Nonlinear Models

and

Hierarchical Nonlinear Models

Assumption of linearity

- The final assumption of linear models that we'll address is that of linearity
  - Recall that linearity of models is wrt parameters

- “Beastiary” of model from lecture 6 (Bolker ch 3)

Assumption of linearity

- Consider any arbitrary function / process model $y = g(x|\theta_m)$
  - Choose a data model

    $$y \sim PDF(\ g(x|\theta_m)\ ,\ \theta_d\ )$$

  - If Bayesian, choose priors on θ_m & θ_d

Fitting nonlinear models

- Rarely an analytical solution
- Likelihood
  - Numerical optimization
  - LRT or Bootstrap error estimates & prediction
- Bayes
  - Metropolis-Hastings

Fitting nonlinear models

- Nothing you haven't seen / done before
- Nothing sacred about linear models

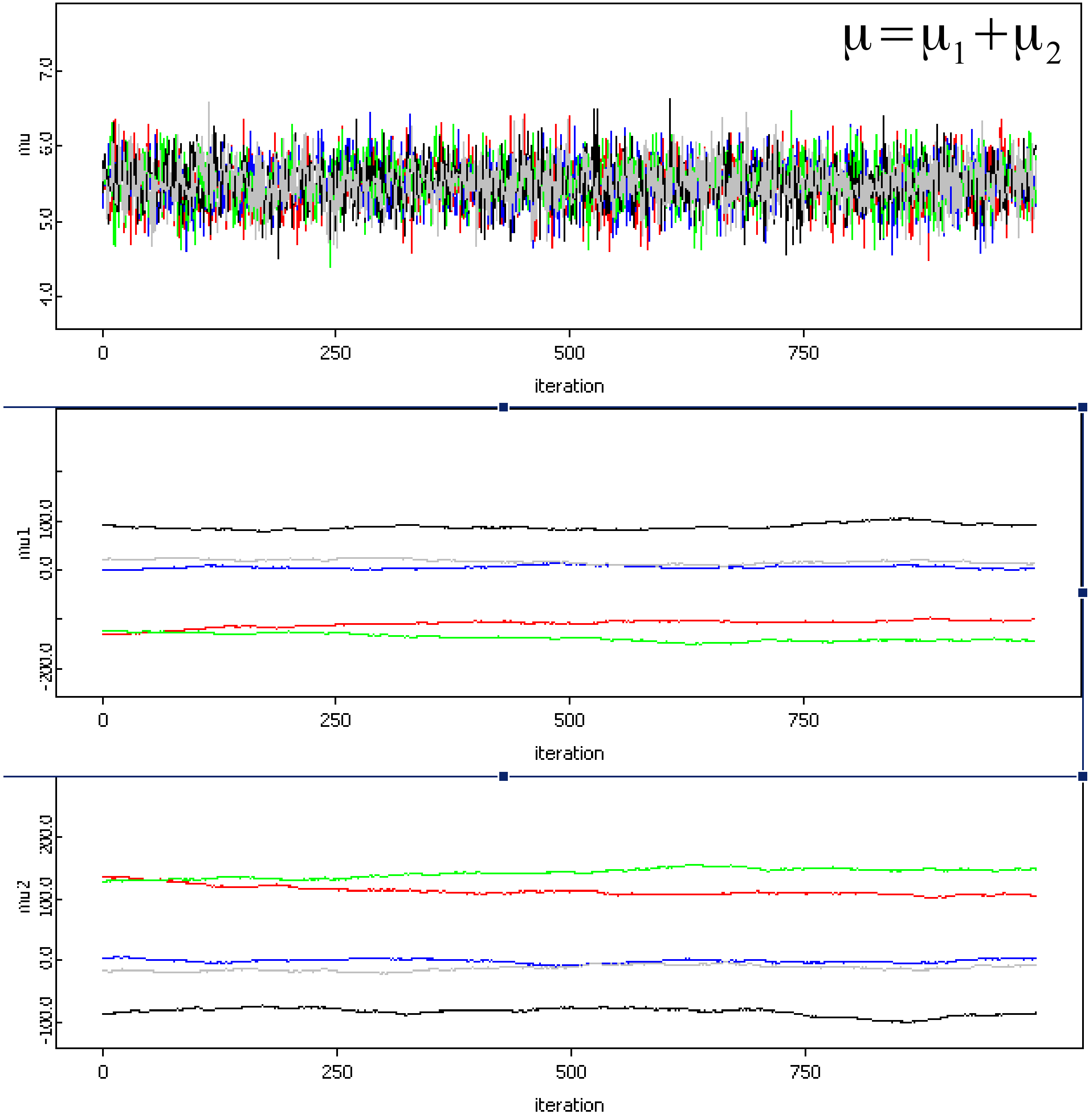
Things to watch for...

- Parameter identifiability
- Redundant parameters

$$Y = \frac{a}{b + c\,X}$$

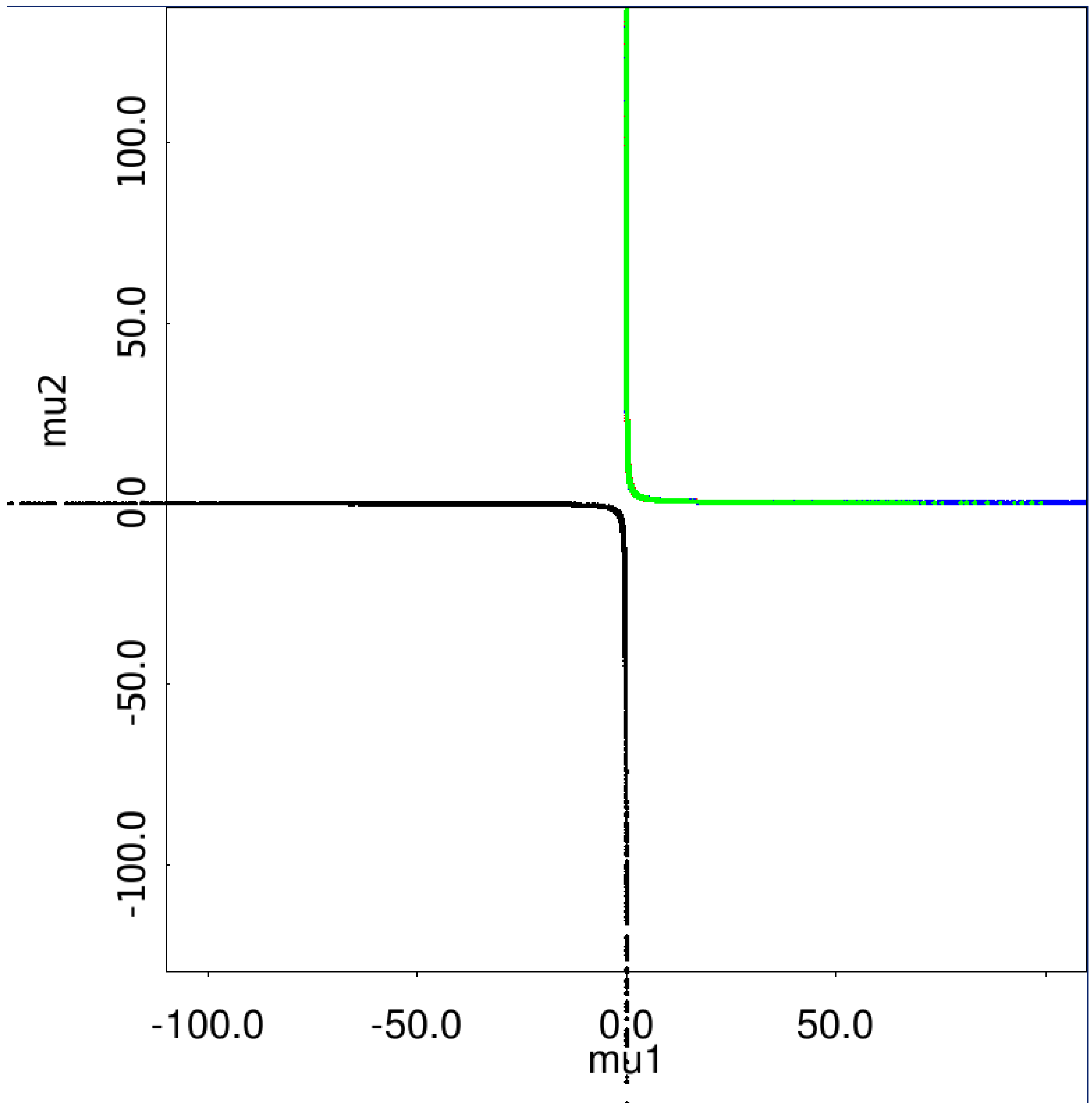
$$\downarrow$$

$$Y = \frac{1}{b' + c'\,X}$$



Things to watch for...

- Odd correlations between parameters



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?

Alternatives

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

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

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

$$r_{i,t} = \frac{s_t}{1/\alpha + s_{i,t}/r_m} e^{\epsilon_{i,t} + \underline{\theta_i}}$$

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

Process model

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

Residual error

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

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

Stream-level parameters

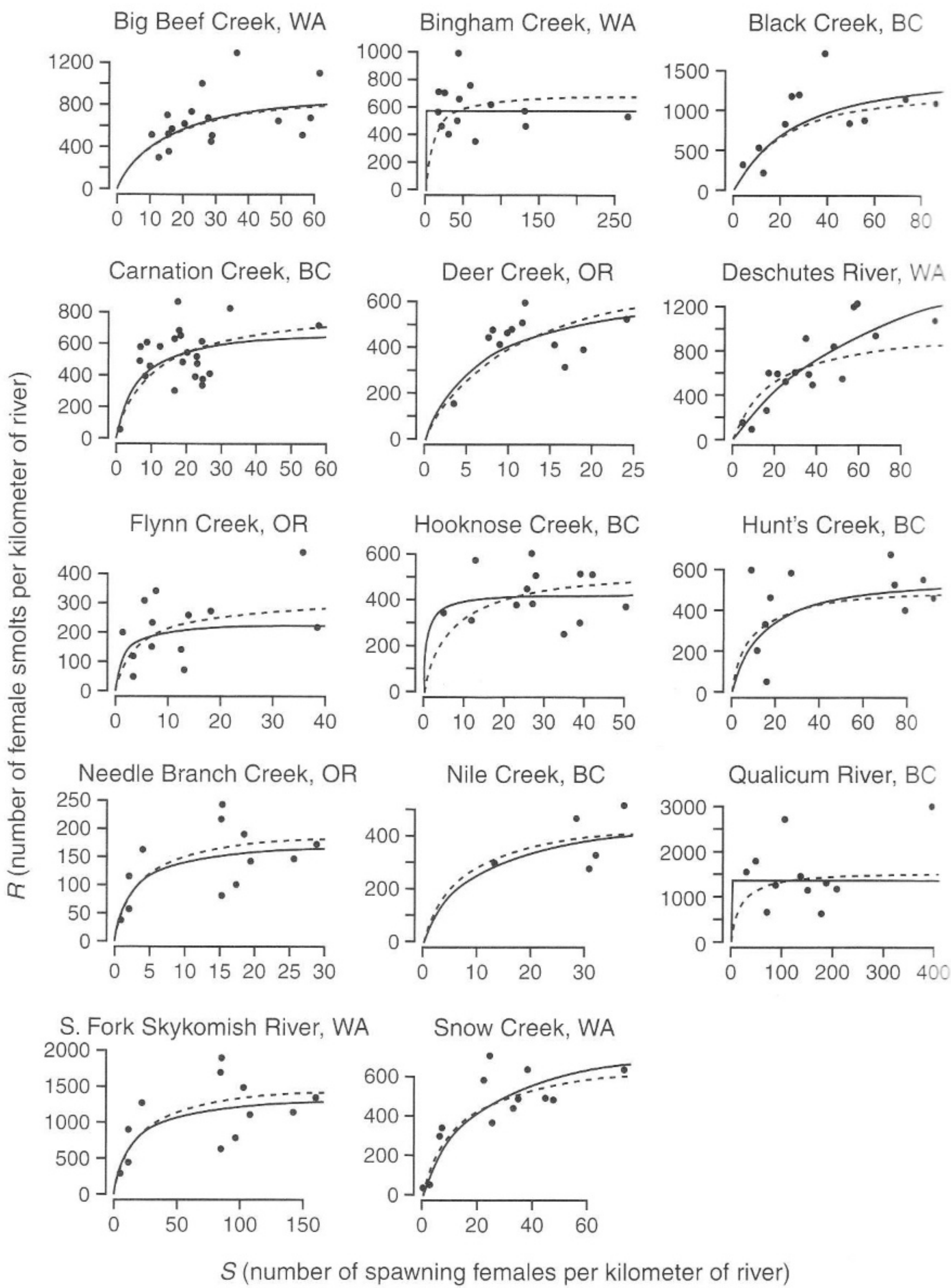
$$\mu_r \sim N(r_{0,} V_r)$$

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

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]

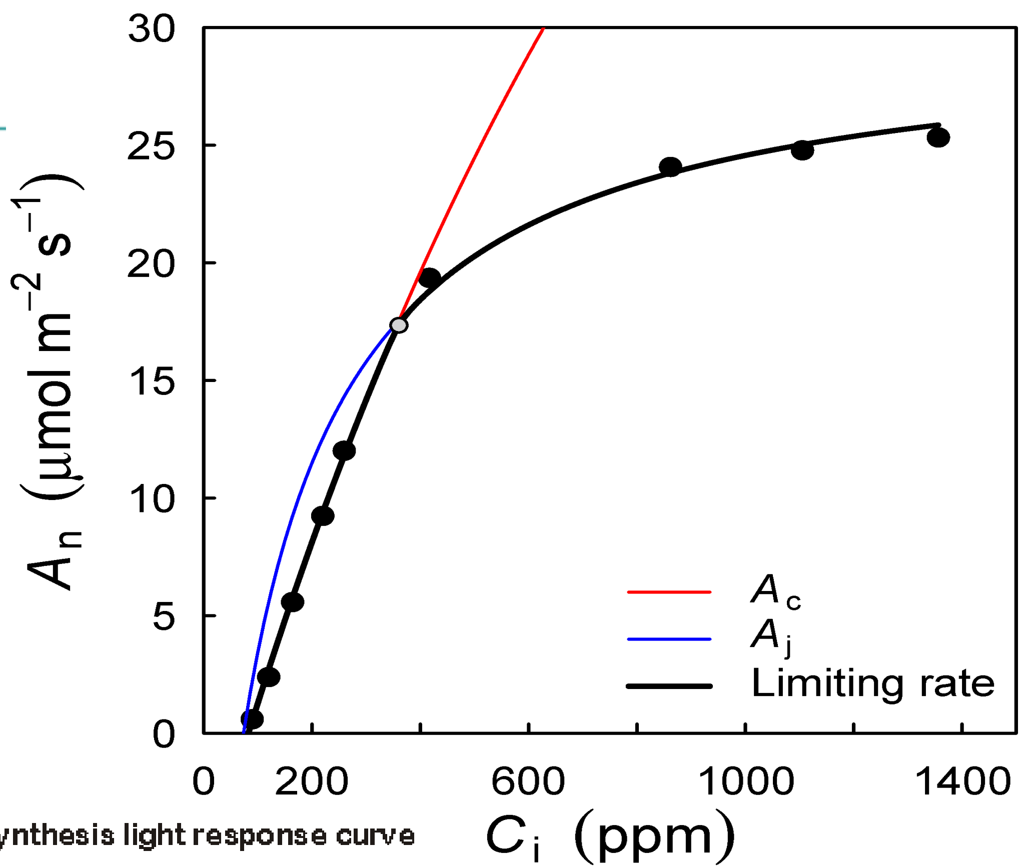
FvCB Model

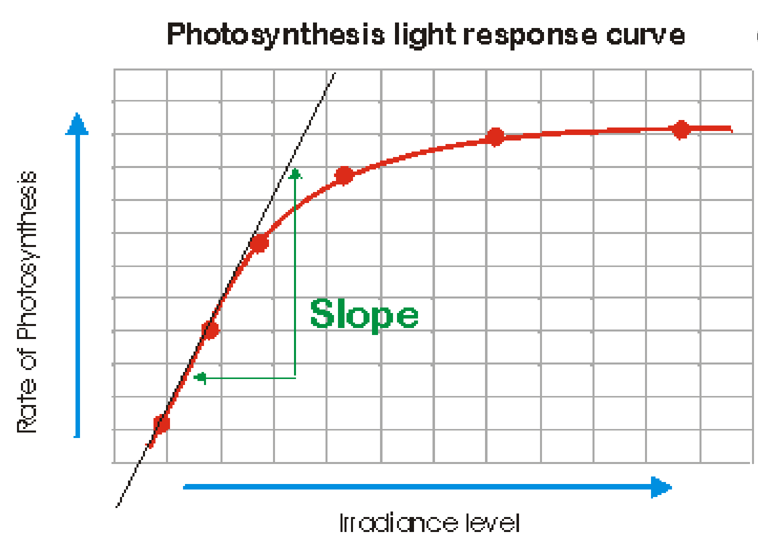
$$A_n^{(m)} = \min\{A_v, A_j\} - R_d$$

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

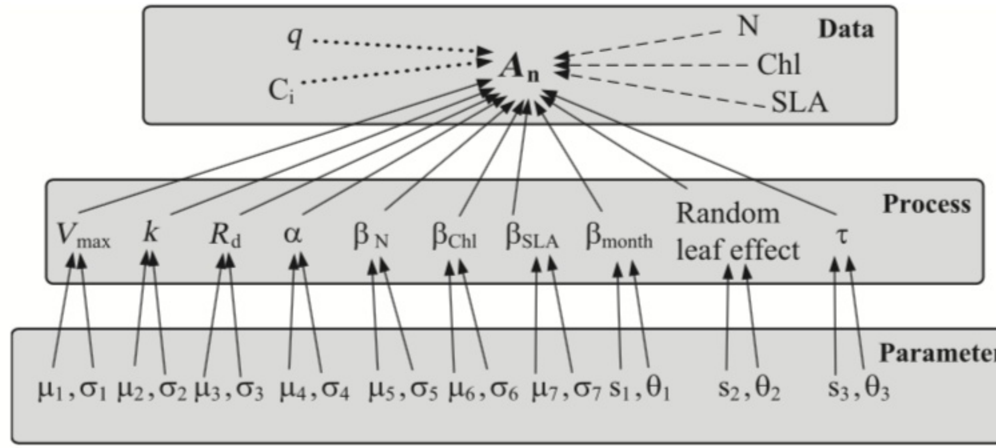
$$A_j = \frac{J(C_i - \Gamma^*)}{4C_i + 8\Gamma^*}$$

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





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



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

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

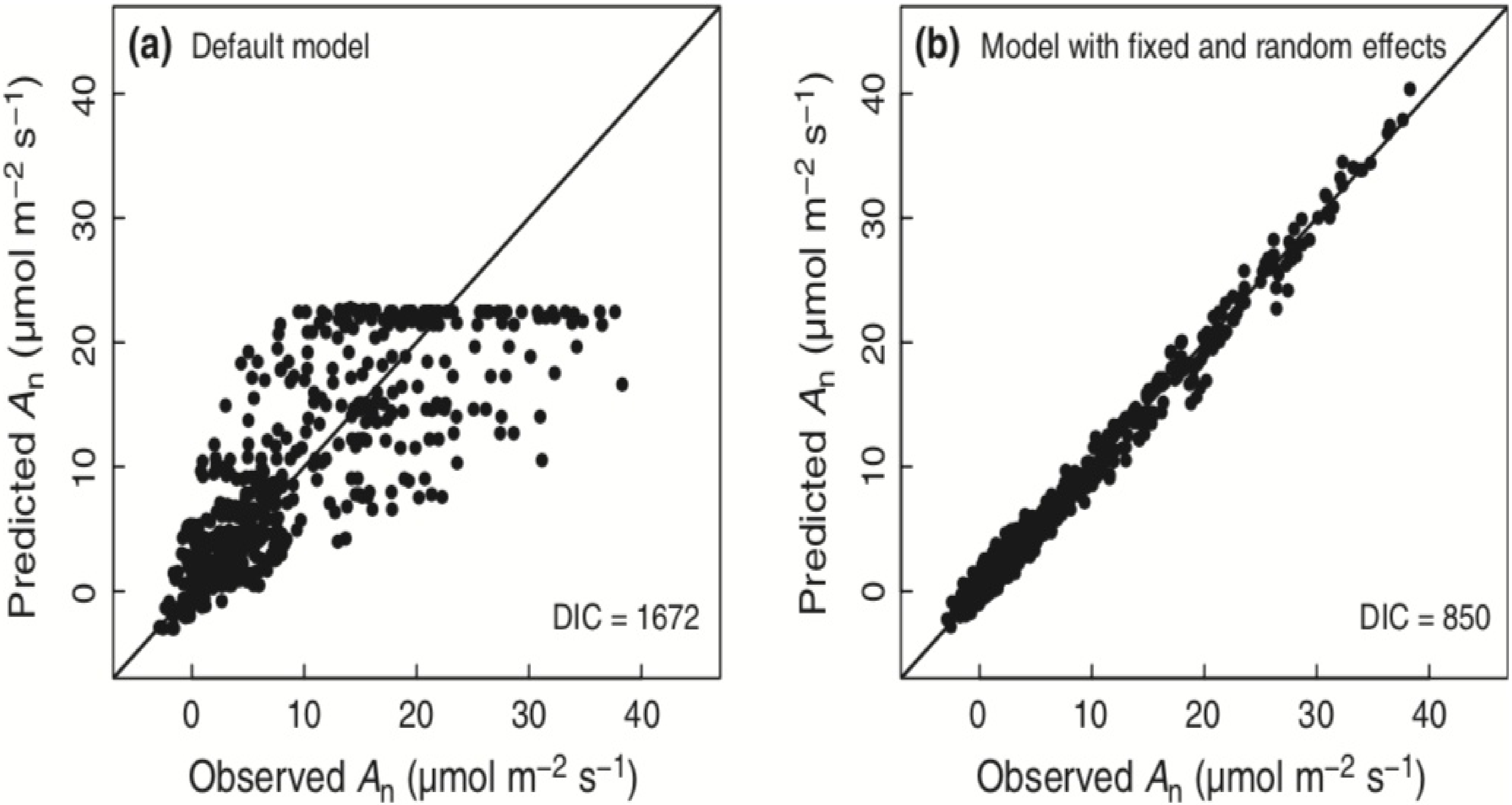
$$A_j = \frac{J(C_i - \Gamma^*)}{4C_i + 8\Gamma^*}$$

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

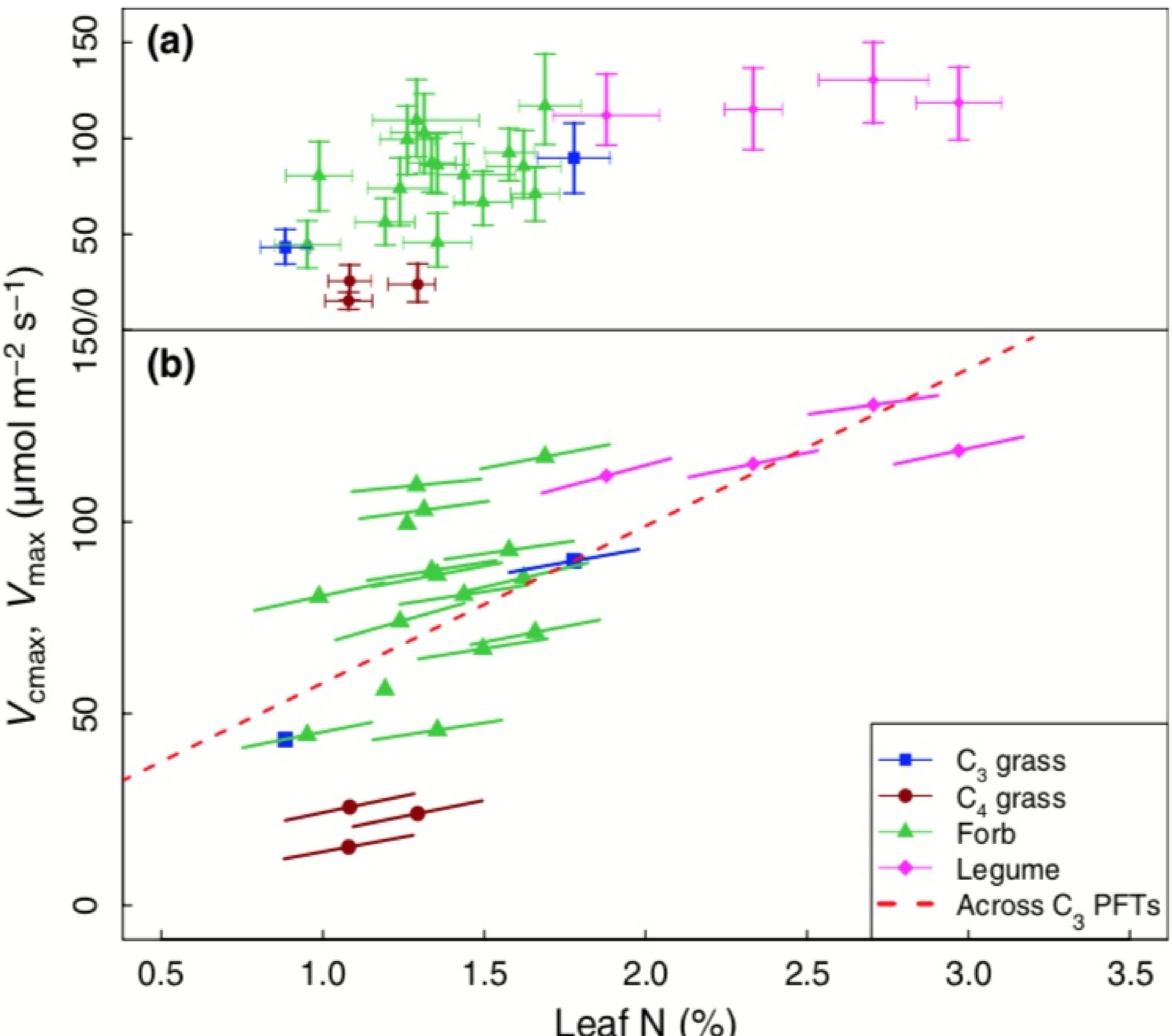
$$J = \frac{\alpha' q}{\sqrt{\left(1 + \frac{\alpha'^2 q^2}{J_{max}{}^2}\right.}}$$

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

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



(a) Default model
DIC = 1672
Predicted A_n (µmol m^{-2} s^{-1})
Observed A_n (µmol m^{-2} s^{-1})
0
10
20
30
40
(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)
V_{cmax}, V_{max} (μmol m^{-2} s^{-1})
150
100
50
150/0
150
100
50
0
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

Example: CO2 effect on tree seedling growth

- i – seedling
- j – plot
- t – year
- l – light
- y - growth

$$y_{i,j,t} = \underbrace{\mu_{i,j,t}}_{\text{mean}} + \underbrace{k_t}_{\text{year}} + \underbrace{\epsilon_{i,j,t}}_{\text{residual}}$$

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

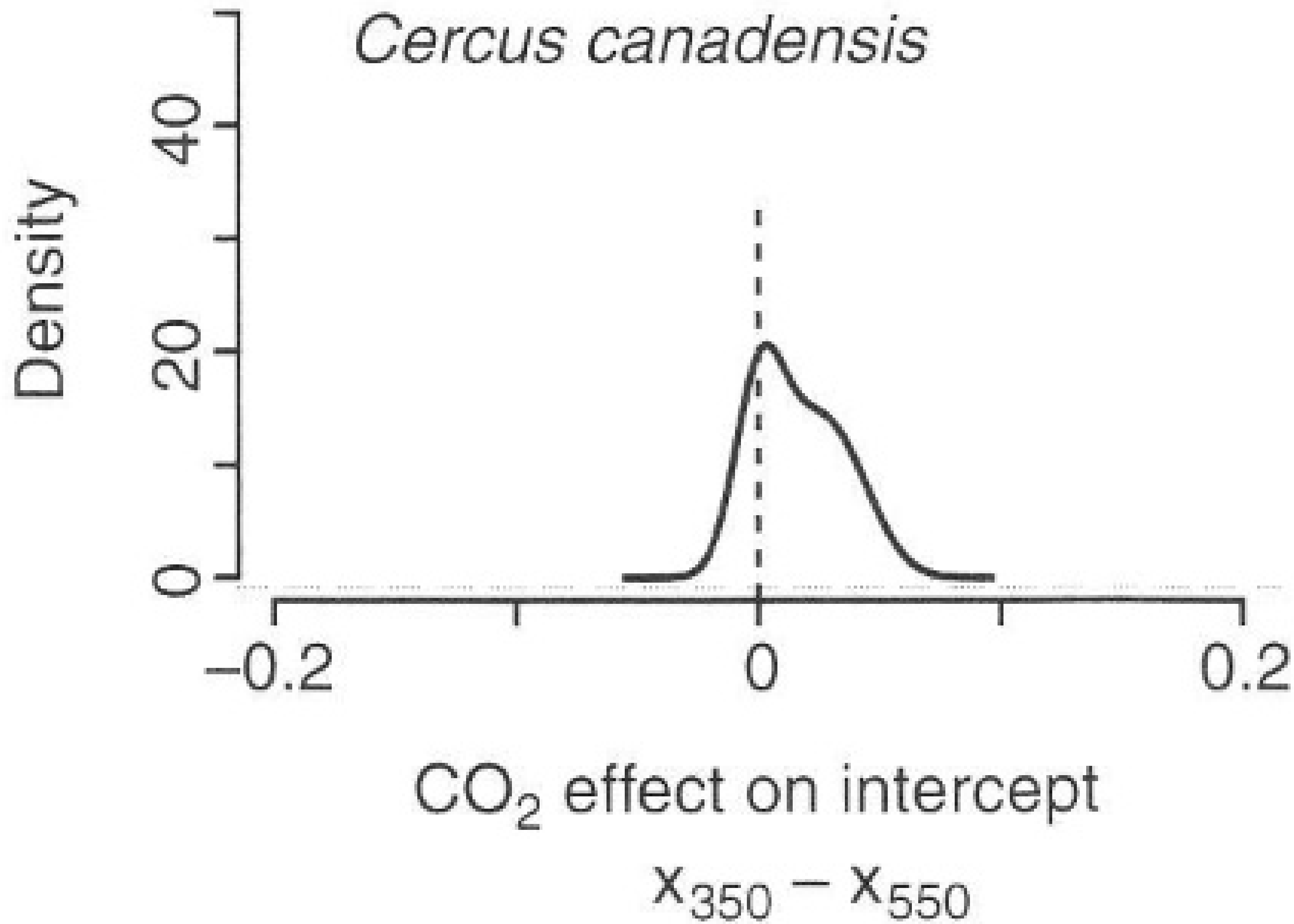
$$k_t \sim N(0, \nu_k)$$

$$\mu_{i,j,t} = g_{i,j} \left(\frac{l_{j,t} - l_c}{l_{j,t} + \theta} \right)$$

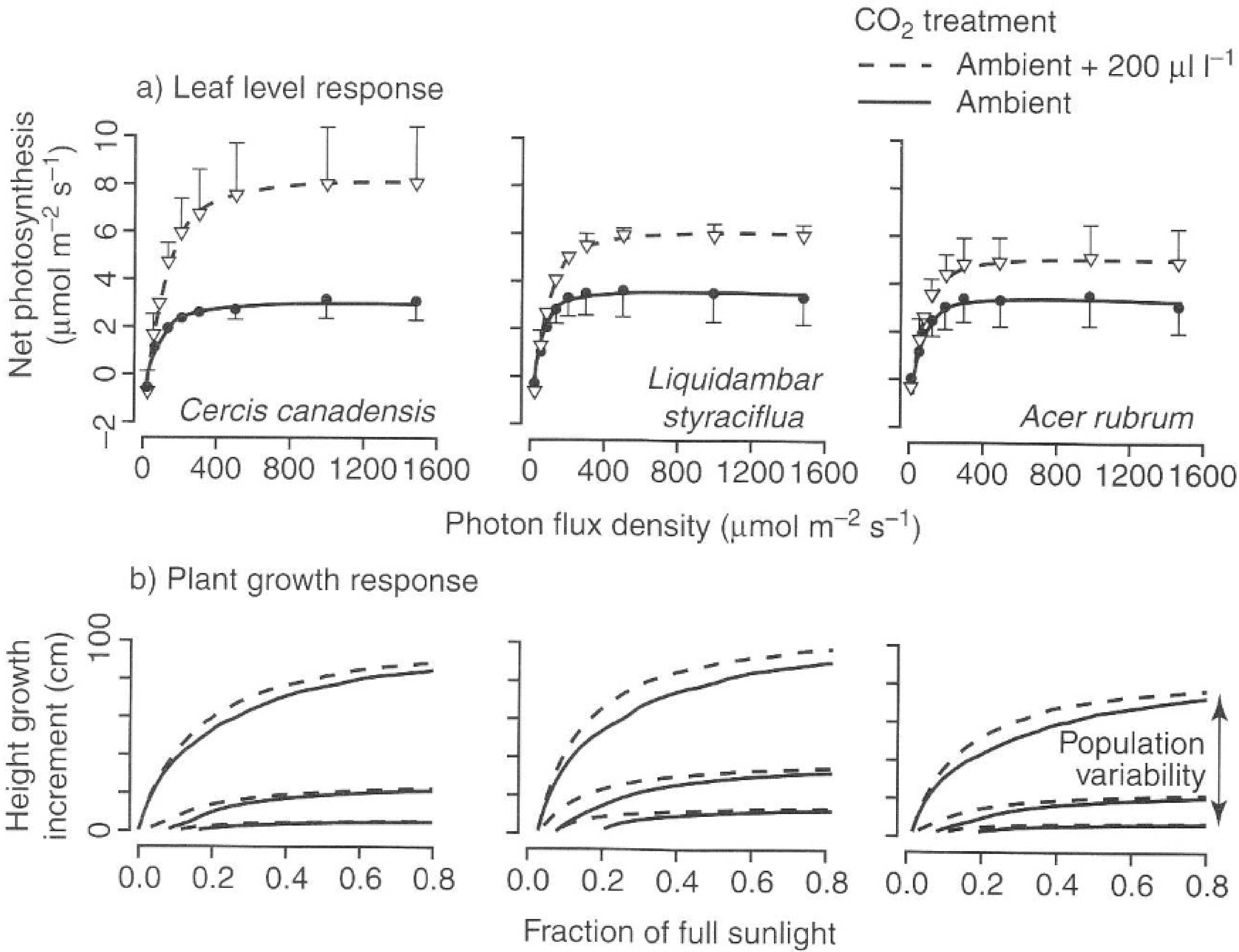
$$g_{i,j} \sim \ln(\alpha, \nu_g)$$

l_c varies w/ CO2, Priors on $\in$, v_g, v_k, σ^2, θ, l_c

CO2 effect



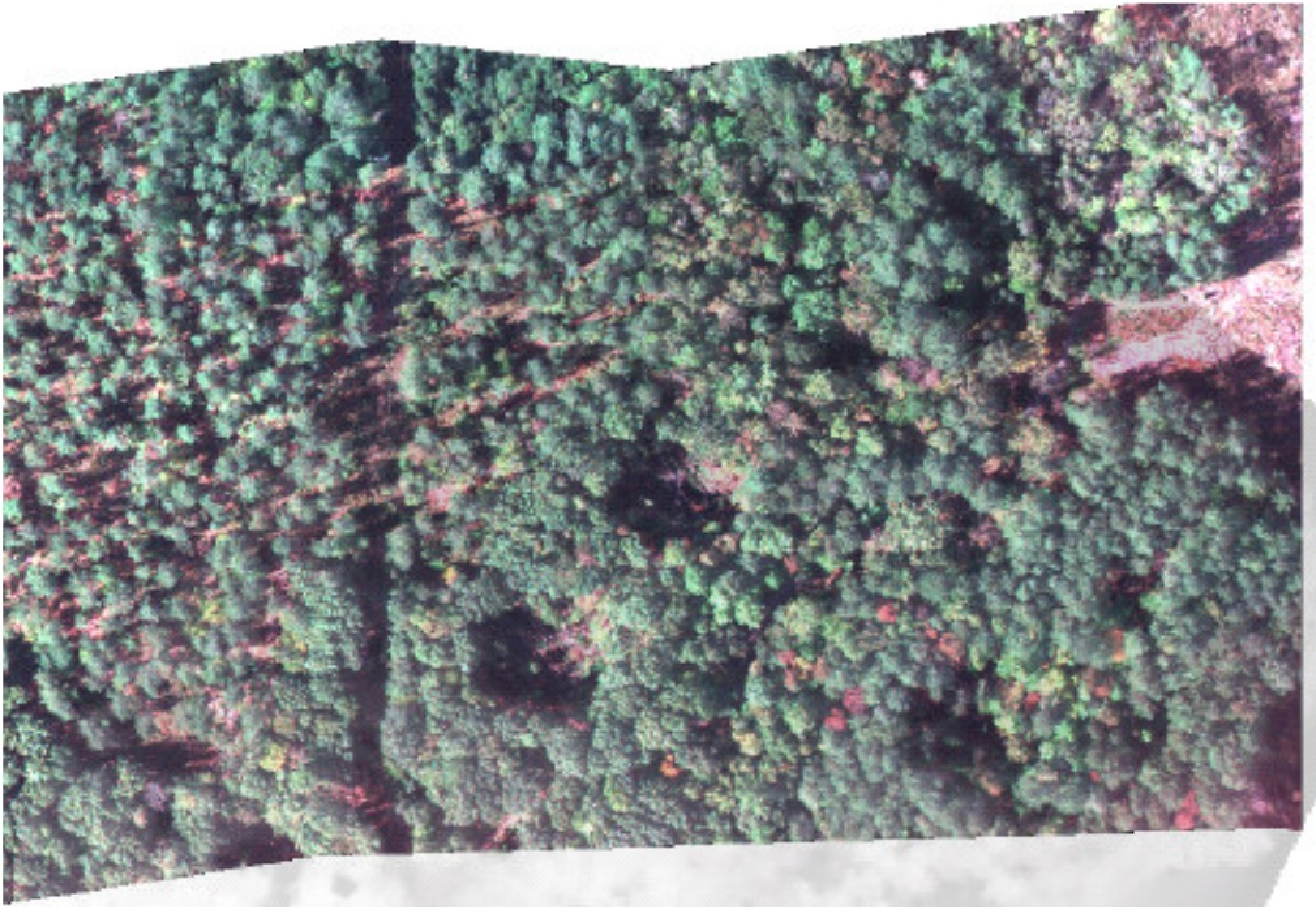
Cercus canadensis
Density
0
20
40
−0.2
0
0.2
CO2 effect on intercept
x350 − x550



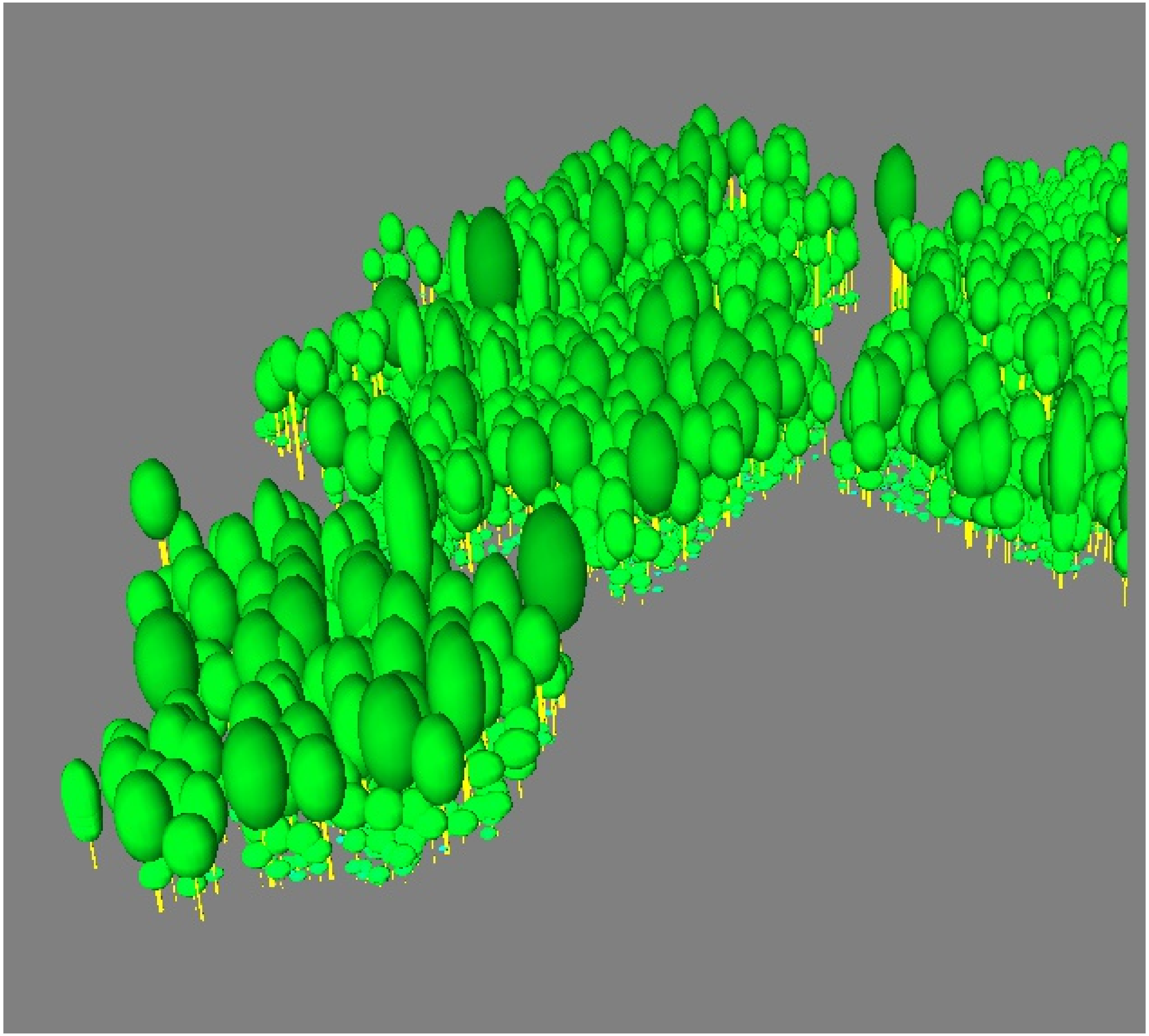
CO2 treatment
Ambient + 200 µl l−1
Ambient
a) Leaf level response
Net photosynthesis (µmol m−2 s−1)
Cercis canadensis
Liquidambar styraciflua
Acer rubrum
Photon flux density (µmol m−2 s−1)
b) Plant growth response
Height growth increment (cm)
Population variability
Fraction of full sunlight

Canopy Light: Synthesizing multiple data sources

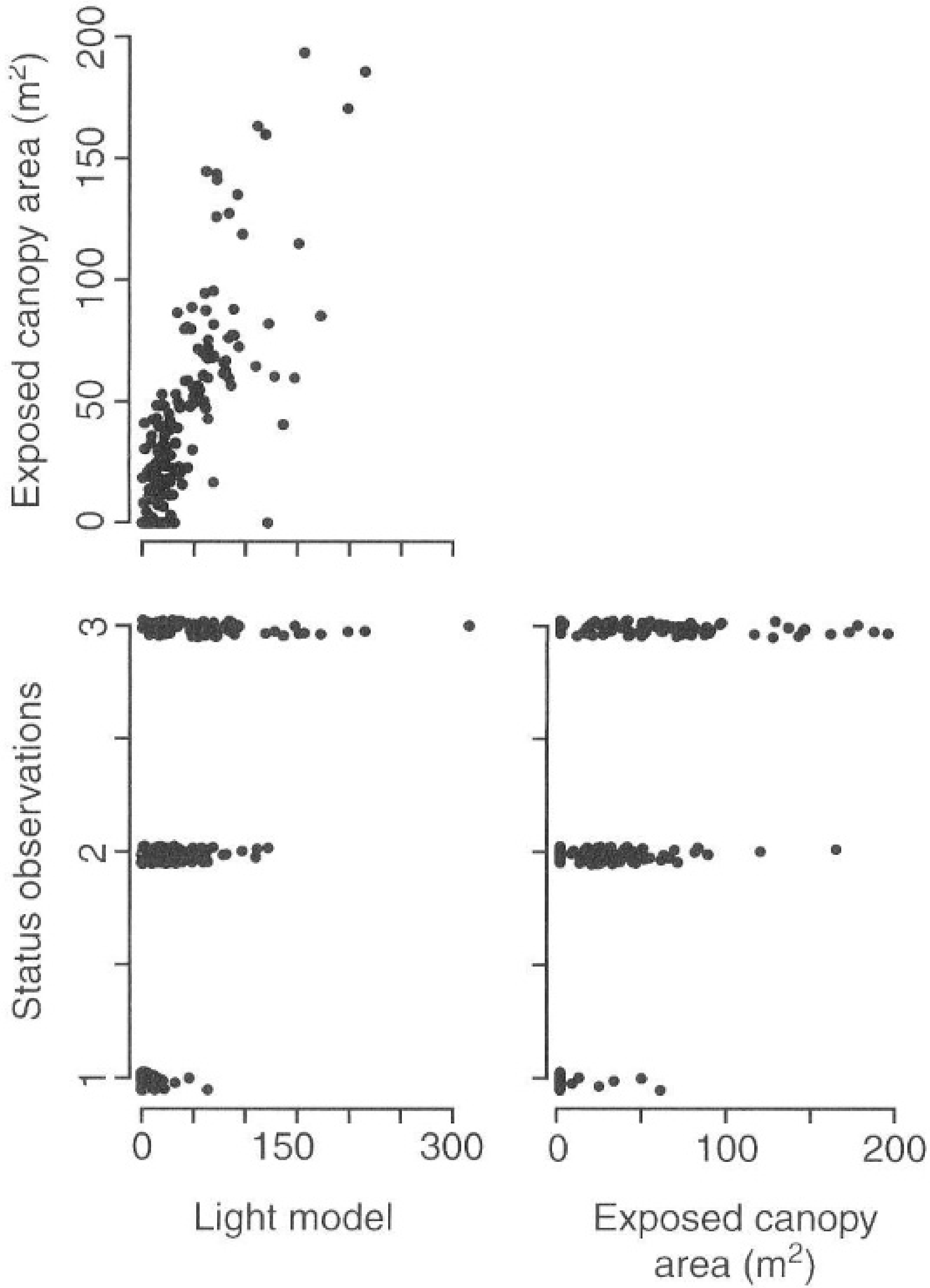
- Plant growth depends upon light (previous example, lab 7)
- Hard to measure how much light an ADULT tree receives
- Multiple sources of proxy data
  - Exposed Canopy Area
    - aerial photography, Quickbird
  - Canopy status
    - suppressed, intermediate, dominant (ex 8.2.2)
  - Light models
    - Allometries, stand map

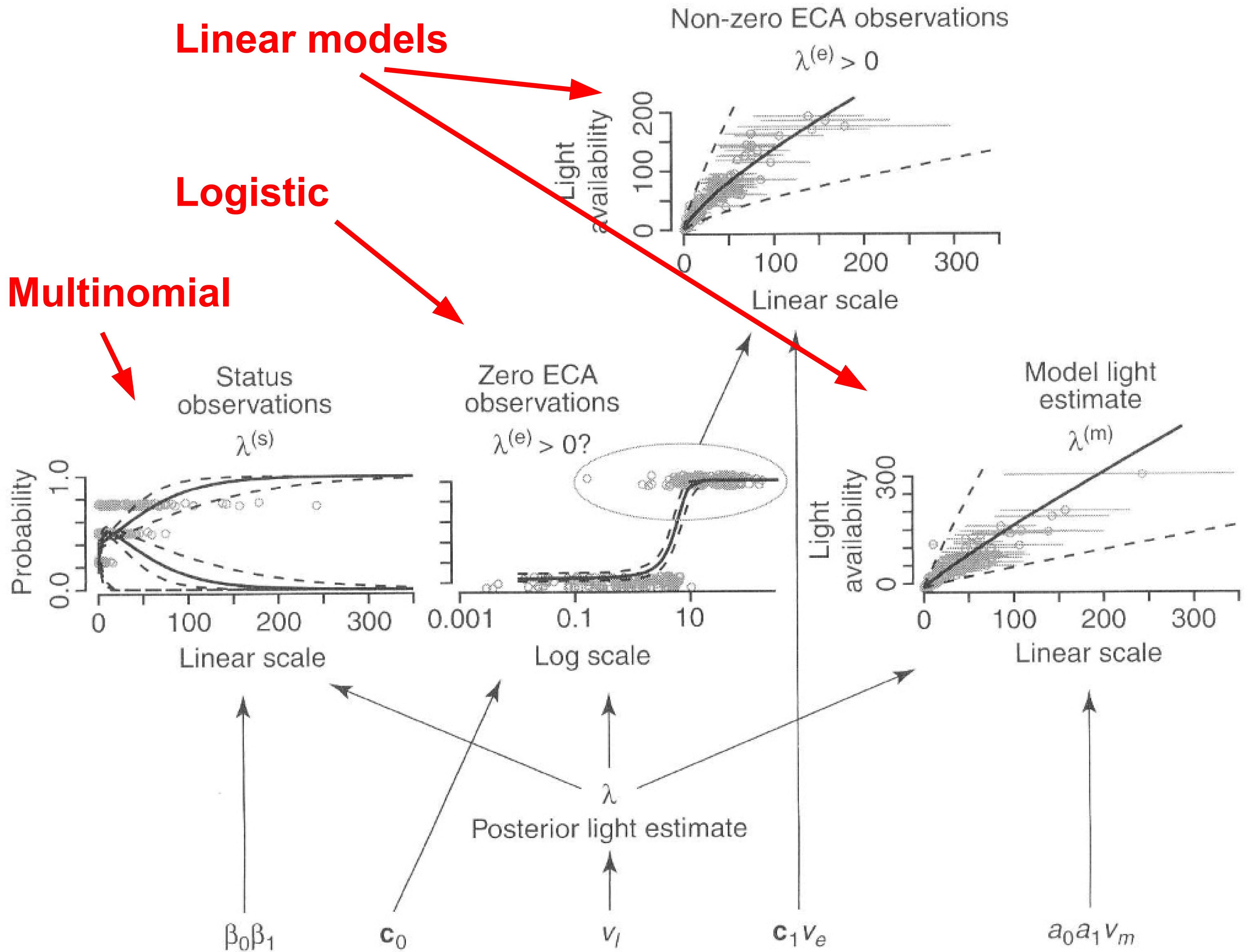


Mechanistic Light Model



- Estimate light levels based on a 3D ray-tracing light model
- Parameterized based on canopy photos, tree allometries


Exposed canopy area (m^2)
0
50
100
150
200
Status observations
1
2
3
0
150
300
Light model
0
100
200
Exposed canopy area (m^2)



Linear models
Logistic
Multinomial
Non-zero ECA observations
λ(e) > 0
Light availability
0 100 200
0 100 200 300
Linear scale
Status observations
λ(s)
Probability
0.0 1.0
0 100 200 300
Linear scale
Zero ECA observations
λ(e) > 0?
0.001 0.1 10
Log scale
Model light estimate
λ(m)
Light availability
0 300
0 100 200 300
Linear scale
λ
Posterior light estimate
β0β1
c0
vl
c1ve
a0a1vm

Exposed Canopy Area

- Error in relationship between "true" light λ and observations λ^e

$$p\left(\lambda_i^{(e)}\right)=\begin{cases} 1-p_i & \lambda_i^{(e)}=0 \\ p_i N\left(\ln\left(\lambda_i^{(e)}\right) \middle| \ln\left(\lambda_i\right), \nu_e\right) & \lambda_i^{(e)}>0 \end{cases}$$

- Probability of observing the tree in imagery increases with "true" light availability

$$logit\left(p_i\right)=c_0+c_1\lambda_i$$

Mechanistic Light Model

- Assume a log-log linear relationship between "true" light and modeled light
- Provides a continuous estimate of light availability for understory trees
  - ECA = 0
  - Status = 1

$$p\left(\lambda_i^{(m)}\right) = N\left(\ln\left(\lambda_i^{(m)}\right) \middle| a_0 + a_1 \cdot \ln\left(\lambda_i\right), \nu_m\right)$$

Model Fitting

- Model fit all at once
- Find the conditional probabilities for each parameter (i.e. those expressions that contain that parameter)
  - Always at least 2 – likelihood and prior
  - Can be multiple likelihoods
- MCMC iteratively updates each parameter conditioned on the current value of all others