INCOMPLETE BLOCK DESIGNS FOR GENETIC TESTING: ACCURACY OF RANKING FAMILIES AND INDIVIDUALS

Yong-Bi Fu[1], Alvin D. Yanchuk[2] & Gene Namkoong[1]**

[1)] Department of Forest Sciences, University of British Columbia, #3041-2424 Main Mall, Vancouver, BC V6T 1Z4 Canada and [2)] Research Branch, B.C. Ministry of Forests, 712 Yates Street, Victoria, BC V8W 9C2, Canada.
** Author to whom all correspondence should be addressed:
Dr. Yong-Bi Fu, Plant Gene Resources of Canada, AAFC Saskatoon Research Centre, 107 Science Place, Saskatoon, SK S7N 0X2, Canada. Tel: 306-956-7642; Fax: 306-956-7246; Email: fuy@em.agr.ca

Received December 6, 1999; accepted September 29, 2000

ABSTRACT

A randomized incomplete block design (ICB) can increase the precision of estimating family means on heterogeneous test sites and consequently increase the probability of having an accurate ranking of families, but the magnitude of the increase in the probability compared with a randomized complete block design (RCB) has been less clear. To address this issue, a stochastic computer simulation of a full-sib progeny trial with 90 families of 10 seedlings per family was conducted on a test site varying with patchiness and gradients in environmental effects and assuming a range of heritabilities. Seedlings were laid out following both RCB and ICB with single-tree plots. Family means were estimated and combined family plus individual selection indices based on estimated genetic parameters were applied for individual trees. Kendall's coefficient of concordance was used to measure the degree of agreement between the true and estimated ranks. Results from the simulation showed that the individual heritability, rather than the block design, had the main impact on the accuracy of ranking both the families and individuals. The higher the true heritability, the more accurate the estimated ranking. Both designs showed a slight increase in accuracy of ranking with increased patchiness and gradients of environmental effects. Generally, ICB showed a slight superiority over RCB in ranking accuracy. The ranking of families was more accurate than ranking of individuals.

Keywords: incomplete block design, spatial variation, heritability, family evaluation, simulation

INTRODUCTION

Forest genetic field trials usually involve large numbers of families that must be grown at wide spacing often on heterogeneous test sites for up to 20 years (LOO-DINKINS 1992, MAGNUSSEN 1993). These features, along with the large operational costs required for the trials, explain at least in part why efficient field designs are desirable for accurate and precise estimates of breeding values. However, this has not prompted much effort over the last three decades to explore alternative field designs in forest progeny trials (MAGNUSSEN 1993), although some stimulating research has been done (e.g. see MCCUTCHAN *et al.* 1985, FRIEDMAN & NAMKOONG 1987, WILLIAMS & MATHESON 1994). The randomized complete block design (RCB), the most commonly applied field design, can provide some control of site variability by simple blocking, but its ability to account for the site variability is quite limited as blocks used are typically large in size (500–1000 trees per block of size 0.3–0.6 hectares). Recent examinations of a series of the British Columbia (B.C.) Douglas-fir progeny trials showed that the applied RCB partitioned only an average of 5% of the site variations displayed over the 66 test sites in tree height, while an effective blocking could remove up to 24% (FU *et al.* 1999a). Given the site variations comparable to those in the Douglas-fir progeny trials, implementation of a randomized incomplete block design (ICB) in the trials would achieve smaller error variance of family mean estimates than using RCB. The relative efficiency (which was commonly obtained as the ratio of the error variance of family mean estimates for RCB to the error variance for ICB) would be 1.25 or higher (FU *et al.* 1998). In their studies on the progeny trial of *Pinus patula* in Zimbabwe, Dr. Richard Barnes and his colleagues observed an average increase of 30% in relative efficiency of the applied triple lattice design over RCB (BARNES & SCHWEPPENHAUSER 1979, BARNES *et al.* 1992a, 1992b). In his analysis of a *Pinus banksianan* Lamb. family test with a cubic lattice design in Manitoba, Canada, Dr. Jerome Klein showed a reduction of the proportion of the plot error variance for 10-years height from 13.9 % (when analyzed as a randomized block

experiment) to 3.1% (KLEIN 1989). Clearly, these studies are encouraging, as they endorse use of ICB in forest genetic field trials.

In this study, we attempted to address the accuracy of ranking both families and individuals with RCB and ICB tests which was one of the issues raised to us during our investigation on implementations of ICB (FU *et al.* 1998, 1999a, 1999b, 1999c). Tree breeders are typically interested in achieving not only an increased precision, but also accuracy, of ranking families and individuals from a forest genetic field trial. On heterogeneous test sites, use of small blocks (or small blocking) can remove more site variation to achieve higher precision of estimating family means, as discussed above. It can be reasoned that such increased precision could increase the probability of having an accurate ranking of families, but less clear is the magnitude of the increase in the probability from an ICB test when compared with an RCB test. A large increase in the probability from an ICB test would be desirable for tree breeders to have as an additional justification of implementing ICB, but the empirical study in the *Pinus patula* progeny trial did not seem to suggest so. Even with an increase of 30% in precision of the ranking from the ICB tests as mentioned above, there were no practically significant shifts in family ranking after both RCB and ICB analyses (BARNES & SCHWEPPENHAUSER 1979, BARNES *et al.* 1992a, 1992b). Clearly, it is of value to examine the magnitude of the difference in accuracy of ranking families and individuals between the two block designs. With a known expectation for the difference, tree breeders could exploit more judiciously the possible advantage from implementations of small blocking.

The objective of this study was to examine by a stochastic computer simulation the accuracy of ranking families and individuals from RCB and ICB tests of 90 full-sib families with 10 seedlings each on a heterogeneous test site and to illustrate the possible magnitudes of the difference in their ranking accuracy. Computer simulation provides a powerful means for such illustration, as the true ranks of families and individuals based on only their genetic values can be known and compared with the estimated ranks. Ideally, the probability of having an accurate ranking of all the 90 families and 900 individuals should be used for the comparison, but realistically the probability with such a large number of families used is very small (i.e., less than 0.0001), thus impossible to be calculated from a simulation with 1000 runs or less. For this reason, Kendall's coefficient of concordance (W) was used to determine the degree of agreement between the true and estimated ranks. This coefficient ranges from 0 to 1; the higher the observed coefficient of concordance, the more accurate the estimated ranking.

METHODS

Genetic trial and block design

We considered a full-sib progeny trial of forest trees aimed at evaluating family and individual performances. This trial comprised of 90 families with 10 seedlings per family and was conducted on a test site of 18 rows (y) and 50 columns (x) that exhibited spatial environmental variations as modeled below. Seedlings were laid out with single-tree plots following the rules of either RCB or ICB. For RCB, one seedling was randomly selected from each of 90 families and randomly allocated to a block of 5 contiguous columns, and this was repeated for the other nine replicates. For ICB, 90 families were randomly allocated into 10 half-column blocks (of size 9) in 5 contiguous columns in the first replicate, and this process was repeated for the other nine replicates. This simple ICB is "no constraint over replicates" called NC as described in FU *et al.* (1998) and is expected to be only slightly less efficient than the Alpha design (a class of generalized lattice Designs; PATTERSON & WILLIAMS 1976, FU *et al.* 1998).

Spatial model of environmental variation

We applied a spatial model that is widely used in geostatistics to describe spatial processes (CRESSIE 1991, CLARKE *et al.* 1997). Under this model, every experimental plot has co-ordinates (x, y) so that the yield (U) from a plot with co-ordinates (x, y) receiving genotype k may be written as

$$U_{xy(k)} = t_k + \gamma x + e_{xy}, \qquad [1]$$

where t_k is the genetic value of the genotype k, γx is the gradient specified in x co-ordinate (γ – the slope of the gradient), and e_{xy} is the error term in the experiment which has a zero mean and co-variances which are functions of distance (CRESSIE 1991, pp13–25). To simulate $U_{xy(k)}$, we needed to specify t_k (defined below), the gradient with γ, and e_{xy} as the joint variance matrix of plot errors (with dimensions of xy × xy; 900 × 900 in this case) because plot errors were correlated. In this study, we assumed that **e** (bold here for matrix) has mean(**0**) and variance(**e**) as:

variance(**e**) = **V** = $\{v_{xy,x'y'}\}$,

where $\{ v_{xy,x'y'}\} = [1 - \frac{3}{2}(\frac{\delta_{xy,x'y'}}{\alpha}) + \frac{1}{2}(\frac{\delta_{xy,x'y'}}{\alpha})^3]\sigma_e^2$

$(x, y) \neq (x', y')$ and $0 \leq \frac{\delta_{xy,x'y'}}{\alpha} < 1$

$$= 0 \qquad (x, y) \neq (x`, y`) \text{ and } \frac{\delta_{xy,x'y'}}{\alpha} \geq 1,$$

where $\delta_{xy,x'y'}$ is the unit distance between plots xy and $x'y'$, α is the maximum distance apart beyond which plot yields are uncorrelated (also called the range parameter; note that the intensity of patchiness is not specified under this model), and σ_e^2 is the fully random error variance in the site. With this 'spherical' covariance matrix of **V**, we generated **e** as:

$$\mathbf{e} = \mathbf{V}^{1/2}\,\mathbf{n},$$

where $\mathbf{V}^{1/2}$ is the Choleski factor or square root of **V**, and **n** is a normally distributed vector (of dimensions $xy \times 1$) with a mean of zero and variance of 1. Details of how this model can generate error distributions over a test site by specifying drift and range parameters γ and α were provided by FU *et al.* (1998). In this study, six scenarios of site variation were specified to represent the spread of typical site variations observed in the B.C. Douglas-fir progeny trials (see FU *et al.* 1999a for the findings in detail). They were (1) $\alpha = 1$ $\gamma = 0$, (2) $\alpha = 6$ $\gamma = 0$, (3) $\alpha = 10$ $\gamma = 0$, (4) $\alpha = 6$ $\gamma = 0.005$, (5) $\alpha = 6$ $\gamma = 0.01$, and (6) $\alpha = 6$ $\gamma = 0.02$. For example, a value of 6 for α shows a patch size of 6 plots across and $\gamma = 0.02$ would mean an one-meter difference (0.02 cm × 50 columns) in height for the trees 50 columns apart.

It should be mentioned that there are many spatial models that one could apply to specify environmental variations (e.g., see LEGENDRE & FORTIN 1989, MAGNUSSEN 1990, CRESSIE 1991). The spatial model used in this study, however, is capable of taking into account the 'spherical' correlation and threshold function (specified by α) that reflect reasonably well most of the spatial patterns. Also, our preliminary assessment of this model with the residuals of tree height at ages 9–12 in the B.C. Douglas-fir progeny trials (FU *et al.* 1999a) supports for the use of this model. For example, there were 50 out of the 66 test sites showing the visual fit of the spherical covariance model for the residuals, after removing the broad trends (i.e., deterministic structures) in row and column directions.

Generation of simulated data

The true genetic values for the 900 trees were generated as:

$$t_k = t_{fw} = \mu + H_f + H_w \qquad [2]$$

where μ was the overall mean of 5 (arbitrary units), H_f ($f = 1..90$ for family) and H_w ($w = 1..10$ for seedling) were normally distributed deviates with means of zero and variances of σ_f^2 (family genetic variance) and σ_{wf}^2 (within-family genetic variance), respectively. For full-sib progeny, $\sigma_f^2 = \sigma_{wf}^2 = [0.5h^2/(1-h^2)]\sigma_e^2$, assuming a pure additive genetic model for h^2 (individual heritability) and σ_{total}^2 (the total phenotypic variance) = $\sigma_f^2 + \sigma_{wf}^2 + \sigma_e^2$. As mentioned above, σ_e^2 (random environmental variance) was assumed to be 1. Three levels of h^2 (0.05, 0.25, and 0.45) were specified. Based on the true genetic values generated, the true ranks were made from 1 to 900 and from 1 to 90, respectively, for the 900 individuals and the 90 families generated. When the individual trees were allocated with the design rules to plots, the three components of plot error values (specified with the above spatial model for gradient, patchiness, and random error) were simply added accordingly to the true genetic values to obtain individual phenotypic values.

Estimation and ranking

To estimate family means under both RCB and ICB designs, a mixed model was used for analysis of the simulated phenotypic values as:

$$U_{ij(k)} = \mu + \tau_k + \beta_i + p(i, j) \qquad [3]$$

where $U_{ij(k)}$ was the phenotypic observation of a trait made on family k ($k = 1..90$) in the jth unit of the ith block ($i = 1..100$, $j = 1..9$ for ICB; $i = 1..10$, $j = 1..90$ for RCB), μ the overall mean, τ_k the fixed effect of family k, β_i the random effect associated with block i [$E(\beta) = 0$, $E(\beta^2) = \sigma_b^2$], and $p(i, j)$ the plot residual associated with the observation $U_{ij(k)}$ [$E(p) = 0$, $E(p^2) = \sigma_e^2 + \sigma_{wf}^2$. This was done with SAS® PROC MIXED which allows for estimation of family means and family mean variance with the generalized least squares estimators in which block and error variance components were estimated by the REML procedure (SEARLE *et al.* 1992). To estimate individual heritability for ranking the individuals, the variance components for family, block, and plot were estimated with a random model as in Equation (3) in which all the three factors were considered random, including the family factor with expected mean zero and variance σ_f^2. The estimation was also done with SAS® PROC MIXED. Note that the estimate of block variance σ_b^2 was included for an unbiased estimation of h^2 as $\hat{h}^2 = 2\hat{\sigma}_f^2/(\hat{\sigma}_f^2 + \hat{\sigma}_b^2 + \hat{\sigma}_p^2)$, although this is not common in practice as block effects are often considered fixed.

Estimated ranks for the families were made based on the estimates of the 90 family means. Estimated ranks for the individual trees were generated using the combined indices. For each of the 900 trees, a combined index (FALCONER 1981: Chapter 13) was calcu-

lated as follows:

$$I_{fw} = \hat{h}_f^2(\hat{\tau}_f - \hat{\mu}) + \hat{h}_w^2(PV_w - \hat{\tau}_f),$$

where $\hat{\tau}_f$ was the estimate of family mean τ_f (f = 1..90), $\hat{\mu}$ was the estimate of overall mean μ, and PV_w was the phenotypic value for the seedling w (w = 1..10) of the family f, and $\hat{h}_f^2$ and $\hat{h}_w^2$ were calculated (see BAKER 1986) as:

$$\hat{h}_f^2 = \frac{1 - 0.5(r-1)}{1 + 0.5(r-1)\hat{h}^2}\,\hat{h}^2 \text{ and }, \hat{h}_f^2 = \frac{0.5}{1 - 0.5\hat{h}^2}\,\hat{h}^2$$

where r (the number of replicates) was equal to 10 in this study.

Kendall's coefficient of concordance (W) and its standard deviation were calculated between the true and the estimated ranks (ZAR 1984: section 20.16) for each of the two designs. This coefficient ranges from 0 (when there is no agreement among the two ranks) to 1 (when there is complete agreement among the two ranks). The higher the coefficient, the less the rank change from the true ranks (i.e., the more accurate the estimated ranking). Roughly speaking, an average W of 0.6 would mean that 60% of the 90 families examined have a match between the estimated and the true ranks. Note that this coefficient is related to the commonly used Spearman rank correlation coefficient r_s by $W = (r_s + 1) / 2$ (see ZAR 1984: section 20.16) but ranges from –1 to 1, so W is more preferred for measuring the degree of agreement between the two rankings.

Simulation procedure

Simulations of the progeny trial were made with a computer program written in SAS® MACRO and IML (SAS INSTITUTE INC 1995). This program included several steps: (1), obtaining the genetic values (**t**) for 900 individuals for a given level of individual heritability as in Equation (2) and keeping constant for all repetitions of simulation with various spatial parameters; (2), generating a spatial error distribution (**e**) for the 900 plots of the site, assuming a multivariate normal distribution with a mean and variance given in Equation (1); (3), obtaining the design matrices **Xt** for treatments and **Xb** for blocks which were pre-determined with a PC PASCAL program based on the described design rules and performing a randomization for the design; (4), generating the observation data $\mathbf{U}_{xy(k)} = \mathbf{X}_t\,\mathbf{t} + \mathbf{e}$; and (5) analyzing the simulated data using SAS® PROC MIXED (SAS INSTITUTE INC 1995) with the mixed and random model given in Equation (3) and REML procedure to estimate family means and family variances for calculation of individual heritability. Ranking was made based on the estimates of the family means and the combined indices for individual trees. Kendall's coefficients of concordance (W) between the true ranks and estimated ranks were calculated as mentioned above for each run of simulation. The average and standard deviation of W were calculated over 200 repeated runs. To evaluate the effectiveness of blocking, a ratio of average block variance versus average plot variance was also calculated. This was repeated for 18 combinations of spatial variation (α = 1–10, γ = 0–0.02) and individual heritability (h^2 = 0.05–0.45) for the two designs. Details of the computer simulation program were given in FU *et al.* (1998) and it is also available from the first author upon request.

RESULTS

The accuracy of ranking the 90 families and 900 individuals measured by the average Kendall's coefficient of concordance is given in Table 1 for the RCB and ICB designs in 18 combinations of site variation and individual heritability. Several patterns are clear. First, as expected, the individual heritability used had a large impact on the accuracy of ranking the families and individuals irrespective of the designs employed. For example, when the heritability was low (i.e., h^2 = 0.05), the average coefficients of concordance were less than 0.58 for ranking the families, but with h^2 = 0.45, these coefficients could be up to 0.94.

Second, comparisons of the two designs for the effectiveness of removing site variation clearly showed ICB outperformed RCB as evident in the ratio of average block variance versus average plot variance (Table 1). However, high effectiveness of ICB contributed only slightly to the difference between the two designs in accuracy of ranking the families and individuals as shown in the coefficient of concordance. ICB showed a slight superiority over RCB in accuracy of ranking the family and individuals, particularly when individual heritability was 0.25 or less.

Third, the ranking of families was generally more accurate than the ranking of individuals, as shown with the average coefficients of concordance. For example, Kendall's coefficient of concordance was up to 0.58 with h^2 = 0.05 for ranking the families and was up to 0.44 for ranking the individuals.

Fourth, the accuracy of ranking the families and individuals varied more according to design and spatial parameters with the lower heritability, as evidenced in the standard error of the coefficient of concordance (Table 1). Such variation was comparable between RCB and ICB.

Table 1. Accuracy of ranking families and individuals measured by W (Kendall's coefficient of concordance) in a simulated field trial with 90 full-sib families of 10 seedlings each on a test site of spatial environmental variations under RCB and ICB designs with single-tree plots.

Heritability / site variation	Ranking families				Ranking individuals			
	RCB		ICB		RCB		ICB	
	$\hat{W}(SD)$	$\hat{\sigma}_b^2/\hat{\sigma}_p^2$	$\hat{W}(SD)$	$\hat{\sigma}_b^2/\hat{\sigma}_p^2$	$\hat{W}(SD)$	$\hat{\sigma}_b^2/\hat{\sigma}_p^2$	$\hat{W}(SD)$	$\hat{\sigma}_b^2/\hat{\sigma}_p^2$
$h^2 = 0.05$								
$\alpha = 1; \gamma = 0$	0.46 (0.08)	0.00	0.46 (0.09)	0.01	0.35 (0.05)	0.00	0.35 (0.06)	0.01
$\alpha = 6; \gamma = 0$	0.49 (0.08)	0.15	0.51 (0.08)	0.46	0.37 (0.05)	0.14	0.39 (0.05)	0.47
$\alpha = 10; \gamma = 0$	0.52 (0.08)	0.33	0.58 (0.07)	0.97	0.40 (0.05)	0.31	0.44 (0.05)	1.00
$\alpha = 6; \gamma = 0.005$	0.48 (0.08)	0.14	0.52 (0.08)	0.49	0.38 (0.05)	0.15	0.40 (0.06)	0.49
$\alpha = 6; \gamma = 0.01$	0.49 (0.08)	0.15	0.51 (0.08)	0.50	0.37 (0.05)	0.16	0.40 (0.05)	0.50
$\alpha = 6; \gamma = 0.02$	0.49 (0.08)	0.23	0.51 (0.08)	0.62	0.37 (0.06)	0.24	0.40 (0.06)	0.60
$h^2 = 0.25$								
$\alpha = 1; \gamma = 0$	0.83 (0.03)	0.00	0.83 (0.03)	0.01	0.67 (0.02)	0.00	0.67 (0.02)	0.01
$\alpha = 6; \gamma = 0$	0.85 (0.03)	0.11	0.86 (0.03)	0.39	0.69 (0.02)	0.12	0.70 (0.02)	0.40
$\alpha = 10; \gamma = 0$	0.87 (0.03)	0.27	0.89 (0.03)	0.76	0.70 (0.02)	0.23	0.72 (0.02)	0.77
$\alpha = 6; \gamma = 0.005$	0.85 (0.03)	0.12	0.86 (0.03)	0.39	0.68 (0.03)	0.12	0.69 (0.02)	0.39
$\alpha = 6; \gamma = 0.01$	0.85 (0.03)	0.14	0.86 (0.03)	0.42	0.68 (0.03)	0.14	0.70 (0.02)	0.44
$\alpha = 6; \gamma = 0.02$	0.84 (0.04)	0.21	0.86 (0.03)	0.50	0.68 (0.03)	0.22	0.69 (0.02)	0.52
$h^2 = 0.45$								
$\alpha = 1; \gamma = 0$	0.91 (0.02)	0.00	0.90 (0.02)	0.02	0.73 (0.02)	0.00	0.73 (0.02)	0.03
$\alpha = 6; \gamma = 0$	0.92 (0.02)	0.08	0.92 (0.02)	0.28	0.74 (0.03)	0.08	0.74 (0.02)	0.28
$\alpha = 10; \gamma = 0$	0.93 (0.02)	0.16	0.94 (0.02)	0.51	0.75 (0.03)	0.18	0.75 (0.03)	0.52
$\alpha = 6; \gamma = 0.005$	0.91 (0.02)	0.09	0.92 (0.02)	0.30	0.73 (0.03)	0.10	0.74 (0.02)	0.30
$\alpha = 6; \gamma = 0.01$	0.92 (0.02)	0.11	0.92 (0.02)	0.33	0.74 (0.02)	0.11	0.74 (0.03)	0.31
$\alpha = 6; \gamma = 0.02$	0.92 (0.02)	0.18	0.92 (0.02)	0.42	0.72 (0.03)	0.18	0.73 (0.02)	0.40

DISCUSSION

Review of simulated results

Our simulation shows that the individual heritability used, not the block designs employed, had the major impact on the accuracy of ranking the families and individuals. The higher the true heritability, the more accurate the estimated ranking. Both block designs showed a slight increase in accuracy of ranking the families and individuals with increased patchy and gradient environmental variations. ICB showed a slight superiority over RCB in accuracy of ranking the families and individuals, particularly when individual heritability was 0.25 or less. These simulated results agree well with the observation of no practically significant shifts in family ranking in the progeny trial of *Pinus patula* after replacing RCB with ICB (BARNES & SCHWEPPENHAUSER 1979, BARNES *et al.* 1992a, 1992b).

Why is there no significant increase in accuracy of ranking families expected between RCB and ICB? The accuracy of ranking family means in a population depends largely on the coefficient of variation (CV) for the population that is determined by family genetic variance (as a function of heritability). This can be conveyed better with an example. Considering two families (A and B) with extreme expected means of 5 and 7, respectively, in a population. After an RCB test, the two families have estimated means of 5 and 7, respectively, but equal estimates of standard error 0.60. With an ICB test, they have the same estimates of family means as in the RCB test, but smaller estimates of standard error 0.5, which gives a relative efficiency of 1.44 for ICB over RCB. In this case, the probability of having overlap between estimates of two means is low. Thus a high accuracy of ranking is expected for both designs. If the case is changed with the extreme expected mean for the family B reduced from 7 to 5.4, the probability of having two mean estimates overlapping is high, even with the same relative efficiency of 1.44 obtained for ICB over RCB. This will result in a

low accuracy of ranking for both designs. The coefficients of variation for these two cases are approximately 17.0 and 3.9 in percentage, respectively (calculated by

$$CV = \frac{E(B) - E(A)}{1.96} / \frac{E(B) - E(A)}{2},$$

where E is the expectation for family mean). This clearly shows the impact of CV (or heritability) on the accuracy of ranking.

Follow-up simulation

To illustrate the impact of increased precision of estimating family means on the accuracy of ranking, an additional simulation was performed, following the cases given above. A normally distributed deviate with a mean specified by CV and an error variance of means (specified by the relative efficiency; i.e., the error variance of estimates of family means for RCB divided by the respective error variance for ICB) was randomly generated for each family. The deviates for the two families were compared to determine if the family A had a smaller mean than the family B. The two deviates were also ranked and Kendall's coefficient of concordance was calculated. This was repeated for 10,000 runs. The probability of having a correct ranking of the families A and B (B > A) was the proportion of the runs that the family A had a smaller estimated deviate than the family B. An average of Kendall's coefficient of concordance was calculated over the runs. In the simulation, the expected mean was fixed to be 5 for the family A and changed for the family B from 5 to 7 with an interval of 0.2 to reflect the increase in CV. The variance for RCB was fixed to be 0.36 (or se = 0.6) and the two relative efficiencies (1.19 and 1.44) were specified to reflect the realistic range of the increased precision from ICB. The results are given in Table 2. As expected, increasing precision from ICB increased the probability of having a correct ranking and the accuracy of ranking, but the impact of increasing precision was much smaller than that with CV.

Concluding remarks

It was our hope that these simulations (either for the full-sib progeny trial or simple reasoning) can provide a better understanding of the accuracy of ranking families and individuals with RCB and ICB tests, so that tree breeders would be in a better position to take the possible advantages from implementations of ICBs. The simulated results are clear. The accuracy of ranking families and individuals can be increased with the replacement of RCB with ICB on heterogeneous test sites, as illustrated in Tables 1 and 2. However, the magnitude of the increase in accuracy is expected to be small, as the accuracy of ranking depends largely on the level of individual heritability for the traits of interest and not much on the increase of precision from ICB. In

Table 2. Impacts of increasing the relative efficiency (*RE*) of estimating family means and the coefficient of variation (*CV*) for the population on the probability of having correct ranking of the two families (A and B) with extreme means and the accuracy of ranking them with Kendall's coefficient of concordance (*W*).

Extreme means		CV (%)	Prob (A<B)			*W* (agreement of ranking)		
			RCB	ICB	ICB	RCB	ICB	ICB
A	B		se* = 0.6	se = 0.55	se = 0.5	se = 0.6	se = 0.55	se = 0.5
			RE = 1	*RE* = 1.19	*RE* = 1.44	*RE* = 1	*RE* = 1.19	*RE* = 1.44
5	5.0	0	0.49	0.50	0.50	0.00	0.00	0.00
	5.2	2.0	0.61	0.61	0.62	0.22	0.23	0.24
	5.4	3.9	0.67	0.67	0.71	0.35	0.34	0.42
	5.6	5.8	0.78	0.77	0.81	0.55	0.54	0.62
	5.8	7.6	0.84	0.84	0.87	0.68	0.69	0.73
	6.0	9.3	0.89	0.89	0.92	0.79	0.79	0.84
	6.2	10.9	0.92	0.94	0.95	0.84	0.88	0.89
	6.4	12.5	0.94	0.96	0.98	0.89	0.92	0.95
	6.6	14.1	0.97	0.99	0.99	0.93	0.97	0.98
	6.8	15.6	0.98	0.99	1.00	0.96	0.98	0.99
	7.0	17.0	0.99	1.00	1.00	0.98	1.00	1.00

* se stands for the standard error of the mean estimate.

practice, this expectation probably holds as most of the individual heritabilities observed for various traits are 0.25 or higher in forest genetic trials (ZOBEL & TALBERT 1984). Considering just this factor alone, one may argue that ICB is not so much more attractive than RCB, but the combined advantages and benefits from implementations of ICBs are substantial. These include the high statistical efficiency of evaluating family performance, the test of more families in single trials, the flexibility in the choice of planting site on irregular surface areas, the sampling of more extreme environmental ranges, and the promise of reducing operational costs in tree improvement programs (see FU *et al.* 1998).

ACKNOWLEDGEMENT

The authors would like to thank Drs. Richard Barnes, Rowland Burdon, Steen Magnussen and Chang-Yi Xie for their helpful discussions on the issue and comments on the earlier version of the manuscript and Drs Rowland Burdon and Carolyn Raymond and an anonymous journal reviewer for their comments on the manuscript. This work was financially supported by Forest Renewal BC research grant No. HQ 96445RE.

REFERENCES

BAKER, R. J. 1986: Selection Indices in Plant Breeding. CRC Press, Boca Raton, FL, 218 pp.

BARNES, R. D. & SCHWEPPENHAUSER, M. A. 1979: *Pinus patula* Schiede and Deppe progeny tests in Zimbabwe Rhodesia: Genetic control of 1.5 year-old traits and a comparison of progeny test methods. *Silvae Genetica* **28**: 156–167.

BARNES, R. D., MULLIN, L. J. & BATTLE, G. 1992a: Genetic control in fifth year traits in *Pinus Patula* Schiede and Deppe. *Silvae Genetica* **41**: 242–248.

BARNES, R. D., MULLIN, L. J. & BATTLE, G. 1992b: Genetic control in eighth year traits in *Pinus Patula* Schiede and Deppe. *Silvae Genetica* **41**: 318–326.

CLARKE, G. P., HAINES, L. M. & FU, Y. B. 1997: Kriging and field experiments. *In*: Proceedings of 51st session of the International Statistical Institute, book 2, pp. 229–232. International Statistical Institute, Istanbul.

CRESSIE, N. A. C. 1991: Statistics for Spatial Data. John Wiley & Sons, Inc. New York, 900 pp.

FALCONER, D. S. 1981: Introduction to Quantitative Genetics. Longman Inc. New York, 340 pp.

FRIEDMAN, S. & NAMKOONG, G. 1987: Estimating family means using unbalanced incomplete blocks. *In*: Proceedings of IUFRO Joint Meeting of Working Parties on Breeding Theory, Progeny Testing, and Seed Orchards, Williamsburg, USA (1986). pp. 457–468. North Carolina State University, Industrial Cooperative Tree Improvement Program, Raleigh, N.C.

FU, Y. B., YANCHUK, A. D. & NAMKOONG, G. 1999a: Spatial patterns of tree height variations in a series of Douglas-fir progeny trials: implications for genetic testing. *Canadian Journal of Forest Research* **29**: 714–723.

FU, Y. B., YANCHUK, A. D. & NAMKOONG, G. 1999b: Incomplete block designs for genetic testing: some practical considerations. *Canadian Journal of Forest Research* **29**: 1871–1878.

FU, Y. B., YANCHUK, A. D., NAMKOONG, G. & CLARKE, G. P. Y. 1999c: Incomplete block designs for genetic testing: statistical efficiencies with missing observations. *Forest Science* **45**: 374–380.

FU, Y. B., CLARKE, G. P. Y., NAMKOONG, G. & YANCHUK, A. D. 1998: Incomplete block designs for genetic testing: statistical efficiencies of estimating family means. *Canadian Journal of Forest Research* **28**: 977–986.

KLEIN, J. I. 1989: Effects of incomplete block design and stand structure on heritability of growth in a jack pine family test in Manitoba. *Canadian Journal of Forest Research* 19: 1044–1050.

LEGENDRE, P. & FORTIN, M. J. 1989: Spatial pattern and ecological analysis. *Vegetatio* **80**: 107–138.

LOO-DINKINS, J. 1992: Field test design. *In*: Handbook of quantitative forest genetics. (ed. L. Fins, S. Friedman & J.V. Brotschol). pp. 96–139. Kluwer Acadamic Publisher, Bos-ton.

MAGNUSSEN, S. 1990: Application and comparison of spatial models in analyzing tree-genetics field trials. *Canadian Journal of Forest Research* **20**: 536–546.

MAGNUSSEN, S. 1993: Design and analysis of tree genetic trials. *Canadian Journal of Forest Research* **23**: 1144 –1149.

MCCUTCHAN, B. G., OU, J. X. & NAMKOONG, G. 1985: A comparison of planned unbalanced designs for estimating heritability in perennial crops. *Theoretical and Applied Genetics* **71**: 536–544.

PATTERSON, H. D. & WILLIAMS, E. R. 1976: A new class of resolvable incomplete block designs. *Biometrika* **63**: 83–92.

SAS® INSTITUTE INC. 1995: SAS® User Guide. Version 6.03 Ed. Cary, NC, 1028 pp.

SEARLE, S. R., CASELLA, G. & MCCULLOCH, C. E. 1992: Variance Components. John Wiley & Sons, Inc. New York, 501pp.

WILLIAMS, E. R. & MATHESON, A. C. 1994: Experimental Design and Analysis for Use in Tree Improvement. CSIRO publication, Australia, 174 pp.

ZAR, J. H. 1984: Biostatistical Analysis. 2nd edition. Prentice-Hall, Inc., New Jersey, 718 pp.

ZOBEL, B. & TALBERT, J. 1984: Applied Forest Tree Improvement. John Wiley & Sons, New York, 505 pp.