Hamrick, J. L.: Isozymes and the analysis of genetic structure in plant populations. *In:* Isozymes in plant biology. *Edited by* Soltis D. E. and P. S. Soltis. Dioscorides Press, Portland, Oregon. pp. 87–105 (1989). — Hicks, M., Adams, D., O'Keefee, S., MacDonald, E. and Hodgetts, R.: The development of RAPD and microsatellite markers in lodgepole pine (*Pinus contorta* var. *latifolia*). Genome **41**: 797–805 (1998). — Knowles, P.: Comparison of isozyme variation among natural stands and plantations: jack pine and black spruce. Can. J. For. Res. **15**, 902–908 (1985). — Lynch, M. and Milligan, B. G.: Analysis of population genetic structure with RAPD markers. Mol. Ecol. **3**: 91–99 (1994). — Maniatis, T., Fitsch, E. F. and Sambrook, J.: Molecular Cloning: A laboratory Manual. 2nd edition, Vol. I, II, III. Cold Spring Harbor Laboratory, Long Island, N.Y. (1989). — Moran, G. F., Bell, J. C. and Eldridge, K. G.: The genetic structure and the conservation of the five natural populations of *Pinus radiata.* Can. J. For. Res. **18**: 506–514 (1988). — Mosseler, A., Egger, K. N. and Hughes, G. A.: Low levels of genetic diversity in red pine confirmed by random amplified polymorphic DNA markers. Can. J. For. Res. **22**: 1332–1337 (1992). — Mosseler, A., Innes, D. J. and Roberts, B. A.: Lack of allozymic variation in disjunct Newfounland populations of red pine (*Pinus resinosa*). Can. J. For. Res. **21**: 525–528 (1991). — Müller-Starck, G.: Genetic differences between tolerant and sensitive beeches (*Fagus sylvatica* L.) *in* an environmentally stressed adult forest stand. Silvae Genet. **34**: 241–246 (1985). — Nkongolo, K. K.: RAPD and cytological analyses of *Picea spp.* from different provenances: genomic relationships among taxa. Hereditas **130**: 137–144 (1999). — Nkongolo, K. K., Klimaszewaska, K. and Gratton, W. S.: DNA yields and optimization of RAPD patterns using spruce embryogenic lines, seedlings and needles. Plant Mol. Biol. Rep. **16**: 285–384 (1998). — O'Malley, D. M., Allendorf, F. W. and Blake, G. M.: Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa*. Biochem. Genet. **17**: 233–250 (1979). — Sholz, F.: Genetic effects of air pollutants. Proc. 18th. IUFRO World Cong., Ljubljana, Vol. I, 286–294 (1986). — Simon, J.-P., Bergeron, Y. and Gagnon, D.: Isozyme uniformity in populations of red pine (*Pinus resinosa*) in the Abitibi region, Quebec. Can. J. For. Res. **16**: 1133–1135 (1986). — Strauss, S. H. and Doerksen, A. H.: Restriction fragment analysis of pine phylogeny. Evolution **44**: 1081–1096 (1990). — Strauss, S. H., Hong, Y. P. and Hipkins, V. D.: High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata, P. miricata,* and *P. attenuata.* Theor. Appl. Gen.: **86**, 605–611 (1993). — Weeden, N. F. and Wendel, J. F.: Genetics of plant isozymes. *In:* Isozyme in plant biology. *Edited by* D. E. Soltis and Soltis, P. S. Dioscorides Press, Portland, Oregon. pp. 46–72 (1989). — Wu, J., Krutovskii, K. and Strauss, S. H.: Nuclear DNA diversity, population differentiation, and phylogenetic relationships in the California closed-cone pines based on RAPD and allozyme markers. Genome **42**: 893–908 (1998). — Yamaguchi, D. K.: A simple method for cross-dating increment cores from living trees. Can. J. For. Res. **21**: 414–416 (1991). — Yeatman, C. W.: Biogeography of Jack pine. Can. J. Bot. **67**: 2201–2211 (1967). — Yeatman, C. W.: P. N. F. I. Genetics and breeding: Genetics of jack pine: 1982–1983. Proc. 19th C. T. I. A. Mtg., Toronto, Ont., Aug. 22 to 26, 1983. Pp. 138–140 (1984). — Yeh , F. C. and Layton, C.: The organization of genetic variability in central and marginal populations of lodgepole pine *Pinus contorta* var *latifolia*. Can. J. Genet. Cytol. **21**: 2201–2211 (1979).

A Case Study of a Provenance/Progeny Test Using Trend Analysis with Correlated Errors and SAS PROC MIXED

By C. Saenz-Romero1)3), E. V. Nordheim2), R. P. Guries and P. M. Crump4)

Department of Forest Ecology and Management, University of Wisconsin – Madison, 120 Russell labs, 1630 Linden Drive,
Madison, WI 53706, U.S.A.

(Received 30th October 2000)

Abstract

A statistical analysis was performed on data from a provenance test of *Pinus banksiana* Lamb. 1.5 year-old open-pollinated seedlings from 47 stands collected in five distinct soil-moisture classes within a pine-barren landscape, in west-central Wisconsin. A large-scale environmental gradient across the nursery bed and small-scale among-microsite variability were evident in seedling responses. We compared three analyses in terms of their capabilities for accounting for such within-experimental area variation: (A) a "standard analysis" using ANOVA for a randomized complete block design, (B) trend analysis in addition to (A), and (C) correlated errors in addition to (B). PROC GLM of the Statistical Analysis System (SAS) was used for analyses (A) and (B); PROC MIXED was necessary for analysis (C). We concluded that analysis (C) was the best option for adequately modeling the data, reducing the error variance and consequently, detecting significant differences among sets of stands grouped by soil-moisture classes. We suggest that the row and column position of each experimental unit in provenance or progeny tests with forest species be recorded in order to allow for the possibility of conducting analysis of this type. As an important caveat, we found an unexpected bimodal likelihood surface when PROC MIXED included a term for correlation among plots; this requires use of the PARMS statement when applying PROC MIXED.

Key words: Trend analysis, correlated errors, spatial autocorrelation, PROC MIXED, provenance test, progeny test, *Pinus banksiana*, genetic variation, within block variability.

Introduction

A typical experimental design for provenance and progeny testing with forest trees is the randomized complete block design (RCBD) (Wright, 1976; Zobel and Talbert, 1984; Gumpertz and Brownie, 1993), with each entry (entries can be provenances, progenies, clones or ramets) appearing in one plot per block and multiple trees (seedlings) per entry per plot (usually 4 to 10 trees per plot) (Bridgwater *et al.*, 1983). This model can be written:

$$[1] \qquad Y_{ijk} = \mu + \beta_i + \tau_j + \beta_i \times \tau_j + \varepsilon_{ijk}$$

where Y_{ijk} = observation on the k^{th} seedling of the j^{th} entry in the i^{th} block, μ = overall mean, β_i = effect of the i^{th} block, τ_j = effect of j^{th} genetic entry, β_i x τ_j = interaction of block by entry (also known as plot error), and ε_{ijk} = within plot error term, i=1,...b, j=1,...t, and k=1,...n_{ij}, where b, t and n_{ij} are the

---

1) Present address: Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolás de Hidalgo, Apartado Postal 2-105, 58041 Morelia Michoacán, México.

2) Also Department of Statistics, University of Wisconsin – Madison.

3) Corresponding author: Phone/fax +(52)(4)327-2350 and 51, csaenz@zeus.ccu.umich.mx

4) Department of Computing and Biometry, College of Agriculture and Life Sciences, University of Wisconsin – Madison, 1675 Observatory Drive, Madison, WI 53706, U.S.A.

number of blocks, entries and seedlings-per-plot for the j^{th} entry in the i^{th} block, respectively.

Typically, blocks are spatial entities; the goal of blocking is for the environment to be homogeneous within blocks and for most of the environmental variation to occur among blocks. Blocking can remove environmental variation by employing "block" as a source of variation in an analysis of variance.

However, the assumption of homogeneity within blocks is often difficult to meet with field experiments of forest trees. Sometimes, the experimental area contains a large and systematic environmental gradient that can cause heterogeneity within (as well as among) blocks. There may also exist scattered microsites, in which some plots may exhibit an exceptional growth (*e.g.* nutrient-rich microsites), and others may show poor growth (*e.g.* a frost pocket or a poorly drained spot). Frequently, poor microsites result in greater mortality, creating additional design imbalance.

One factor that makes it difficult for blocks to be homogeneous in experiments with trees is the large block size necessary due to the large number of entries. The number of entries and the minimum space required for a tree to grow without competition for several years can make block size so large that it is impossible for the within-block variation to be small, particularly if the site has irregular topography. If more than 50 entries per block are tested, within-block variation may become unacceptably large (BRIDGWATER *et al.*, 1983).

Consequently, a RCBD analysis can result in a significant block x entry interaction. This b_i x t_j interaction term is critical for the interpretation of the results because if block and/or entry are considered as random effects (common in forest genetic experiments), the block x entry would be the appropriate error term to test for entry effects. If this term is large, it reduces the chances that the entry effect will be found significant.

An alternative to large non-homogeneous blocks is to use fixed blocks of smaller sizes in an incomplete block design. If large and complete blocks were used in a field test already, eventually smaller incomplete blocks could be redefined for analyses purposes (although such alternative was not developed in the present paper).

Even with a good plot design, a field test may result in positive correlations among neighboring plots (spatial autocorrelation), because plants that are close together share similar microenvironmental conditions. This autocorrelation results in the violation of the assumption of independent error terms for standard analysis of variance; also, in this case within-block variability can not be neutralized by randomization (ZIMMERMAN and HARVILLE, 1991; MAGNUSSEN, 1994). Negative correlations are also possible due to competition among neighboring trees.

We present a case study of a population/progeny test of *Pinus banksiana* LAMB., which manifested a large environmental gradient leading to a large within block variation. Our data also demonstrated autocorrelation among plots (probably due to microsite variation) and moderate mortality. It would not be unfair to call our case study a "messy data" scenario, but one that is fairly common in practice.

We demonstrated the inadequacy of a "standard" RCBD analysis using the popular General Linear Model procedure (PROC GLM) of the Statistical Analysis System (SAS) (SAS 1988), even including a trend analysis. However, the (relatively new) "MIXED" procedure of SAS can incorporate models with correlated errors and has the flexibility to consider various design complexities, including unbalanced and/or missing data, and give the correct result. We believe that the proper use of PROC MIXED addresses simultaneously the problem of large scale variation (gradient) and small scale variation (microsites) within the experimental area.

Methodology

Objectives and study area

The original objective of our research was to address the question of the patterning of genetic variation among *Pinus banksiana* populations within a landscape, in the absence of strong climatic and altitudinal clines and with the likely presence of an intensive gene flow among populations. Our hypothesis was that, even small environmental forces, such as differences in soil and moisture availability, might impose a sufficiently large selective force to create genetic differences among stands within a landscape. We expected that if such differentiation were present, we could find significant differences in early growth traits of seedlings in a progeny test. Our sampled area was an approximately 30 km x 30 km pine barren landscape in the Black River State Forest and Jackson County Forest in west-central Wisconsin. We collected open-pollinated seeds from 86 individual trees from 47 stands (also called populations). Each stand sampled was assigned to one of five categories of a site index that we constructed and called 'soil-moisture classes'. The index includes characteristics of soil texture and soil drainage based on soil surveys [Wisconsin Department of Natural Resources (DNR), Nontechnical Soil Description Report; all Jackson County Soils. DNR, Madison, Wisconsin. 19 p.], and distance to the water table and presence of 'indicator species', based on 'habitat types' (KOTAR and BURGER, 1996). Soil moisture classes take five values, from I (sandy soil, very well drained, dry habitat) to V (sandy-loamy-clayey, very poorly drained, dry-mesic habitat) (*Table 1*). Initially, we grew seedlings from these populations in a randomized complete block design population/progeny test for six months in a greenhouse at the University of Wisconsin – Madison.

Randomized complete block design (RCBD) and field layout

After the greenhouse stage, seedlings were planted in a common garden population/progeny test in a nursery bed at the Wilson State Nursery (Boscobel, Wisconsin) where they grew for one year. This paper considers the analysis of those 1.5-year-old seedlings. The common garden test followed a randomized complete block design, with 4 blocks and 3-row-seedling plots. The 47 stands were unevenly represented by 86 half-sib families (trees): 13 stands were represented by 3 families, 13 stands by 2 families and 21 stands by a single family. Each stand belonged to a unique soil-moisture class. In our analyses, stands were considered to be nested within the five categories of soil-moisture classes.

In each block of the common garden test, the plots were laid out in a 5 x 18 grid. The seedlings were evenly spaced at 19 cm intervals; with all 4 contiguous blocks, the layout became a 5 x 72 grid. We will refer to the five rows of seedlings along the long dimension of the nursery bed as "columns" and the 72 rows perpendicular to the long dimension of the nursery bed as "rows" (*Figure 1*). Four randomly assigned plots per block were filled with randomly chosen seedlings from the same test material but for which data were not recorded and which we call 'dummy' seedlings; this was done to make any competition effects as uniform as possible.

After one year in the nursery, we determined seedling height and timing of budset. Then, we harvested all the seedlings and obtained dry weights for stem, foliage, root, shoot: root ratio and foliage: root ratio. We consider stem dry weight and second-

Table 1. – Soil-moisture classes (I to V) based on soil and habitat characteristics for 47 *Pinus banksiana* stands in west-central Wisconsin.

Soil type			Habitat type	Soil-moisture class
Soil classes	Texture	Drainage		
TrB Tarr Sand, TrC Tarr sand, MaB Mahtomedi sand	Sandy	Extremely well	Dry*	I
TtA Tint sand, RkA Rockdam sand	Sandy	Moderate to poor	Dry	II
RkA Rockdam sand, IxA Ironrun Ponycreek complex, Pv Ponycreek-Dawsil complex	Sandy	Moderate to very poor	Dry-Mesic†	III
LxB Ludington-Fairchild sands, MoB.Merit-Gardenvale silt loams, FeA Fairchild-Elm Lake	Sandy-loamy	Moderate to very poor	Dry	IV
FeA. Fairchild-Elm Lake complex, FaA Fairchild sand, MrA Merrillan-Veedum complex	sandy-loamy to Sandy-loamy-clayey	Very poor	Dry-Mesic	V

* Habitat PVHa *Pinus-Vaccinium-Hamamelis.*

† Habitat PVRh *Pinus strobus-Vaccinium-Rubus hispidus* with ground water influence.

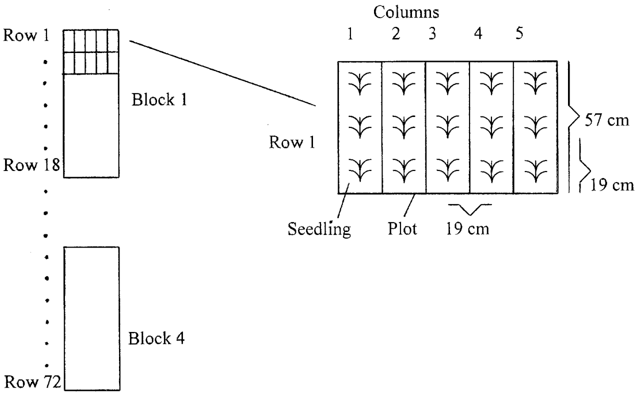

Figure 1. – Plot layout for the *Pinus banksiana* provenance/progeny common garden test.

arily foliage dry weight as response variables to illustrate our analyses in this paper; the results from analyses of other variables will be discussed elsewhere, in the context of drawing genetic conclusions from this study. We estimated the plot average (over seedlings) using SAS PROC MEANS (SAS, 1988). We found it necessary to transform the plot averages for stem dry weight and foliage dry weight by the square root to lead to a satisfactory distribution of residuals. These transformed plot averages constitute our basic data set.

We conducted a standard RCBD analysis of variance using an extension of model [1], but we included soil-moisture classes with stands nested within these soil-moisture classes. We considered but then eliminated terms for "family" because we found that variation among families was not significant in any of the tested models throughout all of our analyses. This non-significance may be due in part to a) the very small number of families per stand, b) an insufficient seedling age for fully expression of the among family genetic differences or c) "noise" due to transplanting effects. Thus, our resulting model was:

[2]

$$Y_{ijkl} = \mu + \beta_i + \alpha_j + \tau(\alpha)_{jk} + \beta_i \times \alpha_j + \beta_i \times \tau(\alpha)_{jk} + \varepsilon_{ijkl}$$

Model (A)

where: Y_{ijkl} = plot average for the l^{th} family in the k^{th} stand nested within the j^{th} soil-moisture class in the i^{th} block, b_i = effect of the i^{th} block, a_j = effect of the j^{th} soil-moisture class, $t(a)_{jk}$ = effect of the k^{th} stand nested within the j^{th} soil-moisture class, $b_i \times a_j$ = effect of the interaction of block by soil-moisture class, $b_i \times t(a)_{jk}$ = effect of the interaction of block by stand (again nested); $i = 1, 2, \ldots 4$; $j = 1, 2, \ldots 5$; $k = 1, 2, \ldots 47$; $l = 1, 2$ or 3, and e_{ijkl} = error.

We considered blocks as random because we viewed them as a sample of a larger population of possible experimental site conditions. We considered stands as random because they are a sample of a larger population of possible stands (*Pinus banksiana* stands within the studied area) for which we wanted to make inferences. Analysis used SAS PROC GLM (SAS, 1988).

Trend analysis

General trends due to environmental gradients within the experimental area can be removed via "trend analysis", a regression technique that uses a predicted "yield value" as a covariate (Kirk *et al.*, 1980; Brownie *et al.*, 1993). The yield value (the expected growth of a seedling due to its position) can be estimated by fitting a quadratic response surface or other polynomial model, using an imposed grid of Cartesian coordinates on the experimental area. Terms of the quadratic-surface model are used for describing plot position effects, making the block term unnecessary (Brownie *et al.*, 1993).

We conducted a trend analysis by augmenting model [2] (without the block terms) with the terms of a quadratic response surface model plus the corresponding interactions, using a grid of 5 units – columns- along the *x* axis and 72 units – rows- along the *y* axis. Then the model becomes:

[3]

$$Y_{jklmp} = \mu + \alpha_j + \tau(\alpha)_{jk} + \beta_1 C_m + \beta_2 R_p + \beta_3 C^2{}_m + \beta_4 CR_{mp} + \beta_5 R^2{}_p + \alpha_j \times C_m + \ldots$$

$$\ldots + \alpha_j \times R_p + \alpha_j \times C^2{}_m + \alpha_j \times CR_{mp} + \alpha_j \times R^2{}_p + \tau(\alpha)_{jk} \times C_m + \tau(\alpha)_{jk} \times R_p + \ldots$$

$$\ldots + \tau(\alpha)_{jk} \times C^2{}_m + \tau(\alpha)_{jk} \times CR_{mp} + \tau(\alpha)_{jk} \times R^2{}_p + \varepsilon_{jklmp}$$

where *m* and *p* index column and row respectively, b_1 to b_5 are regression parameters to be estimated, C_m = column position, R_p = row position, $a_j \times C_m \ldots t(a)_{jk} \times R^2{}_p$ = corresponding interactions of soil-moisture classes and stand with the quadratic response-surface terms, $m = 1...5$, $p = 1...72$, and the other terms are as before. Again, the analysis used PROC GLM. We consistently found that in evaluating the significance of the interaction terms in [3] across numerous models, the only relevant interaction was that between stand (nested in soil-moisture classes) and column. However, the quadratic pattern for column did not fit the data well. Extensive analyses showed that there was a great deal of variability across columns that could not easily be described by a polynomial model. Thus we modified the trend analysis by considering column as a random effect (row remained as fixed effect). Since we consistently found non-significant interaction between rows and columns, we removed the cross-product term. The resulting model becomes:

[4]

$$Y_{jklmp} = \mu + \alpha_j + \tau(\alpha)_{jk} + C_m + \beta_1 R_p + \beta_2 R^2{}_p + \tau(\alpha)_{jk} \text{ x } C_m + \varepsilon_{jklmp}$$

Model (B)

In this model, stand [$t(a)_{jk}$], column (C_m), and stand by column [$t(a)_{jk}$ x C_l] interaction were considered random effects, and soil-moisture class (a_j) and the covariates (R_p and $R^2{}_p$) are considered fixed effects. Note that column no longer had a corresponding regression term since it was now a random "class" variable. It is convenient to center and scale the values of row so that their means are 0 and their maximum and minimum values are 1 and –1, respectively.

Analysis incorporating spatially correlated errors using PROC MIXED

Although model (B) successfully reduced the residual error by modeling environmental trends within the experimental site (using information on columns and rows), we also noted some microsites with particularly good growth and other microsites with poor growth, likely due to uneven distribution of nutrients and drainage. An example of a microsite with good growth is the area contained in rows 48 to 51 in column 3, block 3; an example of a microsite with poor growth is the region described by rows 68 to 71 in column 3, block 4 (*Table 2*). Even after removing the large-scale systematic variability by our modified trend analyses, it was possible that small-scale variation among-microsites could have a statistically significant impact. The variation among microsites at different distances could be approached as a problem of spatial autocorrelation (for example, see the paper of Sokal and Oden, 1978).

One approach to account for the variation among microsites was the use of correlated error models. Such models attempt to account for small-scale variation in soil properties through correlations among the plot errors. For our data, we invoked models which assume that the strength of the correlation between two errors, Corr(e_{mp}, e_{qr}), is greatest for adjacent plots but decreases exponentially as the distance between the plots increases (Brownie *et al.*, 1993):

[5]

$$\text{Cov}(\varepsilon_{mp}, \varepsilon_{qr}) = \sigma^2 \text{ Corr}(\varepsilon_{mp}, \varepsilon_{qr}) = \sigma^2 e^{-\theta d(mp,qr)}$$

where $d(mp, qr)$ = the center-to-center distance between the plots in the m^{th} column and p^{th} row with the plot in the q^{th} column and r^{th} row, and q is a parameter to be estimated. The among-plot correlation [Corr (e_{mp}, e_{qr})], which we call $r_{mp,\, qr,}$ was included in our model as a way to account for the effect of small-scale variation among-plots. Thus, adding this term to those for model (B) allowed for the simultaneous treatment for both large-scale and small-scale environmental variation. This combination of trend analysis together with correlated errors is known as a "random field linear model" because its theory is based on certain stochastic processes known as "random fields" (Zimmerman and Harville, 1991). The resulting model is:

[6]

$$Y_{jklmpqr} = \mu + \alpha_j + \tau(\alpha)_{jk} + C_m + \beta_1 R_p + \beta_2 R^2{}_p + \tau(\alpha)_{jk} \text{ x } C_m + \rho_{mp,qr} + \varepsilon_{jklmpqr}$$

Model (C)

Stand ($t(a)_{jk}$), column (C_m) and stand by column interaction ($t(a)_{jk}$ x C_m) were considered random effects, whereas a_j soil-moisture class, R_p row position, and $R^2{}_p$ squared row were considered fixed effects. The among-plot correlation ($r_{mp,\, qr}$) accounted for the spatial autocorrelation.

Model [6] could be analyzed with the MIXED Procedure of SAS, designed for statistical models with both random and fixed effects with a range of error structures (Littell et al., 1996, SAS, 1996a and b). PROC MIXED uses restricted maximum likelihood (REML), also known as residual maximum likelihood (Searle *et al.*, 1992), for estimation of variance components instead of the usual Least Squares used by PROC GLM (Other options are possible but we feel REML is the most appropriate).

In order to determine if a specific variance component is significant, the model must be fitted twice, once including and once excluding the component; then the change in the likelihood must be evaluated5). For consistent terminology, we will call the model with the component in question the "full model" and the model without this component the "reduced model". For large samples, the difference between minus twice the logarithm of the likelihood (-2RLL) for the reduced and full models results in an approximate chi-square test with one degree of freedom for each parameter tested. This test is called the likelihood ratio statistic (Littell *et al.*, 1996). Thus,

$$[7] \qquad X^2 = (-2RLL_{\text{Reduced Model}}) - (-2RLL_{\text{Full Model}})$$

and the partial p-value for this test is

$$[8] \qquad \text{Prob}\,(\chi^2{}_{d.f.=i} \geq X^2)$$

where i = number of parameters tested.

In *table 3* we provide the SAS program for model [6] – also referred to as (C) – along with brief comments. We also include as an example of a reduced model that omitting stand nested within soil-moisture classes from model [6]; the other reduced models are formulated similarly. For information about the programs corresponding to models (A) and (B), contact the senior author.

Results and Discussion

RCBD analysis

The plot averages of stem dry weight (without transformation to square root) are given in *table 2.* The design is highly unbalanced due to the different numbers of families per stand as well as the (moderate) mortality during winter; this resulted in a total of 61 missing plots from a total of 344 plots. (Implicit

---

5) There is the option COVTEST on the PROC MIXED statement, which provides a z-statistic and the corresponding p-value for testing H_o: $a^2{}_x = 0$. However, this test, is based on asymptotic properties, is not very reliable if the degrees of freedom to estimate the variance component is small (Littell *et al.*, 1996). The test described here is more reliable.

Table 2. – Plot average of stem dry weight (STE) in mg for 1.5-year-old *Pinus banksiana* seedlings. B = block, R = row, P = provenance, F = family. Missing plots are indicated with a period. Provenances and families labeled '0' are dummy seedlings.

B	R	Column 1			Column 2			Column 3			Column 4			Column5		
		P	F	STE	P	F	STE	P	F	STE	P	F	STE	P	F	STE
1	1	31	3	1.92	73	2	0.72	37	1	.	51	3	.	15	9	2.51
1	2	27	2	.	22	2	2.14	3	5	.	64	3	0.47	36	5	2.47
1	3	42	8	1.95	59	9	1.05	58	1	2.6	67	1	3.22	76	10	8.03
1	4	59	1	1.42	11	3	1.25	55	2	.	34	4	10.35	0	0	.
1	5	0	0	.	26	3	.	60	3	0.37	55	1	1.78	25	7	7.34
1	6	23	7	2.56	56	7	3.37	36	1	.	53	1	4.58	57	4	5.36
1	7	67	4	.	27	1	0.08	53	2	4.27	85	5	14.1	75	3	3.53
1	8	11	7	3.07	45	3	1.72	67	7	2.52	54	3	10.63	33	1	3.14
1	9	63	4	3.13	31	8	2.66	22	6	4.55	77	2	11.4	37	8	4.75
1	10	27	6	.	75	1	.	32	10	2.51	35	6	4.25	54	9	10.5
1	11	32	1	.	31	9	.	34	3	2.31	28	7	.	3	10	.
1	12	0	0	.	63	10	.	49	8	.	47	7	2.13	51	2	2.7
1	13	86	2	.	56	11	2.2	46	9	0.46	11	2	.	10	10	5.14
1	14	32	2	3.37	25	9	0.89	19	13	7.59	22	5	6.45	75	7	3.79
1	15	45	6	1.39	49	14	2.27	76	9	1.59	37	6	4.43	64	10	6.43
1	16	3	7	1.65	77	6	1.99	63	5	2.5	16	5	5.98	42	7	.
1	17	38	3	.	36	4	2.95	30	7	7.95	64	8	3.87	51	11	4.31
1	18	60	6	0.19	72	6	0.49	18	7	.	5	2	.	0	0	.
2	19	57	4	0.64	54	9	0.36	64	10	0.73	45	3	2.23	11	2	.
2	20	75	1	2.15	37	1	1.42	38	3	3.48	59	1	7.66	22	5	3.52
2	21	49	8	3.44	10	10	1.9	51	11	7.56	46	9	9.61	56	7	15.6
2	22	77	6	1.17	55	2	3.49	11	7	3.3	33	1	6.95	23	7	8.83
2	23	11	3	2.66	19	13	.	36	5	7.74	31	3	.	42	8	3.12
2	24	25	9	0.46	31	9	0.82	36	4	.	53	2	.	67	4	.
2	25	27	1	0.94	22	6	2.29	26	3	1.73	51	3	.	63	5	.
2	26	56	11	1.37	42	7	.	86	2	.	63	10	12.7	37	6	6.98
2	27	28	7	2.91	22	2	1.3	5	2	.	73	2	.	32	1	5.11
2	28	31	8	2.51	0	0	.	64	8	7.39	27	6	.	59	9	.
2	29	60	6	1.54	76	9	.	3	5	.	64	3	.	30	7	3.22
2	30	60	3	.	72	6	7.93	27	2	9.09	49	14	11.3	77	2	.
2	31	32	10	3.17	55	1	2.4	0	0	.	76	10	10.25	63	4	9.89
2	32	32	2	0.95	37	8	0.75	54	3	5.63	3	10	6.08	16	5	7.83
2	33	34	3	1.94	35	6	4.48	51	2	6.99	0	0	.	58	1	6.68
2	34	25	7	1.8	45	6	0.84	3	7	3.44	15	9	8.26	34	4	5.62
2	35	67	7	1.59	36	1	2.23	53	1	4.18	18	7	7.58	75	7	8.89
2	36	67	1	0.86	85	5	3.06	75	3	5.22	47	7	7.22	0	0	.
3	37	59	9	.	56	7	0.99	64	8	1.48	63	5	15.5	35	6	14.6
3	38	53	2	0.73	15	9	0.51	56	11	2.18	34	4	2.3	34	3	8.66
3	39	51	11	0.56	26	3	0.88	75	7	0.68	63	4	4.18	37	8	3.48
3	40	73	2	0.5	16	5	1.48	27	1	.	22	6	5.5	47	7	6.05
3	41	75	1	.	28	7	0.67	42	7	.	76	9	.	25	7	7.38
3	42	25	9	.	77	6	.	19	13	0.71	3	5	4.42	49	14	8.03
3	43	31	8	0.38	67	4	0.18	59	1	0.14	3	10	1.16	45	6	5.56
3	44	57	4	3.87	45	3	1.3	67	1	1.62	22	2	1.08	86	2	.
3	45	36	5	0.89	31	3	0.25	64	10	0.94	0	0	.	36	4	2.47
3	46	0	0	.	67	7	.	11	3	1.01	22	5	0.81	77	2	8.01
3	47	54	3	1.18	53	1	0.89	42	8	0.64	72	6	0.7	5	2	6.72
3	48	32	10	1.04	75	3	2.68	55	1	6.59	58	1	10.87	37	1	.
3	49	32	2	2.15	0	0	.	54	9	8.41	10	10	5.18	55	2	10.11
3	50	37	6	1.82	38	3	2.11	51	3	6.96	46	9	8.15	32	1	6.54
3	51	51	2	0.45	36	1	.	76	10	5.67	85	5	10.11	27	2	6.55
3	52	33	1	0.44	11	2	1.83	3	7	1.33	49	8	2.19	60	6	8.33
3	53	18	7	2.69	63	10	0.85	23	7	1.55	64	3	1.48	11	7	2.9
3	54	27	6	0.87	30	7	1.21	60	3	1.57	31	9	2.05	0	0	.
4	55	53	2	1.36	34	3	1.08	63	5	1.49	19	13	1.68	22	5	4.28
4	56	75	3	2.83	22	6	4.04	16	5	2.27	45	6	.	31	9	.
4	57	64	3	1.26	51	2	2.33	27	1	0.63	86	2	1.1	36	1	3.82
4	58	85	5	1.42	26	3	1.57	56	7	1.69	11	7	0.84	54	9	6.83
4	59	54	3	2.71	60	3	2.85	47	7	2.5	64	10	3.52	57	4	9.46
4	60	49	14	2.84	5	2	1.7	59	9	0.74	28	7	2.55	60	6	18.35
4	61	38	3	0.7	23	7	.	31	8	.	33	1	0.26	73	2	.
4	62	30	7	.	37	6	0.94	45	3	.	22	2	1.89	18	7	3.55
4	63	72	6	1.74	37	8	1.15	32	1	2.2	55	1	6.16	3	5	8.63
4	64	51	11	1.78	53	1	.	0	0	.	67	1	2.39	49	8	8.65
4	65	37	1	.	75	1	0.71	46	9	0.35	51	3	1.28	32	2	7.11
4	66	3	7	1.96	64	8	0.68	27	6	2.53	36	5	2.42	77	6	10.3
4	67	27	2	1.27	67	7	1.15	36	4	4.21	25	7	2.94	32	10	9.85
4	68	11	3	3.47	63	10	0.46	42	8	0.37	0	0	.	31	3	3.98
4	69	42	7	4.24	25	9	2.02	58	1	0.55	56	11	2.21	15	9	2.09
4	70	76	10	0.2	10	10	0.12	3	10	0.39	67	4	0.85	34	4	2.64
4	71	11	2	0.65	0	0	.	76	9	0.79	59	1	2.71	77	2	4.27
	72	35	6	0.56	63	4	0.61	75	7	2.03	55	2	1.62	0	0	.

in our model is the assumption – which we checked – that the mortality is random). An analysis of variance using PROC GLM, using the RCBD model (A) (equation [2]), suggested that the overall differences among soil-moisture classes were not significant, and that there were no significant differences among stands within soil-moisture classes (*Table 4*). The error

Table 3. – SAS program for trend analysis with correlated errors using PROC MIXED (model C).

SAS PROGRAM	*COMMENTS*
	FULL MODEL
DATA ONE; INFILE 'stemfol.dat' FIRSTOBS=2;	*STAND=Provenance. See explanation of COLUMN and ROW in the*
INPUT BLOCK ROW COLUMN SOIMOIST	*text. SOIMOIST = Soil-moisture class. STEM and FOLIAGE= Plot*
STAND FAMILY STEM FOLIAGE ;	*average of stem and foliage dry weight, respectively.*
IF STAND = '0' THEN DELETE;	*With this command we are excluding the dummy seedlings(labeled as '0').*
SQRTSTEM = SQRT(STEM);	*Transforming to square root.*
R = (ROW – 36.5) / 35.5 ;	*Defining the linear term row position(R) and "centering" and "scaling" its values.*
R2 = R * R ;	*Defining the quadratic term R^2 (R2).*
COL_DIS = COLUMN * 0.19 ;	*0.19 m is the distance between plots across columns.*
ROW_DIS = ROW * 0.57 ;	*Since we pooled the data of all trees within a plot (by estimating the average), the distance between centers of plots is: 3 x 0.19m = 0.57 m.*
PLOT_ID = (ROW * 10) + COLUMN ;	*We need a new class variable for plot to be used in the REPEATED statement.*
PROC MIXED;	*Invoking Procedure Mixed, designed for models with random and fixed effects*
CLASS SOIMOIST STAND COLUMN PLOT_ID;	*R and R2 are not declared in the CLASS statement, since they are used as regression variables.*
MODEL SQRTSTEM = SOIMOIST R R2 / DDFM=SATTERTH;	*This corresponds to model (C). Independent variables declared here are fixed effects. DDFM=SATTERTH specifies estimation of the degrees of freedom for the F-test using the correction by Satterthwaite (Milliken and Johnson 1984). It is recommended with unbalanced and/or missing data.*
RANDOM STAND(SOIMOIST) COLUMN COLUMN*STAND(SOIMOIST);	*The RANDOM statement allow us to specify which variables are assumed to be random effects.*
PARMS (0.3) (0.07) 0.01) (0.03) (0.31);	*PARMS specifies initial values for variance components and correlations. We selected values based on the output of a preliminary run without the PARMS statement. Note that the values are provided in the order of the parameters as specified in the RANDOM statement with the spatial and ERROR terms assumed last: STAND, STAND*C, SP and ERROR.*
REPEATED PLOT_ID / TYPE=SP(POW) (COL_DIS ROW_DIS) SUB=INTERCEPT;	*The REPEATED statement specifies the covariance structure (among-plots in our case). TYPE=SP specifies a spatial structured correlation (other options are available). POW specifies the spatial correlation term given in equation [5]. (COL_DIS ROW_DIS) is needed for constructing a Euclidean matrix of distances among plots. SUB=INTERCEPT is required in order for the spatial autocorrelation to be valid for all data points.*
	REDUCED MODEL WITHOUT STAND(SOIL-MOISTURE)
PROC MIXED DATA=FOUR; CLASS BLOCK SOIMOIST STAND COLUMN COL_DIS ROW_DIS PLOT_ID;	
MODEL SQRTSTEM=SOIMOIST R R2;	*Notice that DDFM=SATTERTH statement is not needed in the reduced models.*
RANDOM COLUMN COLUMN*STAND(SOIMOIST);	*Reduced model without STAND(SOIMOIST).*
REPEATED PLOT_ID / TYPE=SP(POW) (COL_DIS ROW_DIS) SUB=BLOCK ;	

term for testing soil-moisture classes is approximately the Means Square (MS) of Block x Soil-Moisture + MS of Stand(Soil-Moisture) – MS of Block x Stand(Soil-Moisture), using the method of synthesis (MILLIKEN and JOHNSON, 1984). Note that the Block x Soil-moisture interaction is significant; this may contribute to the non-significance of soil-moisture classes.

Trend analysis

A visual examination of the seedlings in the common garden test identified a strong within-block gradient in growth across the five columns (*i.e.* the gradient was perpendicular to the larger dimension; *Figure 1*). This is evident by noting generally larger values of stem dry weight in columns 4 and 5 and smaller values in columns 1 and 2 (*Table 2*). This pattern is moderately consistent across the four blocks. Apparently, the management of neighboring nursery beds (timing of seedling lifting, shade effect and pattern of irrigation) influenced our experimental seedlings. Also, there appeared to be smaller fluctuations across the rows, probably due to drainage variation. Thus, the standard RCBD analysis is inadequate to account for this within-block variation.

Table 4. – Analysis for randomized complete block design (model A), using PROC GLM.

S.V.	DF	M.S.	F	P
BLOCK	3	2.942	2.55	0.099
SM	4	1.331	1.27	0.347
STAND(SM)	42	0.514	0.77	0.831
BLOCK x SM	12	1.206	1.84	**0.048**
BLOCK x STAND(SM)	111	0.674	1.14	0.134
ERROR	110	0.546		

SM = Soil-moisture class. Error term for testing SM is = $0.95MS_{BLOCK\ x\ SM} + 0.84MS_{STAND(SM)}\ 0.79MS_{BLOCK\ x\ STAND(SM)} - 0.004MS_{ERROR}$. Error term for testing STAND(SM) is = $0.938MS_{BLOCK\ x\ STAND(SM)} + 0.062MS_{ERROR}$.

Table 5. – Analysis for modified trend analysis using PROC GLM (model B).

S.V.	DF	M.S.	F	P
SM	4	0.72	2.33	0.067
STAND(SM)	42	0.30	0.70	0.911
COLUMN	4	13.24	30.57	**0.001**
COLUMN x STAND(SM)	116	0.45	1.35	0.055
R	1	1.31	3.95	**0.049**
R2	1	0.55	1.64	0.203
ERROR	114	0.33		

SM = Soil-moisture class. Error term used for SOIMOIST was $0.882MS_{STAND(SOIMOIST)} + 0.035MS_{COLUMN\ x\ STAND(SOIMOIST)} + 0.083MS_E$; for STAND was $0.883MS_{COLUMN\ x\ STAND(SOIMOIST)} + 0.117MS_E$; and for COLUMN was $0.867MS_{COLUMN\ x\ STAND(SOIMOIST)} + 0.133\ MS_E$. Significant P-values are in bold.

The resulting analysis using model (B) (equation [4]) suggests that column position (C*m*) and row position (R_p) are highly significant (*Table 5*). This confirms the effects of systematic row and column contribution to the variability from the RCBD model. Note the substantial reduction of the error mean square from 0.546 in the randomized complete block design (*Table 4*) to 0.333 in the modified trend analysis (*Table 5*). However, the soil-moisture effect is only weakly significant (P-value = 0.067). There was also weak significance for the stand by column interaction (P-value = 0.055) (*Table 5*).

Analysis incorporating spatially correlated errors using PROC MIXED

Table 6 presents the partial P-values for the variance components and correlation using the likelihood ratio test for the random effects and the F-test for the fixed effects. An example of how the partial P-values were estimated is shown in *table 7*. Note that the differences among soil-moisture classes are now strongly significant (partial P-value = 0.005; *Table 6*), as opposed to the results from the previous analysis. We have now accounted for the spatial autocorrelation as well as the systematic variability (note that the error term is now reduced to 0.313) so that the soil-moisture class effect is now found to be significant.

The significance of the column by stand (within soil-moisture class) interaction (partial P-value = 0.002, *Table 6*) suggests that there is a genotype x environment interaction, where some stands behave differently depending on the column in which they are located; (we note that there are important differences in sunlight and moisture availability among columns). The high significance of the among-plot spatial autocorrelation (*Table 6*) confirms the existence in the nursery bed of a small-scale spatial structure among plots. Correspondingly, this speaks to the utility of PROC MIXED, which allows us to take into account spatially-structured correlations.

Alternative models examined

We considered the possibility that the increased significance of the soil-moisture effect using model (C) (equation [6]) might

Table 6. – Edited output of SAS program for model (C), modified trend analysis and correlated errors using PROC MIXED.

Random effects			*Fixed effects*				
Source	Estimation	partial P	Source	NDF	DDF	TYPE III F	Pr > F
STAND(SOIMOIST)	0	**1**	SOIMOIST	4	140	3.94	**0.0046**
COLUMN	0.2966	**0.0001**	R	1	27.1	2.24	0.1464
COLUMN*STAND(SOIMOIST)	0.0718	**0.0022**	R2	1	28.3	3.46	0.0734
SPATIAL AUTOCORRELATION	0.0387	**0.0001**					
ERROR	0.3130						

Model fitting information.

DESCRIPTION	VALUE	*COMMENTS*
Observations	283	*Number of observed plots.*
Res Log Likelihood	-252.6	*Logarithm of the Restricted (or Residual) Maximum Likelihood (REML)*
Akaike's Information Criterion	-256.6	*Log REML minus the number of variance and correlation parameters.*
-2 Res Log Likelihood (-2RLL)	505.3	*-2 Log REML.*

NDF = numerator degrees of freedom, DDF = denominator degrees of freedom. Significant P-values are in bold.

Table 7. – Example of computation of partial P-values using the likelihood ratio statistic.

ESTIMATIONS	*COMMENTS*
-2 Res Log Likelihood (-2RLL) = 514.7	*From SAS output from Reduced Model omitting COLUMN x STAND(SOIMOIST)*
$\chi^2 = -2RLL_{\text{Reduced Model}} - (-2RLL_{\text{Full Model}})$	*Calculation of Chi-square for COLUMN x STAND(SOIMOIST)*
$\chi^2 = 514.7 - 505.3 = 9.4$	
P-value for a χ^2 statistic of 9.4 with 1 d.f. is = 0.0022	

be due to using PROC MIXED instead PROC GLM, since the former might have handled the imbalance better, rather than be the inclusion of the plot spatial autocorrelation term. Thus, we used PROC MIXED for models [A] and [B] and obtained essentially the same results as those found with PROC GLM presented in *Tables 4* and *5*. Thus, we believe that model [C] is better primarely because of the inclusion of the spatial term ($r_{mp,\,qr}$).

We also considered the possibility that the heavy imbalance that resulted in the absence of many stands in some of the columns may be responsible for the basic outline of our results. We removed seven stands that were represented in only one or two columns and reran an analysis (not shown) using model [C]. The results were essentially unchanged. We also considered the 'random coefficient model' (Littell *et al.*, 1996, Chapter 7) for describing the interaction between stand and column. However, this approach did not prove capable of adequately describe the data.

A potential problem in maximizing the likelihood that can occur with spatial correlation

In analyses with the variable square root of foliage dry weight, (not stem dry weight used in the previous sections), we employed a model similar to model [C]. In examining a reduced model (omitting stand nested in soil-moisture classes) for foliage dry weight, we discovered that PROC MIXED computed a larger likelihood value than for the full model (including stand nested in soil-moisture class). This is a violation of the theory that requires that the likelihood of the full model must be at least as large as that with the reduced model. We resolved this inconsistency by finding that the likelihood surface is bimodal (*Figure 2*). During iterations of the REML procedure with the full model, the program converged to a local maximum that was smaller than the global (overall) maximum. The smaller peak in *figure 2* represents the local maximum whereas the larger peak shows the global maximum. This resulted from the starting values that the program automatically assigned to the variance terms and the spatial autocorrelation. These starting values caused the interactions to reach the lower peak, as opposed to the higher one which represents the true REML values. Based on our experience, PROC MIXED has a tendency, on occasion, to converge to a local – but not global – maximum, when spatial correlation structures are included in the model (SAS, 1996b).

This difficulty can be corrected in PROC MIXED programs by using the PARMS statement in both the full and reduced models. This requires providing either single starting values or a grid of starting values for all random components and the spatial autocorrelation of the model (SAS, 1996b). For the stem dry weight we used single starting values (*table 3*). For foliage dry weight a broad grid of values was chosen with the selected values based on estimates obtained without the PARMS statement. The grid of values was (0 to 0.05 by 0.025), (0 to 0.1 to 0.05), (–0.03 to 0.03 by 0.1) and (0.48) for the random terms stand, interaction stand by column, spatial autocorrelation and error, respectively. With the use of the PARMS statement, the likelihood for the reduced model was never larger than that for the full model, in agreement with theory. *Figure 2* was developed from the output of a PROC MIXED program for foliage dry weight similar to that presented in *table 3*. For graphical purposes, we used a fairly fine grid of PARMS values. Thus, when using models including a term for spatial autocorrelation, particularly if the sign of this autocorrelation is negative, we strongly recommend use of the PARMS statement.

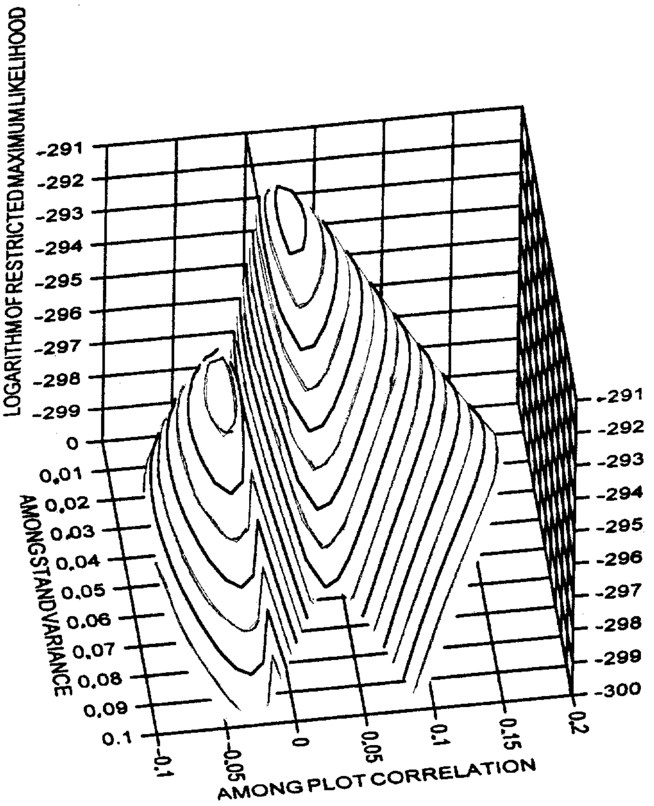

Figure 2. – Surface of the logarithm of the restricted maximum likelihood (RLL) for the square root of foliage dry weight of 1.5-year-old *Pinus banksiana* seedlings. Note the existence of two peaks.

Conclusions

We found that trend analysis was an important addition to the standard RCBD in accounting for large-scale environmental variation within the experimental area. The further addition of correlated errors, using PROC MIXED, also accounted for the apparent small-scale microsite variation. This led to the best description of our data. This model resulted in an error term that allowed us to detect significant differences among soil-moisture classes. Analysis including spatial autocorrelations, may be helpful for other examples with complicated spatial structure. When using common garden studies for provenance testing, information on the row and column position of each experimental unit should be recorded in order to make trend analysis and the inclusion of correlated errors possible.

PROC MIXED is also more versatile than PROC GLM for including several random factors. This can be particularly relevant in nested models. However, we wish to point out two drawbacks to PROC MIXED. First, running PROC MIXED can be very time consuming. Some of our programs took up to 45 minutes of CPU time to run (Hewlett Packard mainframe computer HP9000/H50, 128 Mb memory). The use of a grid of values for the PARMS statement and the use of the DDFM=SATTERTH option tend to be very time consuming. When appropriate, we found it useful for saving CPU time to provide single starting values instead of a grid of values for the PARMS statement (although a grid was sometimes necessary to avoid achieving a local maximum). In addition, it may be useful to omit the DDFM=SATTERTH option whenever it is not necessary (for example, when screening models or for reduced models). Second, we felt that the PROC MIXED release version 6.12 is somewhat cumbersome in that it does not automatically provide the likelihood ratio statistic for testing estimated variance components and spatial autocorrelation. Such tests required further runs with reduced models. Despite these concerns, PROC MIXED is strongly recommended for the types of analysis we considered.

Acknowledgments

The work was supported in part by funds from the Wisconsin Department of Natural Resources, Bureau of Forestry (Project 133-0768) to RPG and a graduate fellowship to CSR from the Mexican Council of Science and Technology (CONACyT). CSR gratefully acknowledges the support from the Centro de Investigación en Genética y Ambiente, Universidad Autonoma de Tlaxcala, Tlaxcala, Mexico, during the completion of this research. We thank ANDREW MONK for seed collection; JIM ADAMS, LYNN HUMMEL and LAURA VAN SLYKE of the University of Wisconsin-Madison Walnut St. Greenhouses, TRENT MARTY and personnel of the Wilson State Nursery for test maintenance. We also thank MICHAEL CASLER, ERIC KRUGER and DON WALLER for valuable suggestions on an early draft of this manuscript.

Literature

BRIDGWATER, F. E., TALBERT, J. T. and ROCKWOOD, D. L.: Field design for genetic test of forest trees. In: Progeny Testing of Forest Trees. Proceedings of Workshop on Progeny Testing, June 15 to 16, 1982, Auburn, Alabama. Southern Cooperative Series Bulletin No. 275. pp. 28–39 (1983). — BROWNIE, C., BOWMAN, D. T. and BURTON, J. W.: Estimating spatial variation in analysis of data from yield trials: a comparison of methods. Agron. J. **85**: 1244–1253 (1993). — GUMPERTZ, M. L. and BROWNIE, C.: Repeated measures in randomized block and split-plot experiments. Can. J. For. Res. **23**: 625–639 (1993). — KIRK, H. J., HAYNES, F. L. and MONROE, R. J.: Application of trend analysis to horticultural field trials. J. Amer. Soc. Hort. Sci. **105**: 189–193 (1980). — KOTAR, J. and BURGER, T. A.: Guide to Forest Communities and Habitat Types of Central and Southern Wisconsin. Department of Forestry, University of Wisconsin-Madison, Madison, Wisconsin. 278 p. (1996). — LITTELL, R. C., MILLIKEN, G. A., STROUP, W. W. and WOLFINGER, R. D.: SAS System for Mixed Models. SAS Institute, Cary, North Carolina. 633 p. (1996). — MAGNUSSEN, S.: A method to adjust simultaneously for spatial microsite and competition effects. Can. J. For. Res. **24**: 985–995 (1994). — MILLIKEN, G. A. and JOHNSON, D. E.: Analysis of Messy Data. Vol. 1: Designed Experiments. Van Nostrand Reinhold, New York. 473 p. (1984). — SAS Institut: SAS/STAT User's Guide; release 6.03 Edition. SAS Institute, Cary, North Carolina. 1028 p. (1988). — SAS Institut: Advanced General Linear Models with an Emphasis on Mixed Models Course Notes. SAS Institute, Cary, North Carolina. 614 p. (1996a). — SAS Institut: SAS/STAT Software. Changes and Enhancements through Release 6.11. SAS Institute, Cary, North Carolina. 1094 p. (1996b). — SEARLE, S. R., CASELLA, G. and MCCULLOCH, C. E.: Variance components. John Wiley and Sons, New York. (1992). — SOKAL, R. R. and ODEN, N. L.: Spatial autocorrelation in biology. 1. Methodology. Biol. J. Linn. Soc. **10**: 199–228 (1978). — WRIGHT, J. W.: Introduction to Forest Genetics. Academic Press, New York. 463 p. (1976). — ZIMMERMAN, D. L. and HARVILLE, D. A.: A random field approach to the analysis of field-plot experiments and other spatial experiments. Biometrics **47**: 223–239 (1991). — ZOBEL, B. and TALBERT, J.: Applied Forest Tree Improvement. John Wiley, New York. 505 p. (1984).

Interactions of Genotype with Site for Height and Stem Straightness in *Pinus taeda* in Zimbabwe

By D. P. GWAZE¹), J. A. WOLLIAMS²), P. J. KANOWSKI³) and F. E. BRIDGWATER⁴)

(Received 10th November 2000)

Summary

The magnitude and importance of genotype x environment interactions (GXE) for stem height and straightness at 2, 10, 14 and 23 years of age were estimated for *Pinus taeda* L. planted at four sites in Zimbabwe. Genetic correlations between heights at different sites varied from low to high (0.18 to 0.95). They were not consistent across the different ages, but were generally high between all sites at younger ages and progressively diminished with age. These results suggest that early growth assessments may not be reliable for predicting GXE at maturity. Genetic correlations between stem straightness scores at different sites also varied from low to high (-0.13 to 0.95); those between sites A and C, and between sites B and D, were consistently high (>0.84) at ages greater than 2 years.

On the basis of the magnitude of genetic correlations, site B was generally the best selection site for height. At 23 years,

---

[1] Texas Forest Service, Forest Science Lab., Texas A & M University, College Station, TX 77843-2585, USA

[2] Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, U. K.

[3] Department of Forestry, Australian National University, ACT 0200, Australia

[4] USDA-Forest Service, Forest Science Lab., Texas A&M University, College Station, Texas 77843-2585, USA