NOTE

NUMBER OF TREES PER EXPERIMENTAL UNIT IS IMPORTANT WHEN COMPARING TRANSPLANT STRESS INDEX VALUES

DAVID B. SOUTH, CURTIS L. VANDERSCHAAF

School of Forestry and Wildlife Sciences and Alabama Agricultural Experiment Station,
Auburn University, Alabama 36849-5418, United States

and C. T. SMITH

Department of Forest Science,
Texas A&M University, College Station, Texas 77843-2135, United States

(Received for publication 29 July 2002; revision 6 December 2002, 10 May 2003)

ABSTRACT

The ability to detect statistically significant treatment differences is dependent on the number of experimental units, the alpha level, and the coefficient of variation for the response variable. Some response variables are inherently more variable than others. As variability increases, a greater number of trees per experimental unit or a greater number of experimental units is required if researchers want to avoid making a Type II error (i.e., accepting a false null hypothesis). In this study (containing 20 experimental units), a significant ($\alpha = 0.05$) treatment effect was obtained for an 8-cm difference in height growth using only 10 *Pinus radiata* D. Don seedlings per experimental unit. However, when Transplant Stress Index (TSI) means were being analysed, 120 seedlings per experimental unit were required before a difference (of 0.25) was declared significant. Because TSI values are inherently more variable than height growth data, the Type II error rate for TSI in some studies may be higher than the error rate for height growth when plots contain less than 100 pine seedlings per experimental unit (when studies contain 20 experimental units or less).

Keywords: transplant shock; experimental design; plot size; statistics; *Pinus radiata*

INTRODUCTION

Bare-root seedlings typically undergo stress after transplanting (Mullin 1963; Sands 1984; Rietveld 1989). Until recently, providing an objective estimate of the intensity or duration of transplanting shock has been difficult. South & Zwolinski (1997) proposed a Transplant Stress Index (TSI) that has proved useful in quantifying stress. They defined TSI as the slope of the linear relationship between shoot height at the beginning of the growth period and height increment. A negative TSI indicates seedlings have undergone transplant shock (the more negative the value, the greater the intensity of transplant stress). Negative

TSI values have been observed for seedlings of *Pinus nigra* Arn. var. *maritima* (Jinks & Kerr 1999), *Pinus radiata* (South & Smith 2000), *Pinus taeda* L. (South *et al*. 2001), and *Pinus halepensis* Miller (Oliet *et al*. 2002). In several trials, negative TSI values occurred only during the first year after transplanting. Unlike height growth or diameter growth, the TSI value can be determined only for a population of seedlings (it cannot be determined for an individual seedling).

The size of the experimental unit can be an important factor in experimental designs (Savage 1956; Stern 1968; Correll & Cellier 1987). When one is analysing for differences in TSI values among treatments, it is important to have a sufficiently large number of seedlings in each experimental unit in order to reduce the inherent variation associated with this variable. If the number of measured trees is too small, it will be difficult to detect real differences in TSI between treatments and a Type II error (Snedecor & Cochran 1967; Foster 2001) may occur. Therefore, South & Zwolinski (1997) recommended that TSI values not be calculated when there are less than 100 measured trees. Since many establishment trials have experimental units containing less than 40 trees, one might ask what effect sample size has on conclusions derived from the analysis of TSI data. This paper examines the effect of tree number per experimental unit on the ability to detect significant treatment effects for one study with *Pinus radiata* (with four replications).

METHODS

A study was installed on a Taupo sandy loam soil at the Kinleith Forest in the central North Island of New Zealand (38°14´ S, 175°58´ E, elevation 490 m). The area has a mild, maritime climate. Bare-root *P. radiata* seedlings (1/0) were hand-planted in August 1992 at a 2 × 2-m spacing. At the initial height (h_0) measurement in November, seedling heights ranged from 4 cm to 52 cm (mean 27 cm). The study was established as a randomised complete block design with four blocks and five treatments. Each plot (i.e., experimental unit) contained 200 trees (10 rows of 20 trees per row). Seedling heights (h_1) were measured 1 year after planting in mid-winter. For each plot, a TSI value was obtained by the slope of the linear relationship between shoot height at the beginning of the growth period (h_0) and height increment (South & Zwolinski 1997).

This study involved an area where five harvest treatments were used prior to replanting. The treatments included: (A) whole-tree harvest and forest litter removed, (B) whole-tree harvest, (C) a stem-only harvest with a V-blade treatment (to facilitate hand planting), and (D) stem-only harvest. Weed competition in these four treatments was virtually eliminated manually and with herbicides throughout the trial. The fifth treatment (E) was also a stem-only harvest but there was no weed control at time of establishment

DATA ANALYSIS

To illustrate the effect of tree numbers on the ability to detect significant differences, 270 analyses were conducted using estimates based on 5, 10, 20, 40, 60, 80, 100, 120, and 140 measurement trees per plot. Data were analysed using the General Linear Model procedure provided by the SAS-PC program (SAS 1988). One-third of the analyses were for survival, one-third were for height, and the remaining third were for TSI. Each analysis contained 20 experimental units (i.e., one response value per experimental unit) and the tree measurements used in each analysis were selected at random (from a total of 200 trees) using the RANUNI uniform generator statement in SAS. Error terms with 12 degrees of freedom (d.f.) were used to test for significance in the block (3 d.f.) and treatment (4 d.f.) factors. The equation for the model of a randomised complete-block design is as follows:

$$Y_{ij} = \mu + \tau_i + \beta_j + \beta\tau_{ij} + \varepsilon_{ij}$$

where: Y_{ij} = the response of treatment *i* in replication *j*
μ = the overall mean
τ_i = treatment effect of the ith treatment
β_j = replication effect of the jth replication
$\beta\tau_{ij}$ = a replication x treatment interaction effect or plot error
ε_{ij} = experimental error

After 10 analyses were conducted for each number of measurement trees, the probability values (Fig. 1) and sums-of-square values (Table 1) were averaged and a power test was conducted (Nemec 1991; Foster 2001). Mortality data were transformed prior to analysis using a square root transformation in an attempt to enhance normality. Dead seedlings were not included in the determination of TSI values.

To illustrate the variation in mean values that can occur, one analysis was randomly selected from each group of 10 analyses (Table 2). When the overall treatment effect for an analysis was significant ($\alpha = 0.05$), means were compared using Duncan's multiple-range test. As a comparison, an additional analysis included 200 trees per experimental unit.

RESULTS AND DISCUSSION

When there were only five trees per experimental unit, there were no significant treatment differences ($\alpha = 0.05$) in height, mortality (transformed), or TSI (Fig. 1). For

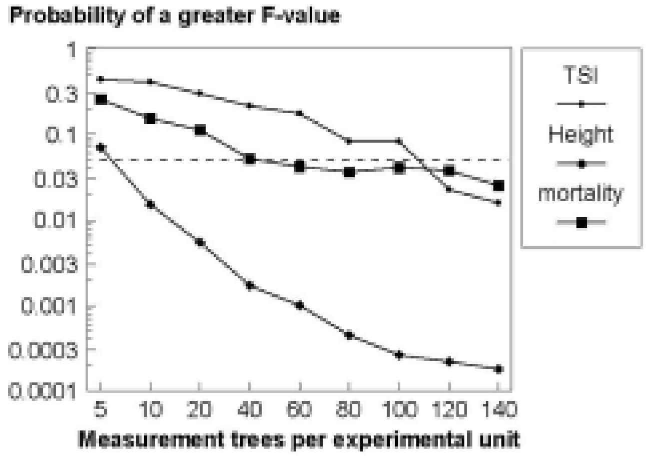

FIG. 1–The effect of the number of measurement trees per experimental unit on the probability of detecting a significant treatment effect on mortality, height growth, and transplant stress index (TSI). The dashed line indicates the 5% level of probability of a greater F-value for treatment effect. Each dot represents an average of 10 probability values. This example contains 20 experimental units.

TABLE 1–The average sums-of-squares for TSI, first-year height, and first-year mortality (transformed) for 20 experimental units. Sums of squares are the means from 10 analyses. Values in parentheses indicate the number of analyses that resulted in a significant (α = 0.05) treatment effect.

Source	df	Number of trees per experimental unit								
		5	10	20	40	60	80	100	120	140
TSI										
Block	3	14.15	0.99	0.41	0.17	0.11	0.06	0.06	0.03	0.02
Treatment	4	15.21	1.68	0.74	0.44	0.41	0.41	0.39	0.38	0.36
Error	12	49.43	3.87	1.23	0.70	0.49	0.32	0.33	0.24	0.20
Total	19	(2)	(2)	(2)	(1)	(3)	(6)	(6)	(9)	(10)
Height										
Block	3	404.9	340.1	327.8	312.0	307.2	304.9	297.1	290.4	285.9
Treatment	4	526.4	581.7	525.0	514.7	522.7	503.6	491.2	483.5	475.4
Error	12	500.0	318.6	209.6	157.9	141.8	125.6	109.7	104.6	99.4
Total	19	(6)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)
Mortality										
Block	3	17.68	14.69	12.69	13.12	10.43	9.65	8.03	6.93	7.10
Treatment	4	31.33	33.57	25.65	26.03	24.88	23.48	21.55	20.12	20.01
Error	12	54.12	35.27	28.28	19.56	16.56	15.18	14.58	13.83	13.18
Total	19	(2)	(4)	(2)	(6)	(6)	(8)	(7)	(7)	(9)

mortality and TSI, the power was less than 40% (Fig. 2). However, when there were 140 trees, the average F-value for treatment was significant for all variables. Therefore, the number of trees measured per experimental unit affected the Type II error rate for all three variables.

When five trees were used per experimental plot, Type II errors were made in 40%, 80%, and 80% of 10 analyses for height, mortality, and TSI, respectively (Table 1). The power of the test increased as the number of measurement trees increased (Fig. 2).

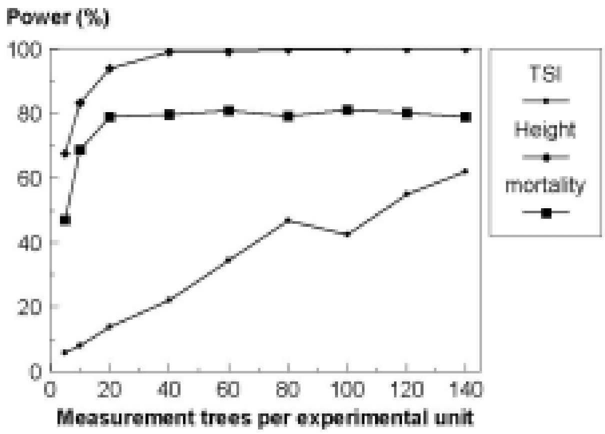

FIG. 2–Power curves for the Kinleith study showing the effects of the number of measurement trees on statistical power (1–β) when detecting a 5-cm difference in height, an 8-percentage point difference in mortality, and a 0.1 difference in TSI (α = 0.05).

TABLE 2–Examples of the variation in transplant stress index (TSI) values, first-year heights, and statistical significance that can occur when using output from a single GLM procedure.

Treatment		Number of trees per experimental unit									
		5	10	20	40	60	80	100	120	140	200
TSI	A	–0.5	0.1	–0.3	–0.1	–0.1	–0.2	–0.2c	–0.2c	–0.2b	–0.2c
	B	0.2	0.4	0.1	0.0	0.0	0.1	0.2a	0.1ab	0.1a	0.1ab
	C	0.5	0.5	0.1	–0.1	–0.1	–0.1	–0.1bc	0.0bc	–0.1b	0.0bc
	D	0.0	–0.1	0.2	0.3	0.2	0.1	0.1ab	0.2a	0.2a	0.1a
	E	1.3	0.4	0.2	0.1	0.1	0.2	0.1ab	0.1ab	0.2a	0.1ab
	LSD	1.74	0.98	0.35	0.45	0.31	0.32	0.20	0.20	0.19	0.17
Height (cm)											
	A	46.0	49.9b	52.0b	48.5b	49.7b	49.9b	49.5b	49.5b	49.7b	49.7b
	B	47.7	44.6b	47.9b	47.0b	46.6b	46.6b	47.1b	46.3b	46.8b	46.8b
	C	62.2	60.6a	60.1a	60.5a	61.0a	60.0a	60.6a	59.7a	59.8a	60.1a
	D	45.7	48.5b	48.6b	49.4b	48.4b	48.7b	48.4b	48.4b	48.5b	48.5b
	E	50.0	46.7b	45.8b	47.9b	47.2b	46.7b	47.5b	47.3b	47.1b	47.3b
	LSD	12.28	7.95	7.09	5.86	5.06	5.42	4.36	4.64	3.97	4.44
First-year mortality (%)											
	A	0	0	1	2bc	0	1b	1	1c	1b	1b
	B	10	7	5	4ab	7	5b	5	5ab	4ab	5ab
	C	0	0	3	2c	1	3b	2	1bc	1b	2b
	D	15	13	10	14a	14	14a	15	14a	14a	14a
	E	10	7	11	10a	7	10a	10	10a	10a	10a
	LSD (transformed)	3.58	2.83	1.95	1.44	1.89	1.49	2.06	1.45	1.76	1.56

Parameter values in the same column and followed by the same letter are not significantly different (p=0.05) according to Duncan's Multiple Range Test.

Significant differences in height could be consistently detected when plots contained 10 trees, but for mortality 60 trees were required before the treatment effect was large enough to be significant (Fig. 1). Values for TSI were more variable and therefore 120 trees were required to detect a significant treatment effect. For TSI, the average probability level when using 100 trees was p = 0.0825 (Fig. 1). Even with 100 trees per plot, the power to detect a 0.1 difference in TSI was less than 50%. These data suggest that, compared to measurements of height or mortality, TSI will require more trees per plot to obtain an equivalent Type II error rate.

The precision of the analysis (i.e., the ability to detect differences) increased as the number of trees per experimental unit increased. For this study, a difference of 8 cm in height could be detected with only 10 trees per experimental unit. Likewise, a 5% difference in mortality could be detected (most of the time) when using 80 trees per plot. A 0.2 difference in TSI could be detected 90% of the time when using 120 trees per plot (Table 1).

Since these results are specific to the Kinleith study, the absolute treatment differences and required tree numbers should not be extrapolated to other species or trials. For example, when using 100 trees per plot, a 0.35 difference in TSI was declared significant in a similar study in the South Island but a 0.9 difference in TSI was not significant in a trial on the west coast of the North Island (South & Smith 2000). With *Pinus taeda* in the United States, a 0.35 difference in TSI was significant using 80 measurement seedlings per plot (South *et al*. 2001).

The TSI values are easily affected by variance associated with small samples. For example, the estimated TSI value for treatment C ranged from +0.5 to –0.1 when there were less than 60 trees per experimental unit. Therefore, it is important to measure enough trees so the regression equation (per experimental unit) remains relatively stable. This is the basic reason why more trees per plot are required for TSI values than for mortality or height values.

CONCLUSIONS

Some researchers may want to test the hypothesis that regeneration practices affect TSI values. These individuals should be aware that TSI values are inherently more variable than other measures of seedling performance such as mortality and height growth. To avoid making Type II errors, tests should be designed so they could detect a TSI difference of 0.2. A study with this level of precision may require more trees per plot than typically used in outplanting trials.

ACKNOWLEDGMENTS

The authors are grateful for the constructive comments made on this manuscript by Dr Greg Somers of Auburn University and Dr Richard Woollons of the University of Canterbury. This work was financially supported by the New Zealand Forest Research Institute Ltd and by the Auburn University Southern Forest Nursery Management Cooperative.

REFERENCES

CORRELL, R.L.; CELLIER, K.M. 1987: Effects of plot size, block size and buffer rows on the precision of forestry trials. *Australian Forest Research 17*: 11–18.

FOSTER, J.R. 2001: Statistical power in forest monitoring. *Forest Ecology and Management 151*: 211–222.

JINKS, R.L.; KERR, G. 1999: Establishment and early growth of different plant types of Corsican pine (*Pinus nigra* var. *maritima*) on four sites in Thetford Forest. *Forestry 72*: 293–304.

MULLIN, R.E. 1963: Planting check in spruce. *Forestry Chronicle 39*: 252–259.

NEMEC, A.F.L. 1991: Power analysis handbook for the design and analysis of forestry trials. *British Columbia Ministry of Forestry, Research Branch Victoria, B.C. Biometrics Information Handbook No. 2*. 26 p.

OLIET, J.; PLANELLES, R.; ARIAS, M.L.; ARTERO, F. 2002: Soil water content and water relations in planted and naturally regenerated *Pinus halepensis* Mill. seedlings during the first year in semiarid conditions. *New Forests 23*: 31–44.

RIETVELD, R.J. 1989: Transplanting stress in bareroot conifer seedlings: its development and progression to establishment. *Northern Journal of Applied Forestry 6*: 99–107.

SANDS, R. 1984: Transplanting stress in radiata pine. *Australian Forest Research 14*: 67–72.

SAS INSTITUTE INC. 1988: "SAS User's Guide: Statistics". SAS Institute Inc., Cary, N.C. 584 p.

SAVAGE, M.B. 1956: What about the size of the sample plot? *Forestry Chronicle 32*: 75–76.

SOUTH, D.B.; RAKESTRAW, J.L.; LOWERTS, G.A. 2001: Early gains from planting large-diameter seedlings and intensive management are additive for loblolly pine. *New Forests 22*: 97–110.

SOUTH, D.B.; SMITH, C.T. 2000: Harvest treatments and fertilizer application affect transplant stress index of *Pinus radiata. New Zealand Journal of Forestry Science 30*: 308–314.

SOUTH, D.B.; ZWOLINSKI, J.B. 1997: Transplant stress index: a proposed method of quantifying planting check. *New Forests 13*: 315–328.

SNEDECOR, G.W.; COCHRAN, W.G. 1967: "Statistical Methods". Iowa State University Press, Ames. 593 p.

STERN, K. 1968: Optimum plot size in field experiments with forest trees. *Silva Fennica 2*: 248–260.