

- Consider
  - s = # of spawning Coho salmon
  - r = # of recruits
- Reproduction varies by stream?
  - How can we incorporate random stream effect?

$$r_{i,t} = \frac{s_t}{1/\alpha_i + s_{i,t}/r_{m,i}} e^{\epsilon_{i,t}}$$

Process model

$$\epsilon_{i,t} \sim N(0, \sigma^2)$$

Residual error

$$r_{i,m} \sim N(\mu_r, \tau_r^2)$$

$$\alpha_i \sim N(\mu_\alpha, \tau_\alpha^2)$$

Stream-level parameters

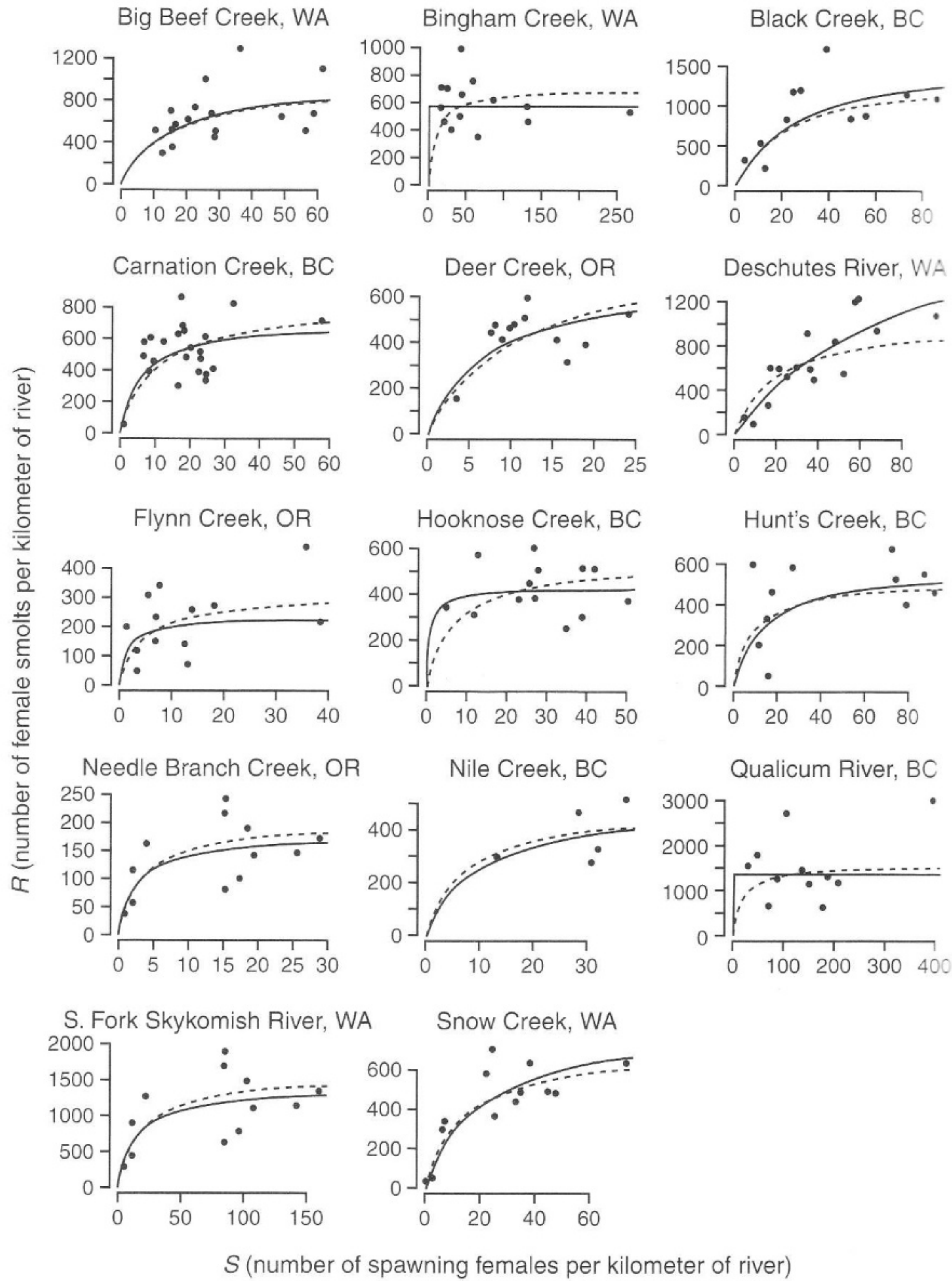
$$\mu_r \sim N(r_{0,} V_r)$$

$$\mu_\alpha \sim N(\alpha_{0,} V_\alpha)$$

Across stream parameters

$$\tau_\alpha, \tau_r \sim IG(s_{1,} s_2)$$

Across stream variance


Big Beef Creek, WA
Bingham Creek, WA
Black Creek, BC
Carnation Creek, BC
Deer Creek, OR
Deschutes River, WA
Flynn Creek, OR
Hooknose Creek, BC
Hunt's Creek, BC
Needle Branch Creek, OR
Nile Creek, BC
Qualicum River, BC
S. Fork Skykomish River, WA
Snow Creek, WA
R (number of female smolts per kilometer of river)
S (number of spawning females per kilometer of river)

Scale dependence in the effects of leaf ecophysiological traits on photosynthesis: Bayesian parameterization of photosynthesis models

Xiaohui Feng[1] and Michael Dietze[2]

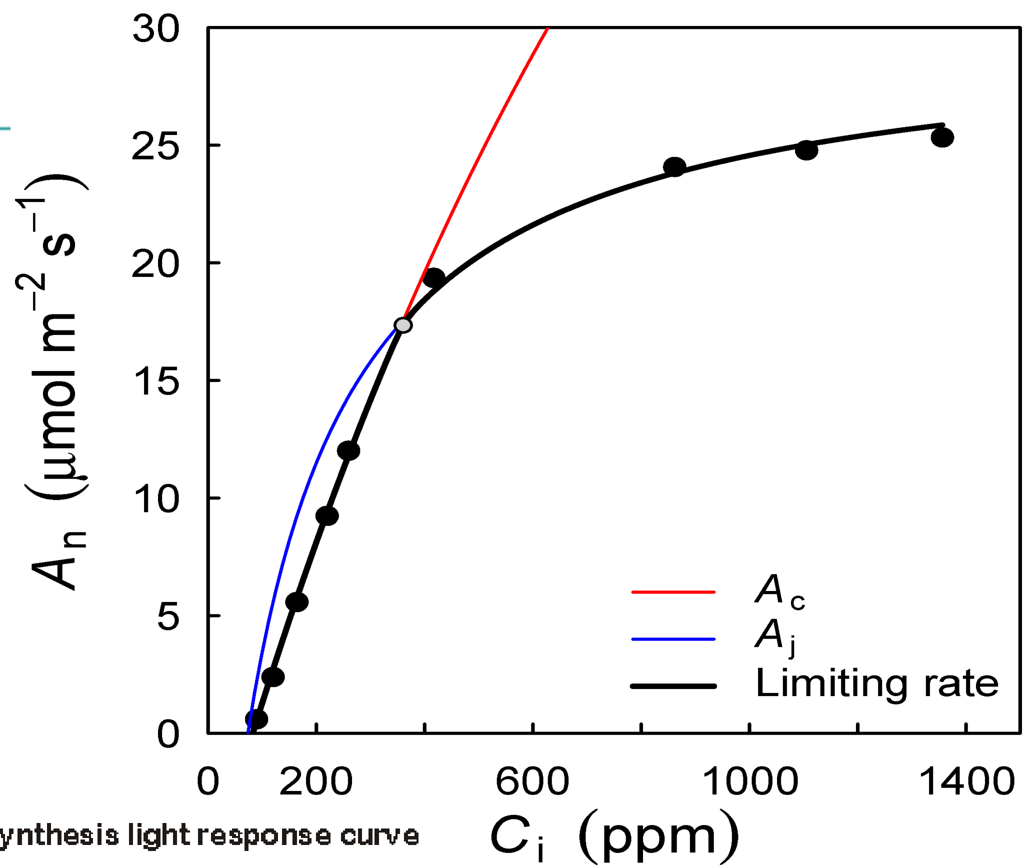
FvCB Model

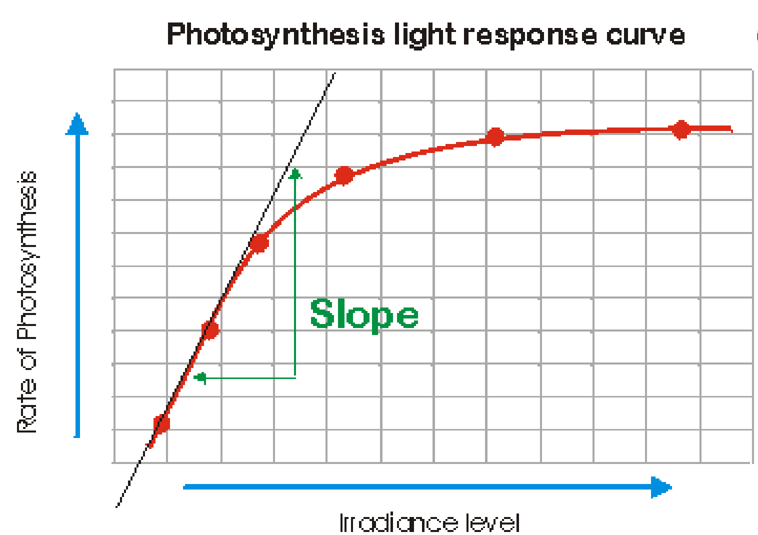
$$A_n^{(m)} = \min\{A_v, A_j\} - R_d$$

$$A_v = V'_{cmax} \frac{C_i - \Gamma^*}{C_i + K_c\left(1 + \frac{O}{K_O}\right)}$$

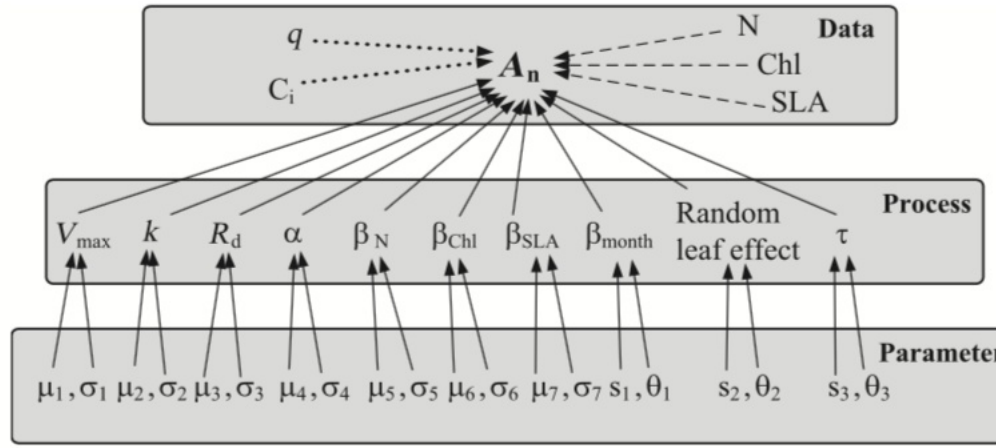
$$A_j = \frac{J(C_i - \Gamma^*)}{4C_i + 8\Gamma^*}$$

$$J = \frac{\alpha' q}{\sqrt{\left(1 + \frac{\alpha'^2 q^2}{J_{max}{}^2}\right)}}$$





$$A_n^{(o)} \sim N(A_n^{(m)}, \tau^2)$$



$$A_n^{(m)} = \min\{A_v, A_j\} - R_d$$

$$A_v = V'_{cmax} \frac{C_i - \Gamma^*}{C_i + K_c\left(1 + \frac{O}{K_O}\right)}$$

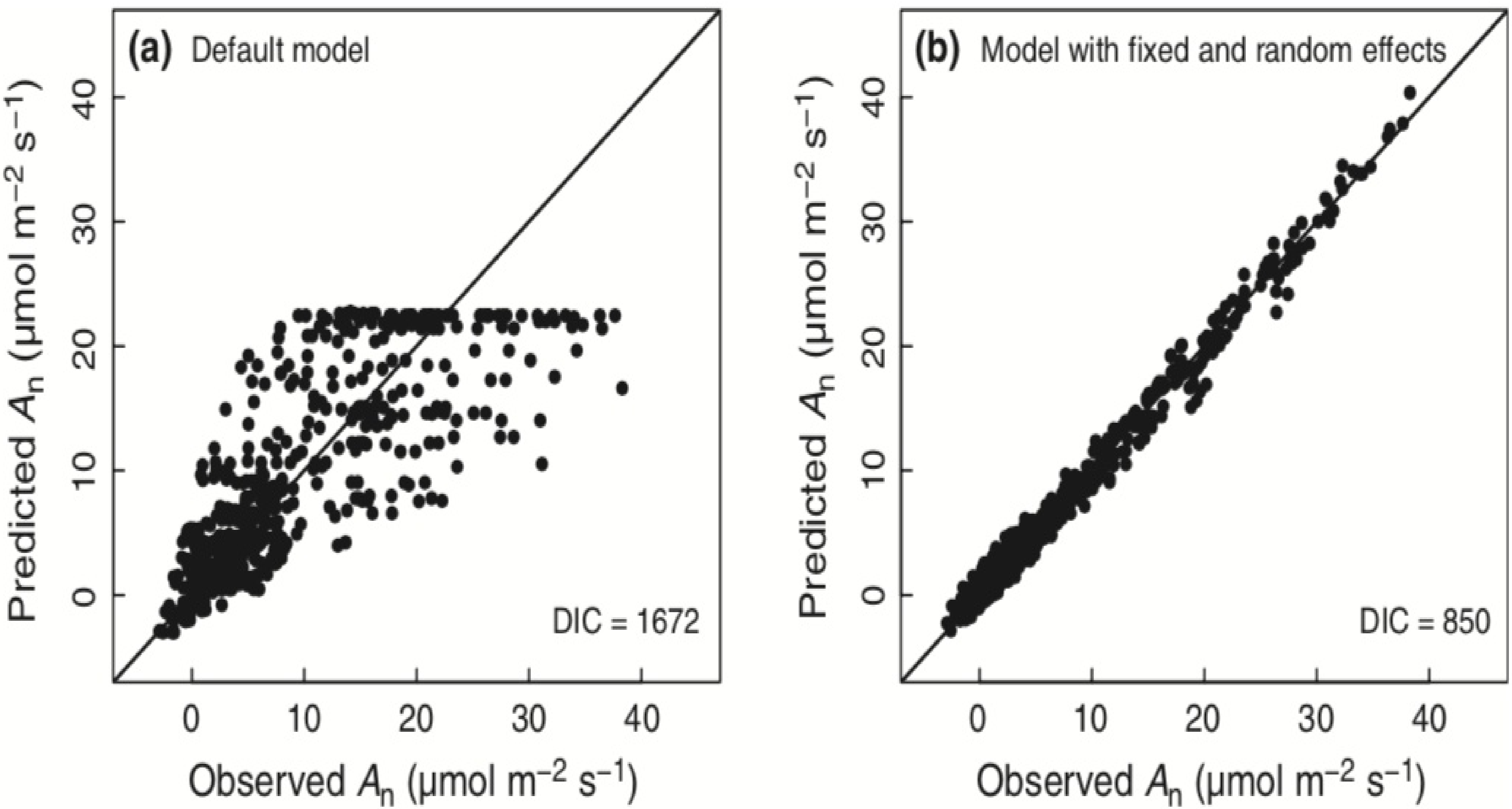
$$V'_{cmax} = \underline{V_{cmax} + \beta_N(N - \overline{N}) + \beta_{mon} + \upsilon_{leaf}}$$

$$A_j = \frac{J(C_i - \Gamma^*)}{4C_i + 8\Gamma^*}$$

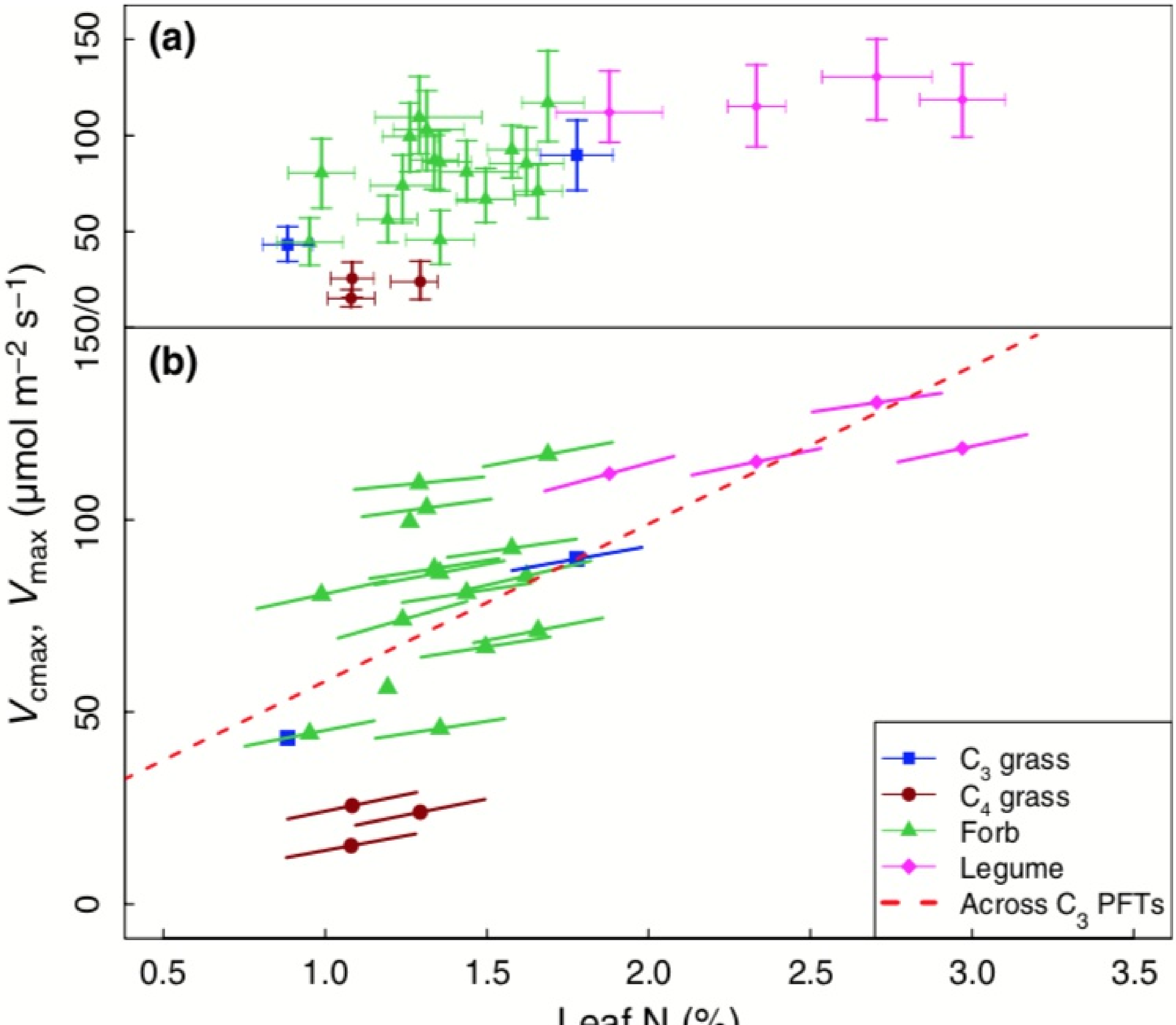
$$J = \frac{\alpha' q}{\sqrt{\left(1 + \frac{\alpha'^2 q^2}{J_{max}{}^2}\right.}}$$

25 prairie species
2 years
Monthly (within growing season)
3-5 replicates/species

$$\alpha' = \alpha + \underline{\beta_{Chl}(Chl - \overline{Chl}) + \beta_{SLA}(SLA - \overline{SLA}) + \alpha_{leaf}}$$



(a) Default model
DIC = 1672
(b) Model with fixed and random effects
DIC = 850
Predicted A_n (µmol m^{-2} s^{-1})
Observed A_n (µmol m^{-2} s^{-1})
0
10
20
30
40



(a)
(b)
150
100
50
150/0
100
50
0
V_{cmax}, V_{max} (μmol m^{-2} s^{-1})
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Leaf N (%)
C_3 grass
C_4 grass
Forb
Legume
Across C_3 PFTs

Hierarchical Models

- Model variability in the parameters of a model
- Partition variability more explicitly into multiple terms
- Borrow strength across data sets

- Details usually in the SUBSCRIPTS
- Hierarchical with respect to parameters