Aplicações de modelos de efeitos mistos em modelagem e análises experimentais

Michel Colmanetti

Campinas/SP - 2019

- Definição de modelos de efeitos mistos
- Variáveis aleatórias
- Ajuste de modelos
 - Modelagem da variância
 - Calibração de modelos
- Análise experimental

MODELOS ESTATÍSTICOS

- Lineares ou não lineares
- Parâmetros são associados com a população

$$Y = \beta_0 + \beta_1 * x + e$$

sendo: β_0 e β_1 são efeitos fixos

MODELOS ESTATÍSTICOS

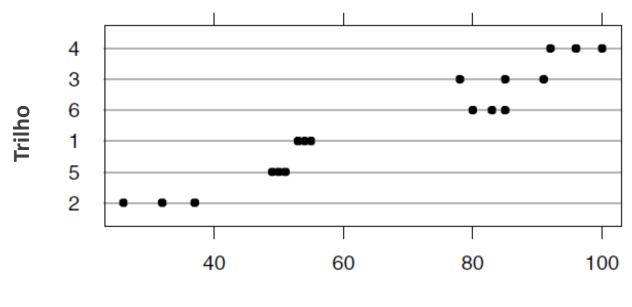
• Parâmetros associados à unidade experimental de uma população

$$Y = (b_0 + b_1) + \beta_0 + \beta_1 *x + e$$

sendo: β_0 e β_1 são efeitos fixos; b_0 e b_1 efeitos aleatórios

• Modelos de efeitos mistos: efeitos fixos e aleatórios

- Exemplo: teste de estresse longitudinal em trilhos (Devore, 2000)
- Experimento
 - Seis trilhos escolhidos aleatoriamente
 - Três medições no tempo



Tempo de viagem (nanosegundos)

- Questões:
 - O tempo médio de viagem para um trilho "típico"
 - Variação do tempo de viagem para cada trilho (variabilidade dentro dos trilhos)
 - Variação do tempo de viagem entre os trilhos (variabilidade entre os trilhos)

MODELOS ESTATÍSTICOS

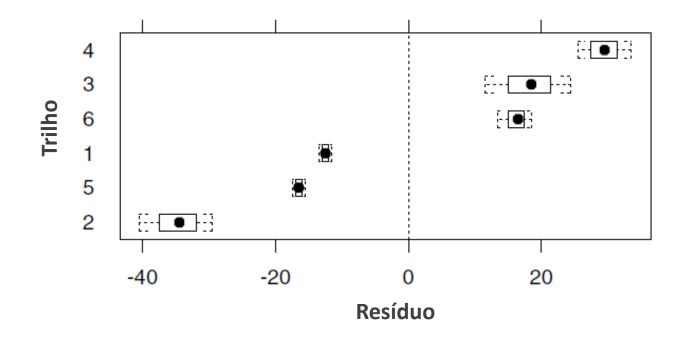
· O tempo médio de viagem para um trilho "típico"

$$y_{ij} = \beta + \epsilon_{ij}$$

Onde y_{ij} corresponde ao tempo de viagem para cada observação j e trilho i; e $\epsilon_{ij} \sim N(0, \sigma^2)$;

β representa a media do tempo de viagem de todos trilhos

- · O tempo médio de viagem para um trilho "típico"
 - $\beta = 66.5$
 - $\hat{\sigma} = 23.645$



MODELOS ESTATÍSTICOS

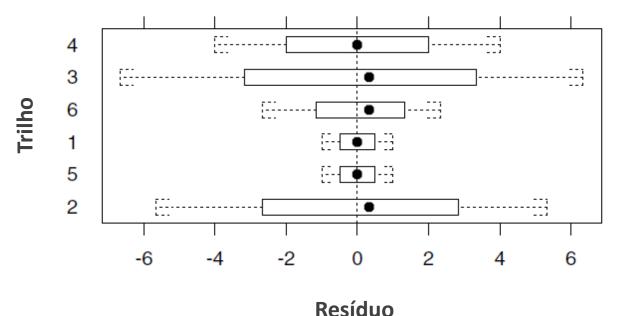
• Variação do tempo de viagem para cada trilho (variabilidade dentro dos trilhos)

$$y_{ij} = \beta_i + \epsilon_{ij}$$

Onde y_{ij} corresponde ao tempo de viagem para cada observação j e trilho i; e $\epsilon_{ij} \sim N(0, \sigma^2)$

 β_i há efeito do trilho i

- Variação do tempo de viagem para cada trilho (variabilidade dentro dos trilhos)
 - $\beta i = média para cada trilho i$
 - $\hat{\sigma} = 4.0208$



- Questões:
 - ✓O tempo médio de viagem para um trilho "típico"
 - ✓ Variação do tempo de viagem para um único trilho (variabilidade dentro dos trilhos)
 - Variação do tempo de viagem entre os trilhos (variabilidade entre os trilhos)

- Questões:
 - ✓O tempo médio de viagem para um trilho "típico"
 - ✓ Variação do tempo de viagem para um único trilho (variabilidade dentro dos trilhos)
 - Variação do tempo de viagem entre os trilhos (variabilidade entre os trilhos)
- Modelos estão condicionados à amostra, mas o objeto de estudo é sempre uma população.

MODELOS ESTATÍSTICOS

• Variação do tempo de viagem entre os trilhos (variabilidade entre os trilhos)

$$y_{ij} = \beta + \mathbf{b}_i + \epsilon_{ij}$$

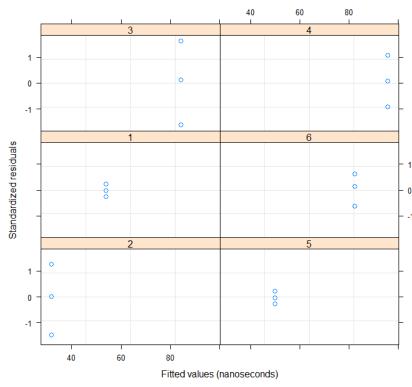
Onde y_{ij} corresponde ao tempo de viagem para cada observação j e trilho i; $\epsilon_{ij} \sim N(0, \sigma^2)$; e $b_{ij} \sim N(0, \sigma_b^2)$

 β é a media da população e b_i o efeito aleatório de cada trilho i, que varia em relação à media da população.

MODELOS ESTATÍSTICOS

Variação do tempo de viagem entre os trilhos (variabilidade entre os trilhos)

- $\beta = 66.5$
- $b_i = -12.4, -34.5, 18.0, 29.2, -16.4, 16.0$
- $\hat{\sigma} = 4.0208$
- $\sigma_b^2 = 24.80547$



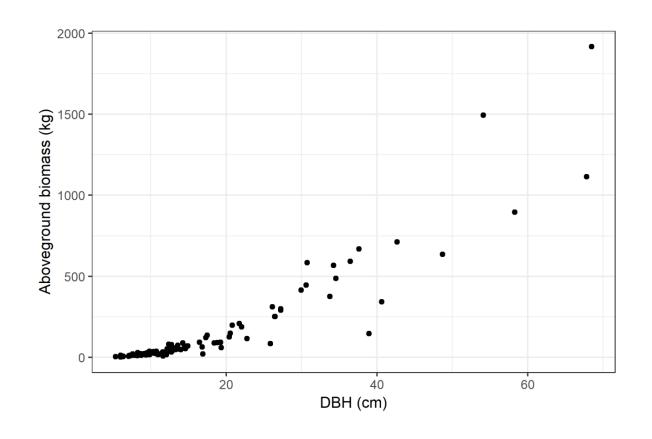
- Modelos de efeitos mistos descreve uma relação entre a variável resposta e covariávies que estão agrupadas, e um ou mais fatores.
- Podem ser:
 - Dados longitudinais (Medidas repetidas)
 - Dados multi-hieráquicos
 - Design em blocos

- **nlme**(model, data, fixed, random, groups, start, correlation, weights, method, ...)
- method (Estimador): ML e REML
 - ML Máxima verossimilhança (Default)
 - REML Máxima verossimilhança restrita
- Distribuição Gaussiana

MODELOS ESTATÍSTICOS

Aplicação em biometria florestal:

- Volume/Biomassa florestal
 - I. Energia
 - II. Madeira
 - III. Celulose
 - IV. Estoque de carbono



MODELOS ESTATÍSTICOS

Modelo florestal clássico para biomassa/volume:

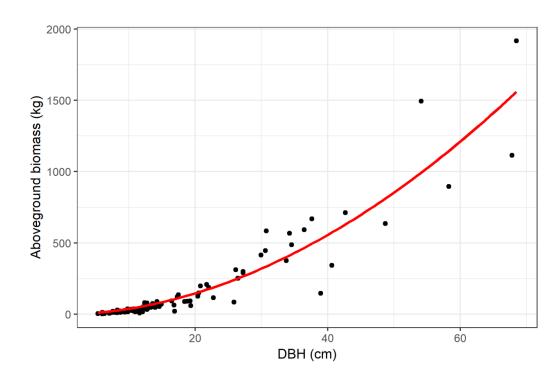
•
$$agb_i = \beta_1 . dbh_i^{\beta_2} . \epsilon_i$$

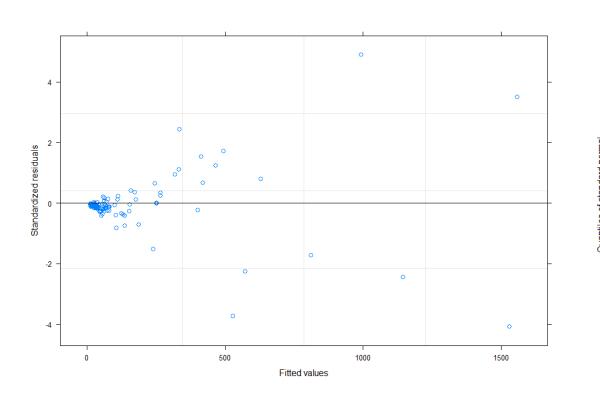
MODELOS ESTATÍSTICOS

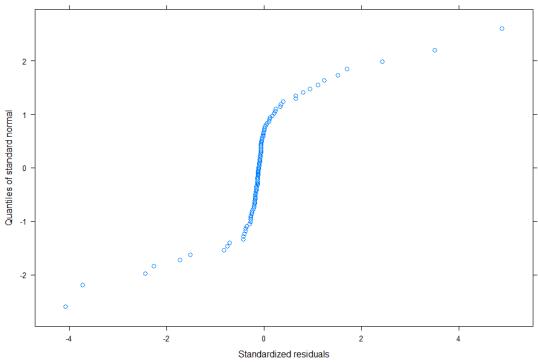
Modelo florestal clássico para biomassa/volume:

• $agb_i = \beta_1 . dbh_i^{\beta_2} . \epsilon_i$

Estimador: OLS; $\epsilon_i \sim N(0, \sigma^2)$







Modelos de efeitos mistos FLORESTAS NATIVAS

- Elevada diversidade de espécies
 - Heterogêneas
 - Inequiâneas

FLORESTAS NATIVAS

LOCAL DE ESTUDO

- 0.8 ha
- 90 espécies
- 16 espécies = 73 % de árvores

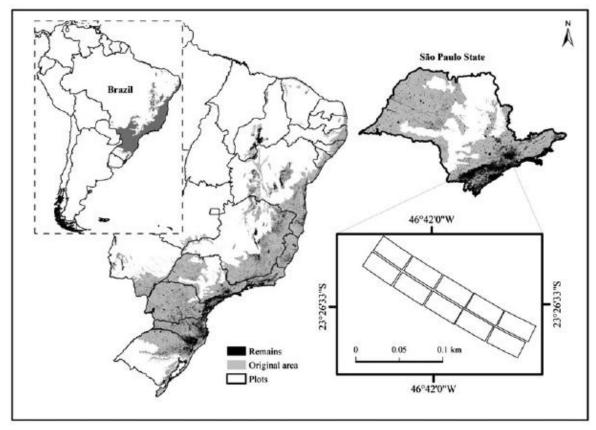


Figure 1. Plots on the study site located in Atlantic Forest near São Paulo, Brazil (souces: SOS Mata Atlântica, 2013; MMA, 2012; IBGE, 2010).

Modelos de efeitos mistos FLORESTAS NATIVAS

Species	dbh (cm)	wsg (g cm ⁻³)	ht (m)	hcb (m)	agb (kg)	Carbon (%)
	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
AlchorneasidifoliaMüll. Arg. ¹	18.8 (12.0)	0.45 (0.01)	9.9 (2.6)	4.2 (1.6)	182.4 (233.4)	45.6 (1.3)
AllophyluspetiolulatusRadlk. ³	13.1 (5.7)	0.53 (0.02)	10.8 (1.4)	6.2 (1.4)	67.0 (71.2)	44.8 (0.7)
Cabraleacanjerana(Vell.) Mart. ³	9.3 (1.9)	0.42 (0.04)	9.3 (1.0)	5.2 (1.8)	17.5 (7.74)	45.4 (1.3)
CaseariasylvestrisSw. ²	13.2 (4.7)	0.55 (0.02)	9.4 (0.9)	4.0 (1.2)	64.1 (48.5)	45.0 (0.7)
Ceibaspeciosa(A. StHil.)	32.1 (20.8)	0.31 (0.06)	12.3 (3.9)	8.4 (2.9)	294.8 (396.0)	42.9 (1.2)
Croton floribundusSpreng.1	16.8 (9.4)	0.53 (0.05)	12.4 (2.3)	7.1 (1.0)	153.6 (161.3)	45.3 (1.2)
CupaniaoblongifoliaMart. ²	12.9 (9.1)	0.68(0.05)	8.0 (1.7)	3.1 (1.2)	100.4 (170.5)	44.1 (1.0)
Jacaranda puberulaCham.2*	12.9 (5.8)	0.34 (0.03)	9.7 (2.1)	5.8 (1.9)	44.2(54.9)	47.5 (1.5)
MachaeriumvillosumVogel ^{3*}	15.3 (7.6)	0.58 (0.02)	11.3 (2.7)	6.8 (1.6)	104.6 (120.6)	45.5 (0.8)
Myrciasplendens(Sw.) DC. ²	11.8 (3.4)	0.51 (0.03)	8.6 (2.4)	4.8 (1.1)	51.7 (27.6)	45.1 (0.8)
NectandraoppositifoliaNess ²	15.4 (7.5)	0.45 (0.05)	11.4 (2.2)	7.3 (1.5)	101.7 (96.9)	47.5 (1.0)
Peraglabrata(Schott) Poepp.	25.0 (23.6)	0.57 (0.06)	10.2 (3.9)	4.6 (1.8)	466.0 (701.8)	44.4 (0.8)
Piptadeniagonoacantha (Mart.) J. F.	20.3 (13.2)	0.58 (0.03)	13.9 (3.7)	8.0 (2.9)	250.1 (293.4)	45.5 (0.6)
$Sesseabra siliensis Toledo^3$	15.3 (4.6)	0.50(0.07)	11.2 (2.4)	6.8 (2.7)	70.0 (42.4)	46.6 (1.7)
TetrorchidiumrubriveniumPoepp. ²	27.1 (20.1)	0.40 (0.03)	14.1 (3.0)	7.7 (1.8)	407.4 (536.1)	47.2 (0.8)
VochysiatucanorumMart. ²	12.9 (4.7)	0.43 (0.03)	11.6 (2.7)	7.4 (2.1)	46.1 (34.6)	44.7 (0.9)
Overall	17.3 (13.0)	0.49 (0.10)	10.9 (3.0)	6.2 (2.4)	158.5 (297.7)	45.4 (1.6)

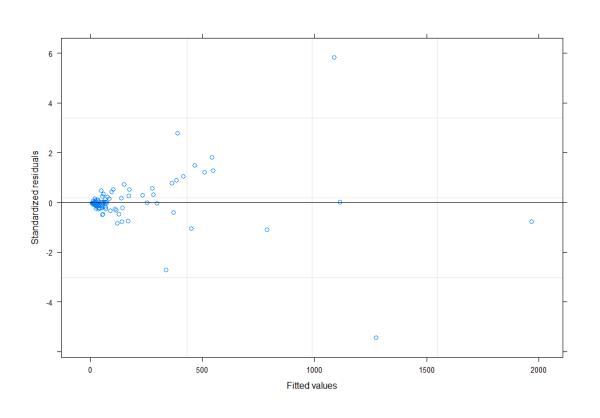
FLORESTAS NATIVAS

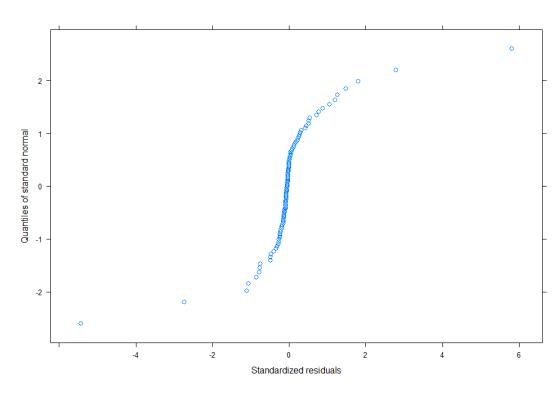
Modelo de efeitos mistos:

$$agb_{ij} = \beta_1.dbh_{ij}^{\beta_2}.\epsilon_{ij}$$

- Estimador: nlme; $e_{ij} \sim N(0, \sigma^2)$;
- Sendo i = árvore; j = variável aleatória em nível de espécies.

FLORESTAS NATIVAS





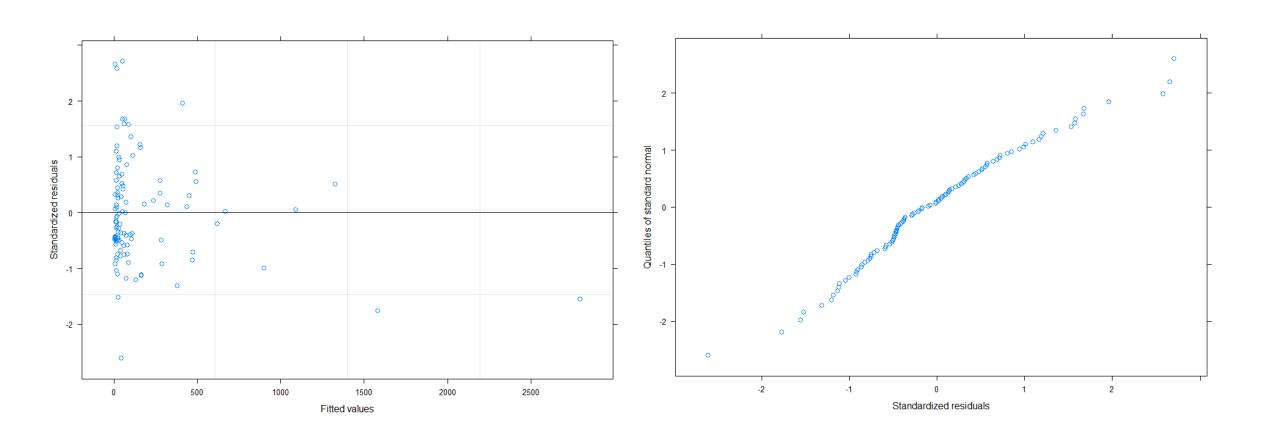
FLORESTAS NATIVAS

Modelo de efeitos mistos com modelagem da variância:

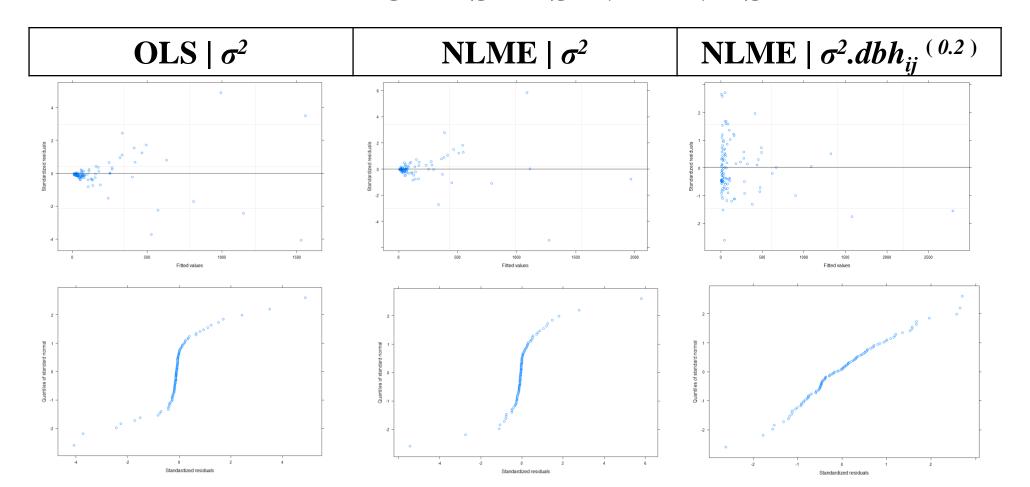
$$agb_{ij} = \beta_1.dbh_{ij}^{\beta_2}.\epsilon_{ij}$$

- Estimador: nlme; $e \sim N(0, Var(e_{ij}))$
- $Var(e_{ij}) = \sigma^2.dbh_{ij}^{(0.2)}$

FLORESTAS NATIVAS



FLORESTAS NATIVAS



FLORESTAS NATIVAS

Akaike Information Criteriom (AIC)

$$AIC = 2k \cdot 2 \ln(\hat{\mathcal{L}})$$

• Sendo k o número de parâmetros, $ln(\hat{\mathcal{L}})$ o de log-verossimilhança do modelo.

FLORESTAS NATIVAS

OLS σ^2	NLME σ^2	NLME $\sigma^2 dbh_{ij}$ (0.2)		
AIC = 1285.7897	AIC = 1225.6837	AIC = 910.1343		
4 -	6			
2 - Sometime of the control of the c	2 Consequence of the following of the fo	Z - C O O O O O O O O O O O O O O O O O O		

Modelos de efeitos mistos FLORESTAS NATIVAS

- Elevada diversidade de espécies
 - Heterogêneas
 - Inequiâneas
 - Legalmente protegidas

Modelos de efeitos mistos FLORESTAS NATIVAS

- Elevada diversidade de espécies
 - Heterogêneas
 - Inequiâneas
 - Legalmente protegidas
 - Calibração de modelos

Calibração de modelos mistos:

•
$$agb_{ij} = \phi_1.dbh_{ij}^{\phi_2}.\epsilon_{ij}$$

•
$$\ln(agb_{ij}) = \phi_1 + \phi_2 \cdot \ln(dbh_{ij}) + \epsilon_{ij}$$

$$\phi = \begin{bmatrix} \phi_{1i} \\ \phi_{2i} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \beta + b$$

$$b^{\sim} N(0, \Psi), \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

Calibração de modelos mistos:

•
$$ln(agb_{ij}) = (\beta_0 + b_{0j}) + (\beta_1 + b_{1j}) \cdot ln(dbh_{ij}) + \epsilon_{ij}$$

BLUP $(b_j) = \hat{b}_j = DZ'(ZDZ' + R)^{-1} (y - \mu)$
 $var(\hat{b}_j - b_j) = D - DZ'(ZDZ' + R)^{-1} ZD$

Sendo: D a matriz de variância e covariância dos efeitos aleatórios; Z matriz design nxp (n=número de árvores e p o número de parâmetros); R a variância dos efeitos aleatórios vezes a matriz I; y_j a média das novas observações; e μ a média para as mesmas observações preditas pelo parâmetros fixos da equação antes da calibração.

FLORESTAS NATIVAS

LOCAL DE ESTUDO

- 0.8 ha
- 90 espécies
- 15 espécies = amostra
 - Calibração para nova espécie

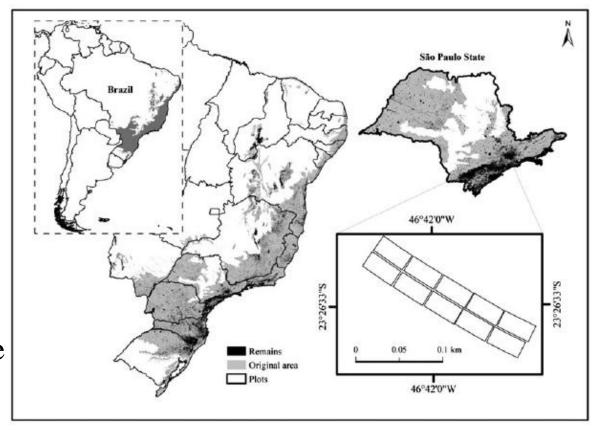
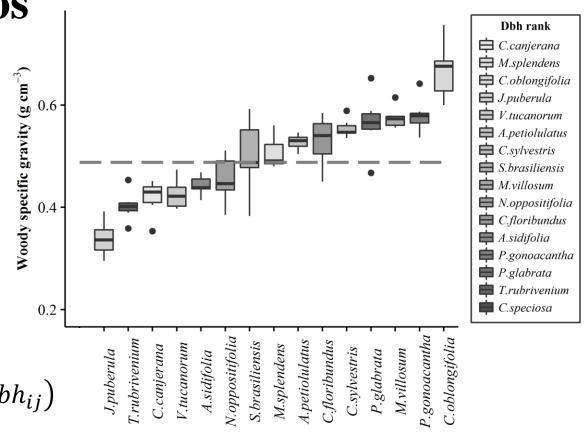


Figure 1. Plots on the study site located in Atlantic Forest near São Paulo, Brazil (souces: SOS Mata Atlântica, 2013; MMA, 2012; IBGE, 2010).

Calibração de modelos mistos

Exemplo:

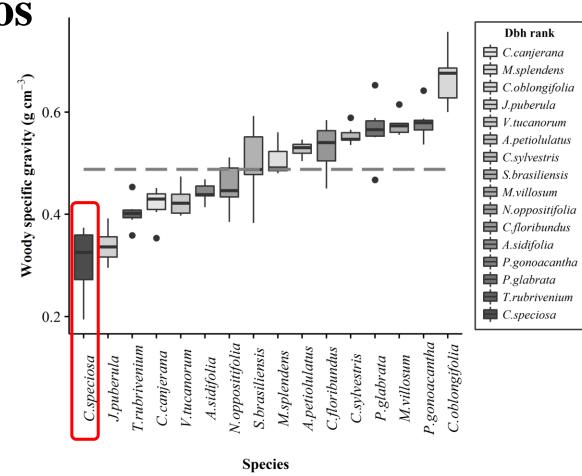
- 15 espécies
- $\rho = 0.34 0.68 \ g \ cm^{-3} (0.5 \ g \ cm^{-3})$
- $DBH = 5.4 68.5 \ cm \ (16.1 \ cm)$
- $ln(a|ij) = (2.178 + b_{0j}) + (2.356 + b_{1j}) . ln(dbh_{ij})$



Calibração de modelos mistos

Exemplo:

- Ceiba speciosa (paineira)
- $\rho = 0.31 \ g \ cm^{-3}$
- $DBH = 6.0 67.8 \ cm \ (32.1 cm)$



Calibração de modelos mistos

$$ln(a|ij) = (2.178 + b_{0i}) + (2.356 + b_{1i}) \cdot ln(dbh_{ij}) + \epsilon_{ij}$$

• Predizendo um efeito aleatório com três árvores de C. speciosa:

$$dbh = \begin{bmatrix} 6.00 \\ 25.85 \\ 40.65 \end{bmatrix}$$
 cm, $agb = \begin{bmatrix} 2.3 \\ 85.77 \\ 342.42 \end{bmatrix}$ kg

Calibração de modelos mistos

• O efeito aleatório pode ser obtido:

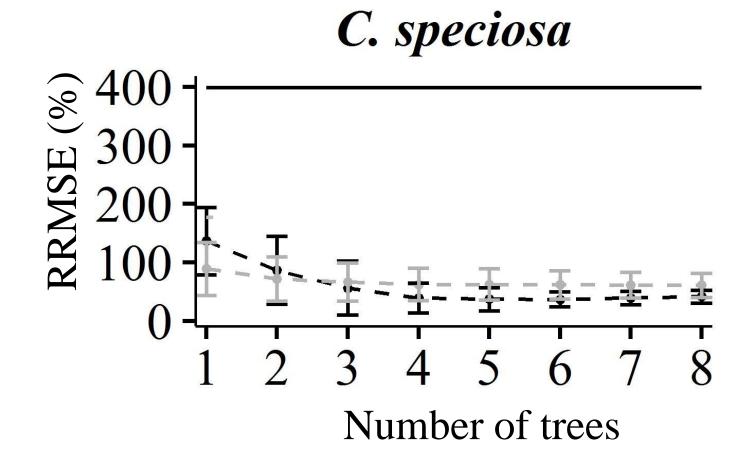
$$\left(\begin{bmatrix} 1 & 6.00 \\ 1 & 25.85 \\ 1 & 40.65 \end{bmatrix}^t \cdot \begin{bmatrix} 0.0476 & 0 & 0 \\ 0 & 0.0476 & 0 \\ 0 & 0 & 0.0476 \end{bmatrix} \cdot \begin{bmatrix} 1 & 6.00 \\ 1 & 25.85 \\ 1 & 40.65 \end{bmatrix} \cdot \begin{bmatrix} 0.03318 & 0 \\ 1 & 25.85 \\ 1 & 40.65 \end{bmatrix}^t \cdot \begin{bmatrix} 1 & 6.00 \\ 1 & 25.85 \\ 1 & 40.65 \end{bmatrix}^t \cdot \begin{bmatrix} 0.0476 & 0 & 0 \\ 0 & 0.0476 & 0 \\ 0 & 0 & 0.0476 \end{bmatrix} \cdot \begin{bmatrix} \log(2.3) - 2.053 \\ \log(85.77) - 5.479 \\ \log(342.42) - 6.541 \end{bmatrix} = \begin{bmatrix} -0.69981 \end{bmatrix}$$

• Modelo calibrado para *C. speciosa*:

[-0.0050]

$$ln(a|ij) = (2.178 - \mathbf{0.6998}) + (2.356 - \mathbf{0.0050}) \cdot ln(dbh_{ij}) + \epsilon_{ij}$$

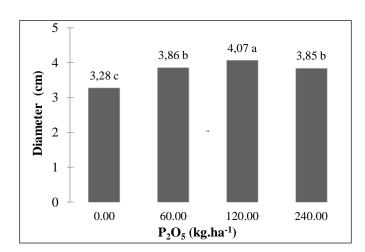
Calibração de modelos mistos

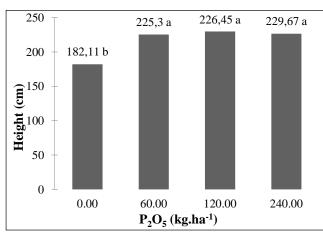


• Análise experimental, comparação de modelos e o uso de modelos de efeito misto.

Schizolobium amazonicum Huber ex. Ducke (Paricá)

- Rápido crescimento
- Production of plywood





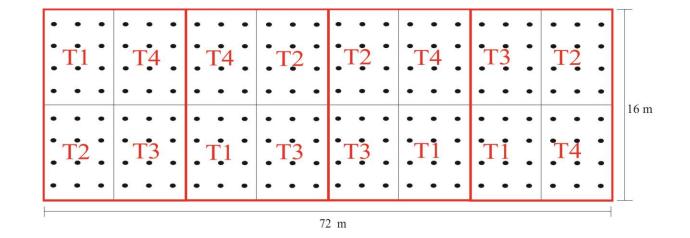
12 months after planting





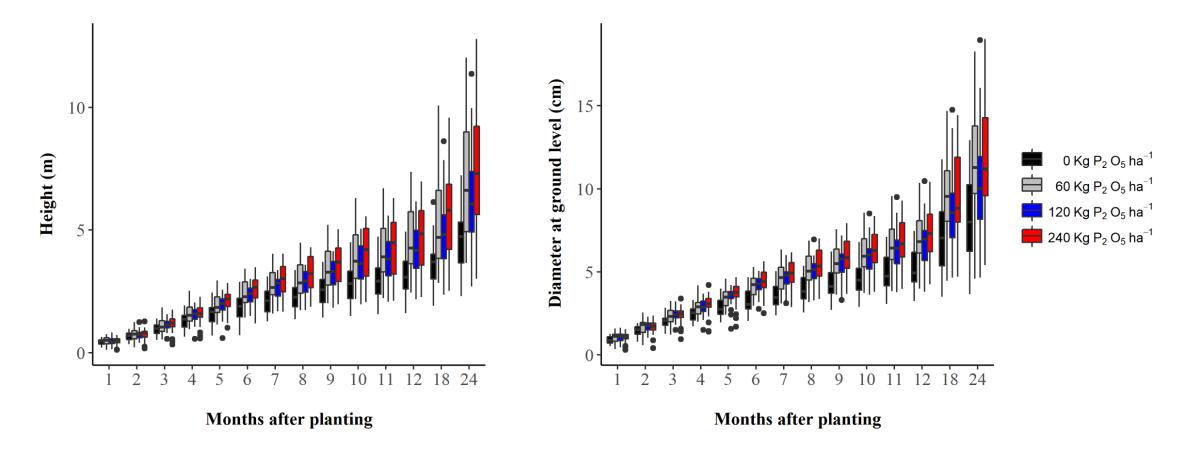
Design experimental

- Four treatments
 - 0 Mg P_2O_5 ha⁻¹
 - $0.06 \text{ Mg P}_2\text{O}_5 \text{ ha}^{-1}$
 - $0.12 \text{ Mg P}_2\text{O}_5 \text{ ha}^{-1}$
 - $0.24 \text{ Mg P}_2\text{O}_5 \text{ ha}^{-1}$



- Four Blocks
- 14 measurements of diameter and height

• Análise experimental e comparação de modelos



MODELOS DE CRESCIMENTO:

• Shumacher equation

•
$$y_{ijk} = \phi_{1ijk} \exp\left[-\phi_{2ijk}/t\right] + \epsilon_{ijk}$$
 (eq. 1)

General logistic equation

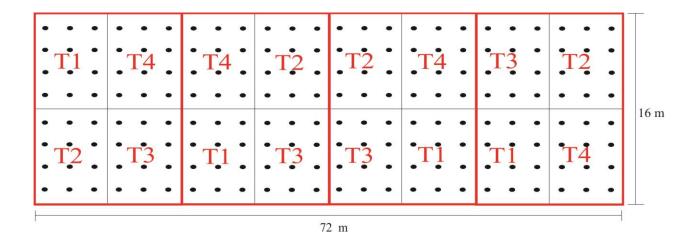
•
$$y_{ijk} = \frac{\phi_{1ijk}}{1 + exp\left[-(t - \phi_{2ijk}) / \phi_{3ijk}\right]} + \epsilon_{ijk}$$
 (eq. 2)

• Monomolecular equation

•
$$y_{ijk} = \phi_{1ijk} \cdot (1 - \phi_{2ijk} \cdot ex \, p[-\phi_{3ijk} \cdot t]) + \epsilon_{ijk}$$
 (eq. 3)

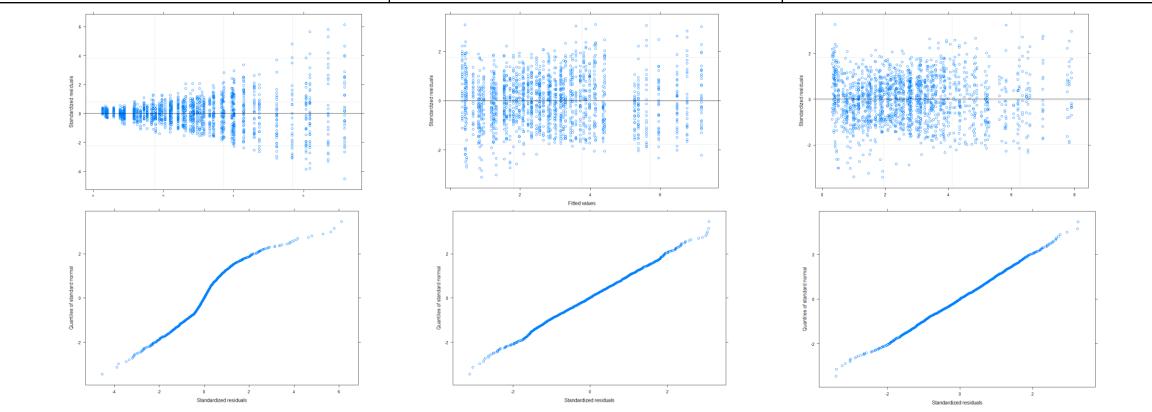
- $b_{ijk} \sim N(0, \Psi)$, $\epsilon_{ijk} \sim N(0, \sigma^{**})$
- $\sigma^{**} = Var(\epsilon_{ij}) = \sigma^2 |months|^{2\delta}$
- Pacote *nlme*
- Níveis de efeitos aleatórios
- Variável indicativa entre os tratamentos para comparar os modelos

- Mais de um nível de efeito aleatório: modelos multiníveis
- Os efeitos aleatórios são "aninhados" (nested) em níveis hierárquicos
- Efeito aleatórios: bloco / tratamento



• Modelagem da variância e níveis hierárquicos de efeitos aleatórios para o diâmetro de paricá

NLME trat σ^2	NLME trat σ^2 .months (2δ)	NLME bloc/trat σ^2 .months (2δ)
AIC = 4766.921	AIC = 3319.933	AIC = 3092.573



• Comparação de modelos

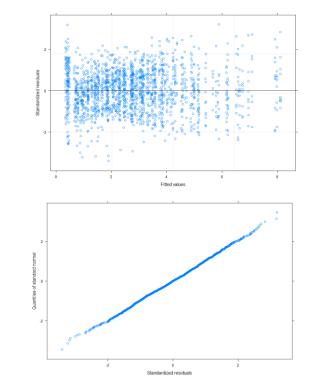
	AIC	1	
Model	Height	Diameter	
Monomolecular (eq.3)	3066.968	4467.239	
Logistic (eq.2)	3245.770	4838.033	
Shumacher (eq.1)	3948.514	5849.342	

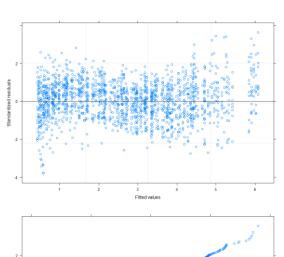
Modelagem da altura (m) de S. amazonicum em diferentes concentrações de fósforo.

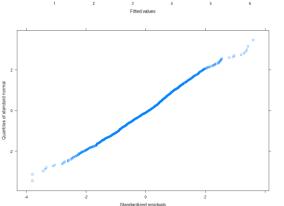
Model	eta_1	eta_2	eta_3	[kg P ₂ O ₅ ha ⁻¹]	Indicator variable	F-value	p-value
		0.982	0.048	0	-	-	-
Manamalagulan	0.014			60	0.029	0.500	0.470
Monomolecular	Monomolecular 9.014			120	0.021	0.200	0.625
				240	0.033	6.200	0.013
		6.502	3.049	0	-	-	-
Tanialia	5 220			60	1.231	9.482	0.002
Logistic 5.230	5.230			120	0.766	1.528	0.217
				240	1.292	55.631	<.0001
Shumacher 6				0	-	-	-
	C 227	4.076		60	1.811	10.069	0.002
	6.237	4.276	-	120	0.989	0.115	0.735
				240	1.973	47.238	<.0001

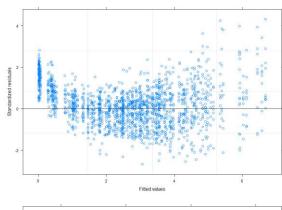
Comparando modelos com variáveis indicativas para altura

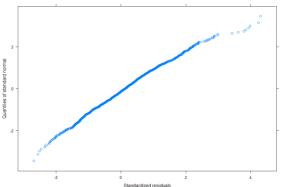
Monomolecular bloc/trat σ^2 .months (2δ)	Logistic bloc/trat σ^2 .months $^{(2\delta)}$	Shumacher bloc/trat σ^2 .months $^{(2\delta)}$
AIC = 3066.968	AIC = 3245.770	AIC = 3948.514









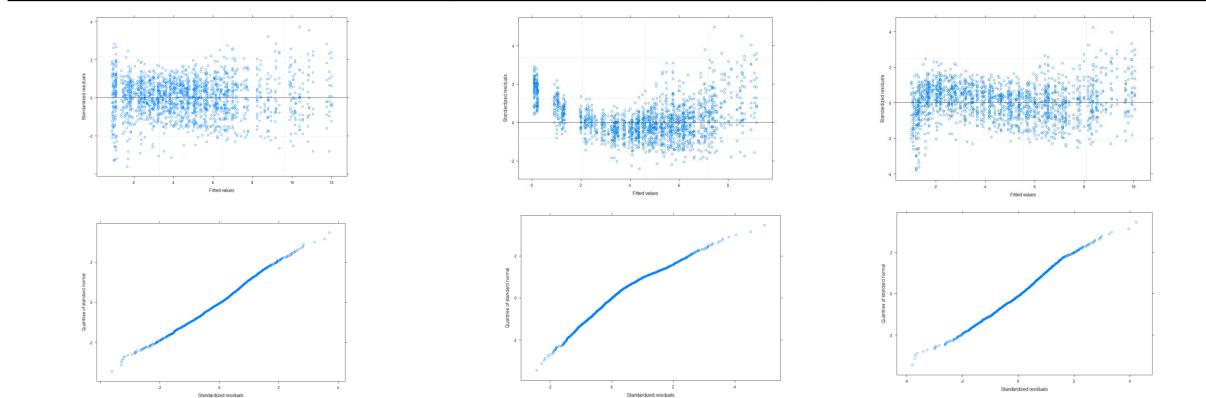


Modelagem do diâmetro (cm) de S. amazonicum em diferentes concentrações de fósforo.

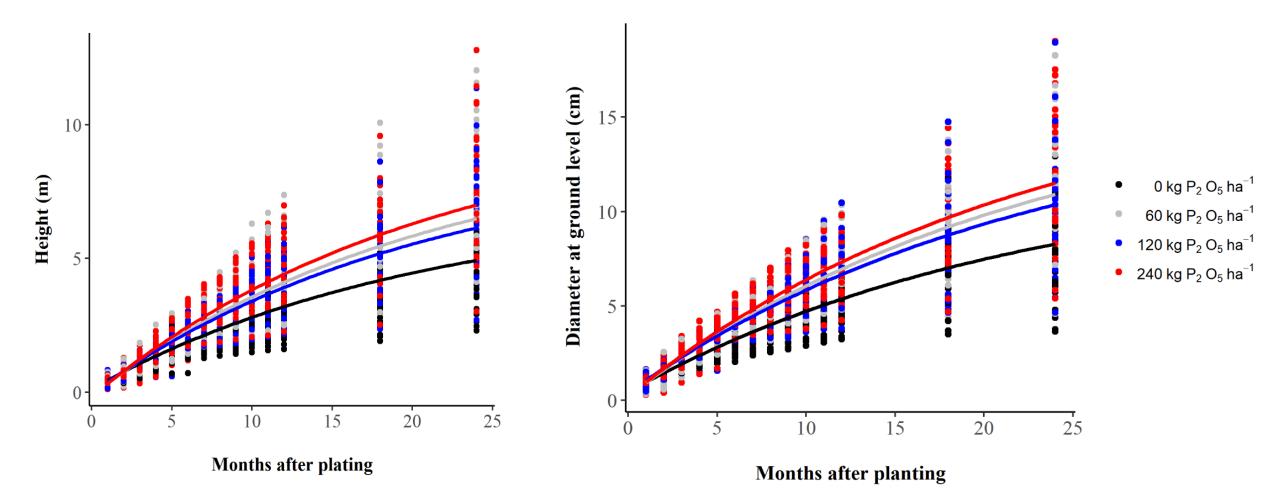
Model	eta_1	eta_2	eta_3	[kg P ₂ O ₅ ha ⁻¹]	Indicator variable	F-value	p-value
Monomolecular 15.563	15.562	0.964	0.044	0			
				60	0.02074	0.900	0.331
	15.505			120	0.009	0.500	0.461
				240	0.022	2.900	0.090
Logistic 10.094	10.004	8.147	4.269	0			
				60	0.995	9.351	0.002
	10.094			120	0.161	2.612	0.106
				240	0.771	8.869	0.003
Shumacher 12	10 471			0			
		c 1 40	142 -	60	1.547	26.347	<.0001
	12.471	0.142		120	0.242	5.416	0.020
				240	1.119	19.488	<.0001

Comparando modelos com variáveis indicativas para diâmetro

Monomolecular bloc/trat σ^2 .months (2δ)	Logistic bloc/trat σ^2 .months (2δ)	Shumacher bloc/trat σ^2 .months $^{(2\delta)}$
AIC = 4467.239	AIC = 4838.033	AIC = 5849.342



• Ajuste com modelos de efeito mistos



Referências

- Lappi, J. 1991. Calibration of height and volume equations with random parameters. For. Sci. 37: 781–801.
- Pinheiro and Bates, 2000. Mixed-effects models in S and S-plus.
- Burkhart, E.H.; Tomé, M. 2012. Modeling Forest Trees and Stands. Springer, New York, NY, USA.
- Colmanetti, MAA. 2018. Aboveground biomass and carbon estimate at Brazilian Atlantic Forest.
- R Development Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.