Spatiotemporal patterns of genetic variation for growth and fertility in Scots pine

Johan Kroon
Faculty of Forestry
Department of Forest Genetics and Plant Physiology
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2011
Spatiotemporal patterns of genetic variation for growth and fertility in Scots pine

Abstract
Genetic testing plays an important role in tree breeding. Understanding basic genetic parameters for growth and fertility is a prerequisite for developing a successful breeding strategy to balance between short- and long-term gains. It is both impractical and inefficient to wait until trials are several decades old to make selection decisions. It is therefore of great value to evaluate a sample of long-term experiments to confirm selection efficiency at ages closer to rotation time. This thesis investigates the genetic expression of field performance in Scots pine for growth and fertility in long-term experiments over time and across sites.

The data revealed large variation in genetic expression over time and across sites. The additive effect was the most important source of genetic variation for growth, while estimates of dominance were small, site-specific, and decreased much with age. Thus, there is little benefit in attempting to explore dominance through specific combining ability to improve genetic productivity in northern Sweden.

Progeny of plus-trees selected in northern Sweden showed faster growth compared to unimproved controls at age 30, as well as genetic differences in stem shape, such that improved trees were more slender. The genetic relationship between height and diameter in northern Sweden demonstrated the importance of considering diameter in selection to obtain greater genetic gain in volume.

Overall, the results show higher heritability at older ages and genetic correlations that reveal important rank changes with time and across environments.

Clonal variation in female fertility in mature seed orchards is rather small and varies as much within clones as among clones, and is heavily dependent on year of assessment. The prospects for early selection of clones for future seed cone production based on a single-year observation are low.

Finally, this thesis illustrates the importance of subjecting data from long-term field tests to a multi-trait, multi-site analysis accounting for environmental effects.

Keywords: Early testing, breeding stock, trait expression, age-age correlation, G-matrix, longitudinal data, best linear unbiased prediction, multi-environment trials, spatial analysis, competition.

Author's address: Johan Kroon, SLU, Department of Forest Genetics and Plant Physiology. P.O. Skogforsk, Box 3, SE-918 21 Sävar, Sweden
E-mail: johan.kroon@skogforsk.se
To my family

Essentially, all models are wrong, but some are useful.

George Box
Contents

List of Publications 7

1 Introduction 9
   1.1 Forestry and tree breeding 9
   1.2 Scots pine breeding in Sweden 12

2 Aims of the thesis 15

3 Genetic review, model considerations and concepts 17
   3.1 Genetic parameters 17
   3.2 Multi-trait selection 19
   3.3 Age-specific genetic expression 21
   3.4 Site-specific genetic expression 22

4 Materials and methods 25
   4.1 Overview of populations and field material 25
   4.2 Multiple trait genetic analysis 26
   4.3 Growth in progeny tests (I-IV) 28
   4.4 Fertility in clonal seed orchards (V-VI) 33

5 Main results and discussion 35
   5.1 Performance and genetic variation in improved trees 35
   5.2 Additive and non-additive genetic variation in growth 37
   5.3 Age trends in genetic expression of growth 38
   5.4 Environmental trends in genetic expression of growth 39
   5.5 Clonal variation in female fertility 40

6 Conclusions and implications for breeding 43

References 47

Acknowledgement 57
List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:


III Kroon, J., Ericsson, T., Persson, T. & Andersson, B. Additive and non-additive genetic variation in height of Pinus sylvestris changes with age. (Manuscript).


Papers I, II and IV–VI are reproduced with the permission of the publishers.
1 Introduction

1.1 Forestry and tree breeding

Tree breeding programmes have the objective to improve the profitability of forest management by improving the genetic constitution of populations in the context of an uncertain future silviculture and markets. Biomass production is of primary interest for operational forestry, although other characteristics are important for specific end-uses. Industrial requirements are hardly predictable a century ahead, so it is most appropriate that breeding objectives are general in nature, such as stem volume production (per unit area), adaptation, stem/wood quality, and robustness over environmental gradients (Rosvall et al., 2011). To date, long-term breeding of pines has focused mainly on growth, controlling for stem quality, mortality and susceptibility to damage (e.g., Wilhelmsson & Andersson, 1993; Xiang et al., 2003b; McKean et al., 2008; Bouffier et al., 2009).

Early evaluation of improved pines in multi-trial analyses shows that they are superior to unimproved trees with respect to height growth (e.g., Buford and Burkhart, 1987; Li et al., 1999; Lambeth, 2000; Andersson et al., 2003). Realised-gain calculations from field-test data are important to verify further breeding achievements. For Scots pine (Pinus sylvestris L.) in Sweden, the use of genetically improved regeneration materials is very profitable, due to low investment cost and the substantial increase in forest growth (Simonsen et al., 2010). No other silviculture practice will increase growth of Swedish forests as much as planting genetically improved seedlings (Rosvall, 2007; Rosvall & Lundström, 2011).

While the value of tree breeding is well established, prediction of full-rotation volume and value production requires further characterisation and quantification. Especially in northern Sweden, genetic tests should be
evaluated at ages older than 12-16 year from regeneration. After this period, mortality of less cold-hardy trees decreases and sensitivity to environmental disturbances is lower as trees become taller (Ståhl & Andersson, 1985; Persson & Ståhl, 1993).

Success of long-term recurrent tree breeding rests with the genetic variation and heritability of important traits. Understanding basic parameters is a prerequisite for developing a successful breeding strategy to balance short- and long-term gains (e.g., Rosvall, 1999). Examining the genetic architecture of the breeding population, at different ages or in different environments, gives information about optimal selection and how much genetic gain can be expected. Improved seeds from today’s Scots pine breeding effort are delivered to operational forestry through seed orchards, but it may be several decades between selection of orchard parents and actual deployment of stock to plantations. The realised gains from these orchards will also depend on the reproductive success of contributing clones and on the correlation between reproductive success and the breeding value for growth/quality of the clones.

Field testing is currently the most important tool for evaluating phenotypic expression in quantitative traits and for inferring genetic value from the breeding stock. It can provide information on inheritance patterns (genetic parameters) and genetic value of individual trees. Genetic testing thus plays an important role in the breeding cycle for selection of individual trees whose genes should be transferred to the next generation, and for deployment to plantations (Fig. 1).

Figure 1. The breeding cycle illustrated over time. BP=Breeding populations. The size of the arrows represents the amount of genetic variation at different phases of the cycle.
The breeding cycle involves selecting good trees and mating (crossing) these for creation of candidate trees, which are then field tested. It is both impractical and inefficient to wait until trials are several decades old to make selection decisions. It is therefore of great value to evaluate a sample of long-term experiments to confirm selection efficiency at ages closer to rotation time. As improved material for operational forestry is planted on land with large environmental heterogeneity, the selected trees should perform well across a range of sites covering the intended production area.

While phenotypic expression of stem and wood quality are demanding traits to assess operationally, wood properties are now commonly considered for selection on pines (e.g., Ivković et al., 2006; Isik et al., 2008; Rosvall et al., 2011). Research has shown the potential for improvement of certain quality traits, e.g., wood density and fibre morphology, as these are usually under moderate to strong genetic control (Zobel & Jett, 1995; Ericsson & Fries, 1999; Jayawickrama, 2001; Hannrup et al., 2003; Fries & Ericsson, 2006, 2009; Baltunis et al., 2007; Gaspar et al., 2008; Wu et al., 2008). At the same time, adverse genetic correlations between traits raise concerns that genetic gain in quality could be negatively affected when selecting for growth (e.g., Sánchez et al., 2008; Wu et al., 2008; Gaspar, 2009).

Using standard growth-and-yield functions for prediction of full-rotation genetic gain in volume production based on early height assessments will be biased if the genetic correlation between height and diameter is low. Stem shape (relationship between height and diameter) is thus a fundamental trait to evaluate in bred material, as it affects individual-tree volume, as well as other stem and wood-quality traits, such as branch growth dynamics (Mäkinen, 1999) and wood stiffness (Waghorn et al., 2007). It differs among tree species (e.g., La Farge, 1972; Wang et al., 1998; Rymer-Dudzinska & Tomusiak, 2000; Poorter et al., 2003) and among populations (shown for Scots pine by Wright, 1976; Rymer-Dudzinska, 1992a). It has a practical impact on tree stability, snow damage and wind breakage (Weihe, 1977; Erteld, 1979; Abetz, 1976; Konopka et al., 1987; Rymer-Dudzinska, 1992b; Richter, 1996; Wang et al., 1998; Pretzsch, 2009). Growing conditions are, however, the most important factors determining the shape of an individual tree, e.g., spacing, biomechanical factors (such as wind loads), competitive status, site quality and age (Abetz, 1976; Konopka et al., 1987; Niklas, 1995; Wang et al., 1998; Ilomäki et al., 2003; Pretzsch, 2009). Selection for height alone could result in favouring stems that are more slender than average; this may or may not be favourable, but the correlated response will also affect volume production. The height-diameter relationship is therefore an important variable, and few studies have been conducted in this area.
Seed orchards are the source of genetically improved seed for Scots pine. Seed yield is important for the economics of seed orchards and for determining the areas required. Prediction of average seed yield is crucial for determining the productive life of orchards and their yield per unit area over time. Early flowering would obviously enhance the harvest yield in seed orchards. One way to increase production is to select clones that are prolific seed producers, although flowering has generally not been considered when orchard parents are selected (Kang et al., 2001).

Variation in female fertility (e.g., cone production) has consequences both for breeding and genetic variation in seed crops. If trees reach flowering competence early, the time to mate selected trees and produce a new candidate population can be shortened. Additionally, information about flowering capacity can help optimise the composition of clones (female and males) in a new seed orchard, and could be used for enhanced orchard management (e.g., genetic thinning, selective harvest, etc.). While yearly variation in flowering is well recognised, the genetic component of this variation in Scots pine has not been thoroughly studied. A previous study has reported considerable variation in fertility among clones and ramets in young Scots pine orchards across years (Jonsson et al., 1976).

1.2 Scots pine breeding in Sweden

Sweden was one of the first countries to organise breeding efforts in nationally financed tree improvement programmes with strong support from the forest industry. This support has continued and tree breeding is regarded as a profitable investment, as evidenced by the establishment of a third round of Scots pine seed orchards (Rosvall et al., 2002). Breeding programmes for the commercially most important native tree species including Scots pine were initiated in the mid-20th century. The present programme was developed during the 1980’s and is organised as a large, national meta-population with a number of breeding populations for various environments (Wilhelmsson & Andersson, 1993). Around this same time, breeding populations were expanded through additional selection of superior “plus-trees”.

A period of intensive activity took place during the 1970’s when most of today’s advanced mature field trials were established. These tests arose from ambitious mating schemes (partial diallel crossing design) between plus-trees grafted in first-generation seed orchards. The trials were designed as randomised, single-tree plots (which have been shown to be efficient for
early evaluation by Ericsson, 1997, 1999a; Jansson et al., 1998) and are still actively managed by Skogforsk as part of the long-term programme.

The original selection of plus-trees in Sweden targeted the improvement of height growth, timber quality and vigour (cf. Werner et al., 1981). The use of improved seed for forest regeneration is standard operational practice in Swedish forestry today. Consequently, the demand for orchard seed has increased. Early evaluation of Scots pine genetic field trials has so far focussed on tree height as the most important trait. Survival is of great importance for stand productivity in boreal conditions, where site harshness affects tree vitality and mortality rates. Previous investigation of Scots pine provenances has shown that genetic differences evolved in different environments with a continuous (clinal) variation in both tree height and field survival (Langlet, 1936; Persson & Ståhl, 1993; Persson, 1994). Survival of improved material is of importance if genetic gains for biomass production are to be realised.

A recent review of the breeding plan identified a need for increased effort on changing environmental and management conditions, in addition to the more general breeding objectives, focusing on tree adaptation and adaptability to changing conditions (Rosvall et al., 2011). The review attached great importance to further field testing (preferably by clonally-replicated testing), where response to variation in site conditions could mimic the capacity for tree-to-stand temporal changes. In addition, management and research for shortening the period to achieve flowering were also emphasised.
2 Aims of the thesis

The overall objective of this thesis was to evaluate the genetic expression of field performance in growth and fertility for Scots pine in long-term experiments. The thesis should provide information to facilitate selection decisions for breeding and deployment objectives in the Swedish breeding programme. The following questions were addressed:

- How is performance and genetic variation in growth expressed at maturity, considering competition? It was of special importance to determine: (i) if gains in volume for improved versus unimproved progenies at age 30 years correspond to differences in height at age 10 (I); (ii) if improved progenies differed from unimproved with respect to stem shape (I); and (iii) if genetic differences in stem shape are important to consider when breeding for improved stem volume (II).

- How is additive and non-additive genetic variation in growth expressed over time and across environments? It was of special importance to determine: (i) if selection based on genetic differences in early height could predict tree productivity at maturity (III); (ii) if selection based on genetic differences in growth expressed at single sites could predict tree productivity across environments (II and III); and (iii) if environmental patterns in genetic parameters, of importance for planning breeding programs, could be identified (VI).

- How is clonal variation in fertility expressed through the life of a seed orchard? It was of special importance to determine: (i) if genetic variation in female fertility could be a selection criterion when selecting clones for a seed orchard (V); (ii) if selection based on early differences in fertility could be used to improve clonal balance in a seed orchard (VI); and (iii) if
single-year observations could predict clonal productivity over the life of the orchard (VI).
3 Genetic review, model considerations and concepts

3.1 Genetic parameters

Phenotypes for most ecologically important traits (e.g., tree height) have a continuous distribution with complex inheritance patterns. Quantitative genetic theory rests on the assumption of the infinitesimal model (Fisher, 1930), where the genetic control of a trait is by many additively inherited genes of small effect, in linkage equilibrium and summing up to the breeding value \( A \). Thus, the standard model used to describe phenotype \( P \) in quantitative genetics is

\[
P = G + E
\]

where \( G \) is the genotypic value (assuming that \( G = A \)), \( E \) is the non-genetic or environmental value, and the variance of \( P \) is \( \sigma_P^2 = \sigma_G^2 + \sigma_E^2 \). Although this is a simple model, it has been remarkably useful for understanding genetic data and for predicting genetic responses in breeding programmes (Thompson et al., 2005; Hill, 2010), under the assumptions that: (i) genetic and environmental factors contribute independently to phenotype; (ii) no interaction between the genetic and environmental effects; (iii) no general environmental effects confounded with genotype (e.g., maternal effects); and (iv) the absence of epistatic and dominance gene effects.

The proportion of additive genetic effects in phenotypic expression varies for different traits. As an example, tree height exhibits typical quantitative expression and has been shown to be under moderate additive genetic control in many tree species. Cornelius (1994) investigated published observations of additive genetic control in various tree species for different characteristics (Fig. 2). He found that almost all traits (except specific gravity)
exhibited estimates of narrow-sense heritability \( (h^2) \) less than 0.4 (most often 0.1-0.3) and estimates of the additive coefficient of variation \( (CV_A) \) below 15\% (except for stem volume). This is in general agreement with estimates reported for traits in Scots pine populations (e.g., Haapanen et al., 1997; Hannrup et al., 1998, 2000). Differences among traits shown by Cornelius (1994) are in agreement with the suggestion that traits more closely associated with fitness have low heritability, but high additive genetic coefficient of variation (Mousseau & Roff, 1987; Falconer & Mackay, 1996).

The standard additive genetic model can be extended to involve more complex non-additive genetic effects (Falconer & Mackay, 1996; Lynch & Walsh, 1998), partitioning the phenotypic variance as:

\[
\sigma_p^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2.
\]

where \( \sigma_D^2 \) is the dominance variance due to effect of interaction between genes at homologous loci \((D)\) and \( \sigma_I^2 \) is the epistasis variance due to effect from interactions among genes at different loci \((I)\).
Deployment practices capturing the non-additive genetic components through mass production of full-sib crosses or vegetative propagation can increase genetic gain in the improved material (Yanchuk, 1996; Xiang et al., 2003b; Danusevičius & Lindgren, 2002; Wu & Matheson, 2004). The expression of non-additive genetic variance (dominance and epistasis) for growth in pine field trials has been very variable and often of less importance compared to the additive variance (Foster & Bridgwater, 1986; Dieters et al., 1995; Paul et al., 1997; Isik et al., 2003). The importance of non-additive variance also changes over time, depending on the trait. Decreasing importance of dominance with age has been observed for tree height in longleaf pine (Pinus palustris Mill.) (Stine et al., 2001), and in loblolly pine (Pinus taeda L.) (Balocchi et al., 1993; Xiang et al., 2003a). In addition, diameter is more influenced by environmental conditions, such as competition. In an across-sites analysis of diameter in radiata pine (Pinus radiata D. Don), the importance of non-additive variance varied among sites, from non-significant to non-additive variance accounting for all genetic variation among full-sib families (Wu & Matheson, 2004).

Improvement of vegetative propagation techniques in conifers (Park et al., 1998) offers gain potential from capture of dominance effects, by way of clonally replicated testing and selection for clonal deployment. Common group effects (“C-effects”, in sensu Lerner, 1958), including effects carried over from the mother (maternal effects), could bias the estimate of additive variance and can introduce effects interpreted as dominance effects (Libby & Jund, 1962; Mullin & Park, 1992).

Field testing with clonal replication by cuttings has been suggested to capture gains from the additive effect better than trials with seedlings. The improved accuracy of estimates compensates for the fact that additive effects are confounded with a portion of the non-additive effects, which is usually rather small. Attractive as clonal replication might be, the difficulties associated with vegetative propagation in Scots pine today dictate that most testing be done with zygotic seedlings from different mating designs for full-sib and half-sib families.

### 3.2 Multi-trait selection

Selection acts on specific characters, but these may be genetically correlated with other traits. While selecting on many traits simultaneously (multi-trait selection) can increase the total economic value of the gain, it also reduces the gain for each character. Combining traits in a selection index always gives more profit than sequential selection on different traits (Searle, 1965).
Genetic correlations measure the strength of the heritable association between traits and arise from two phenomena: (i) genes having more than one phenotypic effect (pleiotropy); and (ii) genes for different traits that are in close proximity on a chromosome and thus more likely to be inherited together (linkage) (Falconer & Mackay, 1996). Genetic correlations can be used to predict favourable or unfavourable correlated responses, and consequently help to decide which traits to include in the breeding goal and whether indirect selection can be applied. In general, phenotypic response to selection is a function of selection intensity and its efficiency or accuracy (e.g., $h^2$) (Lynch and Walsh, 1998). Multi-trait selection, in matrix notation, is defined as,

$$\Delta z = G P^{-1}s$$

where $\Delta z$ is the vector of change in response caused by one generation of selection, $s$ is the vector of selection differentials, and $G$ and $P$ are the additive genetic and phenotypic variance and covariance (further on expressed as (co)variance) matrices, respectively. Multi-trait models can be used to predict $G$, which involves genetic linkages between individuals and genetic correlations between traits (Mrode and Thompson, 2005). In an evolutionary context, differences in $G$ between generations and in different environments gives general ideas about the “evolvability” of each trait (Lande, 1979; Houle 1992; Kirkpatrick, 2009).

In applying practical selection, the tree breeder needs to specify feasible selection criteria, depending on time and costs, which are correlated with the breeding objective traits. Multi-trait selection thus maximises an aggregate breeding objective ($I$) for highest profit, defined as

$$I = \hat{g}^T v$$

where \( \hat{g} \) is a vector of predicted breeding values for the objective traits, and $v$ is a vector of economic weights. It follows that the breeding objective traits can be obtained from (e.g., Schneeberger et al., 1992):

$$\hat{g} = G_{SS}^{-1} G_{SO} \hat{u}$$

where $\hat{u}$ is vector of predicted BLUPs (Best Linear Unbiased Predictor) for the selection criteria, $G_{SO}$ is the genetic (co)variance matrix between the objective traits and selection criteria, and $G_{SS}$ is the genetic (co)variance matrix between the selection criteria. Berlin et al. (2010) identified several knowledge gaps where research is needed for efficient multiple-trait Scots pine breeding. The most important of these were: (i) relationships between objective trait changes and aspects of the production system objective; (ii) a
flexible growth-and-yield simulator to better model growth and other traits; and (iii) the genetic parameters necessary for performing selection for the suggested breeding objectives ($G_{SO}, G_{SS}$).

### 3.3 Age-specific genetic expression

Efficiency in tree breeding is dependent on assessing individuals at comparatively young ages to make predictions of performance at full rotation and final harvest. A complement to field-testing is early testing. The definition has not been used consistently, but is understood as an early evaluation of trees in the nursery or other artificial environment (Lambeth et al., 1983; Lundkvist, 2000). Unfortunately, early testing has not proven to be very effective (e.g., Lundkvist, 2000). Genetic gain in objective traits by selecting on traits measured at early ages (selection criteria) has three major restrictions: (i) heritability of both traits; (ii) genetic associations between traits (age-age correlation); and (iii) genotype×environment interaction (G×E).

Trends in genetic parameters are difficult to detect in many situations, owing to deficiency of data and large variability among genetic parameter estimates at any given age (e.g., Cornelius, 1994; in Scots pine, e.g., Haapanen, 2001; Jansson et al., 2003). At an early phase of stand development, heritability and additive genetic variance in pines seem to increase with age (e.g., Foster, 1986; Balocchi et al., 1993; Apiolaza et al., 2000; Xiang et al., 2003b; Jansson et al., 2003). Inter-tree competition is proposed to be an influential factor for genetic expression in field tests (Franklin, 1979; Foster, 1986; Hodge and White, 1992; Balocchi et al., 1993; Lambeth & Dill, 2001; Sierro-Lucero et al., 2002; Gwaze & Bridgwater, 2002), but general trends for development of heritability and genetic variance over time are not consistent in the literature (e.g., Cotterill & Dean, 1987; Smith et al., 1993).

A common finding is that measurements of a trait at ages closer to each other, within a specific test site, show stronger correlation than do those further apart (e.g., Lambeth, 1980; Foster, 1986; McKeand, 1988; Hodge & White, 1992; Matheson et al., 1994; Xie & Ying, 1995; Greaves et al., 1997; Lambeth & Dill, 2001; Stine et al., 2001; Haapanen, 2001; Jansson et al., 2003; Li & Wu, 2005). Assessments at older ages give stronger correlation (Johnson et al., 1997), which is reasonable if a cumulative trait is considered. Aside from genetics (e.g., heritability, ontogeny, pleiotropy, site-specific genetic expression), other circumstances also affect the correlation between measurements at different ages, such as mislabelling in the field or other

3.4 Site-specific genetic expression

An assumption of the additive genetic model is violated if there is site-specific genetic expression leading to G×E. This site-specific genetic expression results mainly from differences in performance across environments (heterogeneity of genetic variances) or from rank changes between environments (White et al., 2007). More specifically, G×E could result from a changing correlation pattern between traits across environments. From the point-of-view of a breeder making selections, a rank change across sites is the most important impact. G×E should be incorporated in the genetic model if data arise from multiple sites.

Similar site conditions give rise to high genetic inter-site correlations representing low G×E (e.g., Xiang et al., 2003a). Strong G×E is to be expected when sites are separated by large geographic or ecological distance, and when material is transferred to unsuitable conditions. Low to moderate inter-site correlations, high G×E, for growth have been reported in *Pinus* spp. (e.g., Hodge & White, 1992; McKeand et al., 2006, Gapare et al., 2010), while others have reported low G×E (e.g., Hannrup et al., 2008; McKeand et al., 2008). G×E has often shown to cause effects that are not repeatable, probably arising from accidental effects.

Since Scots pine is planted over a wide area with large variation in site conditions, G×E is an important effect to consider for genetic evaluation (e.g., Persson, 2006). Rank change is typically evaluated by genetic correlations between environments in a multi-trait model approach (e.g. Costa e Silva et al., 2005). The additive genetic correlation is then estimated directly by considering the trait expression on two different sites as two different traits. This approach is computationally demanding as the number of traits increases. A less demanding approach estimates an additive genetic variance and a variance for specific genetic entries across sites, e.g. using a random site×parent effect. This model approach is more general, but has limitations, e.g., genetic expression for specific sites cannot be investigated more thoroughly.

Type-B genetic correlation is a well-accepted measurement of G×E for the genetic expression of a trait in different environments (Yamada, 1962; Burdon, 1977). For determining the degree of bias in the heritability estimate due to G×E, the average Type-B genetic correlation of additive genetic effects across several sites can be calculated as:
where $r_B$ is the additive genetic correlation, $\sigma^2_A$ is the additive genetic variance and $\sigma^2_{AE}$ is the interaction variance for the same trait expressed on different sites. A large Type-B correlation means that there is a little genotype by environment interaction, such as demonstrated by Xiang et al. (2003a) in field tests with loblolly pine. On the other hand, Type-B correlations in slash pine (*Pinus elliottii* Engelm.) are reported to be low, which suggests that large site-index differences between progeny test sites and commercial production land will decrease the reliability of breeding values (Hodge and White, 1992).
4 Materials and methods

4.1 Overview of populations and field material

The materials reported in this thesis originate from different sets of field-test assessments made for inference about genetic variation (additive and non-additive) in Scots pine populations. All material originates from established experiments where crossing and field design was planned beforehand. This thesis involves both recent and historic assessments in the trials, since time trends were to be studied. The material represents genetic entries in the Swedish breeding stock of Scots pine propagated for: (i) progeny testing (I-IV), established during the 1970-90s in northern and southern Sweden, and covering a wide geographic and climatic gradient over areas used for commercial forestry in Sweden; and (ii) a now-mature clonal seed orchard experiment (V and VI). The trials are maintained by the Swedish Forestry Research Institute (Skogforsk), and have reached an age where many growth and wood-quality traits can be assessed.

As the progeny trials are designed mainly for early evaluation with single-tree plots, their initial planting density may contribute to interference for resources among trees (e.g., competition for light); this made it appropriate to conduct analyses with models accounting for competition (I and II). While the main focus is on performance in northern Sweden (I-III, VI), general inferences of genetic parameters in Swedish Scots pine are made in IV and V, as these also included southern material.

The material used for estimation of genetic differences and genetic (co)variation of growth comprise of: (i) improved trees, represented by progenies (controlled crosses) from the first round of phenotypically selected plus-trees (selected mainly for superior height growth), and unimproved trees originating from unselected natural stands (I-III); and (ii) data from all
available reports on breeding values estimated from field tests of Scots pine between 1992 and 2006 (IV). Data for female fertility expressed by cone production (V and VI) originate from experiments and operational clonal seed orchards. Both tree growth and fertility were considered to be complex composite traits with polygenic inheritance.

4.2 Multiple trait genetic analysis

Estimates of genetic (co)variances (G) and breeding values (u) are required for selection of parents for further mating and deployment. Patterns of genetic expression in field trials can be obtained from proper genetic (constructing genetic relationship between groups of individuals) and environmental designs (explored in Scots pine by Ericsson, 1997, 1999b; Jansson, 1998; Haapanen, 2001), and by large experiments with many families (Hodge & White, 1992). By contrast, “nuisance” factors in field trials disturb the genetic inference due to: (i) spatial dependence among trees; (ii) G×E; and (iii) selection arising from field mortality.

The statistical analysis of genetic data has developed from analysis of variance methods (ANOVA) (Thompson et al., 2005). Methods, such as the multivariate maximum likelihood estimation procedure, can utilise information from all traits contained in the evaluation, including correlations between traits and genetic relationships between individuals. If the population is under selection pressure and the data has non-random missing records, genetic (co)variance estimates from a univariate analysis are likely to be biased (e.g., Meyer and Thompson, 1984; Meyer, 1991). Applying multivariate statistical methods with all information describing the selection may account for such selective culling and thus yield more accurate estimates (e.g., Schaeffer et al., 1998; Persson & Andersson, 2004).

The mixed-model equation (MME) framework has been developed to statistically interconnect the phenotypic observations from the biological model \( P=G+E \) into casual inference of genetic (co)variances and patterns of selection of individuals. MME is designed for unbiased estimation of variance components and prediction of random effects (BLUP) (Henderson, 1984). For individual tree-based multitrait analysis, a linear mixed-model for \( q \) traits (which determines the dimension of \( G_0 \) and \( R_0 \) below) is generally written as

\[
y = Xb + Zu + e,
\]

with
\[
\text{var}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R} = \mathbf{V}, \quad \text{var}(\mathbf{u}) = \mathbf{G} = \mathbf{G}_0 \otimes \mathbf{A}, \quad \text{var}(\mathbf{e}) = \mathbf{R} = \mathbf{R}_0 \otimes \mathbf{I}
\]

where \( \mathbf{y} \) is the vector of observations, \( \mathbf{X} \) and \( \mathbf{Z} \) are design matrices relating observations to the fixed and random effects in vectors \( \mathbf{b} \) and \( \mathbf{u} \), and \( \mathbf{V}, \mathbf{G} \) and \( \mathbf{R} \) are the (co)variance matrices with corresponding sub-matrices \( \mathbf{G}_0 \) and \( \mathbf{R}_0 \), of \( \mathbf{y}, \mathbf{u} \) and \( \mathbf{e} \), respectively, \( \mathbf{A} \) is the relationship matrix associated with the studied individuals, \( \mathbf{I} \) is the identity matrix, and \( \otimes \) is the Kronecker product. The corresponding MME are

\[
\begin{pmatrix}
\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\
\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}
\end{pmatrix}
\begin{pmatrix}
\hat{\mathbf{b}} \\
\hat{\mathbf{u}}
\end{pmatrix}
= 
\begin{pmatrix}
\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\
\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}
\end{pmatrix}.
\]

When population (co)variances are not known, the generally accepted strategy is to estimate necessary (co)variances using REML and to use these estimates for BLUP of breeding values (\( \hat{\mathbf{u}} \)) (e.g., Thompson et al., 2005). REML is a modified maximum likelihood approach that takes into account the loss of degrees of freedom resulting from estimating fixed effects (cf. McCulloch et al., 2008). Also, a Bayesian approach has shown to be useful in forest genetic analysis (Cappa & Cantet, 2006, 2007), though differences in variance and parameter estimates in a practical situation may be small compared to conventional methods (Waldmann & Ericsson, 2006).

In the framework of MME a number of statistical methods, such as spatial analysis, have developed and shown to account effectively for “nuisance factors” (e.g., Costa e Silva et al., 2001; Dutkowski et al., 2002, 2006; Zas 2006; Cappa & Cantet, 2007; Ye & Jayawickrama, 2008; Ding, et al., 2008). A methodological investigation of breeding value prediction for Scots pine in northern Sweden showed that spatial analysis improved prediction accuracy by reducing the error variance by 10% for health, 15% for height and 5% for diameter (Dutkowski et al., 2007).

The software package Asreml (Gilmour et al., 2006) estimates variance components in a general linear MME framework using REML with an average-information algorithm (Gilmour et al., 1995). It was the main statistical tool employed in this research for estimation of the (co)variance component in spatial and multi-environment multivariate analyses (II and III), and repeated univariate analyses of repeated measures (VI) and at varying sites (IV and V). In some of the work, the SAS statistical software package (SAS Institute Inc. 1999) was used for statistical analysis (I) and for generating competition indices (I and II).
4.3 Growth in progeny tests (I-IV)

The first study (I) compared patterns of population differentiation observed among progenies from controlled crosses of selected material (improved) to those observed among progenies raised from unselected seeds collected in natural stands (unimproved). This study inferred differences among genetic entries in volume (and area-based production) in comparison with measurements of tree height assessed at an earlier age in the same trials (at age 10) (reported by Andersson et al., 2003). The study material comprised a balanced sample of improved and unimproved Scots pine in 36 north Swedish field tests, covering broad geographic and climatic gradients (latitude 62.3–67.8°N; temperature sum 496–1056 degree days). Analysis of tree growth, survival and damage at a second measurement of the trials was done at age 19–33 years (except for some trials with additional measurements in between). Stem shape (height-diameter relationship) is an important characteristic for stem volume and the trials studied were old enough to allow diameter to be included in the assessment.

Material comprising three sets of Scots pine progeny (seed orchard test series) in ten field trials (a subset of the 36 analysed in I) was used to further evaluate genetic variation in growth, including inferences about the pattern of correlations between growth traits (height, diameter and volume) at the late measurements (about 30 years of age) (II), and for evaluating the genetic correlations in growth expression between early and later measurements (III). This material allowed a complete analysis based on all trees in the trials, since they were completely remeasured at age 26–30 years, and had been measured at least one additional time at about 10 years. The trials were originally established for the purpose of progeny testing of plus-trees selected in natural stands (latitude 61.8–65.7°N; altitude 110–570 m.a.s.l.) and mated in accordance with Kempthorne and Curnow’s (1961) circulant design (Fig. 3) in three commercial first-generation seed orchards.

In order to reduce spatial dependence among trees, distance-dependent competition and height indices based on neighbouring trees were used to adjust for competition and site variability among single-tree plots. This modelled the interference between adjacent units in a nearest-neighbour approach, with distance-dependent indices used as covariates (spatial covariate indices), controlling both competition and micro-site effects (Draper & Guttman, 1980; Besag & Kempton, 1986; Magnussen, 1994).
Figure 3. An example of the partial-diallel “circular” mating scheme; shown here for trial 263 (x=successful and f=failed crosses).

Each orchard test series consisted of field trials on three or four sites (representing unique sets of plus-tree progenies). Originally, trees were planted in a fully randomised experimental design as single-tree plots (eight trials) or in randomised blocks with four-tree square plots (two trials) at 2 or 2.2 m square spacing. About 40 trees per family were planted at each site, with little overlap of parents in the mating schemes (Table 1).
A second study (II) more closely examined the genetic (co)variance of growth traits in northern Scots pine, with special emphasis given to the relationship between height and diameter for the late assessment at about 30 year of age. As inference about genetic variances and relationships were targeted at stability across test series, an interaction model for capturing G×E was used, which allowed for estimation of the site-specific additive genetic variation. Statistical consideration regarding spatial dependence differed from that employed in I; the covariate-based approaches for modelling spatial dependence for competition were further developed in II. This included: (i) further partitioning of the phenotypic interference into genetic and non-genetic competition components of a given focal tree in relation to its neighbours; and (ii) an iterative approach that used family and residual effects from the previous iteration to redefine the nearest-neighbour covariates, in order to satisfy the assumptions of linear mixed models for which unbiased parameter estimation via standard REML can be applied.

A third study (III) examined whether the pattern of genetic (co)variances for growth documented in the earlier studies at later assessments (at about age 30 years) was the same as younger ages (about age 10 years). This study focused on correlations between sites (amount of G×E) and age-age correlations (obtained in the G-matrices) with respect to tree height at about age 10 and 30 for three populations of northern Scots pine. Tree vitality was included to account for non-random selection due to mortality in the trials. The analyses enabled a full (co)variance structure where assessments of all traits within sites were considered as unique traits, thus enabling all combinations of correlations between traits and sites at the two ages (10 and 30) to be estimated. Spatial dependence from competition was neglected, since only tree height was studied. The statistical procedure adjusted for
fertility trends, by first obtaining the correlated residual structure in a
univariate analysis with an AR1 (first-order autoregressive) variance model at
the residual level, and secondly subtracting this structure from the
phenotypic trait values. Adjusted values were used in the final multi-trait
mixed-model analysis, where the amount of G×E and age-age relationships
were expressed as additive genetic correlations between tree heights at
different sites, for all possible combinations of trials within a population.

Parameter estimates for growth in field tests over the 15-year period
1992-2006 were compiled to evaluate relationships between parameter
values and test environments in study IV. The aim was to utilise available
information to elucidate regional differences, but also to determine any
other pattern in the parameters that might indicate novel conditions for tree
breeders to take into account. All data originated from seedling progenies of
the first round of plus-trees, selected in natural forests over Sweden (Fig. 4).

![Figure 4](image_url)

*Figure 4. The distribution of genetic tests included in IV for evaluation of genetic parameters collated from available reports on breeding values estimated from field tests of Scots pine (filled circles) and Norway spruce (Picea abies (L.) Karst.). The field trials are aimed at testing of plus-trees in seedling progeny tests and clonal tests.*
A multiple regression analysis was used in IV, where estimates of $h^2$ and CV$_A$ were used as dependent variables in the trend analysis. Preliminary analysis identified the following variables as important: site (latitude of the test site, °N); orig (mean latitude of origin for parents of the tested trees, °N); age (age of the plantation, years in test); grows (number of measured growing seasons/internodes, frequently equal to tree age); and hgrow (mean recorded height growth, m, frequently equal to tree height). The linear relationships between the investigated variables indicate that other combinations of variables could also explain trends in the genetic parameters (Table 2). In addition, information on survival (surv) was frequently missing and had to be discarded. Only a few CV$_A$ estimates were missing and therefore the 201 otherwise complete records were used for further investigation of the single-site data. Information on latitude of origin in the multiple-site data was limited and allowed for 62 complete records to be used.

Table 2. Correlations in the progeny test data where non-zero numbers are significant estimates (p < 0.05). Single + or – denote possible positive or negative tendency although non-significant (zeros are unmistakably non-significant; na = not applicable). ‘site’ (latitude of the test site, °N), ‘orig’ (mean latitude of origin for parents of the tested trees, °N), ‘age’ (age of the plantation, years in test), ‘grows’ (number of measured growing seasons) ‘hgrow’ (mean recorded height growth, m), ‘surv’ (mean survival), ‘transf = orig – site’.

<table>
<thead>
<tr>
<th>Item</th>
<th>orig</th>
<th>transf</th>
<th>age</th>
<th>grows</th>
<th>hgrow</th>
<th>surv</th>
<th>$h^2$</th>
<th>CV$_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-site data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>site</td>
<td>0.93</td>
<td>–0.21</td>
<td>0.34</td>
<td>0.38</td>
<td>0</td>
<td>–0.63</td>
<td>–0.27</td>
<td>–</td>
</tr>
<tr>
<td>orig</td>
<td>–0.56</td>
<td>0.19</td>
<td>0.26</td>
<td>0</td>
<td>–0.43</td>
<td>–0.19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>transf</td>
<td>0.26</td>
<td>0.18</td>
<td>0.31</td>
<td>0</td>
<td>–</td>
<td>–0.26</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>age</td>
<td>0.79</td>
<td>0.80</td>
<td>–0.49</td>
<td>0</td>
<td>–0.19</td>
<td>0</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>grows</td>
<td>0.86</td>
<td>–0.44</td>
<td>0</td>
<td>0</td>
<td>–</td>
<td>–0.26</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>hgrow</td>
<td>–</td>
<td>0.13</td>
<td>–</td>
<td>0</td>
<td>–0.33</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>surv</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$h^2$</td>
<td>0</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
</tr>
</tbody>
</table>

Multiple-site data

<table>
<thead>
<tr>
<th>Item</th>
<th>orig</th>
<th>transf</th>
<th>age</th>
<th>grows</th>
<th>hgrow</th>
<th>surv</th>
<th>$h^2$</th>
<th>CV$_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>site</td>
<td>0.97</td>
<td>–0.35</td>
<td>0.23</td>
<td>0.41</td>
<td>–</td>
<td>na</td>
<td>–0.45</td>
<td>0</td>
</tr>
<tr>
<td>orig</td>
<td>–0.55</td>
<td>0</td>
<td>0.33</td>
<td>–0.27</td>
<td>na</td>
<td>–0.41</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>transf</td>
<td>0.28</td>
<td>+</td>
<td>0.40</td>
<td>na</td>
<td>0</td>
<td>–</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>age</td>
<td>0.75</td>
<td>0.74</td>
<td>na</td>
<td>0</td>
<td>na</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>grows</td>
<td>0.72</td>
<td>na</td>
<td>0</td>
<td>0</td>
<td>na</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hgrow</td>
<td>na</td>
<td>0.37</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$h^2$</td>
<td>0</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0.58</td>
</tr>
</tbody>
</table>
4.4 Fertility in clonal seed orchards (V-VI)

Female fertility in mature Scots pine seed orchards was evaluated in an initial study (IV), using data compiled from new assessments and from previously published sources (mainly on cone production). Annual variation in seed production among trees in natural forest stands and clones in seed orchards is known to be large (e.g., Schmidtling, 1983; Kang et al., 2001). By comparison, variation is likely to be smaller in seed orchards due to a more uniform environment (Kang et al., 2001).

The compiled sources of data were used to estimate the variance components and broad-sense heritability in fertility data. The sources of data varied in their definition and assessment of fertility, so the data were standardised to facilitate comparison among orchards and assessments. The analyses apply more correctly to observations of relative fertility rather than fertility, and the results were expressed as a percentage of average fertility. Additionally, a more thorough study (V) of the development of the genetic variation in cone production over time was done in the Sävar experimental seed orchard (also included in IV, where data were provided from age 30).

The experimental seed orchard was established 1969-77 at the Skogforsk research station at Sävar (latitude 63°54’N; longitude 20°33’E; altitude 10 m.a.s.l.) and comprising about 4 hectares. It holds different treatments for spacing, pruning and fertilisation, arranged in 16 large blocks (Fig. 5). The genetic correlation pattern over time was evaluated using a bivariate approach for paired estimates between years, and by using all data simultaneously in a full-fit model with autoregressive genetic structure.
<table>
<thead>
<tr>
<th></th>
<th>Grafts 625 st/ha (thinned)</th>
<th>Fertilizer</th>
<th>Discing</th>
<th>Pruning, yearly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>331 st/ha</td>
<td>Fertilizer</td>
<td>Discing</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>2.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Discing</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>3.</td>
<td>625 st/ha</td>
<td>No Fertilizer</td>
<td>Discing</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>4.</td>
<td>156 st/ha</td>
<td>No Fertilizer</td>
<td>Discing</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>5.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>Pruning, 3rd year</td>
</tr>
<tr>
<td>6.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>No pruning</td>
</tr>
<tr>
<td>7.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>Pruning, 3rd year</td>
</tr>
<tr>
<td>8.</td>
<td>625 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>9.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>10.</td>
<td>156 st/ha</td>
<td>Fertilizer</td>
<td>Discing</td>
<td>No pruning</td>
</tr>
<tr>
<td>11.</td>
<td>625 st/ha</td>
<td>Fertilizer</td>
<td>Discing</td>
<td>Pruning, yearly</td>
</tr>
<tr>
<td>12.</td>
<td>331 st/ha</td>
<td>Fertilizer</td>
<td>Discing</td>
<td>No pruning</td>
</tr>
<tr>
<td>13.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>Discing</td>
<td>Pruning, 3rd year</td>
</tr>
<tr>
<td>14.</td>
<td>331 st/ha</td>
<td>Fertilizer</td>
<td>Discing</td>
<td>Pruning, 3rd year</td>
</tr>
<tr>
<td>15.</td>
<td>156 st/ha</td>
<td>No Fertilizer</td>
<td>Mowing grass</td>
<td>No pruning</td>
</tr>
<tr>
<td>16.</td>
<td>331 st/ha</td>
<td>No Fertilizer</td>
<td>No treatment</td>
<td>No pruning</td>
</tr>
</tbody>
</table>

Figure 5. The design of the experimental seed orchard at Skogforsk, Sävar. The original planting.

Research station at Sävar The breeding archive road
5 Main results and discussion

5.1 Performance and genetic variation in improved trees

Growth of improved trees (from phenotypically selected parents) was superior to that of unimproved trees 19-33 years of age (I). The result supports earlier findings for the same set of trials at around age 10 years by Andersson et al. (2003). The relative differences in growth between improved and unimproved progenies at ages 10 and 30 seem to develop in accordance with standard site-index and yield functions, and were stable across sites. The covariate-based approach used for competition and fertility adjustment improved the estimates. Furthermore, when grown at a site south of their geographic origin, the overall pattern was the same (i.e. northern sources grow less, but have higher survival), although survival rates increased more and height growth decreased less in improved trees than in unimproved trees. The improved trees were also more slender, which motivated further investigation of the genetic relationship between height and diameter at about age 30, especially as this relationship might affect volume at full rotation and is related to stem quality (e.g., wood stiffness and density).

Selected trees showed a considerable amount of heritable variation for growth and slenderness at about age 30 (II). Genetic correlations between height and diameter across populations were in the range 0.35-0.68, and those in all but one population were significant ($p < 0.05$). Gain for volume from indirect selection on tree height was less than selecting directly on volume. Thus, not considering diameter as a selection criterion would decrease possible gain in tree volume, and it would therefore be desirable to include both height and diameter in selection indices to obtain maximum gain.
Taller trees are more difficult and costly to measure than smaller ones, whereas ease of diameter measurement is essentially independent of tree height. It remains, however, that height is more independent of competition and has higher heritability than diameter. The genetic correlation between height and diameter could help to decide if measurement of only diameter is an alternative to use in older field tests.

Available data of additive genetic height–diameter genetic correlation compiled from progeny tests in Scots pine originating from the investigation in IV (but unpublished) showed a scattered pattern with an average value of 0.58 (Fig. 6). Corresponding estimates compiled for Norway spruce (Picea abies (L.) Karst.) showed an average value of 0.87 (all were single-site estimates).

Figure 6. Estimates of additive genetic correlations ($r_A$; mean = 0.58) between height and diameter measurements from study IV. The smallest circles denote single-site estimates while the others show pooled estimates; circle size corresponds to the number of sites: 1, 2, 3, 4 or 5 with 0.3–2.8° span in latitude. Years of growth, represented time span 9–30. Mean number of trees per estimate 10,000, 3 estimates are open-pollination or common testers with mean 40 families/estimate, 8 estimates partial diallel crosses with mean 86 families/estimate.
5.2 Additive and non-additive genetic variation in growth

Study IV considered a large number of observations across Swedish trials, and thus gives a more general view of genetic parameters than estimates for a specific population at a given point in time. The investigations showed large variation in parameter estimates from different field experiments, highlighting the importance of field testing over multiple sites. Average estimates for \( h^2 \) and CV\(_{A}\) were 0.23 and 0.08, respectively, from single sites, and 0.22 and 0.09, respectively, over multiple sites. G\( \times \)E estimates for both single and multiple sites were low. Both estimates of \( h^2 \) and CV\(_A\) varied considerably among ages and over a latitudinal cline, which also determined the major large-scale trends found in the genetic parameters. Estimates of \( h^2 \) increased with age and southward transfer, whereas estimates of CV\(_A\) increased with southward transfer and more southerly test latitude (IV). As the materials analysed were progenies from the first round of phenotypic plus-tree selection, the genetic parameters should correspond fairly well to those in natural populations. Experimental settings, however, usually have lower environmental variation than forests or commercial plantations.

The additive genetic effect was much larger in II than the dominance or specific combining ability variance associated with full-sib families. The average dominance ratio across populations for height at maturity was 0.03 (II), considerably lower than estimates of \( h^2 \). The importance of non-additive variance was also dependent on the trait. Diameter was more variable and in general expressed less non-additive genetic variation (II).

The importance of additive to non-additive genetic effects increased with time with large variation among specific test sites (III), in agreement with published experience (e.g., Jansson et al., 2003). There was a doubling of the additive to non-additive genetic effect from age 10 to 30. This doubling also refers to the number of undefined estimates due to zero estimates of non-additive variance. Estimates of this ratio suffered from relatively low precision, as the standard error was moderate to high. One possible explanation for this finding is that the dominance effects are not true genetic effects, but rather “C-effects” arising from grouping of seedlings in the nursery and vanishing with time. Jansson et al. (2003) identified other practical factors that could also lead to overestimation of dominance: mistakes in labelling, or inadvertent mixture of pollen from different sources during mating.
5.3 Age trends in genetic expression of growth

The results showed, as already observed by many others (Haapanen, 2001; Jansson et al., 2003; Persson et al., 2010), that estimates of $h_i^2$ for tree height clearly increased with age (III).

The large-scale trend of genetic parameters across-sites showed that older trees expressed larger estimates of $h_i^2$ and $CV_A$, but the general pattern studied in IV did not show any direct influence of tree size on estimates of $h_i^2$ (except from likely confounding effects with age). On the contrary, the regression equation in IV revealed a behaviour of $CV_A$ in tending to drop as the trees grew taller, which reflects that faster-growing tests exhibited lower estimates of $CV_A$ for height. This might be caused by damaging factors correlated with genetic expression, which led to greater genetic variation for slower-growing (maladapted) material.

Estimates of $CV_A$ on the other hand, tended to decrease with height growth and age in III, but no clear trend was found. Decreasing estimates of $CV_A$ for tree height with age has been seen by others (e.g., Burdon et al., 1992; Haapanen, 2001; Jansson et al., 2003), but this pattern is not consistent in the literature. Persson et al. (2010), for example, found no clear pattern of estimates of $CV_A$ decreasing with increasing height in three series of progeny trials with Scots pine in northern Sweden. However, the phenomenon with decreasing estimates of $CV_A$ might also be a matter of scale. In the very early stages growth can conform, in part, to an exponential pattern, such that small variations in an exponential parameter can generate much larger coefficient of variations (CVs). Later on, as trees get taller (and older) height growth corresponds closer to a linear pattern, such that similar (or even slightly larger) relative variation in the linear growth parameter can be accompanied by declining observed CV.

Scots pine growth is better expressed in older trials around 30 years of age. Genetic correlations between early and late measurements of tree height varied greatly for some sites in III. Furthermore, genetic correlations between sites for tree height revealed greater $G\times E$ at age 10 (III) than at age 30 (II and III). The pattern of genetic correlations between tree height and tree vitality (survival) varied over time. Tree vitality was primarily used to account for selection due to mortality, but did help interpretation of the underlying causes for the different inter-site correlations obtained at ages 10 and 30.

While longer testing times will favour higher correlations and thus greater response to selection, maximising genetic gain per unit time in the breeding programme is greatly influenced by the rate of generation turnover. Delays in establishing new orchards will result on a lower rate of
increase in realised gains. Furthermore, gain potential is increased each time the breeding population is mated.

5.4 Environmental trends in genetic expression of growth

In general, genetic parameters vary across environments. This seems to be especially important for the genetic expression of growth in northern Sweden that is much dependent on cold adaptation, expressed by survival. The large-scale trend was that both estimates of $h^2$ and $CV_A$ increased from north to south (IV) (Fig. 7), although not as much for estimates of $CV_a$ as for $h^2$. The finding in IV agrees with the general clinal latitudinal pattern already established for Scots pine. Higher estimates of genetic variation obtained in southern Sweden might be attributed to a transfer effect, or the inclusion of foreign provenances and provenance hybrids.

If Scots pine is transferred southward from northern Sweden, it has higher survival and greater genetic variation, indicated here by estimates of $CV_A$ increasing with $orig$ (IV). Additionally, mortality may decrease environmental variation and thus change the contributions of genetic effects to phenotypic expression, as seen by larger estimates of $h^2$ with $orig$ (IV).

The genetic correlation pattern obtained within and between sites for height and survival resulted from the genetic expression of height assessed in harsh environment (with low survival) differing from that assessed in milder environments (III). On harsh sites, tree health is correlated strongly with tree height at early ages, but the correlation becomes less strong during later development. In milder areas, such as southern Sweden, frost hardiness has less impact on growth, and studies have shown low to moderate levels of $G\times E$ for early measurement of height (age 5-17) (e.g., Hannrup et al., 2008).

The results presented in III support the findings of Persson (2006), who used other north Swedish test material, which showed a contrasting covariance pattern of growth and survival between sites depending on the degree of mortality at the two sites. He suggested that this was an expression of different genetic factors in different environment. It follows that genetic interpretation of phenotypic expression in field tests partly depends on the definition of the trait.

The contrasting covariance patterns for growth and vitality in northern Sweden, leading to high $G\times E$, are important to consider when selection corresponds to deployment over wide areas. Using several sites for testing would be feasible for early evaluation, as a more precise measure of the plasticity for the improved material could then be obtained.
5.5 Clonal variation in female fertility

Variation for female fertility in mature Scots pine seed orchards was rather small and varied as much within clones as among clones (V). In addition, the variation was highly dependent on year of assessment. Harvestable levels of cones should start from about age 10. Once most trees have reached flowering competence, there is little difference among clones in seed set, making it unlikely that a few clones have a large impact on the genetic constitution of the seed crop in the long run (V). At younger ages, when not
all clones have reached reproductive competence, there will be more imbalances in the clones contributing to the crop (Almqvist, 2001). This was shown by the sharp decrease by estimates of the genetic (clonal) coefficient of variation ($CV_G$) for cumulative cone yield ($VI$), which later stabilised as all clones reached reproductive competence. As results from young orchards are more often published than those for mature orchards, there is undoubtedly a trend that seed-set differences are overestimated in the literature. This was an expected outcome as there are two main factors contributing to the variation among clones: (i) variation in the onset of cone production, and (ii) fecundity of mature clones; only the latter cause variation in fertility as the orchard matures.

In contrast to the result for estimates of $CV_G$, the low estimates of broad-sense heritability ($H^2$) obtained at early ages was unexpected, but could be caused by environmental variation among small young grafts, or selected clones being already reproductively competent at time of grafting, so that cone production was more a matter of graft development than onset of competency.

Age-age genetic correlations indicated that early forecasts by clone of cone production at older ages are unreliable ($VI$). Better predictions (moderate correlations) are achieved only if based on rather mature grafts, 14 or more years after establishment; even then, if there is a year with poor cone set, the prediction accuracy will be low (Fig. 8). Efficiency of selection and predictions improved if based on cumulative cone count over many years. Nevertheless, the limited variation in mature female fertility suggests that this should not be a criterion when selecting clones for a seed orchard ($V$).

The phenotypic correlations are probably more indicative of what might be possible if early data are used for roguing the seed orchard. The age-age correlations indicate that thinning of ramets with poor seed set could improve seed yield per hectare.
Figure 8. Pairwise genotypic correlations between single-season cone yield per clone (CY) at ages 11 to 25, and CY at age 30. The dotted line shows a first-order regression trend line ($R^2$=0.63).
6 Conclusions and implications for breeding

Genetic differences among trees are the key to success in tree improvement programmes. Most programmes make selection decisions long before rotation age to maximise the rate of genetic progress in long-term breeding. This thesis examines the genetic expression of field performance in long-term experiments over time and across environments in the Swedish Scots pine breeding stock. Based on evaluation of many field experiments at both early age and near maturity, it provides novel information of genetic differences among Scots pine trees for field performance to facilitate selection decisions for breeding and deployment objectives.

Genetic control of growth in Scots pine is expressed differently across sites and over time, but a sufficient amount of additive genetic variance forms the basis for successful breeding. Within the Swedish breeding stock, there is great potential for genetic adaptability and for further breeding for growth, with mean estimates of $h^2$ and $CV_A$ at the magnitude of 0.2 and 9%, respectively. Genetic differences were better expressed on favourable sites and at older ages.

Progeny of first-generation selected plus-trees for northern Scots pine show faster growth with about 20% gain over unimproved material in unit-area production at 30 years of age, with no differences in survival or quality. The predictions of differences between improved and unimproved material on full-rotation volume of about 10% based on 30 year height measurements is in accordance to predictions made from 10 year height measurements. Genetic correlations indicate a moderate relationship between height and diameter in northern Sweden, and there was also a genetic difference in stem shape. In conclusion, both diameter and height need to be considered to obtain maximum gain in volume production.
The additive effect was the most important source of genetic variation. By comparison, estimates of the dominance part were small, site-specific, and decreased much with age. In conclusion, there may be little benefit in attempting to explore dominance through specific combining ability to improve productivity in northern Sweden.

Genetic differences are better distinguished at age 30 than at age 10, due to heritability increasing with time. Selection at younger ages is most likely to be efficient when differences between trees are not due to damage or vitality. Currently, selection of Scots pine in northern Sweden is primarily based on vitality and height at 10-15 years of age. At this time, the trees have been exposed to environmental conditions responsible for most of the mortality throughout the life of the stand. Using early growth as a predictor of height at later ages is much dependent on site conditions, but is also dependent on non-additive genetic effects. Evaluation of breeding values will be more accurate if testing times are longer, but this obviously needs to be balanced with time and cost and the increase in gain potential that is created when gametes from new selections are recombined through mating.

In general, there is large variation in parameter estimates from different field experiments as the degree of expression of genetic differences in growth varies across environments, highlighting the importance of field testing over multiple sites. This seems to be especially important for growth in northern Sweden that is much dependent on cold adaptation, primarily expressed by survival. The contrasting covariance patterns for growth and vitality in northern Sweden, leading to high G×E, are important to consider when selection is intended for deployment over wide areas.

Strong genetic correlation between sites at around 30 years of age suggests that growth is expressed with less additive G×E in older trials. This means that there is less site-specific additive genetic variation for growth at age 30 compared to age 10, which might be due to that: (i) “C-effects” being less influential and not biasing the additive genetic variance; (ii) field assessments of tree height corresponding better to the underlying biological processes for growth (as environmental factors determining growth is lower); or (iii) a shift in genetic control with age (e.g., ontogeny or pleiotropy). It follows that early selection could be more effective for deployment over larger areas, including harsher sites, than expressed by the varying covariance pattern at age 10 across all trials. Selection for growth on favourable sites, where the degree of expression of genetic variation for growth is high and age-age correlation is stronger, would capture more of the genetic potential for growth expressed across sites at later ages. Additionally, assessment of height on low-survival test sites should be recognised with genetic
correlations between sites or should be avoided in the genetic analysis for predicting general breeding values for growth across sites.

Competition among trees inflates genetic differences and is an important environmental factor to consider for unbiased genetic interpretation of growth in older field tests. When complicated statistical models are used to adjust for spatial dependency in single-tree plot trials, there is a need for validation in comparative realised-gain trials with large multiple-tree plot designs.

The age at which flowering begins in selected trees determines the earliest time for generation turn-over and thus affects genetic progress in long-term breeding. The productive lifespan of a seed orchard depends, in part, on the genetic value of the seed crop, and on the rate of progress offering better trees. The genetic value of a seed crop is dependent on the clone frequency, variation in clonal fertility and differences in breeding values among clones in the orchard.

Clonal variation in female fertility in mature Scots pine seed orchards is rather small and varies as much within clones as among clones, and is heavily dependent on year of assessment. This yearly variation in cone production needs to be accounted for when the value of an early seed crop is determined or for a year with a low seed set. When trees have reached sufficient age of flowering there are no large differences among clones in seed set making it unlikely that a few clones have a large impact on the genetic constitution of the seed crop. This is especially true if harvesting in poor crop years is avoided. Harvestable levels of cones should generally start from about age 10. The prospects for early selection of clones for future seed cone production based on a single observation are low. Reliable predictions can be done on rather mature grafts, 14 or more years after establishment.

In general, the approach of genetic field testing and evaluation in forest tree breeding should correspond to variation in site conditions in an uncertain future. Using several sites for testing would be feasible for genetic evaluation as more precise measure of the plasticity for the improved material.

Finally, this thesis illustrates the importance of examining long-term field tests with multi-trait, multi-site analyses accounting for environmental effects.
References

Abetz, P. (1976) Beiträge zum Baumwachstum - Der h/d-Wert mehr als ein Schlankheitsgrad. [A comment on tree growth - Highlighting the slenderness coefficient]. Der Forst- und Holzwirtschaft, 31, 389–393. [In German]


Erteld, W. (1979) Groesse und Entwicklung des h/d-Wertes in Kiefernbestandes. [Diameter and the development of the h/d-ratio in pine stands]. *Allgemeine Forst und Jagdzeitung*, 150, 72–75. [In German]


Acknowledgement

In finalizing this thesis I am sending my sincere gratitude to all persons, co-authors and founding organizations that have been involved, for making this work possible, as this project was not free of cost, neither from personnel “software”, nor from financial “hardware”.

I greatly appreciated the privilege to join the Research School in Forest Genetics and Breeding “företagsforskningskolan i skogsgenetik och förädlning”, and to be guided by its superb group of advisors. It has been a great journey with the best leadership from its board and from the executive group with Bengt Andersson, Jan-Erik Nilsson, Gunnar Jansson, and “Rishi” Bhalerao. I have enjoyed every part of it and especially the friendship with all the students: Jon, Sara, Maria, Ellinor, David, Henrik, Mats, Lars, Jenny and David. The school was set up as a joint venture between the forest industry (one year of practical training) and the academia, with a five year obligation to fulfil. Due to some practical breeding work in between, I could hang on for a couple of years extra, before I had to finalize the thesis.

The projects within the school were supported by the KK-foundation, Föreningen Skogsträdsväxthållning, Kempefïstelserna, SLU, L-E Thunholm’s foundation, and by the forest industry, through host companies, which in my case were Skogforsk and SSP (Svenska Skogsplanter AB).

I thank my supervisor Bengt Andersson for continuing support and stimulating guidance throughout the project. He initiated the work and helped me pass the line at the goal. I am looking forward to becoming a full time working member of the tree breeding group at Sävar.

I thank my supervisor Dag Lindgren for financial support before I started my graduate studies and for introducing me to his group of seed orchard collaborators. His travelling accompany around the world and endless belief in the importance of forest genetics have provided me with happiness throughout these years.
I am particularly thankful to Tore Ericsson and Torgny Persson who provided me most advanced knowledge in statistics, genetics and scientific writing. Their skilful and patient guidance and practical help have been the most valuable for accomplishing the task for my project. Thanks Tore for challenging me towards more mindful writing.

I am also very thankful to “Tim” Mullin, who worked with me on all the manuscripts and on this thesis. He is a skilful teacher and his feedback and language-correcting on my writing was both motivating and most helpful. The imperfections in the thesis are purely my responsibility.

I sincerely appreciated the stimulating discussions and the fun together with Jon Hallander. Also, thanks to Patrik Waldmann for bringing new statistical methods and genetic concepts into my mind. Stefan Löfmark for computer issues, the staff at “my” department at SLU, and all people at the Skogis library who kindly provided me with the tools I needed for accessing scientific publications.

Finnvid Prescher gave me the kindest guidance to the practical experience at SSP. Except for challenging me with stimulating seed orchard tasks, I am thankful that he allowed me to carry out a molecular lab project for a month and a half to sort out the rate of true hybrids of *L. × eurlepis* in a yearly seed batch. The work was carried out at UPSC and was kindly supervised by Yoshinari Morigushi. I appreciated the time at the lab and would like to thank them all, particularly Sara Abrahamsson, Rosario Garcia Gil and Inga-Britt Carlsson for providing me with all equipment and lab space.

Thanks to Ola Rosvall, for being a most encouraging and supporting person highly devoted to forestry and tree breeding and for sharing your inspiring ideas with us, and for making Skogforsk at Sävar a place where everybody feels welcome. And so do I.

Thanks to Skogforsk, and all its personnel. Jan Fryk, Ola and Bengt for present and future work, Sven Westerling and his staff for the computers they have provided for me, Gunnar Jansson for being my co-author, Johan Westin for discussions, my collaborators at Sävar who have patiently accepted my distracted and divided focus on the work at Sävar, simultaneously being a graduate student and running the lodgepole pine programme during the last years. Especially thanks to Ulfstand, Berit, Tomas, Margareta, Monica, Mikael, Jörgen, Anita, Jan, Erik, Martin, Göran, Mats, Åsa, Marie, folks at Brunsberg field station, either for assistance or for supporting data to the thesis, or for carefully operating the lodgepole pine breeding activities.
I am grateful to my parents and parents in law for all support and help with family issues during these years.

Finally, I am in debt to my family, my devoted wife Lovisa and our sons Adam, Olof and Gustav, for making me laugh and reminding me that they are the most important ones.