Selection of parents for the Sitka spruce breeding population in Britain and the strategy for the next breeding cycle

S.J. LEE

Tree Improvement Branch, Forestry Commission Research Agency, Northern Research Station, Roslin, Midlothian EH25 9SY, Scotland

Summary

Simple techniques were used to overcome the problems associated with estimating breeding values for a number of traits prior to calculating parental multi-trait index values for over 800 plus trees within the Sitka spruce General Breeding Population (GBP). Progeny from different trees were planted in half-sib tests in different years. It took an 11-year period involving numerous sites of differing quality to complete the test establishment. The estimation of parental breeding value was restricted to across-sites analysis involving the mean performance of open-pollinated progeny collected from each plus tree planted over a maximum of three representative sites; in each case the data from each site were weighted by the family heritability for the trait and site. Standard parameter estimates for family heritability and phenotypic and genetic correlations, and standard values of the control common to all experiments, were calculated by averaging estimates over all across-sites analysis. Ranking of re-selected parents to form the GBP was based on index values calculated using multi-trait index selection. Two hundred and forty reselected parents have been allocated to one of four sub-lines of equal mean index value. Plus trees have been ranked within each sub-line based on index value. Full-sib controlled pollinations according to an amended assortative mating procedure are now planned between selections within each sub-line. The second generation breeding population will be created by selecting the best individuals within the best families.

Introduction

Sitka spruce (*Picea sitchensis* Bong. Carr) is the main plantation species in Britain and represents nearly 30 per cent of the area of commercial forestry (Forestry Industry Council of Great Britain, 1998) equivalent to around 40 million planted trees per year.

Fletcher and Faulkner (1972) outlined the methods to be employed in the improvement of Sitka spruce by selection and breeding. Breeding value estimation of economic traits for Sitka spruce plus trees has been progressing well since the first progeny tests were planted in 1967 (Lee, 1994). Wind (or open) pollinated seed was collected from all plus trees. The trees raised from

the seed are called open-pollinated progeny and collectively the progeny of one parent are referred to as a family. Since the mother was known but the pollen contribution was from a large number of surrounding trees, the progeny are considered half-siblings (half-sibs) to each other. Sitka spruce is an out-crossing species and the possibility of self pollination is assumed to be zero.

By measuring individual progeny for different traits in replicated comparative tests established in units of around 50 families, it is possible to calculate the breeding value (or genetic value) of the parent plus tree for each of the measured traits. Parental breeding values (BVs) are a function of genetic quality which can be passed from one generation to the next in a predictable manner. These BVs are usually expressed in terms of deviation relative to the experimental mean or to a standard control. It took 27 years to establish all the tests required to estimate the BVs of the 1800 plus trees thought to originate from the Queen Charlotte Islands (QCI), British Columbia (Canada). The last tests were planted in spring 1993.

The economic traits under selection within the Sitka spruce breeding programme are: end of rotation volume; stem form; and wood density. The earliest indirect selection ages for these traits are respectively 6-year height, 6- or 7-year stem-form and 15-year density assessed using a Pilodyn (see Lee, 1992). A negative correlation between 15-year diameter and wood density ($r = -0.34$ to -0.69; Wood, 1986) ensures that any selection programme weighted towards diameter without regard to wood density will indirectly lead to a reduction of wood density. This would be considered as detrimental to the strength of British-grown Sitka spruce which already has a marginal wood density for construction timber (approximately 350 kg/m^3 when air-dried to 12 per cent moisture content; Lavers, 1983). If breeders are to maximize diameter and stem form without decreasing wood density it follows that data are required on all three of these economic traits prior to any reselection. This would restrict reselection of plus trees for the breeding population to those represented by progeny in tests at least 15 years old from which wood density data have been collected.

An internal review of the Sitka spruce breeding strategy (Lee, 1994) concluded that it was unacceptable to delay the next cycle of Sitka spruce breeding until estimation of breeding values of all traits for all plus trees had been complete (1993 + 15 years = 2008 at the earliest). It was then decided that a preliminary Sitka spruce breeding population should be assembled, consisting of the best reselections available amongst the original plus trees for which breeding value estimates of diameter, stem form and wood density were available. Once constructed, progress could be made into the next cycle of breeding and selection by making controlled crosses between members of the breeding population. Reselections based on breeding value estimations of plus trees represented in later-established progeny tests would be incorporated into the breeding population at a future date.

Further recommendations (Lee, 1994) were that resources would be most efficiently used in the next generation by carrying out positive assortative mating (PAM; Mahalovich and Bridgewater, 1989) and concentrating a disproportionate effort on superior clones (Lindgren, 1986). PAM means that parents are not crossed randomly but according to some pre-arranged grouping of the parents. An extreme form of PAM would be to mate first and second ranked entries, third and fourth ranked entries and so on. A milder form of PAM involves sub-dividing parents into broader groups across predefined rank ranges and then restricting crosses to parents within each group. Once parents have been allocated to groups, increased genetic gain may be possible as a result of increased selection intensity by using higher ranked parents in a larger number of crosses.

In order to carry out PAM, it is necessary to be able to compare estimated breeding values for each plus-tree in test so that some sort of ranking can be performed for all the selected population. Plus tree breeding values are calculated as the product of twice the progeny-mean deviation from the standard control and the estimated family heritability (Hodge and White, 1992). The factor two is introduced since only half the genetic make-up of the originally selected parent plus tree contributed to the progeny.

Family heritability is a function of the ratio of genetic variance which can be transferred through generations to the total variance operating within a population (i.e. genetic and environmental

variance). Family heritability can therefore vary with site and can be influenced by progeny test design. Ranking of plus trees based on breeding values relative to a standard control would be most straightforward if all progeny from all plus trees had been planted in tests of the same design over the same number of limited sites in the same year or years, and specific traits had been assessed on all trees including the control at the same time. The reality is far from this.

The objective of this paper is twofold. Firstly it is to outline the relatively straightforward procedures adopted to minimize the effects due to different planting years and different planting sites of varied quality, in the multi-trait reselection of 240 parents for the General Breeding Population (GBP) of the British Sitka spruce breeding strategy from an original population of over 1000. The second objective is to outline the strategy to be adopted in the next cycle of Sitka spruce breeding.

Materials and methods

Progeny representing over 1000 plus trees had been planted in half-sib tests on 80 different sites in nine different years between 1967 and 1977. The collection of sites onto which the same group of families was planted is referred to as a series. The number of series planted in each year varied from one to seven. There were 19 series in all and the number of families planted in each series varied around 50. Progeny from a particular plus tree tended to be unique to a given series; only rarely were progeny collected from the same plus tree planted in more than one series. Diameter at breast height (cm), stem form and wood density had each been assessed on most sites at an age of 15–22 years from planting. Occasionally a site may have been rejected for assessment of one or more of the three traits due to a number of factors including widespread frost damage, animal browsing and generally poor survival.

Stem form was assessed using a 1 = good to 6 = bad subjective score (Lee, 1992). Wood density was measured indirectly by firing a blunt pin from a Pilodyn 6J once into each tree at or about breast height, avoiding any whorls of branches. The penetration of the pin (mm) gives an indirect assessment of wood density. Previous studies (Wood, 1986; Lee, 1997a, b) have found there to be a good correlation between whole tree density as measured using X-ray densitometry techniques and the much faster (indirect) method using the Pilodyn.

In accordance with the objectives of the GBP, reselection was dependent on selecting a population of original plus trees which on average would have a mean wood density breeding value no less than the standard QCI control, whilst maximizing improvements for stem diameter and stem form. The negative genetic correlation between diameter and wood density means that very few trees would qualify for the breeding population if every parent had to equal or exceed the control for wood density, whilst satisfying certain minimum criteria for diameter and stem form. Multi-trait selection index (Smith, 1936; Hazel, 1943) assists the tree breeder in selecting a population of trees which *on average* will satisfy predefined selection criteria and was therefore used in this study to meet the required objectives.

Much of the theory associated with the multi-trait selection of Sitka spruce is outlined in Lee (1999) although relevant parts are repeated here for ease of reference.

Index theory

If a population is to satisfy certain predefined selection criteria *on average* it follows that not all individuals within the population will satisfy the minimum selection requirements for each trait. Therefore, certain individuals within the reselected population may have (for example) wood density BVs less than the control but well above average BVs for diameter, whilst others will have good positive wood density BVs but more modest diameter BVs. Identification of the parent trees which should make up a population according to the predefined criteria is a particularly difficult process, especially when negative correlations are involved. Multi-trait index selection theory allows the tree breeders to meet these criteria as efficiently as possible.

An index value (I) for each parent is calculated, based on the sum of progeny mean performance for each trait (P) weighted by an index coefficient (b). The index coefficient for each trait is estimated using information about trait variation, genetic and phenotypic correlations between each

pair of traits, trait heritability and the relative economic weight applied to each trait.

$$I = b_1P_1 + b_2 P_2 + b_3 P_3 \qquad (1)$$

where I = multi-trait index value, b = index coefficient, P = observed progeny mean value and subscripts 1, 2, 3 = three economic traits such as diameter, wood density and stem form.

Parents are ranked according to the multi-trait index value calculated from the mean performance of their progeny. A top-ranked proportion of the original population can then be reselected to form a breeding population. The mean breeding values of the reselected population are then calculated and expressed as a percentage deviation from the standard control in order to measure individual trait gains at the population level. The gains can then be compared to the selection objectives.

Prior to analysis, all input variables are generally known apart from the relative economic weights to be applied to each trait. It is also possible to apply restrictions to certain traits, known as 'Kempthorne Restrictions' (Kempthorne and Nordskog, 1959), which prevent changes in selected traits (e.g. wood density) while maximizing others (e.g. diameter and stem straightness).

Problems associated with analysis

The main considerations when attempting to draw together multi-trait BVs based on across-site performance of half-sib progeny for comparison and ranking leading to reselection for a breeding population were: (1) the comparison of BVs when progeny representing parents had been established on varying numbers of sites across series; (2) the adjustment of progeny mean values for the varying quality of sites within series; (3) equating index values across sites and series using a comparable scale; and (4) standardizing the inputs required to estimate the index coefficients across sites and series.

Varying number of sites

The number of sites within a series varied from three to nine. When genetic testing of Sitka spruce plus trees first started, open pollinated progeny from each plus tree were planted on up to nine sites across Britain (1967–71). Following an

investigation into genotype by site interaction (Johnstone and Samuel, 1978) this was reduced to just three for all tests planted after the mid-1970s. Typically, the three sites chosen had to be fertile, uniform sites to which QCI origin Sitka spruce would be considered well adapted within the areas of northern Scotland, southern Scotland or northern England, and Wales. A further complication was that not all families were planted at every available site within a given year due to plant shortages in the nursery. Although in the latter half of this testing phase an operational decision was made that progeny from a plus tree had to be planted on at least two sites, it did mean that progeny from some plus trees were planted on nine sites, whilst progeny from others were planted on just one.

If different families are represented in differing numbers of tests across years and are then ranked by the across-site family mean performance, then those families with the highest (and lowest) rankings tend to be those tested on the smallest number of sites (White and Hodge, 1989). This problem occurs because of the increasing variance of grand family means with decreasing numbers of tests, and the subsequent effects of different variance structures on selection. This would tend to put those plus trees represented by progeny in series which consist of a large number of sites at a disadvantage.

It was decided to restrict all series to a maximum of three sites. Over 50 per cent of the families were already only present at a maximum of three sites. If data were available from more than three sites then the three were selected which best represented the areas of highest economic importance for Sitka spruce; as much as possible these tended to be the three areas previously mentioned. Families present on just two sites were still included in the analysis to avoid complete loss of genetic information for certain plus trees although care was taken in reselection of such trees for the breeding population due to the problems identified earlier; progeny representing a plus trees on just one site were discarded from further analysis.

Varying quality of sites

In the calculation of breeding values for earlier, provisional breeding populations, each progeny

test site within a series had been given equal weighting in the subsequent across-sites analysis. Family mean performance was estimated as the average performance across all the sites within a series on which progeny from a plus tree had been planted. No weighting was given to a site according to whether it was a good progeny test site (fast growth, few environmental influences, good site preparation and layout) or poor progeny test site (poor or varied growth, poor site preparation or layout).

White and Hodge (1989) explain how across-site family means with equal weight to each site are most affected by individual site means from those tests with larger variances. This leads to the selection of a greater proportion of clones from sites with greater environmental variation – these are often the poorer sites.

It was decided to weight family mean values for a given trait by the family mean heritability for that trait at each site when estimating across-site family mean performance within a series. This is roughly equivalent to the 'best linear predictor' equation adopted by White and Hodge (1989) which serves to weight observations in proportion to the ratio of additive to phenotypic variance. It also relates to the multi-site index selection approach suggested for across-sites analysis by Burdon (1979).

Comparable scores across years

If data across all 19 series were to be brought together such that a superior proportion could be selected, then there has to be a common scale for each trait across series and years which at the same time takes into account the relative superiorities of families within each site. For example, if site mean 15-year diameter at one site is 14 cm and the best family is 20 cm, how can this be equated with another site where the site mean value is 10 cm and the best family is 16 cm if the objective is to select a given proportion from the whole population independent of site, series and year? This problem had to be addressed before traits were combined to give an index score, since reselection of a proportion of original plus trees for the breeding population would be based on the multi-trait index value.

An important strength of the Sitka spruce progeny testing programme in Britain was the planting of a standard QCI control in *all* tests. This standard seedlot represents the average genetic quality of unselected QCI origin seedlots planted in Britain and is therefore a suitable yardstick against which to measure family mean performance and estimate parental breeding values. Comparison of family mean performance to the standard control means that despite the variation in series, planting years and site quality it is always possible to assess the genetic quality of plus trees as a percentage deviation to the same control. In this way the control can be considered a 'link' between sites, series and years.

The series mean value of the control will vary, so the mean control value across all series was estimated for each trait, weighted by the number of families within each series. By substituting the overall mean control value within each series and multiplying this by the percentage deviation for each series mean family value, all family values were brought onto the same scale which was comparable across series.

Different variance components at different sites

Estimates of variance components, heritabilities and genetic and phenotypic correlations are all dependent on site quality and the size and degree of selection involved for the trees under test (Falconer, 1981). Heritabilities are site specific yet the material selected based on data collected from those sites will be planted on many different sites all over Britain. In order to extrapolate the mean progeny performance to a 'mean' site representing the average performance of Sitka spruce in Britain, there is a need to derive standard variance component estimates across all sites.

Standard values for family mean heritability, and genetic and phenotypic correlations were estimated by finding the mean of these functions across all the different progeny tests series. These standard values were then used in the estimation of national breeding and index values. The concept of using standard parameters in index calculation has previously been recommended by Cotterill and Dean (1990) and Woolaston *et al.* (1990) to better reflect the overall inheritance patterns for a species across a wide range of sites.

Analytical methods

Restrictions of survival and minimum standards for representation of progeny within a series (outlined above) reduced the number of parent trees eligible for multi-trait evaluation to just under 800. All tests were a randomized complete block design. The mean number of progeny within each series was 52 (range 23–134); plot size varied from 4 tree line plots to 7×7 (49) tree block plots (mean = 11) and the number of replications from 2 to 6 (mean = 4).

Diameter, wood density and stem-form data for each site had been analysed previously according to the following linear model:

$$Y_{ijk} = \mu + R_i + f_j + fr_{ij} + W_{ijk} \qquad (2)$$

where
Y_{ijk} = individual tree observation;
μ = a general mean;
R_j = fixed effect of replicate j; $j = 1,2 \ldots r$;
f_k = random effect of family k; $k = 1,2 \ldots f$; Var = σ_f^2;
fr_{ij} = random interaction between family k and replicate j; Var = σ_{fr}^2;
W_{ijk} = random error of tree l in plot ijk; $l = 1,2 \ldots n$; Var = σ_e^2;
$Var = \sigma^2$ = variance.
Also family mean heritability (h_f^2) has been estimated for each trait at each site according to:

$$h_f^2 = \frac{\sigma_f^2}{\sigma_f^2 + \dfrac{\sigma_{fr}^2}{r} + \dfrac{\sigma_e^2}{rfn}} \quad \text{(Wright, 1976)} \quad (3)$$

The linear model used to analyse data across test sites for each progeny series was:

$$Y_{ijkl} = \mu + E_i + R_{ij} + f_k + fs_{ik} + fr_{ijk} + W_{ijkl} \quad (4)$$

where
Y_{ijkl} = individual tree observation (equation 4);
μ = a general mean;
E_i = fixed effect of test environment i; $i = 1,2 \ldots s$;
R_{ij} = fixed effect of replicate j in test i;
f_k = random effect of family k; Var (f_k) = σ_f^2;
fs_{ik} = random interaction of family k and test i; Var (fs_{ik}) = σ_{fs}^2;
fr_{ijk} = random interaction of family k and replicate j in test i; Var (fr_{ijk}) = σ_{fr}^2;
W_{ijkl} = random error of tree l in plot i; Var (W_{ijkl}) = σ_e^2.
Family mean heritability (h_f^2) for each trait within each series was estimated as follows:

$$h_f^2 = \frac{\sigma_f^2}{\dfrac{\sigma_e^2}{nrs} + \dfrac{\sigma_{fr}^2}{rs} + \dfrac{\sigma_{fs}^2}{s}} \quad \text{(Wright, 1976)} \quad (5)$$

Phenotypic and genetic correlations for each trait within each site were calculated as follows (Falconer, 1981):
Phenotypic correlation:

$$r_{P1,2} = \frac{COV_{P1P2}}{\sigma_{P1}\,\sigma_{P2}} \qquad (6)$$

Genetic correlation:

$$r_{A1,2} = \frac{COV_{A1A2}}{\sigma_{A1}\,\sigma_{A2}} \qquad (7)$$

where
COV = covariance of trait 1 and trait 2;
σ_{P1}, σ_{P2} = phenotypic variance of trait 1 and trait 2;
σ_{A1}, σ_{A2} = additive variance of trait 1 and trait 2.
A two-way table constructed in GENSTAT (Payne, 1993) according to the above linear model allowed for the weighting of each site by the number of blocks and family mean heritability at each site. All functions of variance components (h_f^2, r_P, r_A) were averaged across series to give the standard values for heritability, and genetic and phenotypic correlations used in the estimation of index coefficients.

The across-sites model also generated series mean family values which were converted to percentage deviation relative to the series value for the QCI control. This exercise was repeated for all three economic traits within each series.

The mean performance of *all* the QCI controls planted in each series was calculated for each economic traits. Within-series family mean values were then recalculated for each trait by multiplying the percentage differential relative to the control by the overall mean control performance. Adjusted values for plus trees were then combined into one large data set for each economic trait. These three data sets were then merged by family identity.

The Restricted Selection Index (RESI4) software package compiled by Jackson, Cotterill and Dean (Cotterill and Dean, 1990) was used to obtain the necessary index coefficients using the standard parameter data. The only unknown variable required for the RESI4 package was relative economic weights to be placed on diameter, wood density and stem form.

Economic weights

Economic weights are defined as the additional profit that may be expected from one unit increase in one trait (e.g. 1 cm increase in d.b.h.) relative to one unit increase in another (e.g. 1 mm less penetration of the Pilodyn pin). The Sitka spruce breeding programme is not yet sophisticated enough to give these relative data. In the absence of such definitive data, a form of arbitrary weighting must be employed. Since a larger diameter value is perceived to be an improvement, the weighting on this trait was positive in sign and was fixed at +1. Conversely a lower Pilodyn score indicates higher density (which is an improvement) and so it was envisaged that a negative weighting would be appropriate. As this was the first time the relative importance of diameter and density had been investigated for Sitka spruce in Britain the weighting on density was varied from –10 through 0 to +10 in steps of 1. This was a very arbitrary way of comparing diameter and density but at least it would give a rough indication of the relative importance of the two traits.

The objective ultimately was to obtain the correct relative weightings which achieved the breeding objectives when all other variables were fixed.

The result was 21 sets of index coefficients. Index values were calculated by substituting index coefficients in equation 1. Plus trees were then ranked based on index values and genetic gains estimated for each trait following selection of the top 10 per cent according to:

$$\text{Genetic gain} = 2 \frac{(S \; x \; h_f^2)}{(\overline{X}_{QCI})} \; x \; 100\% \qquad \text{(Wright, 1976) \quad (8)}$$

where
S = Selection differential which is calculated as the difference between the mean of selected population and the QCI overall mean;
h_f^2 = Family heritability;
$\overline{X}_{QCI}$ = Overall performance of the QCI control across all series.

The factor '2' was introduced as the family mean values represent only half the estimated breeding values.

Investigation of the relative gains from these first 21 pairs of index values highlighted the likely range of values which should result in acceptable

gains for diameter without appreciable loss in wood density. A further nine sets of economic values were investigated, including stem-form in addition to diameter and wood density, with a Kempthorne restriction applied to wood density.

Results

The standard parameter values for family heritability, phenotypic variances and phenotypic and genetic correlations used in the generation of index coefficients are given in Table 1 (also presented as Table 3 in Lee, 1999). The overall control values for diameter, wood density and stem straightness calculated as an average of each series, weighted by the number of families within each series, are given in Table 2.

Figure 1 gives the percentage gains for diameter and wood density relative to the control following selection of the top 10 per cent of index-ranked families for each of the 21 pairs of economic values investigated using the index selection software. Major changes of genetic gain result from altering the economic weight for wood density from –4 to +4 when the economic weighting for diameter is fixed at 1. Positive gains for both diameter and wood density result when the economic weighting for wood density varies between +0.5 and +2.0. Diameter gains are maximized without a fall in wood density when respective economic weightings are +1.0 and +0.5.

When the Kempthorne restriction was applied to prevent a fall in wood density and stem-form was introduced as an additional trait, the maximum gain achieved for diameter without any appreciable fall in wood density was achieved at an economic weighting of +0.1 for density and +1.0 for stem form. This combination of economic weights was referred to as the 'Preferred Index'. Predicted genetic gains for each of the traits under selection for both the breeding (top 240 plus trees) and production population (top 40 plus trees) when ranking is based on index values derived from the preferred index, are given in Table 3. Predicted genetic gains of alternative production populations based on reselection within the same breeding population are given in Lee (1999).

Table 4 gives estimated BVs expressed as a

Table 1: Standard values for heritabilities, phenotypic variances, and phenotypic and genetic correlations used in estimating the index coefficients

	Diameter	Density	Stem-form
Family heritabilities:	0.70	0.80	0.70
Phenotypic variances:	0.7744	1.2321	0.0961
Phenotypic correlations:			
Diameter	1.0000		
Density	−0.5800	1.0000	
Stem-form	0.1100	0.0400	1.0000
Genetic correlations:			
Diameter	1.0000		
Density	−0.6600	1.0000	
Stem-form	0.0400	0.0000	1.0000
Phenotypic standard deviations:	0.8800	1.1100	0.3100
Genetic standard deviations:	0.7363	0.9928	0.2594
Phenotypic co-variances:	0.7744		
	0.5665	1.2321	
	0.0300	0.0138	0.0961
Genetic co-variances:		0.5421	
	0.4824	0.9857	
	0.0076	0.0000	0.0673

Table 2: Overall control values for diameter, wood density and stem straightness

	Diameter (cm)	Density (mm)	Stem straightness (score)
Overall control value	11.84	14.59	2.63

Density was measured as penetration of the Pilodyn pin (mm). Stem straightness was measured on a subjective 1–6 score where 1 = best to 6 = worst (Lee, 1992). Measurements were carried out over a period of 15–22 years from planting.

percentage deviation from the overall QCI control, of each trait for the top 40 parent plus trees reselected for the Sitka spruce breeding population. It shows the wide variation of individual BVs within the population which may well compose a future production population although the mean breeding value of the population satisfies the original selection criteria.

Discussion

The techniques described here outline the relatively simply methods employed to tackle the problem of 'messy data' (Milliken and Johnson, 1984) within the first generation of genetic testing of the Sitka spruce breeding programme. The analysis carried out in this study to derive a multitrait Sitka spruce breeding population adjusted for series and sites differences was carried out in 1993 before the use of mixed model analyses (MMA) became more routine in tree breeding. Today, all of the adjustments described above could be carried out using restricted maximum likelihood (REML; Patterson and Thompson, 1971) to estimate variance components, and best linear unbiased predicator (BLUP; Henderson, 1949) to estimate BVs. In both cases blocks, series and years would be considered fixed whilst families would be random (White and Hodge, 1988). These novel statistical techniques were developed by cattle breeders who were faced with very 'messy data' involving progeny raised on many different farm environments spanning a great number of years, and even generations (Mrode, 1996).
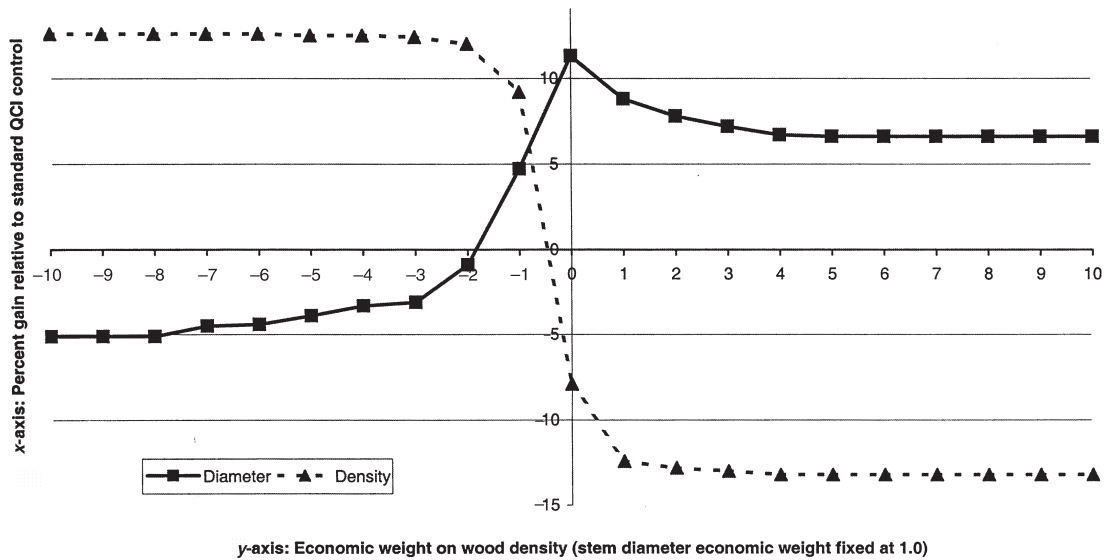
Figure 1. Plot of expected percentage gains for wood density and stem diameter with changing economic weights for wood density.

Table 3: Percentage genetic gains relative to the overall unimproved QCI control associated with the 'Preferred Index'

	Diameter	Density*	Stem-form
Production population	22.0	−0.5	16.3
Breeding population	13.0	−2.0	8.2
Economic weight	1.0	−0.1	1.0

* Kempthorne Restriction applied to wood density.

A major component of BLUP estimates of breeding values involves weighting the observed data in proportion to the ratio of additive to phenotypic variance (White and Hodge, 1989). So, for example, with families observed to have high phenotypic variances relative to the amount of additive genetic variances (such as families from imprecise tests), the predicted breeding values will be regressed more towards zero than would families from more precise sites, with a higher additive to phenotypic variances ratio. The effect is to weight the family mean values by the trait family mean heritability which is the same as the procedure adopted in this study. It is considered therefore, that the simple techniques employed to derive breeding values here effectively simulate the principles employed using REML and BLUP.

No standard errors have been attached to any of the estimated breeding values, which is a restriction of the simple techniques employed here relative to standard BLUP software (e.g. ASREML; Gilmour *et al.*, 1999). There is scope for error in a number of different areas effecting both BVs and variance components. It is inevitable that by the time BVs and variance components are brought together to calculate an index score for each family, errors will have been compounded which in turn could lead to an over-estimate of genetic gain. Another technique may well arrive at a different ranking but it is unlikely that errors will be so great that what has been selected for the production population (top 40), should really not even be in the breeding population (top 240).

Table 4: Diameter, density and stem straightness breeding values for the top 40 parents reselected under the 'Preferred Index' for the Sitka spruce breeding population

Parent	Diameter BVs	Density BVs	Straightness BVs	Index value
1	30	−12	44	2.9572
2	28	4	20	2.9003
3	23	19	12	2.8855
4	34	−10	15	2.7190
5	15	15	28	2.5508
6	22	−6	31	2.4669
7	14	12	30	2.4405
8	18	19	3	2.4329
9	35	−18	6	2.4301
10	23	10	0	2.4297
11	22	−3	24	2.4269
12	18	17	4	2.4168
13	22	5	11	2.4114
14	22	−10	32	2.3727
15	28	−8	10	2.3541
16	19	−10	33	2.3528
17	25	0	6	2.3447
18	13	9	27	2.3100
19	18	8	12	2.2963
20	17	21	4	2.2947
21	17	4	20	2.2800
22	24	−22	40	2.2782
23	23	−16	35	2.2685
24	18	10	8	2.2627
25	21	−12	25	2.2605
26	26	−3	10	2.2559
27	20	5	7	2.2514
28	22	−11	25	2.2288
29	28	−13	9	2.2216
30	18	6	10	2.2145
31	25	5	−3	2.2092
32	26	−13	16	2.2034
33	34	−5	−2	2.1963
34	17	2	20	2.1956
35	17	2	20	2.1952
36	17	−1	22	2.1738
37	31	−12	−5	2.1662
38	20	−5	18	2.1636
39	17	5	13	2.1597
40	27	−7	−5	2.1596
Mean	22	−0.5	16.0	

This table presents actual data but parent identification numbers have been changed. BVs are presented as percentage deviations relative to the overall QCI control.

Determining the final numbers of reselected parents which will form a breeding population is often a difficult decision based on a number of variables, many of which cannot be fully quantified. There have to be sufficient parents to allow a number of unrelated selections to go forward in subsequent generations without severely restricting within-generation genetic gain. And yet the number should not be so large that it is 'unmanageable' in terms of time and resources within the

chosen breeding strategy. Furthermore, inbreeding (or co-ancestry) is to be avoided and yet there ought to be a top limit to the degree of genetic diversity within the population beyond which addition unrelated parents merely adds to operational overheads and does not significantly add to genetic gain.

The decision was made to reselect the top 240 parents based on index value to form the General Breeding Population (GBP). It is anticipated that the final number of reselected parents will be 300–360 once additional parents are chosen following similar analysis of progeny tests planted during the period 1977–93. This range of 240–360 parents is similar to other breeding programmes in the world e.g. *Radiata* pine in Australia (White *et al*., 1999).

The products of the General Breeding Population could be planted wherever trees of Queen Charlotte Island origin would be well adapted. Although seed lots from the GBP will perform well, and better than unimproved QCI on sites best suited to more southerly origins (e.g. Washington origin material growing in Wales or SW England), it remains to be seen if they will out-perform the products of the Southern Breeding Population currently under development (Lee, 1997a).

The evaluation described here has enabled breeders to rank trees within a reselected population according to a multi-trait index score which it is perceived most closely optimizes economic return in the forest. Since the population has been composed and ideas of scope for error have become apparent, it has been possible to construct a strategy for the future breeding of Sitka spruce.

Proposed future strategy

Objectives of the second cycle

There are two objectives in the second cycle of breeding and selection for the Sitka spruce breeding programme in Britain:

1 to select outstanding families composed of reselected parents which can be re-created using controlled pollinations and subsequently be included in family mixtures, sold to commerce and multiplied up using established vegetative propagation techniques;

2 to select the best individuals in a selected proportion of the best unrelated families, which will constitute the next breeding population.

Use of sub-lines

The 240 top-ranked plus trees selected to date for the breeding population based on their index value using the index coefficients from the preferred index have each been allotted to one of four sub-lines. Each sub-line therefore contains 60 reselected plus trees. It is envisaged that the reselection of plus trees based on analysis of progeny tests planted after 1977 will lead to the creation of a fifth, and possibly a sixth, sub-line. The mean index value of each sub-line is approximately equal. Division of the breeding population into sub-lines permits an element of co-ancestry in the breeding population, which is avoided at the production population level by selecting unrelated trees from across the different sub-lines (van Buijtenan and Lowe, 1979). Controlled crossing can now take place within each sub-line according to a pre-arranged design.

Amended positive assortative mating

A form of amended positive assortative mating (PAM) will be used to select parents for mating in the next cycle of breeding. Plus trees in each sub-line have been ranked 1 to 60 based on their index values. Those plus trees in the top 30 will be crossed four times (as either a male or a female) with other plus trees in the top 30; those in the bottom 30 will be crossed three times with other plus trees in the bottom 30. Pair combinations are effectively random within each half, although some additional corrective mating is also planned between, for example, plus trees of high vigour but low density, and plus trees of lower vigour but high density.

The advantage of this sort of mating is that by mating the 'best with the best', it more readily meets the objective of increasing the frequency of favourable alleles required under recurrent selection, allows an element of selection between families which can further increase genetic gain (King and Johnson, 1993) and yet enables the possibility of selecting families whose component parents have demonstrated positive specific combining ability (see 'Non-additive genetic variance'

below). PAM works the best when BVs are known with little error attached. This is not thought to be the case for the parents selected here hence only the mild form of assortment (top half; bottom half). The relatively large number of crosses from which to select families helps to compensate for imperfectly estimated breeding values (King and Johnson, 1993).

To avoid a repeat of the first cycle of genetic testing being spread over a long time period, all the crosses made within a given sub-line will be planted out to field tests in the same year. All crossings within sub-line 1 have now been completed and field planting took place in spring 1999. Crossings within sub-lines 2, 3 and 4 continue and it is planned that these will have been completed and established in field tests by 2005.

Non-additive genetic variance

When two parents of known BVs are combined, the mean of the progeny ought to be equal to the mean of the parental breeding values. BVs are therefore additive. Occasionally, two parents may interact in a positive or negative direction to give a family mean performance much greater or much poorer than the prediction based on parental BVs. This is referred to as specific combining ability (SCA). Such interactions cannot be predicted and occur if there is an element of non-additive genetic variation within a population. Samuel (1991) concluded that there could be as much non-additive genetic variation as there is additive operating in Sitka spruce for height and stem straightness up to 13 years from planting, but that diameter is little effected by non-additive genetic variance by 15 years from planting.

Whilst these conclusions were based on just a small population of eight parents and therefore could be considered as uncertain, and studies elsewhere with other species indicate the magnitude of SCA may decrease with age (Balocchi et al., 1993) it remains that there could be establishment advantages to early expression of SCA. The amended PAM design along with other existing field trials will allow further investigation into the magnitude of non-additive relative to additive genetic variance within the selected Sitka spruce population. In the meantime combining the parents with the highest estimated BVs seems the best way forward since Cotterill et al. (1987) favoured this approach even when non-additive variance in the population equals additive variance.

The second cycle of breeding and selection will allow breeders to identify families which behave well above the mean of the parental BVs for height and stem straightness and yet retain good wood density. Such families could then be re-created by controlled pollinations and introduced into family mixtures for vegetative propagation by commerce.

Even in the absence of non-additive genetic variance, there remains the opportunity to deploy a limited number of high value pair-wise crosses of known genetic quality to commerce which could then be planted over large areas following vegetative propagation.

Experimental design

Each family will be represented by 30 trees planted at a standard 2 m × 2 m spacing (2500 ha^{-1}) over each of three homogeneous sites across the country from north Scotland to Wales (90 trees in all). Experimental design within each site will be single tree plots with 30 randomized, complete replications. There will be approximately 100 families plus controls in each test. Mixed model analysis techniques are being developed to enable accurate assessments of family mean performance within and across sites, and by removing site and block effects, to allow selection of the best individuals within each family. Around 50 per cent of the poorest families will be rejected and the best individuals within the remaining crosses will form the next breeding population. Since each parent will have been used in either three or four families the possibility of family selection exists without reducing the genetic diversity although the ultimate number of unrelated selection that go forward to the next generation cannot exceed $n/2$ where n = size of breeding population (240–360).

Assessment ages

The time required to complete the second cycle of breeding and selection will be much reduced relative to the first cycle due mainly to a reduction in the ages at which trees are considered to be reliable indicators of end of rotation performance.

Lee (1997b) found that juvenile wood density is an excellent indicator of mature wood density and can be reliably assessed after 9 years from planting. The optimum age for height assessment at the family level is around 5 years from planting (Gill, 1987; Lee, 1997b).

It is anticipated in future that initial screening for height will be carried out around 5 years from planting. This will allow identification of promising families for vigour. Final individual tree selections will be delayed until after 9-year height and wood density, at which time stem straightness and branching quality will also be assessed. Multi-trait analysis will again be used to select those individuals within the selected families which best combine all the desired traits under selection. The progeny test will then be considered closed.

Bottleneck to faster breeding

The new selections will be grafted into seed orchards where they may inter-mate to produce the next generation of improved seed. Delays may be experienced at this point since Sitka spruce is a late-flowering species and a seedling does not usually flower until around 20 years old (Phillipson, 1983). This represents an unnecessary delay between making final selections in progeny tests and the age of flowering for the selected tree. New research is required to induce flowering in more juvenile individuals and thereby increase the genetic gains per year. This will involve bringing potted grafts into polythene houses and using a combination of drought induction and stem injection of gibberellin (Phillipson, 1983).

Prospects for clonal forestry

Clonal forestry is the practice of establishing forests with tested clones. Clonal forestry is currently not being practised due to technical problems regarding physiological ageing which prevents the production of commercial quantities of tested clones either by tissue culture or cuttings once they have been identified in field tests beyond 6 years from planting (John, 1992). Clonal forestry should not be confused with the bulk propagation of tested families currently being practised by some commercial nurseries in Britain, and advocated in this paper.

Investigative clonal tests show that there is considerable scope for increasing gain if the ageing problems could be overcome (Lee, 1997a). Ways have to be found of retaining juvenility amongst copies of each clone whilst testing takes place, or alternatively reversing the physiological changes which have taken place within tested clones. Cryopreservation in liquid nitrogen of some of the clonal tissue seems the most likely way forward. Research continues into micro-propagation under sterile conditions and somatic embryogenesis which mimics *in vitro* the normal embryo development within the seed (John *et al.*, 1994). Both techniques will be necessary to produce successfully the millions of plants required from the limited amount of plant material kept in storage.

Marker aided selection

Marker aided selection (MAS) promises the possibility of identifying outstanding genotypes for selected traits in the laboratory or nursery at a very early age. The technique involves finding DNA markers which are close to the genes on the chromosomes responsible for controlling the traits of interest to breeders. A good deal of development is still required in this area, but progress is now being made including establishing new field-based experiments designed to determine future correlations between field performance and molecular markers found in the laboratory. If research is successful in this area then field-based genetic testing could become a thing of the past; only occasional verification trials would be required. Also the speed with which selected material is passed onto commerce could be improved as screening would take place in the laboratory, not in lengthy field trials, and these early selections could be clonally multiplied using existing technology.

Conclusion

The problems associated with combining across-site family mean data at a multi-trait level for progeny established over a varied number of sites and series over an 11-year period, have been addressed in this paper. A simple method of calculating the breeding values of original plus trees for each trait relative to the overall performance

of the standard control is outlined. Methods described by Cotterill and Dean (1990) to devise standard genetic parameters as inputs to the index selection software have also been adopted and are recommended as a means of representing genetic parameters for Sitka spruce in Britain.

The next cycle of breeding, testing and selection within the General Breeding Population will involve an amended form of positive assortative mating. Crosses will be carried out within the top and bottom half of four sub-lines each consisting of 60 of the 240 reselected plus trees which constitute the General Breeding Population. The number of sub-lines will increase to five or six following similar analysis of progeny tests established between 1977 and 1993.

The objectives of the next cycle of breeding are to identify families demonstrating positive non-additive genetic variance which can be re-created and introduced in to family mixtures, and the selection of the best individuals in the better families to form the next breeding population.

The age at which final selections are made in future genetic tests will be much reduced relative to the first cycle of breeding due to improved efficiencies in age : age correlations; in particular, wood density can now be screened at 9 years (Lee, 1997b) from planting. All the families associated with a given sub-line will be established in the field in the same year. This is in contrast to the 25 years required to establish the genetic tests associated with the first cycle of selection and breeding.

Acknowledgements

References

Balocchi, C.E., Bridgewater, F.E. and Bryant, R. 1993 Selection efficiency in a non-selected population of loblolly pine. *For. Sci.* **39**, 231–251.

Burdon, R.D. 1979 Generalisation of multi-trait selection indices using information from several sites. *N. Z. J. For. Sci.* **9**, 145–152.

Cotterill, P.P. and Dean, C.A. 1990 *Successful Tree Breeding with Index Selection*. CSIRO, Australia.

Cotterill, P.P., Dean, C.A. and van Wyke, G. 1987 Additive and dominance genetic effects in *Pinus pinaster*,

P. radiata and *P. elliottii* and some implications for breeding strategy. *Silv. Genet.* **36**, 221–232.

Falconer, D.S. 1981 *Introduction to Quantitative Genetics*, 2nd edn. Longman, Essex.

Fletcher, A.M. and Faulkner, R. 1972 *A Plan for the Improvement of Sitka Spruce by Selection and Breeding*. Forestry Commission Research and Development Paper 85. HMSO, London.

Forestry Industry Council of Great Britain 1998 *Handbook 1998*. FIGGB, Stirling, Scotland.

Gill, J.G.S. 1987 Juvenile–mature correlations and trends in genetic variances in Sitka spruce in Britain. *Silv. Genet.* **36**, 189–194.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. 1999 *ASREML*. Anonymous ftp: ftp.res.bbsrc.ac.uk

Hazel, L.N. 1943 The genetic basis for constructing selection indexes. *Genetics* **28**, 476–490.

Henderson, C.R. 1949 Estimation of changes in herd environment [abstract]. *J. Dairy Sci.*, **32**, 709.

Hodge, G.R. and White, T.L. 1992 Concepts of selection and gain prediction. In *Handbook of Quantitative Forest Genetics*. L. Fins, S.T. Friedman and J.V. Brotschol (eds). Kluwer, Dordrecht, pp. 140–194.

John, A. 1992 *Studies on the Rejuvenation of Sitka Spruce*. Forestry Commission Report on Forest Research 1992. HMSO, London.

John, A., Drake, P.M. and Selby, C. 1994 Somatic embryogenesis in Sitka spruce. In *Somatic Embryogenesis in Woody Plants. Vol. 3, Gymnosperms*. S.M. Jain, P.K. Gupta and R.J. Newton (eds). Kluwer, Dordrecht, pp. 125–144.

Johnstone, R.C.B. and Samuel, C.J.A. 1978 The interaction between genotype and site: its influence on tree selection programmes in Britain. *Forestry Commission R&D Paper* **122**, Forestry Commission, Edinburgh.

Kempthorne, O. and Nordskog, A.W. 1959 Restricted selection indices. *Biometrics* **15**, 10–19.

King, J.N. and Johnson, G.R. 1993 Monte Carlo simulation models of breeding population advancement. *Silv. Genet.* 42, 68–78.

Lavers, G.M. 1983 *The Strength Properties of Timber*. Building Research Establishment Report. Department of the Environment, HMSO, London.

Lee, S.J. 1992 Likely increases in volume and revenue from planting genetically improved Sitka spruce. In *Super Sitka for the 1990s*. D.A. Rook (ed)., Forestry Commission Bulletin 103, HMSO, London.

Lee, S.J. 1994 Breeding strategy for Sitka spruce in Britain. In *Proceedings of the Nordic Group for Tree Breeding*. Edinburgh, pp. 95–109.

Lee, S.J. 1997a The genetic improvement of Sitka spruce. In *Report on Forest Research 1997*, Edinburgh, pp. 24–28.

Lee, S.J. 1997b *The genetics of growth and wood*

density in Sitka spruce estimated using mixed model analysis techniques. Ph.D. thesis, University of Edinburgh.

Lee, S.J. 1999 Improving the timber quality of Sitka spruce through selection and breeding. *Forestry* **72**, 123–133.

Lindgren, D. 1986 How should breeders respond to breeding values? In *Proceedings of IUFRO Conference on Tree Breeding Theory, Progeny Testing and Seed Orchards*. October 13–17, Williamsburg, VA.

Mahalovich, M.F. and Bridgewater, F.E. 1989 Modelling elite populations and positive assortative mating in recurrent selection programs for general combining ability. In *20th Southern Forest Tree Improvement Conference*, Charleston, South Carolina, pp. 43–49.

Milliken, G.A. and Johnson, D.E. 1984 *Analysis of Messy Data. I. Designed Experiments*. Lifetime Learning, Belmont CA.

Mrode, R.A. 1996 *Linear Models for the Prediction of Animal Breeding Values*. CAB International, Wallingford.

Patterson, H.D. and Thompson, R. 1971 Recovery of interblock information when block sizes are unequal. *Biometrika* **58**, 545–554.

Payne, R. 1993 *The Genstat 5 Release 3 Reference Manual*. Clarendon Press, Oxford.

Phillipson, J.J. 1983. The role of gibberellin A4/7, heat and drought in the induction of flowering in Sitka spruce. *Journal of Experimental Botany* **34**:291–302.

Samuel, C.J.A. 1991 The estimation of genetic parameters for growth and stem-form over 15 years on a diallel cross of Sitka spruce. *Silv. Genet.* **40**, 67–72.

Smith, F.H. 1936 A discriminant function for plant selection. *Ann. Eugen. (Lond.)* **7**, 240–250.

van Buijtenan, J.P. and Lowe, W.J. 1979 The use of breeding groups in advanced generation breeding. In *Proceedings of the 15th Southern Forest Tree Improvement Conference*, Mississippi State University, pp. 59–65.

White, T.L. and Hodge, G.R. 1988 Best linear prediction of breeding values in a forest tree improvement programme. *Theoret. Appl. Genet.* **76**, 719–727.White, T.L. and Hodge, G.R. 1989 *Predicting Breeding Values with Application in Forest Tree Improvement*. Kluwer Academic, Dordrecht.

White, T.L., Matheson, A.C., Cotterill, P.P., Johnson, R.G., Rout, A.F. and Boomsma, D.B. 1999 A nucleus breeding plan for *radiata* pine in Australia. *Silv. Genet.* **48**, 122–133.

Wood, P.E. 1986 *Variation and inheritance of wood properties of Sitka spruce*. M.Sc. thesis, Oxford Forestry Institute.

Woolaston, R.R., Kanowski, P.J. and Nikles, D.G. 1990 Genetic parameter estimates for *Pinus caribaea* var. Hondurensis in coastal Queensland, Australia. *Silv. Genet.* **39**, 21–28.