GENETIC PARAMETER ESTIMATES FOR HEIGHT AND STEM STRAIGHTNESS IN *PINUS TAEDA* LINNAEUS AND IMPLICATIONS FOR BREEDING

by

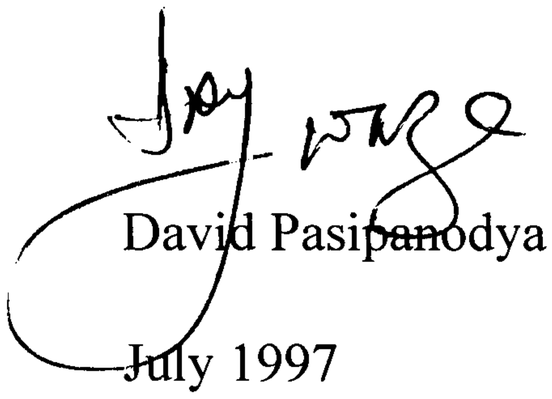
David Pasipanodya Gwaze

Thesis submitted for the degree of Doctor of Philosophy, University of Edinburgh.

1997

Declaration

I hereby declare that this thesis is my own composition and is an account of the analyses performed by me whilst studying for the degree of Doctor of Philosophy at the University of Edinburgh.



David Pasipanodya Gwaze

July 1997

Dedicated in memory of my dad

ABSTRACT

Pinus taeda L., loblolly pine, is an important exotic plantation species in Zimbabwe, and other southern African countries. A major constraint to the efficient breeding of this species in these countries is the lack of genetic parameter estimates to assist informed decisions on the most appropriate breeding strategy and, more generally, to monitor genetic progress. This thesis addresses this deficiency using data from, primarily, the Zimbabwean breeding populations of *P. taeda*. The study focuses on univariate and bivariate estimation of genetic parameters, prediction of optimum selection age, and evaluation of the magnitude and importance of genotype x environment interaction (GE).

Genetic parameters for height and stem straightness were estimated using innovative methods used by animal breeders (individual models), and have recently been used by tree breeders. The data originated from four genetic tests representing 140 full-sib families, assessed at four ages: 1.5, 9.5, 13.5 and 22.5 years. Results suggest height is under moderate to high genetic control, which peaked at 9.5 years, while straightness is under weak genetic control at very young ages, which increases to moderate levels with age. Dominance variance was less than additive variance for both traits, except for straightness at very young ages. Analysing data pooled across sites resulted in biased estimates of heritability at each site for both traits. Age-age genetic correlations for height were high; those for straightness were moderate, apart from correlations involving straightness at very young ages, which were negative. Generally, the genetic correlations between height and straightness were low and positive, which suggests that selecting on height alone will improve straightness.

Annual genetic gain and optimum selection age for height were predicted using the estimated genetic parameters for *P. taeda*. Results suggest that choice of model for predicting trends in age-age genetic correlations is critical for accurate estimation of gain and of optimum selection age. Models based on phenotypic correlations underestimated the annual genetic gain, and needlessly delayed selection. Annual genetic gain was maximised by selection at 10 years; if the species could be induced to flower at 3 years, the annual gain could be maximised by selection at 3 years and increased by 100%, indicating the promise of artificially inducing flowering with *P. taeda* in Zimbabwe.

The implications of GE for breeding strategy was evaluated using genetic correlations, parental rank changes and efficiencies of selection. Results show that GE for both traits was brought about primarily by a change in rankings of genotypes among the locations. Since GE in height was unpredictable, it would be difficult to use GE to advantage in the multiple population breeding strategy for *P. taeda* in Zimbabwe. Evaluating the effect of early selection on one site for predicting mature age performance at another, as suggested in this study, appears to be an efficient approach for GE evaluation in forest trees, since early selection within sites is a common practice in tree breeding programmes. A Bayesian approach, Gibbs sampling, was used to estimate efficiencies of selection across sites and their variances, and the probability that the efficiency of selection lies between certain values. This is the first use of Gibbs sampling for making decisions about optimal selection environment, and results show that it is an efficient approach. Using this approach progeny tests for *P. taeda* in Zimbabwe should be located at site C. Since one should expect GE within a region, it is necessary to consider dispersing at least three progeny tests within the site C region, in order to get results appropriate to commercial progress.

ACKNOWLEDGEMENTS

I wish to thank my supervisors, John Woolliams (*Roslin Institute*), Peter Kanowski (*Australian National University*) and Bill Hill (*University of Edinburgh*) for their support and guidance. The numerous discussions with John greatly improved my understanding of quantitative genetics. Thank you, John, for the patience and encouragement. And thank you Peter for recommending I come to Edinburgh University. Thank you Bill and Peter for the thorough reviews.

I acknowledge those colleagues who have provided both advice and stimulating discussions. In particular, I am indebted to Karin Meyer for providing her DFREML programs and for adapting them to my needs; Ricardo Pong-Wong, for Fortran programming; Frank Shaw, Ron Crump and Neil Cameron for their assistance on computer programmes; Claire Williams, Richard Barnes and Jackie Birks for valuable discussions during the early part of my studies.

The encouragement and friendship provided my fellow PhD students, and staff at Roslin Institute and at the Institute of Cell, Animal and Population Biology, Edinburgh University are gratefully acknowledged.

I wish to express my gratitude to Uncle George for his selfless contribution to my education, without which undertaking these studies would have been unthinkable.

Thanks are due to the Zimbabwe Forestry Commission for releasing me to undertake the studies, for making the data for this study available, and for technical support during my field visits. I am grateful to the Commonwealth Scholarship Commission in the United Kingdom for funding my studies.

Finally, I wish to thank my family for their moral support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS

Gwaze, D. P., Woolliams, J. A. and Kanowski, P. J. (1996). Predicted genetic gain in height in 23-year old progeny tests of *Pinus taeda* L. in Zimbabwe, and inferences for the optimum age of selection. *In* Tree Improvement for Sustainable Tropical Forestry, Dieters, M. J., Matheson, A. C., Nikles, D. G., Harwood, C. E. and Walker, S. M. (*eds.*). pp. 133-137. Proc. QFRI-IUFRO Conf., Caloundra, Queensland, Australia, 27 October-1 November, 1996 (Queensland Forestry Research Institute, Gympie).

Gwaze, D. P., Woolliams, J. A. and Kanowski, P. J. (1997). Genetic parameters for height and stem straightness in *Pinus taeda* L. in Zimbabwe. *Forest Genetics* (submitted).

Gwaze, D. P., Woolliams, J. A. and Kanowski, P. J. (1997). Optimum selection age for height in *Pinus taeda* L. in Zimbabwe. *Silvae Genetica* (submitted).

Gwaze, D. P. and Woolliams, J. A. (1997). Additive genetic covariance function for height in *Pinus taeda*. Poster presentation, Southern Forest Tree Improvement Conference, 9-12 June 1997, Florida, USA.

Gwaze, D. P., Bridgwater, F. E., Byram, T., Williams, C. G. and Woolliams, J. A. (1997). General prediction model for age-age genetic correlations in *Pinus taeda*. Oral paper presentation, Southern Forest Tree Improvement Conference, 9-12 June 1997, Florida, USA.

Gwaze, D. P. and Woolliams, J. A. (1997). Decision-making in progeny test location using Gibbs sampling. Oral paper presentation, Southern Forest Tree Improvement Conference, 9-12 June 1997, Florida, USA.

Chapter 1

INTRODUCTION

1.1 General

The contribution of conifers to world wood production is considerable. In addition to extensive natural stands, especially in the northern hemisphere, 44% of the estimated total of 8.3 million ha of industrial plantations are conifers (1980 figures; FAO 1992). The high growth rates achieved by conifers in exotic environments have resulted in large areas being established for wood production. In Zimbabwe, for example, conifer species comprise 80% of the total plantation area. *Pinus taeda* Linnaeus is one of the major commercial conifer species in the USA, where it occurs naturally, comprising more than 50% of the standing volume of pine (Baker and Langdon 1990). It is also important as an exotic in Zimbabwe and South Africa, comprising 15% of the plantation area in the former. The success of *P. taeda* as an exotic is due to its fast growth rate and wide adaptability, although it is limited on some sites by problems such as susceptibility to drought and damage from pests such as *Pinues pine, Cinara cronartii* and baboons. With the increasing demand for wood products globally (Sharma *et al.* 1992), maximising wood production on available land resources is of major importance. The high growth rate of *P. taeda,* the variation evident in natural and exotic stands, and the need to increase production per unit area led to the establishment of breeding programmes in the USA, Zimbabwe and South Africa. The major constraint to the efficient breeding of this species in Zimbabwe and South Africa is the lack of genetic parameter information to guide decisions on the most appropriate breeding strategy and, more generally, to monitor genetic progress. This thesis addresses this deficiency using data from, primarily, the Zimbabwean breeding populations of *P. taeda*.

1.2 Taxonomy of *P. taeda*

The taxonomy of *P. taeda* can be summarized as:

Family: Pinaceae

Genus: Pinus

Sub-section: Australes

Botanic name : *Pinus taeda* Linnaeus

Most used common name: Loblolly pine

Synonyms: *Pinus lutea*, *Pinus heterophyll*

The species is placed in sub-section Australes with *P. palustris, P. echinata, P. glabra, P. rigida, P. serotina, P. pungens, P. elliottii, P. caribaea, P. occidentalis* and *P. cubensis* (Vidakovic 1991). Eight of these species occur in southern parts of the USA and three in Central America. The sub-section is characterised by species with mostly 2-3 needle fascicles and spring shoots with mostly two or more branch whorls. Cones are symmetrical and open when ripe (Vidakovic 1991).

The species hybridizes naturally with *P. palustris* to produce a hybrid known *as P. sondereggeri* H. H. Chapman (Dorman 1976, Fowells 1965, Mirov 1967), and with *P. serotina, P. echinata, P. rigida* and *P. elliottii* (Dorman 1976).

1.3 Natural distribution of *P. taeda*

P. taeda has a wide natural distribution throughout the south and south eastern United States of America. It extends through fourteen states from Delaware to Texas, but does not grow naturally in the lower gulf coastal plain in Florida and in the Mississippi river flood plain (Figure 1.1). Isolated populations are found in North Carolina, Arkansas, Louisiana and Texas. This range corresponds to between 28^0N to 39^0 21'N longitude and 75^0W to 97^0 30'W latitude (Critchfield and Little 1966, Fowells 1965). Its altitudinal distribution ranges from near sea level to 250 m, and rarely to 600m (Vidakovic 1991).

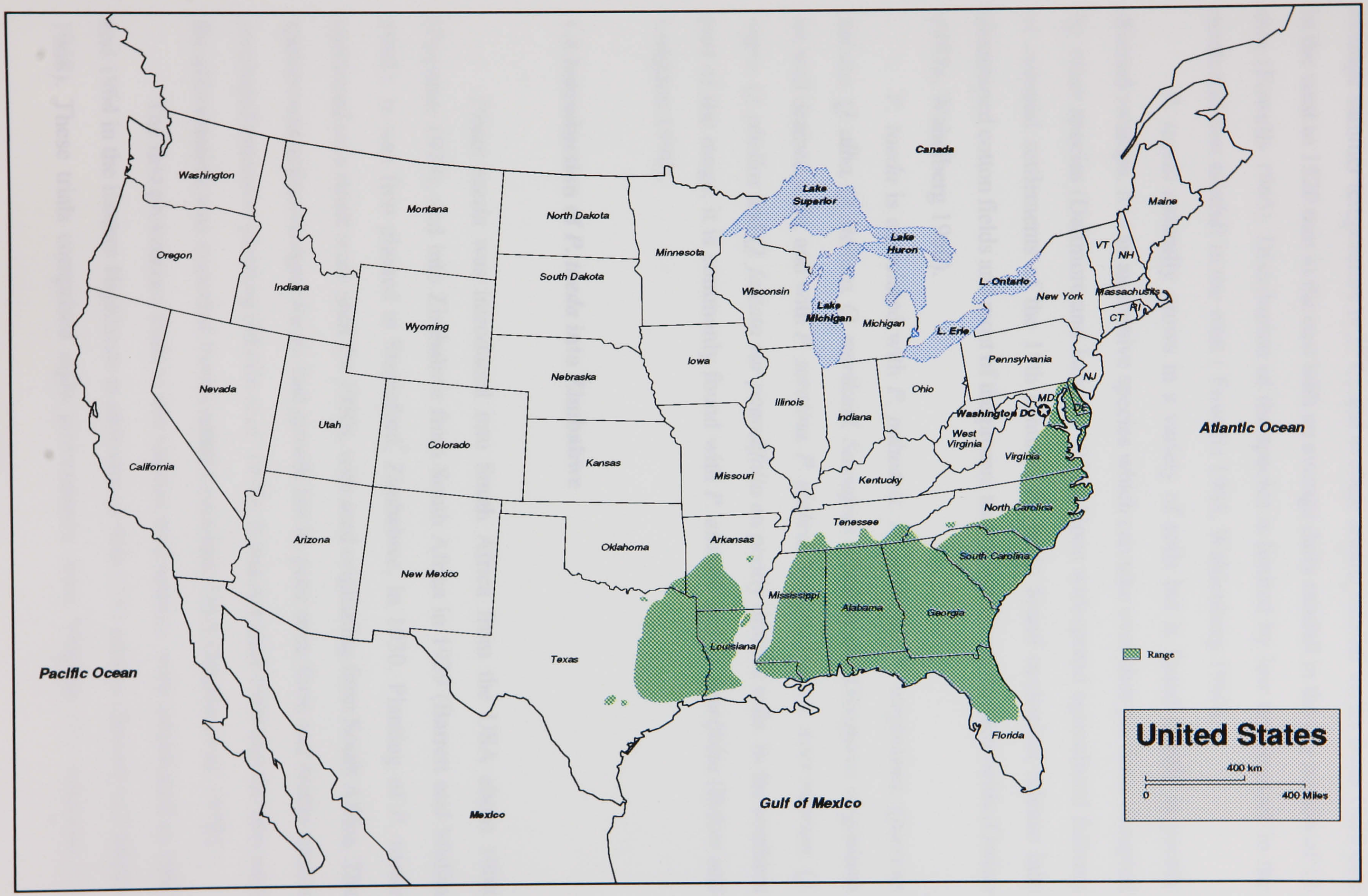


Figure 1.1 Natural distribution of *Pinus taeda* (after Critchfield & Little, 1966)

The climate over most of its range is humid warm temperate, with long hot summers and mild winters. Winter temperatures vary within the range 4-16^{0}C, the average summer temperature is 27^{0}C, and average annual rainfall varies from 1020 mm in the west to 1520 mm in the east with an average daily rainfall in the dry season of 13 mm (Fowells 1965). Distribution of the species is limited by low temperature in the north and low rainfall in the west (Fowells 1965, Wahlenberg 1960).

P. taeda naturally grows in a variety of soils but is found mostly in poorly drained swamps. It is an aggressive species which can take over sites previous occupied by other species (Dalimore and Jackson 1954). Upon widespread agricultural failures of colonial settlements of the 19th century, *P. taeda* found ecological release into abandoned cotton fields and most of the present natural stands are on such fields (Owino 1977a, Wahelberg 1960).

P. taeda is associated with *P. palustris, P. echinata, P. virginiana, Quercus falcata, Q. alba, Q. stellata, Q. mariland, Sassafras albidum* and *Diospyros virginiana* on well drained soils, and with *P. serotina, P. glabra, Nyssa sylvatica, Acer rubrum, Q. nigra, Q. phellos* and *Q. falcata var pagodifolia* on poorly drained soils. In the southern part of the range, it is commonly found with *P. elliottii* and *Q. laurifolia* (Baker and Langdon 1990).

1.4 Introduction of *P. taeda* into Zimbabwe

Pinus taeda was introduced into South Africa from the USA about 1900 (Poynton 1979), and into Zimbabwe from South Africa in 1929 (Barrett and Mullin 1968). It was first planted at Stapleford, Zimbabwe, in 1930. Planting of *P. taeda* continued on a small scale until the 1950s, with seed originating from South Africa. The species was acknowledged for its fast growth but its poor stem form and brittle timber precluded extensive planting (Mullin *et al.* 1978). It was not until 1960 that the first seed for afforestation was imported from central Louisiana, USA (Mullin *et al.* 1978).

The first provenance trials to test various seed sources were established in 1963 and 1964 in the Eastern Highlands at altitudes of 950-1250 metres (Barrett and Mullin 1968). These trials comprised eight provenances, from Virginia (3), Arkansas (1),

Florida (1), Georgia (1), Louisiana (1) and Texas (1). The most comprehensive provenance trials were established in 1965 and 1966 with twenty-two provenances, from Georgia (4), Mississippi (4), Alabama (3), North Carolina (3), Florida (2), Louisiana (2), Arkansas (1), South Carolina (1), Tennessee (1) and Virginia (1). Significant differences in growth rate among provenances became apparent, with Florida and Louisiana provenances superior. It became evident from the early trials that the best provenances of *P. taeda* planted on ideal sites out-performed the commonly grown commercial species, *P. patula*. This realisation, and the initiation of the tree improvement programme, resulted in greater interest in the species by the forest industry.

In Zimbabwe, *P. taeda* is used mainly for structural timber, so breeding for wood quality is important in addition to breeding for growth. Due to a shortage of poplar timber, *P. taeda* is also being used for matchwood. Both growth and quality are important breeding goals for matchwood.

1.5 Genetic improvement of *P. taeda* outside Zimbabwe

P. taeda is a major commercial species in the USA and South Africa and has been planted commercially on a small scale or on experimental basis in Angola, Malawi, Mozambique, Swaziland, Zambia (Poynton 1979), China and Brazil (Haines 1994).

Large-scale genetic improvement of *P. taeda* in the USA started in the 1950's (Williams *et al.* in press). Genetic progress has been achieved in all the main economic traits; extensive genetic studies have been undertaken, and genetic parameter estimates for economic traits are summarised in Chapter 2. However, the extrapolation of genetic parameters from populations in USA to those in Zimbabwe is inappropriate as they apply only to the defined population and test environment (Falconer 1989).This study will provide genetic parameters appropriate to the populations in Zimbabwe.

The tree breeding programme in South Africa was started in 1959 (Poynton 1979): more than 100 plus trees have been selected; progeny tests have been established but results have not been reported; seed orchards have been established, which meet local seed demand.

1.6 Genetic improvement of *P. taeda* in Zimbabwe

The improvement programme of *P. taeda* in Zimbabwe began in 1958 when the first (plus) trees were selected from local stands (Barrett and Mullin 1968) and cloned into seed orchards. The first seed orchard was established at John Meikle research station (1200m asl) and started producing seed for local planting in 1968 (Mullin *et al.* 1978). The first progeny tests were established in 1966. Seed of 31 half-sib families was obtained from USA for breeding purposes in 1966. The programme was aimed primarily at improving stem form, which was very poor. Slow progress in breeding *P. taeda* was due to its poor form in all existing stands, making selection of plus trees difficult. The most comprehensive series of progeny tests in Zimbabwe was established in 1972 on four sites, to test the genetic worth of second generation selections and to provide material for the next generation of breeding. Of the 231 plus trees selected by 1996, 198 were represented in the breeding programme. In 1996 there were 8 seed orchards of *P. taeda* in Zimbabwe.

Analysis and interpretation of data from the second generation trials are still outstanding. Also outstanding is the analysis of data from another major series of progeny tests which were established on five sites in 1976, in Zimbabwe (2), USA (2), and the Republic of South Africa (1). The objective of the later trials was to broaden site coverage for estimating genotype x environment interactions (GE).

Up to 1981, a classical breeding strategy was used in Zimbabwe. Under this strategy, progeny tests were conducted to estimate genetic and environmental parameters, including combining ability for thinning clonal seed orchards, and to provide material for the next generation of selection. It was realised that the programme was becoming too cumbersome and complex to handle, due the large number of species in the breeding programme (9) and the large number of plus trees involved which were more than 1000 (Barnes 1989). It was believed that a breeding programme which addressed the above problems, and also allowed the utilization of GE and minimised the risk of inbreeding, would be more efficient.

In 1981, a multiple population breeding strategy (MPBS) (Namkoong *et al.* 1980) was adopted (Barnes 1981). Under this strategy, a number of populations are kept

separate in order to create genetic differences between the populations through drift and sampling effects. Also, by imposing different selection criteria on the sub-populations, divergence is increased or maintained (Burdon and Namkoong 1983). Variation could be restored at any time by crossing the different populations. Central to implementation of the strategy is the breeding seed orchard (BSO), which combines seed production, progeny testing and forward selection (Barnes 1986). In Zimbabwean BSOs, trees are planted at close spacing (1.0 x 1.0m) in order to increase the selection intensity, and thinned progressively by 50% from two years of age, until at six years of age only one tree remains per plot. From this stage, the trial is managed as a seed orchard. Generally, the multiple population strategy simplifies long term pedigree control, as control can be at the sub-population level instead of the individual tree level, and facilitates the use of GE. The first *P. taeda* BSO was established in 1986 and in 1996 there are 11 BSOs of the species.

The Zimbabwean breeding programme was reviewed in 1994 (Arnold and White 1994). The review highlighted some problems associated with the MPBS, which included the high costs of maintaining a large number of sub-populations; because each sub-population was planted at only one site, it was impossible to assess which families performed the best on different sites and because the populations did not have any families in common, it was difficult to rank families across sub-populations to find the very best families to use for commercial seed production. The review recommended modifying the MPBS to a composite breeding seedling orchard (CBSO) strategy. In the CBSO, the breeding population for each species is not sub-divided into sub-populations. The entire population is managed as one large population and is replicated. The CBSO may reduce costs because the total area under the tests could be reduced and additional genetic gain may be realised because the best families will be allowed to interbreed in each generation. The very best selections may be crossed to increase genetic gain and cloned into a seed orchard to increase gains in plantations.

1.7 Objectives of the study

The work reported here draws from progeny tests established under the classical breeding strategy. It is focussed on developing information which would be of benefit under any breeding strategy.

Improvement of timber yield and timber quality are important breeding objectives for *P. taeda* in Zimbabwe. Therefore, a knowledge of genetic parameters of growth and quality traits are essential to estimate accurate breeding values, to combine different traits in selection, to appraise the rate and magnitude of improvement by selection, and to determine the optimum age for selection. Generally, genetic parameters are important for determining the consequences of a breeding strategy, and hence facilitating decisions about the most effective breeding strategy (Allard 1960). Traits which are of most relevance to these objectives for the case of *P. taeda* in Zimbabwe are height, diameter, volume, stem straightness and wood density. This study will focus on height and stem straightness.

The specific objectives of this thesis are to:

a) estimate the following genetic parameters for height and straightness:

1. Additive genetic variance;
2. Non-additive genetic variance;
3. Heritability;
4. Age-age genetic and phenotypic correlations;
5. Trait-trait genetic and phenotypic correlations;

b) estimate genetic gain and optimum selection age for height;

c) determine the magnitude of genotype x environment interactions;

d) based on these results, assess the implications for breeding *P. taeda* in Zimbabwe.

The genetic parameters will be estimated using innovative methods used by animal breeders. These methods include those currently not used, and those recently used, by tree breeders. The application of these methods will be developed to address specific issues pertaining to, but not limited to, tree breeding.

Chapter 2

LITERATURE REVIEW:
GENETIC PARAMETERS OF *P. TAEDA*

2.1 Additive and non-additive genetic variance

Genetic variance can be partitioned into additive and non-additive components. The presence of the additive genetic variance contributes to general combining ability of selected parents (Falconer 1989). Non-additive genetic variance can be partitioned into dominance and epistasis variance according to whether the interactions between genes are at the same locus (dominance) or at different loci (epistasis). The presence of non-additive genetic variance contributes to specific combining ability of selected parents (Falconer 1989). Performance of offspring of parents with a trait with high non-additive genetic variance is more difficult to predict and depends on the specific matings conducted.

Response of a single trait to selection upon an index depends on the selection intensity and on the additive genetic contribution to phenotypic variance in the trait. When the selection criteria involves measurements on more than one trait, which is often the case, the change in any one trait will also depend on the strength of the additive genetic and phenotypic covariances as well as the variances of the traits under selection.

There have been few reports of the relative magnitude of additive and non-additive genetic variance for *P. taeda*, or of their changes over time. Additive variance for height or volume appears to increase with age (Balocchi *et al.* 1993, Foster 1986, Franklin 1979, Lambeth *et al.* 1983), with that for height peaking at 20 years (Franklin 1979). Non-additive effects appear strongest at early ages: Balocchi *et al.* (1993) reported dominance variance for height to be 4.4 times greater than additive variance at

age 6 years; the non-additive variance continued to exceed the additive variance up to age 12, after which the latter predominated. This result is consistent with Foster and Bridgwater's (1986) finding that the non-additive variance component was at its greatest at 5 years (Vd/Va=2.5 for height, Vd/Va=1.6 for diameter, Vd/Va=7.8 for volume). These results are consistent with reports for other *Pinus* species (Burdon *et al.* 1992a, Cotterill *et al.* 1987, Dean 1990).

2.2 Heritability

Heritability is an expression of the relative contribution of genetic variation to the total phenotype variation. The parameter narrow sense heritability is used in tree breeding programmes relying on recurrent selection and sexual recombination in seed orchards where additive gene effects are most important; the parameter broad sense heritability, which includes all genetic variation, is appropriate for vegetatively propagated material (van Buijtenen 1992).

Heritability estimates for growth, stem straightness and wood properties of *P. taeda* have been estimated mainly from experiments in the USA, using correlations between half-sibs. Tables 2.1-2.5 summarise previous estimates of individual tree and family heritabilities. Height was more frequently measured than diameter or volume, which may be attributed to the fact that height is easy to measure, is correlated with volume, and is less affected by thinning. Heritability estimates in the forestry literature were sometimes expressed on a family mean basis; these estimates need to be converted to an individual tree basis following the formula by Falconer (1989) for comparison purposes.

The individual tree heritability estimates reported for height ranged from 0.10 to 0.69 and those on a family basis ranged from 0.58 to 0.79 (Table 2.1). The range of individual tree heritability estimates reported for height was higher than that reported for forest tree species (0.1-0.40) by Cornelius (1994). Although estimates reported by Li *et al.* (1991) and Williams and Megraw (1994) were from the same test, they varied, which the latter authors attributed to using less families in their analyses. The results at ages 1 and 2 show that heritability estimates derived from better management practices

or high site qualities (ie. fast growth rates) are low. Results at these ages also indicate that heritability estimates from short-term tests where planting is at very close spacing (eg. 0.6 x 0.6m), are higher than those from conventional tests. Short-term tests accelerate stand development and hence lead to different changes in variances with age compared to conventional tests (Franklin 1989). Therefore, age may not be a good guide when comparing estimates from conventional and short-term tests. Also, height is unlikely to be a better guide, since the magnitude of the heritability estimates is likely to be affected by the phase of development (Franklin 1979). Heritability estimates for height were, in general, higher than those for diameter, and consistent with findings of Cornelius (1994) for forest tree species. The individual tree heritability estimates reported for diameter were low, ranging from 0 to 0.10, while those on family mean basis ranged from 0 to 0.60 (Table 2.2). The heritability estimates reported for volume on an individual tree basis were moderate, ranging from 0.15 to 0.35, and those on a family mean basis were high, ranging from 0.58 to 0.75, with 75% of these being greater than 0.70 (Table 2.3).

Stem straightness is an important trait which influences the quality of poles, sawlog grade and wood recovery (Barnes and Gibson 1986). There is a lack of heritability estimates reported for straightness, probably reflecting the fact that this trait is relatively difficult and costly to measure. The most used methods are based on a relative scale where the score is relative to the trees in a single test site rather than relative to all trees in the genetic tests (e.g. Williams and Lambeth 1989), and on an absolute scale (e.g. Barrett and Mullin 1968). Heritability estimates originating from the relative scale are, generally, higher than those from the absolute scale (Cotterill *et al.* 1987, Raymond and Cotterill 1990). Heritability estimates for straightness of *P. taeda* reported in the literature were all from tests assessed with a relative scale. Individual tree heritability estimates ranged from 0.13 to 0.55, while those on a family basis ranged from 0.71 to 0.80 (Table 2.4). In forest trees, the heritability estimates of growth traits are generally lower than those for stem straightness (Cornelius 1994). This is not apparent from literature on *P. taeda,* at least in part because of the small number of estimates for stem straightness.

Table 2.1. A summary of narrow sense heritability estimates (h^2) reported for height in *P. taeda*.

Age (yr)	No. of families	Height (m)	h^2 (s.e.)	Source
Individual tree heritability estimates				
1	183	-	0.05 (0.04)	Balocchi *et al.* (1993)
1	11	0.63	0.09	Foster (1986)
1	25	0.37	0.12 (0.17)	*Li *et al.* (1991)
1	25	0.32	0.32 (0.12)	"
1	25	0.35	0.10 (0.07)	"
2	183	-	0.05 (0.04)	Balocchi *et al.* (1993)
2	11	1.65	0.15	Foster (1986)
2	25	1.46	0.38 (0.14)	*Li *et al.* (1991)
2	25	0.83	0.55 (0.18)	"
2	25	1.15	0.35 (0.13)	"
2	16	1.46	0.42	*Williams and Megraw (1994)
2	16	0.83	0.69	"
3	183	-	0.04 (0.05)	Balocchi *et al.* (1993)
3	16	3.03	0.33	*Williams and Megraw (1994)
3	16	1.91	0.65	"
3	25	3.03	0.39 (0.13)	*Li *et al.* (1991)
3	25	1.91	0.59 (0.19)	"
3	25	2.46	0.31 (0.13)	"
7	11	9.56	0.09	Foster (1986)
8	183	-	0.07 (0.06)	Balocchi *et al.* (1993)
10	11	13.31	0.28	Foster (1986)
12	16	12.80	0.25	*Williams and Megraw (1994)
13	183	-	0.19 (0.09)	Balocchi *et al.* (1993)
15	11	18.19	0.41	Foster (1986)
16	183	-	0.25 (0.10)	Balocchi *et al.* (1993)
25	183	-	0.18 (0.07)	"
Family mean heritability estimates				
1	11	-	0.58	Foster (1986)
15	11	-	0.79	"

*The heritability estimates by *Li et al.* (1992) and Williams and Megraw (1994) were from four different treatments at a single location. This was a short-term test planted at close spacing.

Table 2.2. A summary of narrow sense individual tree heritability estimates (h_I^2) and family mean heritability estimates (h_F^2) reported for diameter in *P. taeda* (Foster 1986), for a population of 11 families.

Age (years)	h_I^2 (s.e.)	h_F^2 (s.e.)
3	0.10	0.54
4	0.09	0.60
5	0.05	0.41
6	0.04	0.39
7	0.02	0.23
8	0.00	0.00
10	0.03	0.21
15	0.04	0.28

Table 2.3. A summary of narrow sense heritability estimates (h^2) reported for volume in *P. taeda*.

Age (years)	No. of families	h^2 (s.e.)	Source
Individual tree heritability estimates			
11	-	0.30	*Bridgewater and Stonecypher
12	-	0.29	(1979)
13	-	0.35	"
14	-	0.17	"
12	-	0.23	"
13	-	0.17	"
14	-	0.15	"
Family mean heritability estimates			
3	11	0.58	Foster (1986)
4	"	0.70	"
5	"	0.73	"
6	"	0.73	"
7	"	0.69	"
8	"	0.71	"
10	"	0.75	"
15	"	0.75	"

*The estimates were from two sites; number of families was not stated.

Table 2.4. A summary of narrow sense heritability estimates (h^2) reported for straightness in *P. taeda*.

h^2 (s.e.)	age (years)	Families	Source
Individual tree heritability estimates			
0.13	8	53	Williams and Lambeth (1989)
0.24	8	59	"
0.55	11	-	*Bridgwater and Stonecypher
0.40	12	-	(1979)
0.29	13	-	"
0.24	14	-	"
0.35	11	-	"
0.27	12	-	"
0.26	13	-	"
0.25	14	-	"
Family mean heritability estimates			
0.80	8	53	Williams and Lambeth (1989)
0.71	8	59	"

*The estimates were from two sites; number of families was not stated.

The most widely measured trait related to wood properties is wood density or specific gravity, because it is well correlated with major strength properties of sawn timber and with pulp and paper properties (van Buijtenen 1969). Heritability estimates

for wood density in *P. taeda* were high (Table 2.5), ranging from 0.42 to 1.00 on an individual tree basis, and from 0.44 to 0.80 on a family basis. These estimates are within the range reported for forest trees (0.3-1.0) (Cornelius 1994), and were higher than those for growth traits. Loo *et al.*'s (1984) results were consistently high compared to results from other sources. Other than their results, a heritability estimate for wood density of around 0.5 on an individual tree basis might be considered typical for *P. taeda*. Loo *et al.*'s (1984) individual tree heritability estimate at 22 years of age was reported as greater than 1, and this result was reported as 1 (Table 2.5) since the additive genetic variance cannot be greater than the phenotypic variance. However, care should be taken in interpreting the constrained estimates.

Table 2.5. A summary of narrow sense heritability estimates (h^2) reported for wood density in *P. taeda*.

h^2 (s.e.)	age (years)	Families	Source
Individual tree heritability estimates			
0.77 (0.30)	2	15	Loo *et al.* (1984)
0.55	2	16	#Williams and Megraw (1994)
0.51	2	16	"
0.79	3	16	"
0.53	3	16	"
0.82 (0.35)	4	15	Loo *et al.* (1984)
0.85 (0.25)	6	15	"
0.89 (0.35)	8	15	"
0.87 (0.35)	10	15	"
0.42	13	16	#Williams and Megraw (1994)
1.00* (0.37)	22	15	Loo *et al.* (1984)
0.44 (0.14)	12	18	Jett *et al.* (1991)
Family mean heritability estimates			
0.80 (0.12)	2	15	Loo *et al.* (1984)
0.74 (0.14)	4	15	"
0.76 (0.17)	6	15	"
0.76 (0.17)	8	15	"
0.76 (0.17)	10	15	"
0.44 (0.16)	20	15	Talbert *et al.* (1983)
0.45 (015)	20	15	"
0.80 (0.14)	22	15	Loo *et al.* (1984)

*estimate >1.
#Results from short-term tests.

2.3 Trait-trait genetic correlations

The genetic correlation between two traits reflects the number of genes that influence both traits, and also the distribution of relative strength of effects of the genes (Falconer 1989). A high positive genetic correlation between two traits means that if selection made on one trait, it will lead to change in the other. Also, a highly heritable trait that is strongly correlated with a poorly heritable but economically important trait can be used as a criterion for indirect selection to maximise gain. For this reason, the estimation of trait-trait genetic correlations can be of considerable importance in breeding programmes.

The literature on *P. taeda* reveals that growth traits (height, diameter and volume) are positively correlated, with the genetic correlations ranging from 0.25 to 0.70 (Foster 1986); growth and wood density can either be negatively correlated, with the genetic correlations ranging from -0.39 to -0.46 (Loo *et al.* 1984), or positively correlated (0.26 and 0.50) (Williams and Megraw 1994). Both positive and negative genetic correlations between growth traits and wood density are common in forest tree species (Zobel and Talbert 1984). The genetic correlations between growth traits and wood density depend on the stage of development: as trees get older, the negative correlations diminish (Zobel *et al.* 1969). Both negative genetic correlations between density and diameter or volume (Burdon and Low 1992, Dean 1990, Magnussen and Keith 1990, Vargas-Hernandez and Adams 1991), and positive genetic correlations between density and height, have been reported in other conifers (Burdon and Low 1992, Dean 1990, Magnussen and Keith 1990).

2.4 Age-age correlations

The rotation age for many forest species is long. Breeders cannot afford to wait until maturity (e.g. 25 years for a number of tropical pines) to select the best trees/families. Therefore, the juvenile-mature (age-age) correlation for quantitative traits plays an important role as an indicator of opportunities for early selection (Burdon 1989, Lambeth 1980). Early selection results in higher rate of genetic improvement due to

shortening of the breeding cycle (Burdon 1989, Eldridge *et al.* 1993, Lambeth 1980). However, the breeding cycle can only be shortened, in most cases, at the expense of gain per generation, since performance at young ages is not precisely related to that at maturity (Lambeth 1980).

Tables 2.6-2.9 summarize age-age correlations reported for growth traits and wood density of *P. taeda*. Genetic correlations reported for height at young ages (less than 4 years) tended to be lower (range, 0.07-0.33) than those between later ages (range, 0.61-1.00) (Table 2.6). The result is consistent with those reported for other pine species (Lambeth 1980), suggesting that traits at very young ages may not be good indicators of later performance. This may be attributed to the different environments at different ages, perhaps to differential expressions of the trait over time, and possibly to maternal effects which later diminish. It might be expected that competition in the short-term tests might inflate the genetic correlations, since changes in family ranks, which otherwise might have been present, may be suppressed. However, this was not evident from literature on *P. taeda.* In some cases (e.g. McKeand 1988), family mean correlations were used as approximations of genetic correlations. These genetic correlations are likely to be biased since error covariance and error variances are unlikely to be zero. There were no estimates of age-age correlations reported for diameter at young ages, but the same trend as that for height could be expected. The age-age correlations reported for diameter ranged from 0.82 to 0.99 (Table 2.7). Genetic correlations reported for volume were all high, ranging from 0.76 to 0.98, and correlations for young ages were also not available for this trait (Table 2.8). The age-age correlations reported for wood density were high for all ages (range, 0.76-1.00) (Table 2.9), suggesting that selection for wood density could be made as early as 2 years after planting. Generally, the correlations for wood density were higher than those for height and diameter.

Opportunity also exists for selections to be carried out at the nursery or green house stage, further reducing the generation interval. For example, Robinson *et al.* (1984) found the genetic correlations between nursery height of *P. taeda* and 5-year measurements of each of height, diameter and volume were 0.63 in all cases.

Table 2.6. A summary of age-age genetic correlations (r_g) reported for height in *P. taeda.*

Younger age (yrs)	Older age (years)	Height (m) younger age	Height (m) older age	r_g	Source
2	15	1.7	18.2	0.25	Foster (1986)
2	12	1.5	12.8	0.09	#Williams and Megraw (1994)
2	12	0.8	12.8	0.22	"
3	12	3.0	12.8	0.07	"
3	12	1.9	12.8	0.33	"
3	25	-	-	0.17	Franklin (1979)
4	15	4.5	18.2	0.66	Foster (1986)
4	12	-	-	0.74	McKeand (1988)
4	16	-	-	0.61	"
5	25	-	-	0.34	Franklin (1979)
5	20	4.7	17.6	1.00*	Lambeth *et al.* (1983)
5	20	4.2	17.6	1.06	"
5	20	4.0	16.4	0.89	"
5	20	4.7	17.7	0.68	"
5	20	4.3	17.6	0.79	"
6	15	7.7	18.2	0.71	Foster (1986)
7	25	-	-	0.40	Franklin (1979)
8	12	6.8	-	0.89	McKeand (1988)
8	16	6.8	-	0.83	"
8	15	11.2	18.2	0.85	Foster (1986)
10	15	13.3	18.2	0.96	"
10	25	-	-	0.47	Franklin (1979)
10	20	-	17.6	0.98	Lambeth *et al.* (1983)
10	20	-	17.6	0.84	"
10	20	-	16.4	1.00*	"
10	20	-	17.7	0.99	"
10	20	-	17.6	0.99	"
12	16	-	-	0.92	McKeand (1988)
15	20	-	17.6	1.00*	Lambeth *et al.* (1983)
15	20	-	17.6	0.94	"
15	20	-	16.4	0.97	"
15	20	-	17.7	0.94	"
15	20	-	17.6	0.96	"
15	25	-	-	0.88	Franklin (1979)
20	25	-	-	0.87	"

*estimates >1.

#Results from short-term tests.

Table 2.7. A summary of age-age genetic correlations (r_g) reported for diameter in *P. taeda.*

Juvenile age (years)	Mature age (years)	r_g	Source
4.2	9.3	0.87	Hagedorn (1994)
4.2	12.5	0.82	"
4.2	15	0.78	"
9.3	12.5	0.96	"
9.3	15	0.96	"
12.5	15	0.99	"

Table 2.8. A summary of age-age genetic correlations (r_g) reported for volume in *P. taeda.*

Juvenile age (years)	Mature age (years)	r_g	Source
4.2	9.3	0.94	Hagedorn (1994)
4.2	12.5	0.83	"
4.2	15	0.76	"
9.3	12.5	0.94	"
9.3	15	0.94	"
10	25	0.79	Franklin (1979)
12.5	15	0.98	"
15	25	0.93	"
20	25	0.96	"

Table 2.9. A summary of age-age genetic correlations (r_g) reported for wood density in *P. taeda.*

Juvenile age (years)	Mature age (years)	r_g	Source
2	25	0.96	Loo *et al.* (1984)
2	12	0.90	Williams and Megraw (1994)
2	12	0.83	"
3	12	0.83	"
3	12	0.76	"
4	25	0.97	Loo *et al.* (1984)
6	25	1.00	"
7	20	0.94	Talbert *et al.* (1983)
7	20	0.82	"
8	25	1.00	Loo *et al.* (1984)
10	25	1.00	"
10	20	0.77	Talbert *et al.* (1983)
10	20	0.68	"

2.5 Genotype x environment interaction

When genotype x environment interaction (GE) exists, genotypes do not respond in a similar manner in all environments. GE may come about in two ways. The relative performance of genotypes may vary across environments but the rank order of genotypes is unchanged, or both the relative performance and rank order of genotypes may vary (Falconer 1989). A significant GE effect where the rank order changes across environments will influence the breeding strategy since no single genotype may be superior in all environments. In such circumstances, the breeder may select for different environments. However, selecting for different environments may limit the number of available genotypes per site and may lead to increased rates of inbreeding. On the other hand, this strategy offers highest gains in the short/medium term (Matheson 1978). Selection for different environments can be justified only when the GE is so large that potential gains would be reduced to a degree regarded as practically serious if alternative breeding populations structure were adopted. A number of methods have been suggested to evaluate the practical importance of GE by predicting the effect of interactions on genetic gain (Matheson and Raymond 1984b, Pederick 1990).

A single trait measured in two environments can be considered as analogous to two traits in a single environment. Hence the genetic correlation across environments for a trait measured in two environments provides a measure of the magnitude of GE. The additive genetic correlation between two environments is given by (Falconer 1989)

$$r_A = \frac{Cov_A(x,y)}{\sigma_{Ax}\sigma_{Ay}} \qquad (2.1)$$

where r_A = additive genetic correlation, $Cov_A(x,y)$ = additive genetic covariance of the trait in environment x, and environment y, σ_{Ax} = additive genetic standard deviation of the trait in environment x and σ_{Ay} = additive genetic standard deviation of the trait in environment y. A genetic correlation of 1 indicates that the underlying genetic structure of the trait is similar across environments. Low genetic correlations indicate the presence

of GE, suggesting that phenotypes may be encoded by largely different suite of genes in alternative environments.

Considering the fact that forest sites where *P. taeda* is planted are heterogenous due altitude, rainfall and soils, it is surprising to find only a few studies of genotype x environment interaction in *P. taeda*. Owino (1977b) found no genotype x environment interaction for height in *P. taeda* families in the USA, but that all genotypes responded well to higher site quality. Owino and Zobel (1977) also found no genotype x environment interaction associated with heterosis across diverse sites. However, considerable GE was reported by Douglas *et al.* (1993) in volume and by Jett *et al.* (1990) in wood density across diverse field sites. Of the above, only Jett *et al.* (1990) carried out further analysis to determine if the interaction was of practical significance to justify subdivision of the breeding population; they found that it was not. These results are consistent with those reported for *P. radiata* (Carson 1991, Johnson 1992) and for the *P. elliottii* population in Zimbabwe (Pswarayi *et al.* in press), but differ from those for the *P. elliottii* populations in the USA where GE was found to be of practical significance (Hodge and White 1992).

2.6 Summary and conclusion

As the genetic parameters were estimated using relatively few families in all cases, except those of Balocchi *et al.* (1993), estimates of heritability and of genetic correlations are imprecise. Estimates of genetic correlations require large sample size and a minimum of 400 families (2 offspring per family) are recommended for estimating the genetic correlation of 0.40 with a standard error of 0.30 for two traits with a heritability of 0.20 (Klein *et al.* 1973). In order to increase precision of genetic parameter estimates and increase genetic gain, greater emphasis is being placed in the USA on the control of environmental variation through better experimental designs and layout in forest genetic tests (Weir and Goddard 1986).

Although the genetic parameter estimates of *P. taeda* from the literature are confounded with such factors as management, competition, and design and layout of the field tests, some trends emerged. For example, heritability estimates for wood density

were highest (0.5), followed by those for height, volume and straightness (0.3); those for diameter were the lowest (0.05). Also, correlations between growth traits (height, diameter and volume) were moderate to high and positive, indicating that these traits may be under the influence of similar genes and a genetic change in one trait is expected to accompany a change in the other. A high and positive genetic correlation between growth traits is desirable, because a selection based on an easily measured trait such as height will automatically improve a trait such as volume which is difficult to assess. This will result in rapid genetic progress at less cost. Genetic correlations between ages older than four years and mature ages in *P. taeda* were generally high (greater than 0.7) for the growth traits, suggesting that early selection after four years of age can be efficient in this species. Wood density at two years was well correlated with that at mature ages (genetic correlations greater than 0.8), indicating that selection at 2 years is feasible in this trait.

Generally, there was reasonable consistency in the estimates of genetic parameters, probably due to the fact that many of these results originate from the south-eastern USA. These estimates may differ from those obtained from other breeding populations elsewhere, from which independent estimates are necessary to make breeding decisions.

The literature review highlighted the following major information gaps:

1. There appears to be few heritability estimates for straightness assessed on an absolute scale;
2. There appears to be no estimates of age-age genetic correlations for stem straightness;
3. There is a lack of correlation estimates between stem straightness and other traits;
4. There is a lack of genetic parameter estimates for all traits of *P. taeda* grown in tropical regions;
5. The corresponding standard errors were not reported for most of the estimates of genetic correlations in the literature, making it difficult to judge their reliability. Also, some of the genetic correlations were equated to family mean correlations, which are likely to underestimate genetic correlations. Therefore,

reliable estimates of genetic correlations between growth traits, and between growth traits and wood density, are required.

Chapter 3

UNIVARIATE PARAMETER ESTIMATES FOR HEIGHT AND STEM STRAIGHTNESS

3.1 Introduction

Variance and heritability estimates in tree improvement programmes are important for estimating gain from selection and for devising the best breeding strategy. Heritability estimates for height reported in *Pinus taeda* in the USA are high (Balocchi *et al.* 1993, Franklin 1979, Lambeth *et al.* 1983). Heritability estimates for height for populations in Zimbabwe, and generally in southern Africa, are lacking, and estimates for stem straightness are lacking in both the USA and southern Africa. This lack of genetic parameter estimates for these economically important traits has potentially adverse consequences for realizing genetic progress in *P. taeda* breeding programmes.

Traditionally, genetic parameters in forest tree breeding programmes have been estimated using analysis of variance (ANOVA), least square methods such as Harvey's programs (Harvey 1987) on SAS (SAS Institute, Inc. 1988), and restricted maximum likelihood (REML) (Patterson and Thompson 1971) written in GENSTAT (Genstat 5 Committee 1987). An important advantage of programs than simple ANOVA is that they can efficiently analyse unbalanced data, a typical problem in tree breeding. Unfortunately, genetic models fitted by these programs allow only sib covariances; they do not allow use of covariances of other relatives, a particularly important aspect as breeding programs advance to second and subsequent generations. Also, where information is available from both the male and female parents, such as in controlled cross mating designs, a problem is encountered as to how best pool the two resultant heritability estimates, especially when the two estimates are not of equal reliability and are correlated. The model which appropriately incorporates information on genetic relationships between trees is the individual tree model, which can be fitted in REML

using derivative-free algorithms (DFREML, Meyer 1989). The individual model includes a random effect for the additive genetic merit or breeding value of each tree, both for trees with records and those that are represented as parents, and incorporates all known relationship information in the analysis. The additive genetic variance is then estimated as the variance of trees' additive genetic merit instead of estimating it from the variance between parents. The individual model has become the method of choice in animal breeding because of its desirable properties, and has recently been applied to tree breeding (*Eucalyptus* - Borralho *et al.* 1995).

The aim of the study reported in this Chapter was to undertake univariate analyses, using individual tree model DFREML, for height and stem straightness. This is the first application of an individual tree model to estimate variance components for *Pinus taeda*.

3.2 Materials and Methods

3.2.1 Genetic Material

The mating design included a diallel between 6 parents, excluding selfs, and a factorial of 7-8 males on 15 females. The design is represented diagrammatically in Table 3.1. The actual number of families planted was 140, 121, 100 and 100 at sites A, B, C, and D, respectively. The 23 parents represented were selected phenotypically from unimproved plantations in Zimbabwe and South Africa. Little is known of their origin or degree of relatedness, but it is assumed they are unrelated. The seed from which the plantations in Zimbabwe were established originated from South Africa.

3.2.2 Field sites

The genetic tests were established in 1972 at four sites in Zimbabwe, at Tarka (A), Stapleford (B), Martin (C) and Nyangui (D). Details of the sites are given in Table 3.2.

Table 3.1. Factorial and diallel crosses (surrounded by a bold line) included in the four genetic tests. A, B, C and D refer to crosses present in tests A, B, C and D, respectively, and - represents crosses absent in a particular test.

Female	Male							
	8	**10**	**13**	**95**	**162**	**164**	**171**	**196**
10	- - - -		A B C D	A B - D	A B - D	A B - D	A B - -	- - - -
13	- - - -	- - - -		A B - -	A B - -	A B - D	A B - -	- - - -
95	- - - -	A B C D	A B C D		A B C D	A B D	A B - -	- - - -
162	- - - -	A B C D	A B C D	A B - -		A B - -	A B - -	- - - -
164	- - - -	A B C D	A B C D	A B C D	A B C D		A B - -	- - - -
171	- - - -	A B C D	A B C D	A B C D	A B C D	A B C D		- - - -
59	A B C D	A B C D	A B C D	A B C D	- B C D	A B C D	- - - -	- - - -
60	A B C D	A B C D	A B C D	A B - -	- B C D	A B C D	A B C D	- - - -
68	A B C D	A B C D	A B C D	A B C D	A B - -	A B C D	A B C D	- - - -
69	A B C D	A B C D	A B C D	A B - -	A B C D	A B C D	A B C D	- - - -
70	A B C D	A B C D	A B C D	A - - -	- B C D	A B C D	A B C D	- - - -
96	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	- - - -
99	A B C D	A B C D	A B C D	A B - -	A - - -	A B - -	A - - -	- - - -
102	A - - -	A B - -	A B C -	A - - -	A - - -	A B C D	A B C D	A - - -
160	A B C D	A B C D	A B C -	- - - -	- - - -	A - - -	- - - -	- - - -
161	A B C D	A B C D	A B C D	A B C D	A B - -	A B C D	A B C D	- - - -
163	A B C D	A B C D	A B C -	A B C D	A B - -	A B C D	A B C D	- - - -
165	A B C D	A B C D	A B C D	A - - -	A B C D	A B C D	A - - -	- - - -
167	A B C D	A B C D	A B C D	A B C D	A B C -	A B C D	A B C D	- - - -
168	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	- - - -
170	A B C D	A B C D	A B - -	A B C D	A B C D	A B C D	- - - -	- - - -

Table 3.2. Synopsis of progeny tests established in 1972 in Zimbabwe.

Site (code)	Tarka(A)	Stapleford(B)	Martin(C)	Nyangui(D)
Region	Chimanimani	Penalonga	Chimanimani	Nyanga
Latitude	19°59'S	18°44'S	19°59'S	17°58'S
Longitude	32°56'E	32°49'E	32°56'E	32°47'E
Altitude (m)	1005	1745	1250	1882
Rainfall (mm)	2156	1836	1016	2364
Soils	Dolerite/alluvial -derived; reddish brown clays; well drained	Dolerite-derived; brown red clays; well drained	Dolerite/siltstones -derived; reddish brown clays; well drained	Dolerite-derived; red-reddish brown clays; well drained

3.2.3 Field design

Trees were planted at 2.4 x 2.4 m spacing and each plot comprised ten trees. The tests comprised three replicates and ten to twelve blocks per replicate, in a triple lattice design.

3.2.4 Silviculture

The first systematic thinning, removing 50% of the trees, was carried out in 1981 at age 9.5 years, and the remaining trees were pruned to one-third total tree height. A second pruning was carried out in 1983 to one-third height. In 1986, at age 13.5 years, a second systematic 50% thinning was carried out in all the tests, leaving two trees per plot.

3.2.5 Assessments

All four trials were assessed for height at 1.5 (HT1.5), 9.5 (HT9.5), 13.5 (HT13.5), and 22.5 (HT22.5) years of age, using height rods at age 1.5, and optical instruments (hypsometers) at 13.5 and 22.5 years. At 9.5 years, height was assessed on thinned trees using measuring tapes.

Stem straightness was assessed at 1.5 (ST1.5), 9.5 (ST9.5), 13.5 (ST13.5) and 22.5 (ST22.5) years of age, using the score in Table 3.3 (Barrett and Mullin 1968).

Table 3.3. Stem straightness visual score used for assessing the tests.

Score	Description
7	No deviation which would appreciably reduce utilisation value
6	Basically as 7 but with one minor defect slightly affecting utilisation or slightly inferior to 7 but no specific defect
5	Basically as 7 but with one major defect appreciably reducing utilisation value or as in 6 but with a minor defect
4	As 6 but with either a major defect or severe minor defects
3	As 6 but with at least two major defects or a series of minor defects
2	Inferior to 3 and with specific defects such that utilisation would be confined to short lengths
1	As 2 but with major and minor defects restricting utilisation to fuel and industrial wood

3.2.6 Summary of data

A summary of the data used in the analyses is shown in Table 3.4. Fewer trees and fewer families than usual were measured at age 9.5 years in test B because trees in only the first nine blocks of the first replicate were measured. The total number of families common to all the four tests was 85.

Table 3.4. Summary of data analysed for tests A, B, C and D.

Age	Test	No. of males	No. of females	No. of families	No. of trees
1.5	A	8	21	140	4162
	B	7	21	121	3557
	C	7	20	100	2878
	D	7	21	100	2907
9.5	A	8	21	140	1862
	B	7	21	94	460
	C	7	20	100	1381
	D	7	21	100	1419
13.5	A	8	21	140	2101
	B	7	21	121	1042
	C	7	20	100	1260
	D	7	21	100	1392
22.5	A	8	21	140	427
	B	7	21	121	708
	C	7	20	100	257
	D	7	21	100	589

3.2.7 Statistical and genetic analyses

Descriptive statistics were derived using Minitab (Minitab Release 7.2 Sun-4 version, 1989). The aim of the descriptive analysis was to ascertain that the data had a normal distribution, and to obtain the mean and the scale of the data for each trait. It was necessary to test assumption of normality because parametric methods used to analyse the data are based on this assumption.

The sib covariance model available in GENSTAT REML and individual tree model available in DFREML were used to estimate the variance components and heritabilities. The first method was employed primarily to highlight some of the major problems of a program commonly used in forestry, and to obtain initial estimates for the DFREML runs.

The analysis based on sib covariance was performed on individual tests and the following model was used:

$$Y_{ijklm} = \mu + R_i + B_{ij} + F_k + M_l + FM_{kl} + \epsilon_{ijklm} \qquad (3.1)$$

where: Y_{ijklm} is the observation on the mth tree in the ith replicate and jth block, and of the kth female and lth male parents,

μ is the overall test mean,

R_i is the fixed effect of the ith replicate,

B_{ij} is the random effect of the jth block in the ith replicate,

F_k is the random effect of the kth female parent,

M_l is the random effect of the lth male parent,

Fm_{kl} is the random effect of the klth family (female x male parent interaction), and

ϵ_{ijklm} is the within plot error (residual), assumed to be normally distributed with mean 0 and variance σ^2.

The narrow sense heritability was calculated on an individual tree basis using the paternal variance component as:

$$h^2_M = 4\sigma^2_M/\sigma^2_P \qquad (3.2)$$

and using the maternal variance component as:

$$h^2_F = 4\sigma^2_F/\sigma^2_P \qquad (3.3)$$

where the phenotypic variance is estimated as:

$$\sigma^2_P = \sigma^2_B + \sigma^2_F + \sigma^2_M + \sigma^2_{FM} + \sigma^2_E \qquad (3.4)$$

The component σ^2_{FM} is an estimate of the covariance of full sibs, less the covariance of both maternal and paternal half-sibs, and estimates 1/4 of the dominance genetic variance.

Standard errors of the heritability estimates were calculated using the vfunction option in GENSTAT. The two heritability estimates obtained from the analyses were pooled, weighting each estimate by the inverse of its standard error:

$$h^2_{Pool} = wh^2_F + (1-w)h^2_M \qquad (3.5)$$

where:

$$w = \frac{se^2_M}{se^2_F + se^2_M} \qquad (3.6)$$

and se_F is the standard error of the female estimate and se_M standard error of the male estimate. The standard error of the pooled estimate was calculated as:

$$se_{Pool} = \sqrt{\frac{se^2_M se^2_F}{se^2_M + se^2_F}} \qquad (3.7)$$

This method outlined above will bias the pooled heritability estimate downwards by giving more weight to the smaller heritability with a relatively smaller variance. since the standard error of a heritability estimate depends on the heritability. The estimates of the pooled heritability were corrected for this bias through an iterative procedure in which estimates of the pooled heritability were used to reestimate the standard error of the heritability for the female and the male parents, and the procedure was repeated until the pooled heritability estimates converged.

The following models were used in the analysis of individual site data using an individual tree model:

Model 1: $Y_{ijkl} = \mu + R_i + B_{ij} + FM_{ijk} + A_l + \epsilon_{ijkl}$ (3.8)

Model 2: $Y_{ijklm} = \mu + R_i + B_{ij} + FM_{ijk} + Pt_m + A_l + \epsilon_{ijklm}$ (3.9)

where: Y_{ijkl} is the observation on the lth tree in the ith replicate and jth block and in the kth family,

μ is the overall test mean,

R_i is the fixed effect of the ith replicate,

B_{ij} is the random effect of the jth block in the ith replicate,

FM_{ijk} is the random effect of the kth family (female x male parent interaction),

A_l is the additive genetic effect of the lth tree, and

ϵ_{ijkl} is the within plot error (residual), assumed to be normally distributed with mean 0 and variance σ^2.

The difference between the two models is that model 2 fitted plot (Pt_m) as an additional random effect in order to measure common environment effect.

The following models were used in the analysis of pooled data across sites using individual tree model DFREML:

Model 1: $Y_{ijklm} = \mu + S_i + R_j + B_{jk+} FM_{jkl} + A_m + \epsilon_{ijklm}$ (3.10)

Model 2: $Y_{ijklmn} = \mu + S_i + R_j + B_{jk+} FM_{jkl} + Pt_m + A_n + \epsilon_{ijklmn}$ (3.11)

The difference between the two models is as described above, where (Pt_m) estimates common environment effect, and site (S_i) was fitted as an additional fixed effect.

The general model in matrix notation was:

$$y = \mathbf{X}b + \mathbf{Z}a + \mathbf{W}c + e \qquad (3.12)$$

where: y = the vector of observations,

b = the vector of fixed effects,

X = the incidence matrix for fixed effects,

a = the vector of additive genetic effects for trees,

c = the vector of additional random effects,

W,Z = the incidence matrices for random effects, and

e = the vector of random residual errors.

The (co)variance structure for the analysis can be described as:

$$V(a) = \sigma^2_A \mathbf{A}$$

$$V(c) = \sigma^2_c \mathbf{I}$$

$$V(e) = \sigma^2_e \mathbf{I}$$

where: σ^2_A is the additive genetic variance,

σ^2_c the variance of the additional random effect (block, family or plot),

σ^2_e the error variance,

A is the numerator relationship matrix between the trees, and

I the identity matrix.

Vectors **c** and **e** were assumed to be uncorrelated with all other effects. The initial values for the DFREML runs were taken from the GENSTAT REML analyses. Different starting values from the initial ones were used to confirm that a global rather than a local maximum had been reached.

Comparisons between the individual tree models were made by likelihood ratio tests (Meyer 1993), which consist of subtracting the maximum log likelihood for the model with fewer parameters from the value corresponding to the model with more parameters, and then multiplying the difference by 2. The test statistic is distributed asymptotically as a Chi-Square random variable with degrees of freedom equal to the difference in the number of parameters estimated for the two models. The test is

appropriate where parameters in one model are a subset of parameters in the other.

The appropriateness of pooling data across sites was determined using the joint likelihood of the four independent analyses. Joint likelihood of the four analyses is simply the sum of the individual log likelihoods, asymptotically distributed as Chi-Square with degrees of freedom equal to the sum of the parameters (i.e. df = 12 for model 1 and df = 16 for model 2). The difference between the joint likelihood and the analyses of pooled data is compared with Chi-Square distribution with degrees of freedom equal to the differences in the number of parameters between them.

The additive genetic coefficient of variation (CV_A) was calculated as:

$$CV_A = 100(\sigma_A/\mu) \qquad (3.13)$$

where:

σ_A is the additive genetic standard deviation, and

μ is the phenotypic mean for the trait.

The importance of the dominance variance was evaluated in two ways:

(i) dominance as a proportion of additive variance:

$$D_A = \sigma^2_D/\sigma^2_A, \text{ and} \qquad (3.14)$$

(ii) dominance as a proportion of phenotypic variance:

$$D_P = \sigma^2_D/\sigma^2_P. \qquad (3.15)$$

D_A was used to assess the relative size of dominance to additive variance, and hence its contribution to the genetic variance. However, a high D_A may be inconsequential to a trait if the dominance variance is small compared with the phenotypic variance. Therefore, D_P was also calculated.

The standard errors of the heritability estimates were calculated using DFREML. At convergence, DFREML attempts to estimate the standard errors of heritability estimates by fitting a quadratic function to the likelihood surface using points evaluated during the search for the likelihood. As pointed out by Meyer (1993), little is known about the likelihood surface and the quadratic surface may not provide a good fit. Therefore, in order to get accurate confidence intervals for the heritability estimates from the pooled data, likelihood profiles were plotted by fixing the heritability to

different values, and the likelihood maximised with respect to all the other parameters. The 95% confidence interval was obtained by dropping 1.92 ($0.5\chi^2_{1,\ 0.05}$) from the maximum log likelihood (Wetherill 1981).

3.3 Results

3.3.1 Descriptive statistics

The descriptive phenotypic statistics derived from using Minitab are shown in Tables 3.5 and 3.6. The overall site mean height at 1.5 years of age ranged from 0.71-1.13 m, with the Stapleford (B) having the best performance and Nyangui (D) the worst (Table 3.5). This trend changed over time, and at 22.5 years Nyangui site had the best performance and Tarka (A) the worst, with a range of 22.46-25.05 m. The trees with the best form at 22.5 years were found at Stapleford (B). The sites with the worst height growth had the worst straightness scores at all ages.

Table 3.5. The minimum, maximum, mean, standard deviation and coefficient of variation of height (m) at four sites (A-D) at 1.5, 9.5, 13.5 and 22.5 years.

Trait	Site	Min	Max	Mean	SD	CV%
HT1.5	A	0.1	1.9	0.997	0.255	25.6
	B	0.1	2.2	1.13	0.291	25.8
	C	0.2	2.2	1.01	0.296	29.3
	D	0.1	1.6	0.704	0.227	32.2
HT9.5	A	7.5	19.5	14.7	1.60	10.9
	B	2.4	15.2	12.1	1.68	13.9
	C	6.2	19.2	14.8	1.73	11.7
	D	3.9	16.0	13.2	1.50	11.4
HT13.5	A	6.2	29.6	19.4	2.68	13.8
	B	3.4	22.4	17.2	2.37	13.8
	C	6.3	24.7	19.9	2.01	10.1
	D	3.1	23.0	19.1	2.01	10.5
HT22.5	A	8.6	26.8	22.5	1.92	8.53
	B	5.7	29.2	24.2	2.00	8.26
	C	14.7	28.4	24.5	1.42	5.80
	D	19.2	30.6	25.1	1.49	5.94

Table 3.6. The minimum, maximum, mean, standard deviation and coefficient of variation of stem straightness (score 1-7) at four sites (A-D) at 1.5, 9.5, 13.5 and 22.5 years.

Trait	Site	Min	Max	Mean	SD	CV%
ST1.5	A	2	6	4.00	0.600	15.0
	B	2	6	3.85	0.484	12.6
	C	2	6	3.95	0.540	13.7
	D	1	6	3.59	0.755	21.0
ST9.5	A	3	6	4.51	0.682	15.1
	B	1	5	3.64	0.866	23.8
	C	2	6	4.42	0.640	14.5
	D	2	6	4.49	0.740	16.5
ST13.5	A	2	6	3.91	0.678	17.3
	B	1	6	3.54	0.746	21.1
	C	2	7	4.68	1.04	22.2
	D	1	7	4.17	0.800	19.2
ST22.5	A	3	6	4.27	0.645	15.1
	B	5	7	6.62	0.513	7.75
	C	3	7	4.75	0.693	14.6
	D	4	7	5.89	0.764	13.0

Both traits showed substantial variation (Tables 3.5 and Table 3.6). The standard deviations for height did not differ appreciably among the sites (ranges: 0.23-0.30 at 1.5 years, and 1.63-1.97 at 22.5 years). This indicates that data can be pooled over sites without seriously violating homogeneity assumptions. The coefficient of variation for height was high at 1.5 years (27%), decreasing progressively with age: at 9.5 (12%), 13.5 (12%) and 22.5 years (8%). The correlation between mean height and variation (standard deviation) was high (0.84), indicating that variation within a site was well predicted by mean height. This relationship was not evident for straightness (correlation= 0.25).

3.3.2 Analysis based on sib covariances

Variance components

The magnitude of the variance component estimates for height and straightness are shown in Tables 3.7 and 3.8, respectively. The major source of variation in both traits was the residual variance which accounted for more than 60% of the variation, while the two smallest were the block and the family variance. In general, the male variance was larger than the female variance for height, and the converse was true for straightness. The standard errors of the female and male variance estimates were large, with those of the former being larger for similar magnitudes of variance. For height, both the residual and the phenotypic variances peaked at 13.5 years. A few negative variance estimates were obtained for straightness, probably reflecting some of the problems associated with estimating variance components from a limited set of parents.

Additive and dominance variances

Tables 3.9 and 3.10 give the estimates of the additive and dominance variances, and ratios of the dominance to the additive for both female and male: D_{AF} and D_{AM}, respectively, and dominance to the phenotypic variances (D_P). Trends of the additive variances followed the trends of the respective variance components. Dominance variance for height as a proportion of the additive variance was less than 1, except at 1.5 years at site A, at 13.5 years at site D, and at 9.5 years at sites B and D. This indicates that additive variance was more important than the dominance variance for height. In contrast, dominance was more important than additive variance for straightness, particularly at age 1.5 where dominance was as much as 14 times greater than the additive variance. However, at this age D_P was less than 0.10, indicating that environmental effects were the major determinant of straightness at this age. The parameter D_P for height ranged from 0.01 to 0.36, and there was no relationship between its magnitude and age.

Table 3.7. Estimates of variance components ($x10^2$) and their standard errors (in brackets) for height at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	σ_M^2	σ_F^2	σ_B^2	σ_{FM}^2	σ_E^2	σ_P^2
1.5	A	0.423 (0.265)	0.268 (0.118)	0.250 (0.71)	0.492 (0.092)	5.07 (0.115)	6.51
	B	0.973 (0.577)	0.480 (0.183)	0.664 (0.188)	0.297 (0.078)	6.06 (0.147)	8.47
	C	0.746 (0.468)	0.761 (0.306)	0.242 (0.91)	0.558 (0.132)	6.47 (0.174)	8.77
	D	0.140 (0.098)	0.180 (0.078)	0.198 (0.069)	0.098 (0.045)	4.55 (0.122)	5.16
9.5	A	42.2 (24.4)	38.6 (13.1)	7.70 (2.80)	5.60 (2.50)	163 (5.60)	257
	B	28.2 (19.6)	24.5 (13.3)	8.70 (8.40)	25.5 (12.3)	196 (14.5)	283
	C	30.2 (18.4)	42.2 (15.3)	5.20 (2.90)	5.40 (3.70)	216 (8.60)	299
	D	10.2 (7.00)	26.8 (10.4)	9.90 (3.90)	11.4 (4.00)	168 (6.60)	226
13.5	A	48.8 (29.2)	35.5 (14.3)	17.4 (7.30)	18.7 (8.30)	598 (19.4)	719
	B	69.6 (41.4)	29.6 (13.3)	58.1 (18.7)	17.0 (9.80)	387 (18.3)	561
	C	36.1 (22.2)	50.8 (18.6)	15.6 (6.70)	2.40 (4.60)	298 (12.4)	402
	D	25.4 (16.9)	25.6 (12.7)	5.60 (4.00)	30.9 (9.00)	319 (12.7)	407
22.5	A	44.9 (28.7)	29.4 (14.5)	18.9 (10.8)	13.3 (15.6)	261 (22.9)	367
	B	24.0 (16.5)	21.4 (11.2)	56.4 (18.5)	23.3 (11.2)	273 (16.3)	398
	C	24.0 (16.4)	13.4 (8.70)	20.5 (10.7)	0.700 (11.7)	142 (16.9)	201
	D	28.5 (17.9)	23.7 (9.90)	18.8 (7.40)	3.60 (5.30)	148 (9.70)	222

Table 3.8. Estimates of variance components ($x10^2$) and their standard errors (in brackets) for stem straightness at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	σ_M^2	σ_F^2	σ_B^2	σ_{FM}^2	σ_E^2	σ_P^2
1.5	A	-0.030 (0.050)	0.260 (0.160)	1.74 (0.510)	0.390 (0.220)	33.7 (0.760)	36.0
	B	0.040 (0.070)	0.180 (0.140)	0.100 (0.090)	0.570 (0.200)	22.6 (0.550)	23.4
	C	0.090 (0.110)	0.800 (0.350)	1.05 (0.380)	0.250 (0.210)	26.9 (0.730)	29.1
	D	0.130 (0.240)	0.370 (0.330)	1.30 (0.530)	0.560 (0.410)	54.6 (0.015)	57.0
9.5	A	2.27 (1.43)	2.93 (1.14)	1.50 (0.590)	1.18 (0.580)	38.6 (1.33)	46.5
	B	4.60 (3.45)	2.86 (2.25)	3.31 (0.260)	2.33 (2.82)	61.9 (4.58)	75.0
	C	0.840 (0.640)	2.01 (0.89)	0.860 (0.480)	0.510 (0.550)	36.7 (1.46)	41.0
	D	0.870 (0.790)	3.94 (1.76)	0.590 (0.490)	3.20 (1.09)	46.1 (1.81)	54.7
13.5	A	1.22 (0.690)	1.84 (0.730)	0.610 (0.320)	0.070 (0.401)	42.2 (1.36)	45.9
	B	3.14 (2.12)	1.74 (1.09)	3.62 (1.36)	3.49 (1.32)	43.7 (2.06)	55.7
	C	5.50 (3.47)	6.36 (2.52)	-0.520 (0.460)	-1.47 (1.05)	98.1 (4.08)	108
	D	1.11 (0.910)	4.88 (1.98)	0.610 (0.550)	1.35 (0.930)	56.0 (2.22)	64.0
22.5	A	0.790 (0.970)	2.27 (1.49)	1.86 (1.26)	3.90 (2.23)	32.8 (2.89)	41.6
	B	1.89 (1.27)	2.16 (0.980)	0.710 (0.470)	1.43 (0.760)	20.2 (1.20)	26.4
	C	4.71 (3.40)	12.38 (5.12)	0.41 (1.10)	4.95 (2.75)	25.6 (2.75)	48.1
	D	2.95 (2.17)	5.05 (2.39)	3.08 (1.56)	2.24 (1.78)	45.1 (2.95)	58.4

Table 3.9. Estimates of additive and dominance variances ($x10^2$), and the ratio of dominance to additive female and male variances, and dominance to phenotypic variance for height at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	σ_{AM}^2	σ_{AF}^2	σ_D^2	D_{AM}	D_{AF}	D_P
1.5	A	1.69	1.07	1.97	1.16	1.84	0.30
	B	3.89	1.92	1.19	0.31	0.62	0.14
	C	2.98	3.04	2.23	0.75	0.73	0.25
	D	0.560	0.720	0.392	0.70	0.54	0.08
9.5	A	169	154	22.4	0.13	0.15	0.09
	B	113	98.0	102	0.90	1.04	0.36
	C	121	169	21.6	0.18	0.13	0.07
	D	40.8	107	45.6	1.12	0.43	0.20
13.5	A	195	142	74.8	0.38	0.53	0.10
	B	278	118	68.0	0.24	0.57	0.12
	C	144	203	9.60	0.07	0.05	0.02
	D	102	102	124	1.22	1.21	0.30
22.5	A	180	118	53.2	0.30	0.45	0.14
	B	96.0	85.6	93.2	0.97	1.09	0.23
	C	96.0	53.6	2.80	0.03	0.05	0.01
	D	114	94.8	14.4	0.13	0.15	0.06

Table 3.10. Estimates of additive and dominance variances ($x10^2$), and the ratio of dominance to additive female and male variances, and dominance to phenotypic variance for stem straightness at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	σ_{AM}^2	σ_{AF}^2	σ_D^2	D_{AM}	D_{AF}	D_P
1.5	A	-0.120	1.04	1.56	-13.00	1.50	0.04
	B	0.160	0.720	2.28	14.25	3.17	0.10
	C	0.360	3.20	1.00	2.78	0.31	0.03
	D	0.520	1.48	2.24	4.31	1.51	0.04
9.5	A	9.08	11.72	4.72	0.52	0.40	0.10
	B	18.40	11.44	9.32	0.51	0.81	0.12
	C	3.36	8.04	2.04	0.61	0.25	0.05
	D	3.48	15.76	12.8	3.68	0.81	0.15
13.5	A	4.88	7.36	0.280	0.06	0.04	0.01
	B	12.56	6.96	13.96	1.11	2.01	0.25
	C	22.00	25.44	-5.88	-0.27	-0.23	-0.05
	D	4.44	19.52	5.40	1.22	0.28	0.08
22.5	A	3.16	9.08	15.60	4.94	1.59	0.37
	B	7.56	8.64	5.72	0.76	0.66	0.22
	C	18.84	49.52	19.80	1.05	0.40	0.41
	D	11.80	20.20	8.96	0.76	0.44	0.15

Heritability estimates

Results from sib covariance analyses are presented in Tables 3.11 and 3.12. Heritability estimates for height were moderate to high, ranging from 0.11 to 0.66 for those estimated from male parents and 0.14 to 0.60 for those estimated from female parents. Where the estimates were similar, those from female parents had a lower standard error, as they were estimated with more degrees of freedom. As expected, the pooled estimates for height had lower standard errors than the individual estimates. The pooled heritability estimates for height peaked at 9.5 years in all tests except D.

Apart from an usually high heritability estimate for straightness at 22.5 years at site C, heritability estimates for stem straightness were much lower than those for height, with those estimated from male parents ranging from 0 to 0.39, and those estimated from female parents from 0.03 to 0.35. Estimates for this trait from female parents were consistently higher than those from male parents. Estimates at 1.5 years of age were low, reflecting the difficulty of straightness assessment at this early age.

Table 3.11. Heritability estimates for height at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	h^2_M (se)	h^2_F (se)	h^2_{Pool} (se)
1.5	A	0.26 (0.15)	0.17 (0.07)	0.19 (0.07)
	B	0.46 (0.24)	0.23 (0.08)	0.28 (0.09)
	C	0.34 (0.20)	0.35 (0.13)	0.34 (0.11)
	D	0.11 (0.07)	0.14 (0.06)	0.13 (0.05)
9.5	A	0.66 (0.32)	0.60 (0.18)	0.62 (0.16)
	B	0.40 (0.25)	0.35 (0.18)	0.37 (0.14)
	C	0.40 (0.22)	0.56 (0.18)	0.52 (0.14)
	D	0.18 (0.12)	0.47 (0.16)	0.38 (0.12)
13.5	A	0.27 (0.15)	0.20 (0.08)	0.22 (0.07)
	B	0.48 (0.26)	0.21 (0.09)	0.29 (0.10)
	C	0.36 (0.20)	0.51 (0.17)	0.46 (0.13)
	D	0.25 (0.16)	0.25 (0.12)	0.25 (0.09)
22.5	A	0.49 (0.28)	0.32 (0.15)	0.38 (0.13)
	B	0.24 (0.16)	0.22 (0.11)	0.22 (0.09)
	C	0.48 (0.29)	0.27 (0.17)	0.35 (0.15)
	D	0.45 (0.28)	0.43 (0.16)	0.45 (0.14)

h^2_F = individual tree heritability calculated using the female variance component.

h^2_M = individual tree heritability calculated using the male variance component.

h^2_{Pool} = pooled estimate of heritability.

Table 3.12. Heritability estimates for stem straightness at four sites and four ages, based on sib covariance analysis.

Age (years)	Site	h^2_M (se)	h^2_F (se)	h^2_{Pool} (se)
1.5	A	0 (0.01)	0.03 (0.02)	0.01 (0.01)
	B	0.01 (0.01)	0.03 (0.02)	0.02 (0.01)
	C	0.01 (0.01)	0.10 (0.05)	0.08 (0.03)
	D	0.01 (0.01)	0.03 (0.02)	0.02 (0.01)
9.5	A	0.20 (0.12)	0.25 (0.09)	0.24 (0.07)
	B	0.25 (0.17)	0.15 (0.12)	0.19 (0.10)
	C	0.08 (0.06)	0.20 (0.08)	0.16 (0.06)
	D	0.06 (0.06)	0.29 (0.12)	0.20 (0.08)
13.5	A	0.11 (0.06)	0.16 (0.06)	0.14 (0.05)
	B	0.23 (0.14)	0.13 (0.08)	0.16 (0.07)
	C	0.20 (0.12)	0.24 (0.09)	0.23 (0.07)
	D	0.07 (0.06)	0.31 (0.12)	0.23 (0.08)
22.5	A	0.08 (0.09)	0.22 (0.14)	0.15 (0.09)
	B	0.29 (0.18)	0.33 (0.14)	0.31 (0.11)
	C	0.39 (0.26)	1.03 (0.34)	0.79 (0.24)
	D	0.20 (0.14)	0.35 (0.15)	0.29 (0.11)

3.3.3 Analysis using the individual tree model

Model without fitting a common environment effect

The magnitude of variances estimated using the individual tree model, without fitting a common environment effect (Model 1), are shown in Tables 3.13-3.16. The corresponding trends over time in pooled analysis are shown in Figures 3.1 and 3.2.

For height, there was a large increase in all variances from age 1.5 to 9.5 years, with the additive variance peaking at 9.5 years, and the phenotypic, environment and dominance variances peaking at 13.5 years. The phenotypic variance in straightness followed the same trend as did that for height (Figure 3.1), but the additive variance continued to increase with age (Figure 3.2). For height, dominance variance was less than the corresponding additive variance, except at age 1.5 years in test A and age 13.5 years in D. When the data were pooled over sites, additive variance was greater than the dominance variance at all ages.

For straightness, dominance variance was much higher than additive variance at 1.5 years in three of the tests. When the data were pooled over sites, dominance variance was five times more than the additive variance at 1.5 years and equal to the additive variance at 9.5 years, after which the additive component predominated.

The additive genetic coefficient of variation (CV_A) for height decreased with age, while that for straightness peaked at 13.5 years and appeared to remain constant thereafter (Figure 3.3). The results show that there is no relationship between heritability and CV_A in either trait.

The magnitude and trends of dominance and additive variances were similar to those of the analyses based on sib covariance, but the residual variances were always less.

Table 3.13. Model 1 estimates of variance components ($x10^2$), and importance of dominance variance for height at four sites and four ages, based on analysis using the individual tree model.

Age (years)	Site	σ_A^2	σ_D^2	σ_E^2	D_A	D_P
1.5	A	1.50	1.75	4.37	1.17	0.27
	B	2.70	1.17	4.72	0.43	0.14
	C	3.10	2.28	4.91	0.74	0.26
	D	0.730	0.450	4.16	0.62	0.09
9.5	A	198	21.0	62.1	0.11	0.08
	B	114	96.6	137	0.84	0.34
	C	193	22.2	120	0.11	0.07
	D	94.3	50.6	120	0.54	0.21
13.5	A	171	80.3	511	0.45	0.11
	B	172	63.0	301	0.37	0.12
	C	209	13.9	194	0.07	0.03
	D	93.7	130	273	1.39	0.32
22.5	A	149	44.4	191	0.30	0.12
	B	106	82.8	221	0.78	0.21
	C	78.6	4.07	102	0.05	0.02
	D	85.9	34.2	104	0.40	0.16

Table 3.14. Model 1 estimates of variance components ($x10^2$), and importance of dominance variance for stem straightness at four sites and four ages, based on analysis using the individual tree model.

Age (years)	Site	σ_A^2	σ_D^2	σ_E^2	D_A	D_P
1.5	A	0.64	1.89	33.27	2.95	0.05
	B	0.27	2.66	22.41	9.85	0.11
	C	2.24	1.22	25.83	0.54	0.04
	D	1.16	2.47	54.05	2.13	0.04
9.5	A	11.83	4.71	32.72	0.40	0.10
	B	16.26	8.48	53.90	0.52	0.11
	C	6.54	2.26	33.45	0.35	0.05
	D	9.99	16.83	41.19	1.68	0.30
13.5	A	7.53	0	38.44	0	0
	B	9.73	14.51	38.80	1.49	0.26
	C	25.85	2.36	83.06	0.09	0.02
	D	15.28	7.38	48.36	0.48	0.11
22.5	A	6.78	18.93	28.64	2.79	0.45
	B	8.74	5.30	15.85	0.61	0.20
	C	45.57	17.63	2.95	0.39	0.33
	D	18.97	7.89	35.71	0.42	0.13

Model fitting a common environment effect

Tables 3.17-3.20 show the variances estimated fitting a common environment effect (Model 2). The dominance variances were lower than those estimated from fitting Model 1, indicating that some of the common environment effects were confounded with the dominance variance. The additive and phenotypic variances estimated by the two models were similar, as expected. This resulted in lower ratios of dominance and additive variance from Model 2. When the data were pooled over sites.

dominance variance for height was less than 40% of the additive variance at all ages. In the case of straightness, dominance variance was lower than the additive variance at all ages except 1.5 years, when it was two and a half times greater.

Table 3.15. Model 1 estimates of variance components ($x10^2$), importance of dominance variance, and additive genetic coefficient of variation (CV_A) for height using data pooled across sites at each of the four ages, based on analysis using the individual tree model.

Age	σ_A^2	σ_D^2	σ_E^2	D_A	D_P	CV_A
1.5	1.73	0.540	4.98	0.31	0.07	14.0
9.5	139	29.0	121	0.21	0.10	8.62
13.5	130	55.1	380	0.42	0.10	6.04
22.5	73.5	37.5	195	0.51	0.12	2.98

Table 3.16. Model 1 estimates of variance components ($x10^2$), importance of dominance variance, and additive genetic coefficient of variation (CV_A) for stem straightness using data pooled across sites at each of the four ages, based on analysis using the individual tree model.

Age	σ_A^2	σ_D^2	σ_E^2	D_A	D_P	CV_A
1.5	0.345	1.29	34	3.74	0.04	1.53
9.5	5.43	5.80	41.8	1.07	0.12	5.46
13.5	5.98	1.89	58.2	0.32	0.03	5.99
22.5	9.52	3.63	30.4	0.38	0.09	5.74

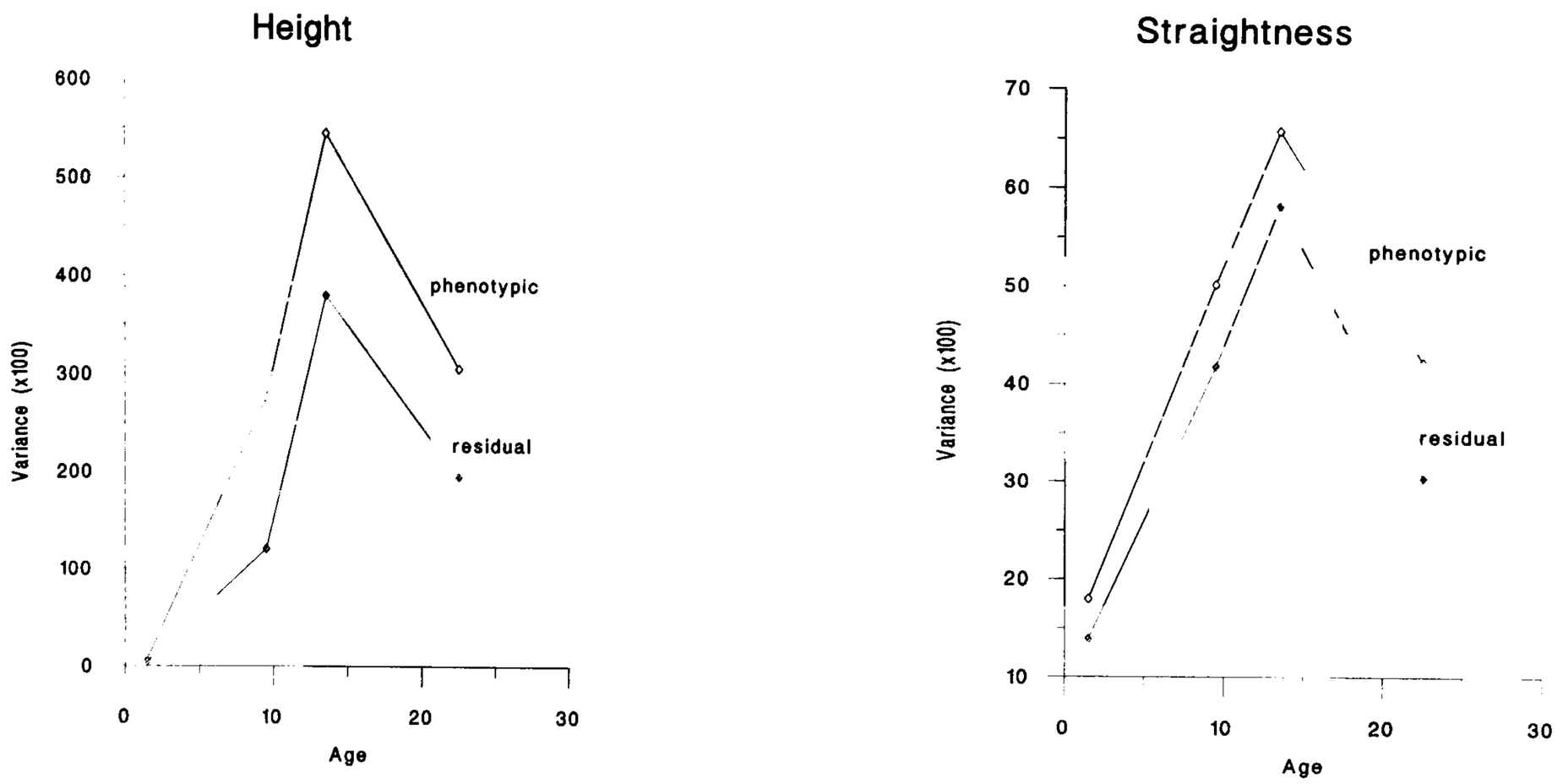


Figure 3.1. Phenotypic and residual variances for height and straightness over time.

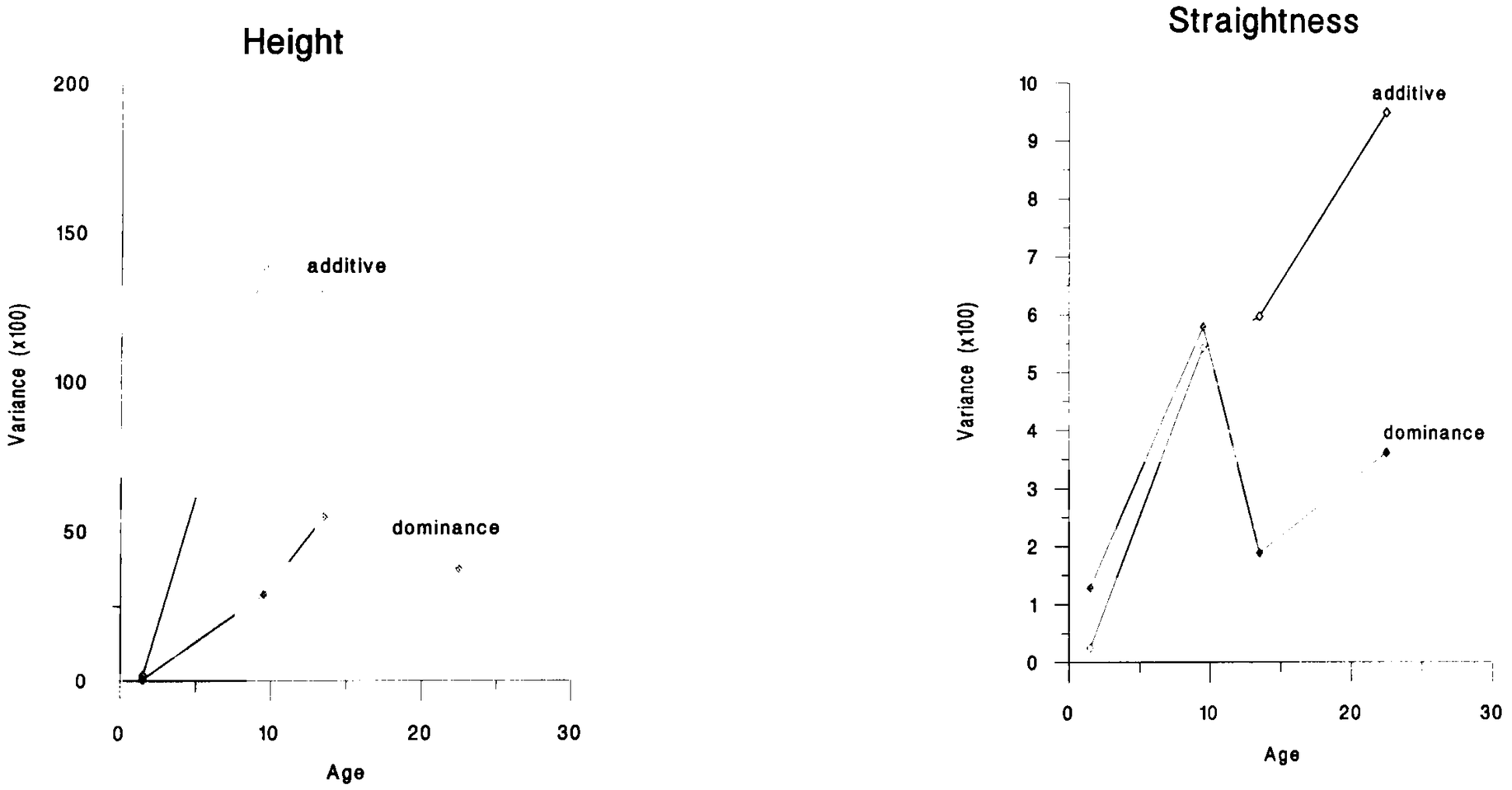


Figure 3.2. Additive and dominance variances for height and straightness over time.

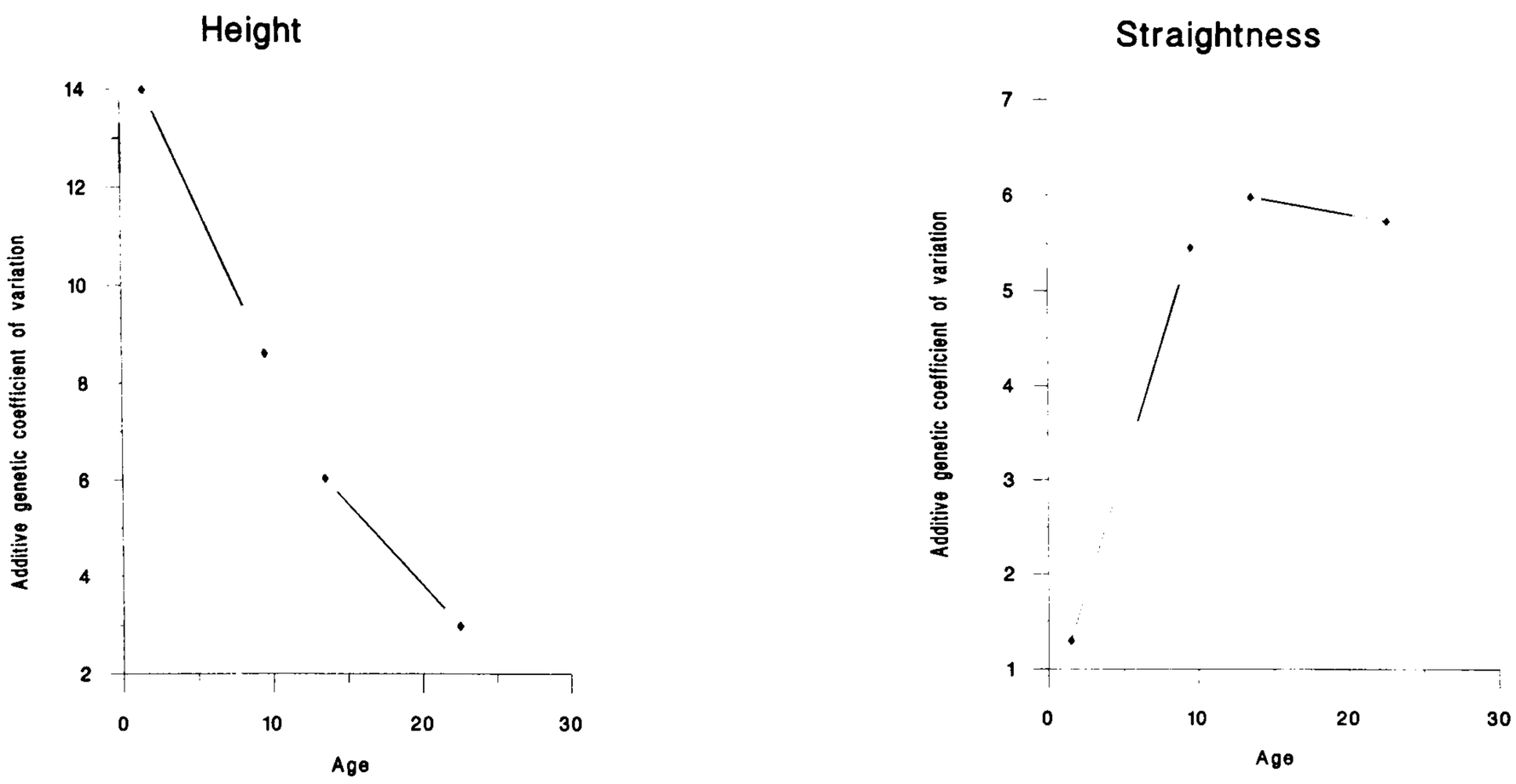


Figure 3.3. Additive genetic coefficient of variation for height and straightness over time.

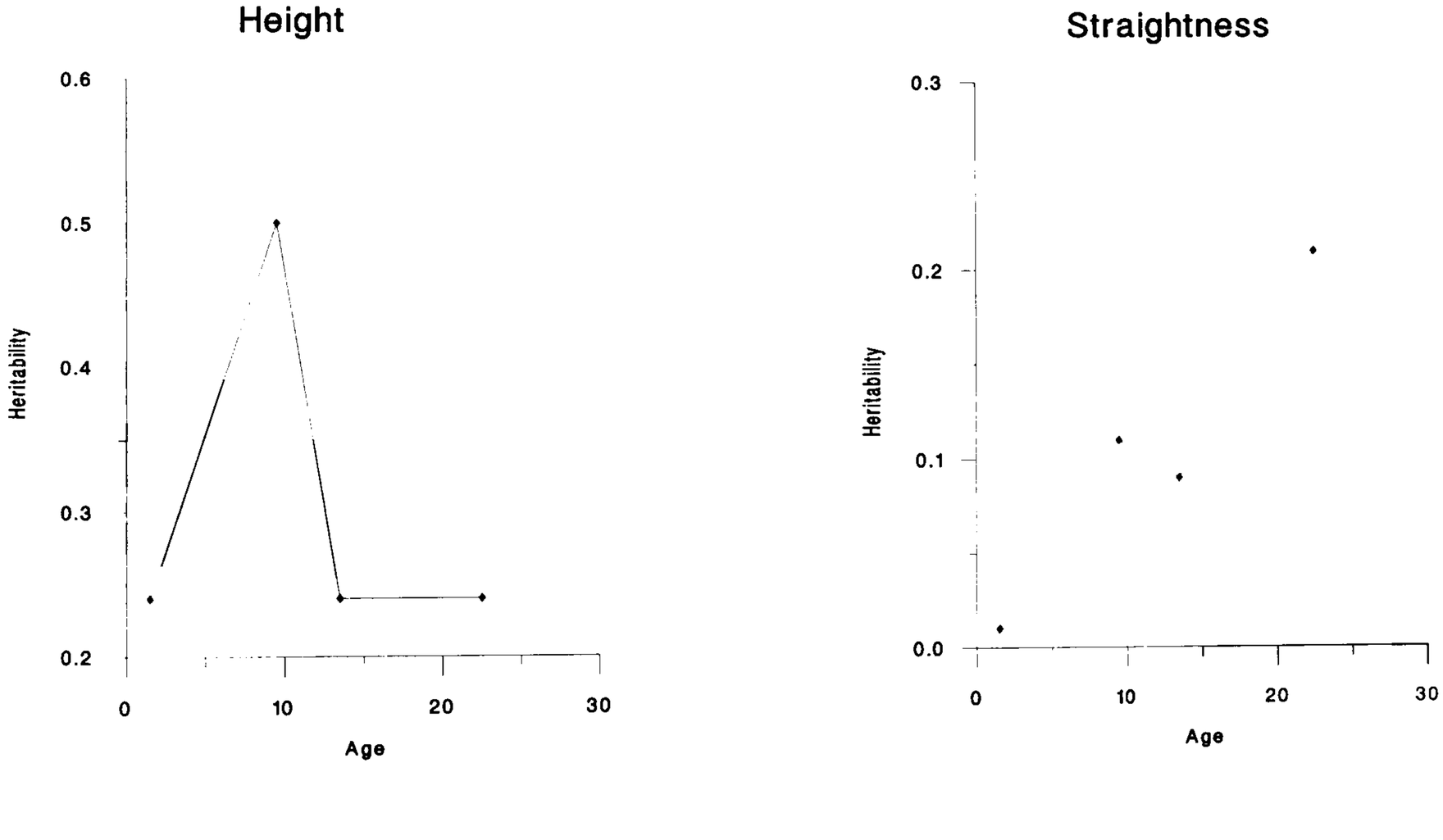


Figure 3.4. Heritability estimates for height and straightness over time.

Table 3.17. Model 2 estimates of variance components ($x10^2$), and importance of dominance variance for height at four sites and four ages, based on analysis using the individual tree model.

Age (years)	Site	σ_A^2	σ_D^2	σ_E^2	D_A	D_P
1.5	A	1.48	1.32	4.13	0.89	0.20
	B	2.68	0.751	4.54	0.28	0.09
	C	3.10	1.67	4.64	0.54	0.19
	D	0.729	0.290	4.10	0.40	0.06
9.5	A	200	16.6	59.7	0.08	0.06
	B	116	49.1	138	0.42	0.17
	C	190	20.1	119	0.11	0.06
	D	93.7	51.1	121	0.55	0.22
13.5	A	173	87.7	508	0.51	0.12
	B	173	60.6	296	0.35	0.11
	C	205	14.0	195	0.07	0.03
	D	93.2	122	268	1.31	0.30
22.5	A	153	5.07	191	0.03	0.01
	B	105	65.19	209	0.62	0.16
	C	77.0	0.601	95.2	0.01	0.002
	D	86.7	16.4	90.6	0.19	0.08

Table 3.18. Model 2 estimates of variance components ($x10^2$), and importance of dominance variance for stem straightness at four sites and four ages, based on analysis using the individual tree model.

Age (years)	Site	σ_A^2	σ_D^2	σ_E^2	D_A	D_P
1.5	A	0.540	0.790	32.6	1.46	0.02
	B	0.260	2.04	22.1	7.85	0.09
	C	2.22	0.550	25.5	0.25	0.02
	D	1.17	1.50	53.6	1.28	0.03
9.5	A	11.6	1.76	30.9	0.15	0.04
	B	16.2	0	54.0	0	0
	C	6.49	0.720	32.5	0.11	0.02
	D	9.80	15.9	40.5	1.62	0.29
13.5	A	7.51	0	38.5	0	0
	B	9.78	11.2	36.6	1.15	0.20
	C	30.3	0.080	78.6	0	0
	D	15.3	7.09	48.5	0.46	0.11
22.5	A	6.77	18.8	28.6	2.78	0.45
	B	8.67	4.46	15.08	0.51	0.17
	C	50.0	20.2	1.12	0.40	0.34
	D	18.9	5.72	34.1	0.30	0.10

Table 3.19. Model 2 estimates of variance components ($x10^2$), importance of dominance variance, and additive genetic coefficient of variation (CV_A) for height using data pooled across sites at each of the four ages, based on analysis using the individual tree model.

Age (years)	σ_A^2	σ_D^2	σ_{Pt}^2	σ_E^2	D_A	D_P	CV_A
1.5	1.74	0.239	0.621	4.47	0.14	0.03	14.0
9.5	140	23.3	12.0	112	0.17	0.08	8.66
13.5	128	45.4	20.3	364	0.35	0.08	6.00
22.5	71.5	28.6	30.6	168	0.40	0.09	2.94

Table 3.20. Model 2 estimates of variance components ($x10^2$), importance of dominance variance, and additive genetic coefficient of variation (CV_A) for stem straightness using data pooled across sites at each of the four ages, based on analysis using the individual tree model.

Age (years)	σ_A^2	σ_D^2	σ_{Pt}^2	σ_E^2	D_A	D_P	CV_A
1.5	0.341	0.860	1.07	33.13	2.53	0.02	1.51
9.5	5.31	4.03	4.02	38.56	0.76	0.08	5.40
13.5	5.94	0.95	2.90	55.76	0.16	0.01	5.97
22.5	9.43	2.97	4.37	26.46	0.31	0.07	5.71

The heritability estimates for height and straightness using the individual tree model are given in Tables 3.21 and 3.22, respectively. The individual site estimates were, in general, higher than the individual site estimates from sib covariance analyses. This might indicate less bias from the use of all available information by the individual tree model. The estimates for height ranged from 0.14 to 0.73. Those for straightness generally ranged from 0.01 to 0.33, except for an unusually high estimate of 0.85

obtained at site C at 22.5 years of age (Table 3.22). For each trait, estimates across the four sites at each assessment age are unlikely to be significantly different from each other, since their differences are less than two standard deviations, other than in the case of straightness at 22.5 years.

Table 3.21. Estimates of narrow sense heritability for height at four sites, ratio of plot to phenotypic variances (c^2) and log likelihood ratio test for differences between the two individual tree models.

Age (yrs)	Site	Model 1	Model 2		LRT
		h^2 (se)	h^2 (se)	c^2 (se)	
1.5	A	0.23 (0.08)	0.23 (0.08)	0.06 (0.01)	44.3**
	B	0.32 (0.09)	0.32 (0.09)	0.03 (0.01)	16.5**
	C	0.35 (0.10)	0.35 (0.12)	0.05 (0.01)	26.6**
	D	0.14 (0.05)	0.14 (0.05)	0.02 (0.01)	5.2*
9.5	A	0.73 (0.14)	0.73 (0.27)	0.02 (0.01)	1.5ns
	B	0.40 (0.14)	0.41 (0.14)	0.04 (0.05)	0ns
	C	0.59 (0.14)	0.59 (0.12)	0.00 (0.02)	0ns
	D	0.40 (0.12)	0.39 (0.13)	0.00 (0.02)	0ns
13.5	A	0.24 (0.08)	0.24 (0.08)	0.00 (0.02)	0.1ns
	B	0.32 (0.10)	0.32 (0.11)	0.01 (0.03)	0.7ns
	C	0.49 (0.12)	0.49 (0.13)	0.00 (0.02)	0ns
	D	0.23 (0.08)	0.23 (0.11)	0.01 (0.04)	0.7ns
22.5	A	0.40 (0.14)	0.42 (0.16)	0.01 (0.01)	0.5ns
	B	0.26 (0.10)	0.26 (0.11)	0.04 (0.05)	1.0ns
	C	0.39 (0.15)	0.39 (0.19)	0.04 (0.09)	2.6ns
	D	0.40 (0.12)	0.40 (0.13)	0.09 (0.05)	3.4ns

Table 3.22. Estimates of narrow sense heritability for stem straightness at four sites, ratio of plot to phenotypic variances (c^2), and log likelihood ratio test for differences between the two individual tree models.

Age (yrs)	Site	Model 1	Model 2		LRT
		h^2 (se)	h^2 (se)	c^2 (se)	
1.5	A	0.02 (0.01)	0.02 (0.02)	0.03 (0.01)	11.8**
	B	0.01 (0.01)	0.01 (0.02)	0.02 (0.04)	6.3*
	C	0.08 (0.03)	0.08 (0.03)	0.02 (0.01)	2.7ns
	D	0.02 (0.02)	0.02 (0.02)	0.01 (0.01)	1.3ns
9.5	A	0.25 (0.07)	0.25 (0.08)	0.06 (0.02)	10.4**
	B	0.22 (0.10)	0.22 (0.17)	0.03 (0.08)	0ns
	C	0.16 (0.07)	0.16 (0.14)	0.03 (0.02)	2.2ns
	D	0.18 (0.08)	0.18 (0.08)	0.02 (0.02)	0.7ns
13.5	A	0.16 (0.05)	0.16 (0.05)	0.00 (0.02)	0ns
	B	0.17 (0.07)	0.18 (0.07)	0.06 (0.03)	3.3ns
	C	0.24 (0.13)	0.27 (0.15)	0.03 (0.02)	7.0**
	D	0.23 (0.08)	0.23 (0.12)	0.00 (0.03)	0.1ns
22.5	A	0.16 (0.10)	0.16 (0.13)	0.00 (0.20)	0ns
	B	0.33 (0.11)	0.33 (0.12)	0.04 (0.05)	0.7ns
	C	0.85 (0.20)	0.83 (0.19)	0.06 (1.19)	1.8ns
	D	0.32 (0.11)	0.32 (0.11)	0.04 (0.05)	0.5ns

Trends over time in heritability estimates are shown in Figure 3.4. The heritability estimates for height peaked at 9.5 years. At 13.5 years, it decreased to the same level as that at 1.5 years and remained constant thereafter. The peak coincides with a peak in additive genetic variance and the low in dominance variance. The

heritability estimates for straightness increased sharply between 1.5 and 9.5 years and again between 13.5 and 22.5 years, but were comparable at 9.5 and 13.5 years.

Correlations between heritability estimates and standard deviation or mean were less than 0.5 for both traits, indicating that heritability could not be predicted well from the corresponding variance and mean.

Log likelihood ratio tests

The log likelihood ratios (Tables 3.21 and 3.22) suggest that differences between the two models at each site were significant only at age 1.5 years for height. Pooling data over sites at each age resulted in highly significant differences between the models (Tables 3.23 and 3.24), indicating that Model 2, which included the common environment effect, better fitted the data.

Table 3.23. Estimates of narrow sense heritability for height, ratio of plot to phenotypic variances (c^2), and log likelihood ratio test for differences between the two individual tree models, using pooled data over sites at each age.

Age (yrs)	Model 1	Model 2		LRT
	h^2 (se)	h^2 (se)	c^2 (se)	
1.5	0.24 (0.07)	0.24 (0.11)	0.09 (0.02)	358.5**
9.5	0.50 (0.15)	0.50 (0.12)	0.04 (0.01)	26.4**
13.5	0.24 (0.06)	0.24 (0.06)	0.04 (0.01)	16.0**
22.5	0.24 (0.08)	0.23 (0.07)	0.10 (0.03)	11.3**

Table 3.24. Estimates of narrow sense heritability for stem straightness, ratio of plot to phenotypic variances (c^2), and log likelihood ratio test for differences between the two individual tree models, using pooled data over sites at each age.

Age (yrs)	Model 1	Model 2		LRT
	h^2 (se)	h^2 (se)	c^2 (se)	
1.5	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	40.8**
9.5	0.11 (0.04)	0.11 (0.04)	0.08 (0.01)	54.2**
13.5	0.09 (0.01)	0.09 (0.03)	0.04 (0.01)	17.7**
22.5	0.21 (0.07)	0.22 (0.08)	0.10 (0.04)	10.2**

All the log likelihood ratio tests for the difference between the joint likelihood and analyses of pooled data across sites were significant (Table 3.25). This implies that fitting a model to the pooled data is not appropriate, as a consequence of heterogeneity of variances between the sites. The most appropriate fits, therefore, are those for individual sites.

Table 3.25. Difference between the joint likelihood of the four individual site analyses (sum) and maximum likelihood values for the analyses of data pooled across the sites, and the log likelihood ratio test for the adequacy of pooling data across sites.

Age	Height		Straightness	
	Model 1	Model 2	Model 1	Model 2
1.5	250.1**	117.1**	382.6**	373.2**
9.5	59.5**	47.0**	101.0**	80.6**
13.5	163.2**	155.9**	256.2**	245.4**
22.5	52.9**	50.5**	90.9**	87.3**

Likelihood profiles

Table 3.26 gives the confidence intervals from the DFREML approximation and that from the likelihood profile. The results show that the likelihood profile was not symmetrical, as assumed by the DFREML approximation. Therefore, biases are introduced by the DFREML approximation.

Table 3.26. Confidence intervals for heritability estimates obtained from using the standard error from the DFREML (i.e. the second derivative of the likelihood) and those from likelihood profiles.

	Confidence interval	
Heritability	Second derivative	Likelihood profile
0.24	0.10-0.38	0.14-0.34
0.50	0.20-0.80	0.31-0.81
0.24	0.12-0.36	0.14-0.43
0.24	0.08-0.40	0.13-0.46

Maternal and paternal effects

Maternal and paternal effects were fitted as additional random effects; these were found to be very small ($c^2 < 1x10^{-5}$ for both traits), and were therefore not included in the model.

3.4 Discussion

The advantage of using the individual tree model in DFREML is that it gives one estimate of heritability by combining all sources of genetic information appropriately. In contrast, models available in GENSTAT REML only allow fitting of sib covariance and, therefore, give two estimates, and combining them appropriately when they are correlated and include the same genotypes is complex. A further problem is that pooling heritability estimates using the sampling variances, as in this study, biases the pooled estimate towards the smaller estimate since the sampling variance of heritability depends on the heritability itself: the bigger the heritability, the bigger the

sampling variance of the heritability. The estimates from individual tree model were consistently higher than the pooled estimates from sib covariance analyses, indicating less bias with the individual tree model. A potential problem of using maximisation methods like DFREML is convergence to points other than the global maximum. In this study, different starting values were used to ensure a global maximum.

Despite the fact that this was the first reported use of the individual tree model in *P. taeda*, the heritability estimates for height were within the range reported previously for this species (Table 2.2) and consistent with those reported for other pine species (Barnes 1992a and b, Cotterill *et al.* 1987, Pswarayi *et al.* 1996). However, this is to be expected from results in early generations, and in the absence of information to correct for the effects of selection of parents. Here, heritability and additive variance for height increased with age from 1.5 years to 9.5 years, and then decreased with age. Although the trend differed from those observed by Franklin (1979) and Foster (1986) in *P. taeda*, it agreed well with that reported by Balocchi *et al*. (1993) in *P. taeda*, who found heritability peaking at 14 years of age in a slower growing progeny test. Whilst this study and that of Balocchi *et al.* (1993) differ in the age of maximum heritability, the estimates were maximum at the same mean height, suggesting a possible link between mean height and heritability estimate. This is consistent with findings of Borralho *et al.* (1992a), but contrasts with results reported by Borralho *et al*. (1992b) and Woolaston *et al*. (1990) who found no relationship between trends in heritability estimates for height and growth rate in *Eucalyptus globulus* and *P. caribaea*, respectively. The change in heritability in long rotation crops such as tree is not surprising since genes involved in growth may change with age (Namkoong *et al*. 1988), and these changes may be related to different growth phases (Franklin 1979). In animals, this change in heritability with age was also attributed to the fact that the trait may change genetically with age (Visscher *et al.* 1991), probably related to different growth phases as reported for trees. These growth phases might be due to changing influences of maternal effects in animals and to a lesser extend in trees and to nursery or competition effects in trees. Changes in heritability with age here may also be attributed to thinning and other management practices. For example, height at 9.5 years was assessed on felled trees, and hence more accurately assessed than that at 13.5 and 22.5

years which was assessed on standing trees.

Heritability estimates for straightness in this study increased with age. a trend also reported by Shelbourne and Stonecypher (1971) in *P. taeda*, although their estimates were much higher. The estimates reported in this study are consistent with those reported by Barnes (1992a and b), Pswarayi *et al.* (1996), Cornelius (1994) and Cotterill *et al.* (1987) but lower than those reported by Woolaston *et al.* (1990) in other conifer species. In general, the estimates of heritability of stem straightness were lower than those of height. Parameter estimates for stem straightness have not been reported as frequently as growth traits, and estimates in the literature are variable: some heritability estimates for straightness have been higher than those for height in *P. taeda* (eg. Matziris and Zobel 1973), whereas others found them to be lower (Barnes 1992a and b, Cotterill *et al.* 1987, Pswarayi *et al.* 1996, Raymond and Cotterill 1990) or equivalent (Burdon *et al.* 1992a, Cornelius 1994) in other conifers. Therefore, results from this study and those from many other studies do not support the contention (e.g. Zobel and Talbert 1984) that heritability estimates of growth traits are lower than those for stem straightness. Estimates depend on how straightness was measured (Cotterill *et al.* 1987, Raymond and Cotterill 1990). Estimates originating from use of an absolute scale, as in this study, are lower than those originating from a site specific scale. The low estimates reported in this study are consistent with those reported by Barnes (1992a and b) and Pswarayi *et al.* (1996) using the same absolute scale on *P. patula* and *P. elliottii*, respectively. Major problems related to the use of the relative scale are that if a trait is poorly expressed at a site this method will indicate large genetic differences, when if fact they are absent, as reported by Williams and Lambeth (1989), and results are not comparable across sites adversely affecting deployment decisions. Estimates also depend on the number of categories with scales with few categories such as 3-point scale having lower heritability estimates than those with moderate categories such as 6-point scale (Raymond and Cotterill 1990). The low heritability estimates with a 3-point scale was attributed to the limited range of the scale which failed to detect genetic differences. Raymond and Cotterill (1990) also found that assessor error was increased with a scale with many categories (9-point scale) resulting in low heritability estimates. Another possible reason for the low estimates of heritability might lie in that the scale

used had a mid value which assessors might use very frequently for what appears to be average trees, reducing the heritability. There was very little evidence of additive or dominance variance for straightness at 1.5 years, indicating that environment effects were the major determinant of straightness at this age. This can be attributed to large measurement error because straightness is difficult to score at a young age. Also, trees at this young age are likely to be more affected by environmental variation, resulting in lower heritability estimates than at older ages. The high heritability estimates for straightness at 22.5 years, and the low estimates at 1.5 years, appear to indicate that stem straightness might be easier to measure on large trees than on small ones, an observation also made by Dean (1990).

The phenotypic variance was expected to increase with age, mainly due to scale effects. Therefore, the decrease in the phenotypic variance between 13.5 years and 22.5 years was unexpected. Although the thinning was systematic, the number of trees remaining per plot at 22.5 years was only 1-2 compared to 10 at 1.5 years and 5 at 9.5 and 13.5 years. This reduction in the number of trees reduced the within plot variance substantially, and would explain the marked reduction in the phenotypic variance between 13.5 and 22.5 years. This results confirms Matheson and Raymond's (1984a) observation that the phenotypic variance may be affected by thinning. Trees recovering from growth restriction such as those in the intermediate and suppressed classes might exhibit more rapid (compensatory) growth than trees in the dominant and codominant classes upon thinning, thereby reducing the phenotypic variance. As canopy closes, trees in the intermediate and suppressed classes might be expected to grow slower than those in the dominant and codominant classes due to competition for light and nutrients- thereby increasing the phenotypic variance. Therefore, thinning effects could partly explain the trend of phenotypic variance with age.

The variance components and heritability estimates presented from data pooled over sites showed evidence of heterogeneity of variances over sites. Heritability estimates were slightly more accurately estimated from pooled data because they were estimated from a larger sample, as evidenced by the smaller standard errors associated with them. However, pooling data may give biased estimates in the presence of heterogeneity of variances. The log likelihood ratio tests (Table 3.25) indicate that

pooling data from the different sites is likely to give heritability estimates which are substantially biased as a result of heterogeneity of variances and/or genotype x environment interactions. The lower heritability estimates from pooled data imply that predicted gain would be less than that from the individual site estimates. Therefore, estimates from the pooled data will only be used subsequently in BLUP evaluation after correcting for heterogeneity of variances if there are no genotype x environment interactions due to rank changes using appropriate methods such as that described by Visscher *et al.* (1991). However, if GE is present, and it is due to rank changes, combining the data as in the pooled analyses will remain inappropriate. Presence of genotype x environment interactions due to rank changes will be investigated in a subsequent chapter.

The ratio of dominance to additive variance for height was less than one at all ages in the pooled analysis, indicating that dominance variance was of lesser importance. These ratios differed from those of Balocchi *et al.* (1993) and Foster and Bridgwater (1986), who reported dominance variance to be considerably higher than additive variance at young ages (less than 6 years) in *P. taeda*. The ratios were consistent with those reported for *P. elliottii* grown in Zimbabwe (Pswarayi *et al.* 1996). For straightness, dominance variance was much greater than additive variance at 1.5 years; thereafter the relationship was reversed. The pattern of the ratios in straightness differed from those reported by Pswarayi *et al.* (1996), who found dominance to be less than additive variance at 5 and 15 years of age but more than the additive variance at 8 years of age.

Common environmental effects for height and straightness were substantially less than the heritability estimates for both traits, indicating a low intra-class correlation between trees within a plot. Excluding the plot term from the model resulted in inflated estimates of dominance variance, but the heritability estimates and their standard errors were not sensitive to inclusion of plots in the model. Therefore, the true level of dominance variance was obtained when plot was included as an additional uncorrelated random effect. This is because family groups were planted within single plots at each site.

Chapter 4

BIVARIATE PARAMETER ESTIMATES FOR HEIGHT AND STRAIGHTNESS

4.1 Introduction

A knowledge of genetic parameters of growth and wood quality traits are necessary to estimate accurate breeding values, to combine different traits in selection, to predict genetic response to selection and to predict the optimum selection age. To achieve these objectives, genetic and phenotypic correlations are of major importance. For example, juvenile-mature genetic correlations are essential for estimation of genetic gain from early selection and trait-trait correlations are important for construction of multi-trait selection indices. Early selection is an indirect selection where performance at a young age is used as an indicator of mature age performance (Burdon 1989). Early selection may result in shortening of the generation interval, increase in gain per unit of time, may lead to easier incorporation of changes in market demands and to savings in testing. On the other hand, multi-trait selection involving height and straightness, for example, may improve the recovery of the wood (i.e. the economic value) if the traits are positively correlated. However, if the traits are negatively correlated a restricted multi-trait selection index may be appropriate.

Most studies in *P. taeda* have indicated high positive age-age genetic correlations for height (Foster 1986, Franklin 1979, Lambeth *et al.* 1983). Further, Lambeth (1980) developed a generalized model which showed that phenotypic correlations between heights at different ages were predictable based on the natural logarithm of the ratio of the ages. This model and similar logarithm type of models have been widely used by tree breeders to make decisions on the optimum selection age (e.g. Lambeth 1980, King and Burdon 1991, Magnussen 1988, McKeand 1988, Riemenschneider 1988, Xie and Ying 1996). Lambeth (1980) suggested that age-age

phenotypic correlations were fair estimates of their genetic counterparts. Recent work has cast doubt upon the critical assumption made by Lambeth concerning the similarity of genetic and phenotypic correlations, and there is increasing evidence which indicates that genetic correlations are higher than phenotypic correlations (Lambeth *et al.* 1983. Newman and Williams 1991, Riemenschneider 1988, Pswarayi *et al.* 1996), implying that the Lambeth model underestimates genetic correlations. This underestimation of genetic correlations will result in underestimates of gain, and may consequently affect predictions of the optimum selection age.

In contrast, there is absence of reports of age-age correlations for straightness in *P. taeda*. Studies in other pine species indicated that there were strong positive age-age genetic correlations in straightness (Pswarayi *et al.* 1996). Age-age correlations for straightness are important because further improvement in quality of end product might be obtained by including straightness in an index. Inclusion of straightness in an index will depend on its heritability and its genetic correlation with height. Unfortunately, the estimates of genetic correlations between height and straightness are few in *P. taeda*. Those reported for other pine species in Zimbabwe were generally positive and lower than 0.5 (Barnes 1992a and b, Pswarayi *et al.* 1996), and those reported for other tropical pine species were also low (Woolaston *et al.* 1990).

The objective of this Chapter is to estimate age-age correlations of height and straightness and the trait-trait correlation between them. Height was chosen as a selection criteria because it is a good predictor of volume even at young ages (Foster 1986, Lambeth *et al.* 1983). The estimated age-age correlations for height will be modelled over time using regression models. Further, genetic correlations will be estimated using covariance functions (Kirkpatrick *et al.* 1990).

4.2 Materials and Methods

Statistical analyses

Data were analysed using bivariate individual tree model DFREML (Meyer 1989). Two programs were available for bivariate analyses. The first program performed

analyses when the same model of analysis was applied to both traits: same fixed and random effects, while the other program allowed models differing in both the fixed and random effects to be fitted to the two traits. The first program was used for estimating correlations between height and straightness since both traits were measured on the same trees and the same models were applied to both traits. The other program was used to estimate age-age correlations because there were different levels of random tree effects at different ages due to thinning. The assessments at different ages within a trait were treated as different traits. Analysis were performed within each site, and on data pooled over sites in order to improve the precision of the estimates. For the bivariate analyses involving 9.5 years and older ages, the two traits were assessed on separate, but genetically related trees since assessments at 9.5 years were carried out on trees that were thinned. Therefore, genetic covariances exist between the traits but there are no environmental covariances.

The following bivariate tree model was used:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \qquad (4.1)$$

where: $\mathbf{y}_1$, $\mathbf{y}_2$ = the vector of observations for traits 1 and 2, respectively,

$\mathbf{b}_1$, $\mathbf{b}_2$ = the vector of fixed effects for traits 1 and 2, respectively,

$\mathbf{a}_1$, $\mathbf{a}_2$ = the vector of random tree (additive genetic) effects for traits 1 and 2, respectively,

$\mathbf{c}_1$, $\mathbf{c}_2$ = the vector of additional uncorrelated random effects for traits 1 and 2, respectively,

$\mathbf{X}_1$, $\mathbf{X}_2$ = the incidence matrix for fixed effects for traits 1 and 2 respectively,

$\mathbf{W}_1$, $\mathbf{W}_2$ = the incidence matrix for additional random effects for traits 1 and 2, respectively,

$\mathbf{Z}_1$, $\mathbf{Z}_2$ = the incidence matrix for additive direct effects for traits 1 and 2, respectively,

$\mathbf{e}_1$, $\mathbf{e}_2$ = the vector of residual effects for traits 1 and 2, respectively.

The variance-covariance structure of the random effects of the bivariate tree model was as follows:

$$V\begin{bmatrix} a_1 \\ a_2 \\ c_1 \\ c_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} A\sigma^2_{a1} & Acov_{a12} & 0 & 0 & 0 & 0 \\ Acov_{a21} & A\sigma^2_{a2} & 0 & 0 & 0 & 0 \\ 0 & 0 & I\sigma^2_{c1} & Icov_{c12} & 0 & 0 \\ 0 & 0 & Icov_{c21} & I\sigma^2_{c2} & 0 & 0 \\ 0 & 0 & 0 & 0 & I\sigma^2_{e1} & Icov_{e12} \\ 0 & 0 & 0 & 0 & Icov_{e21} & I\sigma^2_{e2} \end{bmatrix} \quad (4.2)$$

where: σ^2_{a1}, σ^2_{a2} = the direct additive genetic variances for traits 1 and 2, respectively,

σ^2_{c1}, σ^2_{c2} = the additional random effect variances for traits 1 and 2, respectively,

σ^2_{e1}, σ^2_{e2} = the residual variances for traits 1 and 2, respectively,

cov_{a12} = the direct genetic covariance between traits 1 and 2,

cov_{c12} = the additional random effect covariance between traits 1 and 2, and

cov_{e12} = residual covariance.

For the individual site analyses, the fixed effect was the replicate, and for the pooled analyses, the fixed effects were replicate and site. For all analyses, the additional uncorrelated random effect was the family. While for univariate analyses, blocks were fitted as additional random effect, it was not possible to do so in the bivariate analyses because the programme restrictions allowed fitting only one extra random effect. This problem was overcome by pre-adjusting the data for the block effects before each bivariate analysis. The first starting values of variances and covariances were estimated from the univariate analyses. Different starting values from these initial ones were used to ensure that a global rather than a local maximum had been reached.

The phenotypic correlations are the sum of the genetic and environmental components, the correlations being weighted by the heritability, and the proportion of variance due to environment and non-additive genetic effects, respectively:

$$r_P = h_1 h_2 r_A + e_1 e_2 r_e \quad (4.3)$$

where r_P = phenotypic correlation, r_A= additive genetic correlation, r_e= environment

correlations, h_1, h_2 = the square root of heritability, e_1, e_2 = square root of the proportion of phenotypic variance due to environmental effects plus non-additive effects, for traits 1 and 2 (Falconer 1989). The sampling errors of the genetic correlations were derived using the method of Robertson (1959):

$$se(r_A) = \frac{(1 - r_A^2)}{\sqrt{2}} \text{ x } \frac{\sqrt{se(h_1^2)se(h_2^2)}}{h_1 h_2} \qquad (4.4)$$

where:

$se(r_A)$, $se(h^2_1)$, $se(h^2_2)$ are standard errors of the genetic correlation and heritability estimates for traits 1 and 2, respectively,

h_1, h_2 are square root of the heritability estimates for traits 1 and 2, respectively.

Robertson's formula gives an approximate estimate of the standard error of the genetic correlation. Therefore, in order to get accurate confidence intervals for the genetic correlation estimates from the pooled data, likelihood profiles were plotted by fixing the genetic correlation to different values, and the likelihood maximised with respect to all the other parameters. The 95% confidence interval was obtained by dropping 1.92 ($0.5\chi^2_{1, 0.05}$) from the maximum log likelihood (Wetherill 1981). Estimation of standard errors of genetic correlations in this way is computationally demanding, and therefore only a few likelihood profiles were plotted to verify the estimates from Robertson's formula.

A predictive model was fitted by regressing phenotypic correlations for height on the natural logarithm of the ratio of the younger age to the older age (LAR). Six models for predicting genetic correlations were tested. The relationship between genetic correlations, and the natural logarithm of the ratio of the younger age to the older age (LAR), was developed by Lambeth (1980); r_P was equated to r_g, and r_P predicted by:

Model 1: $r_P = 1.02 + 0.308\log_e$ (younger age/older age) (4.5)

The equality of r_P and r_g was removed:

Model 2: $r_g = \beta_0 + \beta_1$ (LAR), (4.6)

where β_0, β_1 were derived from estimates of r_g at 1.5, 9.5, 13.5 and 22.5 years.

In order to improve Model 2, Model 3 was derived by including additional genetic correlations involving pooled data from two progeny tests established in Zimbabwe in 1976, and assessed at 7.5 and 12.5 years of age (Gwaze 1995):

Model 3: $$r_g = \beta_0 + \beta_1 (\text{LAR}), \qquad (4.7)$$

where β_0, β_1 were derived from estimates of r_g at 1.5, 7.5, 9.5, 12.5, 13.5 and 22.5 years.

Model 4 was derived as for Model 2, but log of height ratio used as a predictor:

Model 4: $$r_g = \beta_0 + \beta_1 \log_e (\text{Height ratio}) \qquad (4.8)$$

Model 5 was derived as for Model 2, but age difference was used as a predictor:

Model 5: $$r_g = \beta_0 + \beta_1 (\text{Age difference}) \qquad (4.9)$$

Additive covariance functions for height, using orthogonal polynomials with symmetric coefficients (Kirkpatrick *et al.* 1990) reported in Chapter 5, were also used to predict trends of genetic correlations. The additive covariance matrix which was used to estimate the covariance functions was estimated from the bivariate DFREML runs. The covariance between records taken at ages t_1 and t_2 is (Model 6):

Model 6: $$\mathrm{T}(t_1,t_2) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \hat{C}_{ij} \phi_i(t_1) \phi_j(t_2), \qquad (4.10)$$

where T is the covariance function, k is the order of fit, $\hat{\mathrm{C}}$ symmetric coefficient matrix associated with the covariance function, and ϕ orthogonal polynomials.

Then,

$$r_g = \frac{cov_A(t_1,t_2)}{\sqrt{var_A(t_1)}\sqrt{var_A(t_2)}} \qquad (4.11)$$

For bivariate analyses, more than one heritability estimate is obtained for each trait at each age. A simple average of the estimates for each trait was obtained.

4.3 Results

Age-age correlations at individual sites

Tables 4.1 and 4.2 show the age-age genetic and phenotypic correlations, respectively, for height. Genetic correlations between early height growth (1.5 years) and mature age (22.5 years) were lowest; the lowest values were at sites A and B, where they were 0.10 and 0.31, respectively. Genetic correlations at other sites and between other ages were of moderate to high magnitude (>0.6), with the majority being greater than 0.8. Phenotypic correlations were consistently lower than the genetic correlations. Estimation of phenotypic covariance assumes that the two traits are measured on the same individual. Because height at 9.5 years was assessed on trees that were thinned, and hence heights at 9.5 years and older ages were assessed on different individuals, the phenotypic correlations could not be estimated. No site had consistently high or low correlations.

Table 4.1. Estimated age-age genetic correlations for height at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
HT1.5,HT9.5	0.88	0.73	0.96	0.81
HT1.5,HT13.5	0.80	0.61	0.87	1.00
HT1.5,HT22.5	0.10	0.31	0.75	0.60
HT9.5,HT13.5	0.81	1.00	0.97	0.94
HT9.5,HT22.5	0.77	0.99	0.82	0.61
HT13.5,HT22.5	0.96	0.77	0.96	0.96

Table 4.2. Estimated age-age phenotypic correlations for height at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
HT1.5,HT9.5	0.47	0.46	0.45	0.37
HT1.5,HT13.5	0.21	0.37	0.42	0.23
HT1.5,HT22.5	0.08	0.14	0.26	0.18
HT9.5,HT13.5	-	-	-	-
HT9.5,HT22.5	-	-	-	-
HT13.5,HT22.5	0.58	0.53	0.78	0.84

Tables 4.3 and 4.4 show the age-age genetic and phenotypic correlations, respectively, for straightness. The genetic correlations between straightness at 1.5 and the older ages ranged from -0.89 to 0.21. Those between straightness at 9.5 years and 22.5 years were greater than 0.7 at all sites. Those between straightness at 13.5 and 22.5 years were greater than 0.8 at all sites, except at site C, which was 0.28. The phenotypic correlations for straightness were generally lower than corresponding genetic correlations. Like the genetic correlations, the phenotypic correlations between 1.5 years and other ages were very low (less than 0.11), but all were positive. Phenotypic correlations between other ages were variable, ranging from -0.01 to 0.67.

Table 4.3. Estimated age-age genetic correlations for straightness at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
ST1.5,ST9.5	0.04	-0.09	0.21	-0.07
ST1.5,ST13.5	-0.89	0.08	0.14	0.20
ST1.5,ST22.5	-0.55	-0.11	0.11	0.04
ST9.5,ST13.5	-0.01	0.10	1.00	0.79
ST9.5,ST22.5	0.71	0.73	0.77	0.95
ST13.5,ST22.5	0.94	0.84	0.28	1.00

Table 4.4. Estimated age-age phenotypic correlations for straightness at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
ST1.5,ST9.5	0.03	0.06	0.08	0.09
ST1.5,ST13.5	0.04	0.01	0.11	0.10
ST1.5,ST22.5	-0.01	0.01	0.05	0.06
ST9.5,ST13.5	-	-	-	-
ST9.5,ST22.5	-	-	-	-
ST13.5,ST22.5	0.38	0.42	0.15	0.67

Correlations between height and straightness

Genetic and phenotypic correlations between height and straightness are shown in Tables 4.5 and 4.6, respectively. Genetic correlations ranged from -0.64 to 0.64. They were lowest between height and straightness at 1.5 years of age, at all sites, and highest between height and straightness at age 22.5 years at sites A and B. Phenotypic correlations between height and straightness were all positive, indicating a favourable relationship between straightness and height, and of low magnitude (range, 0.04 to 0.42).

Table 4.5. Estimated trait-trait genetic correlations for height and straightness at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
HT1.5,ST1.5	-0.64	0.04	-0.42	0.13
HT9.5,ST9.5	0.51	-0.04	0.28	0.38
HT13.5,ST13.5	0.13	0.47	0.12	0.19
HT22.5,ST22.5	0.64	0.63	0.30	0.12

Table 4.6. Estimated trait-trait phenotypic correlations for height and straightness at four sites and four ages, based on analysis by bivariate individual tree model DFREML.

Traits	Site A	Site B	Site C	Site D
HT1.5,ST1.5	0.21	0.20	0.13	0.20
HT9.5,ST9.5	0.28	0.42	0.24	0.34
HT13.5,ST13.5	0.21	0.23	0.35	0.39
HT22.5,ST22.5	0.37	0.24	0.19	0.04

Correlations using pooled data over sites

Results of analyses of pooled data over sites are shown in Tables 4.7 and 4.8. Age-age genetic correlations for height were high, ranging from 0.76 to 0.97. As the age interval increased, the genetic correlations for height decreased. The precision of the genetic correlations increased with increasing values of the correlation itself, as expected. The phenotypic correlations were lower than the genetic correlations (range, 0.21 to 0.80). In general, the phenotypic correlations were lower than 0.50 except that between 13.5 and 22.5 years which was 0.80. In general, age-age genetic correlations for straightness were lower than those for height (range, -0.05 to 0.94) with those involving 1.5 years being lowest (range, -0.05 to 0.21). Phenotypic correlations for straightness were low to moderate (range, 0.02 to 0.55), and lower than genetic correlations.

Genetic correlations between height and straightness were low to moderate (range, -0.28 to 0.66) with those involving straightness at 1.5 years being mainly negative. The highest correlation at one age was between height and straightness at 22.5 years (0.52); and the highest correlation between ages was between height at 22.5 years and straightness at 13.5 years (0.66). The other genetic correlations were below 0.45. Phenotypic correlations between height and straightness were also low (range, -0.19 to 0.48), and were not significantly different from the genetic correlations.

Table 4.7. Estimated genetic (below the diagonal) and phenotypic correlations (above the diagonal), and heritability estimates (on the diagonal) for height and straightness using data pooled across sites at the four sites, based on analysis by bivariate individual tree model DFREML. Standard errors of the genetic correlations are in parenthesis.

	HT1.5	HT9.5	HT13.5	HT22.5	ST1.5	ST9.5	ST13.5	ST22.5
HT1.5	**0.22**	0.48	0.31	0.21	0.27	0.06	0.12	0.06
HT9.5	0.93	**0.50**	-	-	0.09	0.12	0.48	-0.19
	(0.03)							
HT13.5	0.85	0.96	**0.22**	0.80	0.09	0.35	0.33	0.19
	(0.07)	(0.01)						
HT22.5	0.76	0.85	0.97	**0.26**	0.05	-0.04	0.16	0.21
	(0.10)	(0.05)	(0.01)					
ST1.5	-0.08	-0.28	-0.17	0.04	**0.01**	0.06	0.06	0.02
	(0.48)	(0.32)	(0.34)	(0.36)				
ST9.5	0.14	0.13	0.17	0.36	0.20	**0.12**	-	-
	(0.28)	(0.21)	(0.21)	(0.19)	(0.41)			
ST13.5	0.38	0.45	0.25	0.66	0.11	0.94	**0.09**	0.55
	(0.24)	(0.16)	(0.19)	(0.12)	(0.40)	(0.03)		
ST22.5	0.14	0.22	0.28	0.52	-0.05	0.66	0.92	**0.20**
	(0.28)	(0.20)	(0.20)	(0.16)	(0.43)	(0.15)	(0.04)	

The environmental correlations ranged from low negative to low positive (Table 4.8). As for phenotypic correlations, environmental covariance assumes that the two traits are measured on the same individuals. Therefore, environmental correlations between 9.5 years and subsequent years could not be estimated since height was assessed on different individuals in different years.

Table 4.8. Estimated environmental correlations for height and straightness using data pooled across sites at the four sites, based on analysis by bivariate individual tree model DFREML.

	HT9.5	HT13.5	HT22.5	ST1.5	ST9.5	ST13.5	ST22.5
HT1.5	0.25	0.16	0.04	0.31	0.05	0.08	0.03
HT9.5		-	-	0.16	0.14	0.57	-0.44
HT13.5			0.74	0.11	0.38	0.35	0.15
HT22.5	0.05	-0.15	0.05	0.09			
ST1.5		0.06	0.07	0.03			
ST9.5			-	-			
ST13.5				0.49			

Loglikelihood profiles

Two loglikelihood profiles were plotted for genetic correlations and, in all cases, the standard errors of the genetic correlations estimated using Robertson's formula were very similar to those estimated using the 2nd derivative of the loglikelihood. For example, the difference between the standard error of the genetic correlation between straightness at 13.5 years and height at 22.5 years was only 0.02, with that estimated using the Robertson's formula being lower. The loglikelihood profile was not symmetrical, although its departure from symmetry was not great. This resulted in the confidence interval from the loglikelihood profile being slightly different from that estimated using Robertson's formula.

Modelling age-age correlation for height using Lambeth type model

Details of the results of fitting age-age genetic and phenotypic models are presented in Table 4.9. The following regression model was obtained for genetic correlations (fitting Model 2) (Figure 4.1):

$$r_A = 0.98 + 0.065\ (LAR) \qquad (R^2 = 0.61). \qquad (4.12)$$

The intercept was less than unity and this might be attributed to random measurement errors. Since the genetic correlations between heights assessed at the same age can be expected, on statistical grounds, to be unity, the regression model was forced to have a constant equal to unity and the new model was:

$$r_A = 1.00 + 0.076\ (LAR). \qquad (4.13)$$

For Model 2, constraining the genetic correlations between height assessed at the same age to be 1 resulted in a fit which was not significantly different from the unconstrained one (i.e.$\beta_0 = 1$, $P > 0.05$). The estimated regression slope was significantly lower than that of Lambeth's ($P < 0.05$). As a result, predictions of the genetic correlations from the Lambeth model declined more rapidly than those from the genetic correlation model as the age differential increased (Figure 4.1). Table 4.10 shows the observed genetic correlations together with the predictions from the genetic correlation model and that of Lambeth. The good fit with the new model would be expected as it is obtained from the data presented, but the large divergence between the observed and predictions from the Lambeth model is clear, particularly for genetic correlations involving 1.5-year height assessment.

The following model between phenotypic correlations and the log of age ratio was highly significant:

$$r_p = 0.94 + 0.274\ (LAR) \qquad (R^2 = 0.99), \qquad (4.14)$$

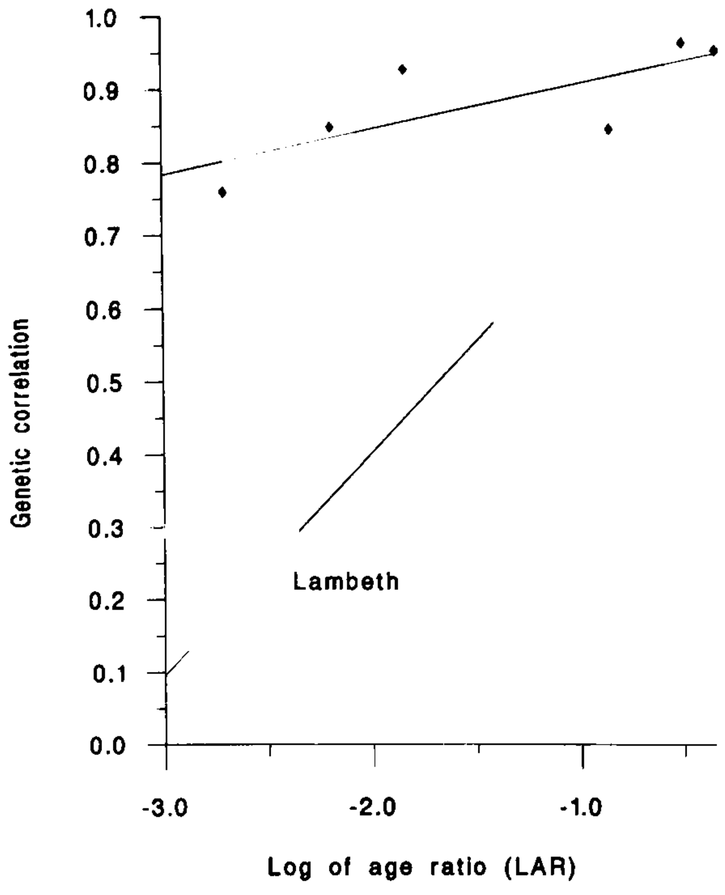
and not significantly different from that of Lambeth (Figure 4.1). Unlike Lambeth's (1980) findings, genetic and phenotypic correlations between height at age 1.5 and at subsequent ages followed the same linear relation with LAR as did other age combinations.

Table 4.9. Results of fitting genetic and phenotypic correlation models for height.

Type of Model	Regression coefficient		Residual	df	R^2
	Intercept±sd	Slope±sd	mean sq.		
Logarithmic (age ratio)-r_p	0.94±0.030	0.274±0.015	0.0006	2	0.99
Model 2	0.98±0.044	0.065±0.026	0.0032	4	0.61
Model 2 (constrained)	1	0.076±0.013	0.0028	5	
Model 3	0.91±0.026	0.042±0.021	0.0044	13	0.24
Logarithmic (height ratio)					
Model 4	0.95±0.045	0.037±0.022	0.0047	4	0.43
Linear (age difference)					
Model 5	1.03±0.031	-0.013±0.003	0.0011	4	0.87
Model 5 (constrained)	1	-0.011±0.001	0.0011	5	

Table 4.10. Predicted age-age genetic correlations for height at four ages by the Lambeth's model (Model 1) and age-age genetic correlation model (Model 2), and differences with observed estimates (%).

Traits	Observed	Model 1		Model 2	
	r_A	Predicted	Diff.	Predicted	Diff.
HT1.5,HT9.5	0.93	0.45	52	0.86	7
HT1.5,HT13.5	0.85	0.34	60	0.83	2
HT1.5,HT22.5	0.76	0.19	75	0.79	-4
HT9.5,HT13.5	0.96	0.91	5	0.97	-1
HT9.5,HT22.5	0.85	0.75	12	0.93	-9
HT13.5,HT22.5	0.97	0.86	11	0.96	1



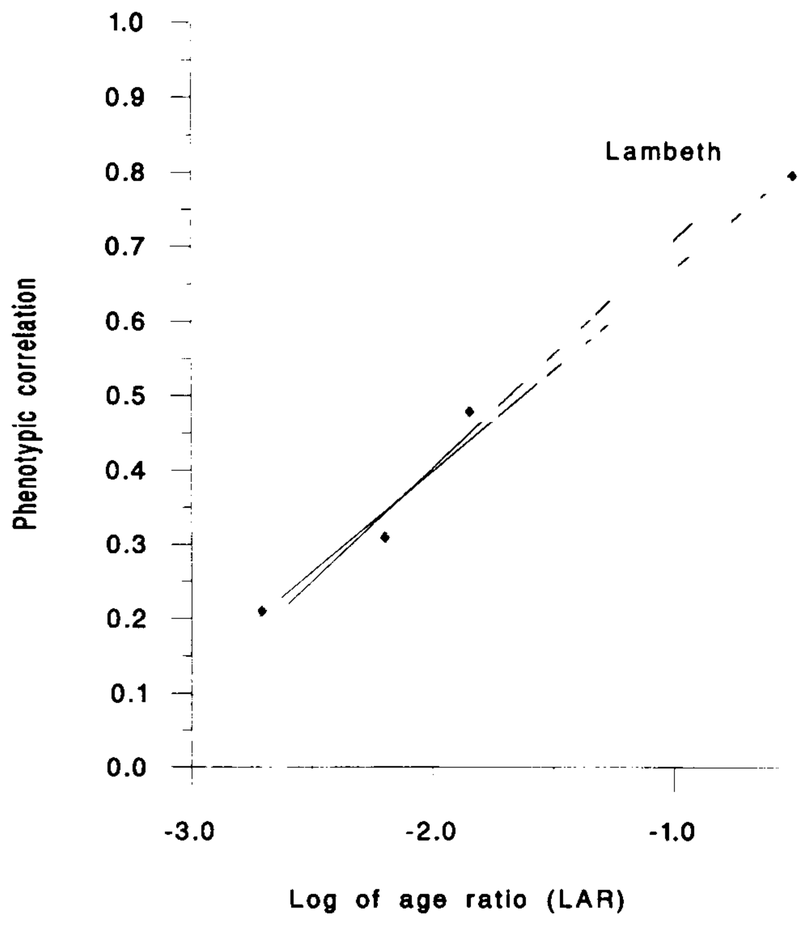


Figure 4.1. Relationship between age-age correlations for height and natural logarithm of the ratio of the younger to the older age (LAR). The Lambeth model is shown.

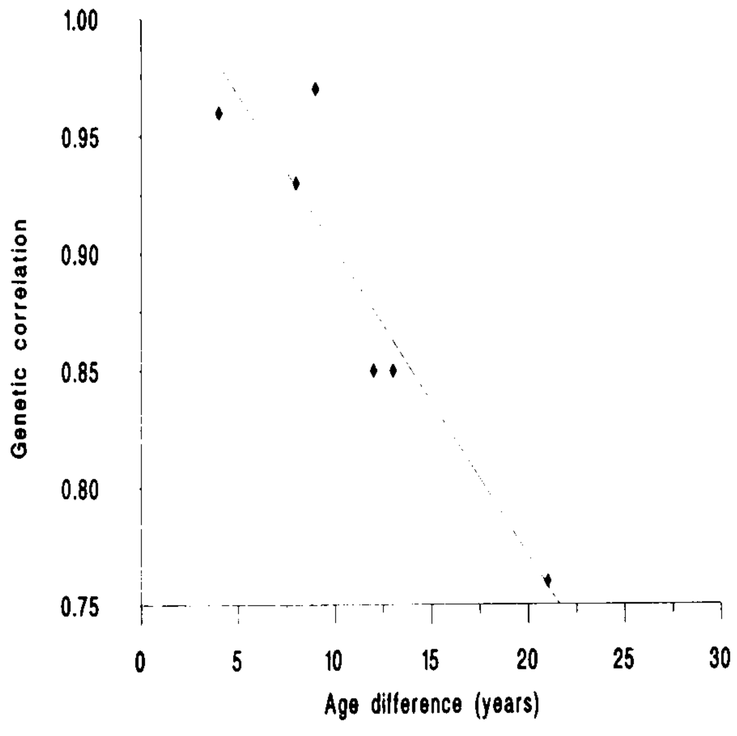


Figure 4.2. Relationship between age-age genetic correlations for height and the age differences.

Effect of fitting Model 3

The 1976 genetic tests were assessed for height at 7.5 and 12.5 years, and the genetic correlations involving these assessments are presented in Table 4.11. When data involving the 1976 assessments were included in fitting a Lambeth type model, the following model was obtained (Model 3):

$$r_A = 0.91 + 0.042\ (LAR) \qquad (R^2 = 0.24) \qquad (4.15)$$

The slope of this model was not significantly different from zero and this was attributed, partly, to a lower than expected genetic correlation (0.72) between 12.5 and 22.5 years, which resulted in this point having a large standardized residual. Although removing this point caused the regression to be significant, there was no biological rationale for this point to be excluded from the analysis. Since the variance increased with decreasing genetic correlation, the genetic correlations were weighted by the inverse of their corresponding variances. Although the weighted regression was significant, the plot of the residuals revealed that our model had changed. Therefore, weighting and removing the outlier were not viable options. The model resulting from forcing the regression to have a unity intercept differed significantly from the unconstrained model. The fact that the fit of the age-age genetic correlation model for height was poor when the genetic correlations involving the 1976 assessments were included indicates possible genotype by environment interaction between the 1972 and 1976 data. The presence of genotype by environment interaction is also indicated by lower genetic correlations between sites than within sites (Table 4.11).

Modelling age-age genetic correlation using Models 4-6

Since the genetic correlation prediction model using the Lambeth type model was not a good fit, other types of models were fitted to the data. Firstly, a model using age differences as the predictor of genetic correlation was fitted. This model was fitted because the genetic correlations tend to decrease as age difference increases. Secondly, a prediction model was fitted using the natural logarithm of the ratio of height at the younger age to height at the older age. This prediction model is believed to fit better than

the ratio of the ages (Williams[1] pers. communication). Thirdly, genetic correlations were predicted from (co)variance functions estimated in Chapter 5.

Table 4.11. Predicted age-age correlations for height at six ages by Lambeth's model (Model 1) and the covariance functions (Model 6), and the differences with observed estimates (%).

Ages	Observed	Model 1		Model 6	
	r_A	Predicted	Diff.	Predicted	Diff.
1.5, 7.5	0.91	0.52	43	1.02	-12
1.5, 9.5	0.93	0.45	52	0.95	-2
1.5, 12.5	0.79	0.37	53	0.84	-6
1.5, 13.5	0.85	0.34	60	0.80	6
1.5, 22.5	0.76	0.19	75	0.78	-3
7.5, 9.5	0.96	0.95	1	1.00	-4
7.5, 12.5	0.90	0.86	4	1.00	-11
7.5, 13.5	0.90	0.84	7	0.94	-4
7.5, 22.5	0.85	0.68	20	0.90	-6
9.5, 12.5	0.90	0.94	-4	1.00	-11
9.5, 13.5	0.96	0.91	5	1.00	-4
9.5, 22.5	0.85	0.75	12	0.97	-14
12.5, 13.5	0.86	1.00	-16	1.00	-16
12.5, 22.5	0.72	0.84	-17	0.98	-36
13.5, 22.5	0.97	0.86	11	0.99	-2

The model with log of height ratio as the predictor of genetic correlation (Model 4) was a poorer fit than that based on the log of age ratio predictor, as shown by the higher residual mean square and lower R^2 (Table 4.9). The best predictor of genetic correlation was the age difference with the lowest residual mean square, and highest R^2 (Table 4.9

[1]C. G. Williams, Texas A&M University, Texas, USA.

and Figure 4.2). For Model 5, constraining the genetic correlations between height assessed at the same age to be 1 resulted in a fit which was not significantly different from the unconstrained model (i.e.$\beta_0 = 1$, $P > 0.05$). The genetic correlations calculated using the covariance function were more consistent with the observed than those estimated by the Lambeth model (Table 4.11).

4.4 Discussion

All age-age genetic correlations estimated for height using pooled data were high indicating that early selection in *P. taeda* in Zimbabwe will be effective. Age-age genetic correlations estimated for height are in close agreement with estimates reported by Lambeth *et al.* (1983) and McKeand (1988) for *P. taeda,* and Barnes (1992a and b) and Pswarayi *et al.* (1996) for other pine species in Zimbabwe. Foster (1986), Franklin (1979) and Williams and Megraw (1994) reported weak genetic correlations (less than 0.4) between height at ages younger than three years and ages older than 12 years from *P. taeda* genetic tests in the USA. The results of the present study show an opportunity for selecting at a very young age (1.5 years). However, it may not be possible to take advantage of early selection because the species only starts flowering at age 10 years in Zimbabwe (Barnes[2], personal communication). The difference between age-age genetic correlations from this study and those from USA genetic tests may be a consequence of management and methodological differences.

In some reports, family mean correlations were used as approximations of genetic correlations (e.g. McKeand 1988). Family mean correlations are likely to underestimate genetic correlations because the components of family mean correlation are (Namkoong *et al.* 1979):

$$r_{F1,F2} = \frac{cov_{F1,F2} + cov_{e1,e2}}{\sqrt{\sigma^2_{F1} + \sigma^2_{e1}}\sqrt{\sigma^2_{F2} + \sigma^2_{e2}}} \qquad (4.16)$$

[2]R. D. Barnes, Oxford Forestry Institute, Oxford, UK.

where σ^2_F = family mean variance and σ^2_e = error variance of family means, $cov_{F1,F2}$ = covariance of family means and $cov_{e1,e2}$ = error covariance. While the error covariance can be zero, the error variance is rarely zero. Also, the error variance may contain dominance effects in the case of full sib-families, thereby increasing the bias in the genetic correlation estimate. For example, Lambeth *et al.* (1983) found that the family mean correlations were less than the genetic correlations by as much as 35%. In other studies, genetic tests were not thinned at random (i.e the best trees were retained; e.g. Lambeth *et al.* 1983), while the Zimbabwe tests received a systematic thinning. This non random type of thinning could also lead to biased estimates of genetic parameters as discussed in Chapter 3, Section 4. It has also been suggested that growth rates might affect the genetic correlations, with lower genetic correlations between any two ages in fast growing material (Magnussen 1988), perhaps as a consequence of slower physiological changes in slower growing trees leading to the expression of more similar genes at any two ages than for faster growing trees. The growth rates of trees in the Zimbabwe tests were higher than those in the USA tests. Under this hypothesis, genetic correlations from the Zimbabwe tests may be expected to be lower, but the opposite was true. Hence, differences in growth rates are unlikely to explain the differences in the genetic correlations reported here and those reported for *P. taeda* grown in USA.

Age-age genetic correlations for height were higher than corresponding phenotypic correlations, which is consistent with the literature survey of genetic and phenotypic correlations in *P. taeda* by Newman and Williams (1991) and in other pine species (Barnes 1992a and b; Pswarayi *et al.* 1996; Riemenschneider 1988). The observation that age-age genetic correlations are generally higher than corresponding phenotypic correlations has also been reported in growth traits in animals (Bishop 1992, Fimland 1973, Koch *et al.* 1982). Since similar genes are likely to influence a trait at different ages, age-age genetic correlations can be expected to be high. However, age-age phenotypic correlations can be expected to be low due to measurement errors and other environmental factors. The Lambeth model underpredicted genetic correlations. This was attributed to the fact that the Lambeth model used phenotypic correlations which tend to be lower than the genetic correlations. Results of this study support the assertion that the Lambeth model underestimates genetic correlations, and it is a good

predictive model for the phenotypic correlations. Although use of phenotypic correlations might be more desirable than genetic correlation in tree breeding as the former are estimated with much higher precision, the phenotypic correlations in this study were significantly lower than their genetic correlation counterparts. Despite having a limited number of parents here, the results show that the precision of age-age genetic correlations for height was high because the genetic correlations themselves were high. The Lambeth model was based on data from Douglas fir, Ponderosa pine and Loblolloy pine genetic tests which were not thinned and therefore suffered mortality from competition. The impact of better silviculture on genetic or phenotypic correlations is unknown. Franklin (1979) found that genetic correlations for height measured at two different ages were generally high if both measurements occurred either before or after the onset of competition, whereas the genetic correlations were low or even negative if the onset of the competition occurred between the two measurements. On the other hand, Lambeth (1983) refutes this hypothesis. The results from this study indicate that competition may not have an effect on phenotypic correlations since our model was not significantly different from that of Lambeth. The study reported by Riemenschneider (1988) for a genetic test of *Pinus banksiana* measured up to 7 years, and therefore unaffected by competition effects, supports the hypothesis that phenotypic correlations are not affected by competition. He found that the phenotypic correlation model did not differ significantly from the Lambeth model, although the genetic correlation model did. The economic consequences of using the Lambeth model based on phenotypic correlations may be very high in forest trees which have long generation intervals, and so more efficient methods should be used for predicting genetic correlations.

Our findings disagree with those of Burdon *et al.* (1992b) who found that the logarithmic relation of Lambeth's model represents a valid framework for describing age-age genetic correlations for height. This study indicates that the logarithmic relationship is not necessarily the best model. The best fit for the genetic correlations in this study was provided by the model involving age difference. The regression models assumes consistent and predictable changes of age-age correlations with time, an assumption that has not been verified, and may not be valid in some cases. Therefore, covariance functions (Kirkpatrick *et al.* 1990) might be more efficient, as they do not

make any prior assumptions about the model, and hence are more flexible. We found the genetic correlations predicted by the covariance functions to be more consistent with our data. Fitting covariance functions requires good data sets, which may not be available; computational procedures are more involved than those of simple linear regression, making simple linear regressions more attractive for many situations in forestry. Genetic correlations predicted by the different methods presented here may have wide implications for the expected gain and optimum age for selection, and this will be examined in a subsequent study in this thesis.

Age-age genetic correlations in straightness between 1.5 years and older ages were very low and this could partly be explained by the fact that this trait is difficult to measure at a young age. The low correlation indicates that there were large rank changes in parents between 1.5 years and subsequent ages. Genetic correlations between straightness at older ages were moderate to high, indicating that early selection at ages 9.5 years or older in this trait will result in improvement in straightness at rotation age. Results of this study are consistent with those reported by Pswarayi *et al.* (1996). The results of this study show that it might not be possible to consider straightness scores at different ages as repeated records, rather should be considered as different traits. The repeatability model assumes that the traits are genetically similar, and that variances across ages are equal and heritabilities are equal, which does not appear to be the case here.

Genetic correlations between height and straightness were generally low and positive, except those involving straightness at 1.5 years which were generally negative, indicating that there were few similar genes influencing the two traits at the older ages. Selection on height alone at any age should result in improvement in straightness at harvest age, but selection on straightness alone at 1.5 years would adversely alter straightness at harvest age. There appear to be no genetic correlations between height and straightness reported in *P. taeda* in the literature. Those reported here were lower than those reported by Woolaston *et al.* (1990) in *P. caribaea* in Australia, and by Cotterrill *et al.* (1987) and Pswarayi *et al.* (1996), and higher than those reported by Barnes *et al.* (1992a and b), for other pine species in Southern Africa. In this study genetic and phenotypic correlations between height and stem straightness were not

significantly different. The result is consistent with studies in animals. For example, Cheverud (1988) compared 41 published estimates of pairs of trait-trait genetic and phenotypic correlations from mouse populations, and Koots and Gibson (1996) compared 1015 pairs from beef cattle populations, and both found that the differences between trait-trait genetic and phenotypic correlations were mostly due to sampling errors.

Heritability estimates for height and straightness, with the assessments at the four different ages treated as different traits as in the univariate analysis, may be biased due to selection. Although the bivariate analysis removes some of the bias due to selection, it is not as efficient as multivariate analysis which includes all the ages. Estimates of heritability of bivariate analysis that include the 1.5-year assessments most likely have the smallest selection bias because thinning was carried out after this age. The heritability estimates at ages older than 1.5 years based on bivariate analyses showed a small absolute difference (<0.02) from those based on the univariate analyses, which may indicate that bias due to selection might be small.

Chapter 5

PREDICTION OF COVARIANCE FUNCTIONS FOR HEIGHT

5.1 Introduction

Predictions of time trends in genetic parameters are valuable for estimation of annual genetic gain and for choosing the optimum age for selection. Many forestry genetic tests are characterised by few, and irregular assessment intervals, which make such predictions difficult. The most common method for predicting time trends of genetic parameters is fitting of a growth curve to each tree. For example, Buford and Burkhart (1987) used a linear regression method to describe the *Pinus taeda* height growth curve, but the most commonly used function is the exponential growth curve developed by Richards (1959) (Richards function, e.g. Ballochi *et al.* 1993, Knowe and Foster 1989, Namkoong *et al.* 1972). The Richards function permits each tree to have its own unique growth function of the form:

$$HT = A(1 - e^{-bx})^c \qquad (5.1)$$

where:

HT = height of the tree,

A = the ultimate limiting value for the tree height (asymptote),

e = the base of natural logarithms,

b & c = shape parameters that determine the shape of the curve along the time axis,

x = age in years.

The function requires that measurement data for individual trees is available at all the assessment ages. Therefore, only data for trees available at the final assessment age can

be used. Given that thinning is carried out at various stages of growth, the number of trees at final assessment age can be substantially less than that originally planted, resulting in a large part of the data not being used. Heights estimated at each age using the growth function are then used to estimate the variance components at those ages and covariance components across ages. The advantages of such methods are that they correct for errors, estimate missing data, and provide smooth growth curves (eg. Balocchi *et al.* 1994, Knowe and Foster 1989). One of the major disadvantage of the methodology is that it makes prior assumptions about the form of the curves, and the data may poorly fit the growth curve. Moreover, it does not consider the underlying continuous covariance structure of the data.

The commonly used models by tree breeders for predicting the time trends of genetic correlations for growth traits are linear models with the natural logarithm of the ratio of the younger age and mature age (LAR) as the predictor (e.g. King and Burdon 1991, Lambeth 1980, McKeand 1988); heritability is assumed constant when predicting genetic gain (e.g. Lambeth 1980). These models may not be the best; hence, alternative models such as non-linear models may fit the data better. Variation in heritability which has been reported in many studies (Ballochi *et al.* 1993, Pswarayi *et al.* 1996, Franklin 1979) suggests that models which make it possible to predict genetic correlations as well as age-related trends of heritability will ensure more accurate predictions of gain and optimum selection age. Therefore, modelling of the 'coefficient of genetic prediction' ($h_x r_A h_y$, standardised genetic covariance or coheritability) as proposed by Baradat (1976) or modelling of the heritability itself as suggested by Wei and Borralho (1996) may make possible more accurate predictions of gain and optimum selection age.

An alternative analytical approach is to fit non-linear models rather than linear ones. New methods for predicting time trends in (co)variances, developed by Kirkpatrick *et al.* (1990), fit continuous functions of time to covariance matrices using weighted least squares. The derived covariance function gives the covariance between assessments at any two ages as a higher order polynomial of the ages. The covariance function can be regarded as an infinite dimensional equivalent of a covariance matrix. The covariance functions can provide more accurate estimates of covariances than possible from a few assessment ages because covariances at any two ages can be

improved by using information about the covariances at other ages. The other attractions of these methods are that they are a direct approach to estimating covariances, are easy to use, and do not make any prior assumptions about the form of the curves. The covariance functions account for spacing of ages. Furthermore, by using orthogonal polynomials the additive covariance function can be decomposed into its eigenfunctions and eigenvalues which allow analysis of additive genetic variation. For a given matrix, A, a vector, U (eigenvector), and scalar, λ (eigenvalue), $\boldsymbol{AU} = \boldsymbol{\lambda U}$; similarly, a covariance function can be written in terms of its eigenfunctions and eigenvalues as: $T\psi_i = \lambda_i\psi_i$, where T is the covariance function, λ_i is the eigenvalue associated with eigenfunction ψ_i. Therefore, the eigenfunctions are analogous to eigenvectors associated with matrices. The eigenvalues and eigenfunctions are useful for analysing directions in which mean growth curves are likely to change under selection. The eigenvalue is proportional to the amount of genetic variation in the population corresponding to the particular eigenfunction, with large eigenvalues indicating changes for which there is large genetic variation.

The objective of this Chapter is to predict time trends in (co)variances for height growth using covariance functions. This approach has recently been used in animal breeding (Kirkpatrick *et al.* 1994, Meyer and Hill 1997), and this study demonstrates this methodology using forestry data.

5.2 Materials and Methods

An estimate of the additive genetic covariance matrix of the tree height at ages 1.5, 7.5, 9.5, 12.5, 13.5 and 22.5 years was derived using individual tree model DFREML.

Estimation of covariance functions was made using orthogonal polynomials with symmetric and asymmetric coefficients by fitting continuous functions of time to covariance matrices using least squares. A covariance function is the infinite-dimensional equivalent to covariance matrix for a given number of traits. It gives the covariance between any two traits assessed at given ages as a function of the ages and some coefficients. Using symmetric coefficients approach, the covariance between

heights assessed at ages t_1 and t_2 is:

$$T(t_1,t_2)=\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\hat{C}_{ij}\phi_i(t_1)\phi_j(t_2), \qquad (5.2)$$

where T is the covariance function (phenotypic, additive or coheritability), *k* is the order of fit (k ≤ number of ages), $\hat{C}$ the coefficient matrix associated with the covariance function, and ϕ the orthogonal polynomials. The genetic correlations between the height assessed at ages t_1 and t_2 were estimated as:

$$r_g=\frac{cov_A(t_1,t_2)}{\sqrt{var_A(t_1)}\sqrt{var_A(t_2)}} \qquad (5.3)$$

and heritability at any age t_1 as:

$$h^2=\frac{var_A(t_1)}{var_P(t_1)} \qquad (5.4)$$

The difference between the asymmetric coefficients approach described by Kirkpatrick *et al.* (1994) and that based on symmetric coefficients lie in the coefficient matrix. With the symmetric coefficients approach, the matrix of the coefficients used for estimating the covariance function is symmetric (i.e. $\hat{C}_{ij} = \hat{C}_{ji}$); with the asymmetric coefficients approach, it is not (i.e. $\hat{C}_{ij} \neq \hat{C}_{ji}$). The latter method may lead to smoother and better-behaved estimates because it eliminates the product of two $(k-1)^{th}$ order polynomials (Kirkpatrick *et al.* 1994).

Full and reduced order models were used to estimate the appropriate covariance functions. A full order model, in which the order of orthogonal polynomial models equals the number of ages, estimates the coefficient matrix in such a way that the corresponding covariance function exactly reproduces the estimated (co)variances at the ages that were assessed. These estimates, however, include the sampling errors and therefore are not smooth. Smoothing, using information from adjacent points to average

out the errors, was achieved by fitting reduced order models, in which the order of polynomial models is less than the number of ages. When a reduced model is used, the program requires the error covariance matrix, and this was estimated after providing the program with the phenotypic covariance matrix and assuming that a standard balanced half-sib breeding design with 7 male and 21 female parents, and 20053 residual degrees of freedom was used. The Lambeth model ($r_p = 1.02 + 0.308 \log_e$ (younger age/older age)), together with the observed phenotypic variances, were used to estimate the phenotypic covariance matrix. The Lambeth model was used since not all age-age phenotypic correlations were available and those that were available, showed excellent agreement with it (Chapter 4, Figure 4.2).

The phenotypic covariance function was estimated using a phenotypic matrix estimated as above. For a reduced fit, the additive covariance matrix was supplied in order to estimate the error covariance matrix.

In addition, a continuous function was fitted to the so called 'coefficient of genetic prediction' (i.e. coheritability or standardised additive covariance) matrix. The 'coefficient of genetic prediction' (CGP) is estimated as follows (Baradat 1976):

$$CGP = \frac{Cov_A(x,y)}{\sigma_x \sigma_y} \qquad (5.6)$$

or alternatively as, coheritability:

$$CGP = r_A h_x h_y \qquad (5.7)$$

where:

$Cov_A(x,y)$ = additive genetic covariance between traits x and y,

σ_x, σ_y = phenotypic standard deviation for traits x and y, respectively,

r_A = genetic correlation between the traits,

h_x, h_y = square root of the heritability for traits x and y, respectively.

The coheritability (coefficient of genetic prediction) function allowed the modelling of heritability with time, since the coheritability of assessments observed at the same age

is the heritability of the trait. When a reduced model was used, the error covariance matrix was estimated by providing the program with the appropriate phenotypic covariance matrix and assuming that a standard balanced half-sib breeding design with 7 male and 21 female parents, and 20053 residual degrees of freedom. The appropriate phenotypic covariance matrix was a standardised phenotypic matrix comprising 1 on the diagonal and the phenotypic correlations off diagonal. The phenotypic correlations were estimated using the Lambeth model.

The goodness-of-fit of reduced models was tested using the χ^2 test procedure described by Kirkpatrick *et al.* (1990).

5.3 Results

Additive covariance function

Estimates of additive variances and covariances for height from DFREML and additive covariance function with symmetric coefficients (k = 4) are summarized in Table 5.1. The variances and covariances estimated with the covariance function were consistent with the DFREML ones, except those involving 12.5 years.

A plot of the original additive covariance matrix showed fluctuations which may be attributed to sampling errors. Full estimates (k=6 order polynomial) of the additive covariance function, using symmetric and asymmetric coefficients also showed fluctuations, and data interpolated between the points of the matrix were over-inflated. The symmetric estimates ranged from 0 to 60 m^2, and those using asymmetric coefficients 0 to 10 m^2, while the original data ranged from 0 to 1.4 m^2. Estimates using reduced order models (k=4 order polynomial), using both symmetric and asymmetric coefficients, were much smoother. Figure 5.1 shows a 3-dimensional plot of the reduced model to the order of 4 with symmetric coefficients. The goodness-of-fit tests gave χ^2 (11 df) = 109.9 for the symmetric fit, and χ^2 (11 df) = 121.7 for the asymmetric fit. This indicated that the discrepancies between the smoothed covariance functions using both approaches and the original data were significant, and that use of the symmetric coefficients provided a better model. Therefore, the reduced order model (k = 4) with

the symmetric coefficients (Figure 5.1) was selected. This model gave discrepancies of additive variances of less than 12%, except for 12.5 year variance which it was 50% (Table 5.1). Covariance discrepancies with this model were largest for those involving 12.5-year assessments (11-63%), with the rest being less than 10%.

Table 5.1. Estimates of additive variances and covariances for height (m^2) for ages 1.5. 7.5, 9.5, 12.5, 13.5 and 22.5 years, based on DFREML (DF) and additive covariance function (ACF) with symmetric coefficients (k = 4).

σ^2_A		Cov_A					
DF	ACF	DF (below the diagonal), ACF (above the diagonal)					
0.017	0.019		0.150	0.154	0.138	0.130	0.091
1.002	1.127	0.127		1.238	1.264	1.244	0.854
1.404	1.362	0.147	1.257		1.392	1.371	0.95
0.930	1.423	0.108	0.846	1.253		1.402	0.989
1.285	1.381	0.106	1.132	1.401	0.969		0.980
0.715	0.708	0.071	0.793	1.000	0.607	0.938	

The first eigenvalue explained 99% of the total genetic variation ($\lambda_1 = 20.52$; sum of all eigenvalues = 20.72). Its eigenfunction was positive at all ages and flat (Figure 5.2a), showing that there was large genetic variation for increased mean height growth at all ages. The results indicate that improvement of height at one age will tend to improve height at all ages: genetic correlations for height between different ages are high. The other two eigenfunctions are shown in Figure 5.2b and 5.2c, and the eigenvalues associated with them were all positive and small ($\lambda_2 = 0.25$ and $\lambda_3 = 0.01$).

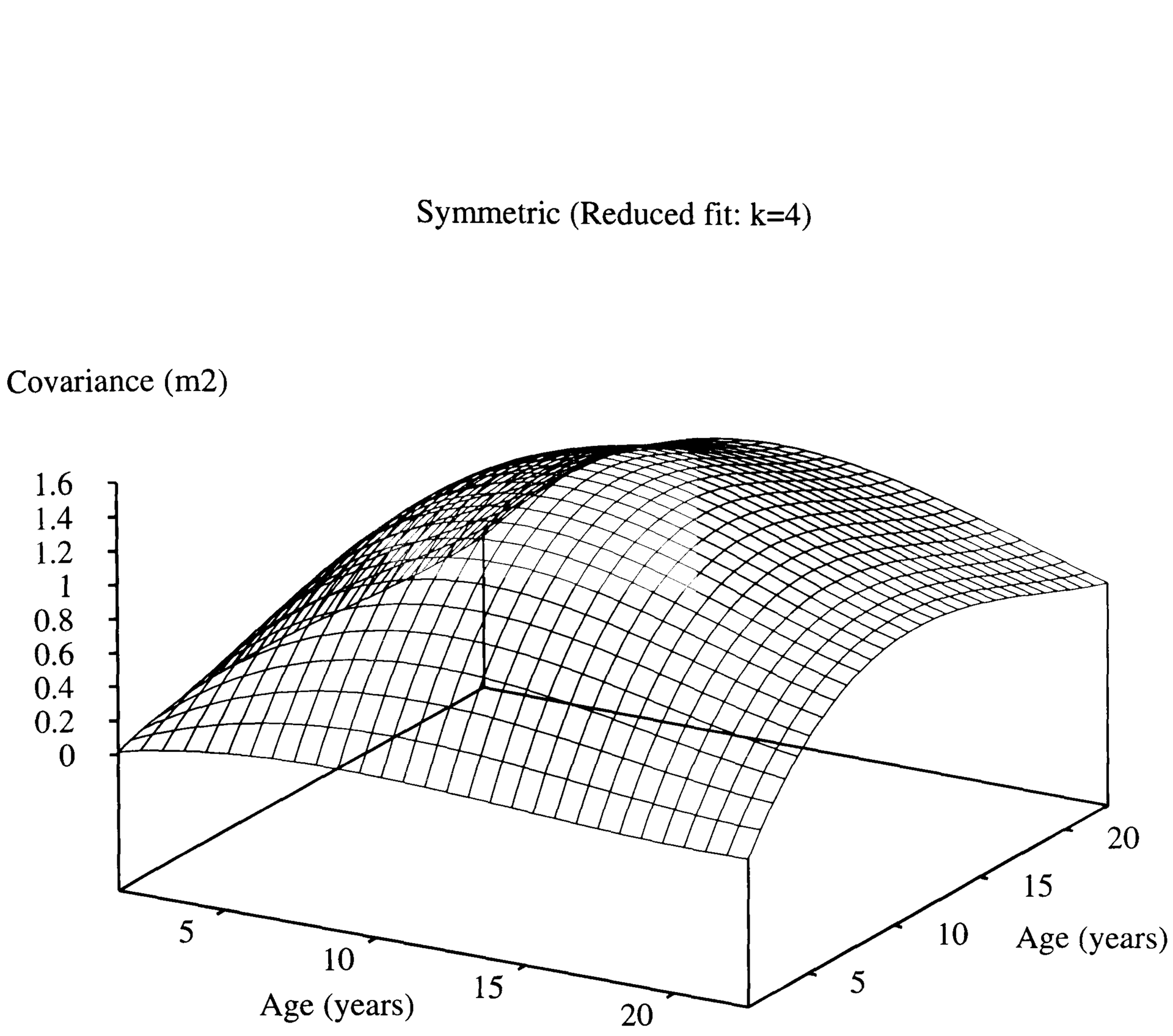


Figure 5.1. A 3-dimensional plot of the additive covariance function.

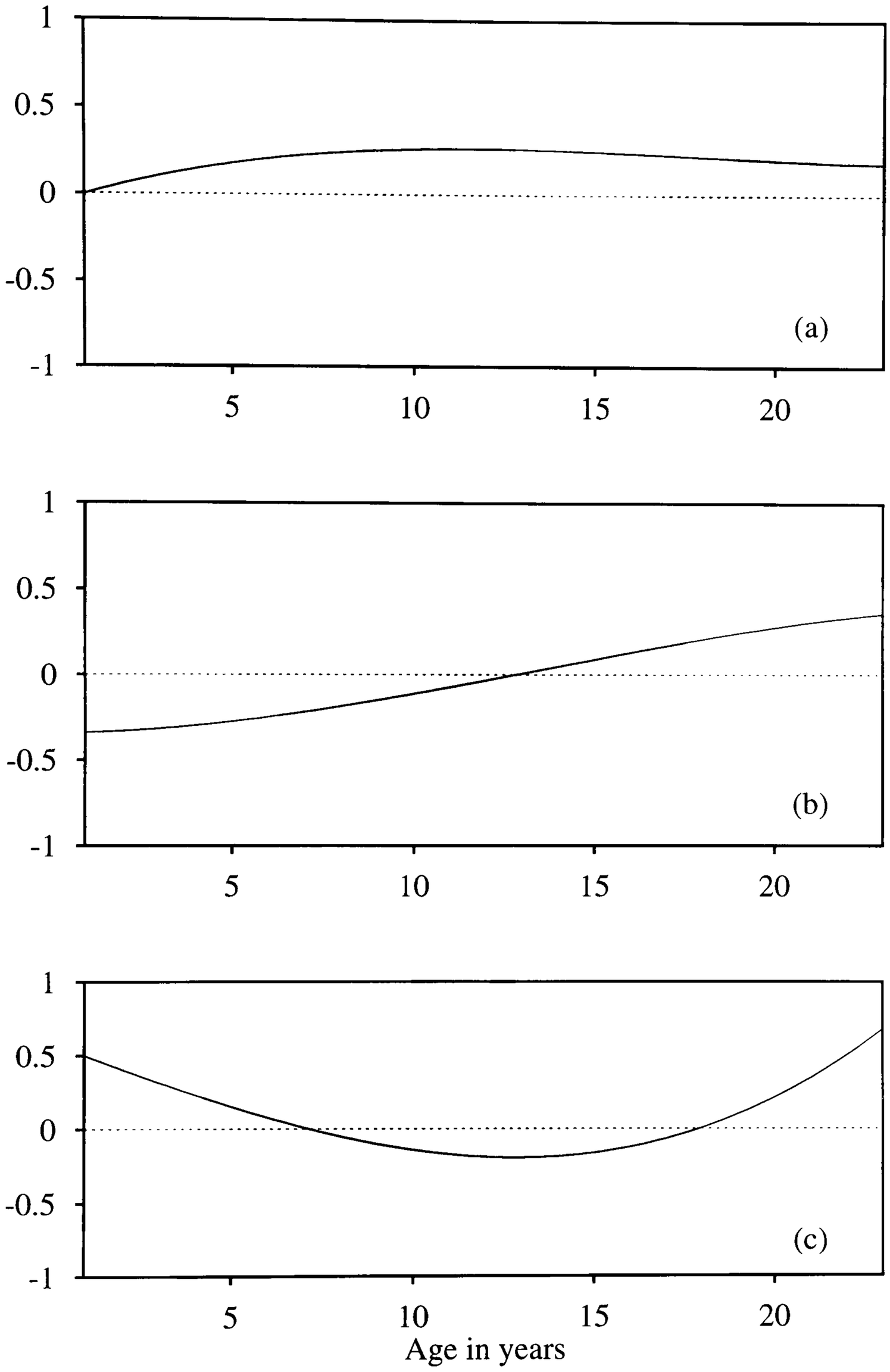


Figure 5.2. Estimates of the 1st eigenfunction (a), 2nd eigenfunction (b) and 3rd eigenfunction (c). Their eigenvalues (m^2) were 20.52, 0.25 and 0.01, respectively.

The eigenvalue associated with the fourth eigenfunction was negative (λ_4 = -0.06), indicating that the additive genetic covariance matrix, hence the additive covariance function, was not positive semi-definite. Although Hayes and Hill (1981) suggested 'bending' the data matrix when the matrix is not positive semi-definite, our negative eigenvalue was very small and unlikely to differ significantly from zero, so bending of the data matrix was not undertaken. The small amount of genetic variation associated with all other eigenvalues, apart from the first, implies that selection to alter the shape of the growth curve (i.e. tradeoffs between early and late height growth) will make very little or no progress.

Table 4.11 in Chapter 4 shows that the genetic correlations estimated using the additive covariance function with symmetric coefficients (k = 4) were consistent with the original estimates, except for heights assessed at 12.5 and 22.5 years.

Phenotypic covariance function

Full estimates using both symmetric and asymmetric coefficients over-inflated values between the assessment ages. For example, the range of the data was 0.02 to 5.46 m^2, while that predicted using the symmetric coefficients was 0 to 400 m^2, and 0 to 60 m^2 using the asymmetric coefficients. All possible reduced models gave negative variances and covariances which were highly inconsistent with the data. Therefore, it was not possible to estimate a phenotypic covariance function.

Coheritability function

Full and all possible reduced models over-inflated values between assessment ages and produced very large discrepancies between 13.5 and 22.5 years. Because the method failed to fit a smooth function consistent with the data, functions were estimated using a reduced data set: data assessed at 1.5, 7.5, 9.5, 12.5 and 13.5 years. Omitting data involving 22.5 years was justified on the grounds that we concentrated on a period when we had more information.

Reducing the data set made it possible to fit a model, and the most appropriate

one was derived using symmetric coefficients (k=4; Figure 5.3). Estimates of heritability and coheritability for height from DFREML and the coheritability function are summarized in Table 5.2. The heritabilities and coheritabilities estimated with the coheritability function were consistent with the DFREML ones. The coheritability model gave discrepancies of heritability of less than 10%, except for 12.5-year heritability which it was 20% (Table 5.2). Coheritability discrepancies with this model were all less than 10%. However, the goodness-of-fit test gave χ^2 (5 df) = 55.0 for the fit, indicating that the discrepancies between the smoothed coheritability function and the original data were significant.

Table 5.2. Estimates of heritability and coheritabilities for height for ages 1.5, 7.5, 9.5, 12.5 and 13.5 years, based on DFREML (DF) and coheritability function (CHF) with symmetric coefficients (k = 4).

h^2		Coheritabilities				
DF	CHF	DF (below the diagonal), CHF (above the diagonal)				
0.22	0.22		0.280	0.309	0.250	0.187
0.43	0.41	0.280		0.441	0.357	0.280
0.50	0.50	0.308	0.445		0.418	0.319
0.45	0.36	0.249	0.396	0.427		0.277
0.22	0.22	0.187	0.277	0.318	0.271	

Symmetric (Reduced fit: k=4)

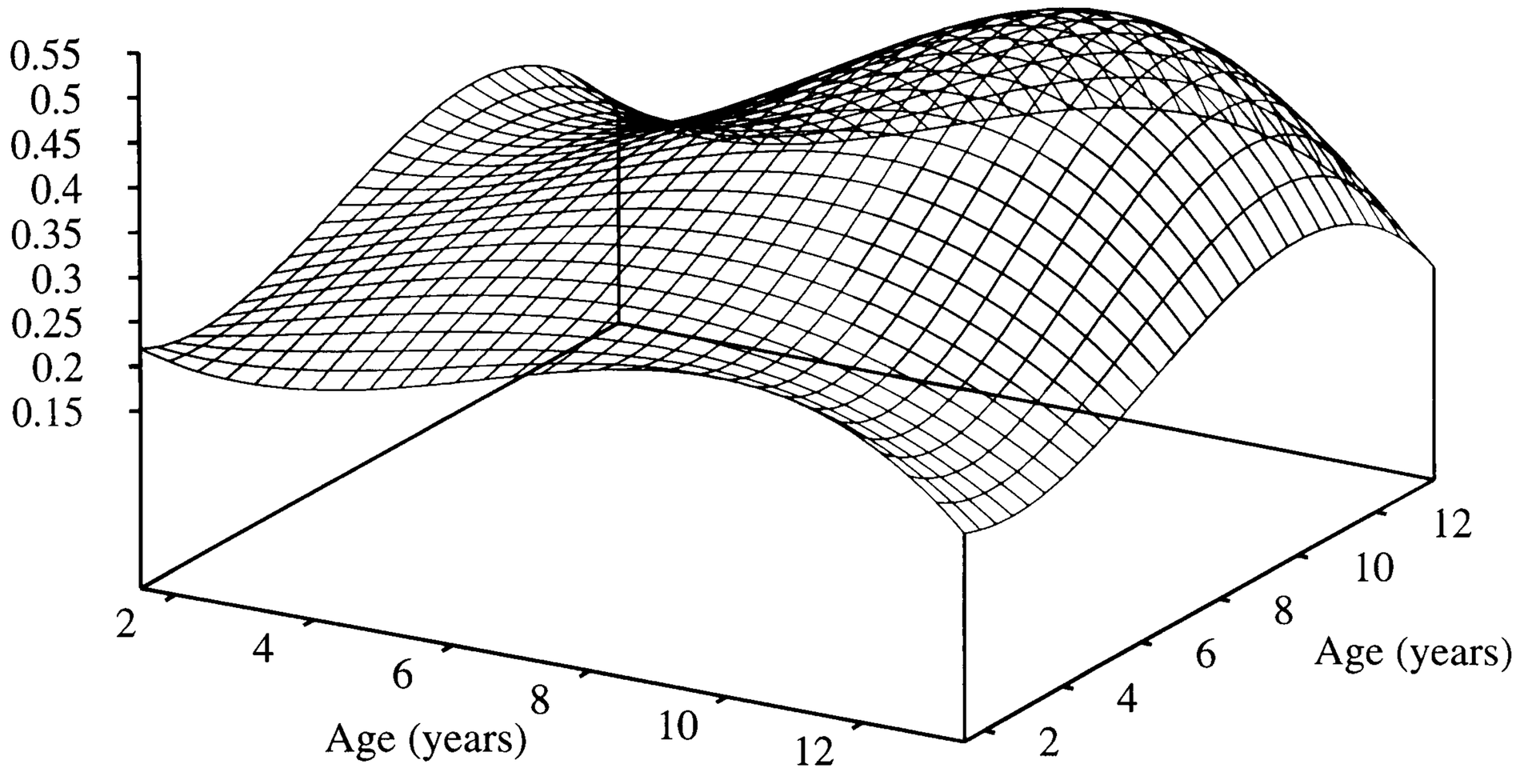


Figure 5.3. A 3-dimensional plot of the coheritability function.

5.4 Discussion

The symmetric coefficients method provided the most appropriate model to the additive covariance matrix using 4 polynomials (i.e. ages to the power 0 to 3). Additive variance and covariance estimates from this model were consistent with the original estimates, except those involving 12.5 years. Although the fourth eigenvalue was negative, indicating that the function was not positive semi-definite, the eigenvalue was very small. The function differed significantly from the original data, but the discrepancies were small. The first eigenvalue explained 99% of the total genetic variation and eigenfunction was positive at all ages and flat, implying that selection for total height growth will be rapid. In contrast, the small amount of genetic variation associated with all the other eigenvalues implies that selection to alter the shape of the growth curve will make very slow progress. These results support the findings in Chapter 4 which showed high positive genetic correlations for height among all ages.

Although the methods were able to fit a function to the additive matrix, they failed to fit an appropriate function based on the phenotypic matrix. This may be attributed to large fluctuations in the phenotypic covariance matrix, requiring more complex functions to be fitted to the data. The phenotypic covariance matrix show that variance increased with age up to age of thinning, decreased thereafter and then increased again until age of thinning, and as pointed out in Chapter 3 this may be due to compensatory growth. Our failure to get a good fit to our data might also highlight the limitations of a small data set. With more points (ages), one increases the choice of the polynomials that can be fitted and increases the chance of getting a good fit to the data.

Since the response to selection of a correlated trait using the coefficient of genetic prediction (CGP) is given by $Gain_y = i_x \bullet CGP_{xy} \bullet \sigma_y$, a CGP (coheritability) function eliminates the need to estimate both the additive and phenotypic covariance functions, if the objective is to estimate gain and optimum selection age. In our study, it was possible to estimate a coheritability function after omitting the data for 22.5 years. This limited use of the function for estimating gain at 22.5 years from early selection, but was useful for predicting trends of heritability up to 13.5 years. Although the annual

response to indirect selection is given by:

$$r_{xy} \bullet h_x \bullet h_y \bullet \sigma_y \bullet L_x^{-1}, \qquad (5.8)$$

h_y and σ_y are constant across ages. Therefore, optimum selection age could be predicted using $r_{xy} \bullet h_x \bullet L_x^{-1}$, assuming the selection intensity is constant across ages.

The precision of the covariance function estimates could be improved by estimating the covariance function directly from the original observations rather than from the matrix (Kirkpatrick *et al.* 1994). Such an algorithm has being developed so that the covariance functions can be estimated using REML (e.g. Meyer and Hill 1997). The advantages of using REML for estimating the covariance functions are that the error covariance will be estimated at the same time as the other covariances, estimates of the covariance functions will be guaranteed to be positive semi-definite, and the likelihood ratio test can be used to determine the minimum order of the model (Meyer and Hill 1997).

Short series of measurements, which are common in forestry, tend to favour linear models or simple covariance functions for predicting trends of age-age genetic correlations for growth traits. These short series may actually fail to reflect actual trends. As tree breeding programmes progress to advanced generations more information will be collected, and the simple linear models may not be appropriate. The results reported here demonstrate that alternative models, using covariance functions, can be used for estimating time trends in (co)variances by directly fitting continuous covariance functions to the matrices. The particular distribution over time of data that are available in forestry genetic tests may not always allow the fitting of growth curves; where they do, the data may poorly fit the growth curves, requiring that alternative methods be identified. The methods demonstrated in this chapter might be more efficient at predicting trends in genetic parameters, and at predicting genetic gain and optimum selection age. The strength of the symmetric and asymmetric coefficient methods compared to the more conventional methods, such as the Richards curve, is that they do not make any assumptions about the form of curves. Furthermore, the eigenvalues and eigenfunction associated with the additive covariance function can be used to make inferences about patterns of genetic variation, and provide information on the directions in which mean growth curves are likely to change under selection. Therefore, the

eigenvalues and eigenfunctions provide a convenient and succinct summary of a large body of data which otherwise is not possible with other non-linear models, such as splines. There are, however, some limitations to using the covariance functions in forestry. The computations are more demanding than those of simple linear models, and the polynomial models are not well behaved outside the range of the data. The inability to make good predictions outside the range of the data implies severe limitations to the estimation of gain and optimum selection age, since data in forestry are rarely available up to mature age.

Chapter 6

GENETIC GAIN ESTIMATES IN HEIGHT AND STRAIGHTNESS, AND INFERENCES FOR OPTIMUM SELECTION AGE

6.1 Introduction

Forest trees have long generation intervals; early indirect selection is preferred as it results in shorter generation intervals, and may lead to increased gain per unit of time, reduced testing costs, and greater adaptability to market changes (Magnussen 1988). Also, early selection offers the means for quicker incorporation of gains into production, as parents to be used for seed production can be selected early, and seed orchards can be culled early.

Optimum selection age is usually defined as the age at which genetic gain per year of breeding cycle is maximized, and is critical to the efficiency of any tree breeding program. Thus, identification of the optimum age for early indirect selection has been of major interest to tree breeders (e.g. Ballochi *et al.* 1993, King and Burdon 1991, McKeand 1988, Riemenschneider 1988, Xie and Ying 1996). Unfortunately, traits are rarely measured at all ages up to harvest. Therefore, optimum selection age can only be estimated using models which make it possible to predict genetic correlations between ages other than those at which assessments were made. This has resulted in the development of predictive models for genetic correlations. Lambeth (1980) used phenotypic correlations as an approximation to genetic correlations, and showed that correlations between heights at different ages were predictable based on the natural logarithm of the ratio of the younger and the mature age (LAR). He suggested that age-age genetic correlations were approximately equal to their corresponding phenotypic correlations, and therefore phenotypic correlations could be used in place of

corresponding genetic correlations. The logarithm-based models have been widely used by tree breeders to make decisions on optimum selection age (e.g. King and Burdon 1991, McKeand 1988, Riemenschneider 1988, Xie and Ying 1996).

Due to poor juvenile-mature correlations reported from some early studies of *P. taeda*, it was concluded that selection could be made reliably only after half-rotation age (e.g. Wakeley 1971). However, recent studies, which have used either a biological (gain per unit of time) or economic (present value of gain per unit of time) criterion, indicate that the optimum selection age for height in *P. taeda* could be as young as 4 years (Newman and Williams 1991), or between 6 and 8 years (Ballochi *et al.* 1993, Lambeth 1980, Li *et al.* 1996, McKeand 1988).

Studies on the prediction of genetic gain and optimum selection age have focussed on growth traits, and there appear to be none on straightness. Results from this study (Chapter 4) indicate that age-age genetic correlations for straightness are moderate to high, and Chapter 3 indicate that heritability is low, and its phenotypic variation moderate (coefficient of variation 14%). This indicates that selection on straightness will make good progress, and opportunity for early selection may exist. The results also indicate that selection for straightness could be less efficient relative to height; as straightness has lower heritability estimates, and lower age-age genetic correlations. Therefore, genetic gain per year in straightness is expected to be lower and optimum selection age to be higher than those for height.

There appear to be no estimates of genetic gain and optimum selection age for height in *P. taeda* grown in the tropics, an issue of concern as fast-growing tree crops exhibit different genetic correlations from slower growing ones at similar ages (Magnussen 1988). Given the fact that growth rates of *P. taeda* in the tropics can be substantially higher than those achieved in temperate regions, potential genetic gain and the optimum selection age may differ between these regions.

Genetic correlations are major determinants of gain and optimum selection age, and the model used in predicting them will affect the predicted gain and may influence the optimum selection age. There is increasing evidence demonstrating that genetic correlations are much higher than corresponding phenotypic correlations (Barnes 1992, Lambeth *et al.* 1983, Pswarayi *et al.* 1996, Riemenschneider 1988), implying that the

Lambeth model underestimates genetic correlations, and hence underestimation of genetic gain and may consequently affect the optimum age of selection. The economic consequences of not estimating optimum ages efficiently may be high in long rotation forest trees, and the Lambeth model has not been tested using data from pines grown in tropical regions. One of the problems with estimates of genetic correlations is that assessments of genetic tests are carried out before maturity (harvest age), and mostly at less than half harvest age. For example, Riemenschneider (1988) used age-age genetic correlations for height from tests assessed up to 7 years in *P. banksiana*, Matheson *et al.* (1994) used age-age genetic correlations for diameter up to 14 years of age in *P. radiata*, and McKeand (1988) used age-age genetic correlations for height for up to 16 years in *P. taeda*. Therefore, the genetic correlations involving mature age traits are extrapolations. The assumption that a linear relationship holds outside the range of the data may not be true; therefore, data that involve the trait at harvest age is preferred.

A constant heritability has been used in predicting the optimum age for selection in *P. taeda* (Gonzalez and Richards 1988, Lambeth 1980), but evidence suggests that heritability estimates for tree height in *P. taeda* increase with age, at least up to half mature age (Balocchi *et al.* 1993, Franklin 1979). Therefore, erroneously assuming heritability is constant will overestimate gain from early selections and may also affect estimates of the predicted optimum age for selection.

One of the major factors which influences the generation interval, and hence optimum selection age, is the age at which the species becomes sexually mature and produces seed. For example, *P. taeda* in Zimbabwe flowers at 10 years of age (Barnes[2], personal communication). Artificially inducing flowering in *P. taeda* has been successful in the USA, with induction achieved at less than 3 years of age (Bramlett *et al.* 1995, Burris *et al.* 1991).

The objective in this Chapter is to estimate annual genetic gain in height and straightness, and to determine the optimum age for selection of height in *P. taeda* using tests assessed up to harvest age. Selection was based on height because height is a good predictor of volume at rotation age (Foster 1986, Lambeth *et al.* 1983), heritability for

---

[2]R. D. Barnes, Oxford Forestry Institute, Oxford, UK.

height is higher than that for diameter in *P. taeda* (Foster 1986), and height is more easily and more accurately assessed at young ages than is diameter or volume. Furthermore, models for predicting age trends of age-age genetic correlations for height exist. The effects of using different models for predicting age-age genetic correlations. taking into account age-related changes of heritability, and of reducing the flowering age (hence generation interval) were also explored here.

6.2 Materials and Methods

Data

The genetic tests have been described in detail in Chapter 3.

Statistical analyses

Gain per year from mass selection was calculated for direct selection at rotation age (assumed to be 22.5 years) and for indirect selection at younger ages. Gain per year from direct selection on trait *y* was given by (Falconer 1989):

$$Gain_y = i_y h_y^2 \sigma_{py} L_y^{-1} \qquad (6.1)$$

and gain per year for trait *y* from indirect selection on trait *x* is given by:

$$Gain_y = i_x h_x h_y r_A \sigma_{py} L_x^{-1} \qquad (6.2)$$

where:

$Gain_y$ = gain per year for trait *y*,

r_A = genetic correlation,

h_x, h_y = square root of individual tree heritability for traits *x* and *y*, respectively,

i_x, i_y = selection intensity for traits *x* and *y*, respectively.

where selection intensity at 1-9 years was 2.665 (1:100), at 10-14 years 2.421 (1:50) and that at 15 and older 2.154 (1:25) in all traits,

σ_{py} = phenotypic standard deviation for trait y (assumed to be 1.732 metre for height, and 0.650 for straightness; Chapter 3),

L_x, L_y = generation interval for traits x and y, respectively.

It was assumed that it takes three years to establish genetic tests for the next generation once the trees are in flower. Since *P. taeda* flowers at age 10 in Zimbabwe, the generation interval for selecting at ten years or younger was flowering age + 3 years (13 years), while that for selection at older ages was selection age + 3 years. Gain and optimum selection age were also estimated assuming that selected trees could be induced to flower at 7, 5 and 3 years of age. In this case, the generation interval for selecting at any of these flowering ages or younger was flowering age + 3 years, while that at older ages was selection age plus 3. Initially, heritability was assumed to be constant with age (0.2); subsequently, changes in heritabilities estimated by Model 4 were taken into account in calculating gain and estimating optimum selection age. Predictions of age to age genetic correlations for height were made using four models: Lambeth model in which r_P is assumed to equal r_g, and r_P predicted by:

Model 1: $r_P = 1.02 + 0.308 \log_e$ (younger age/older age). (6.3)

The equality of r_P and r_g was removed:

Model 2: $r_g = \beta_0 + \beta_1 \log_e$ (younger age/older age) (6.4)

where β_0, β_1 were derived from estimates of r_g at 1.5, 9.5, 13.5 and 22.5 years. Model 3 was derived as for Model 2 but age difference was used as predictor:

Model 3: $r_g = \beta_0 + \beta_1$ (Age difference). (6.5)

Model 4 was derived using covariance functions (Kirkpatrick *et al.* 1990). Time trends of heritability estimates was modelled using covariance functions. The covariance between records taken at ages t_1 and t_2 is:

Model 4: $$T(t_1,t_2)=\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\hat{C}_{ij}\phi_i(t_1)\phi_j(t_2), \quad (6.6)$$

where T is the covariance function, k is the order of fit, $\hat{C}$ is a symmetric coefficient matrix associated with the covariance function, and ϕ are orthogonal polynomials.

Then,

$$r_g=\frac{cov_A(t_1,t_2)}{\sqrt{var_A(t_1)}\sqrt{var_A(t_2)}} \quad (6.7)$$

and heritability at any age (t_i) is estimated as,

$$h_{t_i}^2=\frac{var_A(t_i)}{var_P(t_i)} \quad (6.8)$$

The trend of the heritability estimated by Model 4 is shown in Figure 6.1, and that of the age-age genetic correlations predicted by the four models is shown in Figure 6.2.

Table 6.1. Summary of heritability estimates for straightness and height for four ages, and age-age genetic correlations between straightness at each of the ages and straightness at 22.5 years, and trait-trait genetic correlations between height at each of the ages and straightness at 22.5 years.

Age	Straightness		Height	
	h^2	r_A	h^2	r_A
1.5	0.01	-0.05	0.22	0.14
9.5	0.11	0.66	0.50	0.22
13.5	0.09	0.92	0.23	0.28
22.5	0.21	1.00	0.22	0.52

For straightness, it was not possible to calculate the trends in age-age correlations and heritability, therefore genetic gain and optimum age calculations were based on the ages at which the tests were assessed. The parameters used for predicting gain and optimum selection age for straightness are shown in Table 6.1.

6.3 Results

6.3.1 Height

Genetic gain and optimum selection age, assuming a constant heritability

Figure 6.3 shows genetic gain estimates at different ages using four different flowering ages, with age-age genetic correlations derived using the four models, and assuming a constant heritability at all ages. The optimum ages are shown in Table 6.2.

Flowering age had a large influence on the genetic gain expected and also on the optimum age for selection. Gain at optimum selection age increased with reduced flowering age by as much as 100% when flowering was reduced from 10 to 3 years. Optimum selection age was decreased with reduced flowering age. For each of the flowering ages, the optimum selection age was equal to the flowering age.

Where flowering age was 10 and 7 years, and genetic correlations were estimated using the Models 2-4, the graphs show that selecting between 3-10 years, and between 3-7 years, respectively, could be achieved with little loss in genetic gain. However, selecting after the optimum age resulted in rapid decline in gain at all the flowering ages using Models 2-4. When estimates from Model 1 were used, selecting before the optimum age resulted in large reductions in gain at all the flowering ages.

For each flowering age, the model used for estimating the age-age genetic correlations affected the magnitude of the expected gain but not the optimum selection age. The genetic gain for each of the flowering ages was 30-100% lower when the genetic correlations were estimated by Model 1 than by other three models.

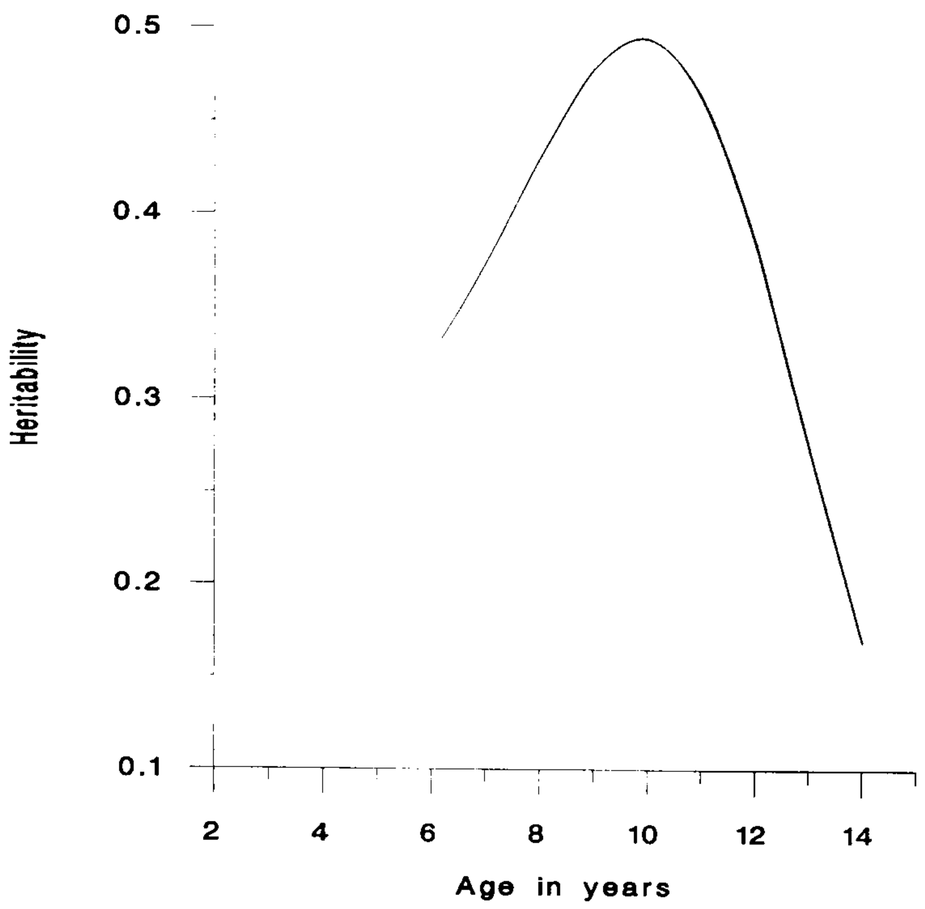


Figure 6.1. Predicted trends of heritability of height over time using Model 4.

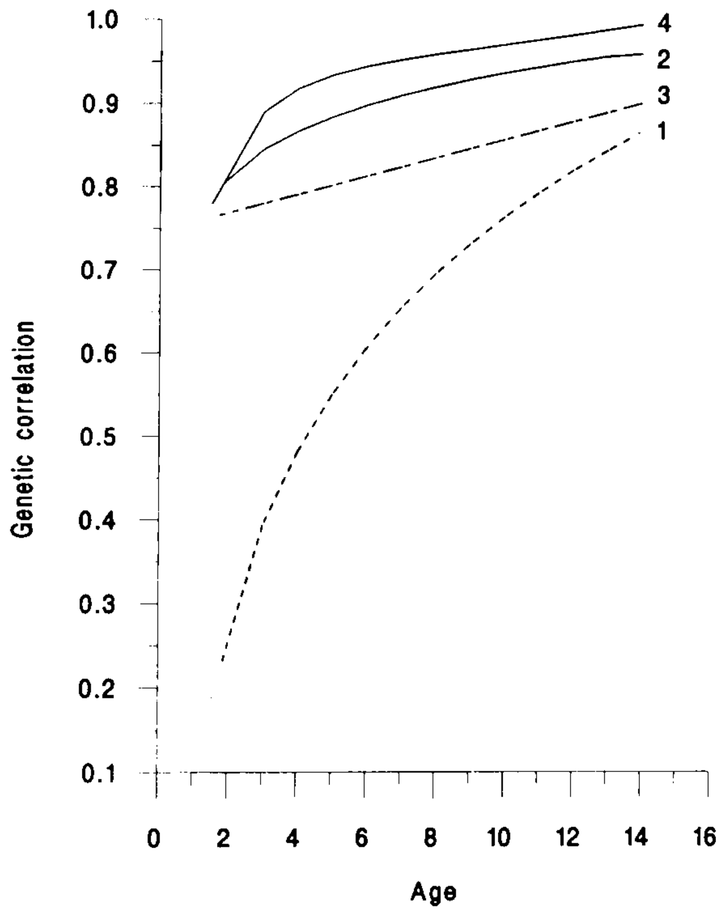
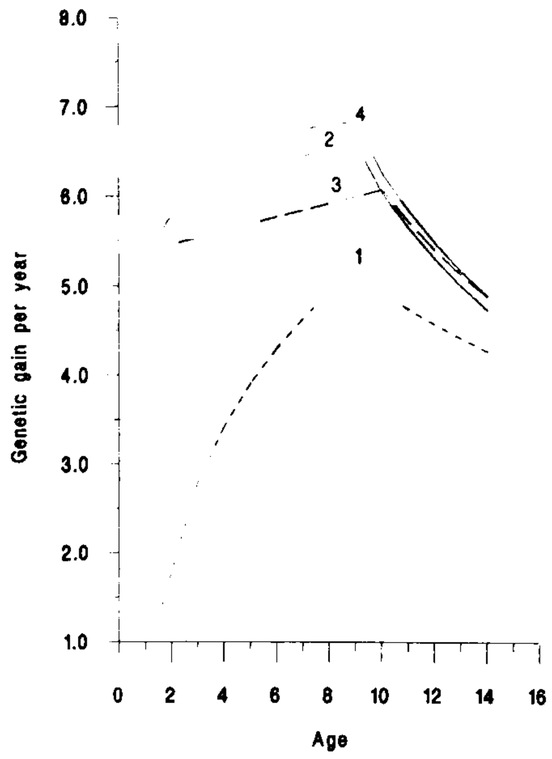
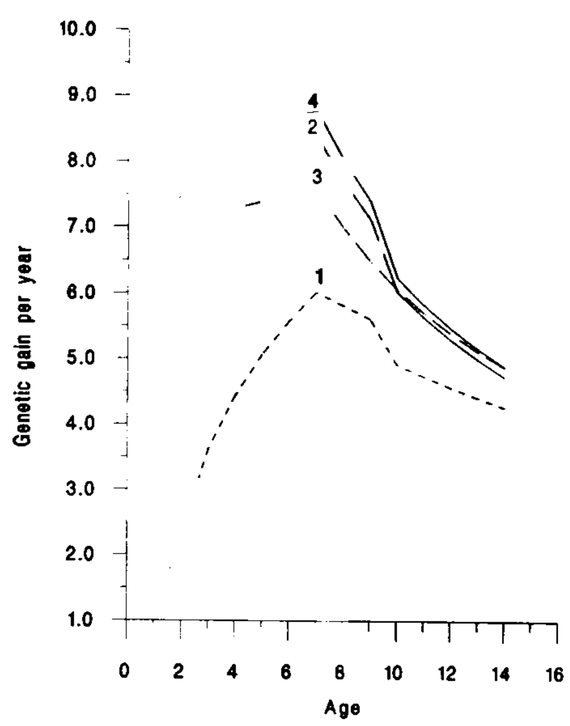
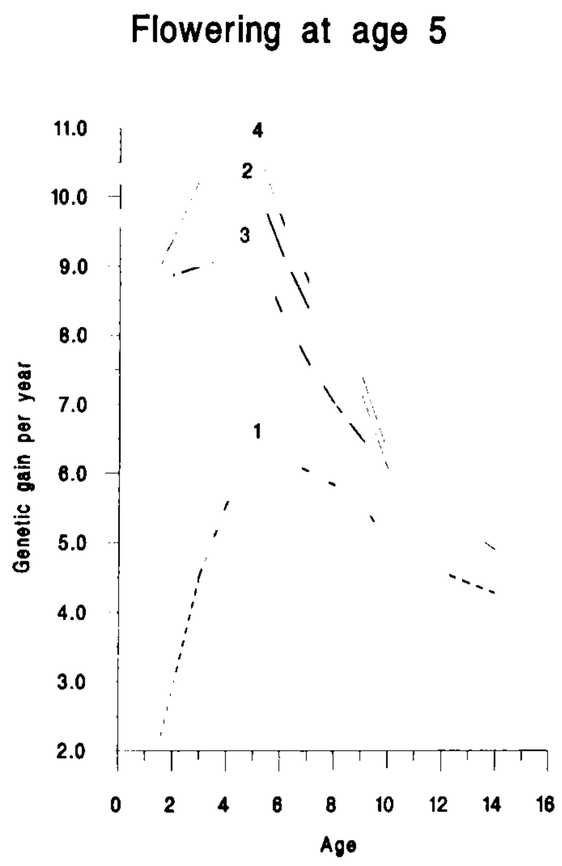


Figure 6.2. Predicted trends with age for genetic correlations for height with 22.5-year data using Models 1-4.







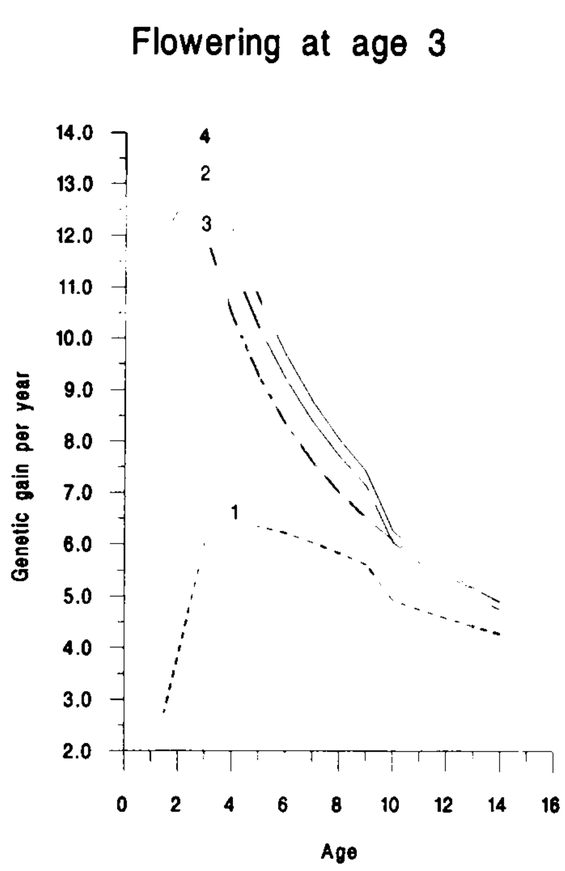


Figure 6.3. Effect of the model for predicting age-age genetic correlations on estimated genetic gain (x100) and optimum selection age for height when heritability is assumed constant with age. Numbers 1-4 refer to Models 1-4, respectively.

Annual genetic gains for height from selection at ages between 3 and 14 years were greater than that at harvest age (0.029), confirming the efficiency of early selection. The lowest genetic correlation at 3 years was 0.4 (Model 1, Figure 6.2), indicating that the genetic correlation for height between young ages and rotation age does not have to be high for early selection to be more efficient than selecting at rotation age.

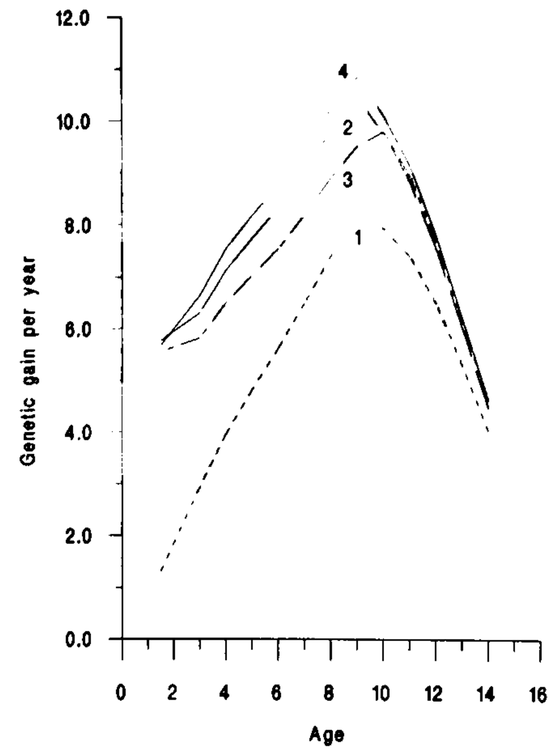
Genetic gain and optimum selection age taking into account age-related changes in heritability

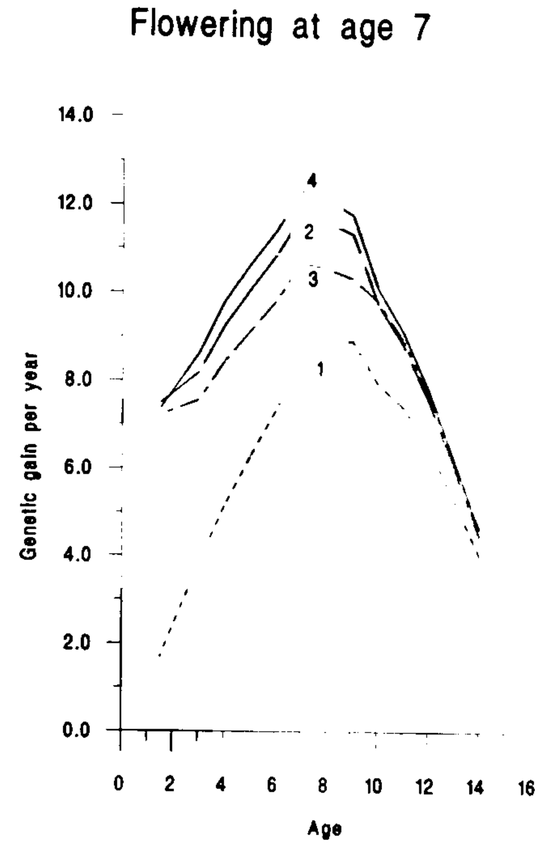
Effects of including age-related changes in heritability on genetic gain and optimum selection age are shown in Figure 6.4 and Table 6.2.

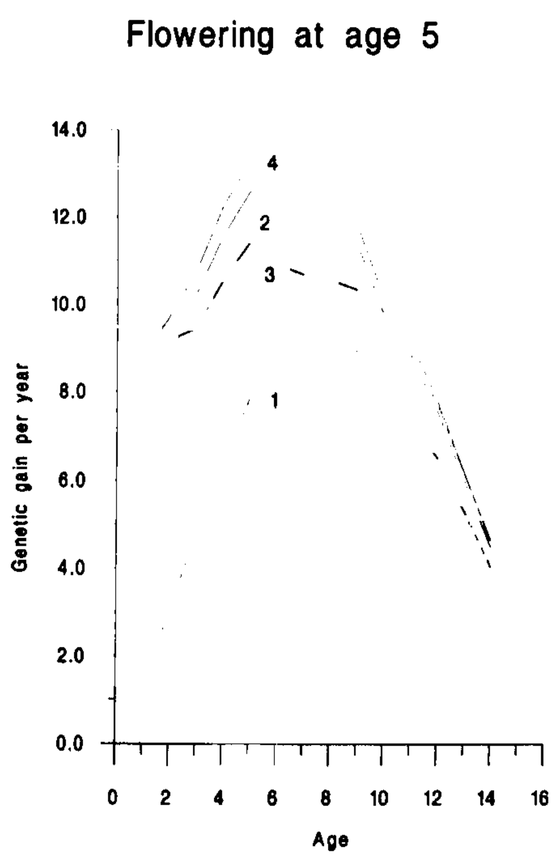
The predicted genetic gain was higher than that obtained using the constant heritability: 37% higher at flowering age of 10 years, 28% higher for flowering age of 7 years, and 5% higher for flowering age of 3 years. This is attributed to higher true heritability estimates than the assumed constant value (Figure 6.1).

When age-age genetic correlations were estimated by Models 2-4, the optimum selection age decreased with reduced flowering age, and was similar to the one estimated assuming constant heritability. However, when the age-age genetic correlations were estimated using Model 1, the optimum selection age was 9 years for each of the flowering ages. This is attributed to a combination of low age-age genetic correlations involving young ages estimated by Model 1 (Figure 6.2) and the low heritability estimates at young ages (Figure 6.1). The difference between optimum selection ages predicted by Model 1 and the other three models varied from 0 to 7 years (Table 6.2).

When age-related changes in heritability were considered, it was critical to select at the optimum age, particularly were the genetic correlations were estimated by Models 2-4.







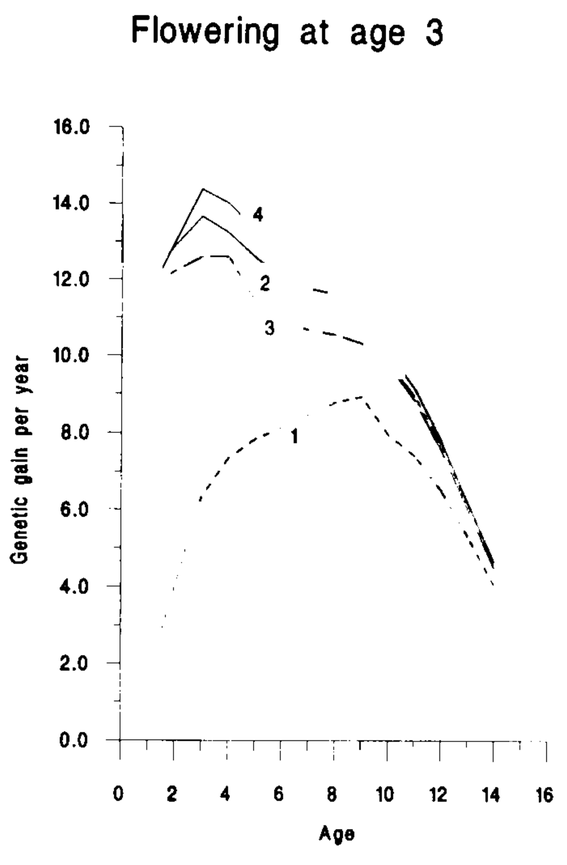


Figure 6.4. Effect of the model for predicting age-age genetic correlations on estimated genetic gain (x100) and optimum selection age for height when age-related changes in heritability with age are considered. Numbers 1-4 refer to Models 1-4, respectively.

Table 6.2. Summary of the optimum selection ages for height for four flowering ages using (1) constant heritability and (2) age-related heritability estimates, and four models for estimating age-age genetic correlations.

	Model number	Flowering age			
		10 years	7 years	5 years	3 years
Constant Heritability	1	9	7	5	3
	2	9	7	5	3
	3	10	7	5	3
	4	9	7	5	3
Variable Heritability	1	9	7	5	3
	2	9	9	9	9
	3	10	7	5	3
	4	9	7	5	3

6.3.2 Straightness

For straightness, genetic gain was maximized at 9.5 years for each of the flowering ages, when heritability was assumed either constant or variable (Table 6.3). The genetic gain per year when selection is made at 1.5 years was negative, reflecting the negative genetic correlation for straightness between this age and rotation age (Table 6.1). Assuming a constant heritability, early selection for straightness was more efficient than direct selection at rotation age, except for selection at 1.5 years. In contrast, under variable heritability, early selection was not always more efficient than selecting at rotation age, due to a much higher heritability at rotation age (Table 6.1).

Table 6.3. Predicted genetic gain per year (x100) in straightness for four flowering ages using a constant heritability (0.2), and age-related heritability estimates.

	Selection age (years)	Flowering age 10 years	7 years	5 years	3 years
Constant Heritability	1.5	-0.13	-0.17	-0.22	-0.29
	9.5	1.76	1.83	1.83	1.83
	13.5	1.75	1.75	1.75	1.75
	22.5	1.23	1.23	1.23	1.23
Variable Heritability	1.5	-0.03	-0.04	-0.05	-0.07
	9.5	1.34	1.39	1.39	1.39
	13.5	1.21	1.21	1.21	1.21
	22.5	1.30	1.30	1.30	1.30

6.3.3 Selecting height for improving straightness

Genetic gain expected in straightness from indirect selection on height was lower than indirect selection of straightness at young ages, except at 1.5 years (Tables 6.4). This is attributable to lower genetic correlations between height and straightness compared with age-age genetic correlations for straightness (Table 6.1). Selecting straightness at 1.5 years is predicted to make negative genetic progress, but selecting height at the same age for improvement in straightness at harvest age would make positive genetic progress.

Under the assumption of constant heritability early selection on height resulted in less gain in straightness at 22.5 years compared to selecting on height at same age. This is attributable to a much higher genetic correlation between straightness and height at harvest age than between straightness at harvest age and height at younger ages. In contrast, when flowering age was reduced to 3 years, the optimum selection age was 1.5 years.

Table 6.4. Predicted genetic gain per year (x100) in straightness for four flowering ages when selection is made on height using (1) constant heritability (0.2) and (2) age-related heritability estimates.

	Selection age (years)	Flowering age 10 years	7 years	5 years	3 years
Constant Heritability	1.5	0.37	0.49	0.61	0.81
	9.5	0.59	0.61	0.61	0.61
	13.5	0.53	0.53	0.53	0.53
	22.5	0.64	0.64	0.64	0.64
Variable Heritability	1.5	0.40	0.52	0.65	0.87
	9.5	0.95	0.99	0.99	0.99
	13.5	0.59	0.59	0.59	0.59
	22.5	0.69	0.69	0.69	0.69

6.4 Discussion

The length of the breeding interval (flowering age) had a strong effect on the genetic gain and optimum selection age when heritability was assumed to be constant: as the breeding interval was reduced, gain increased and optimum selection age decreased. In contrast, when heritability was varied, changing the breeding interval had no impact on the optimum selection age in the presence of low genetic correlations predicted by Lambeth's model (Model 1). Magnussen (1989) also found that the effect of the genetic correlations depended on the size and variation of the heritability estimates. Therefore, using the covariance functions (Kirkpatrick *et al.* 1990), modelling of the 'coefficient of genetic prediction' ($h_x r_A h_y$, standardised genetic covariance or coheritability) as proposed by Baradat (1976), or modelling of the heritability itself as suggested by Wei and Borralho (1996), will ensure more accurate predictions of gain and optimum selection age.

The results (Figure 6.3) showed that at constant heritability, there was little difference in gain from selecting between 3-10 years when flowering age is 10 years, or between 3-7 years when flowering age is 7 years, using Models 2-4. This result supports Lambeth's (1980) observation that there is a range of ages at which selection is nearly as efficient as at the optimum selection age. However, this result was not observed when age-related changes in heritability were considered, and when Models 2-4 were used. Therefore, Lambeth's observation it is unlikely to be appropriate for breeding decisions of *P. taeda* in Zimbabwe, since heritability changes with age.

The consequences of using the Lambeth's phenotypic model was more than four times less predicted gain than that by the other three models, and an overestimation of the optimum selection age by 6-7 years. The underprediction of potential gain at any age will lead to unfavourable investment appraisal of tree breeding programmes, and the conservative prediction of optimum selection age means that breeding programmes would deliver more gain, and are therefore being run inefficiently. Inaccurate predictions of gain and optimum selection age will also limit the ability to identify appropriate research priorities (e.g. the importance of identifying methods which induce early flowering).

Although genetic correlations and the generation interval are major determinants of gain and optimum selection age, heritability and its variation with age, and variation of the selection intensity with age, are also important. When heritability was assumed constant, the Lambeth model underpredicted gain, but gave comparable predictions of optimum selection age to other models. The results show that if the objective is to estimate optimum selection age, and a constant heritability is assumed, the Lambeth model gives good estimates.

The optimum selection age for straightness was 9.5 years, and was insensitive to changes in flowering age when age-related changes in heritability were considered in calculating gain. Due to lack of predictions of trends of age-age correlations and heritability for this trait, the actual optimum age at each of the flowering ages may differ from that predicted here. Gain in straightness at rotation age when selection is made on straightness at 1.5 years was negative, but when selection was on height at the same age, gain was positive. This result reflects the difficulties of assessing straightness at an early

age. The results also indicate that genetic gain per year for straightness is comparably lower and optimum selection age higher than that for height, due to lower heritability estimates, and lower age-age genetic correlations.

Although the study assumes that only one selection age is used, a two-stage approach may be more efficient for improving both height and straightness. With the two-stage approach, early selection would be based on height only and later selection on straightness only.

Economic factors were not considered in this study, and these are likely to reduce the optimum age of selection further (McKeand 1988). To determine the economic optimum, the present values, or better still the net present values, of the gains can be calculated. The net present values which include the costs of more rapid turnover of generations and more frequent seed orchard establishment would be preferred, but information for such an analysis is unlikely to be available for many tree breeding programmes.

Most studies suffer from the problem of not having assessments at maturity, and therefore of relying on extrapolations which may be inaccurate or have large errors (e.g. Matheson *et al.* 1994, McKeand 1988, Riemenschneider 1988). The strength of this study is that assessments at near-harvest age were available, allowing realistic predictions of rotation age gains. However, the study also suffers from the problem of having few point estimates, a problem in many other forestry studies (e.g. 3 points, King and Burdon 1991; 4 points, McKeand 1988).

The reproductive biology of *P. taeda* is a barrier to juvenile breeding of the species in Zimbabwe. Nevertheless, the study has demonstrated that, were early flowering to be induced, optimum selection age would be reduced from 10 to 3 years and annual genetic gain increased by more than 100%. Therefore, the extra cost of flower induction must be compared against these additional gains expected from lowering the breeding interval. Other options, such as the selection of sites with early flowering potential should also be explored in order to reduce the breeding interval.

Chapter 7

GENOTYPE X ENVIRONMENT INTERACTIONS FOR HEIGHT AND STEM STRAIGHTNESS

7.1 Introduction

The presence of genotype x environment interactions complicates tree breeding programmes and may reduce the rate of genetic progress. Heterogeneity between forest sites, due to variation in soil and/or climatic conditions, may cause genotype x environment interactions (GE). The magnitude of the GE affects decisions on testing, selection and deployment. At the species level, GE is used in matching species to sites; at the family or individual level, it influences major elements of the breeding strategy such as the structure of the breeding population, and the selection of parents in the breeding population.

GE may be due to heterogeneity of variances measured at each of the sites, where ranking of genotypes in the various environments is unaffected ('pseudo-interaction', Dickerson 1962). It may also be due to both heterogeneity of variances and rank changes (Dickerson 1962). Breeding strategy will largely be influenced by the latter.

The method most used by tree breeders for determining the magnitude of GE and implications for gain is that of Falconer (1989), in which a trait measured on two environments is considered analogous to two traits in a single environment. Hence the genetic correlation across environments for a trait measured in two environments provides a measure of the magnitude of GE. The additive genetic correlation between two environments is given by the following equation:

$$r_A = \frac{Cov_A(x,y)}{\sigma_{Ax}\sigma_{Ay}} \qquad (7.1)$$

where r_A = additive genetic correlation, $Cov_A(x,y)$ = covariance of the trait in environments x and y, σ_{Ax} = additive genetic standard deviation of the trait in environment x and σ_{Ay} = additive genetic standard deviation of the trait in environment y. Genetic correlations which do not differ significantly from 1 suggest there is no GE, while genetic correlations significantly less than 1 indicate the presence of GE. This approach to detecting the presence of GE, first introduced in forestry by Burdon (1977); referred to as Type B correlations, has been widely used recently in tree breeding (Carson 1991, Matheson and Raymond 1984b, Johnson 1992, Johnson and Burdon 1990, Pswarayi *et al.* in press). A method of determining loss of potential gain is the estimation of efficiency of selecting at one site for planting at another site, using the method of indirect selection proposed by Falconer (1989). It appears that all recent GE studies in forestry have followed this approach, and examined the magnitude of genetic correlations and efficiency of selection across different sites when the trait has been assessed at the same age across sites. Since early selection is normally practised in tree breeding, what may be more important is the effect of early selection of a trait at one site for predicting mature age performance of the trait at another site.

When GE is present, and considered important, the options are to group environments, or to group genotypes (Raymond and Lindgren 1990), or simply to ignore it. The first option implies a multiple population or subline breeding strategy, where environments within which GE is approximately zero are grouped. This option is likely to be the most expensive as more than one breeding population will need to be managed and sites will need to be classified to determine regions of minimum GE. However, this strategy will result in the largest gain in the short term. The second option implies elimination of the most interactive families, and is likely to be less expensive than the first. However, some of the families most productive on average across all sites may be the most interactive. The third option is likely to be the least expensive but may result in largest losses in gain.

Most studies carried out in forest tree species report that whilst there is often statistically significant GE, GE is not of practical significance (Carson 1991, Johnson 1992, Johnson and Burdon 1990, Owino 1977b, Pswarayi *et al.* in press). GE was considered to be of no practical significance when either the genetic correlations were

high, and hence the potential loss of gain from selecting at one site for planting at another was low (e.g. Pswarayi *et al.* in press), or GE was due to heterogeneity of variances across sites (e.g. Owino 1977b).

The environments under which *P. taeda* in Zimbabwe grows vary greatly in terms of rainfall and altitude, and applicability of genetic parameters at one site to another site is unknown. In Chapter 3 it was discovered that pooling data across sites reduced heritability to less than that obtained at each of the individual sites, indicating the possible presence of GE. In addition, Zimbabwe is implementing the multiple population breeding strategy in *P. taeda*, and one of the reasons for adopting this strategy was to take advantage of GE by selecting for specific adaptations (Barnes 1989); however, this strategy was based primarily on the results from *P. patula* genetic tests. Genotype x environment interactions in *P. taeda* in Zimbabwe have not been quantified, and hence the appropriateness of the multiple population strategy for *P. taeda* has not been verified. In a related species planted in Zimbabwe, *P. elliottii*, Pswarayi *et al.* (in press) concluded that GE was not of sufficient practical significance to warrant sub-dividing the population. Therefore, there is a need to quantify the magnitude of GE to assist with making informed decisions on the breeding strategy for *P. taeda*.

Genetic tests of *P. taeda* in Zimbabwe were established across a range of environments, covering the diverse plantation sites. The results should indicate whether or not breeding populations should be sub-divided; if they are to be sub-divided, which parents to include in each of the sub-populations; and whether the same genetic parameters can be used across sites.

This Chapter quantifies GE for height and stem straightness, and its implications for breeding strategy of the Zimbabwean breeding population of *P. taeda*.

7.2 Materials and Methods

Data

Genetic tests located at 4 sites in Zimbabwe were assessed for height and stem straightness at 1.5, 9.5, 13.5 and 22.5 years of age. Details of the tests and assessment

procedures are described in Chapter 3.

The sites differed mainly in altitude and rainfall. Tarka (A) and Martin (C) genetic tests were located on low altitude sites (1005 and 1250 m ASL, respectively) in the Chimanimani area of Zimbabwe, and Stapleford (B) and Nyangui (D) were high altitude sites (1745 and 1882 m ASL). Three of the sites had high rainfall (1836-2364 mm per annum), while Martin had low rainfall (1016mm). Any GE reported here may reflect differential responses of families to different altitudes or rainfall. Location details of the sites are summarized in Chapter 3, Table 3.1.

Statistical analyses

Genetic correlations were estimated using bivariate individual tree model ASREML (Gilmour 1996). ASREML was used instead of DFREML, as for analyses described in previous chapters, because of its faster computational speed (Gilmour *et al.* 1995).

The assessments of the same trait at different sites were treated as different traits. These correlations are referred to as Type "B" correlations (Burdon 1977). For all the analyses, the traits were assessed on separate but genetically related trees across sites. Therefore, genetic covariances exist between any two traits, but there were no environmental covariances.

The following bivariate tree model was used:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \qquad (7.2)$$

where: $\mathbf{y}_1$, $\mathbf{y}_2$ = the vector of observations for traits 1 and 2, respectively,

$\mathbf{b}_1$, $\mathbf{b}_2$ = the vector of fixed effects for traits 1 and 2, respectively,

$\mathbf{a}_1$, $\mathbf{a}_2$ = the vector of random tree (additive genetic) effects for traits 1 and 2, respectively,

$\mathbf{c}_1$, $\mathbf{c}_2$ = the vector of additional uncorrelated random effects for traits 1 and 2, respectively,

$\mathbf{X_1}$, $\mathbf{X_2}$ = the incidence matrix for fixed effects for traits 1 and 2, respectively.

$\mathbf{W_1}$, $\mathbf{W_2}$ = the incidence matrix for additional random effects for traits 1 and 2, respectively,

$\mathbf{Z_1}$, $\mathbf{Z_2}$ = the incidence matrix for additive direct effects for traits 1 and 2, respectively,

$\mathbf{e_1}$, $\mathbf{e_2}$ = the vector of residual effects for traits 1 and 2, respectively.

The variance-covariance structure of the random effects of the bivariate tree model was as follows:

$$V\begin{bmatrix} a_1 \\ a_2 \\ c_1 \\ c_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} A\sigma^2_{a1} & Acov_{a12} & 0 & 0 & 0 & 0 \\ Acov_{a21} & A\sigma^2_{a2} & 0 & 0 & 0 & 0 \\ 0 & 0 & I\sigma^2_{c1} & Icov_{c12} & 0 & 0 \\ 0 & 0 & Icov_{c21} & I\sigma^2_{c2} & 0 & 0 \\ 0 & 0 & 0 & 0 & I\sigma^2_{e1} & 0 \\ 0 & 0 & 0 & 0 & 0 & I\sigma^2_{e2} \end{bmatrix} \quad (7.3)$$

where: σ^2_{a1}, σ^2_{a2} and cov_{a12} are the direct additive genetic variances and covariance for traits 1 and 2, respectively,

σ^2_{c1}, σ^2_{c2} and cov_{c12} are the corresponding additional random effect variances and covariance, and

σ^2_{e1}, σ^2_{e2} are the corresponding residual variances for traits 1 and 2, respectively.

The effects of the replicate and site were considered fixed, and the family (male x female interaction) was considered an additional uncorrelated random effect. Pre-adjusting the data for the block effects was used before each bivariate analysis as for analyses described in Chapter 4. The first starting values of variances were estimated from the univariate DFREML analyses. The standard errors of the genetic correlations were estimated directly by ASREML using the following approximate formula for obtaining variance of a ratio (Gilmour 1996):

$$\mathrm{Var}(r) = r^2\left[\frac{\mathrm{Var}(v(a))}{4v(a)^2} + \frac{\mathrm{Var}(v(b))}{4v(b)^2} + \frac{\mathrm{Var}(v(ab))}{v(ab)^2} + \frac{2\mathrm{Cov}(v(a),v(b))}{4v(a)v(b)} - \frac{2\mathrm{Cov}(v(a),v(ab))}{2v(a)v(ab)} - \frac{2\mathrm{Cov}(v(ab),v(b))}{2v(ab)v(b)}\right] \quad (7.4)$$

where:

$$r = \frac{v(ab)}{\sqrt{v(a)v(b)}} \quad (7.5)$$

Rank changes

In order to identify the parents which were most highly interactive (i.e. least stable), each of the 22 parents was ranked on its estimated breeding value at each site, and also over all sites for each trait. The absolute deviation of the ranking at each site from the overall ranking over all sites was calculated, and these were summed across the 4 sites. Then the mean rank deviation was calculated by dividing the total deviations by 4. The parents with the greatest mean rank deviations were considered most interactive. This method was developed by Matheson and Raymond (1984b), but they used phenotypic values rather than breeding values.

Index values

Index values were calculated to determine the aggregate genetic merit of the 22 parents. The estimated breeding values (BV) for each trait at four sites were combined together into appropriate index values. Each breeding value was weighted by the degree of importance. The degree of importance was determined according to representativeness of the genetic test sites to operational plantation areas. Since 80% of the plantations are found equally on sites similar to A and C, and very few plantations on B (15%) and D (5%), the weights for tests A, B, C, D were assigned as 0.4:0.15:0.4:0.05, and the aggregate genetic merit was determined by:

$$H = 0.4*BV_{site\ 1} + 0.15*BV_{site\ 2} + 0.4*BV_{site\ 3} + 0.05*BV_{site\ 4} \quad (7.6)$$

Efficiency of selection

The efficiency of selecting at site 1 (trait x) for planting at site 2 (trait y), relative to both selecting and planting at site 2 (trait y), was estimated using the following equation, assuming selection intensities and generation intervals are similar (Falconer 1989):

$$E = \left[h_x r_A h_y^{-1} \right] \text{ x } 100\% \qquad (7.7)$$

where r_A is the genetic correlation between traits x and y, h_x and h_y are the square root of the heritability of traits x and y, respectively. These efficiencies indicate the relative loss in genetic gain from selecting at one site for planting at another.

Efficiency of early selection - based on the ratio of genetic gain in the mature trait (22.5 years) at site 2 (trait y) from indirect selection based on an early trait (1.5 and 9.5 years) on site 1 (trait x), relative to gain genetic gain in the mature trait (22.5 years) at site 2 from indirect selection based on an early trait on site 2 (trait z) - may be a better criterion for examining GE since early selection within sites in efficient, and it is normally practised in conifers. For example, optimum selection age for height in *P. taeda* in the USA can be as young as 4 years (Newman and Williams 1991). In this study, calculations of efficiency of early selection were based on selection at 1.5 and 9.5 years only, because selecting earlier than 10 years was found to be most efficient (Chapter 6). Assuming the selection intensities are equal, the ratio of the efficiencies is given by:

$$E = \frac{h_x r_{xy} h_y^{-1} L_x L_y^{-1}}{h_z r_{zy} h_y^{-1} L_z L_y^{-1}} \text{ x } 100\% \qquad (7.8)$$

where r_{xy} = the genetic correlation between traits x and y; r_{zy} = the genetic correlation between traits z and y; h_x, h_z and h_y are the square root of the heritability of traits x, z and y, respectively; L_x, L_z and L_y are the generation intervals of traits x, z and y, respectively. The ratio of the efficiencies of early selection within and across sites reduces to:

$$E = \left[h_x r_{xy} h_z^{-1} r_{zy}^{-1}\right] \text{ x } 100\% \qquad (7.9)$$

7.3 Results

7.3.1 Genetic correlations

Traits assessed at the same ages across sites

Genetic correlations between heights assessed at the same age across sites are shown in Table 7.1. These were high at 1.5 years, ranging from 0.77 to 0.95, and moderately high at 9.5 and 13.5 years (0.62-0.95). At 22.5 years, genetic correlations between site A and B and that between sites D and B were high, 0.90 and 0.73 respectively. The genetic correlation between site A and D was particularly low (0.18) indicating considerable parent x site interactions. The other correlations were moderate (0.51-0.58), indicating significant parental rank changes. The precision of the genetic correlations increased with increasing values of the correlation itself, as expected.

Genetic correlations between stem straightness scores assessed at the same age across sites are shown in Table 7.2. These were high at 1.5 years, except those involving site D which were highly negative. At 9.5, 13.5 and 22.5 years, the genetic correlations were low except those between sites A and C, and sites B and D. This result is expected since the site conditions for sites A and C (low altitude sites) were similar, and those for sites B and D (high altitude sites) were also similar.

Traits assessed at different ages across sites

Genetic correlations for height assessed at an early age at one site and that corresponding to maturity (22.5 years) at the other sites are shown in Table 7.3. The genetic correlations involving site B at 22.5 years and the other sites at 1.5 years were all high (0.71-0.79), indicating that there will be little loss in gain at mature age at site B when selections are made at the other sites at 1.5 years. However, selecting at 1.5

years at site B for planting at the other sites would result in moderate loss in gain (genetic correlations, 0.43-0.64). Genetic correlations involving site D and sites other than site B at 1.5, 9.5 years and 22.5 years were low, indicating that site D differed from sites A and C.

Table 7.1. Estimates of genetic correlations for heights assessed at the same age across four sites. Standard errors of the genetic correlations are in parenthesis.

Age	Site			
		B	C	D
1.5	A	0.9 (0.09)	0.95* (0.13)	0.88 (0.10)
	B		0.95* (0.08)	0.77 (0.13)
	C			0.78 (0.14)
9.5	A	0.89 (0.10)	0.90 (0.05)	0.66 (0.15)
	B		0.94 (0.10)	0.95* (0.09)
	C			0.73 (0.13)
13.5	A	0.64 (0.17)	0.79 (0.11)	0.62 (0.21)
	B		0.65 (0.16)	0.77 (0.14)
	C			0.72 (0.16)
22.5	A	0.90 (0.15)	0.51 (0.25)	0.18 (0.28)
	B		0.58 (0.25)	0.73 (0.17)
	C			0.52 (0.24)

*Constrained by the program to ≤ 0.95.

The genetic correlations between stem straightness scores at one site at 1.5 years and that at 22.5 years at other site were low, other than between sites A and D (0.77) and sites D and B (0.69). The corresponding correlations for 9.5 years and 22.5 years were also low, except between sites B and D (0.8) , sites C and A (0.71), and sites D and B (0.66).

Table 7.2. Estimates of genetic correlations for stem straightness assessed at the same age across four sites. Standard errors of the genetic correlations are in parenthesis.

Age	Site			
		B	C	D
1.5	A	0.86 (0.46)	0.95* (0.23)	-0.81 (0.55)
	B		0.94 (0.24)	-0.95* (1.00)
	C			-0.82 (0.44)
9.5	A	0.35 (0.29)	0.91 (0.08)	-0.11 (0.28)
	B		0.41 (0.30)	0.87 (0.17)
	C			0.14 (0.29)
13.5	A	-0.25 (0.28)	0.88 (0.10)	-0.24 (0.26)
	B		0.05 (0.29)	0.95* (0.05)
	C			0.04 (0.27)
22.5	A	0.66 (0.30)	0.84 (0.19)	0.06 (0.38)
	B		0.44 (0.23)	0.95* (0.11)
	C			0.22 (0.29)

*Constrained by the program to ≤ 0.95.

Table 7.3. Estimates of genetic correlations between height at either 1.5 or 9.5 years at one site and that at 22.5 years at another site. Standard errors of the genetic correlations are in parenthesis.

Age	Site	22.5 years			
		A	B	C	D
1.5 years	A		0.71 (0.18)	0.60 (0.21)	0.59 (0.21)
	B	0.64 (0.18)		0.54 (0.22)	0.43 (0.22)
	C	0.71 (0.17)	0.79 (0.15)		0.45 (0.22)
	D	0.14 (0.29)	0.75 (0.18)	0.39 (0.27)	
9.5 years	A		0.61 (0.18)	0.55 (0.21)	0.39 (0.22)
	B	0.82 (0.20)		0.71 (0.26)	0.62 (0.24)
	C	0.72 (0.15)	0.71 (0.16)		0.39 (0.22)
	D	0.38 (0.25)	0.62 (0.24)	0.49 (0.24)	

Table 7.4. Estimates of genetic correlations between stem straightness at either 1.5 or 9.5 years at one site and that at 22.5 years at another site. Standard errors of the genetic correlations are in parenthesis.

Age		Site			
			22.5 years		
		A	B	C	D
1.5 years	A		-0.12 (0.38)	0.20 (0.36)	0.77 (0.28)
	B	-0.27 (0.57)		-0.71 (0.31)	-0.18 (0.52)
	C	-0.24 (0.36)	-0.25 (0.28)		0.16 (0.30)
	D	0.22 (0.55)	0.69 (0.38)	0.29 (0.41)	
9.5 years	A		0.11 (0.27)	0.46 (0.22)	-0.13 (0.28)
	B	0.49 (0.40)		0.02 (0.34)	0.80 (0.25)
	C	0.71 (0.22)	0.43 (0.24)		0.23 (0.29)
	D	0.06 (0.39)	0.66 (0.20)	0.03 (0.29)	

7.3.2 Parental rank changes

The parental rank changes at each site and over all sites for height and stem straightness are shown in Tables 7.5 and Table 7.6, respectively. The ranks were calculated only for height and straightness at 22.5 years since the genetic correlations between sites were particularly low at this age, indicating considerable GE. In some parents, rank changes for height and straightness were large. For example, parent 36 was ranked 2 in height at site D and 19 at site C, and parent 32 was ranked 22 at site A and 4 at site C (Table 7.5), indicating large parent x site interactions. Similarly for straightness, parent 14 ranked last at site C and was ranked 4 at site D, and parent 20 was ranked 1 at site C and 21 at site D. As shown by the mean rank deviations, the parents varied in their stability of rank across sites. For height, 3 parents were highly interactive with rank deviations of equal or greater than 7. About half of the parents were fairly stable in ranking with rank deviation less than 3. For straightness, 5 parents were considered highly interactive with rank deviations of greater than 5. About half of the parents were considered fairly stable in ranking with rank deviation less than 3 (Table 7.6). The parents with low ranking in height also had a low ranking in straightness. For example, parent 14 was ranked 20 in height and 18 in straightness, parent 37 was ranked 21 in height and 22 in straightness, and parent 31 was ranked 19 in both traits. However, the parents with high ranking in height were not necessarily those with high ranking in straightness.

Although the most interactive parents were of average or poor performance across all sites, a few were of superior performance across all sites. For example, parent 29 was among the most interactive and had the 4th best overall rank across all sites for height; similarly for straightness, parent 36 was the most interactive and had the 3rd best overall rank across all sites. In contrast, parent 8 was ranked the best over all sites in height, and had the lowest rank deviation.

Table 7.5. Ranking of parents at each site for height at 22.5 years, overall rank across all sites, and average rank deviation from the overall rank.

Parent	Site					Mean Deviation
	A	B	C	D	All sites	
8	1	2	1	1	1	0.25
9	13	9	2	19	8	5.75
10	6	8	14	12	9	3
11	9	15	8	16	16	4
12	4	13	7	20	12	5.5
13	15	20	9	4	14	5.5
14	20	18	21	18	20	1.25
15	2	4	5	6	2	2.25
16	7	6	17	3	5	4.25
17	8	7	15	13	7	3.75
20	5	10	11	17	10	3.25
23	18	17	13	11	18	3.25
29	16	1	16	5	4	7
30	3	5	3	9	3	2
31	21	21	6	15	19	5.25
32	22	14	4	7	17	7.75
33	17	22	20	22	22	1.75
34	11	16	18	8	15	3.75
36	19	3	19	2	6	8.25
37	14	19	22	21	21	2.5
39	12	11	12	10	13	1.75
40	10	12	10	14	11	1.5

Table 7.6. Ranking of parents at each site for stem straightness at 22.5 years, overall rank across all sites, and average rank deviation from the overall rank.

Parent	Site					Mean Deviation
	A	B	C	D	All sites	
8	5	16	9	20	14	5.5
9	6	9	13	8	10	2.5
10	13	3	20	10	11	5
11	14	22	10	22	21	5
12	9	15	11	9	12	2.5
13	15	12	6	6	5	4.75
14	22	18	22	4	18	5.5
15	4	6	17	14	9	5.25
16	1	2	2	7	2	1.5
17	3	5	3	1	1	2
20	2	10	1	21	8	7
23	19	13	15	13	16	2.5
29	12	8	5	3	6	3
30	7	7	4	11	4	3.25
31	20	19	7	18	19	3.5
32	21	4	16	12	13	5.25
33	8	11	8	5	7	2
34	11	14	19	15	15	2.25
36	18	1	12	2	3	6.75
37	16	21	14	19	22	4.5
39	10	17	21	16	17	1
40	17	20	18	17	20	2

7.3.3 Index values

Breeding values for height of parents at individual sites, and those from data pooled across sites, with or without weighting based on representativeness to plantation sites, are shown in Table 7.7. When breeding values at the individual sites were not weighted according to importance of sites, their correlations with the aggregate genetic merit were 0.70, 0.90, 0.59 and 0.71, for sites A-D, respectively. The breeding values at the individual sites were weighted by the correlation with the aggregate breeding values, and these weights were 0.24:0.44:0.17:0.28. When the breeding values were weighted according to importance of the sites, the correlations with the new aggregate breeding values were 0.88, 0.69, 0.74 and 0.43, for sites A-D, respectively.

The breeding values for stem straightness are shown in Table 7.8. The correlations between the aggregate breeding values and those at individual sites, when importance of sites was not taken into account, were 0.57, 0.83, 0.58 and 0.74 for sites A-D, respectively. Unlike height, the weights used to determine the aggregate genetic merit for straightness did not appear to depend on the correlations between the aggregate breeding values and those at individual sites. The weights were 0.4:0.39:0.16:0.35 for sites A-D, respectively. When breeding values at each of the sites were weighted according to importance of the site, the correlations between the weighted aggregate breeding values and those at individual sites were 0.75, 0.46, 0.95 and 0.24 for sites A-D, respectively.

The ranking of the parents based on the weighted aggregate genetic merit reveal that parents with high breeding values for height generally had high breeding value for straightness, indicating that selection for height will also improve straightness.

Table 7.7. Estimates of breeding values for height at 22.5 years at four sites, and the aggregate breeding values over the four sites without weighting for site representativeness (H), and with weighting (H^*). Ranks of parents based on the weighted aggregate breeding values are shown.

Parent	Site				H	H^*	Rank
Number	A	B	C	D			by H^*
8	2.08	1.14	1.28	1.6	1.58	1.6	1
9	-0.145	0.253	1.01	-0.707	0.257	0.349	5
10	0.68	0.316	-0.122	-0.007	0.225	0.27	7
11	0.338	-0.569	0.252	-0.41	-0.226	0.13	10
12	1.2	0.02	0.282	-0.811	0.1	0.554	4
13	-0.172	-1.04	0.099	0.686	-0.056	-0.151	14
14	-1.15	-0.795	-0.947	-0.62	-0.945	-0.988	22
15	1.37	0.847	0.429	0.504	0.988	0.873	3
16	0.517	0.663	-0.293	0.881	0.622	0.233	8
17	0.419	0.502	-0.266	-0.029	0.277	0.135	9
20	0.747	0.155	0.04	-0.549	0.187	0.309	6
23	-0.925	-0.696	-0.094	-0.005	-0.718	-0.512	17
29	-0.442	1.48	-0.282	0.644	0.729	-0.035	13
30	1.37	0.694	0.941	0.122	0.799	1.04	2
31	-1.41	-1.04	0.427	-0.138	-0.818	-0.557	18
32	-1.97	-0.34	0.506	0.376	-0.487	-0.619	19
33	-0.467	-1.13	-0.735	-1.457	-1.1	-0.723	20
34	0.08	-0.593	-0.375	0.331	-0.141	-0.19	15
36	-0.974	0.89	-0.486	1.2	0.498	-0.39	16
37	-0.164	-0.889	-1.651	-1.15	-1.07	-0.917	21
39	0.001	0.07	-0.031	0.02	0.08	0	12
40	0.18	0.06	0.07	-0.093	0.102	0.104	11

Table 7.8. Estimates of breeding values for stem straightness at 22.5 years at four sites, and the aggregate breeding values over the four sites without weighting for site representativeness (H), and with weighting (H^*). Ranks of parents based on the weighted aggregate breeding values are shown.

Parent	Site				H	H^*	Rank
Number	A	B	C	D			by H^*
8	0.178	-0.076	0.19	-0.533	-0.111	0.109	8
9	0.165	0.07	-0.055	0.202	0.1	0.06	9
10	-0.026	0.344	-0.583	0.133	0.1	-0.185	18
11	-0.055	-0.614	0.166	-0.652	-0.425	-0.08	14
12	0.07	-0.053	0.08	0.168	0.08	0.06	10
13	-0.062	0.013	0.412	0.295	0.241	0.157	7
14	-0.288	-0.216	-0.873	0.305	-0.234	-0.482	22
15	0.191	0.179	-0.385	-0.035	0.1	-0.052	13
16	0.296	0.366	0.807	0.268	0.425	0.509	2
17	0.239	0.207	0.806	0.679	0.547	0.483	3
20	0.285	0.028	1.17	-0.566	0.141	0.559	1
23	-0.209	-0.035	-0.215	-0.019	-0.194	-0.176	15
29	-0.003	0.154	0.416	0.324	0.218	0.205	5
30	0.13	0.167	0.789	0.08	0.28	0.397	4
31	-0.231	-0.255	0.256	-0.349	-0.248	-0.045	12
32	-0.283	0.271	-0.271	-0.011	-0.053	-0.182	16
33	0.106	0.028	0.246	0.301	0.176	0.16	6
34	0.02	-0.035	-0.525	-0.175	-0.136	-0.215	19
36	-0.156	0.375	0.013	0.607	0.3	0.03	11
37	-0.121	-0.377	-0.133	-0.475	-0.554	-0.182	17
39	0.047	-0.216	-0.623	-0.256	-0.217	-0.276	20
40	-0.147	-0.322	-0.502	-0.31	-0.305	-0.323	21

7.3.4 Efficiency of selection

Traits assessed at the same age across sites

Efficiencies of selection on height at one site for planting at a different site are given in Table 7.9. These ranged from 18 to 123%, and were high at 1.5 and 9.5 years of age and moderate at 13.5 and 22.5 years. The lowest efficiency of selection was for height between sites A and D at 22.5 years (18%). At 1.5 years the efficiency of selection at other sites for planting at D was more than 100%, indicating that it was better to select at these sites to plant at site D than at site D itself. This was attributed to a much lower heritability estimate obtained at site D than at the other sites. This low heritability at site D also caused the efficiencies for selecting at D for planting at the other sites to be low. The results at 1.5 and 9.5 years indicate that selection can be carried out with little loss in gain at any site, other than site D, for planting at the other sites. On the contrary, the results at 13.5 and 22.5 years indicate that for sites that are close geographically, such as A and C, losses in genetic gain from selection at the other sites are high - sometimes even higher - than selections at more distant sites.

Efficiency of selecting for stem straightness at one site for planting at another site are shown in Table 7.10. These range from high negative (-164%) to high positive (122%). At 1.5 years, selecting at site C was better than direct selection on site A and B, and selection on A was better than direct selection on site B. When selections were made on site D for planting at the other sites the efficiencies of selection were highly negative, and were also highly negative when selections were made at the other sites for planting at site D. At older ages, efficiencies of selection for straightness were high between site A and C, and between B and D, indicating similarity between the two groups of sites.

Table 7.9. Estimates of efficiencies (%) of indirectly selecting for height at one site compared with directly selecting at another site, across four sites at four ages.

Age		Site			
		A	B	C	D
1.5	A		76	77	113
	B	106		91	116
	C	117	99		123
	D	69	51	49	
9.5	A		119	100	89
	B	67		78	96
	C	81	113		89
	D	49	94	60	
13.5	A		55	55	63
	B	74		53	90
	C	113	71		105
	D	61	65	49	
22.5	A		114	53	18
	B	71		47	59
	C	49	71		51
	D	18	91	53	

Table 7.10. Estimates of efficiencies (%) of indirectly selecting for stem straightness at one site compared with directly selecting at another site, across four sites at four ages.

Age		Site			
		A	B	C	D
1.5	A		122	48	-81
	B	61		33	-67
	C	190	266		-164
	D	-81	-134	-41	
9.5	A		37	114	-13
	B	33		48	96
	C	73	35		13
	D	-9	79	15	
13.5	A		-24	68	-20
	B	-26		4	82
	C	114	6		4
	D	-29	111	4	
22.5	A		46	37	4
	B	95		27	96
	C	191	71		36
	D	8	94	14	

Traits assessed at different ages across sites

The ratios of efficiency of early selection within and across sites for height are shown in Table 7.11. The results show that, if selections are made at 1.5 years of age, it is better to select at sites other than those where the trees are going to grow to maturity, other than at site C where it was better to select and plant at the same site. The efficiencies were particularly high when selections were made at site B or C for planting at A. This was attributed to much higher genetic correlations between height at age 1.5 years at sites B and C, and that at 22.5 years at site A (0.86, 0.95, respectively).

compared to height at age 1.5 at site A and height at 22.5 at site A which was suspiciously low (0.10). The results show that if selection are to be carried out at 1.5 years of age, the best site to locate the genetic tests is site C since the gain from indirect selection at this site will be higher than that obtained from direct selection on the other sites. If selections are made at 9.5 years of age, the results suggest that indirect selection on site D resulted in a large loss in genetic gain compared to selecting at 9.5 years at each of the other three sites. Therefore, site D should be avoided for establishing genetic tests to make selections for planting at the other sites. However, selecting at 9.5 years of age on any of the other sites (A, B, C) did not reduce genetic gain much compared to selecting on the same site since the efficiencies of selection were all greater than 70%.

The ratios of efficiency of early selection within and across sites for straightness are shown in Table 7.12. At 1.5 years of age, the best sites for selecting for straightness were A and C. At 9.5 years, the loss of potential gain from selecting at site A for planting at C was low, as was that between selecting at site B and planting at D.

Table 7.11. Estimates of efficiencies (%) of early selection across sites for height compared to within site.

Age		Site			
		22.5 years			
		A	B	C	D
1.5 years	A		194	65	126
	B	755		69	108
	C	876	267		119
	D	109	160	33	
9.5 years	A		82	75	86
	B	80		72	103
	C	84	86		78
	D	37	62	49	

Table 7.12. Estimates of efficiencies (%) of early selection across sites for stem straightness compared to within site.

Age		Site			
		22.5 years			
		A	B	C	D
1.5 years	A		154	91	1925
	B	35		-228	-318
	C	87	643		800
	D	-40	-887	132	
9.5 years	A		16	75	-16
	B	65		3	93
	C	80	50		23
	D	7	82	4	

Efficiency of selection at each site to improve average performance across all sites

The efficiencies of selecting for height at each site to improve the average performance across all four sites with and without weighting according to importance of sites, are shown in Table 7.13. When sites were not weighted, all sites had high efficiencies of selection for height; site B had the highest efficiency. Although some of the correlations of breeding values at each site and the aggregate ones were low, the individual site heritability estimates were all higher than the one estimated over all four sites (0.24), resulting in high efficiencies of selection. When sites were weighted, sites A and C had high efficiencies with site A having the highest.

The efficiencies of selecting for stem straightness at each site to improve the average performance across all four sites with and without weighting according to importance of sites, are shown in Table 7.14. When sites were not weighted, efficiencies of selection for stem straightness were high for all sites, except site A which had the lowest heritability estimate and lowest correlation. When sites were weighted, site C had the highest efficiency.

Table 7.13. Estimates of correlations of breeding values for height at each site without (r_A) and with weighting (r_A*) and efficiencies of selecting for height at each site for improving the average without (E) and with weighting (E*).

Site	h^2	r_A	r_A*	E	E*
A	0.40	0.70	0.88	0.90	1.14
B	0.26	0.90	0.69	0.98	0.72
C	0.39	0.59	0.74	0.75	0.94
D	0.40	0.71	0.43	0.92	0.56

Table 7.14. Estimates of correlations of breeding values for straightness at each site without (r_A) and with weighting (r_A*) and efficiencies of selecting for straightness at each site for improving the average without (E) and with weighting (E*).

Site	h^2	r_A	r_A*	E	E*
A	0.16	0.57	0.75	0.50	0.65
B	0.33	0.83	0.43	1.04	0.54
C	0.85	0.58	0.95	1.16	1.91
D	0.32	0.74	0.24	0.91	0.30

7.4 Discussion

In identifying the presence of GE it is important to ascertain whether GE is due to heterogeneity of variances across sites or to rank changes of the genotypes. In this study, the additive variance and heritability estimates varied slightly across sites (see Chapter 3), and genetic correlations across sites for height at 22.5 years were low, indicating rank changes of the genotypes. Therefore, GE in this study was due primarily to rank changes of genotypes across sites.

For height, GE was more pronounced as the trees aged. The genetic correlations and efficiencies of selection were high at 1.5 and 9.5 years, indicating that if selections are made at one site there will be little loss in potential genetic progress at the other site. However, when selections were made at later ages, both the genetic correlations and

efficiencies of selection were low to moderate, indicating that there were significant rank changes between sites. Therefore, maximum gain for a specific site would require that selection be conducted in tests planted at that site. The very low efficiency of selection estimates involving site D, indicate that this site is substantially different from the other sites.

Because of the long rotation periods of forest tree species, results and conclusions regarding GE are usually drawn from tests evaluated at a young age. For example, Pswarayi *et al.* (in press) found across site genetic correlations in *P. elliottii* assessed up to 15 years in Zimbabwe to be higher than 0.7, indicating little GE. Similarly, studies of GE in *P. radiata,* based on tests assessed up to 12 years, found GE not to be of practical importance (Johnson 1992, Matheson and Raymond 1984b, Johnson and Burdon 1990). These conclusions are supported by our results up to 9.5 years, and to a lesser extent by those at 13.5 years. In contrast, results at 22.5 years indicate that, while GE in this population was small at young ages, it was large at mature ages. Therefore, early growth assessments may not be reliable for assessing GE at maturity. This implies a critical need to evaluate GE at mature ages in forest trees. It is therefore important to verify, for the other species where inferences about GE is based on early growth assessments, the trend in GE over time to maturity.

The results describing the effect of early selection for height at one site for predicting mature performance at another site, compared to early selection for mature performance on the same site, shows that if selections are made at 9.5 years of age, selecting at site D will result in large losses in gain at maturity at the other three sites. This result also suggests that site D is a different site to the other sites, and should be avoided for establishing tests from which selections will be made for planting at all sites. On the other hand, selections made early at the other sites would do well at site D.

Genetic correlations and efficiencies of selection for stem straightness at the same age across sites were high between sites A and C, and between sites B and D. The similarity between sites A and C, and sites B and D, was also evident when early selection across sites was made at 9.5 years. Therefore, if straightness is the main selection criterion, sites A and C could form one breeding population and sites B and D another. However, the interpretation of these results is complicated by the fact that sites

A and C were assessed by one team, and sites B and D by another. Although the assessment scale used is meant to be invariant across sites, the possibility remains that differences of interpretation by the two assessment teams could confound these results.

The heritability estimates reported in Chapter 3 for both traits were lower when they were estimated from the data pooled across sites than from individual site data, and the GE found in this study partly explains these results. Since pooling data across site lowers the heritability in the presence of GE, it may reduce the potential genetic gain compared to selection at individual sites. As pointed out by Pedrick (1990), even within a region, one should expect GE. Therefore, results of a single test (normally less than 5 hectares) are likely to overestimate gain if the results are to used for a region (normally more than 2000 hectares), since it ignores possible GE. Given this hypothesis, the most appropriate heritability estimate to use for gain prediction may actually be the one from data pooled across sites. Furthermore, breeding values estimated from a single site are not as precise as those from the pooled data. In order to yield results that are appropriate to commercial progress, it is therefore necessary to consider dispersing progeny tests within the region where most of the commercial plantations are situated.

The overriding importance of altitude and rainfall to performance at the tree species level, which has been used to match species to environments in Zimbabwe, may not be true at a population within species level. Sites A and C were low altitude sites and were less than 5 kilometres apart, while sites B and D were high altitude sites, and more than 100 kilometres from both sites A and C; nevertheless, the genetic correlations for height at 22.5 years between sites A and B was 0.90 while that between sites A and C was 0.51. This result complicates deployment and breeding strategy, as no environmental factor which influenced GE is easily apparent.

Implications for breeding P. taeda *in Zimbabwe*

In Zimbabwe, height is likely to be a more important selection criterion than straightness. Furthermore, there is no significant planting of *P. taeda* at site D, which appears to differ substantially from sites A and C, to warrant a separate breeding population. Also, we could not identify any single environmental factor which

influenced GE, making it difficult to classify environments. Therefore, one breeding population is recommended. Given this recommendation, and that the overall objective is to increase timber production in all regions where the species is grown commercially, rather than to increase production at a particular site, use of the correlation between individual site breeding values and the aggregate breeding values, and efficiency of selection on each site to improve the average, are more appropriate criteria in making decisions regarding the choice of location of progeny test. Results indicate that the selections to improve the average performance across all the four sites should be made at site A, which had the highest correlation between the breeding values at the site and the weighted aggregate genetic merit, and the highest efficiency of selection. Furthermore, efficiencies of across-site early selection at 9.5 years for height at site A were greater than 75%. This strategy will not significantly affect selections for straightness, since selections made at 9.5 years at site A predicted fairly well mature performance of straightness at site C, which is an important region for *P. taeda* plantations. Since results of a single test are likely to overestimate gain across a region, as it ignores possible GE, the tests should be replicated within the region represented by site A. Nursery period in *P. taeda* varies with altitude, being shorter at low altitudes, and therefore locating the nursery and genetic tests at site A, which is the lowest in elevation, should result in savings both in nursery and transport costs.

Since GE may be due to a few families, the most interactive parents could be excised from the breeding programme, as proposed by Matheson and Raymond (1984b). However, this should be done with care since there may be a danger of removing some parents of superior average performance over all sites.

The multiple population breeding strategy is being used in *P. taeda* in Zimbabwe, due in part to the possibility of utilizing GE. Our study failed to identify major environmental factors which may have caused GE, and GE was found to be larger for sites located close to each other than those far apart, limiting our ability to define and classify environments with minimum GE. Our results, therefore imply that efficiency of the multiple population breeding strategy for *P. taeda* in Zimbabwe, in terms of utilizing GE, might be low.

However, the precision of many genetic correlations estimated here were low,

and hence the precision of the efficiencies of selection derived from them would be expected to be low. Therefore, the findings reported here should be verified with better data sets.

Chapter 8

DECISION-MAKING IN PROGENY TEST LOCATION USING GIBBS SAMPLING

8.1 Introduction

Forest sites are heterogenous due to variations in soils, weather conditions, or other factors. The manner in which progeny tests are deployed affects genetic gain in production populations in the presence of genotype x environment interactions (GE). The question faced by breeders is where should one locate progeny tests so as to ultimately maximise gain in production populations. The conventional method of determining the best site to locate progeny tests is to estimate the efficiency of selecting at one site for planting at another site, using the method of indirect selection described by Falconer (1989). In order to estimate the efficiency of selection, heritability and genetic correlations are estimated using standard methods such as REML. A major limitation of this approach is that the distribution or the variance of the efficiencies of selection are unknown, adversely influencing the efficiency of decision-making. A further problem is that genetic correlations are difficult to estimate and can range from -1 to 1 regardless of the true parameter if sample sizes are low. For example, REML analyses in Chapter 7 indicated that the genetic correlation between heights at 22.5 years at sites A and C was low and had a large standard error (0.51± 0.25; Table 7.1), as was that between heights at 9.5 years at site A and those at 22.5 years at site C (0.55 ± 0.21; Table 7.3). The large standard errors associated with the genetic correlation estimates particularly in tree breeding due to few parents, and hence possible large sampling variance associated with the derived selection efficiencies, make it difficult to have confidence in making decisions on whether or not progeny tests should be located at one site.

An alternative approach which might overcome this lack of confidence is to use

a Bayesian approach such as Gibbs sampling. This approach, where random samples from joint distributions are generated, may assist in decision-making regarding location of progeny tests in heterogenous sites in the presence of GE. When calculating the efficiencies of selection, the genetic correlations and heritability estimates from REML are assumed to be known without error. Gibbs sampling is particularly attractive in this context, because the sampling distribution of the efficiency of selection can be obtained. The confidence of choosing a location for planting progeny tests is hard to assess and might be lower than anticipated, if these sampling distributions are not considered. Furthermore, Gibbs sampling enables one to estimate the probability that the efficiency of selection lies between certain values, thereby producing considerably more information on which to base decisions compared to the point estimates from REML. Furthermore, the approach makes use of prior information, and nuisance parameters are integrated out.

The Gibbs sampler has recently been used for estimating variance components in animal breeding applications (Jensen *et al.* 1994, Sorensen *et al.* 1994, Wang *et al.* 1994). In addition, Sorensen *et al.* (1994) used the Gibbs sampler to estimate uncertainty in response to selection. It appears that Gibbs sampling has not been used in tree breeding, but has been used in forest inventory (Green and Strawderman 1996, Green *et al.* 1994).

The present study will investigate use of the Gibbs sampling in decision-making on choice of site for locating progeny tests. The results will be compared to those obtained by REML.

8.2 Materials and Methods

Data

Data for the Gibbs sampler were heights assessed at ages 9.5 years and 22.5 years at sites A and C. These two ages were selected because early selection at 10 years was predicted to be effective in *P. taeda* (Chapter 6). The two sites were selected because both sites represented the region where most of the commercial plantations of *P. taeda* are located, and therefore making decisions regarding the location of progeny

tests between these two sites are critical. Furthermore, the outcome of alternative strategies involving progeny test location using REML estimates was not clear due to the large standard errors of the genetic correlations (see Chapter 7). Details of the tests and assessment procedures are described in Chapter 3.

Overview of Bayes theorem

The objective of the Bayes methods is to compute the posterior distribution of the parameter of interest. To start with, a prior distribution, which represents the belief about the parameter before any data are observed, is assumed. The posterior distribution then represents the updated belief after viewing the data. The posterior distribution is expressed as proportional to the prior distribution of the parameter times the conditional distribution of the data given the parameter (likelihood) (Gilks *et al.* 1996):

$$P(\theta|\mathbf{y}) \propto P(\theta)P(\mathbf{y}|\theta)$$

where $\mathbf{y}$ is the data and θ is the parameter.

Gibbs sampling

Gibbs sampling is a method of numerical integration that allows inferences to be made about joint or marginal distributions of the parameters of interest. The Gibbs sampling algorithm is an updating sampling scheme which requires random independent draws of variables from all of the full conditional distributions. The full conditional distribution is the distribution of a variable given all other parameters in the model. Gibbs sampling integrates out the other parameters leaving the distribution of the parameter in question, conditional on the data (i.e. marginal posterior distribution of the parameter). After obtaining samples from the marginal distributions, means and variances of the distribution can be estimated.

The following quadrivariate tree model was used to estimate covariance components for height across the two sites and the two ages (hence four traits):

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & X_3 & 0 \\ 0 & 0 & 0 & X_4 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 & 0 \\ 0 & Z_2 & 0 & 0 \\ 0 & 0 & Z_3 & 0 \\ 0 & 0 & 0 & Z_4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

where: $\mathbf{y_i}$ = the vector of observations for trait i,

$\mathbf{b_i}$ = the vector of fixed effects for trait i,

$\mathbf{a_i}$ = the vector of random tree (additive genetic) effects for trait i,

$\mathbf{X_i}$ = the incidence matrix for fixed effects for trait i,

$\mathbf{Z_i}$ = the incidence matrix for additive direct effects for trait i,

$\mathbf{e_i}$ = the vector of residual effects for trait i. For i = 1, 2, 3, 4.

The effect of the replicate was considered fixed. The assessments of height at the two sites and two ages were treated as different traits. At 9.5 years the trees assessed were those that were removed after thinning. Therefore, for all the analyses, the traits were assessed on separate but genetically related trees across and within sites. Therefore, genetic covariances exist between any two traits, but there were no environmental covariances.

The conditional distribution of the complete data given the location and scale parameters is assumed to be quadrivariate normally distributed.

Priors were assigned to unknown parameters (the variance components and fixed effects) in the model. Uniform (flat) prior distributions were assumed for fixed effects. As priors for the additive random genetic and residual covariance matrices two types of prior distributions were assumed: inverted Wishart and uniform distributions. The Wishart (IW) distribution is a matrix generalization of the univariate Chi-squared distribution (Sorensen 1997). For a proper prior (IW) distribution , the shape parameter (degree of freedom) should be 2 more than the order of the matrix (Van Tassell and Van Vleck 1995). A shape parameter 9 was used, indicating moderate belief for the prior distributions of the variance components. The additive genetic and residual (co)variance components estimated using REML were used as the starting values. The rationale for

this was that it was difficult to find reports of additive and residual covariances from literature. Sensitivity analyses were conducted to find the effect of changing the starting values on the expected values of the posterior distributions-the REML covariance estimates were multiplied by 2. To use flat priors, a value 0 was used for the shape parameter, indicating no prior knowledge for the distributions of the variance components.

The analysis of the four traits simultaneously was done using the Multi Trait Gibbs Sampling for Animal Models program (MTGSAM, Van Tassell and Van Vleck 1995). The first 5000 iterations were not stored to ensure that samples saved were from the proper stationary posterior distributions. Thereafter, a total of 100,000 iterations were made, and samples stored every 100th iteration to make sure that the samples were nearly uncorrelated, giving a total of 1000 samples of additive genetic and residual covariance estimates stored. From these, heritability, genetic correlations and efficiencies of selection were calculated for each sample, and inferences about efficiencies of selection were made by computing directly summary statistics from the resulting distributions derived from the 1000 samples. The mean of the posterior distribution was taken as the mean of all the samples and the mode of the posterior distribution as the most frequent value of the parameter. The most frequent value was determined from the histogram. The probability that the efficiency of selection was greater than 0.7 and 1.0 were estimated. A probability of 0.7 was selected because efficiencies lower than 0.7 would justify extra costs of having separate progeny tests.

In order to check for convergence the Gibbs sampler was run several times with different starting values and different intervals of saving samples to make sure that the same estimates were obtained each time. Similar estimates were obtained, hence convergence was assumed.

To test independence of samples autocorrelations were estimated. The samples were moderately correlated. For example, the 1st order lag-correlations ranged between 0.45 and 0.74, when every 100th sample was saved, and this was reduced to 0.26-0.60, when every 200th sample was saved. The expected values of the posterior distributions for the efficiencies of selection were similar indicating that autocorrelations were not a major problem. This may be due to the long Gibbs chain used in this study.

To test if the covariance estimates from Gibbs sampling were significantly different from the REML ones, a log ratio test was performed in which the Gibbs sampling estimates were used as initial values in REML and fixed. The difference between loglikelihoods with and without fixing the covariance components were compared with Chi-square distribution with degrees of freedom equal to the number of fixed parameters (i.e. 14).

8.3 Results

Genetic parameters

Using a log ratio test, covariance estimates from Gibbs sampling and those from REML were not significantly different.

Table 8.1. Estimated heritabilities (in bold), and genetic correlations for height based on Gibbs sampling, and standard deviations of the marginal posterior distributions are in parenthesis.

	HT9.5 (site A)	HT9.5 (site C)	HT22.5 (site A)	HT22.5 (site C)
HT9.5 (site A)	**0.81**			
	(0.09)			
HT9.5 (site C)	0.90	**0.66**		
	(0.04)	(0.10)		
HT22.5 (site A)	0.74	0.72	**0.55**	
	(0.12)	(0.13)	(0.13)	
HT22.5 (site C)	0.52	0.80	0.47	**0.56**
	(0.17)	(0.09)	(0.21)	(0.13)

The estimated heritabilities and genetic correlations using Gibbs sampling and REML are shown in Tables 8.1 and 8.2, respectively. The heritability estimates from Gibbs sampling (0.55-0.81) were slightly larger compared to those from REML (0.39-0.73), but the genetic correlations were similar. There is expected positive error covariances between REML and Gibbs estimates, making significance testing difficult.

However, the differences were small, and therefore unlikely to be important.

Selection efficiencies

The estimated selection efficiencies are shown in Table 8.3 and the histograms of their marginal posterior distributions are given in Figures 8.1 and 8.2. While the estimates of efficiencies of early selection across site showed little variation, those for selection at maturity across sites did. The distributions of the selection efficiencies were slightly skewed. The posterior modes were similar to estimates from REML (Table 8.3). This is expected since the maximum likelihood estimate of a parameter is the value that maximises the likelihood (i.e. mode). The probability that the efficiency of early selection at site C for planting at site A was greater than 0.70 was 0.93 (Table 8.3), indicating that early selection at site C would result in little loss in gain at site A at harvest age, compared to early selection at site A. In fact, the probability that more gain would be obtained from early selection at site C compared to site A is 0.2. The high efficiency of selection at site C was attributed to a high genetic correlation between heights at 9.5 years at site A and those at site C at 22.5 years, which was as high as that between 9.5 and 22.5-year heights at site C (Table 8.1).

Due to much lower across-site genetic correlations than within-site correlations, the probability that the efficiency of early selection at site A for planting at site C was greater than 0.70 was only 0.57, and the probability that early selections at site A would result in higher gain at site C at harvest age than early selections at site C was only 0.01. The results suggest that site C is a better progeny test site since selections made here will result in little loss in gain at site A, and may even result in higher gain at site A at harvest age than early selection at site A, whereas early selection at site A would severely reduce progress at site C.

If selection were carried out at 22.5 years of age (harvest age), the probabilities that the efficiencies of selection at alternative sites are greater than 0.7 were all very low: 0.12 for selection at site A for planting at site C, and 0.17 for selection at site C for planting at site A, indicating that these sites were different. Therefore, if selections are to be made at maturity, separate progeny tests should be established for the sites since selections at alternative sites would result in substantial losses in gain at the sites. The

low efficiencies of selection are attributed to low genetic correlation between heights assessed at these two sites at 22.5 years.

Table 8.2. Estimated heritabilities (in bold), and genetic correlations for height based on REML, and their standard errors are in parenthesis.

	HT9.5 (site A)	HT9.5 (site C)	HT22.5 (site A)	HT22.5 (site C)
HT9.5 (site A)	**0.73**			
	(0.14)			
HT9.5 (site C)	0.90	**0.59**		
	(0.05)	(0.14)		
HT22.5(site A)	0.77	0.72	**0.40**	
	(0.08)	(0.15)	(0.14)	
HT22.5 (site C)	0.55	0.82	0.51	**0.39**
	(0.21)	(0.07)	(0.25)	(0.15)

Table 8.3. REML and Gibbs sampling estimates of efficiencies of selection for height, standard deviations (SD) of the marginal posterior distributions, and the probabilities that the selection efficiencies (E) are greater than 0.7 and 1.0.

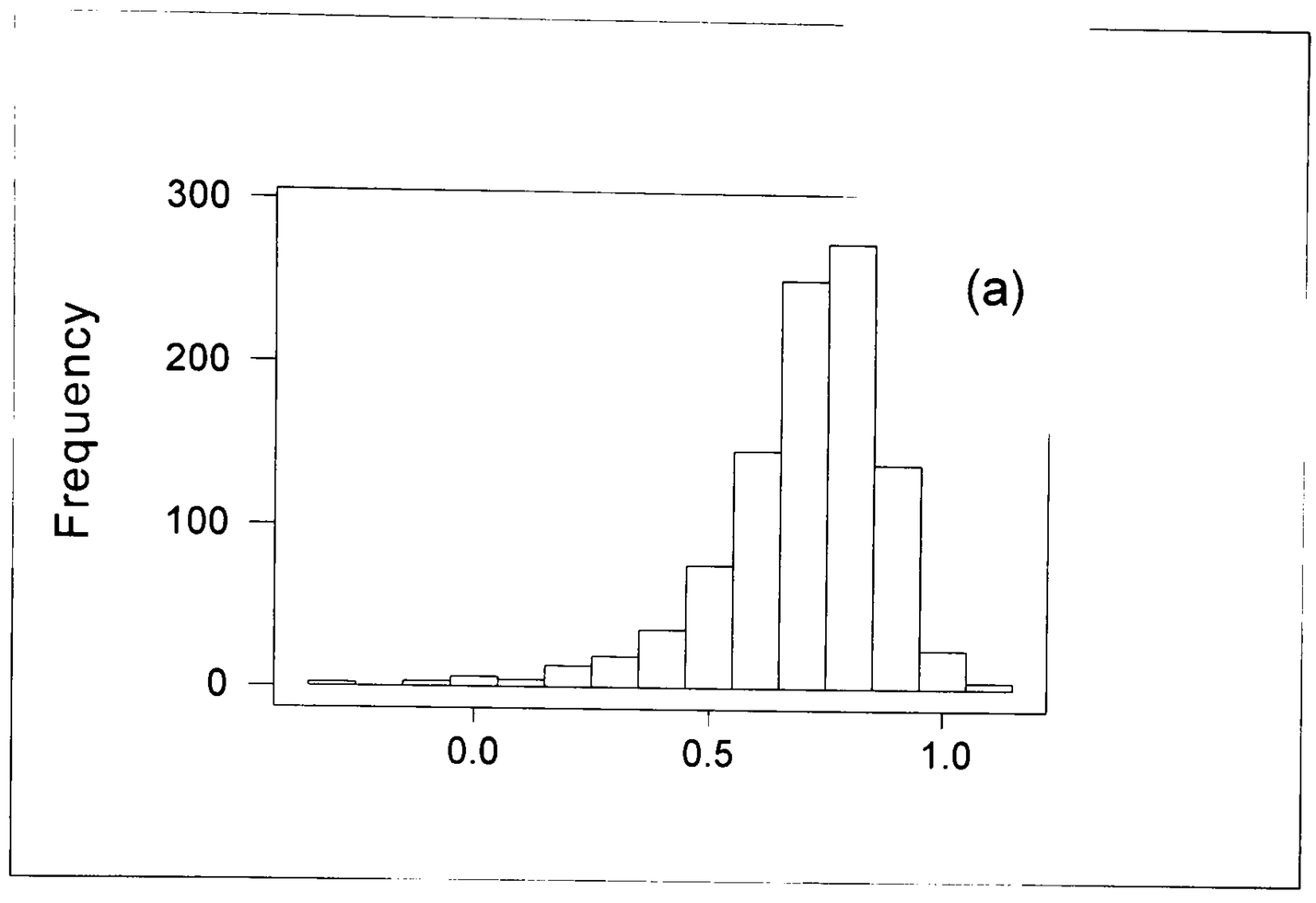
	REML estimate	Posterior mode	Posterior mean	SD	P(E>0.7)	P(E>1.0)
E_{A2c4}*	0.84	0.80	0.70	0.18	0.57	0.01
E_{c2A4}	0.75	0.75	0.89	0.15	0.93	0.20
E_{A4C4}	0.49	0.50	0.46	0.22	0.12	0.00
E_{C4A4}	0.53	0.63	0.48	0.23	0.17	0.01

*E_{A2C4} = efficiency of selecting at site A at 9.5 years for planting at site C compared with early selection at site C at 9.5 years.

E_{c2A4} = efficiency of selecting at site C at 9.5 years for planting at site A compared with early selection at site A at 9.5 years.

E_{A4C4} = efficiency of selecting at site A at 22.5 years for planting at site C compared to direct selection at site C.

E_{C4A4} = efficiency of selecting at site C at 22.5 years for planting at site A compared to direct selection at site A.



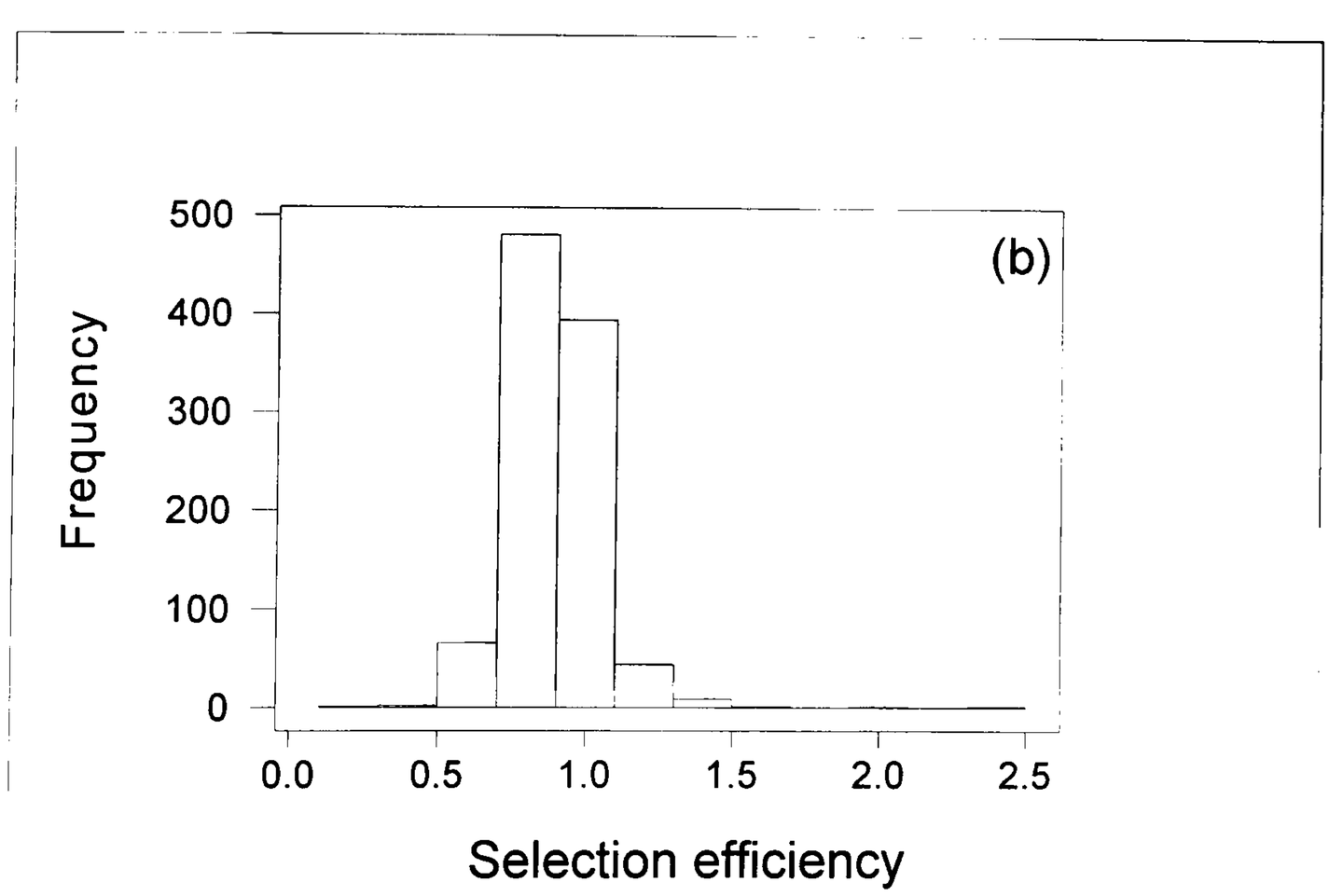
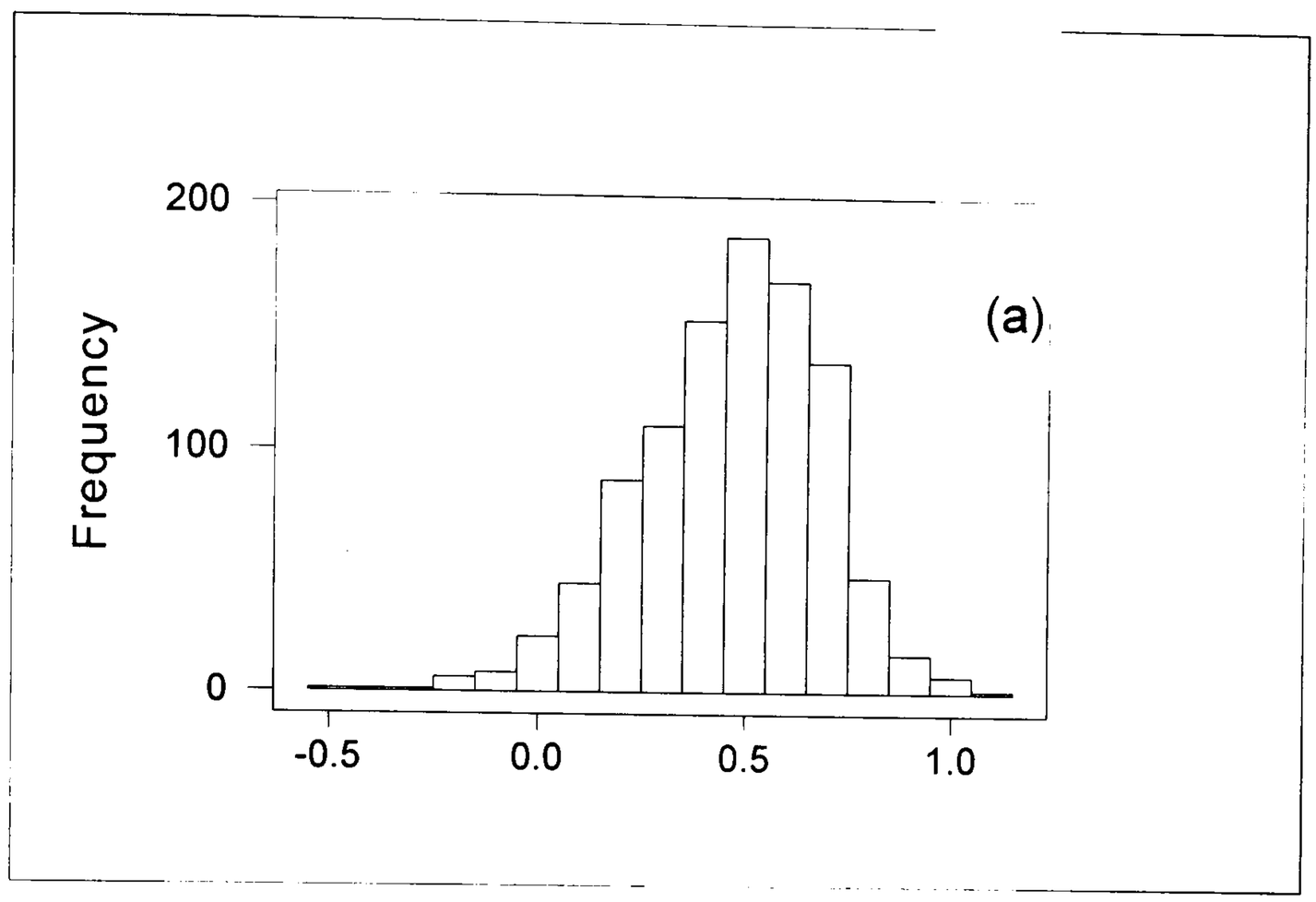


Figure 8.1. Histograms from marginal posterior samples of size 1000 for efficiency of early selection across sites: (a) efficiency of early selection at site A for planting at site C, (b) efficiency of early selection at site C for planting at site A.



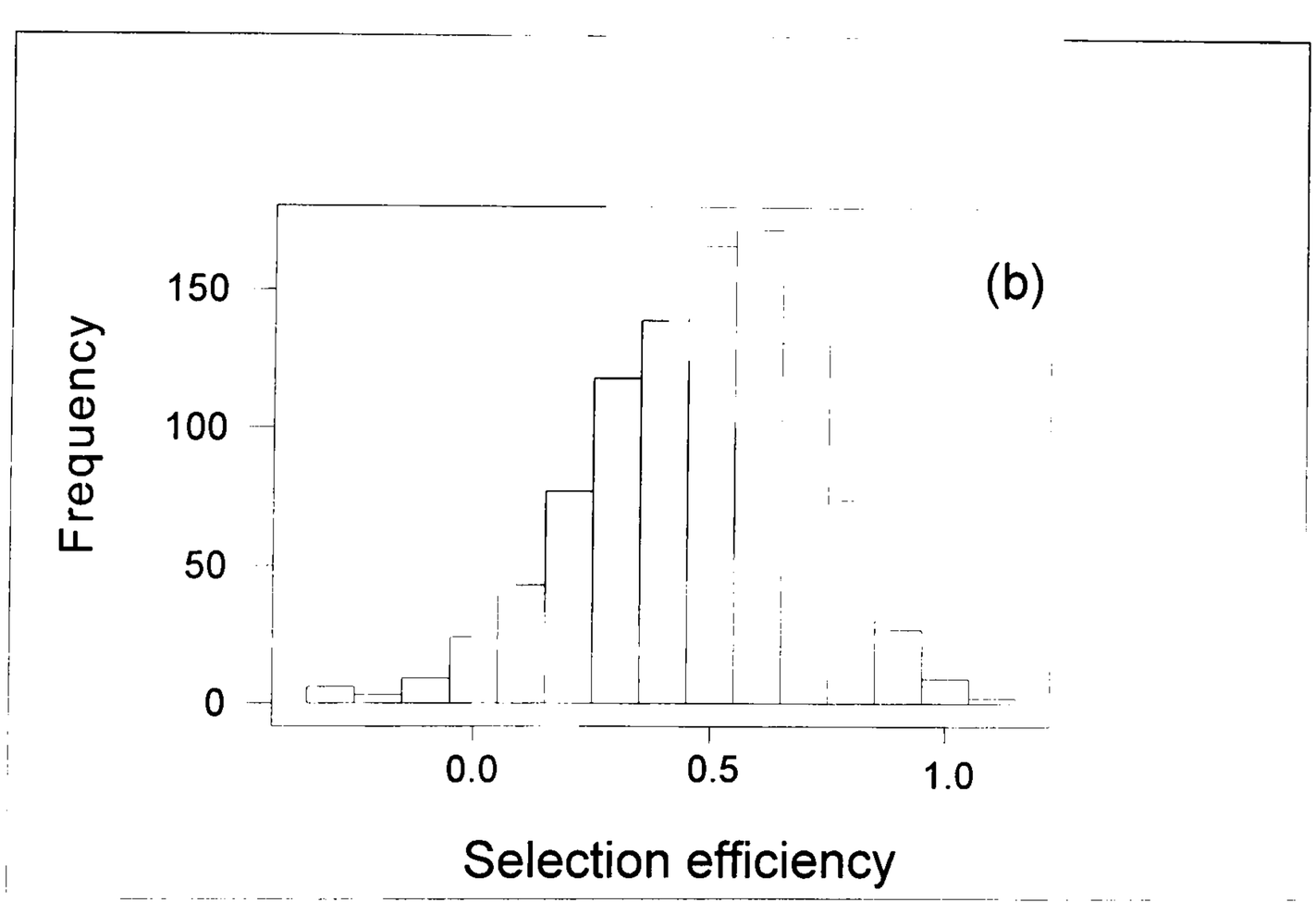


Figure 8.2. Histograms from marginal posterior samples of size 1000 for efficiency of selection at maturity across sites: (a) efficiency of selection at site A for planting at site C, (b) efficiency of selection at site C for planting at site A.

Influence of priors

The results showing the influence of priors are in Table 8.4. When inverted Wishart distributions were assumed and the starting values multiplied by 2, the estimates of the efficiencies of selection and the variance of the posterior distributions were similar. This indicates that the starting values had little effect on the marginal posterior distributions of the efficiencies of selection. However, the results assuming uniform and inverted Wishart distributions as priors differed, particularly in the variances of the posterior distributions which were much higher when the former were assumed as priors (Table 8.4).

Table 8.4. Expected values of the marginal posterior distributions of the efficiencies of selection using inverted Wishart (IW) and uniform distributions (UNI) as priors. Standard deviation of the marginal posterior distributions are in parenthesis.

	IW (starting values from REML)	IW (starting values from REML x2)	UNI
E_{A2c4}	0.70	0.70	0.46
	(0.18)	(0.23)	(6.12)
E_{c2A4}	0.89	0.79	0.92
	(0.15)	(0.16)	(2.64)
E_{A4C4}	0.46	0.40	0.34
	(0.22)	(0.21)	(0.36)
E_{C4A4}	0.48	0.43	0.34
	(0.23)	(0.23)	(0.35)

8.4 Discussion

Use of Gibbs sampling in decision making in progeny test location was demonstrated. Variance components were estimated using MTGSAM. and the estimated components were used to derive heritability estimates and genetic correlations. which

were in turn used to estimate the efficiencies of selection. The study illustrated that the point estimates of the efficiencies of selection were subject to substantial error, particularly those involving selections at maturity. Gibbs sampling provided a method for constructing the posterior distribution of the efficiencies of selection from which the variation of the estimates were obtained, and the probabilities that the estimates were within a specified range were also estimated. A further advantage of the Gibbs sampling approach was that it allowed simultaneously estimation of components of variances and covariances for the four traits, unlike the REML methodology used in Chapter 7. However, heritability and genetic correlation estimates from the bivariate REML analyses and those from multivariate Gibbs sampling analyses were not significantly different, suggesting that bias due to selection was small.

Using this approach, site C emerged as a better site to locate progeny tests than site A if early selection is practised. However, if selections are made at maturity, which is highly unlikely, separate progeny tests should be established for the two sites. The decision regarding selection at maturity is consistent with that obtained using point estimates from REML, but Gibbs sampling allowed the efficiencies of selection to be interpreted with more confidence. The decision regarding early selection differed from that based on REML point estimates. Using REML, the efficiencies of early selection at both site A and site C were greater than 0.7 indicating that either of the two sites could be a suitable location for progeny tests; in contrast, with Gibbs sampling, it was clear that site C was a better site to locate progeny tests. Furthermore, if a choice had to be made between the two sites, site A would be selected using point estimates from REML, resulting in different decisions arising from the two approaches. The difference between the results from Gibbs sampling and REML is attributable to greater information derived using the former method. The study demonstrates the advantage of having some measure of variability associated with the estimates of efficiency of selection.

Influence of priors was studied. Changing the starting values had no impact on the posterior distributions of the efficiencies of selection. Assuming prior distributions were uniform gave different results than assuming priors were inverted Wishart distributions. Sorensen *et al.* (1994) also found that variances were higher when uniform

priors were used. The lack of agreement might be due to improper distributions when uniform distributions are assumed as priors or a weak likelihood due to poor information in the data. The use of uniform priors for variance components was discouraged because the chance of obtaining an improper posterior distribution was high (Van Tassell and Van Vleck 1995).

The results indicate that, even in this simple decision problem, Gibbs sampling can be an attractive approach to decision-making in progeny test location as more information to make inferences about the parameter of interest can be derived from the analyses than is possible from REML. The benefits might be expected to be even greater in more complex decision processes. The advantage of the approach is that it gives a full marginal posterior distribution of a parameter of interest, from which the probability that the parameter lies between specified values can easily be computed (Sorenson *et al.* 1994). Further, marginal posterior distributions for a parameter of interest consider all other parameters as nuisances, and integrates them out, providing a better insight into the parameter of interest. However, the use of the Gibbs sampling procedure may be limited due to its high computational demands.

Chapter 9

CONCLUDING DISCUSSION

Tree breeders must quantify the extent and nature of genetic control of traits of relevance in order to formulate effective breeding strategies. For example, genetic parameters such as heritabilities are necessary for predicting the genetic merit of trees, both heritabilities and age-age genetic correlations are essential for determining the opportunities for early selection, and trait-trait correlations are important for construction of multi-trait selection indices. Similarly, estimates of the genetic relationship between a trait assessed at different sites are necessary for evaluating alternative breeding strategies. The objectives of this study were to estimate genetic and phenotypic relationships between height and stem straightness using an individual tree model, and to determine the optimum selection age for height for a *P. taeda* breeding population in Zimbabwe. Genetic parameters were estimated for height and straightness, important in *P. taeda* breeding as they are associated with the production of timber, and thus of major influence on the profitability of *P. taeda* plantations.

9.1 Genetic parameter estimates

The work reported here is the first use of an individual tree model in *P. taeda*; previous studies have used exclusively sib-covariance models. Given that tree breeding data are expensive to collect, it is important to make best use of them. The individual model is the most efficient for analysis of genetic tests, particularly where full-sib mating designs are used or information from several generations is available, as it appropriately incorporates information on pedigree and genetic relationships between trees. For this reason, it is recommended that the individual tree model should become the method of choice in tree breeding, as the individual animal model has become in animal breeding. The results of this study indicate that height is under moderate to high

genetic control. The heritability and additive variance for height increased with age from 1.5 to 9.5 years, then decreased with age. Despite the fact that this was the first reported use of the individual tree model in *P. taeda*, the heritability estimates for height were within the range reported previously for this species. The heritability estimates here peaked at the same mean height as those reported from studies of *P. taeda* in the USA, suggesting a possible link between mean height and heritability estimate. This suggests that heritability estimates may be better compared across sites on the basis of height rather than age as, at any given age, a slower growing genetic test may be at a different stage of growth and development than a faster growing test. *P. taeda* grows faster in Zimbabwe than in USA, indicating - that at any particular age - heritability estimates may differ partly because of differences in stages of growth. Changes in heritability with age here may also be attributed to changes in the number of trees and in management. The moderate to high heritability estimates for height at all sites may be attributed to good experimental design and good management, and hence better control of environmental factors. Additive genetic variance was found to be more important for height than dominance variance, but the latter contributed substantially to the genetic variance. The phenotypic coefficient of variation for height was high at young ages and decreased to moderate levels at near-harvest age.

Stem straightness is one of the most important traits affecting timber recovery in *P. taeda*, but genetic parameters for this trait are generally lacking. Therefore the results reported here should be of interest to wherever *P. taeda* is grown commercially. Results of this study indicate that straightness is under weak genetic control at very young ages, increasing to moderate levels at maturity. The very low estimates at young ages might be attributed in part to large measurement error, because straightness is difficult to score when the trees are small. Also, trees at a young age are likely to be more affected by environmental variation, resulting in lower heritability estimates than at older ages. Heritability estimates originating from use of an absolute scale, as in this study, are lower than those originating from a site specific (relative) scale. The implications of these results for use of a relative scale are not clear, since results from it cannot be directly compared across sites and it gives biased estimates of genetic differences; however, it has the advantage of generating higher heritabilities than those

from the absolute scale. It would be advantageous to investigate the use of the relative scale in Zimbabwe. The low heritability estimates for straightness might be due to the 7-point scale which has a mid value which assessors might use very frequently for what appears to be average trees, reducing the heritability. With an even scale assessors need to decide whether a tree is below or above average. Use of a even scale such as a 6-point scale should be investigated in Zimbabwe. Additive genetic variance was found to be more important than dominance variance for straightness at older ages, and less important at very young ages. As with height, the phenotypic coefficient of variation was high at young ages and decreased to moderate levels at near-harvest age.

No single site had consistently high heritability estimates, making it difficult to select an environment which facilitates the expression of genetic differences in either trait.

The variance components and heritability estimates presented from data pooled over sites showed evidence of GE and heterogeneity over sites. The results from this study suggest that, even within a region, one should expect GE. Therefore, results of a single test are likely to overestimate gain if the results are to be applied across a region, since they ignore possible GE. Given this fact, the most appropriate estimate to use for gain prediction may actually be that from data pooled across sites. Furthermore, breeding values estimated from a single site are not as precise as those from the pooled data.

The population of *P. taeda* in Zimbabwe has been under selection, and the parents used in this study were not a random sample, but deliberately selected for their superior phenotypic values. Therefore, the genetic variance is expected to be diminished, hence biasing the heritability estimates. The use of the individual tree model in this study should give less biased estimates than would arise from use of the sib covariance model commonly used by tree breeders. Use of the individual tree model is therefore particularly advantageous in advanced generations.

Age-age genetic correlations estimated for height were high, and greater than corresponding estimates reported from some tests in the USA, particularly those involving young ages. This result may be a consequence of management and methodological differences between this and other studies, since the other tests were

unthinned and family correlations were used as approximations of genetic correlations. Family mean correlations are likely to underestimate genetic correlations because the components of family mean correlation include error covariances and error variances. Results of this study support the wide body of data which indicate that the genetic correlations are larger than corresponding phenotypic correlations. Age-age genetic correlations for straightness between 1.5 years and older ages were very low, which could partly be explained by the fact that this trait is difficult to measure at a young age. Therefore, early selection for stem straightness at 1.5 years is not recommended for this *P. taeda* population. Genetic correlations between straightness at older ages were moderate to high, indicating that early selection at ages 9.5 or older in this trait will result in improvement in straightness at rotation age. Genetic correlations between height and straightness were generally low and positive. Selection on height alone should result in improvement in straightness at harvest age.

The precision of genetic parameter estimates in this study varied. Those for heritability estimates were moderate and were slightly improved by pooling data across sites, suggesting that the number of trees per family for individual sites were probably sufficient. However, the precision would have been improved by increasing the number of parents. The precision of genetic correlations were good when the correlations were high, reflecting the fact that the precision depends on the correlation itself. However, when the correlations were moderate or low, the precision was very low due to the low number of parents and the magnitude of the correlations themselves. In future, progeny tests should include many families from many parents in order to estimate genetic parameters with high precision.

The study of the loglikelihood profiles here reveals that the relationships between the loglikelihood and the heritability estimates or genetic correlations were not quadratic. However, the confidence intervals obtained by plotting the loglikelihood profiles and that obtained by assuming a quadratic relationship were not substantially different, implying that the quadratic assumption was adequate.

9.2 Modelling of (co)variances over time

Model choice for estimating age-age genetic correlations is crucial to conclusions one might reach regarding potential gain and optimum selection age. One of the commonly used models for predicting age-age genetic correlations is the generalised predictive model developed for conifers by Lambeth (1980). Lambeth used phenotypic correlations and equated these to genetic correlation. The results of this study support the increasing evidence which indicates that genetic correlations are higher than phenotypic correlations, implying that a Lambeth-type model underestimates genetic correlations. Although phenotypic correlations are estimated with much higher precision than genetic correlations, their use in place of genetic correlations is not recommended because they are significantly lower than corresponding genetic correlations. The underestimation of genetic correlations by the phenotypic models will result in underestimates of gain, and may consequently affect predictions of the optimum selection age. There is accumulation of information on age-age genetic correlations, particularly in USA; the results reported here highlight the need to develop a general prediction model for genetic correlations in conifers. The results of this study indicate that a logarithmic relationship is not necessarily the best model; the best fit for the genetic correlations in this study was the model involving age difference. The models derived in this study (logarithmic and age difference) appear to be the only predictive models based on genetic correlations fitted from young ages to harvest age, and should benefit tropical countries planting *P. taeda* as an exotic. However, it will be important to validate the models with large data sets available from the USA before extending these models to temperate regions. Preliminary results indicate that the logarithm model developed here differs significantly from similar models developed using datasets from old first generation genetic tests, but might agree more closely with models involving advanced generation genetic tests in the USA (Gwaze *et al.* 1997). This suggests that a general prediction model for both the tropical and temperate regions might be possible using data from advanced generation genetic tests in the USA.

The few assessment ages common in forestry, as in this study, tend to favour simple linear models, and may miss subtle but important changes in genetic parameters.

Covariance functions may reflect actual trends in the data because they are more flexible. The difficulties of obtaining a good covariance function with our phenotypic data may be due to variance heteroscedasticity (i.e. expansion in variances over time of most size-related traits). Also, six ages are clearly inadequate for estimating covariance functions, but sufficient to illustrate their potential in forestry for modelling covariances over time, and hence enable estimation of age-age genetic correlations for a trait at any two ages and heritability estimates for the trait at any age. These results highlight the need to conduct further work on the use of covariance functions in tree breeding.

9.3 Predictions of gain and optimum selection age

Forest trees have long generation intervals, and therefore early selection is preferred, as it results in shorter generation intervals, and may lead to increased gain per unit of time, reduced testing costs, and greater adaptability to market changes. Such early selection is necessarily indirect. Early indirect selection also holds promise for quicker incorporation of gains into production, as parents to be used for multiplication can be selected early, and seed orchards or propagation hedges can be culled early. Therefore, sound solutions to reduce generation interval are needed for cost-efficient tree breeding. For a sound evaluation of the effectiveness of early selection and the wise choice of optimum selection age, trends in heritability should be obtained, and model for deriving the age-age genetic correlation has to be selected with care.

This study demonstrates the disadvantages of phenotypic correlation models, such as Lambeth's, for predicting gain and making decisions on optimum selection age. When the phenotypic model was used, gain was only a quarter of those predicted by the genetic correlation models, and optimum selection age was overestimated by up to seven years. One important consequence of inaccurate models is that tree breeding programmes will operate inefficiently. Other related consequences are that the benefits from tree breeding will be undervalued, leading to less favourable appraisal of investment in tree breeding programmes, and misidentification of research priorities (e.g. the importance of identifying methods which induce early flowering).

The predictions of gain and optimum selection age depended on the size and

variation of the heritability estimates, implying that heritabilities also need to be modelled over time. Erroneously assuming heritability is constant will lead to biased estimates of annual genetic gain, and may lead to inaccurate predictions of optimum selection age.

In this study, annual genetic gain was maximised by selection at 10 years, and could be reducing to 3 years if the species were induced to flowering at that age. If the species were induced to flower at 3 years, rather than at the usual age of 10 years, annual genetic gain would be increased by 100%. Methods of artificially inducing flowering in *P. taeda* are available, and tree breeders in Zimbabwe should investigate the economics of promoting early flowering. Other options such as selection of sites with early flowering potential should also be explored in order to reduce the breeding interval. The optimum selection ages found here are likely to change when overlapping generations are considered.

Given that *P. taeda* can be induced to flower at 3 years, and that there is accumulation of information on age-age genetic correlations, particularly in USA, the results highlight the need to develop a general prediction model for age-age genetic correlations in long rotation plantation species.

9.4 Genotype x environment interaction

The evidence presented in this thesis shows that GE for both traits was brought about primarily by a change in rankings of breeding values of genotypes among the four locations. Despite this finding, some parents were found to be consistently better or worse than others at all sites. There was no single environmental factor identified which was causing the interactions in both traits. However, in straightness, interaction may have been partly due to different interpretations of the assessment scale by the different assessment teams. As pointed out by Matheson and Raymond (1984a), GE may be due to other factors which are neither genetic nor environmental, but methodological.

This study appear to be the first for *P. taeda* to evaluate the effect of early selection on one site for predicting mature age performance at another. This suggested approach has merit since early selection is efficient and normally practised within sites

in conifers.

The classical approach to selection of sites to locate progeny tests involve evaluating efficiencies of selection across sites, but these efficiencies have no sampling variations associated with them. Alternative approaches using Bayesian methods such as Gibbs sampling are now available. The study demonstrated the use of Gibbs sampling to estimate variation of the efficiencies of selection and of the probability that each estimate falls within specified interval, enabling more informed decision-making. The results from the Gibbs sampling indicated that Martin (site C) was a better site for locating progeny tests than Tarka (site A) if early selection is considered. Given that most of the plantations are in the region covered by sites A and C, and that early selection is more likely to be practised than selection at harvest age, site C should be selected as the preferred site for locating progeny tests. Results of a single test are likely to overestimate gain if the results are to be used for a region, since it ignores possible GE. In order to yield results that are appropriate to commercial progress, it will be necessary to consider dispersing 3-4 progeny tests within site C region.

There was a tendency for the effect of GE to depend on the age of the trees. For height, GE was more pronounced as the trees aged. This is an important finding because early growth assessments may not be reliable for assessing GE at maturity. Most studies in forestry have examined GE using early growth assessments, less than half rotation age. Results of this study indicate that there is a critical need to evaluate GE up to harvest ages in forest trees.

9.5 Further research

P. taeda flowers at 10 years in Zimbabwe, and consequently, the generation interval is long. Results from this study illustrate that gain would be increased by 100% if the flowering age were reduced to 3 years. Therefore, effort should be put into reducing the flowering age through artificial flower induction or through selection of sites with early flowering potential.

Since the optimum selection age was based on estimates of genetic parameters, there is a risk or an error associated with early selection, and a need to quantify the

uncertainty surrounding the optimum selection age. The Gibbs sampling method presented in this thesis could be a powerful approach to determine the risk associated with the estimation of the optimum selection age.

It is important to verify, with additional genetic tests, whether or not the trends observed in this study are applicable to other tropical countries where *P. taeda* is being planted operationally. More information on genetic parameters of *P. taeda* grown in the tropics could be obtained from tests established in Australia, South Africa and Zimbabwe in 1976. Since similar tests were established in the USA, analyses of these tests would offer direct comparisons of genetic parameters between temperate and tropical regions. Genetic parameters for wood density and branch traits are lacking in *P. taeda* grown in the tropics; these could be derived using the tests used in this study and those from the 1976 series of tests. Pests and diseases in forest plantations result in economic losses. Therefore, a reduction in the frequency of pests and diseases by selection, if practicable, is economically important. In order to improve pests and disease resistance in breeding programmes, it is essential to obtain reliable estimates of genetic and phenotypic relationship between resistance to pests and diseases, and growth and wood quality traits. While information is available on genetic variation in susceptibility to commercially important pests and diseases, particularly fusiform rust, in *P. taeda* in the USA, no information exists on variation in susceptibility to a particularly important pest in Zimbabwe and South Africa, baboons. The economic implications of baboon damage in *P. taeda* plantation stands in these two countries seem to argue strongly for genetic studies in this area.

Results of this study indicate that covariance functions can be derived using forestry data. However, the dataset used in this study was poor, with few point estimates and no phenotypic correlations involving some years. Therefore, the potential of this method needs further evaluation using better forestry data sets.

Although a large body of information now exists on age-age genetic correlation and their use in predictions of optimum selection age, major limitations are the weak information on genetic parameters beyond half rotation age, and reliance on too few assessment ages. The strength of this study is that assessments at near-harvest age were available, allowing realistic predictions of rotation age gains. However, the study also

suffers from the problem of having few point estimates, a problem in many other forestry studies. Therefore, the estimation of optimum selection age warrants further study, but the estimates obtained here could be used as the best available.

9.6 Principal conclusions

The principal conclusions from this study can be summarized as follows:

a) Genetic control of height in *P. taeda* in Zimbabwe is moderate to high, while that of stem straightness is weak, particularly at young ages.

b) Genetic correlations from this study indicate that early section for height to improve height or straightness at harvest age is efficient, but selection for straightness at 1.5 years would adversely alter straightness at harvest age.

c) Although the reproductive biology of *P. taeda* is a barrier to juvenile breeding of the species in Zimbabwe, the study has demonstrated that, were early flowering to be induced, optimum selection age would be reduced from 10 to 3 years and annual genetic gain increased by more than 100%.

d) The large genotype x environment interactions detected in this study indicate that site should be considered when planning progeny tests. The results showed that it was more efficient to make selections at site C, a site in the region where most of the commercial plantations are located.

e) The results suggest that Gibbs sampling is an efficient approach to making decisions about the optimal selection environment in tree breeding programmes.

REFERENCES

Allard, R. W. (1960). *Principles of plant breeding*. Wiley, New York. 485 pp.

Arnold, M and White, T. (1994). *Review of the tree breeding programme of the Zimbabwe Forestry Commission breeding programme*. unpublished. 48 pp.

Baker, J. B. and Langdon, O. G. (1990). *Pinus taeda* L. Loblolly pine. In: *Silvs of North America. Vol. 1*. Confers. USA department of agriculture and forest services handbook. 654 pp.

Balocchi, C. E., Bridgewater, F. E., Zobel, B. J. and Jahromi, S. (1993). Age trends in genetic parameters for tree height in a nonselected population of loblolly pine. *For. Sci.* **39**:231-251.

Baradat, P. (1976). Use of juvenile-mature relationships and information from relatives in combined multitrait selection. In: *Proceedings of IUFRO Joint Meeting on Advanced Generation Breeding*:121-138. Bordeaux, France, June 14-18, 1976.

Barnes, R. D. (1981). *The tree breeding programme in Zimbabwe*. Plans for 1982/83. Report prepared under United Kingdom Overseas Development Administration Technical cooperation scheme.

Barnes, R. D. (1986). Multiple population tree breeding in Zimbabwe. In: *Proceedings of a joint workshop of IUFRO working parties*, Williamsburg, Virginia. North Carolina State University, Raleigh.

Barnes, R. D. (1989). *The forest genetics programme in Zimbabwe. Part I: Philosophy, strategy, techniques and procedures*. Report prepared under United Kingdom Overseas Development Administration Technical cooperation scheme. 129 pp.

Barnes, R. D. and Gibson, G. L. (1986). A method to assess stem straightness in tropical pines. *Common. For. Rev.* **65**:168-171.

Barnes, R. D., Mullin, L. J. and Battle, G. (1992a). Genetic control of fifth year traits in *Pinus patula* Schiede and Deppe. *Silvae Genetica* **41**:242-248.

Barnes, R. D., Mullin, L. J. and Battle, G. (1992b). Genetic control of eighth year traits in *Pinus patula* Schiede and Deppe. *Silvae Genetica* **41**:318-325.

Barrett. R. L. and Mullin, L. J. (1968). *A review of introductions of forest trees in Rhodesia*. Rhodesia Bulletin of Forestry Research no. 1. 227pp.

Bishop, S. C. (1992). Phenotypic and genetic variation in body weight, food intake and energy utilisation in hereford cattle II. Effects of age and length of performance test. *Livestock Production Science* **30**:19-31.

Borralho, N. M. G., Cotterill, P. P. & Kanowski, P. J. (1992a). Genetic control of growth of *Eucalyptus globulus* in Portugal. II. Efficiencies of early selection. *Silvae Genetica* **41**:70-77.

Borralho, N. M. G., Kanowski, P. J. and Cotterill, P. P. (1992b). Genetic control of growth of *Eucalyptus globulus* in Portugal. 1. Genetic and phenotypic parameters. *Silvae Genetica* **41**:39-45.

Borralho, N. M. G., Jarvis, S. F. and Potts, B. M. (1995). Multivariate restricted maximum likelihood estimation of genetic parameters for diameter at 3 to 5 years in *Eucalyptus globulus* base population in Australia. In: *Eucalypt plantations: Improving fibre yield and quality* (eds. B. M. Potts, N. M. G. Borralho, J. B. Reid, R. N. Cromer, W. N. Tibbits and C. A. Raymond). Proc. CRC-IUFRO Conf., Hobart, 19-24 Feb. CRC for Temperate Hardwood Forestry: Hobart. pp. 230-232.

Bramlett, D. L., Williams, C. G. and Burris, L. C. (1995). Surrogate pollen induction shortens the breeding cycle in loblolly pine. *Tree Physiology* **15**:531-535.

Bridgewater, F. E. And Stonecypher, R. W. (1979). Index selection for volume and straightness in a loblolly pine population. In: *Proceedings of the 15th southern conference on forest tree improvement,* Mississippi State University:132-139.

Buford, M. A. and Burkhart H. E. (1987). Genetic improvement effects on growth and yield of loblolly pine plantations. *For. Sci.* **33**:707-724.

Burdon, R. D. (1977). Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. *Silvae Genetica* **26**:168-175.

Burdon, R. D. (1989). Early selection in tree breeding: principles for applying index selection and inferring input parameters. *Can. J. For. Res.* **19**:499-504.

Burdon, R. D. and Low, C. B. (1992). Genetic survey of *Pinus radiata* 6: wood properties, variation, heritabilities and interrelationships with other traits. *N. Z. J. For. Sci.* **22**:228-245.

Burdon, R. D. and Namkoong G. (1983). Short note: Multiple populations and sublines. *Silvae Genetica* **32**:221-222.

Burdon, R. D., Bannister, M. H. and Low C. B. (1992a). Genetic survey of *Pinus radiata*. 3: variance structures and narrow-sense heritabilities for growth variables and morphological traits in seedlings. *N. Z. J. For. Sci*. **22**:160-186.

Burdon, R. D., Bannister, M. H. and Low C. B. (1992b). Genetic survey of *Pinus radiata*. 5:Between-trait and age-age correlations for growth rate, morphology and disease resistance. *N. Z. J. For. Sci*. **22**:211-227.

Burris, L. C., Williams, C. G. And Douglas, S. D. (1991). Flowering response of juvenile selection in loblolly pine. In: *Proceedings of the 21st Southern Forest Tree Improvement Conference*:110-119. Knoxville, Tennessee.

Carson, S. D. (1991). Genotype x environment interaction and optimal number of progeny test sites for improving *Pinus radiata* in New Zealand. *N. Z. J. For. Sci.* **21**:32-49.

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution* **42**:958-968.

Cornelius, J. (1994). Heritabilities and additive genetic coefficients of variation in forest trees. *Can. J. For. Res.* **24**:372-379.

Cotterill, P. P., Dean, C. A and Wyk, van G. (1987). Additive and dominance genetic effects in *Pinus pinaster, P. radiata* and *P. elliottii* and some implications for breeding strategy. *Silvae Genetica* **36**:221-232.

Critchfield, W. B. and Little, E. L. Jr. (1966). *Geographic distribution of the pines of the world*. USDA For. Serv. Misc. publication no. 991. 97 pp.

Dalimore, W. and Jackson, A. B. (1954). *A handbook of Coniferae including Ginkgoaceae*. Edward Arnold publishers, London. 729 pp.

Dean, C. A. (1990). Genetics of growth and wood density in radiata pine. Unpublished PhD thesis, University of Queensland, Australia. 78 pp.

Dickerson, G. E. (1962). Implications of genetic-environment interactions in animal breeding. *Animal Production* **4**:47-63.

Dorman, K. W. (1976). *The genetics and breeding of southern pines*. US Dept. Agric. For. Serv. Agric. Handbook no. 471. 407 pp.

Douglas, S. D., Williams, C. G., Lambeth, C. C., Huber, D. A. and Burris L. C. (1993). Family x environment interaction for sweep in local and non-local seed sources of loblolly pine. In: *Proceedings of the 22nd Southern Forest Improvement conference*. Atlanta, Georgia, June 14-17, 1993.

Eldridge, K., Davidson, J., Harwood, C. and van Wyk, G. (1993). *Eucalypt domestication and breeding*. Clarendon press, Oxford. 288 pp.

Falconer, D. S. (1989). *Introduction to quantitative genetics*. 3rd edition. Longman, London. 438 pp.

FAO (1992). Mixed and pure forest plantations in the tropics and subtropics. FAO Paper 103. 152 pp.

Fimland, E. A. (1973). Estimates of phenotypic and genetic parameters for growth characteristics of young potential AI bulls. *Acta Agric. Scand.* **23**:209-216.

Foster, G. S. (1986). Trends in genetic parameters with stand development and their influence on early selection for volume growth in loblolly pine. *Forest Science* **32**:944-959.

Foster, G. S. and Bridgwater, F. E. (1986). Genetic analysis of fifth-year data from seventeen parent partial diallel of loblolly pine. *Silvae Genetica* **35**:118-122.

Fowells, H. A. (1965). *Silvics of the forest trees of the United States*. USDA For. Serv. Handbook no. 271. 762 pp.

Franklin, E. C. (1979). Model relating levels of development of four north American conifers. *Silvae Genetica* **28**:207-212.

Franklin, E. C. (1989). Theoretical and applied aspects of short-term progeny testing in loblolly pine. In: *Proceedings of the twentieth southern forest tree improvement conference*:259-268. June 26-30, 1989, Charleston, South Carolina.

Genstat 5 Committee (1987). *Genstat 5 reference manual*. Clarendon press, Oxford. 749pp.

Gilks, W. R., Richards, S. and Spiegelhalter, D. J. (Eds.) (1996). *Markov chain monte carlo in practice.* Chapman and Hall.

Gilmour, A. R. (1996). ASREML, A Spatial REML program. NSW Agriculture, Orange, NSW, Australia (Mimeo). 54pp.

Gilmour, A. R., Thompson, R. and Cullis, B. R. (1995). Average information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51**:1440-1450.

Gonzalez, J. S. and Richards, J. (1988). Early selection for wood density in young coastal Douglas-fir trees. *Can. J. For. Res.* **18**:1182-1185.

Green, E. J. and Strawderman, W. E. (1996). Predictive posterior distributions from a Bayesian version of a splash pine yield model. *For. Sci.* **42**:456-464.

Green, E. J., Strawderman, W. E. and Thomas, C. E. (1992). Empirical Bayes development of Honduran pine yield models. *For. Sci.* **38**:21-33.

Gwaze, D. P (1995). Genetic parameter estimates for wood production and quality traits in *Pinus taeda* Linnaeus and implications for breeding. Unpublished. 52pp.

Gwaze, D. P., Bridgwater, F. E., Byram, T., Williams, C. G. and Woolliams, J. A. (1997). General prediction model for age-age genetic correlations in *Pinus taeda*. Paper presented at the Southern Forest Tree Improvement Conference, 9-12 June 1997, Florida, USA.

Hagedorn, S. F. (1994). Results of a *Pinus taeda* progeny/provenance type trial in Natal, South Africa. *S. A. For. J.* **170**:17-25.

Haines, R. (1994). *Biotechnologies in forest tree improvement: with special references to developing countries*. FAO Forestry paper 118. 230pp.

Harvey, W. R. (1987). *User's guide for LSMLW and MIXMDL, PC-1 Version.* Mimeographed paper. Ohio State University. 59pp.

Hayes, J. F. and Hill, W. G. (1981). Modification of estimates of parameters in the construction of genetic selection indices ('bending'). *Biometrics* **37**:483-493.

Hodge, G. R. and White, T. L. (1992). Genetic parameter estimates for growth traits at different ages in slash pine and some implications for breeding. *Silvae Genetica* **41**:252-262.

Jensen, J., Wang, C. S., Sorensen, D. A. and Gianola, D. (1994). Bayesian inference on variance and covariance components for traits influenced by maternal and direct genetic effects, using the Gibbs sampler. *Acta Agric. Scand. Sect. A., Animal Sci.* **44**:193-201.

Jett, J. B., McKeand, S. E. and Weir, R. J. (1991). Stability of juvenile wood specific gravity of loblolly pine in diverse geographic areas. *Can. J. For. Res.* **21**:1080-1085.

Johnson, I. G. (1992). Family-site interactions in radiata pine families in New South Wales, Australia. *Silvae Genetica* **41**:55-62.

Johnson, G. R. and Burdon, R. D. (1990). Family-site interaction in *Pinus radiata*: Implications for progeny testing strategy and regionalised breeding in New Zealand. *Silvae Genetica* **39**: 55-62.

King, J. N. and Burdon, R. D. (1991). Time trends in inheritance and projected efficiencies of early selection in a large 17-year-old progeny test of *Pinus radiata*. *Can. J. For. Res.* **21**:1200-1207.

Kirkpatrick, M., Hill, W. G. and Thompson, R. (1994). Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genet. Res. Camb.* **64**:57-69.

Kirkpatrick, M., Lofsvold, D. and Bulmer, M. (1990). Analysis of inheritance, selection and evolution of growth trajectories. *Genetics* **124**:979-993.

Klein, T. W., DeFries, J. C. and Finkbeiner (1973). Heritability and genetic correlation: standard errors of estimates and sample size. *Behavior Genetics* **3**:355-364.

Knowe, S. A. and Foster, G. S. (1989). Application of growth models for simulating genetic gain of loblolly pine. *For. Sci.* **35**:211-228.

Koch, R. M., Cundiff, L. V. and Gregory, K. E. (1982). Influence of post-weaning gain on estimates of heritability and genetic correlations. *J. Anim. Sci.* **55**:1310-1318.

Koots, K. R. and Gibson. J. P. (1996). Realized sampling variances of estimates of parameters and differences between genetic and phenotypic correlations. *Genetic* **143**:1409-1416.

Lambeth, C. C. (1980). Juvenile-mature correlations in Pinaceae and implications for early selection. *Forest Science* **26**:571-580.

Lambeth, C., van Buijtenen, J. P., McCollough, R. B. and Duke, S. D. (1983). Early selection is effective in 20-year-old genetic tests of loblolly pine. *Silvae Genetica* **32**:210-215.

Li, B., McKeand, S. E. and Weir, R. J. (1996). Genetic parameter estimates and selection efficiency for the loblolly pine breeding in the south-eastern USA. *In* Tree Improvement for sustainable tropical forestry, Dieters, M. J., Matheson, A. C., Nikles, D. G., Harwood, C. E. and Walker, S. M. (*eds*.). pp. 164-168. Proc. QFRI-IUFRO Conf., Caloundra, Queensland, Australia, 27 October-1 November 1996.

Li, B., Williams, C. G., Carson, W. C., Harrington, C. A. and Lambeth, C. C. (1991). Gain efficiency in short-term testing: experimental results. *Can. J. For. Res.* **22**:290-297.

Loo, J. A., Tauer, C. G. and van Buijtenen, J. P. (1984). Juvenile-mature relationships and heritability estimates of several traits in loblolly pine (*Pinus taeda*). *Can. J. For. Res*. **14**:822-825.

Magnussen, S. (1988). Minimum age-to-age correlations in early selections. *For. Sci*. **34**:928-938.

Magnussen, S. (1989). Determination of optimum selection ages: A simulation approach. In: *Proceedings of the 20th Southern Forest Tree Improvement Conference*:269-285.

Magnussen, S. and Keith, C. T. (1990). Genetic improvement of volume and wood properties of jack pine: selection strategies. *Forestry Chronicle* **66**:281-286.

Matheson, A. C. (1978). Genotype x environment interactions. In: *Proceedings of a joint workshop on progress and problems of genetic improvement of tropical forest trees*:227-236. Brisbane, Queensland, Australia, 4-7 April 1977.

Matheson, A. C. and Raymond, C. A. (1984a). Effects of thinning in progeny tests on estimates of genetic parameters in *Pinus radiata*. *Silvae Genetica* **33**:125-128.

Matheson, A. C. and Raymond, C. A. (1984b). The impact of genotype x environment interactions on Australian *Pinus radiata* breeding programs. *Aust. For. Res.* **14**:11-25.

Matheson, A. C., Spencer, D. J. and Magnussen, D. (1994). Optimum age for selection in *Pinus radiata* using basal area under bark for age-age correlations. *Silvae Genetica* **43**:352-357.

Matziris, D. I. and Zobel, B. J. (1973). Inheritance and correlations of juvenile characteristics in loblolly pine (*Pinus taeda* L.). *Silvae Genetica* **22**:38-45.

McKeand, S. E. (1988). Optimum age for family selection for growth in genetic tests of loblolly pine. *For. Sci.* **34**:400-411.

Meyer, K. (1989). Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genet. Sel. Evol.* **21**:317-340.

Meyer, K. (1993). DFREML version 2.1. Programs to estimate variance component by restricted maximum likelihood using a derivative free algorithm. User notes. Animal genetics and breeding unit, University of New England, Armidale, NSW, Australia (Mimeo). 101pp.

Meyer, K. and Hill, W. G. (1997). Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by Restricted Maximum Likelihood. *Livestock Production Science* **47**:185-200.

Minitab Release 7.2 Sun-4 Version (1989). Minitab INC.

Mirov, N. T. (1967). *The genus Pinus*. The Ronald Press Co., New York. 602 pp.

Mullin, L. J., Barnes, R. D. and Prevost, M. J. (1978). *A review of the southern pines in Rhodesia*. Rhodesia Bulletin of Forestry Research no. 7. 328 pp.

Namkoong, G., Barnes, R. D. and Burley, J. (1980). *A philosophy of breeding strategy for tropical forest trees*. Tropical forestry paper no. 16. Commonwealth Forestry Institute, Oxford. 67pp.

Namkoong, G., Kang, H. C. And Brouard, J. S. (1988). *Tree breeding: principles and strategies*. Springer-Verlag, New York. 180pp.

Namkoong, G., Usanis, R. A. and Silen, R. R. (1972). Age-related variation in genetic control of height growth in Douglas-fir. *Theoretical and Applied Genetics* **42**:151-159.

Newman, D. H. and Williams, C. G. (1991). The incorporation of risk in optimum selection age determination. *For. Sci.* **37**:1350-1364.

Owino, F. (1977a). Genetic divergence in selected populations of loblolly pine. *Silvae Genetica* **26**:64-66.

Owino, F. (1977b). Genotype x environment interaction and genotypic stability in loblolly pine. II.Genotypic stability comparisons. *Silvae Genetica* **26**:21-26.

Owino, F. and Zobel, B. (1977). Genotype x environment interaction and genotypic stability in loblolly pine. III. Heterosis and heterosis x environment interaction. *Silvae Genetica* **26**:114-116.

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

Pederick, L. A. (1990). Family x site interactions in *Pinus radiata* in Victoria, Australia and implications for breeding strategy. *Silvae Genetica* **39**:134-140.

Poynton, R. J. (1979). *Tree planting in Southern Africa. Vol. I. The Pines*. Government printer, Pretoria, South Africa.

Pswarayi, I. Z. (1993). Genetic parameters and selection indices for a population of *Pinus elliottii* Engelm. var *elliottii*. PhD Thesis, Oxford University. 258 pp.

Pswarayi, I. Z., Barnes, R. D., Birks, J. S. and Kanowski, P. J. (1996). Genetic parameter estimates for production and quality traits of *Pinus elliottii* Engelm. var. *elliottii* in Zimbabwe. *Silvae Genetica* **45**:216-222.

Pswarayi, I. Z., Barnes, R. D., Birks, J. S. and Kanowski, P. J. Genotype-environment interaction in a population of *Pinus elliottii* Engelm. var. *elliotti. Silvae Genetica* (in press).

Raymond, C. A. and Cotterill, P. P. (1990). Methods of assessing crown form in *Pinus radiata*. *Silvae Genetica* **39**:67-71.

Raymond, C. A. and Lindgren, D. (1990). Genetic flexibility-A model for determining the range of suitable environments for a seed source. *Silvae Genetica* **39**:112-120.

Richards, F. (1959). A flexible growth function for empirical use. *J. Exp. Bot.* **10**:290-300.

Riemenschneider, D. E. (1988). Heritability, age-age correlations, and inferences regarding juvenile selection in jack pine. *For. Sci.* **34**:1076-1082.

Robertson, A. (1959). The sampling variance of the genetic correlation coefficient. *Biometrics* **15**:469-485.

Robinson, J. F., van Buijtenen, J. P. and Long, E. M. (1984). Traits measured on seedlings can be used to select for later volume of loblolly pine. *S. J. A. For.* **8**:59-63.

SAS Institute, Inc. (1988). *SAS/STAT user's guide*. SAS Institute Inc. Cary, N. C. 1028pp.

Sharma, N. P, Rowe, R., Openshaw, K. and Jacobson, M. (1992). World forests in perspective. In: *Managing the world's forests: Looking for balance between conservation and development:* 17-31. N. P. Sharma (ed.). Kendall/Hunt publishing company, Iowa.

Shelbourne, C. J. A. and Stonecypher, R. W. (1971). The inheritance of bole straightness in young loblolly-pine. *Silvae Genetica* **20**:151-156.

Sorensen, D. (1997). Gibbs sampling in quantitative genetics. Unpublished internal report no. 82.1996. Danish Institute of Animal Science, Denmark.

Sorensen, D. A., Wang, C. S., Jensen, J. and Gianola, D. (1994). Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet. Sel. Evol.* **26**:333-360.

Talbert, J. T., Jett, J. B. and Bryant, R. L. (1983). Inheritance of wood specific gravity in an unimproved loblolly pine population. *Silvae Genetica* **32**:33-37.

van Buijtenen, J. P. (1969). Progress and problems in forest tree selection. In: *Proceedings of the 10th southern conference on forest tree improvement*::17-26. Houston, Texas.

van Buijtenen, J. P. (1992). Fundamental genetic principles. In: *Handbook of quantitative forest genetics*: 29-68. Fins, L., Friedman, S. T. and Brotschol J. V. (eds.). Klumer academic publishers, Dordrecht.

van Tassell, C. P. and Van Vleck, L. D. (1995). A manual for use of MTGSAM. A set of fortran programs to apply Gibbs sampling to animal models for variance component estimation. US Department of Agriculture, Agricultural Research Service.

Vargas-Hernandez, J. and Adams, W. T. (1991). Genetic variation of wood density components in young coastal douglas-fir: implications for the tree breeding. *Can. J. For. Res.* **21**:1801-1807.

Vidakovic, M. (1991). *Conifers morphology and variation*. Translated by Maja Soljan. Boro Brekalo (ed.). Graficki Zavod Hrvatske publishers.

Visscher, P. M., Thompson, R. and Hill, W. G. (1991). Estimation of genetic and environmental variances for fat yield in individual herds and an investigation into heterogeneity of variance between herds.
Livestock Production Science **78**:273-290.

Wahlenberg, W. G. (1960). *Loblolly pine: its use, ecology, regeneration, protection, growth and management*. The School of Forestry, Duke University. 603 pp.

Wakeley, P. C. (1971). Relation of thirtieth-year to earlier dimensions of southern pines. *For. Sci.* **17**:200-209.

Wang, C. S., Rutledge, J. J. and Gianola, D. (1994). Bayesian analysis of mixed linear models via Gibbs sampling with application to litter size in Iberian pigs. *Genet. Sel. Evol.* **26**:91-115.

Wei, X and Borralho, N. M. G. (1996). A simple model to describe age trends in heritability in short rotation tree species. *In* Tree Improvement for sustainable tropical forestry, Dieters, M. J., Matheson, A. C., Nikles, D. G., Harwood, C. E. and Walker, S. M. (*eds.*). pp. 178-181. Proc. QFRI-IUFRO Conf., Caloundra, Queensland, Australia, 27 October-1 November (Queensland Forestry Research Institute, Gympie).

Weir, R. J. and Goddard, R. E. (1986). Advanced generation operational breeding programs for loblolly and slash pine. In: *Proceedings of a workshop on advanced generation breeding: current status and research needs*:21-25. June 6-7, 1984, Baton Rouge, Louisiana.

Wetherill, G. B. (1981). *Intermediate statistical methods*. Chapman and Hall Ltd, London. 390pp.

Williams, C. G. and Lambeth, C. C. (1989). Bole straightness measurement for advanced-generation loblolly pine genetic tests: 20 years of results. *Silvae Genetica* **38**:212-217.

Williams, C. G. and Megraw, R. A. (1994). Juvenile-mature relationships for wood density in *Pinus taeda*. *Can. J. For. Res.* **24**:714-722.

Williams, C. G., Hamrick, J. L. and Lewis, P. O. Multiple-population versus hierarchical conifer breeding programs: a comparison of genetic diversity levels. *Theoretical and Applied Genetics* (in press).

Woolaston, R. R., Kanowski, P. J. And Nikles, D. G. (1990). Genetic parameter estimates for *Pinus caribaea var. hondurensis* in coastal Queensland, Australia. *Silvae Genetica* **39**:21-28.

Xie, C. -Y. and Ying, C. C. (1996). Heritabilities, age-age correlations, and early selection in lodgepole pine (*Pinus contorta* ssp. *Latifolia*). *Silvae Genetica* **45**:101-107.

Zobel, B and Talbert, J. (1984). *Applied forest tree improvement*. John Wiley and sons. New York. 505 pp.

Zobel, B., Kellison, R. and Matthias, M. (1969). Genetic improvement in forest trees: growth rate and wood characteristics in young loblolly pine. In: *Proceedings of the 10th southern conference on forest tree improvement:*59-75. Houston, Texas.