METSÄNTUTKIMUSLAITOKSEN TIEDONANTOJA 826, 2002
FINNISH FOREST RESEARCH INSTITUTE, RESEARCH PAPERS 826, 2002

Evaluation of options for use in efficient genetic field testing of *Pinus sylvestris* (L.)

Matti Haapanen

VANTAAN TUTKIMUSKESKUS – VANTAA RESEARCH CENTRE

METLA

METSÄNTUTKIMUSLAITOKSEN TIEDONANTOJA 826, 2002
FINNISH FOREST RESEARCH INSTITUTE, RESEARCH PAPERS 826, 2002

Evaluation of options for use in efficient genetic field testing of *Pinus sylvestris* (L.)

Matti Haapanen

Finnish Forest Research Institute
Vantaa Research Centre

Academic Dissertation in Forest Tree Breeding

To be presented with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public critisism in Auditorium XII of the University main Building, Fabianinkatu 33, on February 22, 2002, at 12 o'clock.

Haapanen, Matti 2002

Supervisor:

Dr. **Pertti Pulkkinen**
Finnish Forest Research Institute
Vantaa Research Centre, Haapastensyrjä Breeding Station

Reviewers:

Dr. **Gunnar Jansson**
Skogforsk
Uppsala, Sweden

and

Dr. **Tapani Pöykkö**
Häme Polytechnic

Opponent:

Dr. **Tore Ericsson**
Skogforsk
Sävar, Sweden

Publisher: Finnish Forest Research Institute, Vantaa Research Centre
P.O. Box 18, FIN-01301 Vantaa, Finland
Accepted by **Kari Mielikäinen**, Research Director, 7.1. 2002.

ISBN 951-40-1806-0
ISSN 0358-4283

Hakapaino Oy
Helsinki 2002

Contents

Abstract

The present study aims at evaluating the impact of various methods and options to genetic field testing of Scots pine (*Pinus sylvestris* L.), assessing the quality of information from various types of ongoing progeny trials, and suggesting appropriate modifications to improve the efficiency of future genetic testing. The study material consisted of large quantities of tree-height data from young Scots pine progeny trials in southern Finland. The test entries investigated were open-pollinated offspring of first-generation plus-trees.

The results suggest there is much scope for improving the efficiency of genetic field testing. Multiple-tree block plots (usually containing 25 trees), employed in the majority of the Finnish trials today, were found to be characterised by low statistical efficiency and inappropriate for precise and cost-efficient progeny testing. The economics of future testing could be significantly improved by substituting them with single-tree or non-contiguous plots offering higher precision and smaller test sizes. Sampling of only a fraction of trees for measurement can be used to cut a part of the extranous costs associated with inefficiently large plots. The various statistics associated with subsamples comprising roughly half of the trees of a 25-tree plot were found to be comparable to those obtained from doubly as laborious complete samples.

The performance of half-sib families across replicated progeny trials showed noticeable volatility. As a result, the estimates of the additive genetic variance and heritability from individual trials were inflated approximately by half. Type-B genetic correlations were around 0.6 on average. In addition, changes in family rankings between parallel trials were commonplace. The high degree of family instability could not be explained by any external factor distinguishing between trials. The true biological significance of the phenomenon was hence judged questionable. To ensure accurate parental rankings, the observed levels of interaction between families and sites make it necessary to improve the control of environmental variability and to distribute the total testing effort into a fairly large number of test sites.

The time trends in some of the key genetic parameter estimates (heritability, type-B correlation) were weak or non-existent over the first two decades of testing. The average level of narrow-sense heritability for tree height was low ($h^2 \sim 0.12$). Age-age genetic correlations for tree height were positive and moderately high, and could be estimated with a reasonable accuracy by a regression analysis using the log of the age ratio as the independent variable. The relative gains from

indirect selection increased with improved information on genetic values. Indirect parental (backward) selection provided significantly more gains per unit time than direct selection for age-20 height, whilst for within-family (forward) selection the corresponding gains were small. Some progeny test conditions consistently favored selection whereas some others proved to be inferior. The closer initial spacing and improved site quality (rapid growth) independently contributed to higher heritability. Accordingly, the highest selection efficiencies were associated with the densely spaced test orchard trials on homogeneous (formerly agricultural) sites. On the contrary, the conventional widely-spaced trials placed on forest sites (the presently most common combination), consistently showed the least response to indirect selection. A proposal is made to reduce the use of heterogeneous forest sites to a minimum in the future testing of Scots pine.

List of original publications

1. **Haapanen, M**. 1992. Effect of plot size and shape on the efficiency of progeny tests. *Silva Fennica* 26(4): 210-209.

2. **Haapanen, M.** 1995. Within-plot subsampling for assessment in progeny trials of Scots pine. *Silva Fennica* 29(1): 13-19.

3. **Haapanen, M.** 1996. Impact of family-by-trial interaction on the utility of progeny testing methods for Scots pine. *Silvae Genetica* 45(2-3): 130-135.

4. **Haapanen, M.** 2001. Time trends in genetic parameter estimates and selection efficiency for Scots pine in relation to field testing method. *Forest Genetics* 8(2): 129-144.

1. Introduction

1.1. The role of genetic testing in tree breeding

Forest tree breeding is an activity aimed at the sustainable improvement in commercially valuable traits by manipulating the frequency of alleles that influence these traits (ZOBEL and TALBERT 1984). To spawn desirable changes in the population genetic make-up, tree breeders employ the same strong forces that drive evolution in nature: selection and recombination. Their artificial counterparts, directional selection and controlled intermating of the selected genotypes, are the key measures used to take advantage of the high levels of genetic variability characterising undomesticated populations of forest trees (WRIGHT 1976). As distinct from natural evolution, however, reproductive success is not the primary target of genetic improvement. The latter is carried out to increase the profitability of plantation forests, and while it has a biological basis, it is aimed at purely economic objectives (HATTEMER 1991). Consequently, a sufficient compensation in terms of genetic progress must be gained to justify expenses given over to selection and other central breeding operations.

The degree of genetic gain due to selection in one generation (ΔG) is a function of selection intensity (i), the amount of genetic variation (here expressed by σ_G, the genetic standard deviation), and the reliability with which individual genetic values can be determined (r_{GP}) (NANSON 1989, HILL et al. 1998). The functional relationship can be presented (under assumptions of normality) in terms of a triple product:

$$\Delta G = i \; \sigma_G \; r_{GP} \qquad [1]$$

In applied tree breeding programmes, selection intensity is fully controlled, whereas the amount of genetic variability must be (in the short term) considered as a fixed constant. From the standpoint of a tree breeder, and also of this study, the accuracy of genetic prediction (r_{GP}), or "geno-phenotypic correlation" as termed by WHITE and HODGE (1991), is the component of most interest since it is directly affected by the decisions put into action in a breeding programme. In the long run, the genetic advance becomes predominantly dependent on the severity with which selection is practised and the management of population coancestry, through their effect on σ_G (NAMKOONG 1974). This issue, actively explored in recent times (e.g., LINDGREN et al. 1995, KERR et al. 1998, ROSVALL et al. 1998), falls, however, beyond the scope of this study.

The accuracy of genetic prediction is reduced by the environmentally determined contribution to the observable variation. In fact, non-genetic factors (summing up to the 'environment') ordinarily play considerably more important role than the genetic make-up in shaping the phenotype. When evaluating the prospects of genetic improvement, it is essential to know the magnitude of the hereditary component of the total variation, measured by the heritability ratio (FALCONER 1981, NYQUIST 1991). For most production traits in forest trees, heritability values tend to be fairly modest (e.g., CORNELIUS 1994, HAAPANEN et al. 1997). This means the phenotype does not predict the genetic merit very well and is thus not a very promising object of selection. To be effective in the longer run, phenotypic selection necessitates the simultaneous maintenance of a high selection differential and an extensive genetic base. Eventually, this is likely to result in prohibitively high costs. Therefore, selection for the phenotype is normally considered as a valid option only when further information on the true genetic values is lacking (ZOBEL and TALBERT 1984). This situation is regularly confronted in the initial stage of breeding when the trees of native stands are selected to the founder population on the basis of their outstanding phenotype.

Chances for improved genetic returns increase when a tree breeding programme moves onward from the mass selection phase to deal with more advanced generations of trees (NAMKOONG 1979). The problem of low heritability (prediction accuracy) can be overcome with more information, which is gained in the process of *genetic testing*. This involves comparing genotypic elements (usually with a known degree of relatedness) in conditions where statistically legitimate experimental designs are used to control the environmental noise. In more general terms, genetic testing could be defined as an application of the scientific method to selection, which relies on the understanding of fundamental principles of quantitative-genetic theory, experimental design and statistical inference.

Thus far, field test plantations have retained their status as the primary device of genetic evaluation in forest trees. Progeny testing, which is often rather narrowly defined as the means of providing the value of a genotype based on the average performance of its offspring (ALLARD 1960), is the prime tree-breeding application employing designed field trials. This is also the genre of genetic testing principally addressed in this study. The concept 'progeny trial' is today well established but essentially a misnomer since the experiments in this category serve breeding in a multitude of ways, of which the estimation of parental genetic values ("GCA testing") is just one. In addition to this main task, progeny trials

provide an insight into the genetic variation structure of the tree population in question. This information is pivotal to the planning of a tree breeding programme, including the need to make well-informed decisions about selection and mating strategies and environments where the genetically improved materials can be safely deployed. Field trials also enable effective monitoring of genetic gains and prediction of forthcoming advances. Verified genetic gains are crucial both for demonstrating the profitability of breeding and fine-tuning routinely used growth models that are customarily based on data from unimproved stands. Last but not least, progeny trials comprise the source of genetic material from which the next-generation selections are usually drawn (ZOBEL and TALBERT 1984, LINDGREN 1991, MIKOLA 1993).

1.2. The demand for efficiency

Testing programmes should be implemented in a way that maximises genetic progress in the short and long term (ZOBEL and TALBERT 1984, BURDON 1986). Ensuring sufficient quantity and quality of data is obviously a top priority. As a further requirement, these data should be obtained at a reasonable cost. These are contradictory demands since genetic field testing, as a rule, requires substantial investments in land, equipment and various activities. The latter comprise, for instance, expenses due to raising the planting stock, soil preparation, planting, weed control, fencing, thinnings and repeated measurements. Altogether, genetic (progeny) testing is the costliest and the most elaborate phase of the breeding cycle (ZOBEL and TALBERT 1984). Consequently, this critical phase needs to be carried out by exploiting the necessary resources as economically as possible (LAMBETH et al. 1983b). The incentives and arguments for cost-efficient testing become, of course, most influential in circumstances of diminishing financial support to tree breeding.

Much of the field-testing research in recent years has revolved around the optimal allocation of resources at the time of trial-installation. In this context, the concept of 'efficiency' typically refers to the amount of information obtained per unit of resources sacrificed, or the number of individuals needed to achieve a defined precision for the estimates of genetic-entry means. Some of the key issues involve the choice of efficient plot type (e.g., WRIGHT and FREELAND 1960, CONKLE 1963, LOO-DINKINS and TAUER 1987, JANSSON et al. 1998), the benefits of incomplete block designs (e.g., MCCUTCHAN et al. 1985, WILLIAMS and MATHESON 1994, FU et al. 1998), and the most effective ways of replicating test materials within and among trials (e.g., LINDGREN 1985, WHITE and HODGE 1992, RUSSELL and LIBBY 1986, RUSSELL and LOO-DINKINS 1993). The risk of losing efficiency due to a

non-appropriate design is evident and real, which explains the generous attention given to these issues. Nevertheless, the decision on an initial trial design represents only one step in the sequence of actions which collectively determine the outcome of testing and selection. Experimental efficiency is a complex issue (WRIGHT and FREELAND 1960) and if significant enhancements are pursued, the whole data generating procedure, and not just a single part of it, must be reviewed. In monitoring the various steps of this process, the collection, the entry, and the analysis of data also need to be given attention. Emphasizing this holistic view, it is more fruitful to comprehend 'efficiency' in this context as the extent to which the information available to the breeder is utilised and successfully incorporated into selection decisions.

The genetic improvement of forest trees has a short history, and the experience of genetic field testing is even more concise. In many parts of the world, applied tree breeding programmes – including that of Scots pine (*Pinus sylvestris* L.) in Finland – are presently completing the first round of field testing and selection, and entering the phase where new screening needs to be done among freshly selected candidates (MATHESON and COTTERILL 1990). This is obviously an appropriate stage for evaluating the returns from the bygone tree breeding activities. The field trials established over the past decades have unquestionably provided a lot of new knowledge on the inherent variability of our tree species, and the potentiality of genetic advance through selection. In some important aspects, however, the results from the first cycle of genetic field testing have not clearly come up to expectations. As MAGNUSSEN (1993) put it, referring to his experiences of genetic field trials, "*...the odds of obtaining unbiased and accurate estimates of genetic variance components and a reliable ranking of entries across several environments are not very encouraging*". Adversities and disappointments are obviously an inevitable part of pioneering work, and should be accepted as such. The true challenge is how successful we will be in turning the information on the causes of failures and successes of the present-day field trials to the advantage of advanced-generation breeding. Evidently, many current ideas on tree selection and field testing will have to be modified and changed. Learning from the past is the most productive way to gather the knowledge crucial for revising the properties of future genetic trials and raising the efficiency of genetic testing to a new higher level.

2. Objectives of the study

The present study, relying on data accrued from contemporary Scots pine progeny trials in Finland, aimed at: 1) quantifying the impact of various experimental options on the efficiency of genetic field testing, and 2) evaluating the value of information from progeny trials in progress. Where appropriate, modifications to current practices, which could amend the efficiency of the future genetic testing, are suggested.

The key questions posed in this study, refering to the Finnish progeny testing of Scots pine, were:

- Are there prospects of improving the efficiency of genetic testing by modifying established procedures of trial installation and data collection?
- How influential are genotype-by-environment interactions?
- What risks may be associated with the test orchard method?
- What tendencies may be distinguished in the development of key genetic parameters over time?
- How effective is indirect selection and to what degree is it affected by the distinctive features (spacing, site fertility) of field testing methods?

The four research papers included focus on the following topics: the efficiency of different types of experimental plots (**I**), the chances of improving the efficiency of measurements by within-plot subsampling (**II**), the importance of family-by-trial and family-by-testing method interactions (**III**), and the time trends in genetic parameter estimates and selection efficiency and their dependency on field testing methods (**IV**).

3. The general framework

3.1. Problems and challenges of forest genetic field testing

Setting up field trials for effective genetic evaluation of forest trees has turned out to be a more challenging task than probably envisaged by the early pioneers of forest tree breeding. In fact, there may not be another group of plants that has given as much concern over genetic testing as trees (LAMBETH 1986, MAGNUSSEN 1993). The adverse features of forest trees as experimental organisms include, in particular, their progressively increasing size and their long economic maturation time (ZOBEL and TALBERT 1984, BURDON 1986). As trees grow bigger, measurements get logistically more complicated. The physical size also makes it necessary to establish trials that occupy large areas of land. The long rotations, in turn, call for long-term maintenance of test plantations. The benefits of large-sized trials run over long periods of time, are, however, compromised by increasing costs and decreasing marginal benefits: genetic gains per unit time are inversely proportional to the length of a selection cycle (BURDON 1989). When faced with limited resources, optimised field test sizes (RUSSELL and LOO-DINKINS 1993) and reduced testing time (LAMBETH 1983, JANSSON 1998) are obviously the key shortcuts to improved efficiency.

The usefulness of genetic evaluation trials can be assessed by two criteria: discrimination ability and prediction ability. The former norm refers to the successfulness in separating between the genetic and environmental effects that contribute to the variation of measured traits (sometimes referred to as 'variates'). Prediction ability, in turn, refers to the indirect nature of genetic testing of forest trees in which the measured traits are customarily not the ones actually targeted, denoting the indicative value, or trustworthiness, of observations made on adolescent trees. For indirect selection to be productive, information on the material being selected should meet both of these requirements. A trial with good genetic resolution in juvenile characteristics may fail to serve selection if the observations are not adequately correlated to the true target traits. Some authors have expressed concerns, for instance, about the value of data from early tests and intensively managed field trials (FRANKLIN 1979, NIENSTAEDT 1984, HODGE and WHITE 1992, WU et al. 1997).

3.2. Discrimination ability

The various objectives of progeny testing are generally, one way or another, conditional to the precise estimation of genetic effects (MIKOLA 1993). The capability of a trial to discriminate between these (often subtle) effects is often assessed by the smallest difference between two family means detected as statistically significant at a chosen level of risk (WRIGHT and FREELAND 1960, CONKLE 1963, LAMBETH et al. 1983a, STEEL et al. 1997). This critical parameter, termed the least significant difference (STEEL et al. 1997), is a derivative measure of statistical precision. The precision is commonly defined as the reciprocal of the variance of an estimate ($1/\sigma^2_{\bar{x}} = n/\sigma^2$). According to this definition, the precision of field testing can be ultimately enhanced either 1) by increasing the number of replications or 2) by eliminating differences extraneous to those being investigated (STEEL et al. 1997).

To produce a successful experiment, there is a need to determine the degree of replication that is required to achieve the required levels of precision. In general, this is complicated due to scant information about the value of σ^2 (WOOLLONS 1980). Eagerness to assure the achievement of test objectives can lead to excessive replication, which is costly and and may, however, fail to reduce error if the implementation of the trial is somehow flawed (STEEL et al. 1997). Thus, refinement of experimental techniques and practices is clearly a more appealing way to improve the discrimination ability than disproportionate increases in the number of replicates. Improved control of nuisance variation may even allow cutbacks in the replication, and thus reduce the costs of testing. This conforms to the general goal of keeping experiments as small as possible with regard to the desired precision (LAMBETH et al. 1983a).

Individual field trials of forest trees, even those with identical designs and materials, can noticeably vary in their ability to detect genetic differences (WHITE and HODGE 1989, **III**, **IV**). The heterogeneity among trials in their margins for error stems from scores of factors, each of which can markedly lower the discrimination ability (BURDON 1986). Edaphic site heterogeneity is usually considered to be the primary contributor to the residual variance (FRANKLIN 1971). Even carefully selected trial sites may comprise ample variation in site topography, moisture, fertility, and the physical texture of soil (ZOBEL and TALBERT 1984, LOO-DINKINS and TAUER 1987). In an investigation of soil heterogeneity of a representative Scots pine dominated site in Finland, all the examined mineral soil properties showed abundant variability over short distances (JÄRVINEN et al. 1993). Some portion of the soil variation, in so far as it shows a pronounced pattern (such as visible topographic or moisture gradients), may be

controlled by blocking. Tree breeding trials most frequently employ randomized complete block designs that are simple to lay out and assess and which allow straightforward statistical analysis (ZOBEL and TALBERT 1984, WILLIAMS and MATHESON 1994). The precision of RCB designs is, however, to some extent dependent on the number of entries. Increasing this number enlarges the space occupied by blocks, making it increasingly more difficult to retain the internal uniformity of blocks. In a rather typical situation, where a high degree of environmental variability is coupled with a large number of families to be tested, incomplete block designs (WILLIAMS and MATHESON 1994, FU et al. 1998) open up opportunities for more efficient control of experimental error. Spatial adjustment techniques, e.g. trend surface smoothing, nearest neighbour correction or post-blocking, may also be used to adjust for environmental gradients (e.g., BONGARTEN and DOWD 1987, MAGNUSSEN 1990, MAGNUSSEN 1994, ERICSSON 1997). The fine-grained part of the microsite variability, however, usually remains immune to any attempts on elimination (MAGNUSSEN 1994). In field progeny trials with multiunit plots, the substantial within-plot variances mostly epitomize this small-scale environmental variability (**IV**).

Inherent variation of experimental material is another important factor contributing to residual variance. SHIUE and PAULEY (1961) pointed out that the relative magnitude of this variation component is much greater in trees than in (often inbred) crop plants in which genetic differences are effectively averaged out due to the large number of plants used to test each entry. The uniformity of test materials can further decrease as a result of improper handling of planting stock, silvicultural management, and occurrence of environmental stresses. The effects of these factors frequently emerge as outliers and missing observations in the data (MAGNUSSEN 1993). Some portion of the unexplained variation evidently arises from various mistakes and mishaps in the collection of experimental data. These comprise, e.g., inaccuracies in measuring and data recording (especially in visual assessment of traits) and misidentification of genetic entries in the field (ERICSSON 1999).

3.3. Prediction ability

3.3.1. Benefits and pitfalls of early testing

The long commercial rotations of forest trees rarely allow field testing to be extended until final harvest (BURDON 1986, LAMBETH 1986). Running progeny trials over long periods of time is not just economically unsound due to vanishing

genetic gains per unit of time, it also entails a risk of exposing test materials to various indiscriminate stresses that may prove fatal to the objectives of the trial (MIKOLA 1985). For these reasons, genetic values of trees are customarily predicted using observations made of their offspring at a fraction of the generation interval. As the trees subject to testing are, as a rule, not allowed to express their full potential, the selection is characteristically indirect by nature (WHITE and HODGE 1991).

Testing and selection for pre-mature performance have obvious benefits over postponed selection for harvest-age traits, which comprise the real target of improvement. Early genetic testing enables, for instance, faster generation turnover, higher selection intensity, easier measurement, and quicker incorporation of genetic gains into forestry. In particular, indirect selection for early performance can be expected to yield the maximum genetic gain per unit of time. Early testing also holds promise for more compact genetic tests (LAMBETH 1980, WU 1998). As a matter of fact, the use of pre-mature information in selection may be seen as the most important method of improving the efficiency of forest tree breeding. Not surprisingly, the various facets of early testing and selection have received major attention in tree improvement research, especially in relation to long-living conifers (e.g., NANSON 1974 and 1989, BARADAT 1976, BARNES and SCHWAPPENHAUSER 1979, FRANKLIN 1979, LOO et al. 1984, KANG 1985, FOSTER 1986, GILL 1987, MCKEAND 1988, BURDON 1989, PHARIS et al. 1991, WHITE and HODGE 1991, HODGE and WHITE 1992, BALOCCHI et al. 1993, COSTA and DUREL 1996, BRIDGWATER and MCKEAND 1997, HANNRUP and EKBERG 1998, JANSSON 1998, WU 1998, GWAZE et al. 2000, JANSSON 2000).

Regrettably, the many advantages of early testing do not come free of charge. Above all, if indirect selection is to capture any genetic gain, the (early) measured traits must be sufficiently related to the mature traits that determine the value of final products. The degree of this association is measured by the genetic correlation (FALCONER 1981, WHITE and HODGE 1991).

Age-age genetic correlations (between expressions of any trait at different ages) are, as a rule, imperfect. On the other hand, it is intuitively apparent that the magnitude of the correlations is inversely related to the time interval between the two ages in question. The form of this relationship remained unspecified until the work by LAMBETH (1980) indicated that age-age correlations tend to be linearly related to the logarithmic ratio of the two ages. This empiric finding was an important step toward more reliable prediction and extrapolation of gains from early indirect selection. As the relationship is linear on a logarithmic scale, it follows that the predictive value of a trait expression improves in an exponential

manner with age. This is a corollary of the cumulative nature of tree growth (LAMBETH et al. 1983b). In accordance with the predictions of the Lambeth model (LAMBETH 1980), the many retrospective studies comparing characteristics of few-year-old nursery-grown seedlings and those on older trees grown in field trials have usually failed to show satisfactory correlation (NIENSTAEDT 1984, GREENWOOD and VOLKAERT 1992, JANSSON 1998, SONESSON et al. 2001) (for contrasting results, see, e.g., RIEMENSCHNEIDER 1988). The weak predictive value of seedling traits has led to a widely approved practice to defer the selection to at least a few years after trial establishment (WU 1988).

Encouragingly, moderately to high age-age correlations have occasionally been reported for conifer field trials (e.g., LAMBETH et al. 1983b, LOO et al. 1984). Many of the published correlations, however, do not involve true mature traits but successive expressions of a pre-mature (growth) trait – yet these estimates are often, confusingly, referred to as 'juvenile-mature correlations'. Despite the well-established theory of indirect selection (NANSON 1974, LAMBETH 1980, FALCONER 1981, BURDON 1989, WHITE AND HODGE 1991, JANSSON 2000), there is a striking shortage of factual evidence on the efficiency of early selection in terms of improved harvest-age productivity. Consequently, the genetic improvement of many long-lived tree species is based on an optimistic premise about adequate genetic correlations. For the most part, the lack of crucial information is due to the immature state of present-day breeding programmes insomuch as the ages of existing genetic trials are compared to the economic rotations of trees. For example, fewer than one percent of the Scots pine progeny-trial stands in Finland have reached *half* of the full rotation (about 80 years in southern Finland). Although the scarcity of data from mature trials does not permit the verification of hypotheses about indirect selection, it does not nullify the importance of early testing as an efficient tree-breeding tool. Evidently, however, the untested value of early differences underscores the need of careful planning and efficient implementation of the genetic field trials.

During the past ten years, much effort has been directed toward introducing tree breeding with molecular marker technologies (NEALE 2001). These novel methods are based on identification of polymorphic markers which show linkage to the traits of interest (WU et al. 2000), offering a promising shortcut to selection for late expressing traits. It is, however, unclear yet as to what degree techniques such as marker-assisted selection could be used to complement or even replace current field testing activities. The practical applications of molecular-based methods in the early genetic evaluation of forest trees are still few. Many practical constraints are known to exist in the way of their effective implementation, including the large size of the conifer genome, linkage disequilibrium, genotype-by-environ-

ment interactions, and large breeding materials (ARNOLD et al. 1990, BISOFFI 1993). It is thus quite possible that, when it comes to forest trees, artificial selection is not the most promising application of these techniques (STRAUSS et al. 1992).

3.3.2. Competitive interactions

Incoherent performance of test entries over time (and space) is one of the main problems with genetic evaluation. This does not necessarily impair statistical power, but makes the biological interpretation of test results more cumbersome, and may lead to faulty selections. In this category, the role of competitive interactions as a significant source of variation in tree-growth traits is well acknowledged (HÜHN 1970, ADAMS et al. 1973, CORRELL and ANDERSON 1983, TUSKAN and VAN BUJTENEN 1986, VON EULER et al. 1992).

Inter-genotypic competition can modify the effects of neighbouring plots and inflate experimental error (CORRELL and ANDERSON 1983). Characteristically, formerly positive correlations between adjacent trees (due to common microenvironment) may turn to neutral and finally, negative when competition for light, water and nutrients begins (KEMPTON 1982, MAGNUSSEN 1994). The increased use of small plots and dense spacing in field testing has apparently magnified these effects (FOSTER 1986, MAGNUSSEN and YEATMAN 1986). Single-tree plots minimize intrafamily competition and maximize interfamily competition, while the reverse is true for large multiple-tree plots. The net result of the single-tree plots is the inflation of among-family variance, which can severely distort genetic parameter estimates (MAGNUSSEN 1995).

Whether it is advisable to remove competition effects or to adjust to their presence depends on the way that the selected material is to be deployed (MAGNUSSEN 1989b). Suggestions have been presented for the elimination of competition effects by means of statistical adjustment of observations (KEMPTON 1982, KEMPTON and HOWES 1982, CORRELL and ANDERSON 1983, MAGNUSSEN 1989b) or by spatial separation of neighbouring plots with additional buffer rows (CORRELL and CELLIER 1987). The major obstacle is that competitive interactions among trees in a field trial are typically instantaneous and non-repeatable, depending on a multitude of factors such as the genotypic composition, spacing, stand age, and experimental design, and the trait in question (FOSTER 1986, FRANKLIN 1979, MAGNUSSEN and YEATMAN 1986, MAGNUSSEN 1989c). Thus, distinguishing competition effects reliably from the other sources of variation has proven very difficult. Furthermore, the competitive interactions witnessed during

field testing can be completely different from those occurring in production stands (VON EULER et al. 1992), which makes the value of mechanical adjustment procedures even more suspicious.

3.3.3. Genotype-by-environment interaction

The inability of genotypes to maintain the same relative performance over several environments and management practices adds to the list of factors that make the analysis of genetic field testing data intricate. This phenomenon, termed as genotype-by-environment interaction (GEI), is commonly interpreted to reflect a disruptive adaptative process which results in genotypes showing different optima in their response to various properties of the environment (KNIGHT 1969). GEI effects are most evident in association with species and provenance trials of trees, but interactions have also been frequently found in situations where genetic differences are less distinct, as in progeny testing (MATHESON & COTTERILL 1990).

In tree breeding, two key strategies, one aimed at genotypes and the other at environments, exist with regard to biologically significant GEI. The former approach consists of distinguishing between those genotypes that retain their relative performance over a wide range of conditions ('generalists') and the ones which excel in few environments but under-perform in the rest ('specialists'), and deploying each genotype according to its verified reaction norm. The complexity of the models used to elucidate GEI varies greatly (SKRØPPA 1984). Rather than explaining GEI directly by the underlying environmental factors, the interactive performance is usually described statistically. Regression analysis, wherein the average (logarithmic) yields obtained in different environments are regressed to those of a genotype (FINLAY and WILKINSON 1963), is one of the most common techniques to determine the interactive behaviour of genotypes. This approach is workable when there are a relatively small number of genotypes or varieties to be characterised. In plant breeding, such situation is fairly common. Tree breeding populations, however, can involve hundreds or thousands of candidates, the detailed characterisation of which would be an unreasonably costly effort. In contrast, the variation encountered in the prominent macro-environmental variables in forestry (temperature, moisture, soil quality) can usually be sampled by relatively few field trials. In the context of breeding, it is apparently more advantageous to focus on the properties of sites rather than genotypes, and determine the contribution of each type of site to the realised GEI (BURDON 1977). A matrix of between-site genetic correlations can be examined to distinguish the most unstable testing environments. This should preferrably be followed by an analysis

of contributing factors, proceeding to characterisation and division of the deployment environment into several well-defined and homogeneous subsets, within which the magnitude of GEI is practically non-significant (BURDON 1977, SKRØPPA 1984). Such subsets (geographic-elevational breeding zones) are essential to species that naturally inhabit geographically large areas (JOHNSON 1997).

For genetic testing, the most problematic kind of GEI is the one in which the performance of the genotypes is affected by some specific characteristics of genetic trials that are not representative of the deployment conditions (**III, IV**). The decreased genetic correlation between selection and deployment environments would reduce predictive ability and might lead to selection of inferior or maladapted genotypes for breeding and production populations. The increased popularity of field testing in near-ideal conditions, for instance, has raised some suspicions about the possibly increased risk of faulty selections (HODGE & WHITE 1992, WU et al. 1997). This particular question was addressed in this study (**III**), paying special attention to the evaluation of the correlation for Finnish test orchard trials (MIKOLA 1985).

3.4. Progeny testing of Scots pine in Finland

3.4.1. General information

Genetic field testing has been the most prominent activity of Finnish tree breeding for several decades. As a manifestation of this, the national forest genetics register today incorporates records on approximately 3700 field trials (YRJÄNÄ et al. 2000). The first (provenance) trials were commenced already in 1931. Most of the tree improvement efforts have traditionally been devoted to Scots pine, the species of greatest economic significance for Finland's forestry and forest industry. Scots pine presently accounts for 47 % of the growing stock volume and is used on 49 % of the hectares annually reforested in Finland (FINNISH STATISTICAL YEARBOOK OF FORESTRY 2001).

The major undertaking in Scots pine breeding has involved the progeny testing of more than 7000 plus-trees selected from native stands since the late 1940's (SARVAS 1953). Many of these trees were subsequently grafted in seed production orchards without any direct evidence of their genetic merits. Sorting this large base material in genetic terms was a natural follow-up to the initial phenotypic selection. The commencement of an extensive progeny testing followed the inau-

guration of the first 10-year nationwide breeding program in 1967 (Fig. 1). Following this turning point, over six million seedlings have been planted in nearly 1400 Scots pine progeny trials covering altogether over 2000 hectares of land (YRJÄNÄ et al. 2000). Up till now, the 1400 progeny trials of Scots pine have been measured about 2500 times. The data are mainly composed of observations of tree height, which is the foremost (in 2390 data sets) and typically the only variate assessed in juvenile progeny trials. Most measurements (~1350 data sets) were conducted at ages of 5, 10, and 15 years (counted from the establishment year) (Note: The statistics presented in this and the next two chapters on genetic field testing were assembled from the databases of the forest genetics register, maintained by the Finnish Forest Research Institute).

While the evaluation of the genetic potential of plus-tree parents has remained as the most important function of the ongoing progeny trials, these trials are also used for other purposes. The top 50 of the progeny-tested parent trees are to form the nuclei of the first-generation breeding populations, the assemblage of which is now underway. Estimates of parental breeding values are also used to assemble material for "1.5 generation" production seed orchards, and to rogue the most poorly performing clones from existing seed orchards. Furthermore, the offspring grown in the present progeny trials constitute a base population from which outstanding individuals are being selected to form the main groups of breeding populations (250 trees per population) (HAAPANEN et al. 1999).

3.4.2. Testing effort

Many of the features of the Finnish progeny trials of Scots pine are beneficial for accurate parental ranking and estimation of genetic parameters. These include simple designs, fairly large numbers of families, replications over multiple test sites and the common use of open-pollinated seed material (roughly 90% of the families). In general, the trials are relatively consistent with regard to the experimental implementation. With very few exceptions, all the progeny trials use a randomized complete-block design (some single-tree-plot trials have been established in a completedly randomized manner). Families are usually replicated in 4 to 6 blocks. Another feature characteristic to the Finnish trials is the widespread use of large multiple-tree plots, typically accommodating 25 trees planted in a square (5-by-5) pattern (**I**, Fig. 1). The 25-tree plot configuration is currently practiced in 768 (55%) of the progeny trials. Square plots with other dimensions, such as 3-by-3, 4-by-4 and 6-by-6 patterns, are also fairly common (used in 75, 143, and 55 trials, respectively). The number of single-tree-plot trials is 133. The

incidences of other types of plot configuration, such as non-contiguous plots (LIBBY and COCKERHAM 1980), are few.

One consequence of the use of large plots is the sizeable number of offspring tested per family. In an average Scots pine progeny trial, the family size exceeds 150. The total effort per parent is even greater as each trial is ordinarily replicated on two to five sites. Furthermore, many candidates have their offspring replicated in more than just one series of trials. A typical plus-tree is represented in six trials (the median). However, the distribution of parental contribution to testing is highly skewed. There are some plus-trees that have over 20000 individual offspring (the median is 495) planted in more than 200 trials. On the other hand, many plus-trees do not have any of their progeny tested (cf. LINDGREN (1991), who suggested that a suitable family size for Scots pine progeny testing should consist of 15 to 50 offspring).

3.4.3. Testing methods

The traditional progeny trials in Finland mimic ordinary pine plantations in terms of site quality, spacing and management. These experiments are referred to as *forestry trials* in this study. Another numerous group of progeny trials constitutes *test orchards* (MIKOLA 1985). The concept of a test orchard is principally synonymous with 'farm-field trial', yet appreciably many of the trials in this category have been placed on forest sites. The major distinctive features of the two methods are the spacing and the choice of site type: Forestry trials are customarily planted on ordinary forest sites at an initial density of 2000–2500 trees per hectare. They are also managed along with the same routines as ordinary Scots pine plantations. Test orchards, in turn, are regularly established as high density stands (8000–10000 trees per hectare) on as homogeneous sites as possible. The method also involves intensive site preparation before outplanting, mechanical and chemical control of competing vegetation, and fencing in order to avoid damage due to moose. All these measures aim at reducing the experimental error and enhancing early manifestation of genetic differences (MIKOLA 1985).

The first test orchards in Finland were established in the mid-1970's. Since the relatively late startup, the method has gained increasing popularity. Today roughly a third of the ongoing Scots pine progeny trials are test orchards. Most of the new field trials in Finland are also designed as test orchards (Fig. 1).

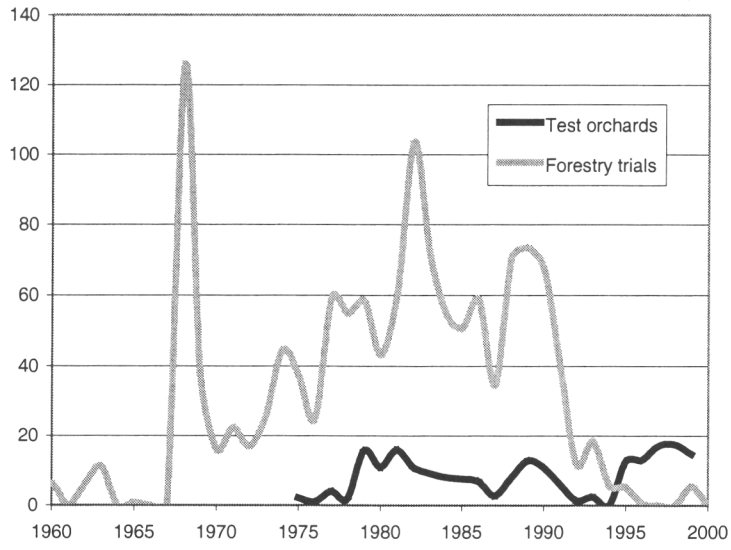

FIGURE 1. The area of land (ha) used annually in forestry and test orchard trials employed for Scots pine progeny testing in Finland between years 1955 and 2000 (The source: Forest genetics register, The Finnish Forest Research Institute).

4. Material And Methods

4.1. General background

4.1.1. Material

This study made use of tree-height data from measurements conducted in progeny trials established in the southern and central parts of Finland (between 60th and 65th latitudes) between the years 1977 and 1983. The ages of the trials at the time of measurement ranged from 4 through 20 years.

The test materials were predominantly open-pollinated offspring of Scots pine plus trees. The seed for the entries had been collected from young clonal archives and seed orchards. Because of free pollination, the pollen parents were assumed to consist of a random sample of unrelated genotypes from the unimproved Scots pine population of the region (southern Finland), i.e., the same population from which the plus-trees were a select sample. Other types of entries, such as 'standard' check-lots and occasional full-sib families, were consistently omitted from the analyses.

All of the trials examined were planned by the Finnish Forest Research Institute, and laid out in a randomised complete block design, which, in most cases, involved multiple-tree (25 most common) plots and four to six block replicates. The number of families varied from trial to trial, but usually from 30 to 60 families per trial.

4.1.2. Analyses

Partitioning of the total variance into independent components was a central course of action throughout the study. The factors into which the variance was attributed were defined by a linear model that was consistent with the design structure of trial. In the standard case of the randomised complete block design with multiple-tree plots and observations on individual trees, the model distinguished family and block effects, their interaction ('plot effect') and the residual part (within-plot deviation). When the analysis involved multiple parallel trials, the underlying model was adjusted to incorporate the effect due to family-by-site interaction (**III, IV**). The component for the among-block variation was usually omitted (or block effects were treated as fixed), as these differences were typically not of any significance to the research questions addressed. The variance-

component analyses were based on the REML (Restricted Maximum Likelihood) method (PATTERSON and THOMPSON 1971) provided by the MIXED procedure of the SAS statistical software package (LITTELL et al. 1996). In translating the statistical variance components into their genetic counterparts, the open-pollinated progenies were assumed to consist of paternally unrelated half-sibs which, on average, have a fourth of their genes in common (FALCONER 1981).

4.2. Individual research papers

4.2.1. Options for greater efficiency (I, II)

The two parallel field trials examined in the plot–size study (**I**) were exceptional since they did not consist of a family structure. The 15 families listed in official trial documents were actually pseudo-entries drawn from a single open-pollinated seed–lot. The most likely reason for concealing the true (lack of) design was a desire to check whether the third-party organisation responsible for the installation of trials in the early 1970's was using an appropriate method of randomisation when assigning test entries to experimental units (VELLING, pers. comm.). Aside from the original purpose of these trials, their genetically homogeneus structure made them particularly useful for studying the patterns of field variation and determine optimal plot sizes in the way of the agricultural research tradition of the early 1900's (SMITH 1938). In forest trees, earlier attempts to apply 'uniformity trial' data to study these issues are few (WRIGHT and FREELAND 1960, CONKLE 1963). Compared to inbred crop plants employed in agricultural plot size studies, open-pollinated trees evidently contain plenty of genetic variation. However, this was not regarded as a significant weakness since 1) the conclusions drawn were founded on variance ratios, not the sizes of absolute variances, and 2), the genotypes were randomly dispersed over the sites, thus excluding any systematic confounding between genetic and environmental effects.

In the first stage of data processing, the row-column coordinates of each tree (fixed on the basis of the known order of measurement) and the respective observations of tree height were merged to create a virtual map onto which 11 alternative plot types were superimposed. The plots examined covered a wide range of sizes, from single-tree plots to large square plots of 49 trees, thus encompassing the ones most commonly used in Finnish progeny testing. Furthermore, three different types of non-contiguous plots (containing 8, 16, and 49 trees) were included for comparison. In the non-contiguous arrangement, the multiple trees of a plot are randomly distributed throughout the block (LIBBY and COCKERHAM 1980).

Each of the computer-generated plot designs was subjected to a variance component analysis in which the total variance was partitioned into among-plot and within-plot components. These components were used to derive an estimates of the environmental portion of family variance and the size of trial needed to detect a difference between two family means as statistically significant (**I**).

Environmental heterogeneity was modelled using the response of the plot-mean variance to the plot size. According to the 'empirical law' discovered by SMITH (1938), these two variables have the following relationship:

$$V_n = V_1 \; n^{-b}$$

where V_1 and V_n denote the variance of plot means of unit and n-sized plots, respectively, and b is a parameter reflecting the site heterogeneity. When expressed on a log-log scale, this relationship may be written and modelled in terms of a simple linear regression

$$\log(V_n) = \log(V_1) - b \log(n)$$

The formula implies that for each relative increase in plot size, there is a corresponding relative reduction in the plot-mean variance. The coefficient of regression *b* takes values from zero to one, and may be viewed as an index of the degree of spatial correlation; a coefficient approaching zero reflects a high correlation between adjacent experimental units while *b* nearing one implies that the neighbouring units are uncorrelated.

The within-plot subsampling study (**II**) made use of 33 complete sets of tree-height data (all the living trees recorded) from as many open-pollinated progeny trials. Thirty of the trials included had been laid out with conventional multiple-tree plots (mostly 25 trees). Three trials employed non-contiguous plots comprising 8 and 12 trees.

To simulate the effects of partial measurement, random subsamples of varying numbers of tree-records (2, 4, 6, 8, 10, 15 and 20) were subsampled from within each plot using a specifically designed computer program. The procedure was repeated 10 to 15 times for each sampling scheme. The resulting subsets of data were subjected to the analyses in which the contributions of the among-plot and within-plot effects (family and block effects eliminated) to the total variance were determined. Estimates of these two residual variance components are needed to construct the sampling variance of a family mean. This statistic is closely associ-

ated with the three criteria used to gauge the statistical effects of subsampling: The smallest statistically significant difference between a pair of family means, family-mean heritability, and the correlation between family mean estimates from the complete sample and a subsample.

4.2.2. Quality of information and selection efficiency (III, IV)

The last two research papers addressed the various qualities of progeny testing data, focusing on two central subjects: the importance of genotype-by-environment interaction (GEI) in Scots pine progeny trials (**III**) and the development of genetic parameters and indirect response to selection over time (**IV**). The measurement data used in these studies were from 30 and 26 progeny trials, respectively. The data for **III** involved height measurements made in each trial at 10 years of age. The latter study comprised 82 successively measured data, obtained at inconsistent intervals from age 5 through age 18 years.

One of the main goals of these two studies was to compare the two main approaches to genetic field testing in Finland – forestry trials and test orchard trials. Therefore, the material for analyses was chosen from among several independent trial series. Each of the series included several trial replicates, comprising a common set of open-pollinated families. All the series encompassed at least one trial representing each testing method.

The variance (and covariance) component estimates enabled calculation of various quantitative-genetic parameters that are essential for predicting genetic gain from selection (**III, IV**). These parameters comprised, e.g., individual narrow-sense heritability, family and within-family heritability, genetic and phenotypic correlations between ages and trials, plus the coefficients of family-mean and additive genetic variation.

Genotype-by-environment interaction (GEI) in Scots pine progeny trials was studied (**III**) by using an approach in which tree height measured on the same families in two replicated trials is interpreted as two distinct traits, and the genetic (type-B) correlation between these traits is estimated (DICKERSON 1962, BURDON 1977). The type-B correlations were calculated in two ways: using heritabilities of family means in the respective sites and their phenotypic correlation across the sites (BURDON 1977), and by means of the family and family-by-site components of variance from an across-site variance component analysis (YAMADA 1962). Environmental and testing method differences between each pair of parallel trials, involving differences in planting density, mean height, and survival percentage,

were subsequently examined to find explanations for differing site-site genetic correlations between pairs of trials.

Responses to both backward (parental) and forward (based on within-family deviations) selection were predicted using parameters estimated for four groups of progeny trials. The grouping was based on two discriminating variables: initial density (narrow vs. wide spacing) and site type (field site vs. forest site) (**IV**). Using predictions of genetic parameters from the time-trend models, selection efficiencies were calculated for all ages from 5 to 20 years. All the genetic gains were anchored to height at age 20, which was deliberately chosen as the target trait for improvement. The time loss associated with postponed selection was accounted for by dividing the predicted selection response by the number of years used to the testing phase and the subsequent crossing of selected genotypes (breeding phase).

To forecast genetic response to early selection, knowledge of the strength of the genetic relationship between the selection and target-age performances is required. In this study (**IV**), trends in age-age correlations were estimated for the full range of ages examined (5 to 20 years) by employing a method based on "the log of the age ratio" (LAMBETH 1980). The variance component estimates from the 82 measurement data sets (**IV**) were also analysed to determine age-related changes of the respective parameters. To explain variation in these estimates, a mixed repeated-measures model (LITTELL et al. 1996) was developed which accounted for age, spacing, site quality, and the interactions of these factors. To eradicate the exponential relationship between age and the variance components, logarithms of both of these variables were taken prior to the analysis. The main outcome of the analysis was a time-trend function which, after statistically insignificant factors were eliminated, allowed to predict the response variable (variance component, heritability or the coefficient of genetic variation) in question in the varying environmental conditions over the first two decades of testing. Finally, the age and site-specific predictions of the different parameters were used as input values for genetic gain formulae.

5. Results and Discussion

5.1. Options for greater efficiency (I, II)

5.1.1. Plot size

The simulations with 11 types of experimental plots (row, rectangular, square, non-contiguous) in the two uniformity trials of Scots pine (**I**) showed that the efficiency is tremendously affected by the size of experimental unit. There was a clear inverse relationship between the plot size and the amount of information for a fixed number of trees tested. The single-tree plots excelled in the comparison, providing by far the best resolution of family means (judged in terms of the environmental portion of the variance of family means). A minor increment in the number of trees comprising the plot resulted in a substantial decline in the statistical efficiency. The relative efficiencies (keeping the efficiency of the single-tree plots as the standard) of the 3-tree row plots in the two trials were only 64% and 81%. This result agreed with WRIGHT and FREELAND (1960) who found similar plots of *Pinus resinosa* to yield 20% to 40% less information per tree than single-tree plots. The plot sizes ranging from 20 to 30 trees were roughly 30% to 40% as efficient as the single-tree-plots (**I**, Fig. 2). The largest plots, comprising 49 trees (7-by-7), extensively added to the family-mean variance and were thus characterised by exceptionally poor relative performance (the relative efficiencies in the two sites were 16% and 31%).

The relative size of test material required to detect a 5% differences between family means, was used as a parallel criterion of experimental efficiency. Based on the same variance components as the calculation of the family-mean variance, the results were in close agreement (the required test size increasing with the variance). The 49-tree plots were found to assume a three to six times larger area to reach the same degree of family-level discrimination as was obtained by the single-tree plots.

The among-plot and within-plot components of variance were fairly insensitive to changes in the number of trees per plot (excluding the single-tree plots) (**I**). *For a fixed number of plots* (replications), larger plots thus give a higher precision than smaller plots. The notion of diminishing variation of plot means as a function of increased plot size was the rationale of some proposals for fairly large optimal number of trees per plot (e.g., EVANS et al. 1961). This reasoning is, however, misleading as it does not take into account the concurrent effects of plot size on

the experiment-wise error control under a given total number of trees, or the financial considerations involved in the plot layout (SMITH 1938).

The substantial loss of information due to increasing plot size, demonstrated by the present (**I**) and a number of previous studies (WRIGHT and FREELAND 1960, CONKLE 1963, LAMBETH et al. 1983a, LOO-DINKINS & TAUER 1987, LOO-DINKINS ET AL. 1990), underlines the importance of optimal allocation of limited resources. For a fixed number of trees available for testing, the number of trees per plot may only be raised at the cost of fewer experimental units. The latter number is, in general, much more critical to the statistical precision of the experiment (COX 1958). Allotting trees to more sizeable experimental units can thus markedly impair the estimation of genetic effects. Conversely, the smallest plots sample the maximum number of micro-sites with a given number of trees, giving the most detailed information for each tree (WRIGHT and FREELAND 1960). Most importantly, single-tree plots ensure the highest degree of local control by minimising the size occupied by each block. A similar effect may naturally be achieved by applying closer initial spacing (**IV**). As decisions on the plot size and stand density are mutually independent, the experimental efficiency can be impoved in the both ways.

5.1.2. Pros and cons of small vs. large experimental plots

The thread of discussion on the most appropriate type of plot for forest-genetic field testing is lengthy. Despite the well recognized statistical advantage of small, much replicated plots (SMITH 1938, WRIGHT and FREELAND 1960, CONKLE 1963, JOHNSTONE and SAMUEL 1974, WRIGHT 1976, COTTERILL and JAMES 1984, ZOBEL and TALBERT 1984, LOO-DINKINS & TAUER 1987), the use of such arrangements, particularly single-tree plots, is not as widespread as one might expect. In North America, field trials commonly employ plots of four to ten trees (WRIGHT 1976, ADAMS et al. 1994, LOO-DINKINS et al. 1987, LOO-DINKINS & TAUER 1990), whereas the Finnish tradition has mainly involved pure contiguous plots of 25 or more trees (**I**). In Sweden, single-tree plots are the prevailing standard (LINDGREN 1991).

In the past, large multiple-tree plots were often considered to be necessary in order to reliably estimate wood production on a per unit area basis (JOHNSSON 1974). Although tree volume is normally one of the goal traits, this argument can hardly bear critical inspection. Firstly, the primary purpose of genetic testing should be in assessing relative performance differences rather than absolute measures of production capacity. The latter information is, in general, of little use

for a tree geneticist. Secondly, reliable assessments of stand volume without disturbing edge effects would most likely require impractically large-sized experimental units and unbearably long testing periods, effectively deteriorating the statistical power of a trial. It is, in fact, possible that whilst the regular 25-tree plots of the Finnish progeny trials are too large to be statistically efficient, they may concurrently be too small to produce trustworthy information on long-term per-hectare production. Even if such information is eventually attained, its usefulness can be questioned since the deployment of the material in pure family plots differs from that in ordinary plantations. To conclude, the estimation of stand volume at harvest should be decisively isolated from other, more relevant objectives of genetic testing. These include ascertaining whether the variates selected for are sufficiently related to the breeding targets (which include production traits). In contrast to some suspicions, large multiunit plots may not be the only source of such variates. In a recent Scots pine study by JANSSON et al. (1998), single-tree plots were found to be strongly genetically correlated ($r_G \sim 0.8$) to estimates of the per unit area volume production obtained from 36-tree plots.

The biological and statistical validity of single-tree plots has been frequently questioned in the tree breeding literature (SHIUE and PAULEY 1960, EVANS et al. 1961, JOHNSSON 1974, WOOLLONS 1980), which has probably, at least to some extent, discouraged tree breeders from applying small-sized experimental units in their field trials. Tree breeders have especially been concerned with implementing single-tree plot designs due to their vulnerability to natural mortality and thinnings, leading to complications in data analysis (e.g., WOOLLONS 1980). In the early days of tree breeding, such fears were quite reasonable as there were no good solutions for handling unorthogonally structured data sets. Such data are obviously commonplace with single-tree-plot trials since, by the nature of the design, the death of any single tree yields a missing plot record (ZOBEL and TALBERT 1984). Only some decades ago, laborious calculations were required to replace missing observations with approximations (WRIGHT and FREELAND 1960). In addition, complete families were occasionally discarded from the analyses because of a few missing plots (MATZIRIS and ZOBEL 1973). Although severe imbalance still puts inferences from an experiment in danger, the evolution of statistical software packages capable of effectively extracting informaton from even highly fragmented data sets, and simultaneous advances in statistical methods (especially variance component estimation), have greatly alleviated the situation (MCCUTCHAN et al. 1985, LINDGREN 1991). In any case, the avoidance of early mortality is obviously much more of an issue with single-tree plots than with their multiunit counterparts.

Doubts have also been cast about the ability of single-tree plots to meet the regular assumptions of normality and homogeneity of residuals (SHIUE and PAULEY 1960, EVANS et al. 1961) underlying the analysis of variance and more generally linear models. Such violations have more recently proved to be less serious than previously suspected (FRANKLIN 1971, JANSSON and DANELL 1993). In most cases, ordinary transformations (particularly logarithmic and square-root) of the response variable suffice to restore the normal distribution and to homogenise the residual variance (JANSSON and DANELL 1993).

Beyond doubt, single-tree plots, non-contiguous plots, and to a lesser extent, row plots, let trees of different genetic identities interact in more diverse ways than large block plots. This has raised speculations as to possible disagreement between family rankings based on small vs. large plots (LAMBETH et al. 1983a). Clearly, if such discrepancy existed, it would seriously complicate selection. To the relief of tree breeders, several studies with different tree species have shown that changes in genotype rankings as a response to competitive stress are occasional and usually not appreciable (e.g., CORRELL and ANDERSON 1983, FOSTER 1989, MAGNUSSEN 1989b, TUSKAN and WILLIAMS 1989, ST. CLAIR and ADAMS 1991, VON EULER 1993). Crown competition seems to be mainly driven by size differences and stand density and the role of the inter-genotypic component is only a minor one (MAGNUSSEN 1993).

Whereas family means and ranks seem relatively insensitive to competition, the same does not apply to estimates of genetic variance and covariance (and thus the derived estimates of heritability and genetic correlation). In particular for tree-volume related traits, estimates of genetic variance from stands influenced by strong competition can be greatly overestimated (HAMBLIN and ROSIELLE 1978, EULER et al. 1992, MAGNUSSEN 1989a, 1989b, 1989c, 1993, 1994). Naturally, such estimates should not be extrapolated to conditions where competition is less severe. Inflated genetic variances from single-tree-plot trials under mild competition may also be rectified to their large plot expectations by using procedures outlined by MAGNUSSEN (1989, 1989b).

In many situations it might be rational to see moderate intergenotypic competition as a passable, and, perhaps, necessary condition of genetic field testing. In Finland, Scots pine plantations are established with open-pollinated bulk seed from seed orchards or seed collection stands. Logically, genotypes deployed as mixtures should be able to contend with scarce resources under at least mild levels of interference. As large monoculture-like plots provide significantly less 'natural' interfering among genotypes than smaller plots, the former ones might, in fact, be more doubtful when the primarily focus is on accurate ranking of parents. In

ordinary stands, on the other hand, inter-tree competition never gets very intense as it is counteracted by means of repeated thinnings. Obviously, duplicating the stand dynamics of ordinary plantations by field trials is a very difficult task (e.g., VON EULER et al. 1992).

When other objectives are involved or even predominate the testing, the choice between plot types becomes more intricate and the criteria emphasizing precise parental evaluation may no longer be appropriate. This concerns, for instance, the situation where individual trees growing in a trial are selected as candidates for the next round of breeding on the basis of their ranks within the family. Small plots may fit poorly to such a task. Forward selection calls for effective sib-comparisons that are best achieved in the common neighborhood offered by multiple-tree plots. This is apparently one of the few situations where sacrificing the general statistical efficacy of small plots might be well-grounded. As a solution to the dilemma which arises from the conflicting demands for plot size, it might be wise to relinguish on the idea of designing trials that would optimally meet all the multiple functions of genetic testing. Single-tree plots could thus be favored in trials aimed at parental selection, and, conversely, multiple-tree plots for selecting among family members in a common environment (LAMBETH 1986, LOO-DINKINS et al. 1990).

The use of large plots might be justifiable also when large differences in growth rate exist among the entries being tested (ZOBEL and TALBERT 1984). Moreover, if the experimental design consists of additional factors, such as fertiliser treatments, there is reasonably some constraint to the minimum size of experimental units to prevent the confounding of treatment effects of neighbouring units (WOOLLONS 1980). Another case for large plots may arise when the total costs (establishment, management, measurement, etc.) per plot are manifold as compared to those per tree. This is evident from the results of WRIGHT and FREELAND (1960) and CONKLE (1963) who applied a formula devised by SMITH (1938) to determine the most cost-effective plot size. Both of the studies found plots of one to four trees to be generally the most economical ones. In the cases where the expected per-plot to per-tree cost ratio was set exceptionally high, a larger optimum was obtained. Unfortunately, this method is very sensitive to estimates of b (the coefficient of field hereogeneity) and to cost parameters involved, all of which are quite difficult to accurately estimate.

5.1.2. Plot shape

The influence of varying plot shape on the efficiency was clearly less distinguished than that of plot size (**I**). Nevertheless, rectangular or row plots appeared to control field variation slightly more effectively than ordinary square plots. In theory, the shape and orientation of plots should matter most when the field variability is more pronounced in one direction than in the other (SMITH 1938, ZOBEL and TALBERT 1984). The plot-shape effects were probably underestimated in this study since there was no attempt to align the simulated row and rectangular plots in a direction where they would have sampled most of the field variation. On the other hand, deliberate *a posteriori* orientating of plots in the direction of the greatest efficiency might have given the wrong idea about the importance of plot shape, since in reality, trends in environmental variability are not easy to detect.

5.1.2. Non-contiguous plots

The results for the non-contiguous plots were in striking contrast to the overall negative relationship between efficiency and plot size found for the contiguous plots (**I, II**). The overall efficiency of the three non-contiguous plots (of 8, 16, and 49 trees) was high and nearly independent of the number of trees comprising the plot. The non-contiguous plots outdid the contiguous ones throughout the range of plot sizes. Their superiority was mainly due to the nearly complete eradication of among-plot differences (which contribute to the 'family-by-block' mean square in the usual two-way analysis of variance for the RCB design) as a consequence of the plot trees being randomly spread across the block instead of being planted by groups as normal. LAMBETH et al. (1983), who compared non-contiguous plots to row plots in eight *Pinus taeda* L. genetic tests, drew a similar conclusion: the non-contiguous plots were very effective in eliminating the family-by-block interaction variance, yet there were no demonstrable differences between the two plot types for the other sources of variance.

The non-contiguous multiple-tree plots were only slightly less efficient than the single-tree plots (**I**). In addition, the former plot type has some important practical advantages over the latter system. Most importantly, non-contiguous-plot trials are intrinsically robust to the effects of mortality (LIBBY and COCKERHAM 1980). The non-contiguous-plot configuration can embrace an additional design stratum, interlocking replicates within each block. In systematical thinnings, these can be entirely removed without destroying the orthogonal structure of the design (LIBBY and COCKERHAM 1980, MIKOLA 1993).

5.1.2. The pattern of environmental variation

The nature of field heterogeneity was assessed using a measure devised by SMITH (1938), which relates the size and the variability of plots in terms of linear regression. Although the approach lacks a solid theoretical basis (PEARCE 1976), it has proven to be broadly applicable to a number of species and conditions. Forest trees may not be an exception: the regression between the two above-mentioned variables was strong in the two sites studied (**I**). The slopes of regression (*b*) for the two sites were approximately 0.4 and 0.5, indicating predominantly systematic, large-grained type of heterogeneity. For forest trees, comparable estimates are in short supply. WRIGHT and FREELAND (1960) reported regression slopes for six (mostly *Pinus resinosa*) stands to range from 0.57 to 0.91. In CORRELL's (1978) study, the corresponding estimates for three radiata pine (*Pinus radiata)* trials ranged from 0.33 to 0.66.

The pattern of field heterogeneity is related to the expected relative efficiencies of different plot configurations (SMITH 1938, LOO-DINKINS and TAUER 1987). The choice of plot type is most critical for sites characterised by strong spatial correlations (small *b* values). For both uniformity trials (**I**), there was a general tendency for decreasing efficiency with increasing plot size. However, the relationship was more pronounced for the trial with the smoother slope of regression (**I,** Fig 1, 2). This emphasizes the importance of choosing uniform sites for field testing, and the inadequacy of large plots on sites which have a high degree of spatial variation.

The applicability of SMITH's (1938) regression method has sometimes been assumed to be restricted to uniformity trials (CORRELL 1978). KOCH and RIGNEY (1951) showed, however, that *b* could also be estimated from specially structured data (e.g., lattice and split-plot designs). Such estimates would allow a better general view on the nature of soil heterogeneity in tree breeding trials established on various types of soil, and to determine the economically optimal plot size (SMITH 1938).

In recent years, there has been a growing interest in eliminating local irregularities in tree-breeding trials by means of nearest neighbour models, trend surface analyses or post-blocking (BONGARTEN and DOWD 1987, THOMSON and EL-KASSABY 1988, MAGNUSSEN 1994, ERICSSON 1997). The emergence of new procedures and advances in computing power now enable environmental trends and correlations between neighbouring plots to be accounted for in analysis. Nonetheless, even the most sophisticated techniques can not substitute for careful planning and site preparation, which still remain as the most effective means of

reducing the biasing effects due to spatial correlation and the need for (sometimes doubtful) statistical corrections (MAGNUSSEN 1990).

5.1.3. Partial measurement of multiple-tree plots

In experimental design terminology, plots are referred to as experimental units that are sampled *n* times (sampling units) to obtain the plot response (STEEL et al. 1997). In data processing, the unit of interest is normally the plot mean rather than the individual-tree record (WILLIAMS and MATHESON 1994).

In routine measurements of genetic field trials, plots are sampled in their entirety, obtaining one record from each tree. However, complete sampling may not always be the optimal way of collecting data, especially if the test materials are arranged in contiguous plots. Firstly, there is the law of diminishing returns which implies that increasing the number of measurements eventually discontinues improving the statistical precision. A rule of thumb, applicable to RCB design, states that there is not much increase in the power of the experiment when *n* exceeds roughly four times the ratio of the within-unit variance component to the among-unit variance component (COX 1958, BERGERUD 1995). Secondly, as uniformity trials (SMITH 1938) have demonstrated, individuals growing in contiguous plots are generally more alike in their productivity (or any other environmentally influenced trait) than might be expected on the grounds of their common genetic background. Owing to the environmental correlation, trees within contiguous plots fail to provide independent information. As a consequence, the amount of information gained per tree diminishes with the increasing number of trees per plot. Therefore, there is a clear potential for improved cost-efficiency by taking records on fewer sampling units than potentially available.

The sample-size simulations for 30 Scots pine progeny trials (**II**) convincingly showed that maximising the number of spatially correlated sampling units gives little advantage in terms of precision. Instead, it led to poorer efficiency per tree measured. The studied parameters (least significant difference, family heritability and family-rank correlation) responded to subsampling most when the number of trees included in the random sample was raised from two to six. The results for sample sizes of 10 to 15 measured trees per plot were basically similar to those obtained by the doubly laborious measurement of all 25 trees. Family rankings were consistent with even less extensive sampling. The results suggest parental selection might be carried out with reasonable accuracy on as few as four to six trees (of a 25-tree plot).

A few other studies focusing this topic have also found subsampling useful. STEVENSON and SAVILL (1976) sampled trees from 36-tree plots in two Sitka spruce field trials, and found that an acceptable precision was in most cases achieved by a sample size of 12 to 16 trees. The effect of including additional buffer rows in multiple-tree plots and neglecting the buffer trees in measurements was studied by CORRELL and CELLIER (1987) who found that little information was lost when only the inner trees of an 8-by-8 block plot were measured. WRIGHT (1970) proposed, for normally distributed variates, measuring only the tallest tree in a multiple-tree plot. In a nursery study by LEE (1974), this shortcut method saved 90% in measurement effort while losing only 4% of the information. APIOLAZA et al. (1999) studied the effects of varying subsampling (ignoring the plot stratum) intensities on genetic parameter estimation, family ranking and estimates of genetic gain, with simulation data comprising 200 families with 30 individuals in each. They found little improvement in most of the parameters after the sample size (number of trees per family) was raised to 15.

As with the plot size considerations, these results are also subject to reservations if the measurement data are not used to assess the average family performance. More extensive sampling or systematic (visual) selection of the tallest trees could be profitable, for example, when the objective of the assessment is primarily to get information on the highest yielding individuals within each family.

5.2. Quality of information and selection efficiency (III, IV)

5.2.1. Importance of genotype-by-environment interaction

The concept of type-B genetic correlation (r_B) (correlation of additive genetic values between two sites) was employed in this study (**III**, **IV**) to assess the stability of family performance on tree height over a range of conditions representative of the Scots pine progeny testing in Finland. In theory, this parameter can take values from zero to one, smaller values being indicative of stronger genotype-environment interaction (GEI) (BURDON 1977). Family rank changes between parallel progeny trials were also examined (**III**).

Different analyses yielded fairly consistent information on the magnitude of genotype-by-environment interaction in Finnish progeny trials: $r_B = 0.58$ (**III**, the mean of 8 multiple-site analyses), $r_B = 0.61$ (**III**, the mean of 39 paired-trial analyses), and $r_B = 0.67$ (**IV**, 22 multiple-site analyses). These average levels signify a degree of inconsistence in family performance which definitely needs to be taken into account in breeding (SHELBOURNE 1972), especially as a significant

portion of the interactions were of the crossover type, i.e., associated with changes in family rank. As such, however, statistical interactions between genetically close entries, such as the plus tree progenies in this study, and experimental sites or even with blocks within a single trial, are not uncommon (JOHNSON and BURDON 1990, MATHESON And COTTERILL 1990, PEDERICK 1990). The published estimates of r_B mostly conform to the those obtained in **III** and **IV**, falling between 0.5 and 0.8 (e.g., HODGE and WHITE 1992, DIETERS et al. 1995, JOHNSON et al. 1997). The last two studies reported increasing type-B correlations with time, indicating the diminishing importance of GEI. No tendency in r_B was observed in the present data (**IV**), suggesting that the magnitude of GEI in Scots pine trials remained more or less unaltered over the first 20 years of progeny testing.

To gain from GEI (by regionalising breeding programmes or by applying unequal weights for data from different types of trials), it is essential to define the particular underlying components of the environment (BURDON 1977). BARNES et al. (1994) concluded that the chances of exploiting interactions in breeding are good only when GEI can be attributed to a single environmental factor with a pronounced effect. HODGE and WHITE (1992) and JANSSON (1998) found parallel trials on sites of distinct fertility to be more weakly inter-correlated than on average. In this study (**III**, **IV**), all attempts to ascribe the large variability in the estimates of r_B to factors distinguishing between the trials, were unsuccessful. Despite gross differences between parallel trials in mean height, initial density and survival percentage, none of these variables showed any statistically significant relation to the estimates of between-site genetic and rank correlation.

The effect of testing method on GEI also appeared to be negligible (**III**, **IV**). The average genetic correlations for pairs of forestry trials and pairs of test orchard and forestry trial, respectively, were exactly the same, $r_B = 0.62$ (**III**). This implies that the test orchard method, as such, is not an important source of GEI. Hence, systematically biased family rankings due to application of field testing methods that involve close spacings and unconventionally fertile sites are not a serious concern. However, the moderately low correlation suggests that the test orchards trials are afflicted with a general problem of inconsistent family performance at least as much as the conventional trials.

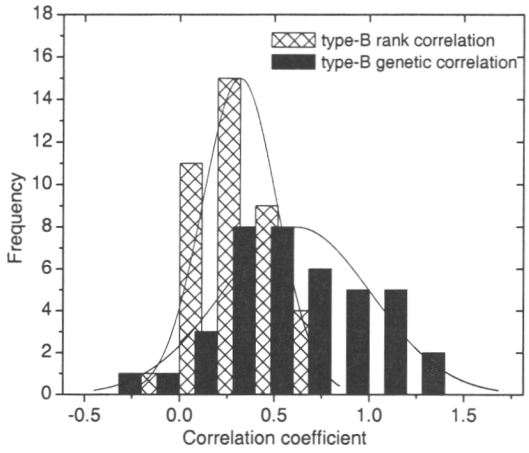

FIGURE 2. Frequencies of paired-site (type-B) genetic and rank correlations, estimated from the data comprising eight series of 10-year-old Scots pine progeny trials (**III**, Table 3)

BURDON (1977), CARSON (1991) and MIKOLA (1993) suggested focusing selection efforts on a single or a few sites which give the best screening for general adaptedness. However, the identification of such environments for Scots pine progeny testing is likely to be difficult due to the highly indeterminate nature of GEI. MATHESON and COTTERILL (1990) pointed out that under random environmental variability (typical for sites of the same region), the nature of family-by-site interactions is also random and should not be expected to be repeatable. It is consequently probable that much of the observed instability between families in parallel progeny trials lacked biological significance. As a noteworthy indication of this, the family-by-block (plot) variance component was frequently of much greater magnitude than the family-by-trial component (**IV**). This suggests that the incoherent performance of families was mainly due to poor control of field variation. This degrades the repeatability of family rankings within any site, and consequently, type-B correlations. Since the pairs of trials with especially poor genetic correspondence were usually associated with low values of family heritability (**III**), a substantial part of the observed interaction might be eliminated simply by improving the statistical precision of family performance within progeny test sites. In the context of Finnish tree breeding, the internal precision of trials could most easily be improved by increasing the number of plot replications at the expense of the currently large plot sizes (**I**).

Irrespective of the underlying reasons contributing to the observed levels of instability, the rank correlations between sites were alarmingly small. Nearly half of the 39 family-rank correlations (**III**) were less or equal than 0.25 (statistically non-significant) (Fig. 2). Evidently, the problem of instability was not merely limited to the pairs of trials with the poorest statistical efficiencies. Although sufficient intra-site replication is essential in achieving the required discriminating ability, the effect of family-by-trial interaction on the variance of the family-mean

is most effectively neutralized by replicating trials in several sites. WHITE and HODGE (1990) showed that, in principle, the allocation of test material to a maximum number of locations and a minimum number of blocks per location should result in maximum efficiency for parental selection. However, increases in costs with increasing number of sites are likely to suggest a smaller optimum. LINDGREN (1984, 1985) provided a formula for determining a cost-efficient number of test sites and suggested that five localities should be adequate for Scots pine progeny testing in Sweden (LINDGREN 1984). JOHNSON (1997) concluded that adding additional sites beyond three only marginally improved genetic gain in Douglas-fir (*Pseudotsuga menziesii*). RUSSELL and LOO-DINKINS (1993) determined four as the minimum number of cloned genetic trials to eliminate the impacts of high GEI. In radiata pine, a random sample of five trials from a total number of 11 trials were found to capture over 80% of the maximum predicted genetic gains obtained when selecting at all 11 sites (CARSON 1991). WHITE and HODGE (1992), in turn, ended up suggesting six to twelve locations for advanced-generation progeny trials of slash pine (*Pinus elliottii*) with non-contiguous plots and four blocks per location.

No analysis was carried out in this study to determine the number of progeny trials appropriate for Finnish conditions, yet such an examination would obviously be of great value in order to optimise the economics of future genetic testing (LINDGREN 1985, WHITE and HODGE 1990) considering the sizeable family-by-trial interaction variance. However, the general uncertainty about the accurate sizes of variance components can become a problem of the analysis. The odds are that the modified designs of future progeny trials, strongly called for in this study, would have a partitioning of the total variance substantially different from that of the present trials. It is questionable to what extent the variance component estimates from the first-generation progeny trials could be applied to determine the optimal number of future test sites. One option might be to carry out a sensitivity analysis, where the values of relevant variance components and cost parameters (LINDGREN 1985) are varied within a reasonable range and the model output observed for each scenario. Such an approach could provide satisfactory approximations, or at least safe minimum estimates, for the number of test sites required to obtain precise family rankings with future progeny trials.

5.2.2. Genetic parameters

One of the primary objectives of genetic field trials is to provide precise and accurate estimates of quantitative genetic parameters for use in, for example,

predicting genetic gains, constructing effective selection indices and comparing different testing strategies (ZOBEL and TALBERT 1980).

The overall additive genetic control of tree height in the progeny trials was found to be weak. The mean of the across-site (i.e., unbiased) estimates of individual heritability (h^2) was only 0.12 (**IV**). Another set of estimates was developed using the variance component estimates for eight series of progeny trials in **III**. As shown in Fig. 3, these estimates were in close agreement with the ones obtained in **IV**. The present values of heritability fall in the lower end of the range of published estimates. A median value of 0.25 was reported for the heritability of height by CORNELIUS (1994), who reviewed 67 papers for a number of genetic parameter estimates. The generally low degree of genetic determination is especially noteworthy as the estimates of heritability reported here (Fig. 3, **III**, **IV**) were, apparently, overvalued rather than undervalued. First of all, the estimates were block-adjusted, i.e., computed using a formula that did not include the among-block component of variance in the denominator of h^2 (COTTERILL 1987). Secondly, they were estimated on open-pollinated materials assuming a truly half-sib family structure. Estimates of genetic variance and heritability from such analyses can be upwardly biased due to selfing and inclusion of full-sibs in the assumedly half-sib families (SQUILLACE 1974).

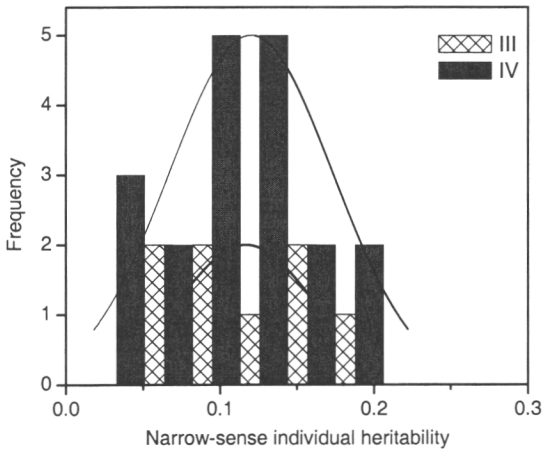

FIGURE 3. Frequencies of unbiased estimates of narrow-sense individual heritability for tree height in Scots pine progeny trials, compiled from the results of across-site analyses conducted in papers **III** and **IV**.

The single-site estimates of heritability were extensively inflated due to the profuse family-by-trial interactions. The relative magnitude of the error in these biased estimates was about 50% (corresponding to $100/r_B - 1$) (**III**, **IV**). The biased single-site estimates were generally below 0.4, conforming to the range (0.1 – 0.4) reported by CORNELIUS (1994). The corresponding estimates of family heritability were, as expected (WRIGHT 1976), markedly higher, varying in a broad range from 0.3 to 0.8 (**III**).

The test orchard trials, as a group, were capable of discriminating individual and family-level genetic differences markedly better than the conventional trials. The impact of site type on the discrimination ability was rather decisive. Test orchards laid out on uniform agricultural soils showed distinctly high levels of heritability as compared to either forestry trials or test orchards placed on less fertile forest sites (**III, IV**).

Results on the development of heritability over time were inconsistent. The time-series of single-site estimates from different trials displayed both increasing and decreasing tendencies. A pooled analysis of heritability estimates, however, failed to show any trend (**IV**). The lack of a systematic time trend contrasts with some earlier studies with conifers in the genus *Pinaceae*, which have usually found heritability to be initially low and then increase with age (FRANKLIN 1979, HODGE and WHITE 1992, BALOCCHI et al. 1993, DIETERS et al. 1995, COSTA & DUREL 1996, JOHNSON et al. 1997). However, stable ratios of family to phenotypic variance have also been reported, for example, by HANNRUP et al. (1998) in Scots pine, and LAMBETH et al. (1983) and FOSTER (1986) in loblolly pine (*Pinus taeda*). The absence of age trends in genetic variance and heritability could be due to the fact that the trial stands examined had not yet entered the critical phase where intense inter-family competition steps in, exaggerating differences among families with inherently different growth rates, as proposed by FRANKLIN (1979). The detection of systematic trends was also complicated by the moderately large standard errors associated with the estimates of heritability. On average, the relative size of the standard error (to that of the estimate) was about 60% (**IV**).

The additive genetic coefficient of variation (CV_A) indicates the potentiality of trait for genetic evolution (HOULE 1992). In this study (**IV**), these estimates were mostly below 15%, conforming to earlier results (e.g. FOSTER 1986).

The patterns of age-age correlations (r_G) for height growth have been inadequately known for Scots pine. Most of the published age-age correlations are from provenance trials (NILSSON 1991) and of doubtful value for progeny testing. The present study (**IV**) confirmed the findings of others, for example, RIEMENSCHNEIDER (1988), JOHNSON et al. (1997) and GWAZE et al. (2000), that age-age correlations can be predicted with a reasonable accuracy by the LAMBETH model (LAMBETH 1980). In this study, the correlation estimates were best fitted by the following equation (**IV**):

$$r_G = 1.02 + 0.423 \log_e(\text{younger age/older age}) \qquad (R^2 = 0.53)$$

The slope coefficient, applicable to the progeny testing of Scots pine, is notably steeper than that of LAMBETH's (1980) 'universal equation' for phenotypic age-age correlations (0.308). Deviations from the Lambeth slope and intercept are not unusual, however (JOHNSON et al. 1997, GWAZE et al. 2000). More often than not, the slopes have been flatter compared to the original estimate (LAMBETH and DILL 2001). The age-age (type-A) correlations examined in this study were not genetic in the strict sense, as they were computed between least-square estimates of family means at each site. This choice was intentional, based on the fact that direct estimates of genetic correlations from tree breeding trials tend to have large standard errors due to small sample sizes (KLEIN et al. 1973, ROFF 1995, HODGE and WHITE 1992). Such a phenomenon was also evidenced in this study (**III**) by the large frequency of irrational estimates ($r_G > 1$) of type-B genetic correlations.

5.2.3. Selection efficiency

The highest responses to selection were consistently associated with the close-spaced test orchard trials (**III**, **IV**). The superiority of this group, especially for parental selection, was associated with high overall levels of heritability compared to the forestry trials. The other component of indirect genetic gain, age-age correlations, did not show divergent tendencies between the two testing methods. Initial spacing and site quality had independent effects on the levels of heritability and indirectly to selection efficiency (**IV**). The widely-spaced forestry trials situated on typical forest soils were characterised by distinctively weak efficiency, with predicted gains for this group being only about half of the level for the other types of trials.

In Finland, ages from 10 to 15 years have been suggested as adequate for selection for growth in test orchard conditions (MIKOLA 1985). In this study, the optimal ages were determined to be 8 to 16 years (from seed) for within-family (forward) selection and 5 to 7 years for parental (backward) selection. The earlier optimal age for the parental mode of selection is consistent with previous findings (LAMBETH et al. 1983b, BALOCCHI et al. 1994, JOHNSON et al. 1997) and was, in fact, expected since backward selection is generally based on more information than forward selection. At the optimal ages, parental selection was found to be 40% to 60% more efficient than direct selection, depending on the type of trial. As for the within-family mode of selection, the gains from the use of indirect information were only few percentage units. Over the first 10 years of testing, selection for within-family deviations proved less profitable than direct selection for height at age 20. In contrast to the evaluation in terms of annual genetic gains,

the cumulative responses to indirect selection (both backward and forward) raised steadily over time (**IV**).

The optimal selection ages in forest tree breeding reported in the literature are noticeably inconsistent. This is due to the diverse nature of the models used, the circumstances under which the genetic parameters are estimated (KANG 1985), and, of course, the biology of tree species in question. The results obtained in this study are also associated with a number of simplifications and uncertainties and should be evaluated with caution. Most importantly, all the predictions of genetic gain were expressed in terms of tree height at age 20 years. This can hardly be considered as a goal trait, since 20 years is equivalent to only one-fourth of the commercial rotation of Scots pine in southern Finland. The genetic association between age-20 and rotation-age performances is obviously imperfect but inestimable. JOHNSON et al. (1997) daringly extrapolated age-age correlations using Douglas-fir data at ages 5 through 25, to rotation age of 60 years. Such an approach, while technically possible, was here considered to be too uncertain to provide reliable results (**IV**).

5.3. Implications for future genetic field testing

5.3.1. Designing better field trials

The results (**III**, **IV**) clearly speak in favor of the continued and augmented use of the test orchard method (MIKOLA 1985) in the Finnish progeny testing of Scots pine. This finding is consistent with a number of earlier papers which suggest that progeny testing resolution can be markedly improved by planting trials on uniform (agricultural) sites at close spacings (FRANKLIN 1979, MIKOLA 1985, MAGNUSSEN & YEATMAN 1986, WOODS et al. 1995, MAGNUSSEN 1995, BRIDGWATER and MCKEAND 1997). Most importantly, the risk of biased selection due to the testing of trees on homogeneous and fertile sites sites with narrow spacings was found to be small (**III**, **IV**). When family heritability and consistency of family performance over sites are considered simultaneously, test orchards appear considerably more attractive choice than the forestry trials. While the validity of the information from the forestry trials may be less controversial than that from the test orchards, test orchard trials appear to more commonly fail in satisfying the main objective of testing, that is, the precise, rapid and cost-efficient ranking of genotypes. The inferior performance of the forestry trials was involved with the generally high degree of site heterogeneity and its less than optimal control. The sparse planting density and the large numbers of trees per

plot together enlarged block sizes and the experimental error. As a final point, there is little reason to continue Scots pine progeny testing as sparsely spaced stands on heterogeneous forest sites (presently the most common combination).

In the future, test orchards will be increasingly laid out on uniform and fertile soils, using sparser spacing than in currently ongoing test orchard trials (HAHL, pers. comm.). Although high initial tree density did have a positive and significant effect on the heritability of height in this study (**IV**), there may be other reasons defending the use of lower densities, such as the easiness of measurement and, in particular, more reliable evaluation of traits related to stem quality. Of site uniformity and close spacing, the former factor seems to have more importance. Field test sites were associated with a significantly better genetic discriminating ability than forest sites, questioning the real worth of the latter for genetic testing purposes. Homogeneous field sites have the further advantage of enabling the application of sophisticated computer-generated experimental designs (such as alpha or row-column designs advocated by WILLIAMS and MATHESON 1994) which may not be easily practised on heterogeneous forest soils. Such non-orthogonal designs have theoretical advantages over the ordinary randomised complete blocks (BARNES and SCHWEPPENHAUSER 1979, WILLIAMS and MATHESON 1994), and their use should be considered, in the upcoming test orchard trials, as a way to neutralise the negative effect of wider initial planting spacing on the control of soil variability. Furthermore, incomplete block designs can accommodate a great number of parents, which is preferable for many purposes of genetic testing (LINDGREN 1991).

Further improvements in testing efficiency are possible through the introduction of more efficient plot configurations (**I**). Most progeny trials in Finland employ sizeable multi-unit plots which have a particularly low efficiency on a per tree basis. Smaller plots allow one to reduce the total land area required for testing while maintaining the precision at the original level. In addition, more families can be tested within any single trial, which permits a higher selection intensity and more reliable estimation of genetic parameters (KLEIN et al. 1973). Large pure-family plots do not seem to offer any real advantage over the most efficient configurations. Thus, for future progeny trials in Finland, large block plots should be predominantly substituted by single-tree and non-contiguous plots.

Progeny trials aimed at early selection are often completed in a couple of decades after which the sites are occupied by new trials. This opens an exciting possibility of utilising the data from the previous trial to delineate a fresh set of blocks which maximise the control of microsite variability. The precision might also be increased by using observations from a past trial as accessory information

(covariates). Although such an analysis might be done on a plot-mean basis, the outcome of the 'information recycling' would certainly be more productive if the row and column position of each and every experimental tree were recorded in the course of the measurement. Such data is presently lacking from most Finnish field progeny trials, except for the few trials established with single-tree plots or non-contiguous plots, as the mapping of trees is a built-in part of these procedures. Due to the typical heterogeneity of forest sites, the option of mapping is probably restricted to the most homogeneous testing sites (in practice, to formerly agricultural sites).

5.3.2. Increasing opportunities of effective selection

The strength of the Finnish Scots pine progeny testing lies in the size of the ongoing programme and extensive measurement data used to rank the first-generation plus-trees. The ranking for backward selection is currently based on an extended version of the 'performance level' method that uses single-site family means as observational units (VENÄLÄINEN 1994). While such straightforward methods are imperfect in several ways, they are likely to rank the plus-tree parents with a sufficient accuracy (COTTERILL et al. 1983, BRIDGWATER and MCKEAND 1997) given the wealth of the data available for analyses. In future progeny testing, the test materials will necessarily be significantly fewer and the inadequacies of experimental design or analytical methods will not be as easily compensated by the generous sample sizes. Incorrect estimates of genetic values that result from incompetent handling of data can nullify all the preceding efforts of testing. Therefore, much emphasis ought to be put into the proper methods of data analysis. Robust methods may entirely change the conclusions of an experiment (MAGNUSSEN 1993). Most importantly, efficient data analysis is, in general, the cheapest of all ways to gain additional genetic information from the target population.

The significant progress made in the analysis of genetic testing data over the past decades is primarily due to animal geneticists who were confronted by the need for appropriate handling of field data with multiple flaws (e.g., non-normality, extensive missing data, unequal representation of genotypes over fixed factors of the model), long before such problems became an issue in forest tree breeding. The most important contributions were the techniques used for reliable estimation of covariances and variance parameters in unbalanced data (HENDERSON 1953, HARTLEY and RAO 1967, PATTERSON and THOMPSON 1971, HARVILLE 1977, HUBER et al. 1994). First-class estimates of these parameters are essential for predicting breeding values with BLUP properties through mixed-model equations

(HENDERSON 1975). In recent years, these methods, largely used for genetic evaluation of livestock, have gained an increasing footing in the analysis of tree-breeding data (e.g., WHITE and HODGE 1989, ERICSSON and DANELL 1995, JARVIS et al. 1995, SORIA et al. 1998).

Even the most powerful analytical methods are incapable of making up for the intrinsic defects in the input data. From this point of view, the unidimensional nature of the information from the first-generation testing population of Scots pine in this study is a clear weakness, affecting both the accuracy of selection and the inferences drawn. Irrefutably, forest tree breeding is a multivariate discipline, so the current dominance of tree height records in the measurement database (over 95% of single data sets) is limiting. There are obvious reasons for the choice of tree height as the dominant variate; it is unambigious, easy to measure and favorably related to vigor. The few sets of multiple-trait data from young Finnish progeny trials indicate that height, compared to diameter and volume, is best genetically correlated to a number of stem–quality traits (HAAPANEN and PÖYKKÖ 1992, HAAPANEN et al. 1997). Furthermore, studies with species of pine (with notably shorter rotations, though) have indicated that early height measurement could be a good predictor of rotation-age volume and vigour (LAMBETH et al. 1983b, FOSTER 1986, NILSSON 1991, COSTA and DUREL 1996) and less prone to competition effects than other growth traits (SAKAI et al. 1968, KREMER 1992, PAUL et al. 1997).

Regardless of the apparently good indicative properties of juvenile height, it is clearly limiting to base selection on just a single trait. In Scots pine, traits related to the branching quality of the butt log become increasingly more important in determining the end-product value as the trees mature. Moreover, selection should somehow account for the capacity of genotypes to withstand various stresses that inhibit the productivity. For these reasons, future measurement procedures need to be diversified to include an optimal set of juvenile variates that are relatively independent, moderately inherited, and satisfactorily correlated to the target traits of economical importance. Scots pine breeding in southern Finland would certainly benefit from more research invested in methods of assessing stem-quality in young progeny trials. This work has already been started (PÖYKKÖ 1993, VENÄLÄINEN et al. 1996). Choosing a set of early traits that should be used to define the best genotypes is a critical step in future breeding, and should be made with consideration and care. It requires the exploration of the genetic determinism of the traits in order to apply effective multiple trait indices in selection (TALBERT 1986, COTTERILL and DEAN 1990). Another crucial decision is that of the breeding objective. In the present tree breeding programme, the primary target of improvement is defined as '*..height growth and the consequent*

growth of stem volume" (PITKÄNTÄHTÄYKSEN... 1989). Instead of such a vague definition, the goal of breeding should be defined in terms of a clearly described tree model (ideotype) (PÖYKKÖ 1993). The theoretical framework of ideotype breeding (DONALD 1968) would be very useful when devising the future selection strategy for Scots pine.

The selection index approach (COTTERILL and DEAN 1990) could be utilised also to combine information on a candidate tree and its relatives in an optimal way (BARADAT 1976, BURDON 1979, HARVEY and TOWNSEND 1985). This variety of the usual multi-trait index would be especially rewarding in the forward selection of individuals, considering the meagre levels of heritability observed in **IV**. Making effective use of the individual's own performance, family performance in the same block as the individual, family performance in other blocks of the trial and in other trials, could lead to significantly better discrimination between individual trees than considering the individual's own phenotype alone (HARVEY and TOWNSEND 1985, WHITE and HODGE 1989).

6. Conclusions

This aim of the study was to examine the impacts of various methods and options on the efficiency of genetic field testing, and to provide an insight into the information obtained from the ongoing Scots pine progeny trials. Based on the results obtained in the four publications included, the following answers could be provided to the specific questions posed in the "Objectives" chapter.

- *Are there prospects of improving the efficiency of genetic testing by modifying the established procedures of trial installation and data collection?*

  Definitely. The types of large block plots, which currently dominate progeny trials in Finland, were found to be notably ineffective with regard to the main purpose of progeny testing (precise and cost-efficient parental evaluation). The economics of field testing could be considerably improved by use of either single-tree or non-contiguous plot configurations. Large plots do not seem to offer any real advantage over efficient configurations. In future progeny testing, large plots should, for the most part, be abandoned, with an expection of trials with special aims (see discussion).

  A portion of the extraneous costs arising from ineffectively large multiple-tree plots can be avoided by measuring only a fraction of the total number of trees. In addition, (beyond the papers included in this study) further gains in testing efficiency could be achieved by applying more sophisticated experimental designs under favourable conditions.

- *How influential are genotype-by-environment interactions (GEI)?*

  The open-pollinated plus-tree families were, as a rule, notably inconsistent in their performance across replicated progeny trials. The between-site instability was especially troubling as it was associated with substantial changes in family ranking. Nevertheless, no underlying factors could be detected to explain the high levels of GEI observed. The unpredictable nature of the family-by-trial interactions casts doubt on the biological significance of the phenomenon, suggesting that little or none of GEI was due to real (repeatable) genotype-by-environment interactions. Cautious interpretation of type-B correlations is therefore recommended. To obtain

precise estimates of parental CGAs, the observed levels of instability make it necessary to improve the discriminating ability of progeny trials and to distribute the testing effort over a fairly large number of test sites. The family-by-trial interactions made the estimation of additive genetic variance and heritability on a single-site basis substantially biased. The single-site estimates of these parameters were inflated by approximately 50%. Along with their relatively high standard errors, the single-site heritability estimates are of poor value in predicting any kind of response to selection.

- *What risks may be associated with the test orchard method?*

  No evidence was found to indicate a significant increase in the risk of maladaptive selections specifically associated with field testing by test orchards. The overall problem of inconsistent family performance, however, concerned the test orchards trials as much as the conventional trials.

- *What tendencies may be distinguished in the development of key genetic parameters over time?*

  The general level of additive genetic control for cumulative tree height in young progeny trials of Scots pine is weak ($h^2 \sim 0.12$). Heritability and the magnitude of GEI (measured by type-B correlation) did not show significant time trends over the first two decades of testing. Over this period, estimates of age-age correlation varied from moderate to high and could be predicted reasonably well by the log of the age ratio.

- *How effective is indirect selection and to what degree it is affected by the distinctive features (spacing, site fertility) of field testing methods?*

  The relative gains from indirect selection increased with improved information on genetic values. Indirect parental (backward) selection at an optimal age of five to seven years provided significantly more gains per unit time than direct selection for age-20 height, whilst for within-family (forward) selection the corresponding gains were nominal. The range of favorable ages for indirect forward selection was much broader (8 to 16 years) than for backward selection, the earliest and latest optimums being due to test orchards and forestry trials, respectively.

Close initial spacing and site fertility (rapid growth) independently augmented the levels of heritability. Accordingly, the highest selection efficiencies were associated with the densely spaced test orchard trials that were placed on former agricultural sites. In contrast, the widely-spaced forestry trials, especially the ones showing slow growth (on forest sites), were inferior with respect to the rest of the trials examined. It seems advisable to abandon or significantly reduce the number of heterogeneous forest sites as environments for the future progeny testing of Scots pine.

Acknowledgements

The long lifespan of forest trees is a problem to all of us whose work relies on measurement data. Had I been forced to set out my research by establishing new field trials from the seed, completing this thesis would certainly have taken at least double the time it actually did. Fortunately, I did not need to start from the scratch. My first and strongest recognition for this goes to a distinguished veteran of Finnish tree breeding, Jouni Mikola, who during his employment in the Finnish Forest Research Institute in the 1970's and 1980's, designed and started nearly all of the progeny trials utilised in this study – plus numerous other experiments which continue producing valuable information for forest tree breeders and geneticists for years to come. As a real professional Jouni has never been possessive about the field trials under his 'copyright'. On the contrary, his generous and encouraging attitude toward my using the data from these trials has been tremendously important to me over all these years. I am also very grateful to Jouni for his willingness to comment my manuscripts.

The route of the measurement data from pine trials to the computer hard disk and, finally, to the tables and graphs of this thesis has involved a large number of people. I wish to express my sincerest thanks to our field staff, who have spent countless days in progeny trials gathering the data and bravely exposing themselves to heat, rain, mosquitos, bears and the bad coffee of roadside gas stations. The personnel of the Finnish forest genetics register are acknowledged for their extremely helpful attitude. I also thank all my fellow workers at Maxwell House, our 'home base' in the Vantaa Research Station, for creating the nice and jovial athmosphere. A special token of gratitude goes to Pirkko Velling for being such a nice person and boss over the years.

My colleagues and long-time rivals in the noble sport of chess, Seppo Ruotsalainen and Martti Venäläinen, have promptly commented many of my manucripts. Thanks!

Professors Max Hagman (the 'father' of the two field trials used in **I**), Veikko Koski, Peter Tigerstedt, and Pertti Pulkkinen, are all thanked for their various contributions to this thesis. Pertti has been my supervisor and constructive discussion partner, whose inexhaustible energy and dedication to research has not ceased to astonish me.

The two official referees, Gunnar Jansson and Tapani Pöykkö deserve the warmest thanks for their careful job, as well as Michael Starr who revised the

language of the summary. I would also like to mention that Tapani was the person who introduced tree breeding to me, and me to tree breeding back in the late 1980's. With this thesis I wish to prove that both these introductions were fruitful.

Finally, I wish to express my sincerest thanks to my wife Anu for her unconditional support, and to our lovely daughter Taru, two and half months old, to whom I dedicate this thesis.

Maxwell House, 15th of January, 2002

References

ADAMS, W., ROBERDS, J., ZOBEL, B. 1973. Intergenotypic interactions among families of loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics* 43: 319-322.

ADAMS, W., WHITE, T., HODGE, G. & POWELL, G. 1994. Genetic parameters for bole volume in longleaf pine: large sample estimates and influences of test characteristics. *Silvae Genetica* 43(5-6): 357-365.

ALLARD, R. 1960. Principles of plant breeding. John Wiley & Sons. 463 p.

APIOLAZA, L., BURDON, R. & GARRICK, D. 1999. Effect of univariate subsampling on the efficiency of bivariate parameter estimation and selection using half-sib progeny tests. *Forest Genetics* 6(2): 79-87.

ARNOLD, S., CLAPHAM, D. & EKBERG, I. 1990. Has biotechnology a future in forest tree breeding? *Forest Tree Improvement* 23: 31-47.

BALOCCHI, C., BRIDGWATER, F., ZOBEL, B. & JAHROMI, S. 1993. Age trends in genetic parameters for tree height in a nonselected population of loblolly pine. *Forest Science* 39(2): 231-251.

BALOCCHI, C., BRIDGWATER, F. & BRYANT, R. 1994. Selection efficiency in a nonselected population of loblolly pine. *Forest Science* 40(3): 452-473.

BARADAT, P. 1976. Use of juvenile-mature relationships and information from relatives in combined multitrait selection. Proc. IUFRO joint meeting on advanced generation breeding, INRA, Cestas: 121-138.

BARNES, R. & SCHWAPPENHAUSER, M. 1979. Genetic control of 1.5-year-old traits in *Pinus patula* Schiede et Deppe and a comparison of progeny test methods. *Silvae Genetica* 28 (4): 156-167.

BARNES, R., BURLEY, J., GIBSON, G. & GARCIA DE LEON, J. 1984. Genotype-environment interactions in tropical pines and their effects on the structure of breding populations. *Silvae Genetica* 33(6): 186-198.

BERGERUD, W. 1995. Power analysis and sample sizes for randomised block designs with subsampling. British Columbia Ministry of Forests, *Biometrics information pamphlet* no. 50: 1-4.

BISOFFI, S. 1993. Marker-assisted selection for the evaluation of forest reproductive material. *In:* Quality of forest reproductive material in the field of the application of European Community rules. Proceedings. Paris, 9-10 December 1993: 61-73.

BONGARTEN, B. and DOWD, J. 1987. Regression and spline methods for removing environmental variance in progeny tests. Proc. Of the 19th Southern Conference on Forest Tree Improvement: 312-319.

BRIDGWATER, F. & MCKEAND, S. 1997. Early family evaluation for growth of loblolly pine. *Forest Genetics* 4(1):49-55.

BURDON, R. 1977. Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. *Silvae Genetica* 25(5-6): 168-175.

BURDON, R. 1979. Generalisation of multi-trait selection indices using information from several sites. *New Zealand Journal of Forest Science* 9: 145-142.

BURDON, R. 1986. Breeding long-lived perennials – frustrations, temptations, opportunities. *In*: Plant Breeding Symposium DSIR, Lincoln, New Zealand (Eds: Williams, T. & Wratt, G.). Agronomy Society of New Zealand. Special publication 5: 137-143.

BURDON, R. 1989. Early selection in tree breeding: principles for applying index selection and infering input parameters. *Canadian Journal of Forest Research* 19: 499-504.

CARSON, S. 1991. Genotype x environment interaction and optimal number of progeny test sites for improving *Pinus radiata* in New Zealand. *New Zealand Journal of Forestry Science* 21(1): 33-49.

CONKLE, M. 1963. The determination of experimental plot size and shape in loblolly and slash pines. North Carolina State University, School of Forestry, Tech. Rep. No. 17. 51 p.

CORNELIUS, J. 1994. Heritabilities and additive genetic coefficients of variation in forest trees. *Canadian Journal of Forest Research* 24: 372-379.

CORRELL, R. 1978. The optimal plot size for *Pinus radiata* establishment trials. South Australian Institute of Technology. Mathematical project G: 1-40.

CORRELL, R. & ANDERSON, R. 1983. Removal of intervarietal competition effects in forestry varietal trials. *Silvae Genetica* 32(5-6): 162-165.

CORRELL, R. & CELLIER, K. 1987. Effects of plot size, block size and buffer rows on the precision of forestry trials. *Australian Forestry Research* 17: 11-18.

COSTA, P. & DUREL, C. 1996. Time trends in genetic control over height and diameter in maritime pine. *Canadian Journal of Forest Research* 26: 1209-1217.

COTTERILL, P., CORRELL, R. & BOARDMAN, R. 1983. Methods of estimating the average performance of families across incomplete open-pollinated progeny tests. *Silvae Genetica* 32(1-2): 28-32.

COTTERILL, P. 1987. On estimating heritability according to practical applications. *Silvae Genetica* 33(6): 203-208.

COTTERILL, P. & JAMES, J. 1984. Number of offspring and plot sizes required for progeny testing. *Silvae Genetica* 37: 138-146.

COTTERILL, P. & DEAN, C. 1990. Successful tree breeding with index selection. CSIRO Australia. 80 p.

COX, D. 1958. Planning of experiments. John Wiley.

DICKERSON, G. 1962. Implications of genetic-environmental interaction in animal breeding. *Animal Production* 4: 47-63.

DIETERS, M., WHITE, T. & HODGE, G. 1995. Genetic parameter estimates for volume from full-sib tests of slash pine (*Pinus elliottii*). *Canadian Journal of Forest Research* 25: 1397-1408.

DONALD, C. 1968. The breeding of crop ideotypes. *Euphytica* 17: 385-403.

ERICSSON, T. 1997. Enhanced heritabilities and best linear unbiased predictors trough appropriate blocking of progeny trials. *Canadian Journal of Forest Research* 27: 2097-2101.

ERICSSON, T. 1999. The effect of pedigree error by misidentification of individual trees on genetic evaluation of a full-sib experiment. *Silvae Genetica* 48 (5): 239-242.

ERICSSON, T. & DANELL, Ö. 1995. Genetic evaluation, multiple-trait selection criteria and genetic thinning of *Pinus contorta* var. *latifolia* seed orchards in Sweden. *Scandinavian Journal of Forest Research* 10: 313-325.

EVANS, T., BARBER, J. & SQUILLACE, A. 1961. Some statistical aspects of progeny testing. *In:* Proc. of the 6th Southern Conference on Forest Tree Improvement, School of Forestry, Gainesville, Florida, June 7-8.: 73-79.

FALCONER, D. 1981. Introduction to quantitative genetics. 2nd. ed. Longman Group Ltd. 340 pp.

FINLAY, K. & WLKINSON, G. 1963. The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research* 14: 742-754.

FINNISH STATISTICAL YEARBOOK OF FORESTRY. 2001. Finnish Forest Research Institute (Ed.: Peltola, A.). 374 p.

FOSTER, G. 1986. Trends in genetic parameters with stand development and their influence on early selection for volume growth in loblolly pine. *Forest Science* 32(4): 944-959.

FOSTER, G. 1989. Inter-genotypic competition in forest trees and its impact on realized gain from family selection. Proc. of the 20th Southern Conference on Forest Tree Improvement: 21-35.

FRANKLIN, E. 1971. Statistical validity of single-tree plots in forest genetics research. *Silvae Genetica:* 73-75.

FRANKLIN, E. 1979. Models relating levels of genetic variance to stand development of four north american conifers. *Silvae Genetica* 28(5-6): 207-212.

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

GILL, J. 1987. Juvenile-mature correlations and trends in genetic variances in Sitka spruce in Britain. *Silvae Genetica* 36(5-6): 189-194.

GREENWOOD, M. & VOLKAERT, H. 1992. Morphophysical traits as markers for the early selection of conifergenetic families. *Canadian Journal of Forest Research* 22: 1001-1008.

GWAZE, D., BRIDGWATER, F., BYRAM, T., WOOLLIAMS, J. & WILLIAMS, C. 2000. Predicting age-age genetic correlations in tree-breeding programs: a case study of *Pinus taeda* L. *Theoretical and Applied Genetics* 100: 199-206.

HAAPANEN, M. & PÖYKKÖ, T. 1993. Genetic relationships between growth and quality traits in an 8-year-old half-sib progeny trial of Scots pine. *Scandinavian Journal of Forest Research* 8: 305-312.

HAAPANEN, M., VELLING, P. & ANNALA, M. 1997. Progeny trial estimates of genetic parameters for growth and quality traits in Scots pine. *Silva Fennica* 31(1): 3–12.

HAAPANEN, M., MIKOLA, J., RUOTSALAINEN, S. & VENÄLÄINEN, M. 1999. Combining genetic gain and diversity by means of open nucleus breeding system. *In*: Climatic adaptation of boreal tree species - Tree breeding in the Nordic countries. Proceedings from the 1998 joint meeting of Nordic Group for the Management of Genetic Resources of Trees and Nordic Arboretum Council, Biri, Norway, June 25-27, 1998. (Ed: Tore Skroppa): 26.

HAMBLIN, J. & ROSIELLE, A. 1978. Effect of intergenotypic competition on genetic parameter estimation. *Crop Science* 18: 51-54.

HANNRUP, B. & EKBERG, I. 1998. Age-age correlations for tracheid length and wood density in *Pinus sylvestris*. *Canadian Journal of Forest Research* 28: 1373-1379.

HARTLEY, H. & RAO, J. Maximum likelihood estimation for the mixed models for large complex pedigrees. *Biometrika* 54: 93-108.

HARVEY,W. & TOWNSEND, A. 1985. Selection on a combination of individual, family and stand merit in provenance tests. *Forest Science* 31(4): 813-823.

HARVILLE, D. 1977. Maximum likelihood approaches to variance component estimation and related problems. *Journal of The American Statistical Association* 72: 320-338.

HATTEMER, H. 1991. Breeding strategies for timber production. *In:* Genetics of Scots pine (Eds.: Giertych, M. & Matyas, C.). Elsevier Science Publishers, Amsterdam, Netherlands: 103-110.

HENDERSON, C. 1953. Estimation of variance and covariance components. *Biometrics* 9: 226-252.

HENDERSON, C. 1975. Best linear unbised estimation and prediction under a selection model. *Biometrics* 31: 423-447.

HILL, J., BECKER, H. & TIGERSTEDT, P. 1998. Quantitative and ecological aspects of plant breeding. Plant Breeding Series 4. Chapman & Hall. 275 p.

HODGE, G. & WHITE, T. 1992. Genetic parameter estimates for growth traits at different ages in Slash pine and some implications for breeding. *Silvae Genetica* 41(4-5): 252-262.

HOULE, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130: 195-204.

HUBER, D., WHITE, T. & HODGE, G. 1994. Variance component estimation techniques compared for two mating designs with forst genetic architecture through computer simulation. *Theoretical and Applied Genetics* 88: 236-242.

HÜHN, M. 1970. The competitive environment and its genetic reaction variations. 2nd meeting of the IUFRO working group on quantitative genetics, August 18-19, 1969, Raleigh, North Carolina.: 62-86.

JANSSON, G. & DANELL, Ö. 1993. Needs and benefits of empirical power transformations for production and quality traits in forest tree breeding. *Theoretical and Applied Genetics* 87: 487-497.

JANSSON, G. 1998. Approaches to genetic testing and evaluation in forest tree breeding. (Doctoral thesis). *Acta Universitatis Agriculturae Sueciae*. Silvestria 56: 1-24.

JANSSON, G., DANELL, Ö. & STENER, L-G. 1998. Correspondence between single-tree and multiple-tree plot genetic test for production traits in *Pinus sylvestris*. *Canadian Journal of Forest Research* 28: 450-458.

JANSSON, G. 2000. Theory of early testing. *In:* Proceedings from the SNS meeting, July 1-3, 1999, Uppsala, Sweden. Dept. of Forest Genetics, Research Notes 55.:16-24.

JÄRVINEN, E., HOKKANEN, T. & KUULUVAINEN, T. 1993. Spatial heterogeneity amd relationships of mineral soil properties in a boreal *Pinus sylvestris* stand. *Scandinavian Journal of Forest Research* 8: 435–445.

JARVIS, S., BORRALHO, N. & POTTS, B. 1995. Implementation of a multivariate BLUP model for genetic evaluation of *Eucalyptus globulus* in Australia", *In:* Eucalypt plantations: Improving fibre yield and quality, Proceedings of the CRCTHF-IUFRO Conference, Hobart, Australia, 19-24 February 1995: 212-216.

JOHNSON, G. 1997. Site-to-site genetic correlations and their implications on breeding zone size and optimum number of progeny test sites for coastal Douglas-fir. *Silvae Genetica* 45(6): 280-285.

JOHNSON, G., SNEIZKO, R. & MANDEL, N. 1997. Age trends in Douglas-fir genetic parameters and implications for optimum selection age. *Silvae Genetica* 46(6): 349-358.

JOHNSON, G. & BURDON, R. 1990. Family-site interaction in *Pinus radiata*: Implications for progeny testing startegy and regionalised breeding in New Zealand. *Silvae Genetica* 39(2): 55-62.

JOHNSSON, H. 1974. Progeny testing. Proc. of the Joint IUFRO Meeting S.02.04.1-3., Stockholm, Sweden: 377-383.

JOHNSTONE, R. & SAMUEL, C. 1974. Experimental design for forest tree progeny tests with particular reference to plot size and shape. In: Proc. Joint IUFRO Meeting S.02.04.1-3., Stockholm, Sweden: 357-376.

KANG, H. 1985. Juvenile selection in tree breeding: some mathematical models. Silvae Genetica 34(2-3): 75-84.

KEMPTON, R. 1982. Adjustment for competition between varieties in plant breeding trials. *J. Agric. Sci. Camp.* 98: 599-611.

KEMPTON, R. & HOWES, C. 1981. The use of neighbouring plot values in the analysis of variety trials. *Appl. Stat.* 30(1): 59-70.

KERR, R., GODDARD, M. & JARVIS, S. 1998. Maximising genetic response in tree breeding with constraints on group coancestry. *Silvae Genetica* 47(2-3): 165-173.

KLEIN, T., DEFRIES, J. & FINKBEINER, C. 1973. Heritability and genetic correlation: standard errors of estimates and sample size. *Behavior Genetics* 3(4): 355-364.

KNIGHT, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica* 19: 225-235.

KOCH, E. & RIGNEY, J. 1951. A method of estimating optimum plot size from experimental data. *Agronomy Journal* 43: 17-21.

KREMER, A. 1992. Predictions of age-age correlations of total height based on serial correlations between height increments in Maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics* 85: 152-158.

LAMBETH, C. 1980. Juvenile-mature correlations in *Pinaceae* and implications for early selection. *Forest Science* 26(4): 571-580.

LAMBETH, C., GLADSTONE, W. & STONECYPHER, R. 1983a. Statistical efficiency of row and noncontiguous family plots in genetic tests of lobloly pine. *Silvae Genetica* 32(1-2): 183-190.

LAMBETH, C., VAN BUIJTENEN, J., DUKE, S. & MCCULLOUGH, R. 1983b. Early selection is effective in 20-year-old genetic tests of loblolly pine. *Silvae Genetica* 32(5-6): 210-215.

LAMBETH, C. 1986. Effective genetic testing – practical and theoretical considerations. Proceedings of a Joint Meeting of Working Parties on Breeding Theory, Progeny testing and Seed Orchards, October 13-17, 1986, Williamsburg, VA: 24–33.

LAMBETH, C. & DILL, L. 2001. Prediction models for juvenile-mature correlations for loblolly pine growth traits within, between and across test sites. *Forest Genetics* 8(2): 101-108.

LEE, C. 1974. Effect of partial measurement on statistical precision and efficiency in a nursery study of *Pinus ponderosa*. *Silvae Genetica* 23(4): 124-126.

LIBBY, W. & COCKERHAM, C. 1980. Random non-contiguous plots in interlocking field layouts. *Silvae Genetica* 29(5-6): 183-190.

LINDGREN, D. 1984. Prediction and optimization of genetic gain with regard to genotype x environment interactions. *Studia Forestalia Suecica.* 166: 15-24.

LINDGREN, D. 1985. Cost-efficient number of test sites for ranking entries in field trials. *Biometrics* 41: 887-893.

LINDGREN, D. 1991. Progeny testing. *In:* Genetics of Scots pine (Eds.: Giertych, M. & Matyas, C.). Elsevier Science Publishers, Amsterdam, Netherlands: 191-203.

LINDGREN, D., GEA, L. & JEFFERSON, P. 1995. Loss of genetic diversity monitored by status number. *Silvae Genetica* 45(1): 52-59.

LITTELL, R., MILLIKEN, G., STROUP, W. & WOLFINGER, R. 1996. SAS System for Mixed Models. SAS Institute Inc., Cary, NC. 633 pp.

LOO, J., TAUER, C. & BUJTENEN, J. 1984. Juvenile-mature relationships and heritability estimates of several traits in loblolly pine (*Pinus taeda*). *Canadian Journal of Forest Research* 14: 822–825.

LOO-DINKINS, J. & TAUER, G. 1987. Statistical efficiency of six progeny test field designs on three loblolly pine (*Pinus taeda* L.) site types. *Canadian Journal of Forest Research* 17: 1066-1070.

LOO-DINKINS, J., TAUER, G. & LAMBETH, C. 1990. Selection system efficiencies for computer simulated test field designs in loblolly pine. *Theoretical and Applied Genetics* 79: 89–96.

MAGNUSSEN, S. & YEATMAN, C. 1986. Accelerated testing of Jack pine progenies: A case study. Proceedings of the joint meeting of IUFRO working parties on Breeding Theory, Progeny Testing, Seed Orchards. Williamsburg, Virginia, October 13-17, 1986. 107-121.

MAGNUSSEN, S. 1989a. Effects and adjustments of competition bias in progeny trials with single-tree plots. *Forest Science* 35(2): 532-547.

MAGNUSSEN, S. 1989b. Bias in progeny trials with single-tree plots. *Forest Science:* 533-547.

MAGNUSSEN, S. 1989c. Inter-plant interactions and their influence on within and among plot variances. *Scandinavian Journal of Forest Research* 4: 369-377.

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

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

MAGNUSSEN, S. 1994. A method to adjust simultaneously for spatial microsite and competition effects. *Canadian Journal of Forest Research* 24: 985-995.

MAGNUSSEN, S. 1995. The impact of genetic variation in relative growth rates on stem volume differentiation: a simulation study. *Silvae Genetica* 44(4): 194-205.

MATHESON, A. & COTTERILL, P. 1990. Utility of genotype x environment interactions. *Forest Ecology and Management* 30: 159-174.

MATZIRIS, D. & ZOBEL, B. 1974. Inheritance and correlations of juvenile characteristics in loblolly pine (*Pinus taeda* L.). *Silvae Genetica* 22(1-2): 38-45.

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

MCKEAND, S. Optimum age for family selection for growth in genetic tests of loblolly pine. *Forest Science* 34(2): 400-411.

MIKOLA, J. 1985. Methods used for the genetic evalution of tree breeding material in Finland. *In:* Crop Physiology of Forest Trees (Eds: Tigerstedt, P., Puttonen, P. & Koski, V.). University of Helsinki, Dept. of Plant Breedng 138-143.

MIKOLA, J. 1993. Progeny testing in conifers. Proceedings of the Nordic Group for Tree Breeding, Edinburgh, Scotland: 1-17.

NAMKOONG, G. 1974. Breeding for future generations. Proc. of the joint IUFRO meeting S.02.04.1-3, Stockholm, 1974: 29-39.

NAMKOONG, G. 1979. Introduction to quantitative genetics in forestry. USDA Forest Service, *Technical Bulletin* No. 1588. 342 p.

NANSON, A. 1974. Some problems of application of quantitative genetics to forest tree breeding. Proc. of the joint IUFRO meeting S.02.04.1-3. Stockholm: 1-15.

NANSON, A. 1989. Genotypic and genetic parameters, early testing and genotype x environment interaction. *In:* Proceedings of the IUFRO working party meeting S2.02-11, Sweden 1988. The Institute for Forest Improvement, Report No. 11.: 1-22.

NEALE, D. 2001. The ten year history of marker breeding in forest tree improvement. Preliminary abstracts of the meeting on 'Tree Biotechnology in the Next Millennium', Stevenson, Wa., USA, July 22-27, 2001: 21.

NIENSTAEDT, H. 1984. Breeding implications of juvenile selection in range-wide blac spruce provenance test. *Canadian Journal of Forest Research* 14:933-939.

NILSSON, J. 1991. The value of early testing. *In:* Genetics of Scots pine (Eds.: Giertych, M. & Matyas, C.). Elsevier Science Publishers, Amsterdam, Netherlands: 255-263.

NYQUIST, W. 1991. Estimation of heritability and prediction selection response in plant populations. *Critical Reviews in Plant Sciences* 10(3): 235-322.

PEARCE, S. 1976. An examination of Fairfield Smith's law on environmental variation..*Journal of Agricultural Science* 87: 21-24.

PATTERSON, H. & THOMPSON, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554.

PAUL, A., FOSTER, G., CALDWELL, T. & MCRAE, J. 1997. Trends in genetic parameters for height, diameter, and volume in a multilocation clonal study with loblolly pine. *Forest Science* 43(1): 87-98.

PEDERICK, L. 1990. Family x site interactions in *Pinus radiata* in Victoria, Australia, and implications to breeding strategy. *Silvae Genetica* 39(3): 134-140.

PHARIS, R., YEH, F. & DANCIK, B. 1991. Superior growth potential in trees: what is its basis, and can it be tested for at an early age? *Canadian Journal of Forest Research* 21: 368–374.

PITKÄNTÄHTÄYKSEN METSÄNJALOSTUSOHJELMA JA TYÖSUUNNITELMA VUOSIKSI 1990-1998. 1989. Metsänjalostusohjelman tarkennustyöryhmä. Helsinki. 127 p.

PÖYKKÖ, T. 1993. Selection criteria in Scots pine breeding with special reference to ideotype. - (Doctoral thesis) *Reports from the Foundation for Forest Tree Breeding* 5.: 1-66.

RIEMENSCHNEIDER, D. 1988. Heritability, age-age correlations, and inferences regarding juvenile selection in jack pine. *Forest Science* 34(4): 1076-1082.

ROFF, D. 1995. The estimation of genetic correlation from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* 74: 481-490.

ROSVALL, O., LINDGREN, D. & MULLIN, T. 1998. Sustainability robustness and efficiency of a multi-generation breeding strategy based on within-family clonal selection. *Silvae Genetica* 47(5-6): 307-320.

RUSSELL, J. & LIBBY, W. 1986. Clonal testing efficiency: the trade-offs between clones tested and ramets per clone. *Canadian Journal of Forest Research* 16: 925–930.

RUSSELL, J. & LOO-DINKINS, J. 1993. Distribution of testing effort in cloned genetic tests. *Silvae Genetica* 42(2-3): 98-104.

SAKAI, K., MUKAIDE, H. & TOMITA, K. 1968. Intraspecific competition in forest trees. *Silvae Genetica* 17(1): 1-5.

SARVAS, R. 1953. Instruction for the selection and registration of plus trees. Helsinki, Valtioneuvoston kirjapaino. (In Finnish with an English summary).

SHELBOURNE, C. 1972. Genotype-environment interaction: its study and implications in forest tree improvement. Proc. of the IUFRO joint genetics symposium, Japanese Forest Experiment Station: 1-28.

SHIUE, C. & PAULEY, S. 1960. Some considerations on the statistical design for provenance and progeny tests in tree improvement programs. *Forest Science* 7: 116-122.

SKRØPPA, T. 1984. A critical evaluation of methods available to estimate the genotype x environment interaction. *Studia Forestalia Suecica.* 166: 3-14.

SMITH, F. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Science* 28: 1-23.

SONESSON, J., JANSSON, G. & ERIKSSON, G. 2001. Retrospective genetic tests of *Pinus sylvestris* L. in growth chambers with two irrigation regimes and two temperatures. *Scandinavian Journal of Forest Research* 16: 21-29.

SORIA, F., BASURCO, F., TOVAL, G., SILIÓ, L., RODRIGUEZ, M. & TORO, M. 1998. An application of Bayesian techniques to the genetic evaluation of growth traits in *Eucalyptus globulus*. *Canadian Journal of Forest Research* 28: 1286-1294.

SQUILLACE, A. 1974. Average genetic correlations among offspring from open-pollinated forest trees. *Silvae Genetica* 23(5): 149-155.

ST. CLAIR, J. & ADAMS, T. 1991. Relative family performance and variance structure of open-pollinated Douglas-fir seedlings grown in three competitive environments. *Theoretical and Applied Genetics* 81(4): 541-550.

STEEL, R., TORRIE, J. & DICKEY, D. 1997. Principles and procedures of statistics – A biometrical approach. 3rd ed. The McGraw-Hill Companies, Inc. 666 p.

STEVENSON, W. & SAVILL, P. 1976. Sampling in forest experiments. *Commonwealth Forest Review* 55: 331-334.

STRAUSS, S., LANDE, R. & NAMKOONG, G. 1992. Limitations of molecular-marked-aided selection in forest tree breeding. *Canadian Journal of Forest Research* 22: 1050–1061.

TALBERT, C. 1986. Multi-criterion index selection as a tool for operational tree improvement. *In* Proceedings of the joint meeting of IUFRO working parties on Breeding Theory, Progeny Testing, Seed Orchards. Williamsburg, Virginia, October 13-17, 1986: 228-238.

THOMSON, A. & EL-KASSABY, Y. 1988. Trend surface analysis of a Douglas-fir provenance-progeny transfer test. *Canadian Journal of Forest Research* 18: 515-520.

TUSKAN, G. & VAN BUJTENEN, J. 1986. Inherent differences in family response to inter-family competition in loblolly pine. *Silvae Genetica* 35 (2-3): 112–118.

TUSKAN, G. & WILLIAMS, C. 1989. Comparison of plot designs for short-term trials. Proc. of the 20th Southern Conference on Forest Tree Improvement, Charleston, SC., June 1989. 8 p.

YAMADA, Y. 1962. Genotype by environment interaction and genetic correlation of the same trait under different environments. *Japanese Journal of Genetics* 37(6): 498-509.

VENÄLÄINEN, M., HAHL, J. & PÖYKKÖ, T. 1996. Assessing the quality of young stems in predicting the total monetary yield of Scots pine progenies. *Canadian Journal of Forest Research* 26(12): 2227-2231.

VENÄLÄINEN, M. 1994. The combined results of 190 progeny tests of Scots pine in southern and central Finland. *In:* Progeny Testing and Breeding Strategies. Proceedings. Nordic Group for Tree Breeding. (Ed.: Lee, S.). Edinburgh, Scotland, 6-10 October, 1993. Forestry Commission, Roslin. p. 36-42.

WHITE, T. & HODGE, G. 1989. Predicting breeding values with applications in forest tree improvement. Kluwer Academic Publishers. 367 p.

WHITE, T. & HODGE, G. 1992. Test designs and optimum age for parental selection in advanced-generation progeny tests of slash pine. *Silvae Genetica* 41(4-5): 293-302.

WHITE, T. & HODGE, G. 1991. Indirect prediction of genetic values. *Silvae Genetica* 40(1): 20-28.

WILLIAMS, E. & MATHESON, A. 1994. Experimental design and analysis for use in tree improvement. CSIRO Publ., Victoria, Australia. 174 pp.

VON EULER, F., BARADAT, P. & LEMOINEN, B. 1992. Effects of plantation density and spacing on competitive interactions among half-sib families of maritime pine. *Canadian Journal of Forest Research* 22: 482-489.

VON EULER, F. 1993. Plot design effects on phenotypic variability and dynamics of field competition between six *Pinus sylvestris* L. varieties. *Scandinavian Journal of Forest Research* 8: 163-173.

WOODS, J., KOLOTELO, D. & YANCHUK, A. 1995. Early selection of coastal Douglas-fir in a farm field test environment. *Silvae Genetica* 44(4): 178-186.

WOOLLONS, R. 1980. Importance of experimental design and plot maintenance in forest field experiments in Australasia. *Australian Forestry Research* 10: 71-82.

WRIGHT, J. & FREELAND. 1960. Plot size and expermntal efficiency in forest genetics research. Michigan State University, Technical Bulletin 280: 1-28.

WRIGHT, J. 1970. An improved record system for forest genetic nursery and plantation studies. *Silvae Genetica* 19: 64-68.

WRIGHT, J. 1976. Introduction to Forest Genetics. Academic Press Inc. 463 p.

WU, H., YEH, F., DHIR, N., PHARIS, R. & DANCIK, B. 1997. Genotype by environment interaction and genetic correlation of greenhouse and field performance in *Pinus contorta* ssp. *latifolia*. *Silvae Genetica* 46: 170-175.

WU, H. 1998. Study of early selection in tree breeding. 1. Advantages of early selection through increase of selection intensity and reduction of field test size. *Silvae Genetica* 47 (2-3): 146–155.

WU, R.. HU, X. & HAN, Y. 2000. Molecular genetics and developmental physiology: implications for designing better forest crops: *Critical Reviews in Plant Sciences* 19(5): 377-393.

YRJÄNÄ, L., KARVINEN, K. & NAPOLA, J. 2000. General statistics on forest tree breeding in Finland 2000. Finnish Forest Research Institute, Research papers 783: 1-45.

ZOBEL, B. & TALBERT, J. 1984. Applied Forest Tree Improvement. John Wiley & Sons. 492 p.

168 references

I

Silva Fennica 1992, Vol. 26 N:o 4: 201–209

Effect of plot size and shape on the efficiency of progeny tests

Matti Haapanen

TIIVISTELMÄ: KOERUUDUN KOON JA MUODON VAIKUTUS JÄLKELÄISKOKEIDEN TEHOKKUUTEEN

Haapanen, M. 1992. Effect of plot size and shape on the efficiency of progeny tests. Tiivistelmä: Koeruudun koon ja muodon vaikutus jälkeläiskokeiden tehokkuuteen. Silva Fennica 26(4): 201–209.

A simulation approach was applied to study the pattern of environmental variability and the relative statistical efficiency of 14 different plot types. The study material consisted of two nine-year-old field tests of Scots pine (*Pinus sylvestris* L.). The area of the test sites was 1.57 and 0.67 hectares. The efficiency was measured as the error variance attached to the estimate of family mean and the total size of a test needed to detect a given, least significant difference between two family means. The statistical efficiency tended to decline along with increasing plot size. The importance of plot shape was negligible compared to plot size. The highest efficiency was obtained with single-tree plots. Non-contiguous plots appeared to be considerably more efficient than block plots of equal size. The effects of intergenotypic competition on the choice of plot type are discussed.

Tutkimuksessa selvitettiin erilaisten simuloitujen koeruutujen avulla ympäristövaihtelun luonnetta sekä vertailtiin 14 erilaisen ruututyypin suhteellista tilastollista tehokkuutta kahdessa yhdeksän vuoden ikäisessä männyn kenttäkokeessa. Kokeet olivat kooltaan 1,57 ja 0,67 ha. Tehokkuutta mitattiin perhekeskiarvojen vaihteluun liittyvän virheen suuruudella sekä kokeen koolla, joka tarvittiin kutakin ruututyyppiä käyttäen kahden jälkeläistökeskiarvon tietyn suuruisen erotuksen osoittamiseen tilastollisesti merkitseväksi. Tulokset osoittivat tilastollisen tehokkuuden laskevan koeruudun koon kasvaessa. Ruudun muodon merkitys oli vähäinen ruutukokoon verrattuna. Yhden puun ruutujen järjestely oli tutkituista ruututyypeistä tehokkain ja hajaruudut osoittautuivat huomattavasti tehokkaammiksi kuin vastaavan kokoiset yhtenäisruudut. Ruututyypin valintaan vaikuttavista tekijöistä tarkastellaan erikseen koe-erien välistä kilpailua.

Keywords: *Pinus sylvestris,* plot size, experimental design, progeny testing, statistical methods, efficiency.
FDC 232.3

Author's address: The Finnish Forest Research Institute, Department of Forest Ecology, P.O. Box 18, SF-01301 Vantaa, Finland.

Accepted January 20, 1993

1 Introduction

The primary objectives of progeny testing are: to determine the true genetic value of parental trees by comparing the performance of their offspring, to select superior trees within the best families, and to obtain estimates of genetic parameters. The testing is conducted in field experiments that cover large areas and take a considerably length of time (10–30 years) to produce useful information. Since the early 1960's more than 1700 progeny tests have been established with several tree species in Finland (Metsänjalostuksen... 1992). A large number of tests demanding continuous management and measurement represents a considerable expense for a breeding program. This means that much emphasis should be paid to designing efficient tests. To be efficient, a progeny test should contain the minimum number of trees to achieve the test objectives (Lambeth et al. 1983).

The statistical efficiency of a progeny test is usually depicted by the number of individuals needed to achieve a defined precision in the estimate of a family mean (Lambeth et al. 1983), or to show that a given difference between two family means is statistically significant (Wright and Freeland 1960, Correll 1978, Cotterill and James 1984). Correll and Cellier (1987) and Loo-Dinkins and Tauer (1987) also based their definition of efficiency on the precision of family means. Several studies have shown that in field tests these parameters, and thus, the efficiency, can be affected by modifying the size and shape of the experimental units, i.e. plots (Smith 1938, Wright and Freeland 1960).

So far, there has been no general agreement on the best type of plot to be used in forest genetic testing. The great majority of Finnish progeny tests have been established using square plots of 25 trees (Fig. 1), a practice motivated by the desire to compare stand level growth of families. In the United States, on the other hand, row plots of 4 to 10 trees have been the most frequently used design (Lambeth et al. 1983, Loo-Dinkins and Tauer 1987).

According to statistical theory and several studies, plots of minimum size, i.e. single-tree plots, give the highest experimental precision within the framework of a fixed set of land area or plant material (Conkle 1963, Wright and Freeland 1960, Loo-Dinkins and Tauer 1987). Single-tree plots provide maximum control over environmental variation by minimizing the block size and thus, the soil heterogeneity within blocks. In an alternative design promoted for forest genetic

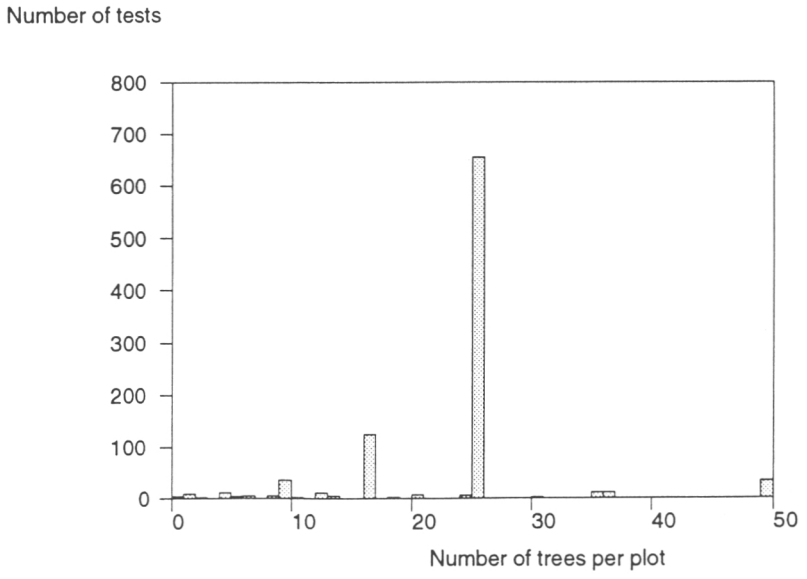

Fig. 1. The frequency distribution of plot sizes in Finnish progeny tests of Scots pine.

testing by Libby and Cockerham (1980), trees belonging to a multiple-tree plot are scattered at random throughout the block. They suggested that the statistical efficiency of the "non-contiguous" plot arrangement is comparable with that of single tree plots, although it avoids some of the major disadvantages of the latter design, including the problem of missing plots arising from natural mortality and artificial thinnings. Despite this, neither of these designs has so far been very popular among Finnish tree breeders.

Two approaches have been commonly used in studies on the plot technique with trees: the analysis of field experiments having sets of blocks with different plot configurations (Johnstone and Samuel 1974, Lambeth et al. 1983, Loo-Dinkins and Tauer 1987), and the simulation of plots in genetically uniform experiments or artificially regenerated stands (Wright and Freeland 1960, Conkle 1963, Correll and Cellier 1987). The simulation approach has the advantage that different plot configurations can all be evaluated under the same environmental conditions. Furthermore, the number of different types of plot that can be studied is practically unlimited.

The objective of this study was to investigate the relative efficiency of different plot types varying in size and shape, including some non-contiguous alternatives. The efficiency of the plot types currently used in Finnish progeny testing was of particular interest. Estimating the amount and pattern of environmental variability in the test sites was the other subject of investigation.

2 Material and methods

The study material consisted of two field tests of Scots pine, located in Ikaalinen (No. 358/1 – 61°54' N, 23°23' E) and Pylkönmäki (No. 358/2 – 62°41' N, 24°42' E). The tests were established in 1971 with a planting density of 2500 (2 × 2 m) trees per hectare. The area of the test sites was 1.57 ha (No. 358/1) and 0.67 ha (No. 358/2). The number of trees included in the analysis was 2869 in test No. 358/1 and 1428 in test No. 358/2. The nine-year height of every tree in both tests was measured.

The planting material used in these tests comprised only one entry, a registered stand origin from Lieksa, eastern Finland (63°04' N, 29°49' E). The entry represented a mixture of open-pollinated seeds collected from several trees within the stand. In this sense these tests were analogous to the large number of "uniformity trials" conducted with agricultural crops during the first decades of this century (Cochran 1937). Following the principles of uniformity trial research, any gradual change in height values within the field – indicating correlation between neighbouring trees – was interpreted as environmental variation (Smith 1938). The genetic differences among individual trees, also contributing to the total variance, were ignored since it was assumed that the expected genetic covariances between neighbouring trees to be zero, i.e. that they were unrelated.

The preparatory work consisted of grouping the data of adjacent trees into plots of varying size and shape. The number of different plot configurations superimposed on the field test data was 14. Plot size ranged from 1 to 49 trees per plot (Table 1). The analysis was based on division of the total variance of the height observations among individual trees into two variance components: 1) variance due to plot effects, $Var_{(plot)}$, and 2) residual variance arising from individual trees within plots, $Var_{(within\text{-}plot)}$. These components were solved for each plot type using the restricted maximum likelihood method in the VARCOMP procedure of the Statistical Analysis System (SAS Institute Inc. 1985). Difficulties in completely filling the test site with varying plot types meant that the number of trees analysed in different cases was not equal. However, because of the large number of trees included in the analysis, this was considered to have an insignificant effect on the precision of the variance components.

Two criteria were applied in comparing the statistical efficiency of different plot configurations: 1) the variance of family means attributable to the environment and 2) the size of the test needed to show that a given difference between two family means ($X_1 - X_2$) is statistically significant. The meaningful difference was chosen as 5 % of the overall test mean. Both measures were eventually converted to a percentage scale by relating them to the respective values ob-

tained for single tree plots.

The variance of family means ($Var_{(F)}$), presented as a sum of the variance components, is

$$Var_{(F)} = Var_{(family)} + Var_{(plot)} / r + Var_{(within\text{-}plot)} / r\,n, \quad (1)$$

where r is the number of replicates per family, and n the number of trees per plot. Since there were no genetically different entries in the studied experiments, the (irrelevant) between family component of variance was zero. The remaining two terms ($Var_{(plot)} / r + Var_{(within\text{-}plot)} / r\,n$) make up *the environmental portion of family variance* (Loo-Dinkins and Tauer 1987).

The size of a test, Z, (see the criterion No. 2) was calculated as:

Table 1. Experiments No. 358/1 (a) and No. 358/2 (b): Plot type and size (n trees), between- and within-plot variance components of height, variance of plot means ($Var_{(P)}$; used to calculate Smith's b-values), relative efficiency (the environmental portion of family mean variance) and relative size of test to detect statistical significance (at 5 % risk level) of differences equal to 5 % of the mean height (when compared to single-tree plots). The symbols of the plot types refer to the number of trees in 'columns' and 'rows', e.g. '7×2' means a rectangular plot with 14 trees arranged in seven columns and two rows; letters 'nc' symbolize a non-contiguous plot.

a) $n_{total} = 2869$

Plot type	n	$Var_{(plot)}$	$Var_{(within\text{-}plot)}$	$Var_{(P)}$	Rel. eff., %	Rel. size, %
1×1	1	18.54	0.00	18.54	100	100
3×1	3	2.47	15.87	9.34	64	151
7×1	7	2.10	15.48	5.11	52	193
8nc	8	0.34	18.20	3.45	67	149
14×1	14	1.63	16.94	3.29	40	249
7×2	14	1.90	16.64	3.53	38	268
16nc	16	0.00	18.54	1.58	73	137
7×3	21	2.02	16.88	3.09	28	353
14×2	28	1.60	17.04	2.43	27	371
6×5	30	1.44	17.39	2.24	28	366
14×3	42	1.53	17.22	2.09	21	481
7×6	42	1.95	17.24	2.43	18	560
7×7	49	1.88	16.69	2.34	16	629
49nc	49	0.06	18.48	0.58	65	156

b) $n_{total} = 1428$

Plot type	n	$Var_{(plot)}$	$Var_{(within\text{-}plot)}$	$Var_{(P)}$	Rel. eff., %	Rel. size, %
1×1	1	29.32	0.00	29.32	100	100
3×1	3	1.39	28.02	11.90	81	122
7×1	7	1.71	27.86	6.25	67	150
8nc	8	0.83	28.54	4.92	75	135
14×1	14	1.43	28.14	3.72	56	180
7×2	14	2.18	27.77	4.44	47	214
16nc	16	0.35	29.03	2.42	76	134
7×3	21	1.90	27.99	3.42	41	248
14×2	28	1.30	28.41	2.46	43	240
6×5	30	1.36	27.90	2.40	40	251
14×3	42	1.07	29.65	1.88	37	277
7×6	42	0.86	28.76	1.64	43	243
7×7	49	1.26	28.12	1.91	31	331
49nc	49	0.45	28.94	1.12	53	195

$$Z = \frac{2\,t^2\,(Var_{(plot)} + \frac{Var_{(within-plot)}}{n_H})\,n}{(X_1 - X_2)^2} \qquad (2)$$

(adapting White and Freeland 1960)

The term n_H stands for the harmonic mean number of trees per plot. The Student's t-value was computed at the 5 % risk level and 2r – 2 degrees of freedom.

"The empirical law of Fairfield Smith" (Smith 1938) was applied to determine the degree of environmental heterogeneity. According to the law, there is a linear relationship between the logarithmic variance of plot means and the logarithmic plot size in any field:

$$\log Var_{(P)x} = \log Var_{(P)1} - b \log x, \qquad (3)$$

where $Var_{(P)x}$ is the variance of plot means of size x (calculated as $Var_{(plot)} + Var_{(within-plot)} / n_H$), and $Var_{(P)1}$ the variance from plots of unit size (a single-tree plot). The regression coefficient, b, the range of which varies from 0 to 1, indicates the nature of the environmental variability. The smaller b is, the higher is the correlation between neighbouring trees, reflecting a non-randomly patterned environment where the adjacent microsites are more similar than more distant ones. Respectively, values of b approaching 1 are obtained under either very homogeneous or *randomly* heterogeneous site conditions. The coefficient is very useful since it only measures the degree of field heterogeneity and is completely independent of the amount of environmental variability (Smith 1938). The plot-mean variances of the 11 contiguous plot types provided the data for estimating the value of b.

3 Results

In the present material, the plot size and the statistical efficiency of the respective test designs were strongly negatively correlated (Table 1). The decrease of efficiency was rapid when the number of trees per plot was less than 20, retarding markedly with larger plot sizes (Fig. 2) The curves in Fig. 2 were drawn to demonstrate the difference in the results from the two test sites; their slight upward tendency on the right side arises from the property of the underlying regression model, and should not be interpreted as if the efficiency had reached a minimum value with around 35 to 40 trees per plot.

Of all plot types examined, the single-tree plot design was the most efficient one. Respectively, the largest contiguous plot type (7 × 7 square plots) gave the poorest result, reaching an efficiency of only 16 and 31 % of that of the single-tree plots. Comparison of the single-tree plots and square plots of 25 trees, showed that the efficiency of the single-tree plots was about three times higher than that of the latter plot type.

The three non-contiguous plot types studied represented an exception to the negative trend between plot size and efficiency in that they displayed a considerably similar efficiency independent of the plot size (Table 1). They also appeared to be clearly more efficient than the contiguous plot types of respective size. The relative benefit of non-contiguous plots increased along with the increasing number of trees per plot. When using square plots of 49 trees, the number of replications needs to be approximately two (358/2) to four (358/1) times larger than with non-contiguous plots of 49 trees, if an equal statistical efficiency is aimed at (Table 1). In other words, the information per tree given by contiguous plots is much smaller than that of non-contiguous plots.

The shape of the plots, compared to their size, appeared to be of minor importance for the efficiency (Table 1). The row plots and long rectangular plots, however, controlled the environmental variability slightly better than the more square-like plots with approximately the same number of trees.

Smith's (1938) measure of field heterogeneity, estimated by the slope of the regression between the logarithmic variance of plot means and the logarithmic plot size (b in equation 3), was 0.405 in test no. 358/1 and 0.507 in test no. 358/2 (Fig. 3). The difference between the coefficients was not statistically significant ($p_{(obs)} < 0.312$). Despite this, the result indicates that the

nature of environmental variability in test No. 358/1 (at Ikaalinen) was slightly more systematic or patterned (smaller b) than in test No. 358/2 (at Pylkönmäki). The relative efficiency differences between the plot types were more emphasized in 358/1 (Fig. 2), which is in accordance with the higher heterogeneity estimated for this test.

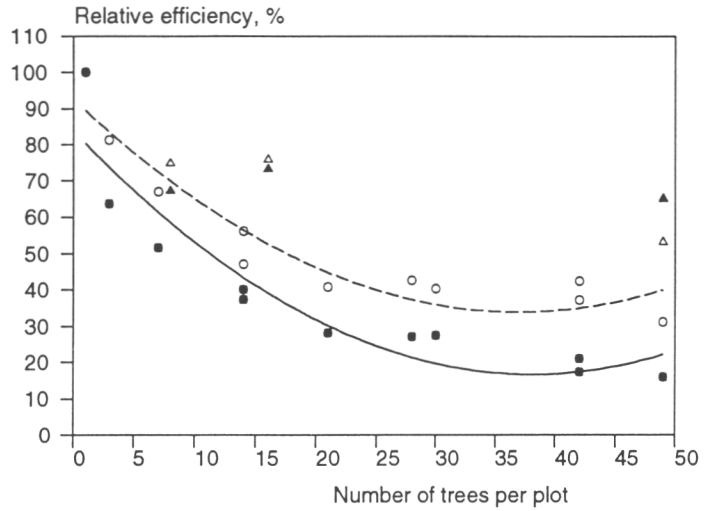

Fig. 2. Relative efficiency of contiguous (circles) and non-contiguous (triangles) plots in relation to plot size in experiments No. 358/1 (filled symbols, straight line) and No. 358/2 (unfilled symbols, dashed line). The regression lines are based on the equation $y = a + b_1x + b_2x^2$ (y representing relative efficiency and x plot size, respectively).

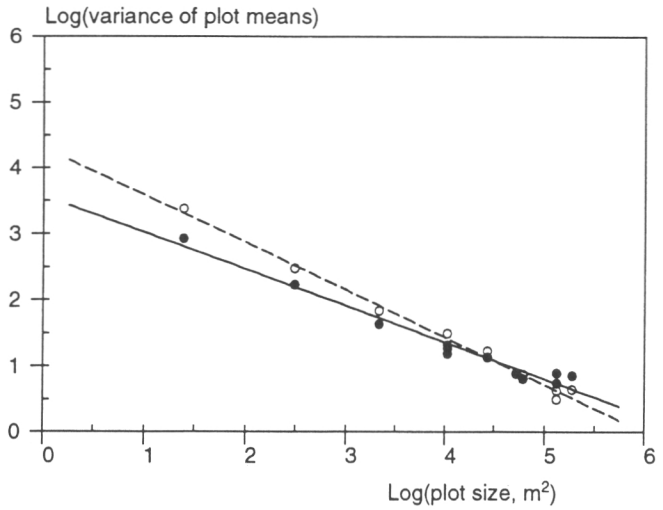

Fig. 3. Comparison of the pattern of environmental variability of the two test sites on the basis of Smith's coefficient of heterogeneity: Logarithmic variance of plot means plotted against logarithmic plot size and the respective regression lines in the experiments No. 358/1 (filled circles, a straight line) and No. 358/2 (circles, a dashed line). For further explanation see the text.

4 Discussion

According to the results, the economics of Finnish progeny testing could be significantly enhanced by decreasing plot size or by introducing non-contiguous plot designs. Even the use of small contiguous plots of, say, 2 to 5 trees would probably result in considerable gains (Fig. 2). The poor statistical efficiency of 5 × 5 plots found in this study is consistent with the findings of earlier studies: Conkle (1963) reported their relative efficiency to vary between 18 % and 50 % of that of single tree plots, whereas the respective measures obtained by Loo-Dinkins and Tauer (1987) ranged from 7 to 40 %.

Introduction of efficient plot techniques allows one to reduce the total land area assigned for testing, while maintaining the precision at the original level. In Finland, the use of large plots in experiments consisting of 4–6 blocks has obviously resulted in the testing of plus trees with unnecessary many offspring. The average family size in Finnish progeny tests of Scots pine is over 160 trees; moreover, many of the plus trees are being tested at two or more locations. For the sake of comparison, Cotterill and James (1984) suggested that a family size of only 10 to 20 individuals is needed to evaluate parental trees accurately enough, provided that single-tree or two-tree, non-contiguous plots are used.

As regards single-tree and non-contiguous plot systems, the latter appears to be a more attractive alternative for progeny testing. Since a non-contiguous plot consists of several trees, the problem of accidental damage resulting in missing plots is not as acute as with single-tree plots. Furthermore, by applying sophisticated planting designs which involve so-called interlocking replications within blocks, systematic thinnings can be carried out without disturbing the statistical orthogonality of the data (Libby and Cockerham 1980). The laborious establishment stage, which involves tagging all the seedlings and recording their exact location in a field, has usually been considered as an inconvenience common to both non-contiguous and single-tree plot design. However, field electronic data recorders have been developed to facilitate field work. In the opinion of Loo-Dinkins and Tauer (1987), the high statistical efficiency of non-contiguous and single-tree plot designs compared to simpler plot designs is sufficient to outweigh the possible additional work required. This may be, however, too simplistic an assumption. In this study the efficiency of small plots (less than 5–10 trees) appeared to be rather independent of whether they were arranged in non-contiguous or contiguous fashion (Fig. 2). Thus, with small plot sizes the selection of plot type may not always be obvious, but requires pre-evaluation of the expected costs and benefits in each individual case.

The situation is different, as far as large plots are concerned. Large plots, whether contiguous or non-contiguous, unavoidably cover large areas in randomized block designs, subjecting trees to different levels of environmental variability. The trees arranged throughout a block, however, sample the within-block variability effectively, all the families tested sharing a relatively equal block effect. This is not the case with contiguous plots which, due to the environmental differences within a block, tend to receive a divergent "environmental treatment". This easily causes relative family performances to vary from block to block. For example, Lambeth et al. (1983) found that the family-by-block interaction variance was almost zero in non-contiguous plots, while significantly high in row plots.

The relative statistical efficiency of various plot types is largely determined by the degree of field heterogeneity (Loo-Dinkins and Tauer 1987). The more significant is the systematic component of field variability (with a low Smith's coefficient of heterogeneity), the more the adjacent trees resemble each other as regards their environment and the less the information obtained per tree from plots of a given size, emphasizing the importance of small plots. On the other hand, little is gained by reducing the plot size or introducing a non-contiguous design if the site is relatively uniform or has a fine-grained mosaic structure with a random pattern (Smith's coefficient approaching 1). Unfortunately, accurate visual estimation of the degree of field heterogeneity can be difficult. Anyway, small plots can be expected to be at least as efficient as large ones, regardless of the amount or type of variability. In other words, although the differences between the plot configurations studied may on other, more homogenous sites be less emphasized than here, the general tendency for decreasing efficiency along with increasing plot size is not likely to change.

The trait studied influences, of course, the importance of plot technique as well. Height, as an indicator of growth, is likely to be affected by environment more than many other traits of interest in breeding, e.g. branch angle or stem straightness. This assumption is supported by many heritability studies with several tree species, demonstrating that quality traits generally show higher heritability than growth traits (Pöykkö 1982, Cotterill 1987). Heritability and efficiency are closely related: the family mean heritability, ranging from 0 to 1, is calculated as the ratio of additive genetic variance and the variance of family means (Falconer 1981). Skrøppa (1987) concluded that increasing the number of replications at the expense of plot size is especially profitable when the heritability of a trait is low. It must be recognized, however, that the heritability estimates themselves may be influenced by the plot design; sacrificing sufficient replication in favour of larger plots may decrease heritability because of the poor control over field variability.

Compared to the question of plot size, plot shape has attracted less attention in the literature. Empirical studies with forest trees have indicated that the orientation of the plots in the field is far more important than their shape (Wright and Freeland 1960, Conkle 1963). This especially concerns long row plots. Smith (1938) stated that they can be either more or less effective than square plots, depending on their elongation in the field: plots directed across, rather than parallel to the major environmental gradient may be very inefficient in reducing the variation, contributing rather to the treatment-by-block variance as discussed earlier.

Together with the statistical aspect considered here, other factors influence the decision concerning the plot type. The intensity of intergenotypic competition after crown closure, for instance, varies considerably by plot size, decreasing as the plot size increases. Whether competition between test entities – or the lack of it – is a disadvantage or not, depends largely on the purpose of the test. This question has been discussed in numerous studies and has not yet been fully resolved. Foster (1989) emphasized that the interactions between trees in the experiment should mimic those in the deployment environment, i.e. in artificially regenerated stands. Considering that production plantations in Finland are established with bulk seed from seed orchards, thus having a random genetic structure, the trees in progeny tests should be subjected to competition with unrelated neighbours. This is an objective that is best achieved with single tree or non-contiguous plots. Under such conditions the correlation between progeny test and stand performance can be expected to be the highest. With large plots the sample of between-family competitive interactions is much smaller than that with small plots (Libby and Cockerham 1980), which may lead to inefficient selection due to biased evaluation of family performances. Large contiguous plots are more justifiable for testing provenances since these are seldom planted in mixtures, and for clonal tests provided that the clones are to utilized as monoclonal plantations. Nevertheless, a single tree design can be useful in the juvenile screening of provenances or clones before the subsequent yield testing with large plots (Shiue and Pauley 1961).

Acknowledgements: Dr. Pertti Pulkkinen, Lic.Sc. Jouni Mikola, M.Sc. Tapani Pöykkö and the two referees, Dr. Claire Williams and Dr. Dag Lindgren are thanked for their valuable comments on the manuscript. Lic.Sc. John Derome is thanked for his revision of the language.

References

Cochran, W. 1937. Catalogue of uniformity trial data. Royal Statistical Society Journal 4: 233–255.

Conkle, M. 1963. The determination of experimental plot size and shape in loblolly and slash pines. North Carolina State University, School of Forestry, Tech. Rep. 17: 1–51.

Correll, R. 1978. The optimal plot size for Pinus radiata establishment trials. Ph.D. thesis. South Australian Institute of Technology. 40 p.

— & Cellier, K. 1987. Effects of plot size, block size and buffer rows on the precision of forestry trials. Aust. For. Res. 17: 11–18.

Cotterill, P. 1987. On estimating heritability according to practical applications. Silvae Genetica 36(1): 46–48.

— & James, J. 1984. Number of offspring and plot sizes required for progeny testing. Silvae Genetica 33(6): 203–208.

Falconer, D. 1981. Introduction to quantitative genetics. 2nd ed. Longman Inc., New York. 340 p.

Foster, G. 1989. Inter-genotypic competition in forest trees and its impact on realized gain from family selections. 20th Southern Forest Tree Improvement Conference, Charleston, South Carolina.

Johnstone, R. & Samuel, C. 1974. Experimental design for forest tree progeny tests with particular reference to plot size and shape. Proceedings of the IUFRO Joint Meeting of Working Parties on Population and Ecological Genetics, Breeding Theory and Progeny Testing, Stockholm 1974: 357–370.

Lambeth, C., Gladstone, W. & Stonecypher, R. 1983. Statistical efficiency of row and noncontiguous family plots in genetic tests of loblolly pine. Silvae Genetica 32 (1–2): 24–28.

Libby, W. & Cockerham, C. 1980. Random non-contiguous plots in interlocking field layouts. Silvae Genetica 29(5–6): 183–190.

Loo-Dinkins, J. & Tauer, C. 1987. Statistical efficiency of six progeny tests field designs on three loblolly pine (Pinus taeda L.) site types. Can. J. For. Res. 17: 1066–1070.

Metsänjalostuksen koeviljelysluettelo. 1991. Metsäntutkimuslaitoksen tiedonantoja 396. 160 p.

Pöykkö, T. 1982. Genetic variation in quality characters of Scots pine. Silva Fennica 16(2): 135–140.

SAS User's Guide: Statistics, Version 5. 1985. SAS Institute Inc., Cary, NC. 817–834.

Shiue, C. & Pauley, S. 1961. Some considerations on the statistical design for provenance and progeny tests in tree improvement programs. Forest Sci. 7: 116–122.

Skrøppa, T. 1987. Statistical efficiencies of experimental designs in progeny testing. A manuscript of a talk given at the Nordic postgraduate course in plant breeding, Honne, January 1987. 15 p.

Smith, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. Journ. Agric. Sci. 28: 1–23.

Wright, W. & Freeland, F. 1960. Plot size and experimental efficiency in forest genetic research. Michigan State University, Agricultural Exp. Stat. Dept. of Forestry. Technical Bulletin 280: 1–28.

Total of 19 references

II

Within-Plot Subsampling of Trees for Assessment in Progeny Trials of Scots Pine

Matti Haapanen

Haapanen, M. 1995. Within-plot subsampling of trees for assessment in progeny trials of Scots pine. Silva Fennica 29(1): 13–19.

Tree height data from 33 progeny trials of Scots pine (*Pinus sylvestris* L.) were used to determine the effect of within-plot subsampling on the magnitude of statistically detectable differences between families, family heritability and correlation of family means based on different sample sizes. The results indicated that, in trials established with a standard plot configuration of 25 trees per plot, measuring only 10–15 trees gives nearly the same precision as with assessment of all the plot trees. Even as few as 4–6 trees assessed per plot may constitute a sufficient sample if families or parental trees of extreme performance are being selected. Trials established with non-contiguous plots were found to be more efficient than those established using multiple-tree contiguous plots.

Keywords non-contiguous plots, plot size, progeny testing, sampling, statistical analysis, efficiency, *Pinus sylvestris.*
Author's address Finnish Forest Research Institute, Box 18, FIN-01301 Vantaa, Finland
Fax +358 0 8570 5711 **E-mail** matti.haapanen@metla.fi
Accepted March 1, 1995

1 Introduction

The genetic testing of forest trees is normally carried out in a large number of field experiments, the establishment, maintenance and measurement of which all represent a considerable financial burden. To reduce the costs incurred in such operations, the experiments should be as small as possible with regard to the desired precision (Lambeth et al. 1983). In practice, however, the ideal size of an experiment is difficult to achieve. It is, of course, important to ensure that the number of replicates is sufficient to reveal meaningful differences between genetic entries. On the other hand, abundant replication, possibly accompanied by inefficiently large plots, easily results in the excessive use of plant material and land. This, in turn, inflates the management and measurement costs.

A part of the extraneous costs arising from

poor experimental design can be eliminated by careful planning of the measurement strategy. Sampling within experimental units is an option worth considering, especially when the units are contiguous (trees belonging to the same unit are planted adjacent to each other) and relatively large multiple-tree plots. This applies to Finland, where most of the progeny trials have been established using square-shaped plots of 25 trees (Mikola 1985, Haapanen 1992). Trees within contiguous plots are generally positively correlated as a result of the partially common microenvironment (e.g. Smith 1938, Hühn 1970). Environmental correlation causes the relative information obtained per tree to decrease along with increasing plot size (Conkle 1963). Consequently, only a subset of all the surviving trees in a large plot needs to be assessed in order to obtain a reasonable precision. This will naturally reduce the total number of trees that have to be measured.

The obvious question that arises here is, how far can one go in reducing the sample size without raising the experimental error above the acceptable limit? To study this, one approach is to generate within-plot samples of the tree records of various size from existing data sets, perform series of analysis of variance, and observe how the residual variance changes (Lee 1974 and 1983, Stevenson and Savill 1976, Correll 1978). This is referred to as a subsampling investigation (Correll 1978, Snowdon and Waring 1982). The size of a subsample is considered representative at the point where an increase in the number of trees sampled within plots no longer provides a significant reduction in the residual variance and consequently, in the chosen measure of efficiency. Such often used statistics include e.g., the magnitude of the difference between two family means found to be statistically significant (*least significant difference*) (Cochran and Cox 1957) and correlation between entry (family, provenance etc.) means based on samples of different size (Lee 1974, Lee 1983, Kung 1977, Cotterill and James 1984).

The efficiency of genetic testing can be further improved by selecting among the trials to be measured. The precision of genetic field trials generally varies considerably due to differences in microsite variability, site preparation, efficiency of blocking etc. Thus, it would be important for a tree breeder to be able to identify and reject those trials that have the least value as a source of genetic information. Estimates on the amount of residual variation from earlier data, if available, undoubtedly provide the most reliable basis for this sort of screening. In addition, information on the lay-out of experiments might also be used (see e.g. Lee 1983, McCutchan et al. 1989).

The objective of this study was to determine the effect of within-plot subsampling on the statistical efficiency in a representative set of progeny trials, as regards the measurement age, plot configuration and number of replication in Finnish conditions. Another object of interest was to study to which degree the efficiency can be predicted by factors related to experimental design. such as block size and number of blocks.

2 Material and Methods

The study material consisted of 33 sets of tree height data from the same number of Scots pine (*Pinus sylvestris* L.) progeny trials (Fig. 1). The most recent data available from every trial, the age of which ranged from 5 to 20 years (median 10 years), was used in the analyses. All surviving trees were measured for total height in each trial.

The main body of the data, 30 trials, represented a random sample of the total number of 784 Scots pine progeny trials established by Finnish Forest Research Institute since the 50's. The experimental design of these trials typically involved 4–6 randomized complete blocks. Trees were most commonly planted at 5 × 5 position in 25-tree block plots (Fig. 2). The standard planting distance was 2.0 m × 2.5 m. Most entries were open-pollinated progeny of selected plus-trees, usually accompanied by 4–8 control seedlots per trial. The remaining three trials were established using non-contiguous plots (Libby and Cockerham 1980), and included subjectively to study the effect of plot configuration on the experimental efficiency. For more detailed information on the trials (Fig. 1) see Pajamäki and Karvinen (1991).

Randomized within-plot subsampling was sim-

Fig. 1. Location of the progeny trials in Finland.

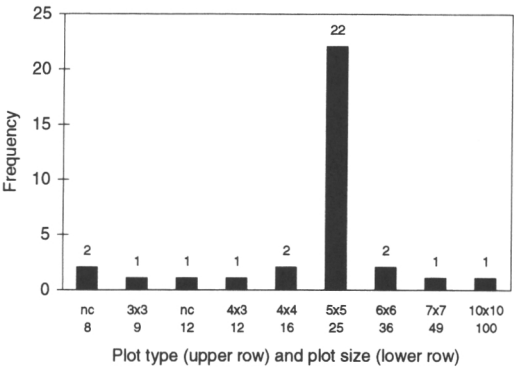

Fig. 2. Frequency of different plot types and plot sizes (number of trees planted per plot) in the studied data set consisting of 33 progeny trials of Scots pine ('nc' = non-contiguous plot design).

ulated using a specifically designed computer program. The sample sizes simulated were 2, 4, 6, 8, 10, 15 and 20 trees per plot. If the average number of measured trees per plot in any trial was less than 15, the 15 and 20 tree-samples were omitted. For each sample size, the sampling procedure was replicated 15 (with samples of 2 and 4 trees) or 10 (with samples of 6, 8, 10, 15 and 20 trees) times per trial. The total number of samples was 2339.

All the samples were analysed on the basis of plot means. A plot mean is usually considered to be the relevant unit of observation when statistical analyses are conducted on progeny trial data. This is due to the genetic and environmental correlation between trees belonging to the same contiguous plot, which violates the assumption of the independence of the within-plot error terms.

The effect of subsampling on statistical efficiency was determined by using three criteria. The least significant difference (LSD, Eq. 1) measures the smallest detectable difference between any two treatment (here family) means at a chosen level of significance (Cochran and Cox 1957):

$$LSD = \sqrt{\frac{2(t_a + t_b)^2 \, s^2_{resid}}{b}} \qquad (1)$$

where b in denotes the number of blocks. The s^2_{resid} is the residual variance, obtained from an analysis in which variance components due to family and block effects were subtracted from the total variance of plot means. The MIXED procedure (method REML) of the SAS statistical package (SAS Technical Report 1992) was used in the estimation of the variance components. The values of Student's t-distribution were calculated using 2 (b – 1) degrees of freedom and 0.05 (t_a) and 0.20 (t_b) probability levels for errors of type I and II, respectively. To permit comparison of the LSD values between different trials, they were converted to a percentage scale (% of the experiment mean).

The two other measures studied were: 2) family heritability h^2_{fam} (Falconer 1981), which is the ratio of the among-family (s^2_f) variance compo-

nent to the total phenotypic variance of family means (Eq. 2), and 3) the correlation between family means based on a complete sample and a subsample of trees (see Kung 1977).

$$h^2_{fam} = \frac{s^2_f}{s^2_f + s^2_{resid} / b} \quad (2)$$

The number of trees sampled per plot affects the chosen efficiency criteria in two ways: At first, residual variance increases along with decreasing sample size, since s^2_{resid} is actually the sum of the between-plot variance component and the n^{th} part of the within-plot variance (Eq. 3) (where n is the average number of trees measured per plot). Secondly, reducing sample size increases the sampling variance of the variance component estimates.

$$s^2_{resid} = s^2_{plot} + s^2_{within-plot} / n \quad (3)$$

The variability of the estimates of efficiency at different sample sizes was measured by the coefficient of among-sample variation (standard deviation in a set of 10 or 15 replicated samples divided by the mean of the sample estimates), which was averaged across the 33 trials.

The relationship between block size and number of blocks, and the estimated statistical efficiency (the LSD and h^2_{fam}) was studied using Pearson's correlation analysis.

3 Results

The efficiency parameters responded to subsampling most significantly when the sample size was increased from 2 to 6 trees. When the number of trees sampled per plot exceeded 10, the average LSD values, for example, improved by 1–3 % (Fig. 3). The curves for family heritability showed a similar, although opposite trend (Fig. 4). The heritability and LSD values showed higher variability among trials than the correlations. The correlations between the family means based on all the plot trees and subsamples of different size were high in all trials (r 0.80), even with sample sizes as small as 4–6 trees per plot (Fig. 5).

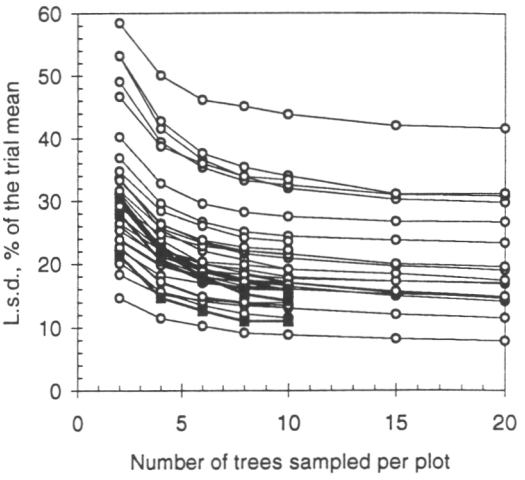

Fig. 3. Average least significant differences (LSD) at different sample sizes in 33 progeny trials. The probabilities of making an error of type I (false rejection) and II (false acceptance) were set to 0.05 and 0.20. The non-contiguous plot trials are denoted by the filled squares, and the others by open circles.

The coefficient of among-sample (residual) variation decreased along with increasing sample size. Family heritability was associated with a high coefficient of variation, especially at small sample sizes (Fig. 6). This was also reflected in the mean heritability values, which did not increase regularly as a function of the sample size in a few cases (Fig. 4).

Independently of the measure of efficiency, the three non-contiguous plot trials showed as high efficiency as the best contiguous plot trials although the number of trees per family planted in these trials was considerably smaller.

The number of blocks established per trial was associated with lower LSD values, and thus, with higher efficiency ($r = -0.43$, $p < 0.02$). The block size, on the other hand, was unfavourably related with the LSD ($r = 0.32$, $p < 0.18$). The trials with the largest blocks showed the poorest efficiency, especially when the number of blocks was small (less than 5). The respective correlations with family heritability values were non-significant (No. of blocks: $r = 0.06$, $p < 0.74$; Block size: $r = -0.13$, $p < 0.46$).

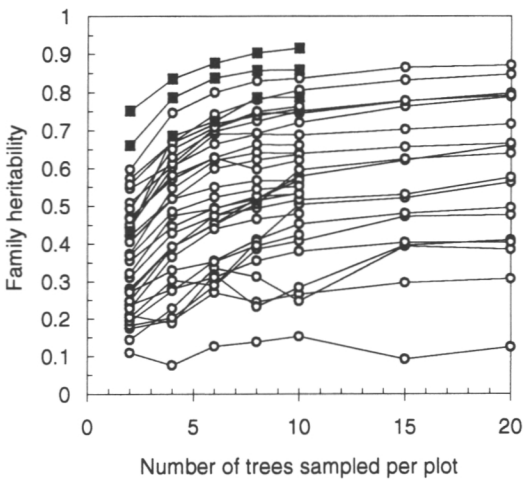

Fig. 4. Average family heritability estimates at different sample sizes in 33 progeny trials. The non-contiguous plot trials are denoted by the filled squares, and the others by open circles.

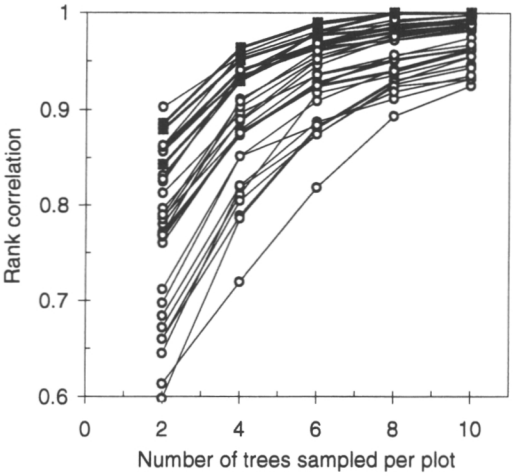

Fig. 5. Average correlations between family means based on complete assessment of all plot trees and a subsample of different size.

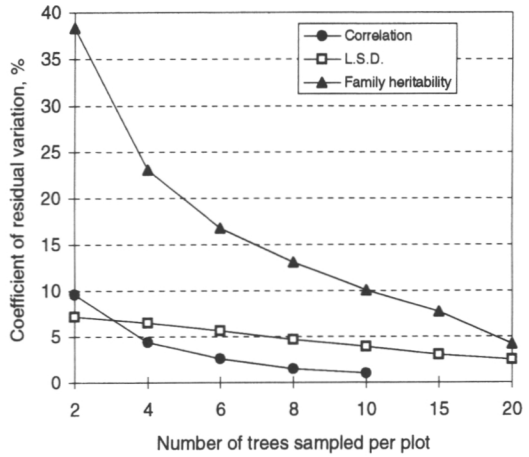

Fig. 6. Average coefficients of among-sample (residual) variation for 1) least significant difference, 2) family heritability and 3) correlation between family means based on complete plot assessment and a subsample.

4 Discussion

The partial assessment of plots seems to be worthwhile especially in the trials established with large multiple-tree plots, as they have a low efficiency on a per tree basis. The results obtained here indicate that measuring only 10–15 trees in the 25-tree block plots is likely to give statistical precision (LSD) that is not significantly different from the costly alternative of complete plot assessment. This is in accordance with the findings of e.g., Stevenson and Savill (1976), who sampled trees from 36-tree plots in two Sitka spruce (*Picea sitchensis* (Bong.) Carriere) experiments. They found that little additional information was achieved by measuring height and girth at breast height on more than 20 trees per plot, 12–16 trees per plot giving results with sufficient precision in most cases. The total savings in costs resulting from reducing the number of measured trees are, of course, primarily dependent on the extent of the testing programme. Taking as an example the number of offspring trees planted in Scots pine progeny trials in Finland since the early 50's, i.e. nearly six million, the use of complete plot assessment as a standard method is likely to lead to substantial accumulated losses in terms of information obtained per time and labour spent.

The high correlations between pairs of family means based on complete samples and subsamples of different size as obtained in this study (Fig. 5), indicate that even as few as 4–6 trees

per plot give reliable information for selecting families or parent trees of extreme performance. However, if the purpose of testing is not just to select the superior families or to cull the poorest ones, more comprehensive sampling will be necessary. Another situation to which the conclusions drawn here do not apply is when the main objective of the assessment is to provide information to be used in selecting the best individuals within families. In that case all the available trees should preferably be measured to make the selection intensity as high as possible. One way to avoid this dilemma is to design different trials for different selection objectives: single-tree plots for family selection and large block plots for selecting sibs in a common environment. (e.g. Lambeth 1986, Loo-Dinkins and Tauer 1990).

The coefficients of residual variation (Fig. 6) indicate the degree of variation among estimates computed from replicated samples (Figs. 3, 4, and 5). In large plots the subsamples of 10–15 trees seem to give both sufficient precision (on the average) and adequately good protection against strikingly poor estimates obtained by chance. It should be noted, however, that all the CV's are more or less underestimated since the individual samples were not mutually independent (due to the infinite number of trees per plot, the different samples were, in part, composed of the same trees).

The reliability of the family heritability estimates was significantly poorer than that of the family mean correlations and LSDs, particularly when the sample size was small. This was obviously due to the fact that the residual variance of heritability (see Eq. 2) is increased by the sampling error of both the residual variance component and the family variance component.

The efficiency of the sampled trials measured in terms of absolute LSD values was rather low. This was the case even with the results from the largest samples. It can be asked whether the ability to detect height differences of 15 to 30 % of the general mean is satisfactory for a tree breeder, if we consider that true genetic differences between families are often much smaller. To conclude, the experiments analysed in this study would have needed greatly increased replication to be truly effective at the chosen levels of error. The increased replication as such will not, however, solve the problem of the inefficiency of large contiguous plots. On the contrary, the efficiency in terms of information per total number of planted trees may even decrease. The results indicate that the non-contiguous arrangement of plot trees which eliminates a large proportion of the residual variance due to spatial correlation of related individuals, (Lambeth et al. 1983), can be recommended for future trials. The results of this study strongly suggest that non-contiguous plot trials need a considerably smaller number of trees per family to yield a statistical precision equal to that of the ordinary trials established using large contiguous plots. More data are, however, needed to give well-founded numerical guidelines for the appropriate family size, number of replication etc. for the former type of trial.

Block size and number of blocks together appeared to indicate the statistical efficiency (LSD) of a trial to some extent, if height is taken as the target trait. Even though the association was not strong, there was a clearly detectable trend: those trials having both the smallest number of blocks and the largest block size proved to be of the least value. Since there is no sense in measuring trials that will probably never give sufficient information, these two variables might be used as a rough tool in differentiating among the existing trials and culling the poorest ones from measurement programs. Further savings might also be attained by omitting some of the families from the measurement on the basis of *a priori* information.

Acknowledgements

I wish to thank the following persons for reading the manuscript and providing many worthy comments on it: Lic. For. Jouni Mikola, Dr. Pertti Pulkkinen, Dr. Tapani Pöykkö and Dr. Pirkko Velling, and especially the official referees, Dr. Steen Magnussen and Dr. Pasi Puttonen. I also thank Mr. John Derome, Lic. For., for his revision of the language.

References

Cochran, W. & Cox, G. 1957. Experimental designs. 2nd ed. Wiley publications in statistics. 611 p.

Conkle, M. 1963. The determination of experimental plot size and shape in loblolly and slash pines. North Carolina State University, School of Forestry, Technical Report 17: 1–51.

Correll, R. 1978. The optimal plot size for Pinus radiata establishment trials. Ph.D. thesis. South Australian Institute of Technology. 40 p.

Cotterill, P. & James, J. 1984. Number of offspring and plot sizes required for progeny testing. Silvae Genetica 33(6): 203–208.

Falconer, D. 1981. Introduction to quantitative genetics. 2nd ed. Longman Inc., New York.

Haapanen, M. 1992. Effect of plot size and shape on the efficiency of progeny tests. Silva Fennica 26(4): 201–209.

Hühn, M. 1970. Untersuchungen zur Konkurrenz zwischen verschiedenen Genotypen in Pflanzenbeständen. IV. Probleme der optimalen Parzellengrösche in Feldversuchen. [Summary in English]. Silvae Genetica 19(4): 151–163.

Kung, F. 1977. Adjustment and interpretation of progeny tests when only the best tree in each plot is measured. Silvae Genetica 26(2–3): 117–119.

Lambeth, C. 1986. Effective genetic testing – practical and theoretical considerations. Proc. of IUFRO Meeting on Breeding Theory, Progeny Testing and Seed Orchards, Williamsburg, Va. p. 24–33.

— , Gladstone, W. & Stonecypher, R. 1983. Statistical efficiency of row and noncontiguous family plots in genetic tests of loblolly pine. Silvae Genetica 32(1–2): 24–28.

Lee, C. 1974. Effect of partial measurement on statistical precision and efficiency in a nursery study of Pinus ponderosa. Silvae Genetica 23(4): 124–126.

— 1983. Statistical efficiency varies with plot size, number of replications and seedlots sampled. Silvae Genetica 32(1–2): 49–52.

Libby, W. & Cockerham, C. 1980. Random non-contiguous plots in interlocking field layouts. Silvae Genetica 29(5–6): 183–190.

Loo-Dinkins, J., Tauer, C. & Lambeth, C. 1990. Selection system efficiencies for computer simulated progeny test field designs in loblolly pine. Theoretical & Applied Genetics 79: 89–96.

McCutchan, B., Namkoong, G. & Giesbrecht, F. 1989. Design efficiencies with planned and unplanned unbalance for estimating heritability in forestry. Forest Science 35(3): 801–815.

Mikola, J. 1985. Methods used for the genetic evaluation of tree breeding material in Finland. In: Tigerstedt, P., Puttonen, P. & Koski, V. (eds.). Crop physiology of forest trees. University of Helsinki, Department of Plant Breeding. p. 138–143.

SAS® technical report P-229. SAS/STAT® software: Changes and enhancements, release 6.07. 1992. SAS Institute Inc., Cary, NC, USA.

Smith, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. Journal of Agricultural Science 28: 1–23.

Snowdon, P. & Waring, H. 1982. Between-tree and between-plot variations of nutrient levels in Pinus radiata foliage and their implications for field sampling intensity. Australian Forest Research 13: 45–56.

Stevenson, W. & Savill P. 1976. Sampling in forest experiments. Commonwealth Forestry Review 55(4): 331–334.

Total of 20 references

III

Impact of Family-by-trial Interaction on the Utility of Progeny Testing Methods for Scots Pine

By M. HAAPANEN[1])

(Received 7th February 1996)

Summary

The magnitude of family-by-trial interaction in the progeny testing of Scots pine (*Pinus sylvestris* L.) was studied by estimating genetic and rank correlations between 40 pairs of progeny trials. The data originated from 20 conventional long-term forestry trials and 15 test orchards, all assessed for total tree height at the age of 10 years. The study material consisted of half-sib progenies of the first generation Scots pine plus trees. Family-by-trial interaction was found to be at least moderate but its pattern was largely unpredictable. None of the studied environmental or design variables (the difference in trial mean height, survival and planting density) explained a significant portion of the variation in across-site correlations. Family by field test design interaction was also absent, i.e. the across-site correlations were rather similar in the groups of test orchards and forestry trials. Site type strongly affected mean height and survival, as well as the heritabilities. The highest estimates of family heritability were derived from test orchards on agricultural land. On forest soil, the test orchards and forestry trials performed comparably, but clearly poorer than the test orchards on agricultural land. The results suggest that the efficient discrimination of genetic differences achieved through intensive field testing on homogeneous sites overrules the interaction bias due to the disparity between test orchard and forestry conditions, which appeared to be negligible.

Key words: Pinus sylvestris, genotype-environment interaction, genetic correlation, heritability, progeny testing, selection, experimental design.

FDC: 232.11; 165.3; 165.5; 165.4; 174.7 *Pinus sylvestris.*

Introduction

The genetic testing of Scots pine (*Pinus sylvestris* L.) in Finland is largely based on information from long-term trials that are managed analogously to operational reforestation sites. During the past 2 decades, however, the use of accelerated trial procedures has become more popular (MIKOLA, 1985; PAJAMÄKI and KARVINEN, 1995). The concept of test orchard, synonymous to 'farm-field trial', was suggested by TIGERSTEDT (1973) to describe a field testing method aimed at rapid and cost-effective screening of genetic entries. Test orchards are regularly established as high density stands (up to 10000 trees/ha), usually on uniform and fertile sites, such as abandoned agricultural land. Weed control and site preparation are commonly used to further reduce the microsite variability. In addition, the trials are occasionally fenced to prevent damage due to browsing animals (MIKOLA, 1985). The purpose of all these measures is 1) to reduce the experimental error, 2) to enhance the early manifestation of genetic differences, and 3) to accelerate the growth of trees and, consequently, to minimize the time period from planting to the onset of root and crown competition. Selection for test orchard performance is, in turn, expected to provide higher genetic gains per unit time than selection carried out in normal forestry trials. In Finland, test orchards are intended for selection for height at a target age of 10 to 15 years (MIKOLA, 1985).

The various hypotheses presented to date about the superiority of accelerated test methods have remained unverified. Furthermore, the obvious disparity between test orchard and operational planting conditions has given rise to concern among tree breeders. If correlation between selection and deployment environments is decreased, the additional selection gain expected from accelerate field trials could be seriously diminished (LINDGREN, 1984; HODGE and WHITE, 1992). In the worst of scenarios, the relative ranking of genotypes in test orchard conditions would be reversed on conventional plantations. The degree to which different testing procedures actually generate genotype-environment interaction (GEI) is poorly known because the subject has received surprisingly little attention in the forestry literature. Many of the interaction studies have focused on determining safe ranges for seed transfer or evaluating the need for delineated breeding zones (MATHESON and RAYMOND, 1984; CARSON, 1990; JOHNSON and BURDON, 1990; PEDERICK, 1990; JOHNSON, 1992). These studies have usually refered to situations where the genetic entries or test environments, or both, differ substantially (e.g. provenances, regions) and interactions are fairly repeatable. In progeny testing, on the other hand, the material being tested is relatively homogeneous, and intended to be deployed within a predefined target environment. In spite of this, family by microsite interaction can be as high as family by macrosite interaction (MATHESON and COTTERILL, 1990). In the latter case, high GEI may be exploited by breeding for specific environments. In the first case, this option is not available; instead, the main concern is the efficiency of indirect selection: How should tree breeders respond to data from test orchards and other unconventional test environments?

The impact of GEI is most conveniently measured in terms of genetic correlation and correlated response for selection. The idea of treating the expression of a trait assessed at 2 sites as 2 distinct traits, and estimating their correlation, was originally presented by FALCONER (1952). *Type-b* genetic correlation (r_B) was first introduced in forestry by BURDON (1977) who claimed that, in the genetic testing of forest trees, the stability of test sites is essentially more important than that of genotypes. Interpretation of r_B in terms of GEI is selfevident, since any degradation of the coefficient of correlation from unity arises from the inability of genotypes to perform equally at twð sites. Genetic across-site correlations possess some favorable statistical properties as compared to ANOVA based estimates of GEI. For instance, genetic correlations are robust against heterogenous site variances, and can be effectively incorporated into selection indices (BURDON,1977; WHITE and HODGE, 1991) and formulae predicting the response to indirect selection (FALCONER, 1981). During the last few years, the use of r_B has become established in forest genetics literature as a measure of stability (MATHESON and RAYMOND, 1984; NIENSTAEDT and RIEMENSCHNEIDER, 1985; CARSON, 1990; JOHNSON and BURDON, 1990; JOHNSON, 1992; HODGE and WHITE, 1992; LAMBETH *et al.*, 1994).

[1]) Finnish Forest Research Institute, Box 18, FIN-01301 Vantaa, Finland, e-mail: matti.haapanen@metla.fi

The principal objective of this paper was to estimate the importance of family-trial interaction in the Finnish Scots pine breeding programme. This was done by studying the magnitude and variation of across-site correlations between families in progeny trials representing different experimental designs and site types. Identification of the factors generating the interactions was also addressed.

Material and Methods

Material

The study material consisted of 10 test orchard trials and 20 long-term forestry trials of Scots pine. The trials were planted in central and southern Finland between 1977 and 1981, and measured for tree height at the age of 10 years. The 30 trials were distributed into 8 series, each consisting of 2 to 6 replicated trials, of which at least one trial was established as a test orchard. Four of the test orchard trials were located on abandoned agricultural land, whereas the rest of the trials were situated on forest sites of varying fertility. The initial spacing ranged from 2 m x 2 m in the forestry trials (2500 trees per hectare) to 0.75 m x 1.5 m or 1 m x 1 m in the test orchards (8888 and 10000 trees per hectare, respectively). All the trials consisted of 4 to 6 randomised complete blocks (*Table 1*).

Most of the entries represented windpollinated offspring of first generation plus trees. Family members were assumed to be true half-sibs. Other types of entry, such as a few full-sib families and standard seed lots (4 to 10 per trial), were excluded from the more detailed analyses. In most series, the family

Table 1. – Description of the Scots pine progeny trials. The first value in the 'Entries' column indicates the total number of entries originally planted, and the second value the number of entries left after the removal of standard seedlots. *Trials No. 573/1 and 573/2 have the same entries as the trials in the 572 series, and are referred to as trials 572/3 and 572/4 in this study.

Trial	Entries	Blocks	Trees/ plot	Trees/ entry	Area, ha	Planting density, trees / ha	Mean height,dm (age 10)	Mean survival, % (age 10)	Trial type	Site type
572/1	45/31	7	18	129	0.64	8888	45.8	98.9	Test orchard	Agricultural land
572/2	38/30	4	16	63	0.97	2500	26.1	60.2	Forestry trial	Agricultural land
*572/3	45/31	10	16	160	2.88	2500	21.6	68.3	Forestry trial	Dryish heath forest land
*572/4	44/31	10	16	158	2.80	2500	24.8	47.0	Forestry trial	Moist forest land
624/1	50/41	5	25	125	0.70	8888	32.0	91.1	Test orchard	Agricultural land
624/2	50/41	5	25	125	2.50	2500	32.7	79.7	Forestry trial	Agricultural land
624/3	50/41	5	25	122	2.45	2500	13.6	61.8	Forestry trial	Dryish heath forest land
698/1	33/23	4	25	86	0.32	8888	43.8	88.2	Test orchard	Agricultural land
698/2	31/22	4	25	97	1.20	2500	22.3	84.8	Forestry trial	Moist forest land
699/1	69/59	6	25	150	1.06	10000	18.3	27.3	Test orchard	Dryish heath forest land
699/3	68/59	6	25	149	4.06	2500	27.1	53.4	Forestry trial	Dryish heath forest land
739/2	88/88	6	25	150	1.49	8888	48.9	71.0	Test orchard	Agricultural land
739/5	46/42	6	25	150	2.76	2500	25.6	44.9	Forestry trial	Dryish heath forest land
740/1	63/53	6	25	149	1.06	8888	20.4	46.5	Test orchard	Dryish heath forest land
740/2	64/54	6	25	150	0.96	10000	19.8	39.1	Test orchard	Dryish heath forest land
740/3	33/28	6	25	150	1.98	2500	25.5	63.6	Forestry trial	Moist forest land
740/4	33/28	6	25	150	1.98	2500	28.9	61.6	Forestry trial	Dryish heath forest land
740/5	29/24	6	25	150	1.74	2500	21.1	57.1	Forestry trial	Dryish heath forest land
740/6	30/25	6	25	150	1.80	2500	23.2	46.6	Forestry trial	Dryish heath forest land
741/1	61/51	6	25	141	0.98	8888	19.1	54.2	Test orchard	Dryish heath forest land
741/2	60/51	6	25	149	0.90	10000	17.4	20.0	Test orchard	Dryish heath forest land
741/3	27/22	6	25	147	1.59	2500	22.5	51.1	Forestry trial	Moist forest land
741/4	26/21	6	25	139	1.45	2500	23.5	55.8	Forestry trial	Moist forest land
741/5	33/28	6	25	150	1.98	2500	25.7	59.9	Forestry trial	Dryish heath forest land
741/6	32/27	6	25	150	1.92	2500	20.9	34.7	Forestry trial	Dryish heath forest land
742/2	66/56	6	25	150	0.99	10000	18.8	46.9	Test orchard	Dryish heath forest land
742/3	32/27	6	25	150	1.92	2500	30.6	69.1	Forestry trial	Dryish heath forest land
742/4	30/25	6	25	135	1.62	2500	28.2	69.5	Forestry trial	Moist forest land
742/5	34/29	6	25	150	2.04	2500	26.3	61.3	Forestry trial	Dryish heath forest land
742/6	34/29	6	25	146	1.98	2500	21.5	40.6	Forestry trial	Dryish heath forest land

composition was identical across the trial replicates. Exceptions were series Nos. 740, 741 and 742, in which only the test orchard replicates consisted of all the families tested, whereas each of the remaining forestry trials contained only half of the families (*Table 1*). For this reason, the analyses of across-trial correlations in these series were defective.

Analysis

The total height of all the living and healthy trees in each trial was measured. Plot mean of the measured trees (y_{ij}) was the basic observation unit used in the analyses. In order to estimate the genetic parameters, total variance was partitioned into statistical variance components in each trial, as well as across all trials within each series using the MIXED procedure in SAS/STAT package (SAS Inc., 1992). The linear models used in the single-site (Eq. 1) and across-site (Eq. 2) analyses where

$$y_{ij} = \mu + f_i + b_j + e_{ij} \qquad [1]$$

$$y_{ijk} = \mu + f_i + b_{j(k)} + t_k + ft_{ik} + e_{ijk} \qquad [2]$$

f_i, $b_{j(k)}$, t_k and ft_{ik} are the effects of family, block, trial site and family by trial interaction, respectively, and e_{ijk} denotes the random plot error. The respective variance components were σ^2_f, σ^2_b, σ^2_t, σ^2_{ft}, and σ^2_e.

The family heritability was calculated for each site as (r denotes the harmonic mean number of blocks):

$$h^2_f = \sigma^2_f \; / \; (\sigma^2_f + \sigma^2_e \; r^{-1}) \qquad [3]$$

Coefficients of family and family mean variation (cv_f, cv_F) were obtained by dividing the square roots of the family variance component and the phenotypic variance of family means, respectively, by the trial mean height (X):

$$cv_f = 100 \; s_f \; / \; X \qquad [4]$$

$$cv_F = 100 \; [\; \sigma_f + (\; \sigma_e \; r^{-1/2} \;) \;] \; / \; x \qquad [5]$$

Genetic correlation was computed between all pairs of replicated trials by dividing the phenotypic correlation of family means in 2 trials ($r_{F(ii')}$) by the geometric mean of the respective single-site family heritabilities (Eq. 6) (BURDON, 1977). The phenotypic type-b family correlations, as well as SPEARMAN rank correlations (r_{rank}), were computed using the CORR procedure of SAS (SAS Inc., 1985). Based on the fact that only genetic covariance contributes to the correlation of phenotypic observations across sites, this method effectively by-passes the need to partition the total covariance into its components through a separate cross-product analysis of variance (BECKER, 1984), a feature that is not directly implemented in most statistical packages.

$$r_B = r_{F(ii')} \; / \; (\; h_{f(i)} \; h_{f(i')}) \qquad [6]$$

PEARSON's product-moment correlation analysis and one-way analysis of variance were used to test whether any relationship existed between r_B and r_{rank} and a few environmental and experimental design variables.

Assuming homogeneity of the variances between trial replications, an alternative set of estimates of type-b genetic correlation (r_{B2}) was computed for each trial series using a formula from YAMADA (1962):

$$r_{B2} = \sigma^2_f \; / \; (\sigma^2_f + \sigma^2_{ft}) \qquad [7]$$

where σ^2_f and σ^2_{ft} are obtained from the across-site analysis.

The selection gain per intensity of unit standard deviation was estimated for both direct (Eq. 8) and indirect (Eq. 9) family selection. The direct gain, relative to trial mean height (μ_i), was obtained as (FALCONER, 1981)

$$G_i = 2 \; h^2_{f(i)} \; cv_{(i)} \qquad [8]$$

When the material selected in trial *i* is deployed in conditions identical to those in trial *i'*, the expected correlated response, relative to the direct gain estimated at the deployment site, becomes (BURDON, 1977):

$$G_{i'/i} = r_B \; h_{f(i)} \; / \; h_{f(i')} \qquad [9]$$

Results

The across-site correlations varied considerably among all the trial series. With a couple of exceptions, all the correlations were positive. Some of the genetic correlations exceeded unity, indicating large standard errors associated with the estimates. The unweighted mean of the 39 type-b genetic correlations was 0.61 (s.d. = 0.39). The genetic correlations based on YAMADA's equation (*Eq. 7*) were of the same magnitude (mean = 0.58). The family-by-trial interaction component was over 80 % of the family component of variance, as averaged over the eight across-site analyses (*Table 2*). The rank correlations were significantly lower than the genetic correlations (mean = 0.30) and less than half of them were statistically significant at the 5% risk level (*Table 3*).

The type-b correlations did not depend on differences between the trials in planting density (r with r_B = –0.09 n.s., r with r_{rank} = –0.21 n.s.), mean height (r with r_B = 0.02 n.s., r with r_{rank} = 0.16 n.s.) or survival percentage (r with r_B = –0.06 n.s., r with r_{rank} = 0.13 n.s.). Neither did the different types of paired trials (forestry trial x forestry trial *vs.* forestry trial x test orchard *vs.* test orchard x test orchard) show significantly different levels of correlation (1-way ANOVA obs. prob. = 0.97). The average genetic correlations for the pairs of forestry trials and pairs of test orchard and forestry trial, respectively, were similar (mean = 0.62 in the both cases). Exceptions were series No. 572 and 624, where a test orchard on agricultural land trial showed better genetic correlation with a replicate trial located on forest land than with another trial on agricultural land.

The estimates of family heritability were also variable between the sites, the maximum range (0.00 to 0.85) being found in series No. 698 (*Table 3*). The influence of site type on the heritability estimates was distinct. The highest family heritabilities were found among the four test orchard trials established on agricultural land (mean h^2_f = 0.83). In contrast, family heritability in the trials laid out on forest sites, including test orchards, was noticeably small (mean h^2_f = 0.50).

The agricultural land trials were distinguished by their excellent height growth which proved to be double that on forest sites. Respectively, the mean survival ranged from about

70% to 100% in agricultural land test orchards, and from 20% to 70% (*Table 1*) in the forest land trials. In the latter set of trials, the test orchards even showed poorer height growth and survival than the parallel forestry trials (*Table 1*). Overall, the test orchards established on forest soil were about as informative as the ordinary forestry trials.

The percentage expected gains for backward family selection, varying from 0 % to 12 %, did not show any clear trend. The highest absolute gains can, however, be expected from the agricultural land test orchards, as suggested by the family variance components (*Table 3*). The estimates of relative indirect gain also appeared to be variable and unpredictable.

Table 2. – Estimates of variance components from the across-site analyses (Eq. 2). Rel. FSI denotes the relative family-by-trial interaction, calculated as $\sigma^2_{ft}\ 100/\sigma^2_f$. The type-b genetic correlation (r_{B2}) was obtained from Eq. 7. Variance components significant at 5% level are underlined.

	REML Variance Components						
Series	Trial	Family	Block	Family x trial	Residual	Rel. FSI, %	r_{B2}
572	120.610	1.311	0.973	0.787	6.268	60.05	0.63
624	116.350	0.450	3.597	0.728	4.671	161.59	0.38
698	232.350	1.521	2.742	2.090	10.471	137.44	0.42
699	38.780	0.384	0.442	0.359	6.503	93.65	0.52
739	273.940	1.890	0.997	1.583	9.055	83.73	0.55
740	13.390	0.964	2.015	0.401	6.398	41.56	0.71
741	8.960	0.583	0.732	0.244	5.955	41.86	0.71
742	24.720	1.175	1.541	0.432	4.986	36.80	0.73
Mean						82.09	0.58

Table 3. – Left: Across-site genetic (upper triangular matrices) and rank correlations (lower triangular matrices) for height at age 10. The shadowed diagonal contains the average genetic correlations. In the calculation of means, estimates exceeding unity were truncated to 1. Middle: Direct selecetion gain, given as percentage of the trial mean height (Eq. 8) on diagonal. The off-diagonal elements are gains from indirect selection, relative to the direct selection gain obtained on site i' (Eq. 9). Right: Family component of variance (σ^2_f), single-site family heritability (h^2_f), coefficients of family (CV_f) and phenotypic family mean variation (CV_F). The underlined rank correlations and family variance components are significant at 5% risk level. The test ochards are distinguished by bold trial numbers.

	Type-b genetic and rank correlations						Direct and indirect selection gains									
	Site i'						Trial site i'									
Trial site i	/1	/2	/3	/4	/5	/6	/1	/2	/3	/4	/5	/6	σ^2_f	h^2_f	CV_f, %	CV_F,
572/1	0.71	0.59	0.69	0.85			7.73	0.65	0.72	0.91			3.80	0.79	4.26	4.86
572/2	0.45	0.55	0.45	0.61			0.54	9.33	0.43	0.60			2.12	0.66	5.58	7.04
572/3	0.39	0.20	0.64	0.78			0.66	0.48	10.45	0.80			1.70	0.73	6.05	7.15
572/4	0.54	0.40	0.51	0.75			0.79	0.62	0.76	7.58			1.17	0.69	4.37	5.47
624/1	0.58	0.39	0.76				8.05	0.60	1.02				1.92	0.86	4.33	4.66
624/2	0.25	0.19	-0.01				0.25	3.18	0.00				0.80	0.36	2.74	4.42
624/3	0.43	0.04	0.38				0.57	0.00	9.28				0.82	0.48	6.68	9.65
698/1	-	n/a					12.25	0.00					7.82	0.85	6.39	7.29
698/2	0.33	-					0.00	0.00					0.00	0.00	0.00	8.94
699/1	0.38	0.38					7.68	0.46					0.96	0.48	5.35	7.91
699/3	0.25	0.38					0.31	3.39					0.58	0.33	2.81	5.08
739/2	0.77	0.77					8.14	1.16					4.75	0.83	4.46	4.88
739/5	0.37	0.77					0.51	5.80					1.52	0.36	4.82	8.01
740/1	0.40	0.73	0.58	0.52	0.29	0.19	10.29	0.68	0.76	0.50	0.29	0.19	2.01	0.54	6.95	9.45
740/2	0.25	0.72	1.15	1.00	0.51	0.34	0.78	8.88	1.39	1.03	0.54	0.37	1.16	0.62	5.43	7.13
740/3	0.16	0.40	0.86	1.31	-	-	0.45	0.72	2.98	0.74	-	-	0.44	0.32	2.61	4.66
740/4	0.40	0.38	0.44	0.84	-	-	0.54	0.98	4.05	7.67	-	-	2.05	0.59	4.95	6.46
740/5	0.11	0.35	-	-	0.46	0.58	0.29	0.48	-	-	5.37	0.59	0.57	0.55	3.58	4.85
740/6	0.15	0.18	-	-	0.25	0.37	0.19	0.32	-	-	0.58	5.79	0.81	0.55	3.89	5.31
741/1	0.87	0.37	0.85	0.93	1.40	1.18	3.82	0.29	0.82	0.82	0.62	0.64	0.52	0.26	3.78	7.37
741/2	0.09	0.25	0.47	0.01	0.28	0.11	0.47	5.90	0.58	0.01	0.22	0.09	0.60	0.42	4.46	6.94
741/3	0.19	0.32	0.30	-0.33	-	-	0.89	0.38	3.11	-0.30	-	-	0.42	0.28	2.90	5.55
741/4	0.36	0.11	-0.08	0.20	-	-	1.06	0.01	-0.37	3.32	-	-	0.43	0.34	2.78	4.91
741/5	0.55	0.15	-	-	0.76	1.15	1.61	0.35	-	-	8.53	1.02	1.78	0.67	5.19	6.37
741/6	0.38	0.02	-	-	0.76	0.70	1.59	0.13	-	-	0.98	9.43	1.46	0.65	5.79	7.22
742/2		0.61	0.26	0.22	1.05	0.95		11.80	0.31	0.24	1.11	1.10	1.61	0.76	6.74	7.78
742/3		0.20	0.41	0.55	-	-		0.22	5.97	0.50	-	-	1.54	0.53	4.06	5.63
742/4		0.06	0.24	0.39	-	-		0.20	0.60	7.59	-	-	1.94	0.63	4.94	6.01
742/5		0.73	-	-	1.00	1.10		0.90	-	-	6.70	0.99	1.25	0.62	4.25	5.38
742/6		0.61	-	-	0.70	0.98		0.86	-	-	1.01	9.04	1.48	0.63	5.65	7.20

As with the genetic correlations, no trend was observed indicating the dependence of indirect gains on any of the factors.

Discussion and Conclusions

Importance and causes of family-by-trial interaction

The genetic correlations between trials, as found in this study, reflect at least moderate family-by-trial interaction. Moreover, the interaction was present in the form of substantial changes in family ranking. This finding agrees with some earlier studies that have demonstrated strong interactions between genetically close entries, like the plus tree progenies in this study, and experimental sites or even with blocks in a single field trial (e. g. JOHNSON and BURDON, 1990; MATHESON and COTTERILL, 1990; PEDERICK, 1990). Although nearly all the correlations were positive in sign, the behaviour of the families was clearly inconsistent enough to deserve notice. On the other hand, GEI appeared to be largely unpredictable and therefore difficult to avoid or exploit in breeding. The variation in correlations could not be attributed to any single external variable. Earlier studies have shown that genotypes can respond differently to biotic damage (HODGE and WHITE, 1986), stocking (CHANNELL, 1982) or site quality (e. g. HODGE and WHITE, 1992). In addition, statistically significant pseudo-interactions resulting, for instance, from heteroscedasticy of trait variances at different sites, are feasible (ROBERTSON, 1959; CAMPBELL and WILSON, 1973; BURDON, 1977; CAMPBELL *et al.*, 1986).

Significant family by fertiliser or soil fertility interactions have frequently been reported in other pine species (BURDON, 1971; JAHROMI *et al.*, 1976; JOHNSON and BURDON, 1990; MATHESON and COTTERILL, 1990). Therefore, it was supposed that the magnitude of height difference between 2 sites might correlate with the frequency of family rank changes. However, no evidence of such an association was found in this study. This contradicts with the recent finding of HODGE and WHITE (1992), who classified pairs of Slash pine (*Pinus elliottii* var. *elliottii*) progeny trials as either 'Same' or 'Different' using a site index (base age 25 years) difference of 2.6 m as the threshold, and found consistently smaller correlations for the pairs of 'Different' trials. The discrepancy in the correlations they found was not large at an age of 10 years (r_B = 0.71 and r_B = 0.63 for the 'Same' and 'Different' groups, respectively). Whether the described relationship would have been observed had the magnitude of the differences between paired trials been greater than in this study, remains open. Alternatively, the lack of site type effects could simply result from phenotypic plasticity, as Scots pine naturally performs well across a fairly wide edaphic gradient (KUUSELA, 1990).

Planting density was another factor that varied strikingly in the sample of trials studied here. The practical importance of spacing-driven interactions is generally estimated to be trivial (FRIES, 1984; MAGNUSSEN and YEATMAN, 1986; GULLBERG and VEGERFORS, 1987; WILLIAMS, 1988; ST. CLAIR and ADAMS, 1991). This is in harmony with the low, nonsignificant correlation found here between planting density and both r_B and r_{rank} values.

Distinguishing between biologically significant and repeatable GEI, and on the other hand, pseudo-interaction due to nuisance factors, is difficult. In the absence of major climatic differences and GEI due to spacing or fertility, the low across-site correlations observed in this study probably reflected the microsite variability and success of experimental design, rather than true biological interactions. In poorly designed or insufficiently replicated trials the precision of observed family performances is low, which naturally degrades the accuracy of family ranking, and consequently, the correlation between other trials. In this study, the pairs of trials with low single-site heritabilities tended to be poorly correlated, suggesting that a part of the interaction can be eliminated simply by enhancing the experimental precision of progeny trials.

It can be further speculated whether the trials studied were at different developmental phases at the time of measurement (10 years old) (see e. g. FRANKLIN, 1979). If the establishment of final family ranking depends on the size rather than on the age of trees, high correlations between trials can not be expected. Further studies are needed to determine the trends that type-b genetic and rank correlations may exhibit in relation to age, stage of stand development and experimental design. Such research would significantly increase our knowledge of the true nature of the interactions, as well as the best age or stage for selection in different conditions.

Efficacy of progeny test methods

The correlations between test orchards and forestry trials and between pairs of parallel forestry trials were of similar magnitude. This result strongly suggests that the test orchard method as such is not likely to be a notable source of GEI, and the risk of biased family ranking in test orchards is small. When both heritability and consistency of family performance with the other trials are considered simultaneously, the test orchards were clearly more effective than the forestry trials. Especially when laid out on fertile agricultural lands, the test orchard trials appeared to discriminate genetic differences considerably better than the normal forestry trials. This probably resulted from both homogeneous growing conditions and the relatively fast transition of the test orchard stands into the inter-tree competition phase, as suggested by MIKOLA (1985). This latter phenomenon was demonstrated in Douglas fir and loblolly pine by FRANKLIN (1978), who showed that the closure of stands was accompanied by an abrupt increase in additive genetic variance and heritability.

The contrasting performance of test orchards established on less fertile forest sites, i. e. poor height and survival even when compared to the forestry trials, could probably be attributed to inter-tree competition for nutrients. The effects of competition logically intensify earlier in dense test orchards stands than in forestry trials with lower densities.

As far as height growth is concerned, the testing of trees on homogeneous and fertile sites with narrow spacings probably involves no great risk. The validity of this conclusion as generalised to cover other important traits, such as branching quality, still remains to be studied. Use of the test orchard method on forest sites, on the other hand, does not seem reasonable, as there may be no additional gain to compensate for the higher management costs. However, statistically less efficient forestry trials are still necessary as they provide the sort of essential long-term data on time trends in genetic parameters, yield development on the stand level etc., that do not come within the scope of intensive field experiments.

When evaluating genetic values, the family-by-trial interaction is best managed by exploiting all available data and incorporating estimates of type-b correlations and heritabilities into the selection index (BURDON, 1979), BLP or BLUP equations (WHITE and HODGE, 1989). Observations from trials showing the lowest heritability, the smallest phenotypic variation and the poorest correlations with other sites are, accordingly, given the least index weights. Although this method theoretically leads to maximum genetic gain, it suffers from erroneous estimates of genetic parameters, especially

genetic correlations (HODGE and WHITE, 1992). The precision mainly depends on the number of families, which was probably too small for this purpose in most of the trials studied. More work is obviously needed, both with experimental designs and the statistical analysis of genetic field trials, to improve the precision of estimates of genetic parameters used as the basis of selection.

Acknowledgements

The author thanks Lic. For. JOUNI MIKOLA, Dr. PERTTI PULKKINEN, Dr. PIRKKO VELLING and an anonymus reviewer for their comments and suggestions.

Literature

BECKER, W.: Manual of quantitative genetics. Academic Enterprises, Pullman, Wa. 195 pp. (1984). – BURDON, R.: Clonal repeatabilities and clone-site interactions in *Pinus radiata*. Silvae Genetica **20**: 33–39 (1971). — BURDON, R.: Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genetica **26**(5–6): 168–175 (1977). — BURDON, R.: Generalisation of multi-trait selection indices using information from several sites. New Zealand Journal of Forest Science **9**: 145–142 (1979). — CAMPBELL, R., ECHOLS, R. and STONECYPHER, R.: Genetic variances and interactions in 9-year-old Douglas-fir grown at narrow spacings. Silvae Genetica **35**(1): 24–32 (1986). — CAMPBELL, R. and WILSON, B.: Spacing-genotype interaction in Douglas-Fir. Silvae Genetica **22**(1–2): 15–20 (1973). — CANNELL, M.: 'Crop' and 'isolation' ideotypes: evidence for progeny differences in nursery-grown *Picea sitchensis*. Silvae Genetica **31**(2–3): 60–66 (1982). — CARSON, S.: The practical importance of genotype x environment interaction for improvement of Radiata pine in New Zealand. Proceeding of the joint meeting of Western Forest Genetics Association and IUFRO Working Parties S2.02-05, 06, 12 and 14, Olympia, Washington, USA, 20–24 August 1990 (1990). — FALCONER, D.: The problem of environment and selection. American Naturalist. **86**: 293–298 (1952). — FALCONER, D.: Introduction to quantitative genetics. 2nd. ed. Longman Group Ltd. 340 pp. (1981). — FRANKLIN, E.: Model relating levels of genetic variance to stand development of four north american conifers. Silvae Genetica **28**(5–6): 207–212 (1979). — FRIES, A.: Spacing interaction with genotype and with genetic variation for production and quality traits in a trial of seedlings and grafted clones of Scots pine (*Pinus sylvestris* L.). Silvae Genetica **33**(4–5): 145–152 (1984). — GULLBERG, U. and VEGERFORS, B.: Genotype-environment interaction in Swedish material of *Pinus sylvestris*. Scandinavian Journal of Forest Research **2**: 417–432 (1987). — HODGE, G. and WHITE, T.: Relationship of rusiform rust and growth traits in open-pollinated slash pine progeny tests. In: Proc. IUFRO Symp. on Breeding Theory and Progeny Testing, and Seed Orchards. Oct. 13 to 17, Williamsburg VA.: 469–481 (1986). — HODGE, G. and WHITE, T.: Genetic parameter estimates for growth traits at different ages in Slash pine and some implications for breeding. Silvae Genetica **41**(4–5): 252–262 (1992). — JAHROMI, S., GODDARD, R. and SMITH, W.: Genotype x fertilizer interactions in Slash pine: Growth and nutrient relations. Forest Science **22**(2): 211–219 (1976). — JOHNSON, I.: Family-site interactions in radiata pine families in New South Wales, Australia. Silvae Genetica **41**(1): 55–62 (1992). — JOHNSON, I. and BURDON, R.: Family-site interaction in *Pinus radiata*: Implications for progeny testing strategy and regionalised breeding in New Zealand. Silvae Genetica **39**(2): 55–62 (1990). — KUUSELA, K.: The dynamics of boreal coniferous forests. Finnish National Fund for Research and Development, Helsinki. 172 pp. (1990). — LAMBETH, C., MASATOSHI, E. and WRIGHT, J.: Genetic analysis of 16 clonal trials of *Eucalyptus grandis* and comparisons with seedlings checks. Forest Science **40**(3): 397–411 (1994). — LINDGREN, D.: Prediction and optimization of genetic gain with regard to genotype x environment interactions. Studia Forestalia Suecica **166**: 15–24 (1984). — MAGNUSSEN, S. and YEATMAN, C.: Accelerated testing of Jack pine progenies: a case study. In: Proc. IUFRO Symp. on Breeding Theory and Progeny Testing, and Seed Orchards. Oct. 13 to 17, Williamsburg, VA.: 107–116 (1986). — MATHESON, A. and COTTERILL, C.: Utility of genotype x environment interactions. Forest Ecology and Management **30**: 159–174 (1990). — MATHESON, A. and RAYMOND, C.: The impact of genotype x environment interactions on Australian *Pinus radiata* breeding programs. Aust. For. Res. **14**: 11–25 (1984). — MIKOLA, J.: Methods used for the genetic evaluation of tree breeding material in Finland. In: Crop Physiology of Forest Trees. (Eds.: TIGERSTEDT, P., PUTTONEN, P. and KOSKI, V.). University of Helsinki, Dept. of Plant Breeding 138–143 (1985). — NIENSTAEDT, H. and RIEMENSCHNEIDER, D.: Changes in heritability estimates with age and site in White spruce, *Picea glauca* (MOENCH) VOSS. Silvae Genetica **34**(1): 34–41 (1985). — PAJAMÄKI, J. and KARVINEN, K.: General Statistics on Forest Tree Breeding in Finland. Metsäntutkimuslaitoksen tiedonantoja (Communicationes of the Finnish Forest Reseach Institute) **565**. 31 pp. (1995). — PEDERICK, L.: Family x site interactions in *Pinus radiata* in Victoria, Australia, and implications for breeding strategy. Silvae Genetica **39** (3–4, 134–140 (1990). — ROBERTSON, A.: The sampling variance of the genetic correlation coefficient. Biometrics **15**: 469–485 (1959). — SAS Institute INC.: SAS User's Guide: Basics, Version 5 Edition. 861–874. (1985). — SAS Institute INC.: SAS Technical Report P-229. SAS/STAT Software: Changes and enhancements. Release 6.07.: 289–366 (1922). — ST. CLAIR, J. and ADAMS, T.: Relative family performance and variance structure of open-pollinated Douglas-fir seedlings grown in three competitive environments. Theoretical and Applied Genetics. **81** (4): 541–550 (1991). — TIGERSTEDT, P.: A proposal for genetic testing of South and Central Finnish plus trees of Scots pine. Mimeograph (in Finnish only) (1973). — WHITE, T. and HODGE, G.: Predicting breeding values with applications in Forest Tree Improvement. Kluwer Academic Pub., Dordrech, The Netherlands. 367 pp. (1989). — WHITE, T. and HODGE, G.: Indirect prediction of genetic values. Silvae Genetica **40** (1): 20–28 (1991). — WILLIAMS, C.: Accelerated short-term genetic testing for loblolly pine families. Canadian Journal of Forest Research **18**: 1085–1089 (1989). — YAMADA, Y.: Genotype by environment interaction and genetic correlation of the same trait under different environments. Jap. Jour. Genet. **37**(6): 498–509 (1962).

IV

TIME TRENDS IN GENETIC PARAMETER ESTIMATES AND SELECTION EFFICIENCY FOR SCOTS PINE IN RELATION TO FIELD TESTING METHOD

Matti Haapanen

Finnish Forest Research Institute, Box 18, FIN-01301 Vantaa, Finland, fax: +358-9-85705711; phone: +358-9-85705545, e-mail: matti.haapanen@ metla.fi

Received November 27, 2000; accepted March 30, 2001

ABSTRACT

Multiple assessments of cumulative tree height in 26 *Pinus sylvestris* L. progeny trials in Finland, ranging between 5 and 18 years of age, were analysed to determine time trends in variance components, heritability, coefficient of additive genetic variation, age-age and site-site genetic correlations, and to estimate the impact of test characteristics on these parameters. Two distinctive methods of field progeny testing were investigated, conventional 'forestry trials' and intensively managed test orchard (farm-field) trials. The effects of contrasting stand densities (2,500 vs. 8,888 trees planted per ha) and site quality (forest vs. field sites) on the levels and trends of variance components and genetic parameter estimates were quantified by means of repeated-measures mixed models. Responses to parental and within-family selection were computed using parameters derived from these models. The trials, which were arranged in groups by spacing and site quality, showed markedly different levels of heritability. The highest estimates occurred on the fast–growing test orchard sites. On average, additive genetic variances and heritabilites from single–site analyses were inflated by 60% due to the family-by-site interaction. No systematic time trends were detected for heritability and type-b correlation. Furthermore, the relative magnitude of the family-by-site interaction was independent of the degree of heterogeneity among the trial sites. Age-age correlations were positive and moderately high, and showed a moderate fit to the log of the age ratio. Selection efficiencies were examined using tree height at age 20 as the target trait. The correlated responses per year to early parental (backward) selection peaked at the age of around 5 to 7 years and were always greater than the gains from direct selection for tree height at the age of 20 years. In within-family (forward) selection, the annual responses were initially low, increasing toward the target age. The optimum age of within-family selection occurred later (at the age of 8, ..., 16 years) than in parental selection. The highest selection efficiencies were consistently associated with the densely spaced test orchard scenarios. Correspondingly, widely spaced trials on sites of poor quality produced the least responses to early selection.

Keywords: early selection, genetic correlation, genetic parameters, heritability, *Pinus sylvestris*, progeny testing, testing method, time trends, type-B correlation.

INTRODUCTION

In forest tree breeding, mature performance is customarily predicted using attributes measured in juvenile field trials. The advantages of pre-rotation selection comprise easier measurement and lower costs per tree, and there is also a quicker incorporation of genetically improved materials into forestry. Above all, however, selection at early ages can be expected to yield higher genetic gain per unit of time than direct selection for harvest–age performance (LAMBETH 1980, LINDGREN 1984). In principle, the use of early testing as an effective screening tool requires sufficient knowledge of the quantitative genetic parameters. The outcome of indirect selection depends on the genetic control in selected traits (heritability), their genetic associations with mature traits (age-age genetic correlation), and the magnitude of genotype-by-environment interaction (site-site correlation). These parameters are thus key ingredients in the planning of efficient breeding, testing and selection strategies, and their estimation is normally a regular part of the analysis of field-testing data.

Genetic parameters may markedly change as trees grow and develop (NAMKOONG & CONKLE 1976, FOSTER 1986, FRANKLIN 1979, BALOCCHI *et al.* 1993, DIETERS *et al.* 1995). To effectively implement a tree improvement programme, there must be a sufficient understanding of the underlying reasons for these changes. There is some evidence suggesting that the genetic control of tree growth is closely related to periodic shifts in the ontogenetic stand development (FRANKLIN 1979, VASQUEZ & DVORAK 1996). FRANKLIN (1979) interpreted sharp changes in genetic variance and heritability as responses to the onset and termina-

tion of intense inter-genotypic competition. However, more recent attemps to verify FRANKLIN's (1979) hypotheses have not always been successful (LAMBETH *et al.* 1983, BOUVET & VIGNERON 1995, FOSTER 1986, GILL 1987, SATO 1994, DANJON 1994) and the issue has not been resolved. This owes much to the fact that the time series of genetic parameters for forest trees are typically sparse and cover only a small part of the rotation. It seems clear that without data from old field trials, any hypotheses about the patterns of genetic parameters are likely to remain controversial. Regrettably, these important mature data are nearly always scarce for long-lived tree species, which have a short history of genetic improvement by comparison to their commercial rotations.

Trends in genetic parameters are in many situations difficult to detect. This is not only because of a deficiency of data, but also because of large variability among genetic parameter estimates at any age. Precise genetic parameter estimates are difficult to obtain from small experiments (HODGE & WHITE 1992). Furthermore, genetic parameter estimates reflect a number of non-genetic factors such as the magnitude and pattern of microsite variability, experimental design, spacing, silvicultural management, and the occurrence of environmental stresses (RINK & CLAUSEN 1989, MAGNUSSEN 1993, XIE & YING 1996). In fact, the common inconsistency of parameter estimates may be seen to be perfectly in line with the fact that population parameters attributed as 'genetic', also reflect the environmental circumstances under which they are estimated (Falconer 1981). Tree breeding trials typically represent diverse environmental conditions and cultural practices since the multiple objectives of genetic testing cannot be optimally met by a single approach (LOO-DINKINS 1992). The accelerated methods adopted by many tree improvement programmes for progeny testing (MIKOLA 1985), add their share to this diversity. It would obviously be important to quantify the effects that the various approaches to field testing may have on the genetic control of traits. This information would be valuable in ensuring the best use of data gathered from different types of field trial, and also when considering the amendment of characteristics of future trials. As yet, however, not too many studies have addressed the effects that test characteristics may have on the patterns of genetic parameters and selection efficiency (FALKENHAGEN 1989, MAGNUSSEN 1991, HODGE & WHITE 1992, WHITE & HODGE 1992, ADAMS *et al.* 1994, WOODS *et al.* 1995, JANSSON *et al.* 1998).

The accelerated field trials used in the progeny testing of Scots pine (*Pinus sylvestris* L.) in Finland are called 'test orchards' (synonymous to 'farm-field trials'). The features of the test orchard method include uniform and often fertile sites, high planting density (up to 10,000 trees per ha), fencing, and intensive site manipulation to minimise weed competition and edaphic heterogeneity. Hence, test orchard conditions often differ markedly from those prevailing at conventional 'forestry trials', which more resemble managed stands with respect to initial spacing (2,000–2,500 trees per ha), site quality (mostly on typical forest soils) and silvicultural management (MIKOLA 1985). In addition, test orchard trials are intended for fairly rapid screening at the age of 10–15 years, whereas forestry trials can sometimes be assessed up until the end of the commercial rotation, to estimate the productivity of genetic entries over several thinnings to final harvest (MIKOLA 1985).

To my knowledge, only four studies have compared these two approaches to field testing using the same set of genetic entries, namely MAGNUSSEN and YEATMAN (1986) with jack pine (*Pinus banksiana* Lamb.), CARLSON (1990) with lodgepole pine (*Pinus contorta* Douglas ex Loudon), WOODS *et al.* (1995) with Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), and HAAPANEN (1996) with Scots pine. While these studies consistently report higher efficiency of selection for test-orchard-like conditions, the findings are based on rather meager data with respect to the number of trials included for comparison or the range of ages observed. To avoid these limitations, this study exploited the large database of routine measurements accumulated from the Scots pine progeny-testing programme in Finland. The objectives of this study were: to determine and model time trends in the genetic parameters of importance to the Scots pine progeny testing, and to find out how these parameters, their tendencies over time and the efficiency of early selection are affected by the distinctive features of the two testing methods.

MATERIAL AND METHODS

Experimental material and measurements

The data were from 9 independent series of Scots pine progeny trials located between 60th and 64th N in southern Finland. Each of the series comprised 2 to 4 parallel trials (at least one forestry trial and one test orchard). Overall, 15 forestry trials and 11 test orchard trials contributed data to this study (Table 1). The 26 trials were sampled from among the nearly 1,300 Scots pine progeny trials established in Finland since the early 1960's.

The trials comprised 4 to 10 (median = 6) blocks laid out in a randomized complete block design. The plots were usually formed of 5 × 5 trees planted at 2 m × 2 m (the forestry trials) or 0.75 m × 1.5 m (the test

Table 1. Description of the progeny trials used in the analyses.

Trial	Type	Lat.	Lon.	Alt., m	Site type	Establ. year	Area, ha	Families	Blocks	Trees/ plot	Planting density, trees/ha	Measurement ages	Mean height at age 12, dm
572/1	Test orchard	61°48'	29°18'	87	Field, mould	1978	0.64	39	7	15	8888	6, 7, 9, 11, 15	50.0
572/2	Forestry trial	60°36'	23°22'	75	Field, clayish	1978	0.97	35	4	16	2500	6, 11, 16	28.9
572/3	Forestry trial	62°32'	24°27'	195	Dryish heath	1978	2.88	39	10	16	2500	7, 11, 17	23.5
572/4	Forestry trial	63°19'	28°50'	185	Moist forest land	1978	2.80	39	10	16	2500	6, 11, 16	27.5
698/1	Test orchard	61°48'	29°17'	83	Field, mould	1980	0.32	23	4	25	8888	5, 7, 8, 11, 14	48.1
698/2	Forestry trial	61°36'	26°18'	110	Moist forest land	1980	1.20	21	4	25	2500	7, 11, 16	24.7
739/1	Test orchard	62°46'	25°42'	140	Dryish heath	1981	1.38	74	6	25	8888	5, 7, 12	20.4
739/2	Test orchard	61°48'	29°17'	81	Field	1980	1.49	78	6	25	8888	5, 7, 9, 12	49.1
739/5	Forestry trial	62°59'	31°13'	178	Dryish heath	1980	2.76	42	6	25	8888	7, 12, 18	26.1
739/6	Forestry trial	60°41'	23°59'	120	Moist forest land	1981	2.70	41	6	25	8888	8, 12, 18	22.9
757/1	Test orchard	60°58'	22°43'	93	Dryish heath	1980	0.76	30	6	25	6944	7, 9, 13, 18	28.2
757/2	Forestry trial	62°00'	25°11'	155	Moist forest land	1981	1.00	30	3	25	2500	7, 12, 18	26.8
803/1	Test orchard	62°44'	25°41'	140	Dryish heath	1982	0.68	50	4	25	8888	7, 10, 12	22.0
803/2	Test orchard	61°48'	29°17'	79	Field	1982	0.66	50	4	25	8888	5, 7, 10, 12, 16	48.8
803/3	Forestry trial	60°21'	25°01'	57	Dryish heath	1982	1.00	21	4	25	2500	7, 12, 18	29.3
861/1	Forestry trial	62°16'	26°55'	122	Dryish heath	1983	2.13	31	6	25	2500	12, 17	24.4
861/2	Forestry trial	62°03'	25°27'	200	Moist forest land	1983	2.16	31	6	25	2500	12, 17	23.4
861/3	Test orchard	62°45'	25°41'	140	Dryish heath	1983	0.41	31	4	25	8888	5, 10, 12, 17	24.3
862/1	Forestry trial	61°35'	28°55'	110	Dryish heath	1983	1.80	25	6	25	2500	12, 17	22.6
862/3	Test orchard	61°49'	27°13'	117	Dryish heath	1983	0.33	25	4	25	8888	5, 10, 13, 17	19.9
864/1	Forestry trial	62°54'	31°33'	190	Dryish heath	1983	2.10	30	6	25	2500	12	20.8
864/2	Forestry trial	62°03'	25°27'	200	Moist forest land	1983	2.10	30	6	25	2500	12, 17	24.8
864/3	Test orchard	62°54'	25°41'	140	Dryish heath	1983	0.39	30	4	25	8888	5, 10, 12, 17	23.4
865/1	Forestry trial	61°35'	28°55'	110	Dryish heath	1983	2.21	32	6	25	2500	12, 17	24.5
865/2	Forestry trial	62°16'	26°55'	132	Dryish heath	1983	2.19	32	6	25	2500	12, 17	24.1
865/3	Test orchard	61°49'	27°13'	117	Moist forest land	1983	0.142	32	4	25	8888	5, 9, 13, 17	19.5

orchards) spacing. The initial spacing varied considerably with the type of trial. The test orchards were planted at a density of 8,888 trees per ha, which is up to 3.5 times higher than that used in the forestry trials (2,500 trees per ha). Differences in site qualities were also large, although not as closely linked to the testing method as the differences in spacing. Most of the trial sites (8 forestry trials, 6 test orchards) were classified as dryish forest land (*Vaccinium* type). Seven trials (6 forestry trials, 1 test orchard) were situated on moderately dry forest sites (*Myrtillus* type). One forestry trial and 4 test orchards were planted on former agricultural land. Apart from these farmland sites, the Podzol soils were typical to the Scots pine dominated natural stands of the region. No thinnings had taken place in any of the trials prior to the measurement.

The families were open-pollinated offspring of first-generation 'plus trees' selected from natural stands of southern Finland. The seed for the progeny trials were collected either from the original plus trees or, most commonly, from their grafts in young seed orchards that had a negligible amount of internal pollen production at the time of collection. In both cases, the pollen parents were presumed to be a random sample of genotypes from the surrounding wild population, and thus share the same origin with the plus trees.

The number of half-sib families in a trial ranged from 25 to 78 (Table 1). As a rule, the parallel trials within any one series consisted of a common set of families. The two exceptions to the rule were the series No. 739 and 803, in which half of the families planted in the test orchard trials were not present in the parallel forestry trials. Altogether the 26 trials consisted of 338 unique half-sib families and approximately 122,000 planted seedlings. In addition, each trial accommodated a few standard check-lots which were omitted in the analyses.

The data comprised measurements of cumulative tree height (to the nearest 1 dm) carried out at inconsistent intervals from age 5 (age from seed) through age 18. In total, the 26 trials provided 82 data sets for analyses (Table 1).

Analyses

Single-site and across-site partitioning of variance

In the first stage, the total variance in each of the 82 data sets was decomposed into additive components due to random family, plot and within-plot effects. Effects due to differences among blocks were interpreted as fixed since they are not relevant to genetic parameter estimation. The single-site analysis was based on the following linear model:

$$y_{jkl} = \mu + B_j + f_k + fB_{jk} + w_{jkl} \quad [1]$$

where: y_{jkl} = height of lth tree in jth block and kth family, $E(y_{jkl}) = \mu + B_j$, $var(y_{jkl}) = \sigma^2_{f(1)} + \sigma^2_{fB(1)} + \sigma^2_{w(1)}$; μ = a general mean; B_j = fixed effect of jth block ; F_k = random effect of kth half-sib family; $E(f_k) = 0$, $Var(f_k) = \sigma^2_{f(1)}$; FB_{jk} = random plot error due to interaction between jth block and kth family; $E(fB_{jk}) = 0$, $Var(fB_{jk}) = \sigma^2_{fB(1)}$; w_{jkl} = random tree error of lth tree in jkth plot; $E(w_{jkl}) = 0$, $Var(w_{jkl}) = \sigma^2_{w(1)}$.

The variance components were estimated using the method of restricted maximum likelihood (REML), available in the MIXED procedure of the SAS/STAT package (SAS 1992, LITTELL *et al.* 1996). The single-site estimates of family variance ($\sigma^2_{f(1)}$) are biased, comprising a variance component that estimates the varying relative performance of families from one site to another. To obtain unbiased estimates of genetic variance and heritability, data from multiple environments are mandatory (ZOBEL & TALBERT 1984). In this study, the measurements made at equal ages were combined over parallel trials within each series (one year difference between the measurement ages was tolerated in order to ensure enough data for each analysis). Before the analysis, the data from each parallel trial were transformed to equal additive genetic variance by multiplying all the observations by a factor σ_{An}/σ_{Ai}, where σ_{Ai} denotes a single-site estimate of the additive genetic standard deviation for the i^{th} parallel trial and σ_{An} is some constant (DANELL 1988, SONESSON & ERIKSSON 2000). In this study, σ_{An} was set equal to the mean of the additive genetic standard deviations for the n parallel trials included in an analysis. The total number of independent across-trial analyses performed was 22, representing ages 5, 6, 7, 10, 11, 12, 16, 17 and 18 (Table 2). The additive model used to estimate variance components across trials was:

$$y_{ijkl} = \mu + S_i + B_{ij} + f_k + fS_{ik} + fB_{ijk} + w_{ijkl} \quad [2]$$

where: y_{ijkl} = height of lth tree in kth family and jth block in ith trial; $E(y_{ijkl}) = \mu + S_i + B_{ij}$, $Var(y_{ijkl}) = \sigma^2_f + \sigma^2_{fS} + \sigma^2_{fB} + \sigma^2_w$; μ = a general mean; S_i = fixed effect of ith trial; B_{ij} = fixed effect of jth block in ith trial; f_k = random effect of kth half-sib family; $E(fk) = 0$, $Var(f_k) = \sigma^2_f$; fS_{ik} = random interaction effect of kth family in ith trial; $E(fS_{ik}) = 0$, $Var(fS_{ik}) = \sigma^2_{fS}$; fB_{ijk} = random plot effect due to interaction between kth family and jth block in ith trial; $E(fB_{ijk}) = 0$, $Var(fB_{ijk}) = \sigma^2$fB; w_{ijkl} = random tree error of lth tree in ijkth plot; $E(w_{ijkl}) = 0$, $Var(w_{ijkl}) = \sigma^2_w$.

Table 2. Across-site estimates of variance components [2], additive genetic variance (unbiased and biased) [7], type-b correlation [6], the ratio of the between-family and family-by-site variance components, and unbiased individual heritability [4].

Trial series	Age	σ^2_f	σ^2_{fS}	σ^2_{fB}	σ^2_w	σ^2_A	$\sigma^2_{A(biased)}$	r_A	σ^2_{fS}/σ^2_f %	$h^2_i \pm$ S. E.
572	6	0.138	0.113	0.397	5.548	0.55	1.00	0.55	82	0.089±0.030
572	11	1.277	0.945	2.058	31.874	5.11	8.89	0.57	74	0.141±0.045
572	16	2.845	1.785	2.384	74.317	11.38	18.52	0.61	63	0.140±0.044
698	7	0.547	0.685	1.311	8.021	2.19	4.93	0.44	125	0.207±0.143
698	11	1.478	2.070	8.385	39.865	5.91	14.19	0.42	140	0.114±0.109
739	7	0.348	0.308	0.993	8.968	1.39	2.63	0.53	89	0.131±0.043
739	12	0.987	1.412	7.642	34.697	3.95	9.60	0.41	143	0.088±0.039
739	18	0.935	2.334	26.021	84.923	3.74	13.08	0.29	250	0.033±0.055
757	7	0.145	0.037	0.785	4.622	0.58	0.73	0.80	25	0.104±0.055
757	12	0.418	0.755	6.291	26.431	1.67	4.69	0.36	181	0.049±0.061
757	18	1.454	3.891	8.135	52.068	5.82	21.38	0.27	268	0.089±0.095
803	7	0.311	0.149	0.381	5.759	1.24	1.84	0.68	48	0.189±0.056
803	12	1.563	0.687	1.486	26.462	6.25	9.00	0.69	44	0.207±0.059
861	12	0.320	0.342	2.494	27.011	1.28	2.65	0.48	107	0.042±0.026
861	17	1.239	0.750	7.075	62.144	4.95	7.95	0.62	61	0.070±0.034
862	12	1.245	1.127	3.531	27.405	4.98	9.48	0.52	91	0.149±0.090
862	17	3.729	3.154	8.316	59.678	14.92	27.53	0.54	85	0.199±0.110
864	12	0.553	0.305	3.275	26.450	2.21	3.43	0.64	55	0.072±0.035
864	16	2.354	0.757	10.627	65.696	9.42	12.45	0.76	32	0.119±0.061
865	12	0.888	0.537	3.511	27.882	3.55	5.70	0.62	60	0.108±0.044
865	17	3.035	1.389	9.018	64.224	12.14	17.70	0.69	46	0.156±0.058

Estimation of genetic parameters

The open-pollinated families were assumed to consist of paternally unrelated siblings. Accordingly, the additive genetic variance on a single-site basis was estimated by multiplying the among-family component of variance by four ($\sigma^2_{A(1)} = 4\ \sigma^2_{f(1)}$), the inverse of the coefficient of genetic relationship for half-sibs (FALCONER 1981). Single-site [3] and across-site [4] estimates of individual heritability and single-site coefficients of additive genetic variation [5] were calculated from the equations:

$$h^2_{i(1)} = 4\ \sigma^2_{f(1)} / (\sigma^2_{f(1)} + \sigma^2_{fB(1)} + \sigma^2_{w(1)}) \quad [3]$$

$$h^2_i = 4\ \sigma^2_f / (\sigma^2_f + \sigma^2_{fS} + \sigma^2_{fB} + \sigma^2_w) \quad [4]$$

$$CV_{A(1)} = 100\ (4\ \sigma^2_{f(1)})^{0.5} / x \quad [5]$$

Standard errors of the heritability estimates were calculated according to an approximation given by DICKERSON (1969).

The influence of the family-by-site component of variance on the single-site additive genetic variance was determined by calculating the ratio of the unbiased and biased estimates of the additive genetic variance, using variance components from the across-site analysis. This ratio [6] estimates the average degree of 'type-B' genetic correlation (r_B) between different individuals of the same genetic group when many environments are involved for testing (DICKERSON 1962).

$$r_B = \sigma^2_f / (\sigma^2_f + \sigma^2_{fS}) = (¼\ \sigma^2_A) / (¼\ \sigma^2_{A(1)}) \quad [6]$$

The type-B correlation is actually a measure of genotype-by-environment interaction (G × E) which can range from 0 to 1. The higher values indicate less interaction (BURDON 1977). In this study, type-B correlation was mainly used to approximate and adjust for the upward bias in the single-site estimates of heritability [9, 10].

Age-age genetic correlations within each trial ('type-A' correlation, r_G) were approximated by Pearson correlation using least-square family means as observational units.

Estimation of time trends

A repeated–measures model was developed [7] and fitted to the time-series data representing single-site estimates of variance components, heritability and coefficient of variation. The aims of the analysis were: (1) to draw statistical inferences on the fixed main effects of testing method (narrow vs. wide spacing) and site quality (forest vs. field site), especially on the interactions of these two main effects with age (modeled as a continuous regression variable), and (2) to set up time-trend functions needed in forecasting the genetic response to early selection. The exponential

Table 3. Parameter estimates of the time-trend model [7] fitted to the natural logarithms of the single-site additive genetic variance, family-by-site variance and residual (within-plot) variance [1], estimates of single-site individual heritability [3], single-site coefficient of additive genetic variation [5], and cumulative mean height. Variances due to the random effects of the model [7] are given in the bottom. Missing estimates (–) denote a statistically non-significant factor which was omitted from the final model.

Factor	Level	Parameter	Parameter estimates and their standard errors					
			$log_e(\sigma^2_{A(I)})$	$\log_e(\sigma^2_{fB(I)})$	$\log_e(\sigma^2_{w(I)})$	$h^2_{i(I)}$	$CV_{A(I)}$	Cumulative height
Intercept			–4.871±0.340	–4.736±0.643	–3.618±0.297	0.372±0.040	26.539±3.065	–18.505±3.061
Age or $\log_e$(age)	(Continuos)	b_1	3.155±0.072	2.031±0.264	2.974±0.133	0.005±0.002	–1.904±0.474	4.576±0.501
Age	(Continuos)	b_2	–	–	–	–	0.044±0.021	0.085±0.022
Method	forestry trial	M_1	–0.493±0.199	2.196±0.767	1.096±0.359	–0.262±0.084	–3.349±1.123	–
	test orchard	M_2	0.000	0.000	0.000	0.000	0.000	–
Site quality	forest	S_1	–1.223±0.240	–3.207±0.782	–0.936±0.362	–0.193±0.046	–7.365±2.977	14.981±2.994
	field	S_2	0.000	0.000	0.000	0.000	0.000	–
Method-by-Site	forestry trial, forest	MS_{11}	–	–	–	0.144±0.091	–	–5.723±2.665
	forestry trial, field	MS_{12}	–	–	–	0.000	–	10.164±5.543
	test orchard, forest	MS_{21}	–	–	–	0.000	–	0.000
	test orchard, field	MS_{22}	–	–	–	0.000	–	0.000
Method-by-Age	forestry trial	$b_{11\cdot}$	–	–0.861±0.298	–0.536±0.170	–	–	–2.419±0.477
	test orchard	$b_{12\cdot}$	–	0.000	0.000	–	–	0.000
Site-by-Age	forest	$b_{1\cdot1}$	–	1.623±0.316	0.076±0.160	–	0.644±0.250	–3.437±0.273
	field	$b_{1\cdot1}$	–	0.000	0.000	–	0.000	0.000
Method-by-	forestry trial, forest	b_{111}	–	–	0.227±0.092	–	–	3.159±0.515
Site-by-Age	forestry trial, field	b_{112}	–	–	0.000	–	–	0.000
	test orchard, forest	b_{121}	–	–	0.000	–	–	0.000
	test orchard, field	b_{122}	–	–	0.000	–	--	0.000
			Variance components					
Among trials		σ^2_g	0.155	0.342	0.017	0.004	0.857	7.968
Within trials		σ^2_e	0.169	0.197	0.049	0.005	13.279	8.745

relationship between the variance components and age was accounted for by analysing these variables on a natural log-log scale.

$$y_{ijkl} = \mu + M_i + S_j + MS_{ij} + (b_{1..} + b_{1i\cdot} + b_{1\cdot j} + b_{1ij})T + b_2T^2 + g_k + e_{ijkl} \quad [7]$$

where: y_{ijkl} = single-site parameter estimate (Table 3) (= $\log_e(y_{ijkl})$, if y is a variance component); M_i = fixed effect of i^{th} testing method (i = 1…2; 1='forestry trial', 2='test orchard'); S_j = fixed effect of j^{th} site quality (j = 1…2; 1='field', 2='forest'); T = age of the trial (=$\log_e(T)$, if y is a variance component); MS_{ij} = interaction between i^{th} testing method and j^{th} site quality; g_k = random subject effect of the k^{th} trial, $E(g_k) = 0$, $var(g_k) = \sigma^2_g$; e_{ijkl} = residual (within-trial) effect, $E(e_{ijkl}) = 0$, $var(e_{ijkl}) = \sigma^2 e$; $b_{1..}$, $b_{1i.}$, $b_{1.j}$, b_{1ij} , b_2 = regression coefficients (i and j refer to levels of the main effects)

Non-significant effects were dropped from the model one-by-one, starting from the highest-order interactions. The analysis was then repeated until all the effects in the model appeared statistically significant ($p < 0.05$). However, those of the main effects that were involved in significant interactions were preserved in the model independently of their own level of significance. These interactions comprised age regressions estimated within subclasses of the main effects, which are considered to provide evidence for the disparity of the linear age trends among the main-effects levels. The time–covariance structure for the residual terms was chosen among four options provided by PROC MIXED, namely compound symmetry (CS), and three spatial structures, SP(POW), SP(SP) and SP(GAU). The structure that gave the best fit, as measured by the Akaike's Information Criterion (LITTELL *et al.* 1996), was used to construct the final model.

The time–trend analysis of the site-site (r_B) and age-age (r_G) correlations consisted of regressing the correlation estimates on age, or the natural logarithm of the age ratio (LAR = $\log_e$(younger age/older age)), respectively. The model for the age-age correlations also accounted for possible interactions between LAR, testing method and site quality. Predictions from these regression equations were subsequently used to estimate genetic gains.

Selection efficiency

Efficiency of early selection was examined by chosing total height at age 20 as the target trait to be improved. The selection schemes considered were: 1) parental (backward) selection (PS), under which parents are chosen based on average performance of their progeny in a single trial, and 2) within-family selection (WFS) among offspring (forward selection), on the basis of deviations of individual phenotypes from family and block averages. Family information is usually combined with individual-level information when trees are selected forward in genetic tests (COTTERILL & DEAN 1990), but this option was not considered here. The correlated responses to early selection (R_{20ij}) were predicted for four groupings of the data (the combinations of two testing methods and two site qualities) as follows (FALCONER 1981):

$$R_{20ij} = q\, i_j\, r_{Gj20}\, h_j\, h_{20}\, cv_{20}, \qquad [8]$$

where q is a constant (1 for WFS and 2 for PS), i_j is the selection intensity at age j, r_{Gj20} is the genetic correlation between cumulative height assessed at ages j and 20, h is the square root of heritability, and CV_{20} is the target-age phenotypic coefficient of variation (cv_{20} = 100 s_{20} / age-20 mean height). The heritability values appropriate for PS (family heritability, h^2_f) and WFS (within-family heritability, h^2_{wf}) were calculated as in Eq. 9 and 10, on the basis of predictions of single-site variance components from [7].

$$h^2_f = r_B\, \sigma^2_{f(1)} / (\sigma^2_{f(1)} + \sigma^2_{fb(1)} / n_B + \sigma^2_{w(1)} / (n_B\, n_P)), \qquad [9]$$

$$h^2_{wf} = 3\, r_B\, \sigma^2_{f(1)} / (c_1 \sigma^2_{fb(1)} + c_2 \sigma^2_{w(1)}) \qquad [10]$$

The terms n_B and n_P refer to the numbers of blocks per trial, and of trees per plot, respectively. Here, these values were fixed to: n_B = 6, and n_P = 25. The coefficients c_1 and c_2 were defined as: $c_1 = (n_B - 1)/n_B$ and $c_2 = (n_B n_P - 1)/(n_B n_P)$.

To adjust for the 'time penalty' associated with postponed selection, the response to selection [8] is often presented on a per year basis (R^y). This requires dividing the response by the sum of the testing phase (j) plus the time required to mate the selected trees and produce the new generation of progeny (b).

$$R^y_{20ij} = q\, i_j\, r_{Gj20}\, h_j\, h_{20}\, cv_{20} / (j + b) \qquad [11]$$

The time required to produce a new progeny generation for WFS was defined as the selection age plus 10 years (b = 10). For PS, the length of the breeding phase was assumed to be zero since the parents selected backward can usually be mated in existing clonal collections. Furthermore, the parental information is immediately applicable for roguing inferior clones in seed orchards.

Relative efficiency of indirect selection (RE_{20ij}) was estimated by dividing R^y_{20ij} by the corresponding response to direct selection on the target trait (R^y_{20i20}).

$$RE_{20ij} = q\, i_j\, r_{Gj20}\, h_j\, h_{20}\, cv_{20}\, (20 + b) / \{ q\, i_{20}\, h^2_{20}\, cv_{20}\, (j + b) \} \qquad [12]$$

Assuming equal intensities of selection ($i_j = i_{20}$), the former equation reduces to a simpler form:

$$RE_{20ij} = r_{Gj20}\, h_j\, (20 + b) / \{ h_{20}\, (j + b) \} \qquad [13]$$

RESULTS

Height growth exhibited large variation among the individual progeny trials. The growth potential of each site was quantified by estimating the mean cumulative height at a base age of 12 years (H_{12}) by means of a regression analysis. The range of this index was from 19.5 to 50.0 dm (Table 1, Fig. 1). The highly differential growth curves were clearly related to edaphic differences among the trial sites. The trials established on former arable land ('field sites') were superior in height growth (mean H_{12}= 45.0 dm) to the trials on forest sites (mean H_{12}= 23.9 dm). On field sites, growth trajectories were nearly identical excluding the single

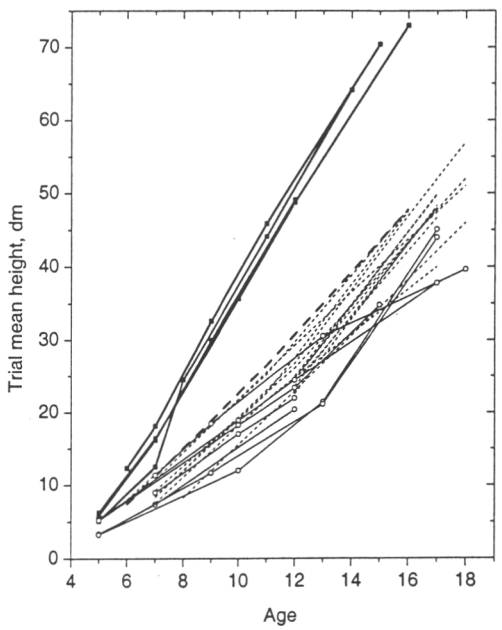

Figure 1. Development of mean height in 26 Scots pine progeny trials representing groups of field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines).

forestry trial, no. 572/2, which showed markedly slower growth (H_{12} = 28.9 dm) than the field-site test orchards. This was probably related to the high clay content of the soil in this trial (Table 1, Fig. 1). The two categories of forest site types, 'dry' and 'moist', appeared to be of similar value from the point of view of tree growth (difference in H_{12} = 0.55 dm), so they were combined in the following analyses.

The largest component of the single-site variance consistently attributed to within-plot residual effects (range 65–95%). The residual variances increased exponentially as the progeny trials became older. The family level sources of variance ($\sigma^2_{f(I)}$ and $\sigma^2_{fb(1)}$) typically remained below 15% of the total variance. The third component in this category, the family-by-site interaction variance from an across-site analysis (σ^2_{fS}), varied from 13 to 295% of the unbiased family variance. The accompanying type-B genetic correlations also showed large variance (range 0.25–0.94). In general, the relative magnitude of the interaction variance did not change much over successive measurements, except for the series No. 757 in which an abrupt decrease in r_B (from 0.86 to 0.25) occurred between ages 7 and 18. When all of the type-B correlation estimates (Table 2) were pooled and regressed on age, no trend was detected [14].

$$r_B = 0.673 - 0.003(\text{age}) \quad (r^2 = 0.01) \qquad [14]$$

Because of the flat slope ($p < 0.405$), the intercept of the regression formula, 0.673 (S.E. = 0.118, $p < 0.001$) was used as an age-independent predictor of r_B. Neither was the type-B correlation related to the heterogeneity of the test environments. Three measures of dispersion were calculated to estimate the degree of among-trial heterogeneity: the variance, the natural logarithm of the variance, and the coefficient of variation of the mean heights of the parallel trials. Pearson correlations between these three statistics and the respective estimates of r_B were all small and non-significant (r = 0.18, 0.13, and, 0.10 respectively).

The estimated level of type-B correlation suggests that the single-site estimates of heritability were overestimated by roughly 50 % (derived as $100/r_B$–100) (Table 2). The mean of the biased heritability estimates over all trials and ages was 0.23 (range 0.00–0.65), whereas for the unbiased (across-site) estimates the mean was 0.12 (range 0.03–0.20). The precision of the heritability estimates varied from low to modest: the ratio of the standard error to heritability value ranged from 0.2 to 6.4 (mean = 0.59). Despite the notable temporal fluctuations in some of the trials (Fig. 2), heritability did not show any systematic tendency over the first 18 years of testing. However, the heritabilities were clearly related to the mean growth rate and, thereby, to the edaphic properties of the test sites. The highest heritability estimates were acquired from the fast–growing test orchards and, conversely, the lowest heritabilities were associated with forest–site forestry trials. Indications of significant testing method-by-site quality interaction were also present (Table 3). For instance, some of the test orchards established on forest soil showed markedly higher levels of heritability as compared to the forestry trials on forest soil (Fig. 2). In general, the effect of the testing method ('high spacing' vs. 'low spacing') on heritability was slightly larger than that of site quality ('forest' vs. 'field') (Table 3).

The coefficient of additive genetic variation, which indicates 'evolvability', the potential for genetic evolution of the mean genotype (HOULE 1992), showed a slightly decreasing tendency over age. The time trends were significantly different for the trials on different type of sites (Fig. 3, Table 3). The most rapid decreases occurred in the fast growing field–site trials. By age 12, most of the CV estimates varied between 5 and 15% of the trial mean height.

The correlations between cumulative heights measured at different ages were all positive and, in most cases, reasonably high. They were also positively and linearly related to the natural log of the age ratio (Fig. 4), decreasing as the time interval between measurement ages increased. The estimated regression slope for the forestry trials indicated slightly slower decrease with increasing age interval, than for the test orchard trials. However, the difference was not statistically significant. The forest vs. field sites also showed nearly identical slopes. Thus, the time trend formula for age-age correlations reduced into a simple regression equation with LAR as the only independent variable:

$$r_G = 1.02 + 0.423 \log_e(\text{younger age/older age}) \quad (r^2 = 0.53) \qquad [15]$$

The model [15] was also tested with age difference as the predictor, but this resulted in a somewhat poorer fit (r^2 = 0.44) than the LAR model.

The time–trend functions for the variance components (Table 3, Fig. 5) were used to predict heritability and correlated genetic gains from early selection (Fig. 6). The correlated gains improved steadily towards the target age of 20 years (Fig. 5), as a result of increasing age-age correlation. Under the premise of 150 progeny per family, parental selection produced three times more improvement in age–20 height, than forward selection for an individual's own (adjusted) phenotype. Independently of the mode of selection, forestry trials

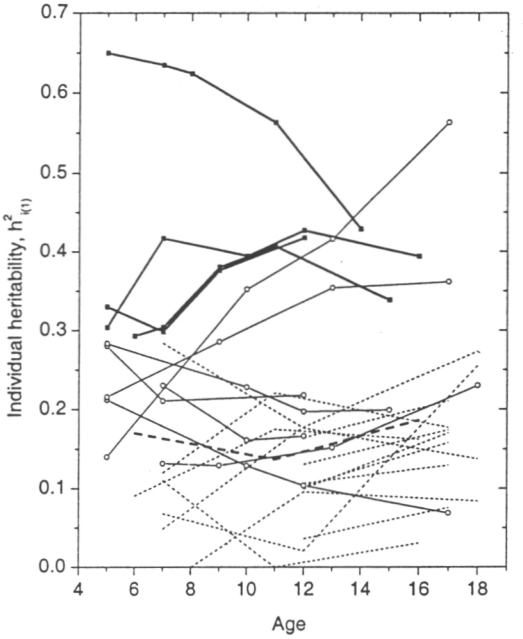

Figure 2. Development of single-site individual-level heritability for tree height in 26 Scots pine progeny trials representing groups of field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines).

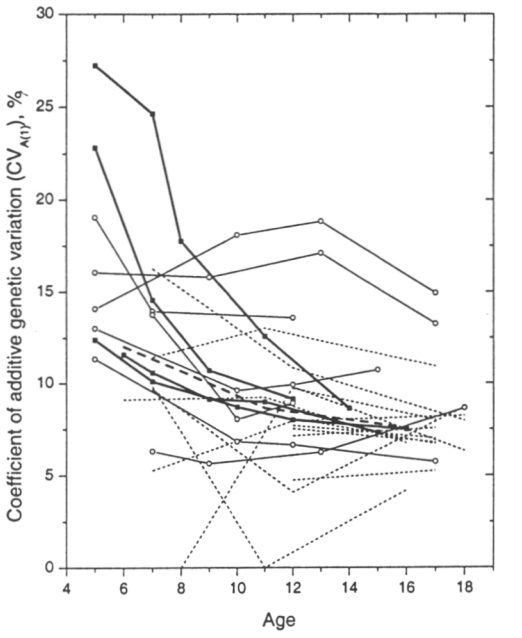

Figure 3. Development of single-site coefficient of variation for tree height in 26 Scots pine progeny trials representing groups of field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines).

on forest soils were significantly less effective than the other types of trial. The inferiority of this combination of testing method and site quality for field testing was evident in all of the scenarios (Fig. 6, 7, 8).

Early parental selection was always more efficient than selection at age 20 in terms of genetic gain per year. The relative efficiency of PS reached the maximum level between ages 5 and 7 (Fig. 7, 8), producing 40 to 60 % more gain per year than direct PS for height at age 20 (Fig. 8). The efficiency of early WFS was initially low and increased over time. However, in contrast to parental selection, the annual responses to WFS were, at best, only slightly greater than responses to direct selection. The age needed to achieve 95% of the gain from direct WFS varied largely in different type of trials (from 8 to 16 years). The relative selection efficiency was found to be fairly sensitive to changes in the regression formula used to predict age-age correlations. When the model presented by LAMBETH (1980) was applied in place of the empiric model [15], the relative efficiencies of early selection were amplified and the maximum correlated responses occurred slightly earlier.

DISCUSSION

Genetic control of tree height

The test orchard trials and forestry trials were associated with distinctly different levels of heritability, providing firm evidence for a significant impact of test characteristics on the magnitudes of genetic and environmental variation. The higher heritabilities of the test orchard group are probably related to the greater edaphic uniformity of the planting sites and to the closer initial spacing. Spacing contributed to the heritabilities through blocking efficiency; when the number of trees in a plot is constant, block size is coupled with planting density. Thus, the close spacing of the test orchards obviously enabled more efficient control of the environmental variability than the four-fold wider spacing of the forestry trials. The effect of spacing on heritability was, in fact, slightly greater than that of site quality, as quantified by the prediction model (Table 3).

The results suggest that estimates of heritability (or any other genetic parameter) are not very informative if reported without giving sufficient reference to the conditions in which they were obtained. For instance,

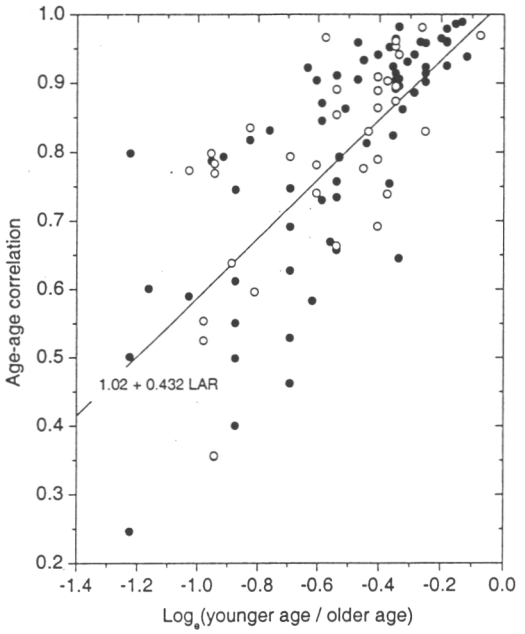

Figure 4. Estimates of age-age correlations vs. the natural logarithm of the ratio of the younger age to the older age (LAR) and the corresponding regression slope (a solid line) from 26 Scots pine progeny trials. The regression slope obtained by LAMBETH (1980) is illustrated by a dashed line. Open and filled dotes denote estimates from forestry trials and test orchards, respectively.

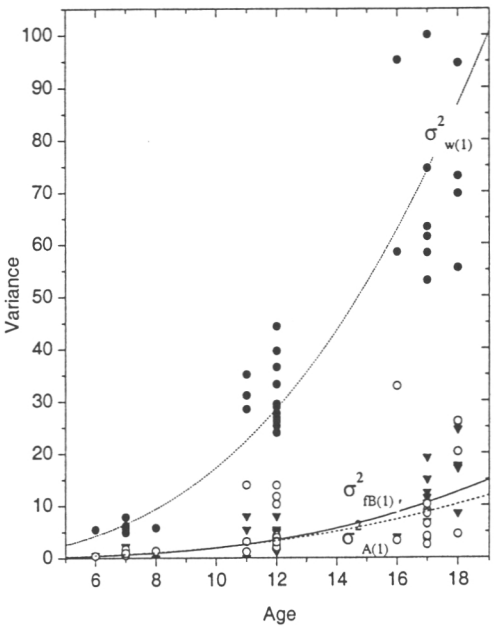

Figure 5. An illustration of single-site variance component estimates [1] and their time-trend models [7] in the group of forest-site forestry trials. Filled triangles, open circles and filled circles denote the single-site estimates of additive genetic variance ($\sigma^2_{A(1)}$), family-by-block variance ($\sigma^2_{fB(1)}$), and residual variance ($\sigma^2_{w(1)}$), respectively. The corresponding estimated time-trends are illustrated with dashed, solid and short-dot lines.

in predicting breeding values for trees, the temporal and environmental heterogeneity of genetic parameters is specifically recognized and taken into account in order to improve the accuracy of predictions (WHITE & HODGE 1989). In the context of forest tree breeding, it is obviously more reasonable to consider heritability as an instant measure of experimental efficiency than a stable parameter that pertains to characteristics of some genetic group (or trait). The predictive models developed in this study for the log-transformed variance components apply to the most common situations of Scots pine progeny testing in Finland. Thus, they could facilitate the prediction of additive genetic values of trees and genetic gain, particularly when there are no data available to allow these parameters to be directly estimated. Nevertheless, updating these models with new data would be beneficial, especially in order to balance the insufficient representation of field–site forestry trials in this study. These type of progeny trials are currently rare in Finland. In future, however, Scots pine progeny trials are to be increasingly laid out on uniform and fertile soils, using significantly wider spacing than in currently ongoing test orchards.

Knowledge of temporal changes in heritability is crucial for devising optimized early selection strategies. Studies with conifers in the genus *Pinaceae* (NAMKOONG *et al.* 1972, FRANKLIN 1979, FALKENHAGEN 1989, HODGE & WHITE 1992, BALOCCHI *et al.* 1993, DIETERS *et al.* 1995, COSTA & DUREL 1996, JOHNSON *et al.* 1997) have found individual heritability for height to be initially low, and to increase with age. For Scots pine, information on age trends in heritability is scarce. However, JANSSON *et al.* (1998) found a similar, slightly increasing tendency for heritability of tree height in three Swedish Scots pine trials measured several times from age 9 up to age 29. In the present study, an increasing pattern of heritability was true for a few of the trials whereas some of the others displayed an opposite decreasing tendency (Fig. 2). In the pooled analysis of the single-site estimates of heritability, however, no clear age trend could be detected (the age coefficient of regression was significant but of negligible size). Corresponding to the finding of this study, stable ratios of family to phenotypic variance with advancing stand age have been reported, e.g., by HANNRUP *et al.* (1998) in Scots pine, LAMBETH *et al.* (1983) and Foster (1986) in loblolly pine (*Pinus taeda*), VASQUEZ and DVORAK (1996) in three species of

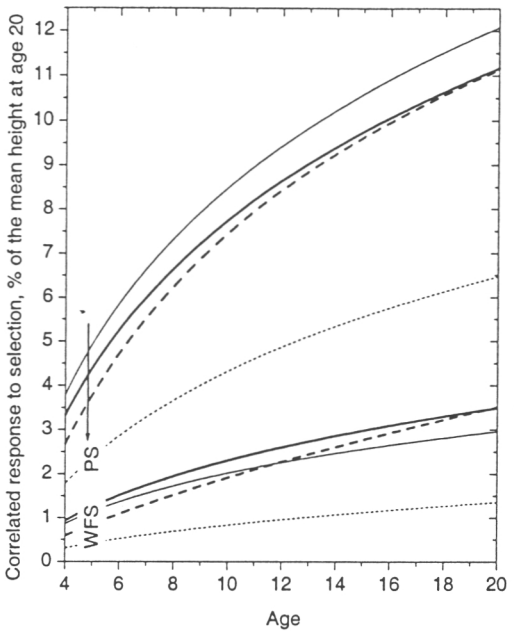

Figure 6. Predicted correlated responses to early parental (PS) and within-family (WFS) selection for field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines). The responses are given per unit of selection intensity and as percentage of the cumulative height at age 20 [8].

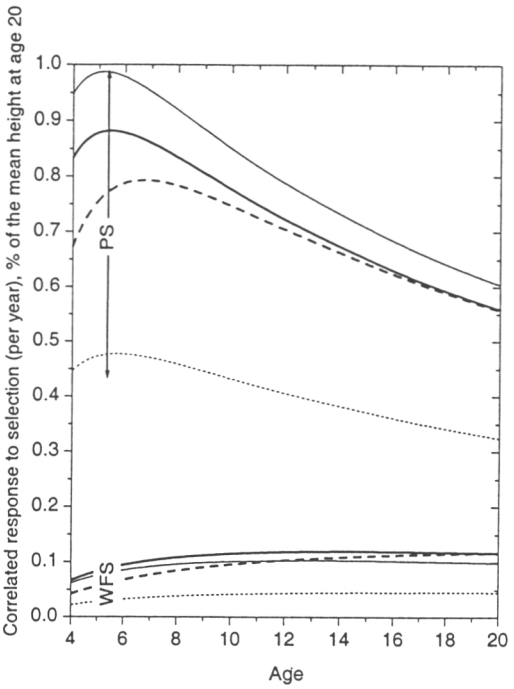

Figure 7. Correlated responses to early parental (PS) and within-family (WFS) selection [11] for field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines). The responses are presented in terms of gain per year and per unit of selection intensity, given as percentage of the cumulative height at age 20.

tropical pines, BASTIEN and ROMAN-AMAT (1990) in Douglas fir, XIE and YING (1996) in lodgepole pine (*Pinus contorta* ssp. *latifolia*), and BENTZER *et al.* (1989) in Norway spruce.

The observed lack of an age trend could be an outcome of two, not mutually exclusive, reasons: (1) the standard errors of the heritability estimates were large enough to conceal possible systematic changes, or (2) inter-family competition had not begun by the time of the latest measurements. According to FRANKLIN's (1979) often-cited hypothesis, heritability can be expected to increase when competition intensifies and begins augmenting differences among families with inherently different growth rates and competitive abilities. The timing of competition is obviously dependent on the mean growth rate (site quality), planting density and changes in spacing due to unplanned mortality. Visual observations made in widely-spaced Scots pine trials suggest that crown closure normally occurs not earlier than 20 to 30 years from planting. Trials with trees at these ages were not represented in this study, and the lack of competition could thus adequately explain the lack of age trend in heritability, at least in the forestry trials. In the dense test orchard stands, canopies typically close much earlier, at around 10 years of age. However, in trials designed to have multiple-tree family plots, the effects of inter-family competition are likely to require some additional years to actualise, so that they were not properly manifested in this study. Furthermore, tree height is less susceptible to crowding than other routinely recorded growth traits (SAKAI *et al.* 1968, KREMER 1992, PAUL *et al.* 1997). The absence of strong competition effects on genetic and phenotypic variances (HAMBLIN & ROSIELLE 1978, FOSTER 1986) in my data apparently simplified the interpretation of the results. MAGNUSSEN (1995) argued that heritability estimates obtained under heavy competition could be severely distorted and lead to false predictions of selection efficiency.

Individual heritability estimates did not couple with the coefficients of additive genetic variation, which diminished with time, in agreement with the commonly noticed inverse relationship between CVs and trait means (HOULE 1992). The values of the CV_As were mostly below 15%, conforming to earlier results (FOSTER 1986, NAMKOONG & CONKLE 1976, CORNE

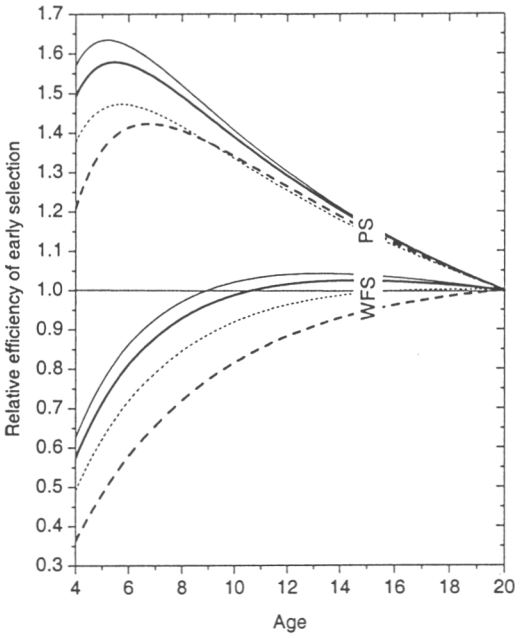

Figure 8. Relative efficiencies of early parental (PS) and within-family (WFS) selection [13] for field – site test orchards (thick solid lines), forest – site test orchards (thin solid lines), field – site forestry trials (thick long-dashed lines), and forest – site forestry trials (thin short-dashed lines).

LIUS 1994, HANNRUP *et al.* 1998).

Impact of family-by-site interaction on genetic parameters and selection

The varying response of the families to the test environments inflated the estimates of the additive genetic variance and heritability by half. As a result, the estimates corrected for the bias were substantially smaller than those usually reported for tree height in conifers. The mean heritability of height in the 67 studies reviewed by CORNELIUS (1994) was 0.28, whereas in this study the mean of the unbiased estimates was less than half of that, 0.12. Type-B correlations were comparable to those found in earlier studies in which the values have mostly fallen between 0.6 and 0.8, corresponding to bias proportions of 66 to 25 %, respectively (ADAMS *et al.* 1994, DIETERS *et al.* 1995, HAAPANEN 1996, JOHNSON *et al.* 1997). DIETERS *et al.* (1995) and JOHNSON *et al.* (1997) found type-B correlations to increase slowly with time, indicating diminishing importance of the G×E. In this study, there was no sign of a time trend for the type-B correlations.

SHELBOURNE (1972) suggested that if family-by-environment interaction variance exceeds the family component of variance by half or more, the interaction can seriously impair selections based on intra-site information. Although this limit was exceeded in many of the across-site analyses, the biological meaning of the observed interaction is not clear. Consequently, it is difficult to estimate the loss of potential gain, the most important outcome of the genotype-by-environment interaction. However, the loss is likely to be less than what mechanical calculations would suggest, if the statistical family-by-site interaction variance were straightforwardly interpreted to be equivalent with the biologically significant G×E. That would hardly be justified in the field testing of forest trees, where large interactions commonly occur between families and sites within the same region, and even between families and block replications at a single trial (MATHESON & COTTERILL 1990). As environmental variability among trial sites is typically random, also the interactions between families and sites are random and not repeatable (MATHESON & COTTERILL 1990). In this situation, breeders should accept the presence of the interaction rather than try to gain from it. In Scots pine breeding, the goal of genetic testing is to find individuals that perform well over a variety of sites within a climatically determined breeding zone. The fairly low correlations between parallel trials underscore the need to account for site-to-site variation in family performances. First of all, each progeny trial should be replicated at a sufficient number of sites. Five test localities, suggested by LINDGREN (1984) for Scots pine testing in Sweden, is probably adequate. In addition, information on candidate trees and their siblings in other trials should be employed in selection. This especially concerns trees selected forward, as parental (backward) selection is, as a rule, based on results from multiple test sites and is therefore less prone to the possible biasing effects due to the family-by-site interaction.

BARNES *et al.* (1984) concluded that explaining and using G×E in breeding is realistic only when a single environmental factor affects an economically important trait in a predictable manner. When a growth trait is considered, large differences in family performances from one site to another could be assumed to be related to overall differences in site productivity. Indeed, HODGE & WHITE (1992) were able to demonstrate that type-B genetic correlations were linked to differences in site quality in slash pine (*Pinus elliottii*). However, no sign of this kind of association was found in this study, even though the trial sites were markedly different in terms of height growth. Most of the variation among the across-site correlations remained unexplained, suggesting that while the magnitude of the family-by-site interaction in young progeny trials is not negligible, its nature is highly random and unpredictable and thus of little use for common breeding purposes. It must be stressed, however, that these trials

were designed with the estimation of GCA values as the primary objective, and are not optimal for the study of G×E. A thorough investigation of this issue would require a large sample of families replicated over several well-defined sites.

Age-age correlations

Sufficiently strong genetic correlation between the selection and the target age is a prerequisite for successful early selection. Considering the importance of age-age correlations to all kinds of predictions of future gain, there is a surprising gap in knowledge with regard to the magnitudes of age-age correlations in Scots pine progeny testing. Encouragingly strong age-age correlations have been reported for performances of full-sib families grown in a growth chamber and in older field trials (JONSSON 2000), and also for some other important Scots pine traits such as wood density (HANNRUP & EKBERG 1998). However, the patterns of field–trial correlations are less well known. This study utilized correlations between least-square family means to approximate genetic correlations (LAMBETH *et al.* 1983, NEWMAN & WILLIAMS 1991). These estimates are contaminated by environmental (co)variances, and are thus potentially biased (JANSSON 2000). However, the effect of the bias diminishes with sample size and is likely be of small importance when the family size exceeds 20 individuals (ROFF & PREZIOSI 1994). As family sizes in Finnish progeny trials are typically large (> 100), family-mean correlations may be expected to be of good accuracy. Furthermore, simulations have shown that true genetic correlations are in many situations estimated with worse precision by direct estimates of genetic correlation than by phenotypic correlations (ROFF 1995), as the former estimates are highly sensitive to the number of families in the sample (KLEIN *et al.* 1973, NAMKOONG 1979, HODGE & WHITE 1992).

As a rule, age-age correlations in growth traits decline with increasing age interval. LAMBETH (1980) reanalysed phenotypic age-age correlations from a number of studies, and found that they could be reliably predicted by means of a linear regression, using the logarithmic ratio of the two ages as the explanatory variable. This same approach has since been frequently used to project age-age correlations (MCKEAND 1988, KING & BURDON 1991, JOHNSON *et al.* 1997). In the present study, the regression method worked acceptably. The slopes were not significantly different between trials on forest vs. field sites. The same was true for the comparison between the forestry and test orchard trials, although the assumption of homogeneity was not quite as clear in this case. Interestingly, the coefficient of regression obtained in this study (0.423) was very close to that reported by GWAZE *et al.* (2000) for young fast-growing *Pinus taeda* L. genetic tests (0.447). Both of these slopes are markedly steeper than that of Lambeth's 'universal equation' (0.308). The discrepancy may be partly due to the fact that the range of the LARs was only half of that in the study of LAMBETH (1980). Family-mean correlations may also more accurately reflect changes in the genetic mechanism controlling the development of tree height than individual-tree phenotypic correlations, which apparently comprised most of the data in LAMBETH's (1980) investigation. Furthermore, many other studies indicate that the function relating age-age correlations to the log of the age ratio is not invariant, in contrast to LAMBETH (1980) who suggested that a single predictive model given in his study would be valid for a wide range of experiments and species. For instance, JOHNSON *et al.* (1997) and GWAZE *et al.* (2000) reported the regression slope to significantly vary depending on the breeding zone. To obtain a realistic view of the true nature of the LAR relationship in various types of Scots pine progeny trials, they should obviously be followed for a longer period of time, probably at least till half-rotation.

Progeny testing methods and selection efficiency

All the scenarios indicated that the highest responses to selection were associated with the test orchard trials. This is consistent with earlier views advocating close-spaced field trials, established on high quality sites, for progeny testing (FRANKLIN 1979, MIKOLA 1985, CAMPBELL *et al.* 1986, MAGNUSSEN & YEATMAN 1986, WOODS *et al.* 1995, MAGNUSSEN 1995, HAAPANEN 1996, BRIDGWATER & MCKEAND 1997). The superiority of the test orchard method was clearly associated with the high average levels of heritability because the age-age correlations in the fast-growing test orchards were not significantly different from the other types of trials. In an investigation of a number of 10-year-old Scots pine progeny trials (HAAPANEN 1996), test orchard trials laid out on agricultural land discriminated genetic differences considerably better than parallel forestry trials. WOODS *et al.* (1995), who compared the corresponding testing methods in Douglas fir, found that selection in 'farm-field' conditions in all cases provided greater genetic gains in stem yield and wood density than selection in 'field trials' (forestry trials). Similarly, CARLSON (1990), in lodgepole pine, found single-tree plot farm-field trials to be more productive and have greater family heritability values than parallel 'wild field sites'. Hence, although the value of long-term data accumulating from traditional forestry trials is indisputable, they appear not to be the optimal choice when the primary goal is the precise and cost-efficient

ranking of genotypes. This especially concerns widely-spaced trials on forestland, which performed inferiorly in comparison to the other combinations of spacing and site quality.

In Finland, ages from 10 to 15 have been suggested sufficient for selection for growth traits in test orchard conditions (MIKOLA 1985). Similar estimates have been presented for other conifers (NAMKOONG *et al.* 1972, SQUILLACE & GANSEL 1974, FRANKLIN 1979). In general, the optimal selection ages reported for various species in the tree breeding literature show a lack of consistency, which, in view of the present results, is likely to owe much to the variable circumstances under which genetic parameters are estimated. The results obtained here suggest that the optimal age for selection occurs later for slow growing than for fast-growing trials. This especially concerning forward selections (Fig. 8). Moreover, the optimal age for WFS and PS was different. The earlier optimum for PS conforms to a number of earlier findings (LAMBETH *et al.* 1983, MCKEAND 1988, BALOCCHI *et al.* 1994, JOHNSON *et al.* 1997) and was, in fact, expected since PS is based on more information than WFS. The form of the relative efficiency function (for PS) was close to that reported by JOHNSON *et al.* (1997) for Douglas-fir. Unlike the relative efficiencies, the absolute responses to early selection kept raising steadily over age. Therefore, if juvenile selection is not urgent, there is nothing to lose if selection is deferred some years later than the optimum (ignoring the loss due to the lower efficiency per year). On the contrary, this may be reasonable considering the possible changes in the ranking of genotypes that can occur between the selection and true economic rotation ages.

The selection efficiencies predicted here are associated with a number of simplifications and uncertainties, and they should therefore be evaluated with prudence. Most importantly, the target age was set at 20 years, which can hardly be considered as a mature age, since it is equivalent to only one-fourth to one-sixth of the commercial rotation of Scots pine in Finnish conditions. Regrettably, there was no alternative for setting the goal age other than at 20 years because of the scarcity of older assessment information. The use of the LAR regression function to extrapolate selection efficiencies to later ages not covered by the data was not considered as a viable option; the results would have been speculative at best, and highly misleading at worst. However, there may be some justifications for choosing an early target. Firstly, prolonging the progeny testing much beyond 20 years is not likely to be very profitable since many important traits in Scots pine are already established by this age. Secondly, measurement of height and, consequently, estimation of bole volume, becomes more laborious and imprecise with the increasing physical size of trees. Finally, earlier studies consistently suggest that genetic gains per unit time peak at a relatively early stage of stand development (LAMBETH 1980, LAMBETH *et al.* 1983, MCKEAND 1988). The importance of the obviously imperfect genetic association between age-20 and rotation age (roughly 80 years in southern Finland) performances is difficult to evaluate, but the bias is likely to be tolerable. For instance, studies with other pines (with notably shorter rotations, though) have promisingly indicated that early height could be a good predictor of rotation-age volume and vigour (LAMBETH *et al.* 1983, FOSTER 1986, COSTA & DUREL 1996). However, tree height is seldom the only selection criterion in the context of teenaged Scots pine trials. Other traits, especially those related to the branching quality of the butt log, become increasingly more important in determining the end-product value as the trees mature. Considering the importance of multi-trait selection, the genetic determinism of these traits should also be explored to assure efficient operation of the Scots pine breeding programme.

REFERENCES

ADAMS, W., WHITE, T., HODGE, G. & POWELL, G. 1994: Genetic parameters for bole volume in longleaf pine: large sample estimates and influences of test characteristics. *Silvae Genet.* **43** (5–6): 357–366.

BALOCCHI, C., BRIDGWATER, F. & ZOBEL, B. & JAHROMI, S. 1993: Age trends in genetic parameters for tree height in a nonselected population of loblolly pine. *For. Sci.* **39**(2): 231–251.

BALOCCHI, C., BRIDGWATER, F. & BRYANT, R. 1994: Selection efficiency in a nonselected population of loblolly pine. *For. Sci.* **40**(3): 452–473.

BARNES, R. BURLEY, J, GIBSON, G. & GARCIA DE LEON, J. 1984: Genotype-environment interactions in tropical pines and their effects on the structure of breeding populations. *Silvae Genet.* **33**(6): 186–198.

BASTIEN, J. & ROMAN-AMAT, B. 1990: Predicting Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) volume at age 15 with early traits. *Silvae Genet.* **39**(1): 29–35.

BENTZER, B., FOSTER, G., HELLBERG, A. & PODZORSKI, A. 1989: Trends in genetic and environmental parameters, genetic correlations, and response to indirect selection for 10-year volume in a Norway spruce clonal experiment. *Can. J. For. Res.* **19**: 897–903.

BOUVET, J. & VIGNERON, P. 1995: Age trends in variances and heritabilities in eucalyptus factorial mating design. *Silvae Genet.* **44**(4): 206–216.

BRIDGWATER, F. & MCKEAND, S. 1997: Early family evaluation for growth of loblolly pine. *Forest Genetics* **4**(1): 51–58.

BURDON, R. 1977: Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. *Silvae Genet.* **25**(5–6): 168–175.

CAMPBELL, R., ECHOLS, R. & STONECYPHER, R. 1986: Genetic variances and interactions in 9-year-old Douglas-fir grown at narrow spacings. *Silvae Genet.* **35**(1): 24–32.

CARLSON, M. 1990: Early testing and selection of lodgepole pine in British Columbia. *In*: Proceedings of the joint meeting of Western Forest Genetics Association and IUFRO Working Parties S2.02–05, 12 and 14. Olympia, Washington, August 20–24, 1990: 4.22. pp. 1–13.

CORNELIUS, J. 1994: Heritabilities and additive genetic coefficients of variation in forest trees. *Can. J. For. Res.* **24**: 372–379.

COSTA, P. & DUREL, C. 1996: Time trends in genetic control over height and diameter in maritime pine. *Can. J. For. Res.* **26**: 1209–1217.

COTTERILL, P. & DEAN, C. 1990: Successful tree breeding with index selection. CSIRO Australia. 80 p.

DANELL, Ö. 1988: Arbetsgång vid bearbetning av contorta-försök. Inst. for Forest. Improv. Arbetrapport 219. Uppsala. [In Swedish].

DANJON, F. 1994: Heritabilities and genetic correlations for estimated growth curve parameters in maritime pine. *Theor. Appl. Genet.* **89**: 911–921.

DICKERSON, G. 1962: Implications of genetic-environmental interaction in animal breeding. *Anim. prod.* **4**: 47–63.

DICKERSON, G. 1969: Techniques for research in quantitative animal genetics. *In*: Techniques and procedures in animal science research. *Am. Soc. Anim. Sci.*, Albany, N.Y.: 36–79.

DIETERS, M., WHITE, T. & HODGE, G. 1995: Genetic parameter estimates for volume from full-sib tests of slash pine (*Pinus elliottii*). *Can. J. For. Res.* **25**: 1397–1408.

FALCONER, D. 1981: Introduction to Quantitative Genetics. Longman Grp Ltd. 340 p.

FALKENHAGEN, E. 1989: Relationships between some genetic parameters and test environments in open-pollinated families of *Pinus elliottii* in South Africa. *Theor. Appl. Genet.* **77**: 857–866.

FOSTER, S. 1986. Trends in genetic parameters with stand development and their influence on early selection for volume growth in loblolly pine. *For. Sci.* **32**(4): 944–959.

FRANKLIN, E. 1979: Models relating levels of genetic variance to stand development of four north american conifers. *Silvae Genet.* **28**(5–6): 207–212.

GILL, J. 1987: Juvenile-mature correlations and trends in genetic variances in Sitka spruce in Britain. *Silvae Genet.* **36**(5–6): 189–194.

GWAZE, D, BRIDGWATER, F., BYRAM, T., WOOLLIAMS, J. & WILLIAMS, C. 2000: Predicting age-age genetic correlations in tree-breeding programs: a case study of *Pinus taeda* L. *Theor. Appl. Gen.* **100**: 199–206.

HAAPANEN, M. 1996: Impact of family-by-trial interaction on the utility of progeny testing methods for Scots pine. *Silvae Genet.* **45**(2–3): 130–135.

HAMBLIN, J. & ROSIELE, A. 1978: Effect of intergenotypic competition on genetic parameter estimates. *Crop Science* **18**: 51–54.

HANNRUP, B. & EKBERG, I. 1998: Age-age correlations for tracheid length and wood density in *Pinus sylvestris*. *Canadian Journal of Forest Research* **28**: 1373–1379.

HANNRUP, B., WILHELMSSON, L. & DANELL, Ö. 1998: Time trends in genetic parameters of wood density and growth traits in *Pinus sylvestris* L. *Silvae Genet.* **47**(4): 214–223.

HODGE, G. & WHITE, T. 1992: Genetic parameter estimates for growth traits at different ages in slash pine and some implications for breeding. *Silvae Genet.* **41**(4–5): 252 –262.

HOULE, D. 1992: Comparing evolvability and variability of quantitative traits. *Genetics* **130**: 195–204.

JANSSON, G. 2000: Theory of early testing. In: Proc. from the SNS meeting, July 1–3, 1999, Uppsala, Sweden. Dept. of Forest Genetics, Research Notes 55.16–24:

JANSSON, G., DANELL, Ö. & STENER, L. 1998: Correspondence between single-tree and multiple-tree plot genetic tests for production traits in *Pinus sylvestris*. *Can. J. For. Res.* **28**(3): 450–458.

JOHNSON, G., SNIEZKO, R. & MANDEL, N. 1997: Age trends in Douglas-fir genetic parameters and implications for optimum selection age. *Silvae Genet.* **46**(6): 349–358.

JONSSON, A. 2000: A decade of early testing research on *Pinus sylvestris* and *Picea abies*. *In*: Proc. from the SNS meeting, July 1–3, 1999, Uppsala, Sweden. Dept. of Forest Genetics, Research Notes 55.25–36.

KING, J. & BURDON, R. 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.

KLEIN, T., DEFRIES, J. & FINKBEINER, C. 1973: Heritability and genetic correlation: standard errors of estimates and sample size. *Behavior Genetics* **3**(4): 355–364.

KREMER, A. 1992: Predictions of age-age correlations of total height based on serial correlations between height increments in maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.* **85**: 152–158.

LAMBETH, C. 1980: Juvenile-mature correlations in Pinaceae and implications for early selection. *For. Sci.* **26**(4): 571–580.

LAMBETH, C., VAN BUIJTENEN, J., DUKE, S. & MCCULLOUGH, R. 1983: Early selection is effective in 20-year-old genetic tests of loblolly pine. *Silvae Genet.* **32**(5–6): 210–215.

LINDGREN, D. 1984: Prediction and optimization of genetic gain with regard to genotype × environment interactions. *Studia Forestalia Suecica.* **166**: 15–24.

LITTELL, R., MILLIKEN, G., STROUP, W. & WOLFINGER, R. 1996: SAS System for Mixed Models. SAS Institute Inc., Cary, NC. 633 pp.

LOO-DINKINS, J. 1992: Field test design. *In:* Handbook of quantitative forest genetics. L. Fins, S. Friedman & J. Brotschol (editors). Kluwer Academic Publ.: 96–139.

MAGNUSSEN, S. 1991: Efficiency of early selections for stem volume and predictions of size distributions of selections in a red pine spacing trial. *For. Sci.* **37**(2): 593–612.

MAGNUSSEN, S. 1993: Design and analysis of tree genetic trials. *Can. J. For. Res.* **23**: 1144–1149.

MAGNUSSEN, S. 1995: The impact of genetic variation in relative growth rates on stem volume differentiation: a simulation study. *Silvae Genet.* **44**(4): 194–205.

MAGNUSSEN, S. & YEATMAN, C. 1986: Accelerated testing of Jack pine progenies: A case study. *In*: Proceedings of the joint meeting of IUFRO working parties on Breeding Theory, Progeny Testing, Seed Orchards. Williamsburg, Virginia, October 13–17, 1986. 107–121.

MATHESON, A. & COTTERILL, P. 1990: Utility of genotype × environment interactions. *Forest Ecology and Management* **30**: 159–174.

MCKEAND, S. 1988: Optimum age for family selection for growth in genetic tests of loblolly pine. *For. Sci.* **34**(2): 400–411.

MIKOLA, J. 1985: Methods used for the genetic evaluation of tree breeding material in Finland. *In*: Crop physiology of forest trees. P. Tigerstedt, P. Puttonen & V. Koski (editors). Helsinki University Press: pp. 225–231.

NAMKOONG, G., USANIS, R. & SILEN, R. 1972: Age-related variation in genetic control of height growth in Douglas-fir. *Theor. Appl. Genet.* **42**: 151–159.

NAMKOONG, G. & CONKLE, M. 1976: Time trends in genetic control of height growth in ponderosa pine. *For. Sci.* **22**: 2–12.

NAMKOONG, G. 1979: Introduction to Quantitative Genetics in Forestry. U.S. Dept. Agric., Tech. Bull. No. 1588. 342 p.

NEWMAN, D. & WILLIAMS, C. 1991: The incorporation of risk in optimal selection age determination. *For. Sci.* **37**(5): 1350–1364.

PAUL, A., FOSTER, G., CALDWELL, T. & MCRAE, J. 1997: Trends in genetic parameters for height, diameter, and volume in a multilocation clonal study with loblolly pine. *For. Sci.* **43**(1): 87–98.

RINK, G. & CLAUSEN, K. Site and age effects on genotypic control of juvenile *Juglans nigra* L. tree height. *Silvae Genet.* **38**(1):17–21.

ROFF, D. 1995: The estimation of genetic correlation from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* **74**: 481–490.

ROFF, D. & PREZIOSI, R. 1994: The estimation of the genetic correlation; the use of the Jacknife. *Heredity* **73**:544 –548.

SAS Institute Inc. 1992: SAS Technical Report P-229. SAS/STAT Software: Changes and enhancements. Release 6.07.: 289–366.

SAKAI, K., MUKAIDE, H. & TOMITA, K. 1968: Intraspecific competition in forest trees. *Silvae Genet.* **17**(1): 1–5.

SATO, T. 1994: Time trends for genetic parameters in progeny tests of *Abies sachalinensis* (Fr. Schm.) Mast. *Silvae Genet.* **43**(5–6): 304–307.

SHELBOURNE, C. 1972: Genotype-environment interaction: its study and implications in forest tree improvement. In: Proceedings of the IUFRO Joint Genetics Symposium, Japanese Forest Experiment Station: 1–28.

SONESSON, J. & ERIKSSON, G. 2000: Genotypic stability and genetic parameters for growth and biomass traits in a water × temperature factorial experiment with *Pinus sylvestris* L. seedlings. *For. Sci.* **46**(4): 487–495.

SQUILLACE, A. & GANSEL, C. 1974: Juvenile:mature correlations in slash pine. *For. Sci.* **20**(3): 225–229.

VÁSQUEZ, J. & DVORAK, W. 1996: Trends in variances and heritabilities with stand development of tropical pines. *Can. J. For. Res.* **26**: 1473–1480.

WHITE, T. & HODGE, G. 1989: Predicting Breeding Values with Applications in Forest Tree Improvement. Kluwer Academic Publishers. 367 p.

WHITE, T. & HODGE, G. 1992: Test designs and optimum age for parental selection in advanced-generation progeny tests of Slash pine. *Silvae Genet.* **41**(4–5): 293–302.

WOODS, J., KOLOTELO, D. & YANCHUK, A. 1995: Early selection of coastal Douglas-fir in a farm field test environment. *Silvae Genet.* **44**(4): 178–186.

XIE, C. & YING, C. 1996: Heritabilities, age-age correlations and early selection in lodgepole pine (*Pinus contorta* ssp. *latifolia*). *Silvae Genet.* **45**(2–3): 101–107.

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