

A new method to compare statistical growth curves: The PL-GMANOVA model and its application

Martin Ricker¹, Víctor Peña², and Dietrich von Rosen³

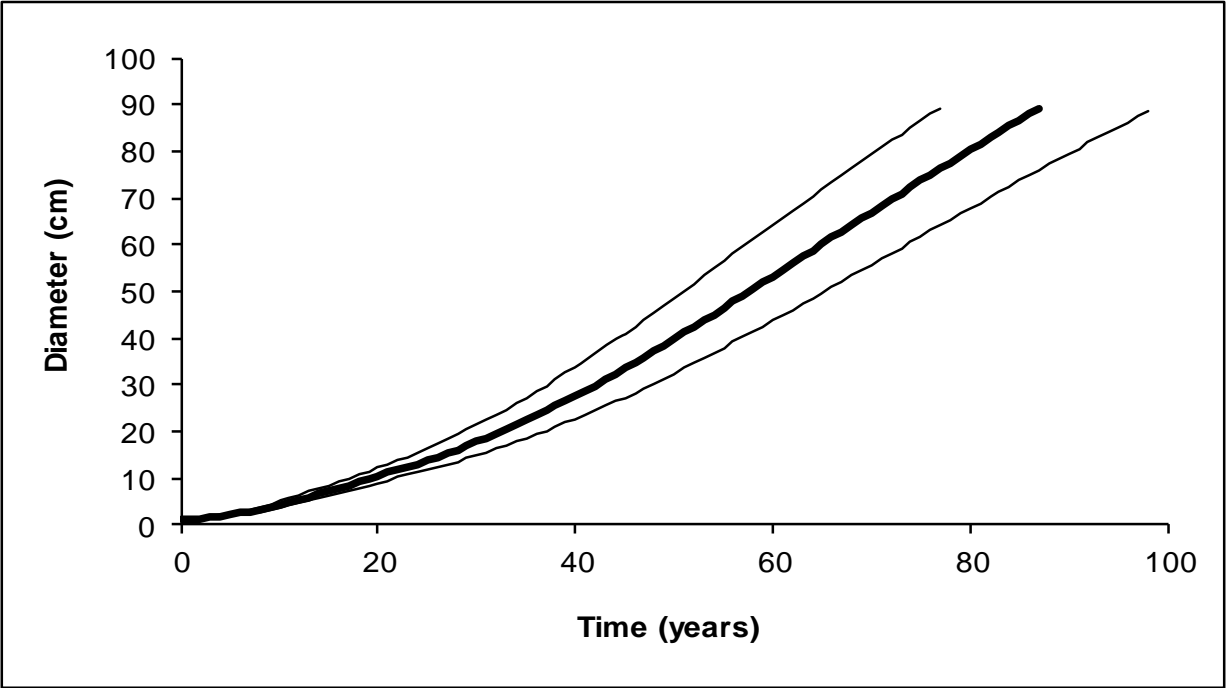
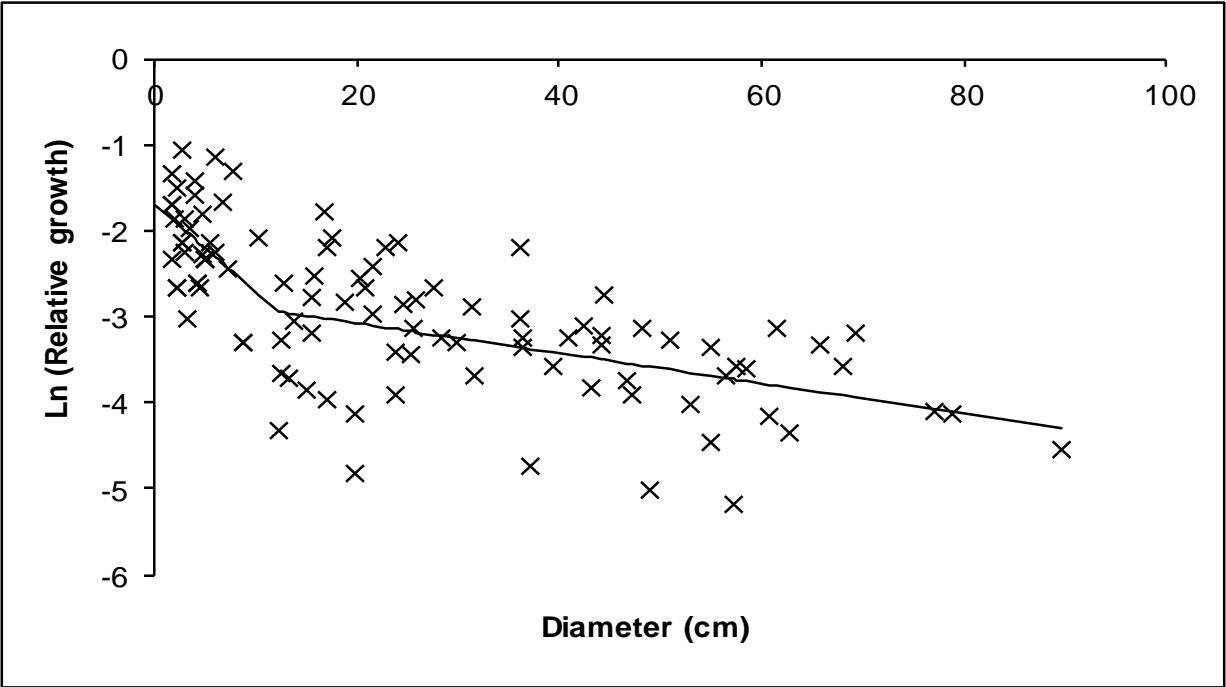
¹ Instituto de Biología, Universidad Nacional Autónoma de México (UNAM), Mexico City <mricker@ib.unam.mx>

² Posgrado en Ciencias Biológicas, UNAM, Mexico City

³ Department of Engineering and Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden



Terminalia amazonia
("tepsúchil")



Piecewise Linear (PL) model: $\ln[(dQ/dt) / Q] = a + b \cdot Q$

(dQ/dt) = instantaneous increment

$(dQ/dt) / Q$ = instantaneous relative growth

$$\Rightarrow t = t_1 + e^{-a} \cdot (Ei[-b \cdot Q] - Ei[-b \cdot Q_1])$$

$Ei(x) = \int (e^x/x) dx$ (all calculations with *Mathematica*)

a = initial logarithmic relative growth

Q_1 at t_1 = one known time-quantity point

with $b < 0 \Rightarrow$ sigmoid growth curve,
where Turning point quantity = $-1/b$

with $b = 0 \Rightarrow$ exponential growth of $Q(t)$

with $b > 0 \Rightarrow$ over-exponential growth

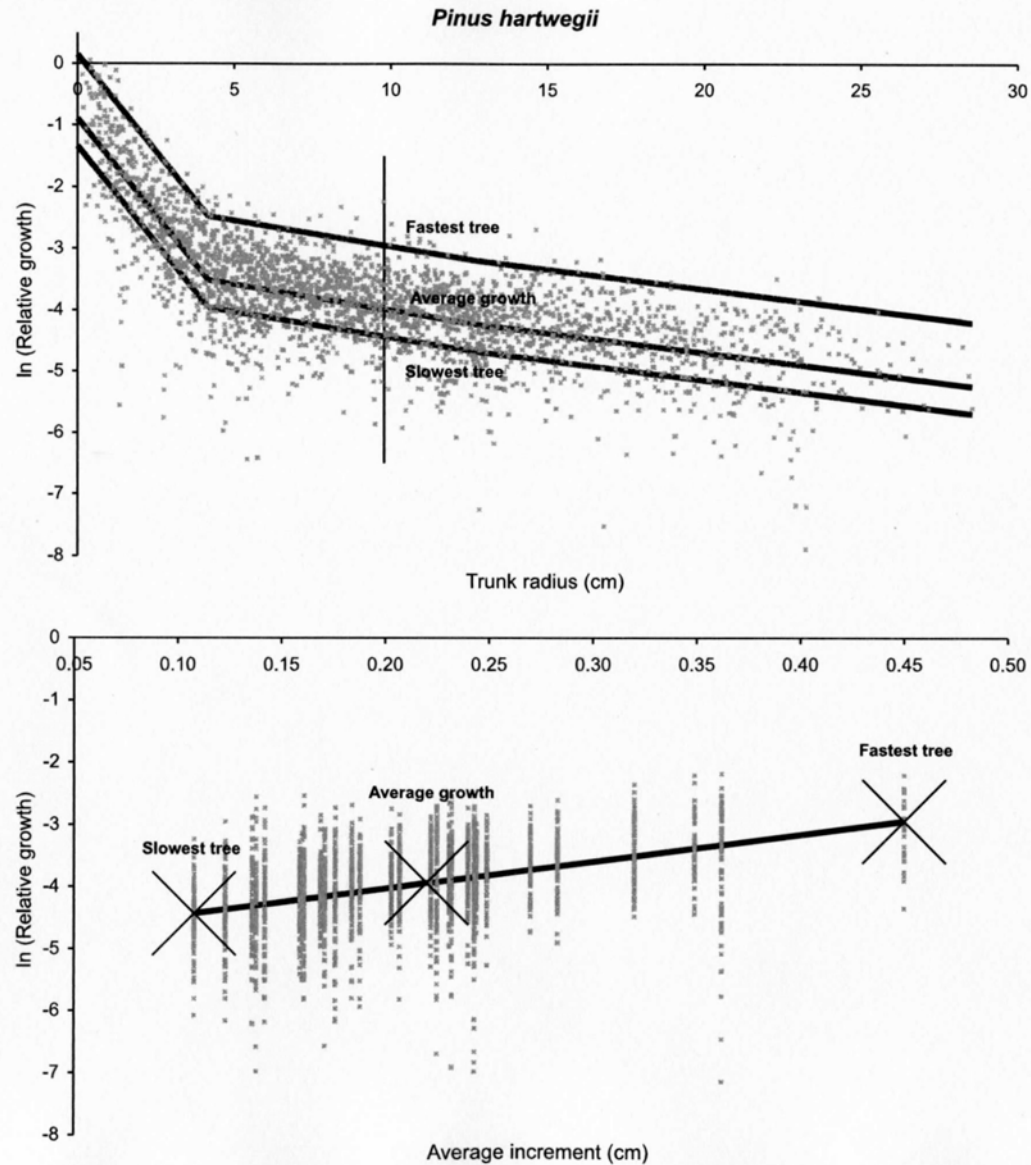
with a = large negative number, the growth curves
becomes almost linear



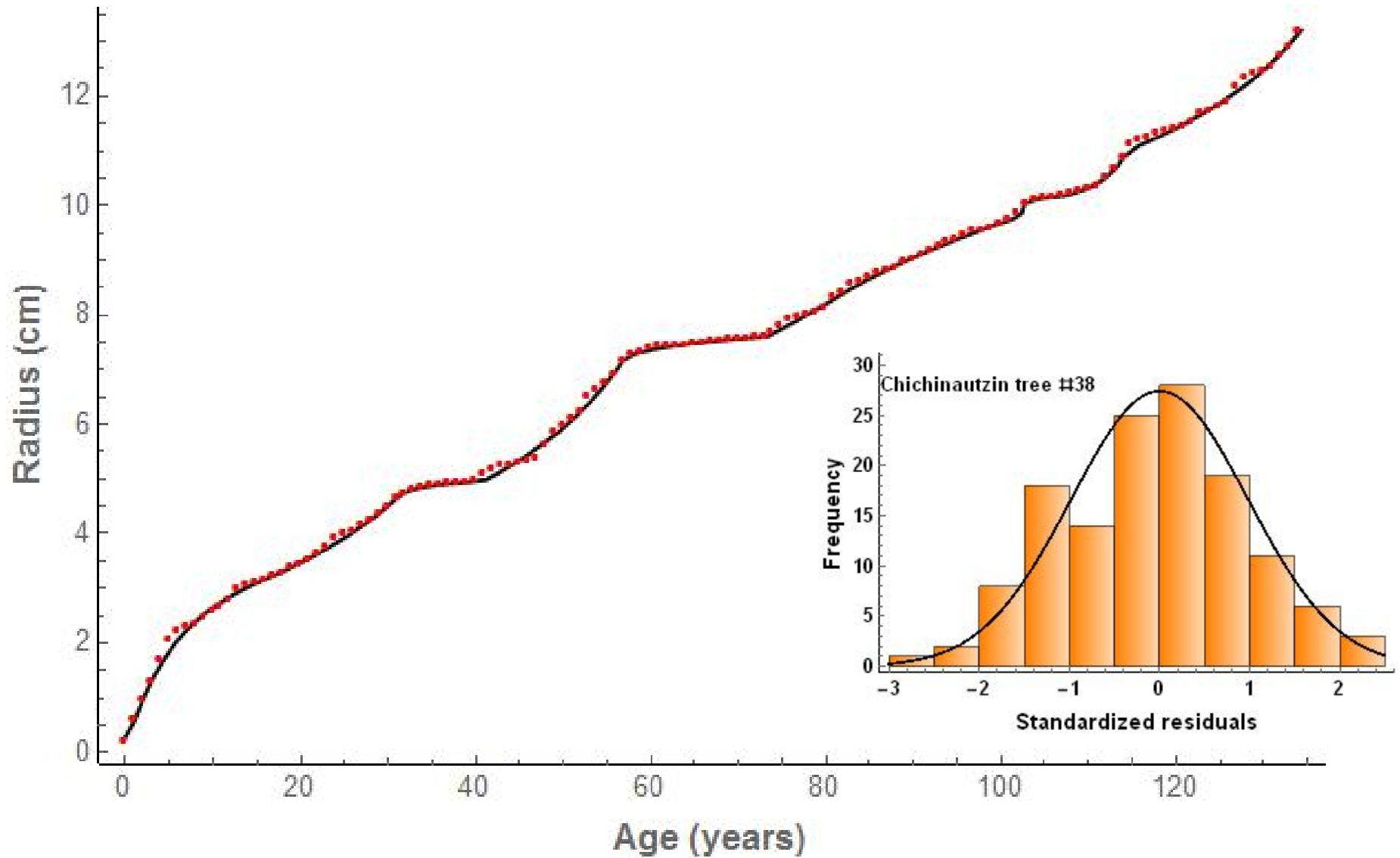
10 volcano sites in Mexico \approx 3100 m above sea level (data from Dr. Víctor Peña, UNAM, doctoral thesis 2013)

235 Pinus montezumae trees: Radius: 9.5 cm to 43.8 cm, Annual radius increment: 0.04 to 0.86 cm

Fig. 2. Piecewise linear regression of logarithmic relative growth as a function of radius, that derive the growth paths in Fig. 1 bottom. The average radius of all data points is indicated as a vertical black line in Fig. 2 top. If we cut through the graph along that black line, and rotate the top graph around by 90° , then we get to see the graph of Fig. 2 bottom. The regression lines of Fig. 2 top are now crosses. All data points are presented as standardized residuals.



Nine knots at 3.1, 4.7, 5.0, 7.2, 7.6, 8.5, 9.8, 10.2, and 11.0 cm



Generalized Multivariate Analysis of Variance (GMANOVA) or
 “classical growth curve model” established by Potthoff & Roy (1964)

$$Q = T \cdot A \cdot G + E$$

Q: $p \times n$ = Data matrix of **quantities to be explained** for p time points (rows) and all n individuals in all groups combined (columns).

T: $p \times 1$ = Within-individuals design matrix for p **time points** and one polynomial term.

A: $1 \times k$ = **Parameter** matrix for one polynomial term and k treatment groups.

G: $k \times n$ = The among-individuals design matrix for k treatment **groups** (rows) and n individuals (columns).

E: $p \times n$ = **Error** matrix with columns that are independent p -variate normal. The errors may be correlated and heteroscedastic.

$Q_{1,1}$	$Q_{1,2} \dots$	$Q_{1,12}$	T_1	$p = 6, n = 12, k = 3$
$Q_{2,1}$	$Q_{2,2} \dots$	$Q_{2,12}$	T_2	$\hat{A}_1 \quad \hat{A}_2 \quad \hat{A}_3$
$Q_{3,1}$	$Q_{3,2} \dots$	$Q_{3,12}$	T_3	
$Q_{4,1}$	$Q_{4,2} \dots$	$Q_{4,12}$	T_4	1 1 1 1 0 0 0 0 0 0 0 0
$Q_{5,1}$	$Q_{5,2} \dots$	$Q_{5,12}$	T_5	0 0 0 0 1 1 1 0 0 0 0 0
$Q_{6,1}$	$Q_{6,2} \dots$	$Q_{6,12}$	T_6	0 0 0 0 0 0 0 1 1 1 1 1

Formulas for GMANOVA model:

$$\hat{\mathbf{A}}_{GLSE}: 1 \times k = (\mathbf{T}' \cdot \mathbf{T})^{-1} \cdot \mathbf{T}' \cdot \mathbf{Q} \cdot \mathbf{G}' \cdot (\mathbf{G} \cdot \mathbf{G}')^{-1}$$

$$D[\mathbf{A}_{GLSE}]: k \times k = (\mathbf{G} \cdot \mathbf{G}')^{-1} \otimes (\mathbf{T}' \cdot \mathbf{T})^{-1} \cdot (\mathbf{T}' \cdot \boldsymbol{\Sigma} \cdot \mathbf{T}) \cdot (\mathbf{T}' \cdot \mathbf{T})^{-1}$$

$$\hat{\mathbf{A}}_{MLE}: 1 \times k = (\mathbf{T}' \cdot \mathbf{S}^{-1} \cdot \mathbf{T})^{-1} \cdot \mathbf{T}' \cdot \mathbf{S}^{-1} \cdot \mathbf{Q} \cdot \mathbf{G}' \cdot (\mathbf{G} \cdot \mathbf{G}')^{-1}$$

$$\text{with } \mathbf{S}: p \times p = \mathbf{Q} \cdot (\mathbf{I}_n - \mathbf{G}' \cdot (\mathbf{G} \cdot \mathbf{G}')^{-1} \cdot \mathbf{G}) \cdot \mathbf{Q}'$$

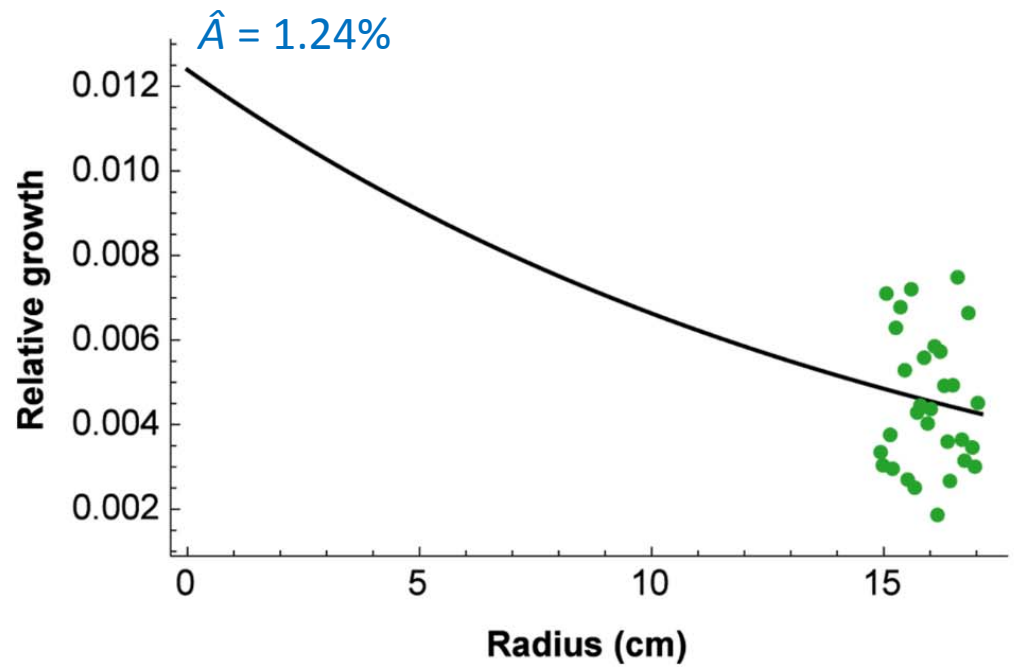
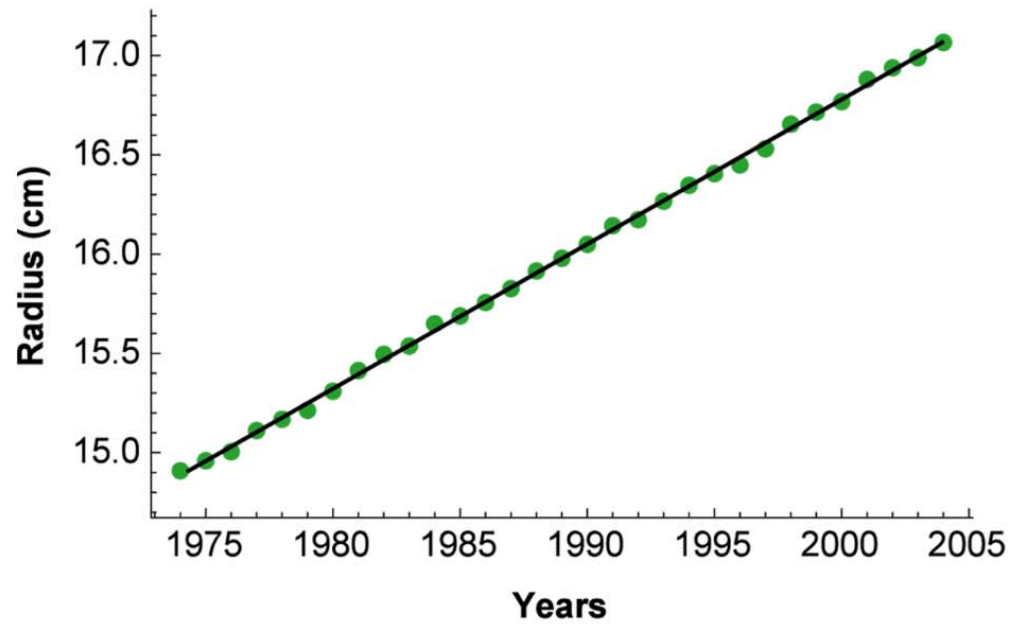
The PL-GMANOVA model:

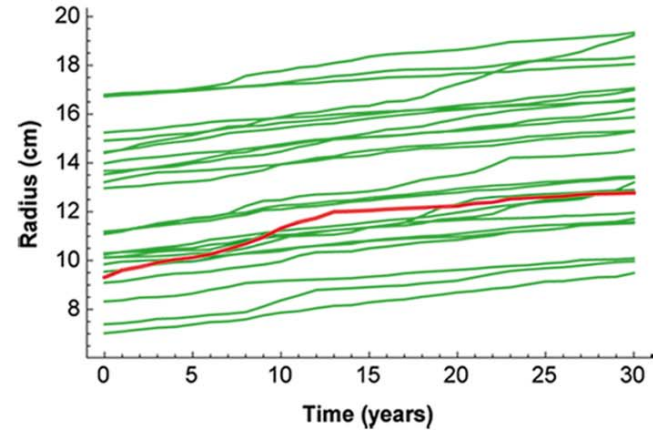
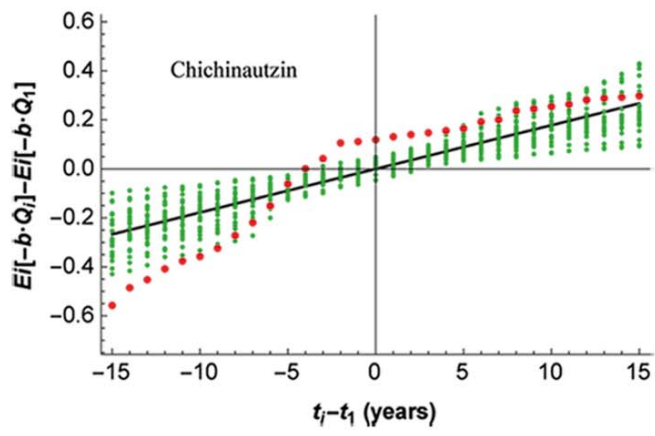
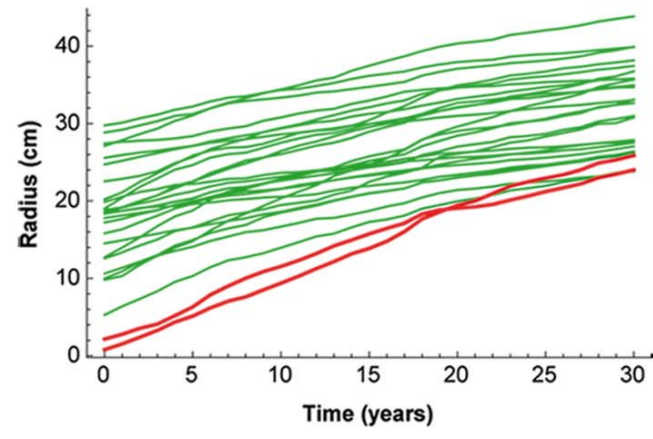
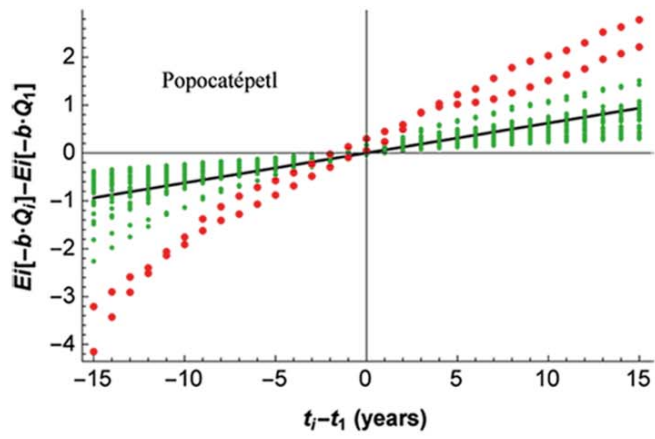
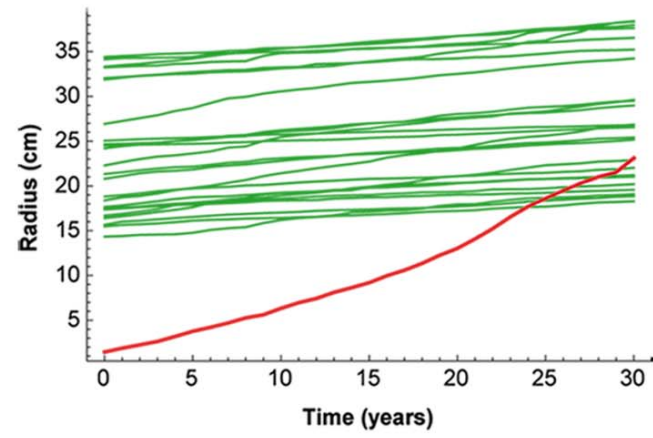
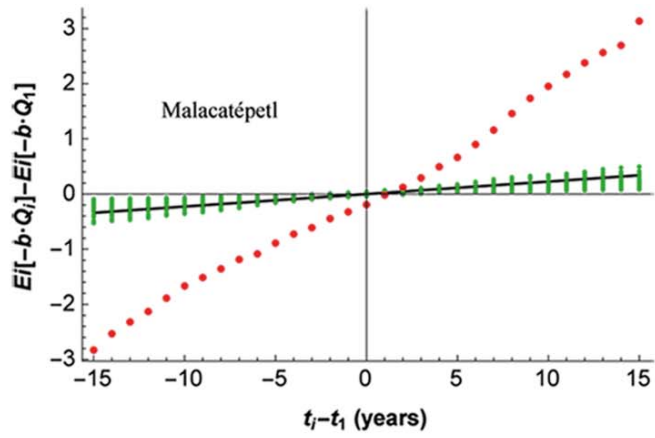
$$Q_{i,j} = A_l \cdot T_i + E_{i,j}$$

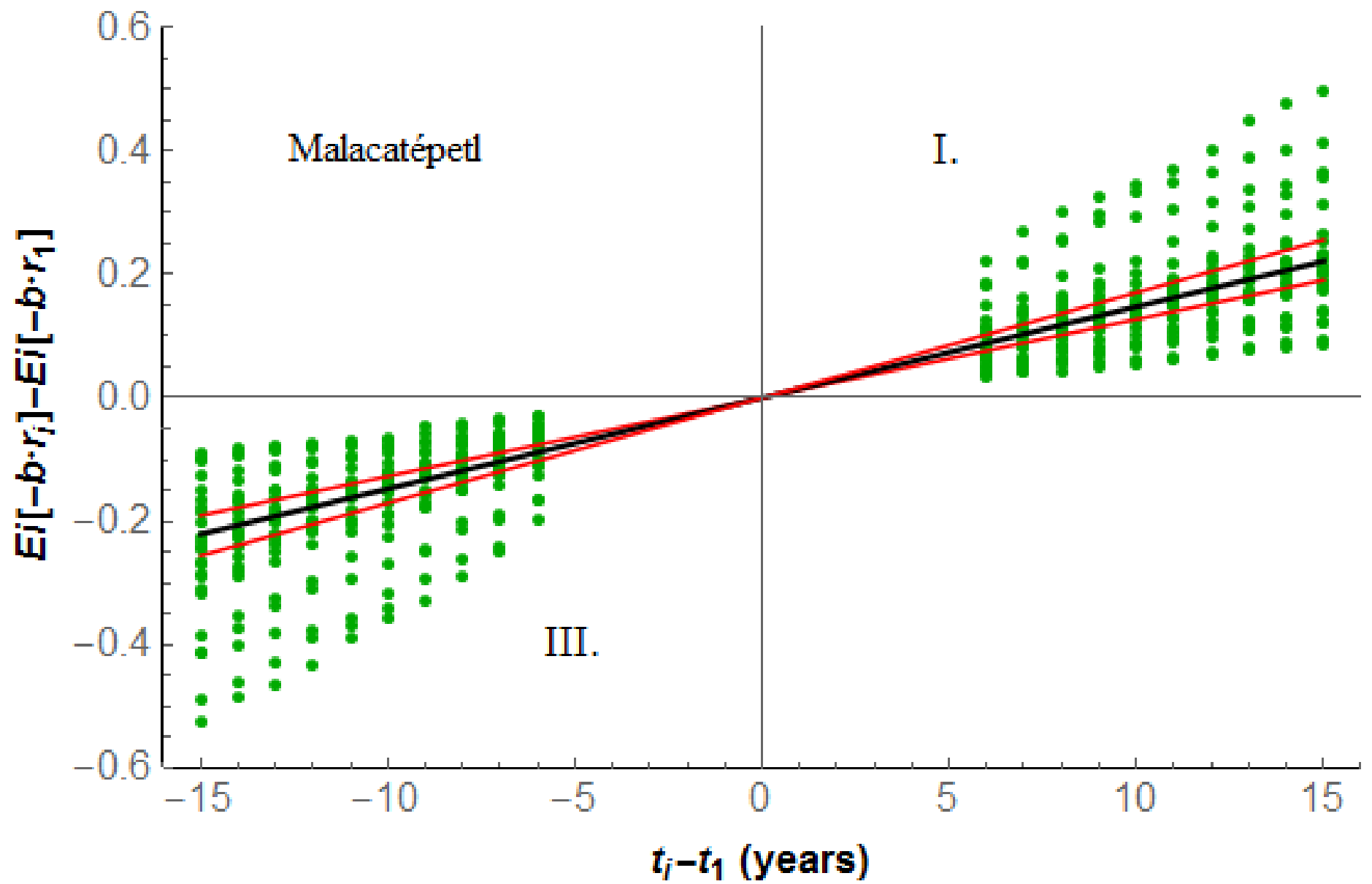
$Q_{i,j} = (Ei[-b_j \cdot r_{i,j}] - Ei[-b_j \cdot r_1])$ at time point i for tree j ; $Q_{i,j}$ has no physical dimension, representing a multiplicative change between radii r_1 and $r_{i,j}$ over the time period $T_i = (t_i - t_1)$.

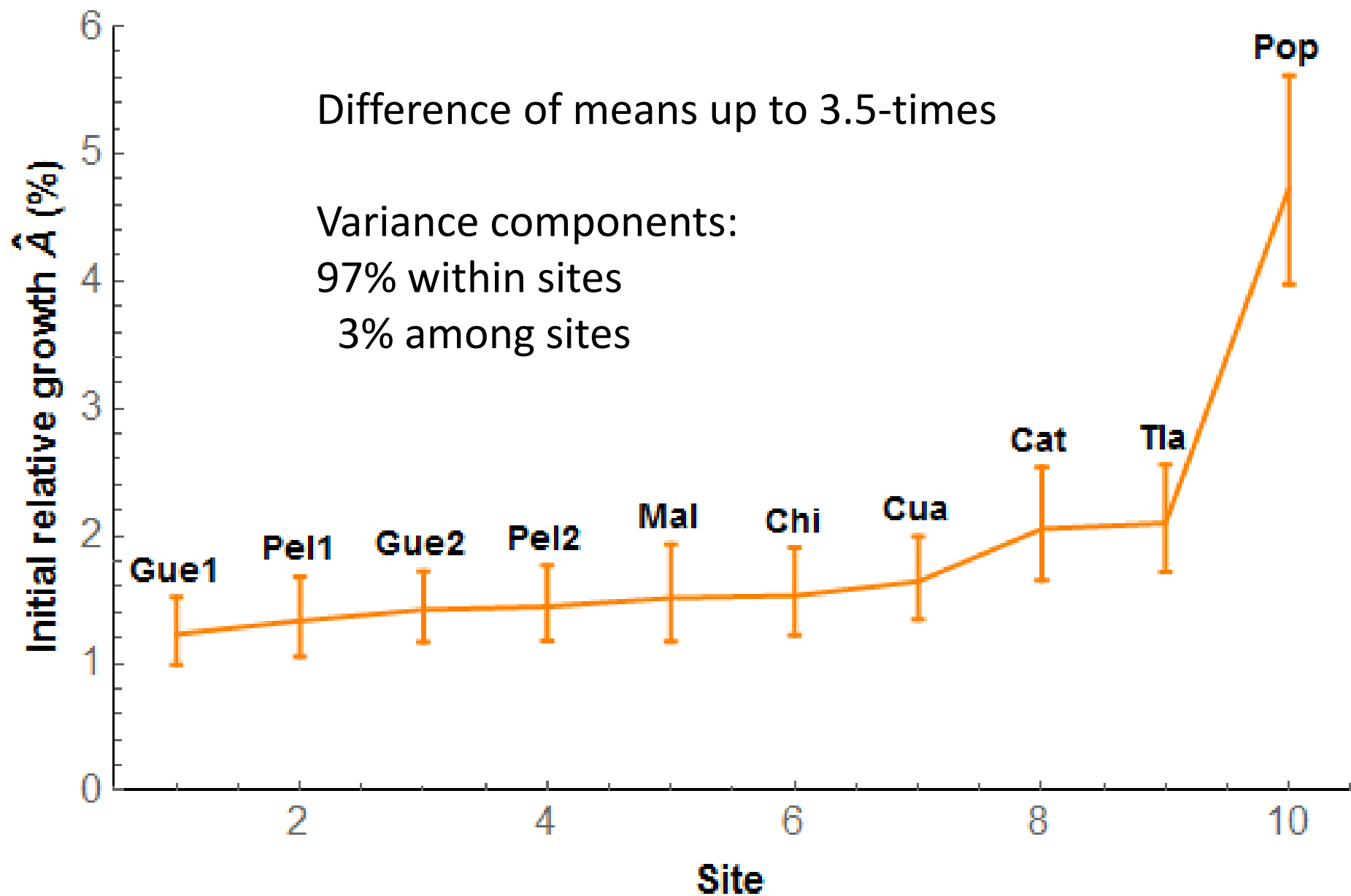
A_l = initial relative growth of the trees in group l .

$E_{i,j}$ = error term data of tree j at time point i .









Advantages of the approach of combining the PL model with the GMANOVA model:

From PL model:

- Highly **flexible and accurate growth curve modeling** (although no direct modeling of negative or cycling growth modeling).
- Clear and useful **interpretation of the coefficients**.
- **Comparison** of growth rates among sigmoid growth curves **with different turning points and starting points**.

From GMANOVA model

- **Residuals** may be **autocorrelated** and **heteroscedastic** .
- **No “arbitrary” choice of the number of polynomial terms** in GMANOVA.
- **Lower standard errors** with GMANOVA, when compared with individually derived growth curves.

From combination approach:

- Hidden **outliers** are easily detectable.
- Statistically **robust** method (regression through origin).

Future research:

Modeling **more than one segment** simultaneously.

Taking **several independent (explaining) variables** into account.

Considering **missing values**.



Pinus lambertiana (Pinaceae)



Ceiba pentandra (Bombacaceae)



Dalbergia retusa (Leguminosae)

