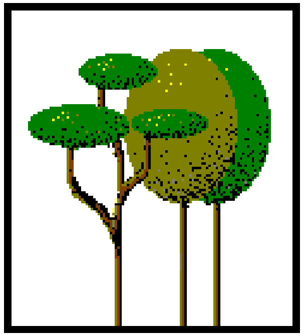
Applications of mixed models in forest growth modelling and breeding value prediction



Salvador A. Gezan
University of Florida

August 2013

OUTLINE

PART 1

Estimation and calibration of height-diameter models for mixed second growth forest in Chile

PART 2

Genetic analysis of tree clonal trials (CCLONES):

a) Combining individual and clonal information.
b) Validating greenhouse experiments over field experiments for rust resistance.

PART 1

Estimation and calibration of height-diameter models for mixed second growth forest in Chile

OBJECTIVES

1) Obtain general H-D models for three hardwood species to be incorporated into a growth and yield simulator.

2) Present and implement the use of calibrated models through the linear mixed model framework.

$$\mathbf{H_{ij} = f(DBH_{ij}, s_i) + e_{ij}}$$

HEIGHT-DIAMETER MODELS

BACKGROUND

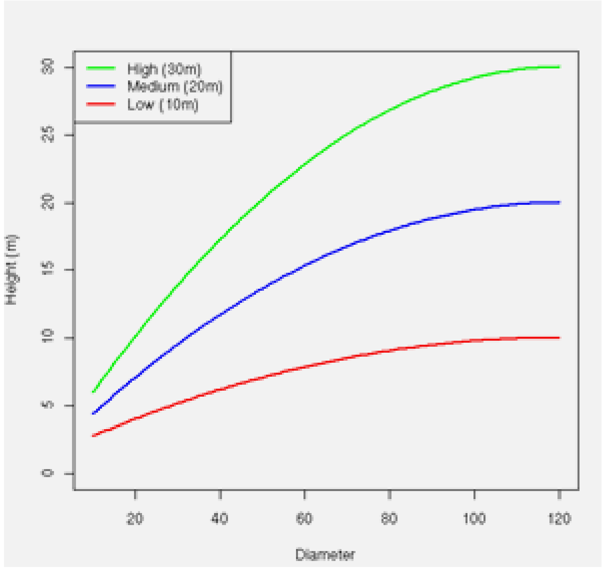
- Height-diameter models (HDM) are constructed to estimate the height of those trees on which diameter (DBH) is known, usually as part of an operational inventory. Data originates from sub-sampling of trees.

- HDM are used to estimate individual (or stand) volume and/or biomass.

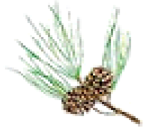
- Characteristics of these curves will depend on geographical conditions but also on stand variables (NHA, G, HD, etc.), species composition, age, atrophic alteration, etc.

- General HDM combine information from several stands into a single model using DBH together with other covariates to adjust predictions according to the characteristics of the stand.

- Future use will be for growth simulators, in prediction of yields or "filling" of data from an inventory



STUDY AREA



- **Sampling area** (709,854 ha)

  37°30’ – 42°00’ S and 71°33’ – 73°3’ W
- **Second growth forests** (mixed or pure)
- **Species composition** (Fagaceae):

Nothofagus alpina (Raulí)	1%
Nothofagus obliqua (Roble)	22%
Nothofagus dombeyi (Coigue)	36%
Mixed forests	41%

- **Age:** 15 - 80 years
- **Growth:** 6 - 10 m^3/ha./year (Natural Forests)

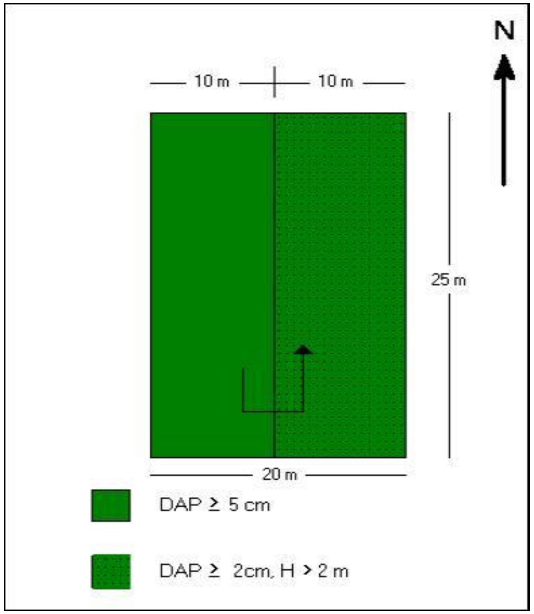
  10 - 18 m^3/ha./year (Plantation)

COLLECTED DATA

FIELD SAMPLING

128 Permanent plots
- 250 or 500 m^2/ha.
- Sample of age and height

60 Felling temporary plots
- 250 m^2/ha.
- 4 felled trees (stem analysis)



VARIABLES (individual)

Species (SP)
DBH
Dominance Status (PS)
Health (S)
Form (F)

VARIABLES (sub-sample)

Increment Core
Total Height (HT)
Beginning Crown Height (HCC)
Upper Diameter (1/3 HCC)

VARIABLES (stand level)

Dominant Height (HD)
Basal Area (BA)
Trees Per Hectare (TPH)
Quadratic Diameter (QD)

HEIGHT-DIAMETER MODELS

DATA ANALYSIS

- Only felled trees were considered.
- A minimum of five trees per plot x species were considered.
- **Fitting dataset:** 90 plots (with 476 trees)
- **Validation dataset:** 86 plots (with 832 trees)

HEIGHT DIAMETER MODEL - Curtis (1967)

$$\ln(H_i - 1.3) = \alpha + \beta \, DBH_i^{p} + e_i$$

where,

H_i is the height of the ith tree, $i = 1, \ldots, n$;
DBH_i is the breast height diameter;
α, β, p are parameters to be estimated;
e_i is the residual error, with $e_i \sim N(0, \sigma^2)$;
n is the total number of observations.

Note

- The parameter p was initially fixed at -0.61 for all models, which was obtained by minimizing the sum of squares for the full dataset.

HEIGHT-DIAMETER MODEL

LOCAL HEIGHT-DIAMETER MODEL

$$\ln(H_i - 1.3) = \alpha + \beta \, DBH_i^{\,p} + e_i$$

GENERAL HEIGHT-DIAMETER MODEL (only fixed effects)

$$\ln(H_{ij} - 1.3) = \alpha + \beta_1 \, DBH_{ij}^{\,p} + \beta_2 \, x_{1,i} + \beta_3 \, x_{2,i} + \beta_4 \, x_{3,i} + e_{ij}$$

GENERAL HEIGHT-DIAMETER MODEL (with random effects)

$$\ln(H_{ij} - 1.3) = (\alpha + a_j) + (\beta_1 + b_j) \, DBH_{ij}^{\,p} + \beta_2 \, x_{1,i} + \beta_3 \, x_{2,i} + \beta_4 \, x_{3,i} + e_{ij}$$

where,

$x_{1,i}$, $x_{2,i}$ and $x_{3,i}$ are stand covariates (HD, TPH, G, or QD)
α, β_1, β_2, β_3, β_4 are fixed parameters;
a_j and b_j are random parameters;
j is the plot/stand index, j = 1, ..., s;
s is the number of plots/stands.

LINEAR MIXED MODELS

Mixed models extend the linear model by allowing a more flexible specification of the errors (and other random factors). Hence, it allows for a different type of inference and also allows to incorporation of *correlation* and *heterogeneous variances* between the observations.

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$$

$$\mathrm{E}(\mathbf{u}) = \mathrm{E}(\mathbf{e}) = 0$$
$$\mathrm{E}(\mathbf{y}) = \mathbf{X}\beta$$

$$\mathrm{Var}(\mathbf{u}) = \mathbf{G}$$
$$\mathrm{Var}(\mathbf{e}) = \mathbf{R}$$
$$\mathrm{Var}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$$

X (n x p) design matrix for fixed effects

β (p x 1) vector of fixed effects

Z (n x q) design matrix for random effects

u (q x 1) vector of random effects

e (n x 1) vector of random errors

G (q x q) matrix of variance-covariance of random effects

R (n x n) matrix of variance-covariance of random errors

Example **yield = μ + block + variety + error**

HEIGHT-DIAMETER MODEL

GENERAL HEIGHT-DIAMETER MODEL (with random effects)

$$\ln(\mathbf{H}_{ij} - 1.3) = (\boldsymbol{\alpha} + \mathbf{a}_j) + (\boldsymbol{\beta}_1 + \mathbf{b}_j)\,\mathbf{DBH}_{ij}{}^{p} + \boldsymbol{\beta}_2\,\mathbf{x}_{1,i} + \boldsymbol{\beta}_3\,\mathbf{x}_{2,i} + \boldsymbol{\beta}_4\,\mathbf{x}_{3,i} + \mathbf{e}_{ij}$$

$$\ln(\mathbf{H}_{ij} - 1.3) = \boldsymbol{\alpha} + \boldsymbol{\beta}_1\,\mathbf{DBH}_{ij}{}^{p} + \boldsymbol{\beta}_2\,\mathbf{x}_{1,i} + \boldsymbol{\beta}_3\,\mathbf{x}_{2,i} + \boldsymbol{\beta}_4\,\mathbf{x}_{3,i} + \mathbf{a}_j + \mathbf{b}_j\,\mathbf{DBH}_{ij}{}^{p} + \mathbf{e}_{ij}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where,

y (n x 1) vector of response, **y** = ln(**H** – 1.3)

X (n x 4) design matrix for fixed effects, **X** = [$\mathbf{1}_n$ $\mathbf{DBH}^p$ $\mathbf{x}_1$ $\mathbf{x}_2$ $\mathbf{x}_3$]

β (4 x 1) vector of fixed effects, **β** = [$\boldsymbol{\alpha}$ $\boldsymbol{\beta}_1$ $\boldsymbol{\beta}_2$ $\boldsymbol{\beta}_3$ $\boldsymbol{\beta}_4$]’

Z (n x 2s) design matrix for random effects, **Z = [**$\mathbf{1}_s$ $\mathbf{DBH}^p$]

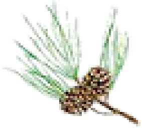
u (2s x 1) vector of random effects, **u = [a b]**’

e (n x 1) vector of random errors

G (2s x 2s) matrix of variance-covariance of random effects

R (n x n) matrix of variance-covariance of random errors

G MATRIX



G is a matrix of variance-covariance of the random effects **u**

$$\mathbf{u} \sim \mathrm{MVN}(0, \mathbf{G}) \qquad \mathrm{E}(\mathbf{u}) = \mathbf{0},\ \mathbf{V}(\mathbf{u}) = \mathbf{G} = \mathbf{G}(\boldsymbol{\theta})$$

Used to define correlations among effects, and/or heterogeneity of variances.

GENERAL HEIGHT-DIAMETER MODEL (with random effects)

$$\mathbf{ln(H_{ij}} - 1.3) = (\boldsymbol{\alpha} + \mathbf{a_j}) + (\boldsymbol{\beta}_1 + \mathbf{b_j})\,\mathbf{DBH_{ij}^{p}} + \boldsymbol{\beta}_2\,\mathbf{x}_{1,\mathbf{i}} + \boldsymbol{\beta}_3\,\mathbf{x}_{2,\mathbf{i}} + \boldsymbol{\beta}_4\,\mathbf{x}_{3,\mathbf{i}} + \mathbf{e_{ij}}$$

$$\mathbf{u} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{b}_1 \\ \mathbf{a}_2 \\ \mathbf{b}_2 \\ \cdot \\ \cdot \\ \mathbf{a_s} \\ \mathbf{b_s} \end{bmatrix} \qquad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{G}_2 & \ldots & 0 \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & \mathbf{G_s} \end{bmatrix} \qquad \mathbf{G_j} = \begin{bmatrix} \boldsymbol{\sigma}_{\mathbf{a}}^2 & \boldsymbol{\sigma}_{\mathbf{ab}} \\ \boldsymbol{\sigma}_{\mathbf{ab}} & \boldsymbol{\sigma}_{\mathbf{b}}^2 \end{bmatrix}$$

HEIGHT-DIAMETER MODEL

STATISTICAL ANALYSIS

- A different model was fitted for each species separately.
- Types of models fitted:
  - a) *Local*
  - b) *With Dominant Height (HD)*
  - c) *Without Dominant Height*
- Models were fitted using REML as implemented in PROC MIXED (SAS).
- Several alternative combinations/transformations of the covariates $\mathbf{x}_k$ were evaluated (selected using a Wald test).
- Significance of random effects was evaluated using a LRT (Wolfinger 1996).
- A ‘fine tuning’ of the **p** parameter was performed by species, in final selected models.
- Baskerville (1972) correction was implemented due to bias originated in the back-transformation of log-transformed response.

HEIGHT-DIAMETER MODEL

COMPARISON OF MODELS

- Statistics obtained for the totality of observations in each species, and for three size classes (DBH 0 to 15, 15 to 25, and 25 to 60 cm).
- Original response used for evaluation of models, using

$$\hat{H}_{ij} = 1.3 + \exp[(\alpha + a_j) + (\beta_1 + b_j)\, DBH_{ij}^{\,p} + \beta_2\, x_{1,i} + \beta_3\, x_{2,i} + \beta_4\, x_{3,i}]$$

STATISTICS FOR FITTING AND VALIDATION DATASETS

$$RECM\,\% = \frac{100}{\bar{\bar{H}}}\left[\frac{1}{n}\sum_{j=1}^{s}\sum_{i=1}^{nj}\left(H_{ij} - \hat{H}_{ij}\right)^2\right]^{1/2}$$

$$BIAS\,\% = \frac{100}{\bar{\bar{H}} \times n}\sum_{j=1}^{s}\sum_{i=1}^{nj}\left(H_{ij} - \hat{H}_{ij}\right)$$

$$r_{emp}^{\;2} = \frac{\sum_{j=1}^{s}\sum_{i=1}^{nj}\left(\hat{H}_{ij} - \bar{H}\right)^2}{\sum_{j=1}^{s}\sum_{i=1}^{nj}\left(H_{ij} - \bar{H}\right)^2}$$

RESULTS

FITTING DATASET

Species	Type	x_1	x_2	x_3	a_j	b_j
RAULI	**Local**	-	-	-	Yes	-
	With HD	ln(HD)	-	-	Yes	-
	Without HD	1/BA	ln(TPH)	-	Yes	-
ROBLE	**Local**	-	-	-	Yes	Yes
	With HD	ln(QD)	QD/HD	-	Yes	Yes
	Without HD	ln(QD)	-	-	Yes	Yes

		WITHOUT RANDOM EFFECTS			**WITH RANDOM EFFECTS**		
Species	**Type**	**BIAS%**	**RECM%**	r_{emp}^2	**BIAS%**	**RECM%**	r_{emp}^2
RAULI	Local	-1.1%	20.5%	0.593	0.2%	9.9%	0.906
	With HD	0.5%	12.8%	0.841	0.0%	9.6%	0.911
	Without HD	-0.1%	15.1%	0.780	0.0%	9.7%	0.908
ROBLE	Local	2.1%	22.5%	0.698	0.1%	7.8%	0.964
	With HD	1.5%	13.9%	0.884	0.0%	8.8%	0.954
	Without HD	0.2%	19.6%	0.771	0.1%	7.8%	0.963

RESULTS

FITTING DATASET

- Rauli required only a random effect for intercept.
- Roble required both, slope and intercept random effects. High negative correlations were found between these parameters (e.g. Without HD, r^2 = -0.81)
- Similar trends of BIAS% and RECM% observed for different size classes (not shown).

	α	β_1	β_2	β_3	s_α^2	s_β^2	$s_{\alpha\beta}$	s^2
RAULI								
Local	3.9131	-3.7381	-	-	0.3518	-	-	0.0154
With HD	1.5537	-3.7686	0.7764	-	0.0833	-	-	0.0135
Without HD	5.4191	-3.7259	-11.2766	-0.1623	0.1520	-	-	0.0141
ROBLE								
Local	3.6763	-3.8144	-	-	0.1289	-0.2892	0.9937	0.0125
With HD	3.0726	-3.7728	0.5886	-1.3477	0.1216	-0.4403	1.5981	0.0128
Without HD	2.5567	-3.6433	0.4025	-	0.0769	-0.2128	1.0097	0.0124

MODEL CALIBRATION

- Calibration is a technique in which the estimated parameters of a previously fitted model are *modified* or *calibrated* using the information available from a new sample of observations.
- It corresponds to a model combining old with new information.
- One approach, within the framework of linear mixed models, involves estimating random effects from a new sample and incorporation of this in to the previously fitted model.
- Assumes the variance components of the original fitted model are known without error.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

$$\begin{aligned}\hat{\mathbf{u}} &= \mathbf{G}\mathbf{Z}'\,\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= [\mathbf{Z}'\,\mathbf{Z} + \hat{\boldsymbol{\sigma}}^{2}\hat{\mathbf{G}}^{-1}]^{-1}\,\mathbf{Z}'\,(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})\end{aligned}$$

MODEL CALIBRATION

$$\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$
$$= [\mathbf{Z}'\mathbf{Z} + \hat{\sigma}^2\hat{\mathbf{G}}^{-1}]^{-1}\mathbf{Z}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

Model with only a random effects

$$\hat{u}_j = \hat{a}_j = \left[n_j + \frac{s^2}{s_\alpha^2}\right]^{-1} \times \sum_{i=1}^{nj}(y_{ij} - \hat{y}_{ij}) = \frac{s_\alpha^2}{s_\alpha^2 + s^2/n_j} \times \frac{1}{n_j}\sum_{i=1}^{nj}(y_{ij} - \hat{y}_{ij})$$

Model with a and b random effects

$$\hat{u}_j = \begin{bmatrix} \hat{a}_j \\ \hat{b}_j \end{bmatrix} = \left[\begin{bmatrix} n_j & \sum_{i=1}^{nj} DBH_{ij}^p \\ \sum_{i=1}^{nj} DBH_{ij}^p & \sum_{i=1}^{nj} (DBH_{ij}^p)^2 \end{bmatrix} + s^2 \begin{bmatrix} s_\alpha^2 & s_{\alpha\beta} \\ s_{\alpha\beta} & s_\beta^2 \end{bmatrix}^{-1}\right]^{-1} \begin{bmatrix} \sum_{i=1}^{nj}(y_{ij} - \hat{y}_{ij}) \\ \sum_{i=1}^{nj} DBH_{ij}^p \times (y_{ij} - \hat{y}_{ij}) \end{bmatrix}$$

RESULTS

VALIDATION DATASET

- Calibration was performed on the validation dataset using original fitted models.
- Predictions based on models with and without calibration.
- Improvements are evident, in all models, however, results are not as good as with the fitting dataset.

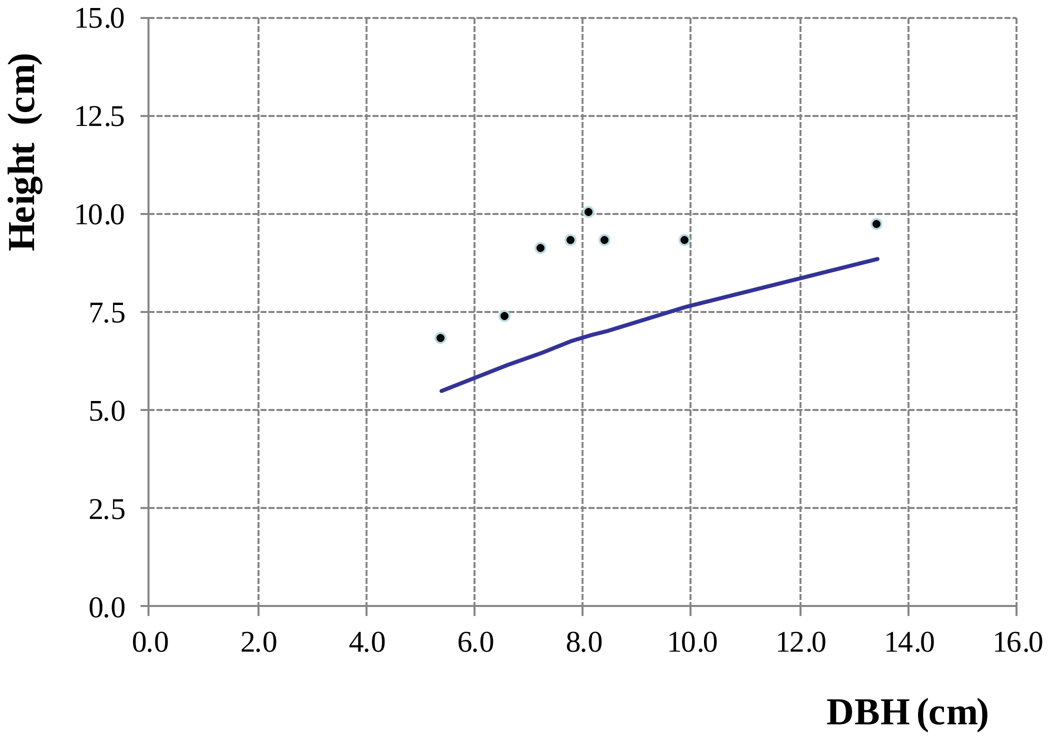
SPECIES	**Type**		**BIAS%**	**RECM%**	r_{emp}^2
RAULI	Without Calibration	Local	-6.0%	24.2%	0.603
		With HD	0.0%	18.3%	0.772
		Without HD	-0.9%	21.8%	0.673
	With Calibration	Local	0.9%	13.3%	0.862
		With HD	1.4%	12.9%	0.869
		Without HD	1.4%	13.1%	0.863
ROBLE	Without Calibration	Local	3.6%	27.2%	0.604
		With HD	0.1%	18.6%	0.814
		Without HD	-1.6%	22.1%	0.738
	With Calibration	Local	1.4%	13.2%	0.881
		With HD	-2.9%	15.4%	0.839
		Without HD	1.4%	13.3%	0.881

EXAMPLE

PLOT 4024

SPECIES: RAULI
BA: 10.76 m^2/ha
TPH: 1920 trees/ha
QD: 8.45 cm
HD: 0.15 m

ID	DBH	HT
61	5.40	6.80
65	6.60	7.35
59	7.25	9.10
60	7.80	9.30
64	8.15	10.00
63	8.45	9.30
62	9.90	9.30
66	13.45	9.70

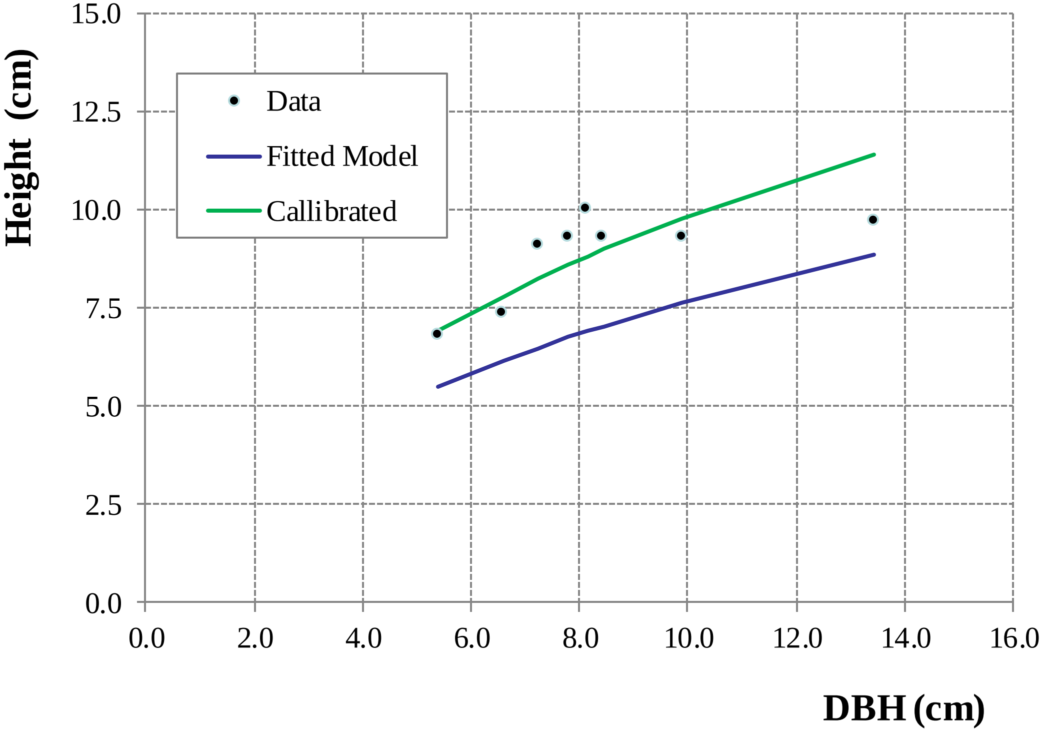


EXAMPLE

PLOT 4024

SPECIES: RAULI
BA: 10.76 m^2/ha
TPH: 1920 trees/ha
QD: 8.45 cm
HD: 0.15 m

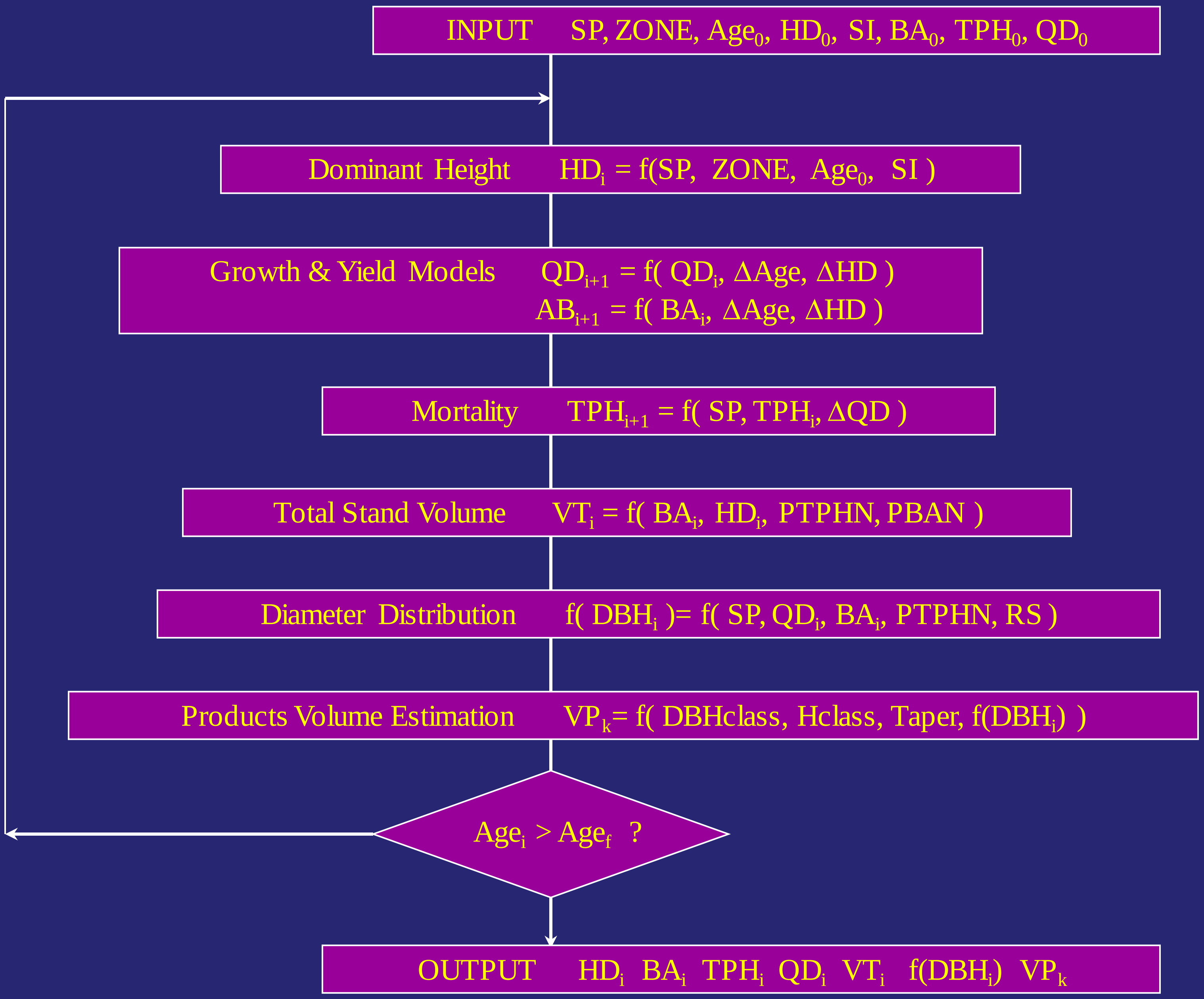
ID	DBH	HT
61	5.40	6.80
65	6.60	7.35
59	7.25	9.10
60	7.80	9.30
64	8.15	10.00
63	8.45	9.30
62	9.90	9.30
66	13.45	9.70



CONCLUSIONS

PART 1

- The incorporation of random effects into the models, allows to properly describe the data at the plot level.
- Stand variables help in describing the shape of the height-diameter curves.
- The use of Barkerville's correction reduced the presence of bias in these models.
- Calibration produced a considerable improvement in the predictions of heights.
- The use of calibration can be easily extended on other applications allowing the use of limited information.



INPUT $SP, ZONE, Age_0, HD_0, SI, BA_0, TPH_0, QD_0$

Dominant Height $HD_i = f(SP, ZONE, Age_0, SI)$

Growth & Yield Models $QD_{i+1} = f(QD_i, \Delta Age, \Delta HD)$

$AB_{i+1} = f(BA_i, \Delta Age, \Delta HD)$

Mortality $TPH_{i+1} = f(SP, TPH_i, \Delta QD)$

Total Stand Volume $VT_i = f(BA_i, HD_i, PTPHN, PBAN)$

Diameter Distribution $f(DBH_i) = f(SP, QD_i, BA_i, PTPHN, RS)$

Products Volume Estimation $VP_k = f(DBHclass, Hclass, Taper, f(DBH_i))$

$Age_i > Age_f$?

OUTPUT HD_i BA_i TPH_i QD_i VT_i $f(DBH_i)$ VP_k

PART 2

CCLONES: Loblolly Pine clonal trials

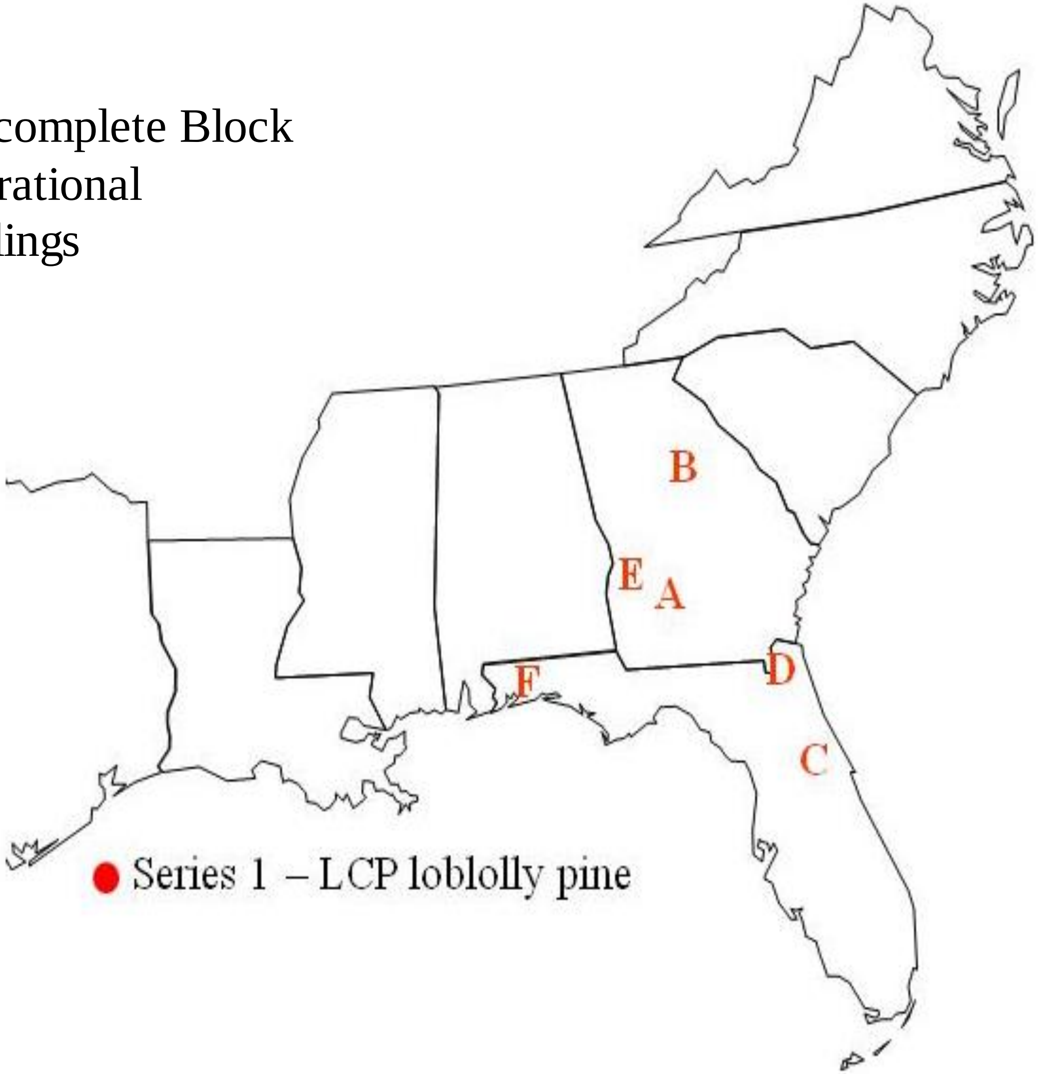
1) Compare the growth and disease resistance of rooted cuttings and seedlings grown under operational and intensive culture.
2) Understand the genotypes (parents, full-sib families & clones) and the genetic architecture of important growth strategies, disease resistance, nutritional efficiency and wood properties.

CCLONES Series 1 Data

FBRC Experiments

Design:	Resolvable Incomplete Block
Treatments:	Intensive/Operational
Propagules:	Cuttings/Seedlings

# sites:	7
# replicates	8
# families	~61
#parents	~32
#clones	~947
#clones/family	~17
#plants	47,040



PART 2(A)

Genetic analysis of tree clonal trials: combining individual and clonal information into ASReml

OBJECTIVES

1) To evaluate statistically if is possible to combine the genetic effects (additive and dominance) estimated using seedlings and cuttings propagules and to understand their level of association.

2) Estimate correlation matrices for phenotypic and genetic effects between a series of size and crown traits in order to facilitate understanding of the their genetic association.

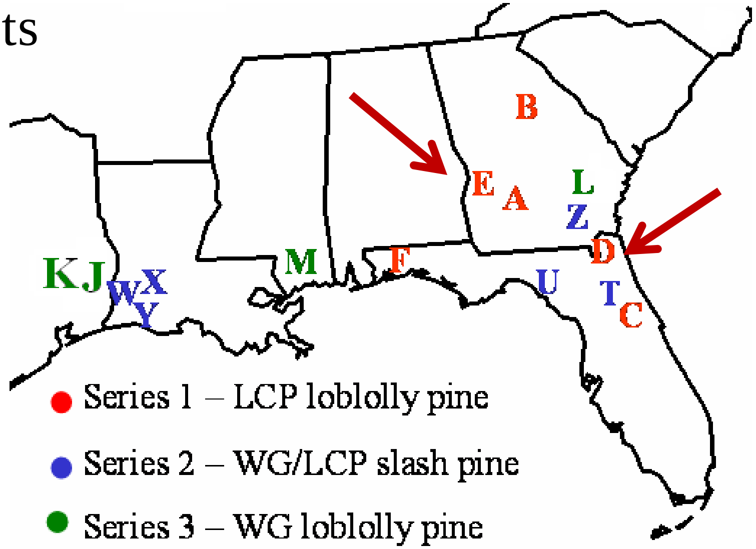
FIELD EXPERIMENTS

Series 1: Growth and canopy measurements collected from sites:

E: Cuthbert

D: Nassau

Trial	Cuthbert	Nassau
Design Type	IB	IB
# Trees	4400	4480
# Reps	4	4
Rep Size	1100	1120
# IBlocks/rep	110	80
Block Size	10	14
# Families	61	61
# Clones	868	947
# Clones/Family	5-20	16-18
Prop. Missing	38.8%	11.8%



VARIABLES

VOL: total stem volume
VCROWN crown volume
MDIAMB diameter of branches
MANG angle of branches
CW crown width
RLC ratio of live crown to total height
RVCR ratio of VCROWN to VOL
LC length of live crown

STATISTICAL ANALYSIS

1. COMPARISON OF GENETIC VALUES FOR PROPAGULE TYPES

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\,\mathbf{b} + \mathbf{Z}_2\,\mathbf{m} + \mathbf{Z}_3\,\mathbf{f} + \mathbf{Z}_4\,\mathbf{mf} + \mathbf{Z}_5\,\mathbf{c} + \mathbf{e}$$

where,

$\mathbf{y}$	vector of n measurements for a given trait,
$\boldsymbol{\mu}$	overall mean,
$\boldsymbol{\beta}$	vector of fixed effects (overall mean, replicate and propagule type),
$\mathbf{b}$	vector of incomplete blocks within replicate random effects, ~ N(0, $\sigma^2_b\,\mathbf{I}_b$),
$\mathbf{m}$	vector of mother random effects, ~ N(0, $\mathbf{I}_f \otimes \mathbf{G}_f$),
$\mathbf{f}$	vector of father random effects, ~ N(0, $\mathbf{I}_f \otimes \mathbf{G}_f$),
$\mathbf{mf}$	vector of family random effects, ~ N(0, $\mathbf{I}_{mf} \otimes \mathbf{D}_{mf}$),
$\mathbf{c}$	vector of clone within family random effects, ~ N(0, $\sigma^2_c\,\mathbf{I}_c$),
$\mathbf{e}$	vector of residual errors, ~ N(0, $\mathbf{R}$),

$\mathbf{G}_f$	matrix of size 2 x 2 of variance-covariance of additive genetic effects between propagule types (cutting and seedling).
$\mathbf{D}_{mf}$	heterogeneous diagonal matrix of size 2 x 2 of variances of family effects (one per propagule type).
$\mathbf{R}$	block diagonal matrix of size *n* x *n* composed of two errors (one per propagule type).

STATISTICAL ANALYSIS

G_f matrix structures

$$G_j = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix}$$

UN unstructured matrix, a matrix that considers three ***different*** (co)variance components (i.e. it allows for model heterogeneity of variances and covariance among effects).

$$\begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

CS compound symmetry matrix, matrix that is restricted to have the ***same variance*** for both propagule types but incorporates a non-zero covariance among them.

$$\begin{bmatrix} a & c \\ c & a \end{bmatrix}$$

DIAGH, heterogeneous diagonal matrix with ***different variances*** for each propagule type but ***independence*** among effects (i.e. it assumes zero-covariance among effects).

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

DIAG, homogeneous diagonal matrix with ***equal variances*** for both propagule types but ***independence*** among effects.

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$$

STATISTICAL ANALYSIS

1. COMPARISON OF GENETIC VALUES FOR PROPAGULE TYPES

- Three hypothesis of interest to compare **G** structures:

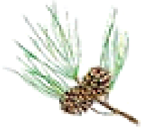
Test A - US vs CS	$\begin{bmatrix} a & c \\ c & b \end{bmatrix}$	vs	$\begin{bmatrix} a & c \\ c & a \end{bmatrix}$
Test B - US vs DIAGH	$\begin{bmatrix} a & c \\ c & b \end{bmatrix}$	vs	$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$
Test C - CS vs DIAG	$\begin{bmatrix} a & c \\ c & a \end{bmatrix}$	vs	$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$

- All analysis done with ASReml v. 3

- For **UN** and **CS** the parameter **ρ** can be obtained, and is a measure of similarity (ranking) in the estimation of additive effects between propagule types.

- Comparisons between models was done using the Akaike information criteria (AIC) and the likelihood ratio test (LRT).

STATISTICAL ANALYSIS



2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

- Traits considered: VOL, VCROWN, MDIAMB, MANG, CW, RLC, RVCR, and LC.
- All pairs of data were analyses separately.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\,\mathbf{b} + \mathbf{Z}_2\,\mathbf{m} + \mathbf{Z}_3\,\mathbf{f} + \mathbf{Z}_4\,\mathbf{mf} + \mathbf{Z}_5\,\mathbf{c} + \mathbf{e}$$

where,

$\mathbf{y}$	n x 2 matrix of measurements for a pair of traits,
$\boldsymbol{\mu}$	vector of overall trait means,
$\boldsymbol{\beta}$	matrix of fixed effects (overall mean, replicate and propagule type),
$\mathbf{b}$	matrix of incomplete blocks within replicate random effects, $\sim N(0, \sigma^2_b\, \mathbf{D}_b)$,
$\mathbf{m}$	matrix of mother random effects, $\sim N(0, \mathbf{I}_f \otimes \mathbf{G}_f)$,
$\mathbf{f}$	matrix of father random effects, $\sim N(0, \mathbf{I}_f \otimes \mathbf{G}_f)$,
$\mathbf{mf}$	matrix of family random effects, $\sim N(0, \mathbf{I}_{mf} \otimes \mathbf{G}_{mf})$,
$\mathbf{c}$	matrix of clone within family random effects, $\sim N(0, \sigma^2_c\, \mathbf{G}_c)$,
$\mathbf{e}$	matrix of residual errors, $\sim N(0, \mathbf{I}_t \otimes \mathbf{R})$,

$\mathbf{G}_x$	matrix of size 2 x 2 of variance-covariance of additive genetic effects between two traits.

STATISTICAL ANALYSIS

2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

- Each of the genetic effects (mother, father, family, etc.) is extended in a 8x1 vector of effects incorporating *correlations* among effects across all traits, e.g. i^{th}-mother will have an effect on each of the 8 traits as:

$$m_i = \begin{bmatrix} m_{i,1} \\ m_{i,2} \\ m_{i,3} \\ m_{i,4} \\ m_{i,5} \\ m_{i,6} \\ m_{i,7} \\ m_{i,8} \end{bmatrix} \qquad \mathrm{Var}(m_i) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} & \sigma_{17} & \sigma_{18} \\ & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} & \sigma_{27} & \sigma_{28} \\ & & \sigma_3^2 & \sigma_{34} & \sigma_{35} & \sigma_{36} & \sigma_{37} & \sigma_{38} \\ & & & \sigma_4^2 & \sigma_{45} & \sigma_{46} & \sigma_{47} & \sigma_{48} \\ & & & & \sigma_5^2 & \sigma_{56} & \sigma_{57} & \sigma_{58} \\ & & & & & \sigma_6^2 & \sigma_{67} & \sigma_{68} \\ & & & & & & \sigma_7^2 & \sigma_{78} \\ & & & & & & & \sigma_8^2 \end{bmatrix}$$

STATISTICAL ANALYSIS

2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

- Causal variance components were related to their corresponding genetic component.

Additive genetic variance	$V_a = 4\sigma_f^2$
Dominant genetic variance	$V_d = 4\sigma_{mf}^2$
Epistatic genetic variance	$V_i = \sigma_c^2 - 2\sigma_f^2 - 3\sigma_{mf}^2$
Total genetic variance	$V_g = V_a + V_d + V_i$
Total phenotypic variance	$V_{tot} = 2\sigma_f^2 + \sigma_{mf}^2 + \sigma_e^2$
Individual tree narrow-sense heritability	$h^2 = \frac{V_a}{V_{tot}} = \frac{4\sigma_f^2}{V_{tot}}$
Individual tree broad-sense heritability	$H^2 = \frac{V_g}{V_{tot}}$

- From the estimated covariance, Type-A correlations (trait-to-trait) can be obtained.

RESULTS

1. COMPARISON OF GENETIC VALUES FOR PROPAGULE TYPES

- According to Test A, most of the traits indicate that the additive variance is the same for both propagule types.
- From Tests B and C, we have indication that the covariance (or correlation) for additive effects among the propagule types is significantly different from zero.

Trial C	**VOL**	**VCROWN**	**MDIAMB**	**MANG**	**CW**	**RLC**	**RVCR**	**LC**
A - US vs CS	0.203	0.240	0.788	0.417	**0.001**	< **0.001**	0.337	0.238
B - US vs DIAGH	**0.003**	**0.003**	**0.004**	< **0.001**	0.517	< **0.001**	< **0.001**	**0.022**
C - CS vs DIAG	**0.001**	**0.002**	**0.004**	< **0.001**	< **0.001**	**0.005**	< **0.001**	**0.009**
Trial N	**VOL**	**VCROWN**	**MDIAMB**	**MANG**	**CW**	**RLC**	**RVCR**	**LC**
A - US vs CS	0.888	0.729	1.000	0.655	0.383	< **0.001**	0.144	0.752
B - US vs DIAGH	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**
C - CS vs DIAG	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**	< **0.001**

Trial C	**VOL**	**VCROWN**	**MDIAMB**	**MANG**	**CW**	**RLC**	**RVCR**	**LC**
Correlation (CS)	0.999 (NA)	0.965 (0.184)	0.999 (NA)	0.999 (NA)	0.999 (NA)	0.974 (0.214)	0.888 (0.133)	0.951 (0.207)
Trial N	**VOL**	**VCROWN**	**MDIAMB**	**MANG**	**CW**	**RLC**	**RVCR**	**LC**
Correlation (CS)	0.997 (0.037)	0.999 (NA)	0.999 (NA)	0.999 (NA)	0.999 (NA)	0.997 (0.062)	0.999 (NA)	0.999 (NA)

RESULTS

2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

ADDITIVE CORRELATIONS (A)

Upper diagonal – C Trial
Lower diagonal – N Trial

	VOL	VCROWN	MDIAMB	MANG	CW	RLC	RVCR	LC
VOL	-	0.648	0.205	-0.351	0.541	0.018	0.414	0.438
VCROWN	0.842	-	0.626	-0.565	0.930	0.644	-0.390	0.736
MDIAMB	0.501	0.076	-	0.192	0.138	-0.162	0.296	-0.224
MANG	-0.508	-0.332	-0.004	-	-0.460	0.231	0.092	-0.061
CW	0.777	0.949	0.716	-0.619	-	0.470	-0.420	0.483
RLC	0.227	0.515	0.169	0.014	0.297	-	-0.829	0.780
RVCR	0.075	-0.443	-0.313	0.166	-0.473	-0.605	-	-0.428
LC	0.669	0.711	0.143	-0.226	0.463	0.818	-0.296	-

- High ***correlation*** between VOL and VCROWN.
- Moderate correlations found among other relevant traits (e.g. VOL and MANG).
- Almost null correlation between MDIAMB and MANG.
- Interesting levels of heritability for MANG, VOL and VCROWN.

RESULTS

2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

GENETIC CORRELATIONS (G = A+D+I)

Upper diagonal – C Trial
Lower diagonal – N Trial

	VOL	VCROWN	MDIAMB	MANG	CW	RLC	RVCR	LC
VOL	-	0.813	0.469	-0.130	0.768	0.292	0.247	0.683
VCROWN	0.764	-	0.549	-0.404	0.963	0.590	-0.340	0.801
MDIAMB	0.370	0.565	-	0.198	0.574	0.396	-0.142	0.322
MANG	-0.330	-0.029	0.048	-	-0.104	0.335	-0.129	0.101
CW	0.716	0.936	0.603	-0.415	-	0.487	-0.354	0.660
RLC	0.273	0.603	0.348	-0.065	0.411	-	-0.667	0.766
RVCR	0.244	-0.394	-0.245	0.176	-0.396	-0.518	-	-0.307
LC	0.667	0.746	0.251	-0.220	0.526	0.807	-0.190	-

- Tendencies are similar to additive effects.

RESULTS

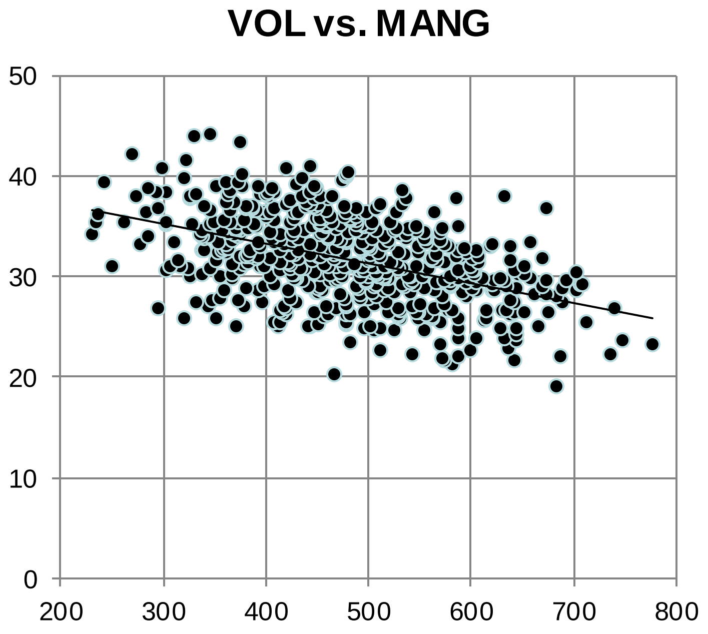
2. MULTIVARIATE ANALYSIS COMBINING PROPAGULE TYPES

FUTURE IMPLICATIONS OF MULTIVARIATE GENETIC ANALYSIS

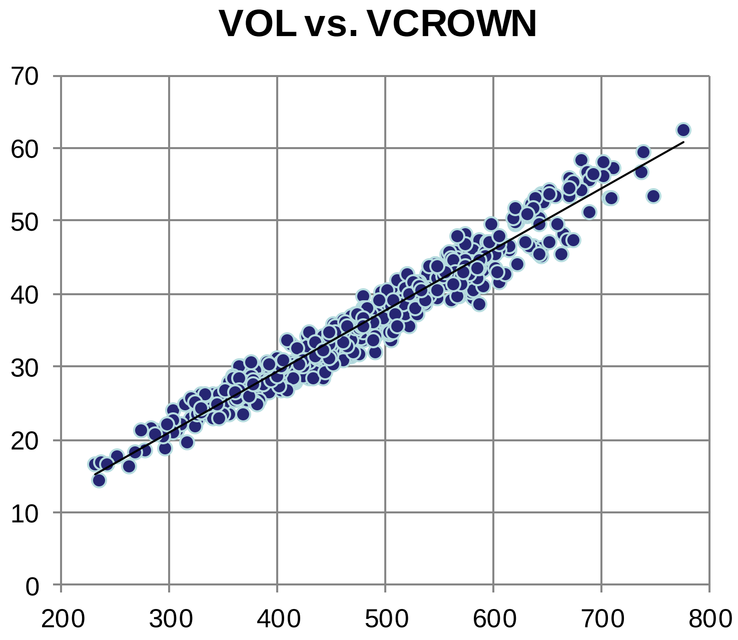
Indirect Selection

$$\Delta G_{A1} = i_2 \times h_1 \times h_2 \times rg_{A(a)} \times \sigma_{P1}$$

Search for correlation breakers



VOL vs. MANG



VOL vs. VCROWN

PART 2(B)

Understanding Rust data: relating greenhouse and trial data

OBJECTIVES

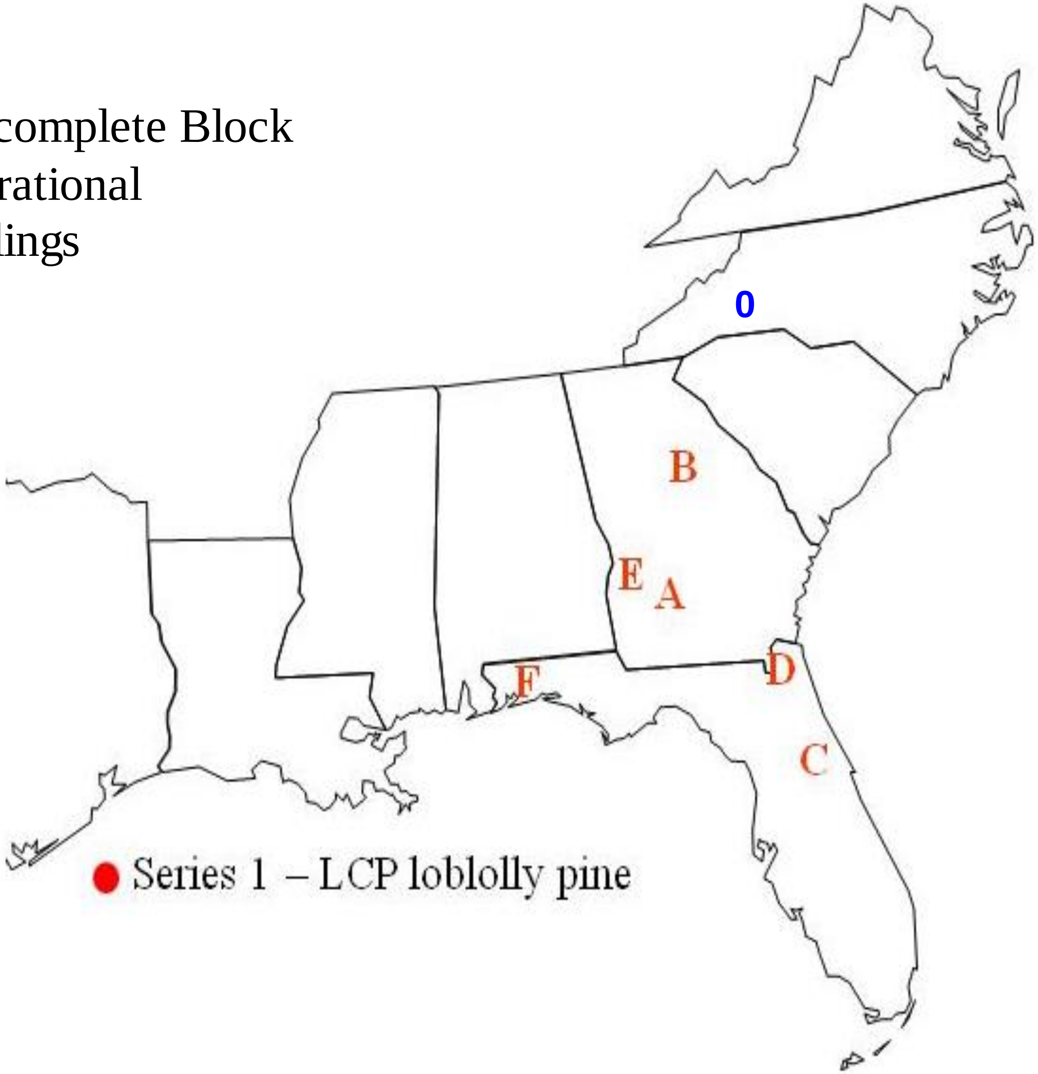
1) Evaluate the use of a different definition of rust incidence.
2) Determine the heritability of rust incidence and when is detectable.
3) Relate the genetic information obtained from a greenhouse experiment to a series of field trial experiments.

CCLONES Series 1 Data

FBRC Experiments

Design:	Resolvable Incomplete Block
Treatments:	Intensive/Operational
Propagules:	Cuttings/Seedlings

# sites:	7
# replicates	8
# families	~61
#parents	~32
#clones	~947
#clones/family	~17
#plants	47,040



CCLONES Series 1 Data

FIELD MEASUREMENTS

Lime Rust / Stem Rust / Stem and Limb Rust / Bush Rust / Nursery Rust

- Different types/levels of rust incidence measured at the individual level.
- Processed to correspond to a binary response 0/1 for each plant.

RESPONSE VARIABLE

RUSTINC: presence of rust at any time

Trial		Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9
BFG	B	x	x	x	x	-	x	-	x	-
CUT	E	x	x	x	x	-	x	-	-	-
NAS	D	x	x	x	x	-	x	-	-	x
OAK	A	x	x	x	x	-	-	-	-	-
PAL	C	x	x	x	x	-	x	-	-	x
REY	G	x	x	x	-	-	x	-	-	-
SAN	F	x	x	x	-	-	-	-	x	-

CCLONES Series 1 Data

Survival after Plantation

- Reference of survival at Year 1 inventory.

Trial	CUTTING	SEEDLING	All Plants
BFG	98.5%	98.1%	98.4%
CUT	98.2%	93.8%	97.3%
NAS	96.8%	97.5%	97.0%
OAK	92.6%	87.1%	91.6%
PAL	97.9%	98.5%	98.0%
REY	90.2%	58.0%	83.1%
SAN	97.9%	97.5%	97.8%
All Sites	96.8%	94.3%	96.3%

RUSTINC

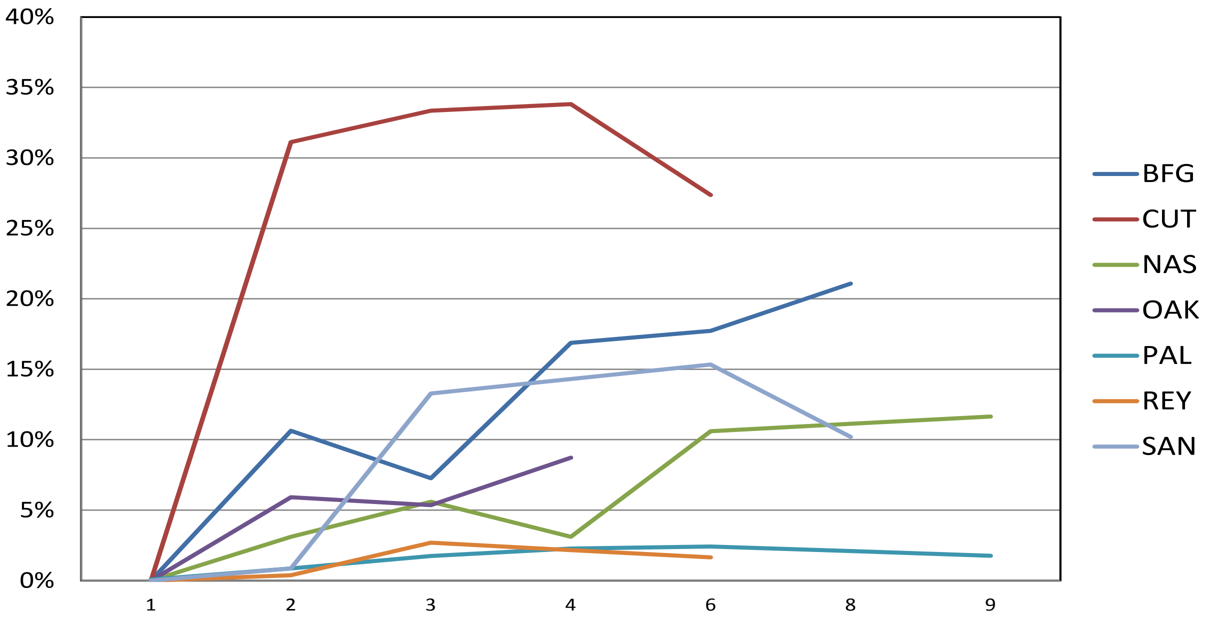
- Proportion from Year 1 that had rust at any of the measurements

Trial	CUTTING	SEEDLING	All Plants
BFG	24.3%	35.4%	**26.3%**
CUT	33.6%	58.4%	**38.6%**
NAS	13.4%	21.7%	**14.9%**
OAK	9.0%	21.4%	**11.2%**
PAL	2.8%	6.0%	**3.3%**
REY	3.8%	8.7%	**4.5%**
SAN	13.8%	23.5%	**16.1%**
All sites	13.1%	22.6%	14.9%

CCLONES Series 1 Data

Rust reported at every year (reference Year 1 inventory)

Trial	1	2	3	4	6	8	9	RUSTINC	Ratio
BFG	0.0%	10.6%	7.3%	16.9%	17.7%	21.1%		26.3%	55.9%
CUT	0.0%	31.1%	33.3%	33.8%	27.4%			38.6%	81.3%
NAS	0.0%	3.1%	5.6%	3.1%	10.6%		11.6%	14.9%	45.7%
OAK	0.0%	5.9%	5.4%	8.7%				11.2%	59.4%
PAL	0.1%	0.9%	1.7%	2.3%	2.4%		1.8%	3.3%	54.1%
REY	0.0%	0.4%	2.7%		1.6%			4.5%	34.5%
SAN	0.0%	0.9%	13.3%			10.2%		16.1%	50.4%



CCLONES Series 1 Data

Effects of Rust over Mortality

- Mortality of individuals at latest available measurement based on Year 1 inventory.
- Restricted to only Cuttings, no distinction on silvicultural treatment.

Trial	No-Rust	Rust	Odds	P-value
BFG	1.2%	10.9%	9.31	< 0.001
CUT	14.3%	18.2%	1.27	< 0.001
NAS	3.6%	7.6%	2.11	< 0.001
OAK	1.9%	3.7%	1.94	0.003
PAL	1.7%	15.7%	9.20	< 0.001
REY	2.3%	7.3%	3.17	0.044
SAN	4.5%	6.9%	1.52	0.007
Average	4.2%	10.0%	2.38	

CCLONES Series 1 Data

Comparison of genetic values for different "sites"

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\,\mathbf{b} + \mathbf{Z}_2\,\mathbf{m} + \mathbf{Z}_3\,\mathbf{f} + \mathbf{Z}_4\,\mathbf{mf} + \mathbf{Z}_5\,\mathbf{c} + \mathbf{e}$$

where,

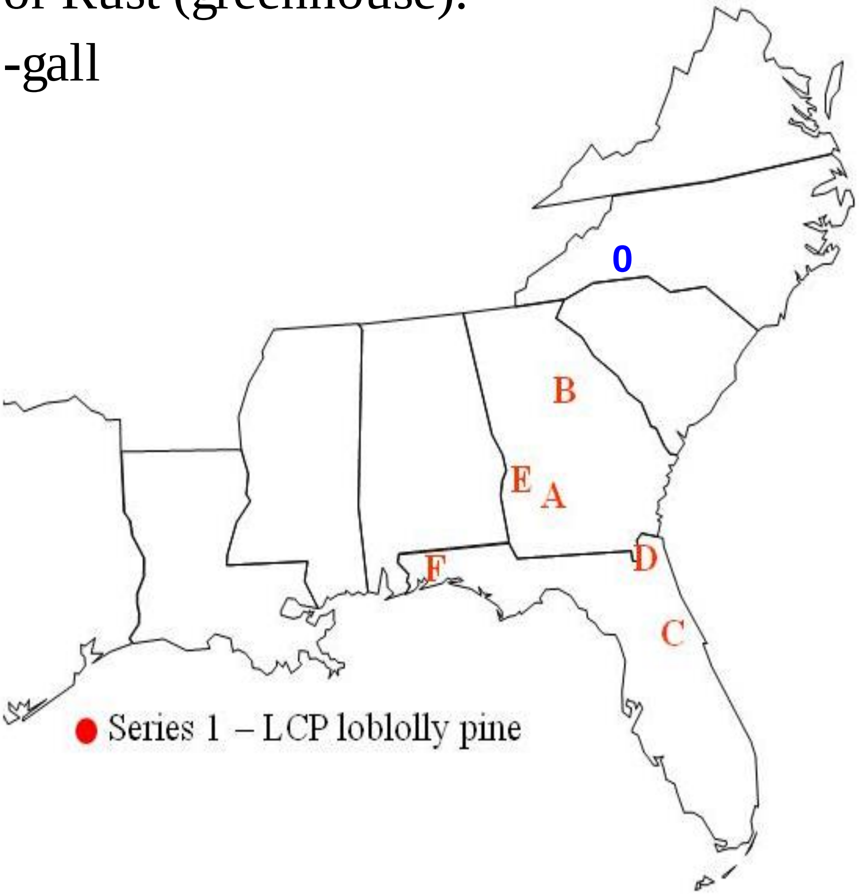
y	vector of n measurements for a given trait,
μ	overall mean,
β	vector of fixed effects (overall mean, replicate and propagule type),
b	vector of incomplete blocks within replicate random effects, ~ N(0, $\sigma^2_b\ \mathbf{I}_b$),
m	vector of mother random effects, ~ N(0, **A**),
f	vector of father random effects, ~ N(0, **A**),
mf	vector of family random effects, ~ N(0, $\sigma^2_{mf}\ \mathbf{I}_{mf}$),
c	vector of clone within family random effects, ~ N(0, $\sigma^2_c\ \mathbf{I}_c$),
e	vector of residual errors, ~ N(0, **R**),

CCLONES Series 1 Data vs Greenhouse

Narrow-sense heritabilities h^2

- Majority of the same families in both types of tests: field and greenhouse.
- Analysis based in RUSTINC (field) or Rust (greenhouse).
- Greenhouse experiments: 1-gall, 10-gall

Trial		h^2	SE	Mean RUSTINC
BFG	B	0.053	0.021	26.30%
CUT	E	0.199	0.068	38.60%
NAS	D	0.005	0.007	14.90%
OAK	A	0.082	0.034	11.20%
PAL	C	0.123	0.052	3.30%
REY	G	0.089	0.039	4.50%
SAN	F	0.000	0.000	16.10%
1-gall	0	0.205	0.064	31.14%
10-gall	0	0.246	0.086	36.33%



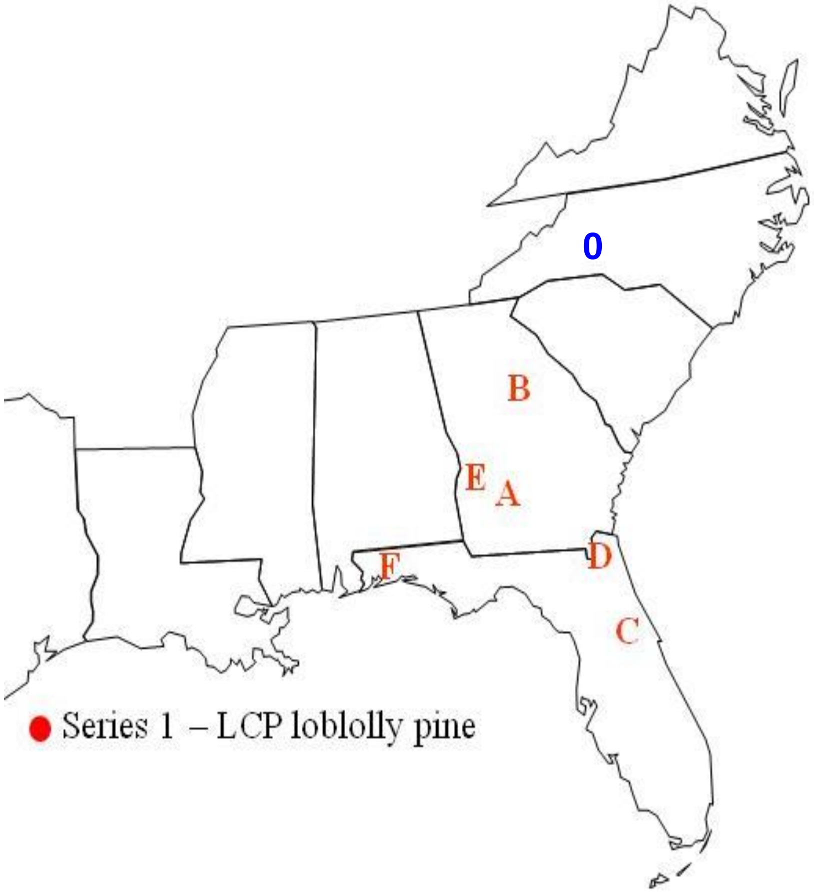
CCLONES Series 1 Data vs Greenhouse

Type B Correlations: Additive effects

- Majority of the same families in both types of tests: field and greenhouse.
- Analysis based in RUSTINC (filed) or Rust (greenhouse)
- Greenhouse experiments:

  1-gall, 10-gall

		1-gall		10-gall		Mean
Trial		**Correl**	**SE**	**Correl**	**SE**	**RUSTINC**
BFG	B	0.535	0.189	0.750	0.120	26.30%
CUT	E	0.774	0.136	0.774	0.136	38.60%
NAS	D	0.457	0.197	0.592	0.160	14.90%
OAK	A	0.650	0.104	0.607	0.112	11.20%
PAL	C	0.000	-	0.067	0.236	3.30%
REY	G	0.000	-	0.454	0.374	4.50%
SAN	F	0.505	0.208	0.549	0.186	16.10%



ACKNOWLEDGMENTS

PART 1

LIBERTAS CAPITUR
UNIVERSIDAD AUSTRAL DE CHILE

GOBIERNO DE CHILE
CONICYT

FONDEF D97I1054
“Software Nothofagus”

- Universidad Austral de Chile
- CONAF
- INFOR
- Forestal Mininco S.A.
- Forestal Cholguán S.A.
- Forestal Tornagaleones S.A.
- Forestal Valdivia S.A.

PART 2

FBRC – University of Florida
CFGRP – University of Florida

UF UNIVERSITY of FLORIDA

This image cannot currently be displayed.

- Dr. Gary Peter
- Dr. John Davis
- Dr. Timothy Martin
- Dr. Dudley Huber
- Dr. Tania Quesada