Applicability of best linear unbiased prediction (BLUP) for the selection of ortets in *Eucalyptus* hybrid populations

by

Marianne Helena Hettasch

Submitted in partial fulfilment of the requirements for the degree

Master Scientiae (Genetics)

in the Faculty of Natural and Agricultural Science
University of Pretoria

Pretoria
(February 2009)

Supervisor: Prof. C.Z. Roux
Co-supervisor: Dr S.D. Verryn

DECLARATION

I delare that the dissertation which I hereby submit for the degree MSc at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature:.  Date: 8 February 2009

ACKNOWLEGDEMENTS

I wish to express my gratitude to the following people:

Prof. Carl Roux, my supervisor, and Dr Steve Verryn, my co-supervisor, for their guidance throughout the duration of the research project,

Prof. Peter Geerthsen for reviewing the dissertation,

Dr Annabel Fossey for her guidance on the writing of a dissertation,

Dr Dominic Kain for his support and guidance during the conceptualization phase of the research project and for teaching me more about BLUP,

my colleagues Cyndi Snedden, Karen Eatwell and Brian Pierce for their technical and moral support.

The research was financially supported by the Australian Centre for International Agricultural Research (ACIAR), the Innovation Fund of the South African Department of Science and Technology, the THRIP Fund of the Department of Trade and Industry the CSIR Parliamentary Grant Fund.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The prediction of the performance of hybrid plants as production material is complicated by the genetic structure of the hybrid population. Fundamental quantitative genetic concepts are defined with respect to disomic, randomly fertilized populations which are in linkage and Hardy-Weinberg equilibrium. Hybrid populations of crosses are, however, in disequilibrium and do not meet these assumptions of quantitative genetic theory. The theory of inheritance in hybrid populations has not been fully developed and the complex models that have been developed have not reached a stage where they have been implemented to adapt selection practices for hybrid populations. Therefore in practice hybrids are often selected using selection methodology that has been developed for pure species despite large differences in the underlying genetic properties of the populations. It is not known to which extent such selections are able to predict the performance of progeny or clones in hybrid populations. This study was based on historical data and investigated the efficiency of BLUP in hybrid populations. As a case study the selection of ortets in three hybrid *Eucalyptus* populations was investigated. Ortet selection in *E. grandis* × *E. saligna*, *E. grandis* × *E. urophylla* and *E. grandis* × *E. camaldulensis* populations was compared with selection in *E. grandis* populations. Clonal performance was predicted from a series of hybrid family trials whereas realised clonal performance was assessed in clonal trials. The predicted and realised clonal performances were correlated to assess the efficiency of the predictions. A series of constructed indices was created that placed a range of weights on family versus individual tree performances to obtain an indication of the range of correlations that could be expected from the data. Different BLUP methods were compared including selection for breeding and various forward selections for clonal forestry ("clonal forward selection"). The clonal forward selections were based on different estimates of the proportion of total genetic variance that is ascribed to non-additive genetic variance. The results of the study indicated that there were no observable differences in the efficiency of the BLUP predictions in the studied hybrids in comparison to prediction in *E. grandis*. There was, however, a decrease in the correlations between predicted and realised clonal performances with increasing genetic distance between the parents. Furthermore, the genetic values were better predictors of

clonal performance than breeding values and genetic values which were based on higher assumptions of non-additive genetic variances (37% of total genetic variance) were more efficient than those based on assumptions of lower proportions (20%). This study indicates that BLUP methods may be efficient in predicting the clonal performance in the three hybrid populations that were investigated.

NOMENCLATURE AND ABBREVIATIONS

ANOVA	Analysis of variance
BLUP	Best linear unbiased prediction
DBH	Diameter at breast height
G x C	*E. grandis* × *E. camaldulensis*
G x S	*E. grandis* × *E. saligna*
G x U	*E. grandis* × *E. urophylla*
F_1	First filial generation
F_2	Second filial generation
h^2	Narrow-sense heritability
H^2	Broad-sense heritability
HWE	Hardy-Weinberg equilibrium
HWD	Hardy-Weinberg disequilibrium
LE	Linkage equilibrium
LD	Linkage disequilibrium
R	Repeatability
RCB	Randomised complete block design
REML	Restricted Maximum Likelihood
s.e.	Standard error
σ^2	Variance
σ_g^2	Genetic variance
σ_a^2	Additive genetic variance
σ_d^2	Dominance genetic variance
σ_{na}^2	Non-additive genetic variance
σ_f^2	Variance attributed to female parents
σ_m^2	Variance attributed to male parents
σ_{fam}^2	Family variance

DEFINITIONS

Backward selection	Selection of parents based on the performance of their progeny.
Breeding value	The breeding value of an individual is the part of the deviation of an individual's phenotype from the population mean that is attributed to the additive effects of alleles and describes the value of an individual as parent.
Clone	A clone is a group of genetically identical descendants that have been propagated from a single mother plant (the ortet). Individual trees within a clone are referred to as ramets.
Clonal prediction	Clonal prediction is the prediction of an individual's performance as a clone, taking both additive and non-additive genetic components into account.
Clonal trial	A clonal trial is a trial in which the performance of clones is evaluated.
Family	A family is a group of related trees as a result of sexual reproduction.
Family trial	A family trial is a trial comprised of half-sib or full-sib families that allow for the evaluation of the genetic worth of the individuals, families and parents.
Forward selection	Selection of the best individual trees based on their own performance.
Full-sibs	A group of trees that share the same mother and pollen parent. Full-sibs usually result from controlled crosses.
Genetic value	The genetic value of an individual is the part of the deviation of an individual's phenotype from the population mean that is attributed to the additive and non-additive effects of alleles.
Half-sibs	A group of trees that share either the same mother or the same father. Half-sibs usually result from open pollination.
Ortet	An ortet is the tree which is reproduced by means of vegetative propagation to produce one or more genetically identical copies or ramets.
Ramet	A ramet is a tree that has been produced by means of vegetative propagation and is genetically identical to its mother plant or ortet.

CHAPTER 1
INTRODUCTION

For more than a century plant hybrids have played an important role in the improvement of agricultural, horticultural and forestry crops. It has been suggested that the exploitation of hybrid vigour has been one of the major successes in plant breeding programmes of the 20th century (Cooper & Merrill 1999). In the past 30 years researchers have also intensified their investigations into the role of hybridization in speciation and have worked towards a better understanding of the unique properties of hybrids (Orr & Presgraves 2000). The need to understand the factors that promote and influence hybridization and how hybrid populations function and evolve has to some extent been driven by an increased concern about potential genetic pollution and unwanted gene transfer between related species. Population, quantitative and molecular studies have contributed greatly to a better understanding of the nature of hybrids, but there are still many aspects that are not well understood (Guo *et al.* 2006; Baack & Rieseberg 2007; Springer & Stupar 2007). It is not known, for instance, whether selection methodologies, which are based upon assumptions of large randomly mating populations, as is assumed to be the case of pure species, are sufficiently robust to be applied to hybrid populations. There are two approaches to investigate the question of selection in hybrid populations. The theoretical approach involves the development of complex models which describe hybrid populations. This study followed an alternative practical approach in which historical data from hybrid trials were analysed to investigate the efficiency of selection procedures in hybrids. As a case study, the prediction of clonal performance was investigated in three *Eucalyptus* hybrid populations.

1.1 Introduction to plant hybrids

The word ‘hybrid’ originates from the Latin word ‘hybridia’ or ‘ibrida’ and originally referred to the offspring of a tame sow and a wild boar (Warren 1884). In breeding the term hybrid refers to the offspring of genetically dissimilar parents. This definition is

broad and includes crosses between different inbred lines, crosses between different provenances as well as crosses between different species or genera. In addition to first filial generation (F_1) hybrids, there are various advanced generations that are derived from F_1 hybrids: these include F_2, F_3, F_4 hybrids and first and second generation backcrosses, BC_1 and BC_2 (Nikles & Griffin 1992). In addition to these nuclear hybrids the technique of somatic hybridization yields cytoplasmic hybrids in which both nuclear and cytoplasmic material of parents is united. In this dissertation the focus will be on interspecific plant hybrids which have resulted from crosses between distinct species.

There are multiple barriers that may prevent the success of interspecific crosses. In nature mechanical, temporal and ecogeographic barriers may prevent crosses (Rieseberg & Willis 2007). In addition to these barriers there may also be pre-zygotic and post-zygotic barriers that prevent such crosses. One species may lack the genetic information to complete the processes which are required for the pre- and post-pollination processes in the other species (van Tuyl & de Jeu 1997). Pre-zygotic barriers may include the failure of pollen germination, slow growth of the pollen tube, poor penetration of the pollen tube or arresting of pollen tube development (Khurana & Khosla 1998). Even if pollination should succeed, the development of the hybrid embryo may be restricted by post-fertilization barriers leading to abortion of the young embryo (van Tuyl & de Jeu 1997). Mitotic irregularities due to large differences in the genomes, deleterious gene actions as well as cytoplasmic effects may play a role in preventing the development of a hybrid (Potts & Dungey 2004).

If an interspecific cross has, however, succeeded, the highly heterozygous product has the potential to outperform either parental species. Since Darwin has noted the phenomenon of heterosis in 1876, breeders have been interested in its exploitation for the improvement of crops and it has been considered to be a ‘miraculous agricultural phenomenon’ (Lippman & Zamir 2007). In forestry, hybrid superiority is thought to be a function of complementarity of parental traits and adaptability rather than true heterosis (Verryn 2000; Volker 2002). In order to improve their capability to unleash this potential of hybrids, breeders are interested in gaining a better understanding of the nature of hybrids.

1.2 Genetics of plant hybrids

Interspecific hybrids reunite differentiated genetic material of genetically distinct populations (Rieseberg *et al.* 2000). Such populations are distinct because genetic isolation has facilitated independent evolution. Factors such as selection pressure, random drift, mutation, population bottlenecks and migrations shape the genetic composition of these evolving populations. Within these distinct populations it is likely that large numbers of alleles and allele combinations of the nuclear and organelle genomes have been selected for their joint effect on fitness, thus leading to their co-adaptation or co-evolution (Falconer 1981; Mráček 2005; Landry *et al.* 2007).

The parent populations of interspecific hybrids may differ in many aspects. These differences include different alleles at many loci (Snyder *et al.* 1985), differences in allele frequencies (Snyder *et al.* 1985), differences in co-adapted gene complexes (Falconer 1981) and differences in regulatory elements of transcriptional networks (Landry *et al.* 2007). In addition there may be differences in chromosome arrangement due to inversions, duplications, deletions or translocations (Snyder *et al.* 1985). Chromosome re-arrangement may lead to differences in genome size and even basic chromosome numbers, while polyploidization events may lead to differences in ploidy. The magnitude of the differences between parental populations will depend on the degree of divergence between the populations.

The differences between parental populations determine the heterozygous nature of hybrids. The impact of the heterozygosity may be described on two levels, the genetic and the chromosomal level. Differences at a genetic level will have an impact on the phenotype of the plants. Although differences at a genetic level may lead to inviability or poor performance, they may also contribute largely to the success of hybrids. Heterosis, a greater adaptability to a range of environmental conditions, and the combination of parental traits may contribute to their survival in nature and to their success as domesticated crops. Differences at a chromosomal level may impact on the fertility and thus the survival of the hybrid genotypes. This in turn may have an impact on the ability to propagate hybrids.

1.2.1 Plant hybrid genomes

Hybrids may be heterozygous at a chromosome level if duplications, deletions, translocations and/or inversions have affected the linear arrangement of the genetic material. Such differences in chromosome organisation as well as differences in ploidy may lead to partial infertility as such hybrids may produce unbalanced and inviable gametes due to irregular meiotic paring (Rieseberg *et al.* 2000; Hegarty & Hiscock 2005). Where parent species have unequal genome sizes there will be genetic material that is hemizygous (Myburg *et al.* 2004). Hemizygosity may lead to a lack of allelic interaction in parts of the genome and potentially a higher proportion of deleterious genes being expressed. It is, however, expected that extreme chromosomal re-arrangements between populations will prevent the formation of hybrids.

The presence of these structural differences in the genetic material of the parents may also lead to a new series of re-arrangements. McClintock (1984) suggested that hybridization may lead to a 'genome shock' which may lead to genomic modifications such as chromosomal re-arrangements and mobilization of repetitive sequences. Some of these changes may continue to arise in subsequent generations (McClintock 1984).

1.2.2 Gene expression in plant hybrids

The combination of elements that regulate transcription from two species may lead to a 'transcriptome shock' (Hegarty *et al.* 2005). Gene expression is regulated by various elements of the transcriptional network including regulatory sequences, basal promoters and protein complexes (Landry *et al.* 2007). Expression is also modified by various epigenetic processes such as DNA methylation, histone modification and chromatin packaging (Hegarty *et al.* 2005). The elements of the transcriptional network, which include transcription factors and their binding sites, need to co-adapt to retain functionality of the genes (Landry *et al.* 2007). Changes that have evolved within one species may not be compatible with changes in another species and combination of the elements of two species may lead to altered gene expression, new expression-phenotypes or malfunctioning of the translation process (Landry *et al.* 2007). Several changes in gene expression have been reported in hybrids, including changes in DNA

methylation (Salmon *et al.* 2005), unequal expression of parental alleles, gene silencing, organ or tissue-specific allelic expression and up and down-regulation of genes relative to the parents (Zhuang & Adams 2007).

Genome-plastome incompatibility is a phenomenon that has been observed in a number of hybrids including in the genera *Oenothera, Trifolium* and *Impatiens* (Mráček 2005). The nuclear genome and the genomes of the organelles, the plastome and the chondriome, have co-evolved as the development and function of the plant cells depend on the interaction of these genomes (Mráček 2005; Turelli & Moyle 2007). Where hybridization is accompanied by the exchange of organelles between species, the intracellular genetic balance may be disturbed greatly, even in closely related species (Herrmann *et al.* 2003; Mráček 2005). If organelles are inherited maternally, the organelles of the maternal species are under the influence of the hybrid nuclear genome, potentially leading to changes in the expression level of the organelle genome (Hegarty *et al.* 2005).

1.2.3 Phenotype of plant hybrids

Burke and Arnold (2001) state that on average, interspecific hybrids tend to perform poorly. There are many reasons for poor performance. As both the superior and the inferior characteristics of the parents are transmitted to the next generation, some poor performance can be expected (Zobel & Talbert 1984). However, some crosses between species lead to outbreeding depression which is expressed as a reduction in fitness. This reduction may be ascribed to a loss in the local adaptation, underdominance (i.e. where homozygotes have and advantage over heterozygotes) or epistatic interactions (Edmands 1999). It has been suggested that the breakdown of co-adapted gene complexes of the parent populations will lead to the breakdown of biochemical and physiological compatibilities of alleles and subsequently lead to outbreeding depression (Siikamäki 1999; Hancock 2005).

On the other hand, some hybrids may show exceptionally good performance in relation to one or more traits. Heterosis refers to the superior performance of a hybrid relative to

its parents or to the superiority of heterozygous genotypes over homozygous genotypes (Rieger *et al.* 1991). Mid-parent heterosis is the improvement of hybrid performance over the mean of the two parents, whereas high-parent heterosis refers to cases where hybrids perform better than the best parent (Lamkey & Edwards 1999). Lamkey and Edwards (1999, p.31) state that "there are hybrids that do not exhibit heterosis but there cannot be heterosis without hybrids." Heterosis is not a phenomenon that applies to an entire hybrid population as it is only a few hybrid genotypes that outperform their parents and usually only under certain conditions as it is environment-dependent (Xu & Zhu 1999; Burke & Arnold 2001; Volker *et al.* 2008). Heterosis is generally considered to be the opposite of inbreeding, in other words that heterosis restores the fitness that has been lost through inbreeding (Falconer 1981). This was also the view of Shull, who coined the term heterosis and defined it as "…increased vigour, size, fruitfulness, speed of development, resistance to disease and to insect pests, or to climatic rigours of any kind, manifested by crossbred organisms as compared with corresponding inbreds, as the specific result of unlikeness in the constitutions of the uniting parental gametes" (Shull 1952, p. 48). There are, however, also reports of heterosis resulting from crosses between heterozygous parents such as crosses between forest tree species (Brewbaker & Sun 1999).

An investigation of 46 studies of plant F_1 hybrids shows that morphological traits are often not predominantly intermediate to the parents, but instead F_1 hybrids often show a mixture of parental and intermediate traits (Rieseberg & Carney 1998). In some case novel phenotypes may be obtained through new combinations of alleles as well as the interaction of different transcriptional networks of the parent species (Landry *et al.* 2007).

Interspecific hybrids may display the combined favourable characteristics of their parents in their phenotype. The superiority of such hybrids relies on the generation of new combinations of genes that may not be achieved through within-population crosses (Kain 2003). It is based on the simultaneous expression of the traits in the F_1 progeny (Verryn 2000). In breeding this combination of traits is referred to as complementarity.

Interspecific hybrids tend to show a higher degree of phenotypic stability or homeostasis across sites as their intrinsic variability lends them the capacity to buffer against fluctuations of the environment (Li & Wu 2000; White 2001). The great diversity of genes may also allow these hybrids to be adapted to a wider variety of environments (Lerner 1954 cited in Kain 2003).

1.2.4 Inheritance in plant hybrids

Segregation in the F_2 populations leads to a larger number of homozygote loci relative to the F_1 generation. The proportion of homozygote to heterozygote loci stabilises by the F_3 generation (Walsh & Lynch unpublished). Some decline in mean trait values is expected in the F_2 relative to the F_1. Such decline was demonstrated in a number of experiments (Walsh & Lynch unpublished).

In intra- and interspecific hybrids deviations from the expected Mendelian segregation ratios have been observed (Zamir & Tadmor 1986; Hall & Willis 2005). There are indications that some genomes are more likely to retain their integrity than others and that, as a consequence, these genomes transmit more alleles to the next generation than expected from Mendelian segregation laws (Zamir & Tadmor 1986). The resulting hybrid progeny is then more similar to one parent than to the other (Zamir & Tadmor 1986). This unequal segregation may be the result of differential inclusion of alleles in the gametes, differential survival of these gametes, differential success of gametes in fertilization or the differential survival of the zygotes (Fishman & Willis 2005). The degree of this transmission ratio distortion is likely to be positively correlated with the level of genomic divergence between the parents (Hall & Willis 2005).

Transgressive segregation is often noted in F_2 or later generation crosses and backcrosses (de Vicente & Tanksley 1993). Transgressive segregation refers to the phenomenon that the phenotypes of some individuals exceed the parental phenotypic values and is therefore similar to heterosis in the F_1 (Rieseberg *et al.* 1999). Such segregation is commonly observed in plants and tends to affect morphological traits (Rieseberg *et al.* 1999). The transgressive phenotypes are reported to be highly heritable

(Rieseberg *et al.* 1999). Many reasons have been proposed for transgressive segregation including complementary gene action of the parental species or unmasking of recessive genes that are normally heterozygous (de Vicente & Tanksley 1993; Rieseberg *et al.* 1999).

1.2.5 Population and quantitative genetics of plant hybrids

The population genetic dynamics of hybrids are influenced by the fact that hybrid populations are in Hardy-Weinberg disequilibrium (HWD) and linkage disequilibrium (LD). If a population is in Hardy-Weinberg equilibrium (HWE) there is a relationship between the gene frequencies and the genotype frequencies which remain constant over generations (Falconer 1981; Suzuki *et al.* 1981). If the allele frequency at a bi-allelic locus is p and the frequency of the other allele is q in the parent population, the genotypic frequency in the progeny will be p^2, 2pq and q^2. Deviations from the HWE are encountered in small populations, or may be the result of differences in fertility, non-random mating, migration and mutation (Falconer 1981). In natural populations hybridization may result from migration events. In breeding, the artificial mixture of populations is usually accompanied by a reduction in population size and assortative mating. Furthermore, in hybrid populations there is an excess of heterozygotes relative to the parental species or populations (Kain 2003). All these factors contribute to the fact that hybrid populations deviate from the HWE. In diploid organisms, only one generation of random mating is required to restore the HWE (Falconer 1981).

Linkage disequilibrium is a deviation from equilibrium with respect to two or more loci (Falconer 1981). Populations are in linkage equilibrium (LE) if alleles at different loci are distributed independently of each other in the gametes that form the next generation. In F_1 hybrids, however, the alleles on one chromosome originate from one species or population and the alleles of the homologous chromosome from the other species or population. The alleles have therefore not been distributed independently in the gametes that formed the F_1 (Kain 2003). Linkage equilibrium is progressively restored over a number of successive generations (Falconer 1981). Strong associations even between

unlinked genes are however sometimes maintained (Li & Nei 1974 cited in Barton 2000).

Most quantitative genetic concepts have been defined for large, randomly mating populations that are in Hardy-Weinberg and linkage equilibrium, however, the genetic organisation and structure of hybrid populations do not meet these assumptions (Gordon 1999; Wu *et al.* 2004). The definition of the additive and dominance genetic variance, for example, differs for hybrid and pure species populations. If a population is in HWE, the genetic variance (σ_g^2) is the sum of the additive (σ_a^2) and the dominance (σ_d^2) genetic variances. However, if a population is in HWD, the genetic variance also contains interaction between gene effects, allele frequency and coefficients of HWD (Wu *et al.* 2004). Gordon (1999) goes further and argues that additive and dominance genetic variances are not defined for hybrids as the concept of allele substitution is questionable in the case of hybrids. He therefore questions whether narrow-sense heritability (h^2) can be defined for hybrid populations. If these concepts are not valid, forward selection for breeding values can not be predicted for hybrid populations as it depends on the population additive genetic variance (Kain 2003).

Hybrids have been successfully produced and deployed in a large number of horticultural, agricultural and forest crops despite a limited understanding of their quantitative, molecular and epigenetic aspects.

1.3 Hybrids in plant breeding

The use of hybrids in plant improvement dates back to the late 19th century (Goldman 1999), but it was the success of intraspecific hybrid maize in the beginning of the 20th century that demonstrated the true potential of hybridization. The first hybrid maize was produced in the USA in the 1920s and yielded in excess of 15% more than the best open pollinated varieties (Duvick 1999). Shull (1908) proposed the use of single crosses between inbred lines but there were practical constraints as the inbred lines were weak. This problem was solved by Jones who in 1918 proposed the use of double crosses (Poehlman 1979). In the 1960s and 70s the single crosses gained increasing popularity

as improved pest control, soil fertility and cultural practices had improved the yields of the inbred lines (Poehlman 1979). By 1965 nearly all maize planted in the USA was of hybrid origin (Duvick 1999). Today approximately 95% of all maize planted in the USA and 65% of the maize worldwide are hybrids (Hochholdinger & Hoecker 2007).

The success of maize hybrids promoted the development of other agricultural, horticultural and forestry hybrid crops, including self-fertilizing crops (Fehr & Hadley 1980; Kuckuck *et al.* 1985; Singh *et al.* 2004). A number of vegetable and ornamental crops are cultivated almost exclusively as hybrids (Duvick 1999). By the end of the 20th century approximately 65% of the maize, sorghum and sunflower crops worldwide were hybrid-based (Lippman & Zamir 2007). It has been estimated that, depending on the crop, increases in yield of between 15% and 50% could be realised through hybridization (Lippman & Zamir 2007).

The seeds of the first hybrid crop species were mainly produced by open pollination, which yielded only 40 to 80% hybrid seeds (Kumar & Singh 2004). The use of self-incompatibility and cytoplasmic male sterility allowed for the large scale commercial production of hybrid crops (Kumar & Singh 2004). These systems have reduced the costs associated with manual emasculation and manual pollination considerably. In tomatoes for example, these activities accounted for 40% of the total cost of seed production (Kumar & Singh 2004). The systems have also ensured that the crosses yield 100% hybrid seed.

In forestry, there are many opportunities for interspecific hybridization as many genera consist of related species that have not developed strong barriers that prevent crossing (Wright 1976). Furthermore, as forestry largely depends on the vegetative yields, a decrease in fertility has less of an impact than in crops that are cultivated for their seeds or fruit (Wright 1976). The oldest known artificial tree hybrids go back to the late 1920s and early 1940s. They include *Quercus*, *Pinus*, *Juglans* and *Larix* hybrids (Khurana & Khosla 1998). Today many of the poplars, eucalypts and pines that are planted in commercial plantations are of hybrid origin (Khurana & Khosla 1998).

A number of hybrid breeding and production strategies have been developed to exploit the benefits of hybrids. The strategies developed for agricultural and horticultural crops generally differ significantly from the strategies that are commonly used in forest tree improvement.

1.3.1 Hybrid production strategies in agricultural and horticultural crops

The two principal reasons for creating agricultural and horticultural hybrid crops are hybrid vigour and introgression. The exploitation of complementarity and adaptability which are important in forest tree improvement seems to plays a smaller role in these crops.

The first commercial plant hybrids were made with the aim of exploiting hybrid vigour and it remains an important reason for the production of intraspecific hybrid crops today. Most hybrid breeding strategies for heterosis involve the crossing of homozygous plants to produce hybrids. Parental homozygosity is, however, not a prerequisite for the expression of heterosis in the offspring, but if the parents are homozygous, the progeny will be uniform (Dodds 1955). As self-fertilizing crops are homozygous at most or all loci by nature, it is possible to cross plants directly for assessment of hybrid vigour without prior inbreeding and production of inbred lines (Dodds 1955). In contrast, cross-fertilizing crops require the production of inbred lines prior to crossing. These inbred lines allow for the production of homozygous lines that can be maintained unchanged (Dodds 1955).

In the 1940s and 50s plant breeders came to realise the potential of introducing traits from wild relatives into related domesticated crops as a way of increasing the utilisable genetic pool (Gur & Zamir 2004, Hajjar & Hodgkin 2007). In this case interspecific hybrids are made between the domesticated crop and the wild relatives and are subsequently backcrossed to the domesticated crop parent to remove undesirable traits (Hajjar & Hodgkin 2007). Pre- and post-zygotic barriers as well as sterility of these hybrids have hampered the success of backcrosses, but the development of ovule culture

and embryo rescue have facilitated introgression from wide crosses (van Tuyl & Lim 2003).

Somatic hybridization allows for the transfer of genetic material, both nuclear and cytoplasmic, between species that cannot be crossed successfully (Waara & Glimelius 1995). Interspecific, intergeneric and even intertribal somatic hybrids have been produced (Waara & Glimelius 1995). In contrast to interspecific hybrids that have been produced by means of sexual reproduction, interspecific somatic hybrids are mostly polyploid as they combine the full genomes of both species. In some cases asymmetric hybrids are produced in which only a limited amount of the genetic information of one parent is transferred (Waara & Glimelius 1995).

An alternative strategy, which has shown success in the past but is not commonly applied, involves the creation of a new hybrid species. The hybrid is stabilised by means of polyploidization. One successful example is the production of Triticale, an intergeneric hybrid between wheat (*Triticum* spp.) and rye (*Secale* spp.) (Larter & Gustafson 1980). The early hybrids were sterile but fertile hybrids were obtained through chromosome doubling.

1.3.2 Hybrid production strategies in forest tree improvement

Hybrid breeding in forestry differs from the strategies that are widely applied in the breeding of crops and ornamental plants. In forestry the development of inbred lines is usually not a viable option, because of long generation intervals, high costs of controlled pollination, the severity of inbreeding depression and a potential temporary shift in selection emphasis to selection for fitness traits which may be at the expense of traits that are of economic value (Williams & Savolainen 1996). There are, however, a few research programmes that have investigated the use of inbreeding in forest trees (e.g. Durel & Kremer 1995; Wang *et al.* 1996; Hettasch *et al.* 2007). Backcrossing followed by selection is rarely applied in forestry because of the long generation intervals. Therefore, introgression is less likely to become a widely adopted strategy. In

forestry thus, hybridization usually involves the crossing of different species or provenances without the use of inbreeding or backcrossing.

Clonal forestry is generally used to capture the hybrid vigour or combination of traits for deployment in plantations (Bradshaw & Grattapaglia 1994). Hybrids are usually deployed as clones because hybrid seed production relies largely on controlled pollinations, which are associated with high costs and low success rates. The use of controlled pollinations is necessitated by the apparent lack of male sterility and self-incompatibility systems in forestry species. Thus far there seems to have been only one study on the use of self-incompatibility systems for the production of eucalypt hybrids (Junghans *et al.* 1998). The study was based on the use of a single highly self-incompatible *E. grandis* clone in a seed orchard. The study showed that the concept is feasible, as isozyme studies showed that more than 95% of the seeds were hybrids. However, the success of such a system relies on the availability of such highly self-incompatible genotypes and because these have not been identified on a larger scale it is not a viable option at this stage. The use of the clonal deployment option, however, tends to limit the use of interspecific hybridization to genera that can easily be multiplied by means of vegetative propagation, such as *Eucalyptus* (some species), *Salix* and *Populus* (Bradshaw & Grattapaglia 1994). In these genera, it is possible to test the hybrid clones before selected clones are deployed on a commercial scale. However, in tree species where true-to-type mass propagation of mature material is technically challenging, as is the case in *Pinus*, clones may be deployed as untested half-sib or full-sib families.

1.4 *Eucalyptus* hybrids – a case study of forest tree hybrids

Eucalyptus hybrids have been successfully deployed in plantations in many countries in the tropics and subtropics. The combination of hybrid breeding and clonal propagation has proved to be very successful in improving wood and wood properties (Grattapaglia 2007). *Eucalyptus* hybrids will be discussed here to illustrate the successes and limitations of hybrid breeding in forestry and to provide a background to the material used in this study.

1.4.1 Introduction to eucalypt species

Eucalypts originate from Australia and the islands north of Australia. They are the most widely grown tree species throughout the world and are grown in most of the tropical and temperate climatic regions of the world between the latitudes 45°S and 40°N (Eldridge *et al.* 1997). Globally, approximately 18 million hectares of *Eucalyptus* forests have been planted in more than 90 countries (FAO 2001). The success of eucalypts lies in their intrinsic capacity for fast growth and their ability to withstand and recover from harsh environmental conditions (Eldridge *et al.* 1997). They are cultivated to meet a wide range of needs including paper pulp, fuel, poles and timber (Eldridge *et al.* 1997). It was however, an increased demand for short fibre pulp, which has lead to an immense increase in the planting of eucalypts during the 20th century (Turnbull 1999).

Eucalypts are evergreen hardwood species that vary in growth habit from shrubs to large trees (Potts 2004). They regenerate by means of seed but have mechanisms of vegetative recovery when damaged (Potts 2004). Although eucalypts preferentially outcross, they have a mixed mating system and a degree of self fertilization does occur (Potts 2004).

1.4.2 Eucalypt classification

The genus *Eucalyptus* was formally named by L'Héritier de Brutelle in 1789 (Brooker 2000). The classification of eucalypts has, however, been challenged in recent years. The first comprehensive classification was published in 1934 by Blakely (Brooker 2000). Later, a new informal classification was published by Pryor and Johnson (1971), who developed a classification system, which groups the species into subgenera, sections, series, subseries, superspecies, species and subspecies. In this classification the authors proposed the recognition of seven subgenera. In 1995 Hill and Johnson proposed a new genus, *Corymbia* Hill & Johnson, which contains the ghost gums and bloodwoods, previously belonging to the genus *Eucalyptus.* The new genus was accepted by many scientists, but some institutions such as the Australian National Herbarium, have followed a more conservative approach, by rejecting the new genus

Corymbia (Euclid n.d.). In 2000, Brooker presented a new classification system based on Pryor and Johnson's classification but combining the genera *Angophora* Cav. and *Corymbia* into one single genus, *Eucalyptus*, consisting of 13 subgenera. The validity of this classification has been disputed by Ladiges and Udovicic (2000). It has become clear that eucalypts form a large group with complex relationships which are, as yet, not completely understood. In this dissertation the term eucalypt is used to recognise the variation in terminology portrayed by the different classification systems.

Depending on the classification between 700 (Delaporte *et al.* 2001) and 800 (Potts *et al.* 2000) *Eucalyptus* species have been recognised. Most plantation eucalypt species belong to the subgenus *Symphyomyrtus* and most are from major sections *Transversaria* (e.g. *E. grandis* and *E. pellita*), *Exsertaria* (e.g. *E. camaldulensis*), *Maidenaria* (e.g. *E. globulus, E. nitens, E. dunnii*), *Bisectaria* (e.g. *E. cladocalyx, E. horistes, E. kochii, E. occidentalis, E. polybractea*) and *Adnataria* (e.g. *E. sideroxylon, E. tricarpa*) (Potts *et al.* 2000).

1.4.3 Eucalypt genome

Eucalypts are diploid and have 11 pairs of chromosomes (Eldridge *et al.* 1997). No natural polyploids have been identified thus far (Eldridge *et al.* 1997). The nuclear DNA content varies among species. Sizes varying from 370 mega base pairs (Mbp) for *E. citriodora* to 710 Mbp for *E. saligna* have been reported (Grattapaglia & Bradshaw 1994). Species from the same sections tend to have a similar nuclear DNA content, whereas species from different subgenera tend to have large differences in genome size (Grattapaglia & Bradshaw 1994). Hybrids tend to have genome sizes that are intermediate to the parental genomes (Grattapaglia & Bradshaw 1994).

A review of the literature has revealed only two reports of studies investigating the correspondence between the genomes of *Eucalyptus* species. Myburg *et al.* (2003) studied genome differences between *E. grandis* (section *Transversaria*) and *E. globulus* (section *Maidenaria*) and the corresponding hybrid. They found general co-linearity of the genomes with no gross chromosomal re-arrangements, and therefore suggested that

incompatibility between these species is genic rather than chromosomal. The two species studied by Myburg *et al.* (2003) differ in genome sizes with the nuclear genome size of *E. grandis* being larger than that of *E. globulus*. Their study of the F_1 hybrid revealed that hemizygous loci were dispersed throughout the *E. grandis* chromosomes. Brondani *et al.* (2006) who studied differences between *E. grandis* and *E. urophylla* (both section *Transversaria*) reported that they had found no evidence of chromosomal re-arrangements in an *E. grandis* × *E. urophylla* hybrid.

1.4.4 Eucalypt species cross compatibility

There are many records of interspecific hybrids in eucalypts, both natural and artificial (Meddings *et al.* 2003). Eucalypt species tend to have weak reproductive barriers (Potts *et al.* 2000). However, the authors did report that there is a tendency for natural hybridization to decrease with increasing taxonomic distance between parents where subgenera tend not to hybridize; intersectional crosses are less likely to be successful than interseries crosses; intraseries tend to be more successful than interseries crosses. On the other hand, both interseries and intersectional F_1 eucalypt hybrids have been artificially produced (Potts *et al.* 2000).

Currently most interspecific eucalypt hybrids that have been developed and tested are from the subgenus *Symphyomyrtus,* from the sections *Maidenaria*, *Exsertaria* and *Transversaria* (Potts & Dungey 2004). Of these the most important commercial eucalypt hybrids include *E. grandis* × *E. urophylla*, *E. grandis* × *E. camaldulensis* and hybrids including at least one of the species *E. saligna, E. pellita, E. exserta* and *E. tereticornis* (Potts & Dungey 2004).

A number of reproductive barriers have been identified in the genus *Eucalyptus.* Potts *et al.* (2003) have identified two major pre-zygotic barriers to hybridization. The first barrier is a unilateral structural barrier which relates to differences in flower size, where pollen tubes from small-flowered species (e.g. *E. nitens*) are not able to reach the ova of large-flowered species (e.g. *E. globulus*). A second pre-zygotic barrier is physiological and results in pollen tube abnormalities and pollen tube arrest in the pistil (Potts *et al.*

2003). Physiological barriers may also act after fertilization and take the form of failure of the embryo to divide, slow embryo development or reduced cellularisation of the endosperm (Potts & Dungey 2004). Knowledge of reproductive barriers has contributed to the development of techniques to overcome physiological and morphological barriers between species when producing artificial hybrids (Potts *et al.* 2003).

1.4.5 Eucalypt hybrid inviability

In the event of successful interspecific hybridization, the resulting F_1 hybrids often produce abnormal plants; referred to as hybrid inviability. The frequency of hybrid inviability is often high in crosses between highly diverged parent species (Potts *et al.* 2003). Such hybrid inviability often manifests shortly after germination; expressed as failure to develop seedling leaves, seedling abnormalities or dwarfism, usually leading to reduced growth and survival at later stages (de Assis 2000). In some cases young seedlings appear to be normal with inviability only manifesting at a later stage (Potts *et al.* 2003); up to two years after establishment in the field (de Assis 2000). The impact of inviability on hybrid crosses is illustrated by Griffin *et al.* (2000) who reported on crosses between *E. globulus* and *E. grandis*. Only 0.15% of the F_1 seed harvested produced normal plants after two years of field growth.

1.4.6 Eucalypt hybrid breeding

Most eucalypt improvement programmes are focused on the improvement of eucalypts for pulpwood production. These programmes aim at improving volume production, wood density and pulp yield (Grattapaglia 2007). Secondary in importance are traits such as disease resistance, pest resistance, and resistance against wind, drought, salinity and frost. In the past decade, however, there has also been interest in improving eucalypts for products such as saw timber, veneer and medium density fibreboard (Grattapaglia 2007). The production of hybrids is one of the breeding options explored by eucalypt breeders to improve the genetic resource.

Many of the early successes of eucalypt hybrids have resulted from spontaneous crosses, rather than from well-designed breeding programmes. The PF1 and HS2 varieties deployed in Congo as well as the *E. urograndis* clones that have been selected from open pollinated orchards in Aracruz, Brazil, are examples of such hybrids (Griffin *et al.* 2000; Vigneron *et al.* 2000). Although in both programmes there have been efforts to improve on the performance of these hybrids, it has proved to be difficult to exceed the performance of the spontaneous hybrids (Griffin *et al.* 2000).

Breeders are increasingly interested in breeding eucalypt hybrids, despite difficulties in improving on the performance of spontaneous hybrids (Volker *et al.* 2008). The authors suggest that eucalypt hybrid development is usually aimed at breeding hybrids that are adapted to marginal sites, where neither parent species thrive, or alternatively to combine economically important traits. There is also interest in the exploitation of heterosis, although hybrid superiority may often rather lie in complementarity and adaptability than true heterosis (Verryn 2000; Volker *et al.* 2008).

Most traits are reported to be intermediate to the traits of the parent species, in particular many of the morphological and physiological traits (Tibbits *et al.* 1991; Delaporte *et al.* 2001, Potts 2004; Volker *et al.* 2008). There are, however, a number of reports of heterosis in eucalypts, but not all cases have been confirmed as many trials are unsuitable for the identification of heterosis as there are no proper pure species controls and therefore the observed heterosis could be a case of adaptability (Volker *et al.* 2008). Heterosis is often associated with growth traits (Bison *et al.* 2006). Reports include better-parent heterosis in *E. camaldulensis* × *E. grandis* and *E. camaldulensis* × *E. globulus* in saline conditions in Australia (Dale & Dieters 2007), mid-parent heterosis in diameter at breast height in *E. grandis* × *E. urophylla* hybrids in Brazil, and *E. saligna* × *E. urophylla* hybrids in South Africa which have displayed superior growth performance to both parental species, even in sites where one would expect the parents to perform best (Pierce 1994 a).

Many eucalypt improvement programmes aim at combining the superior traits of two parent species. Examples include attempts to combine the fast growth of *E. grandis* with high salt tolerance of *E. camaldulensis* (Dale & Dieters 2007); high pulp yield and good

growth of *E. globulus* with salt-water logging tolerance of *E. camaldulensis* (Meddings *et al.* 2003); fast growth of *E. grandis* with higher wood density of *E. saligna* (Gwaze *et al.* 2000) and good wood quality and growth of *E. grandis* with the resistance to canker in *E. urophylla* (Muro-Abad *et al.* 2005).

Two breeding strategies are being used for the production of F_1 eucalypt hybrids. The two strategies, recurrent selection and reciprocal recurrent selection, are discussed in section 1.5.2.2.

High costs of controlled pollination and low seed set because of incompatibility prevent the production of control-pollinated hybrid seed in commercial quantities. Most commercially planted eucalypt hybrids are therefore propagated as clones (Eldridge *et al.* 1997; Shelbourne 2000). The clones are usually propagated from mature ortets which are induced to produce juvenile coppice shoots for propagation purposes (Shelbourne 2000). Some species, for example *E. nitens*, do not root well and are thus vegetatively propagated from very young seedlings (Shelbourne 2000).

1.5 Prediction of performance in plant hybrid populations

The progress of any genetic improvement programme relies on accurate predictions of the genetic value of individuals or families (Bouchez & Goffinet 1990; Lofgren & Stewart 1991). Breeders are, therefore, interested in obtaining the best predictions of the genotypic values possible to allow for the correct identification of the best individuals (Bouchez & Goffinet 1990). When selection traits are highly heritable simpler selection techniques based only on individual tree performance may provide sufficient genetic gains. With low heritability traits, more sophisticated techniques are generally needed for adequate progress as individual tree selection is not expected to be efficient (Cilas 2003; Walsh 2007).

Breeders wanting to select clones or seed parents from hybrid populations need to be able to predict the performance of the progeny or clones. The reliability of the identification of superior genotypes is key to the success of selection, and therefore the

genetic gains that can be achieved through breeding. However, in hybrids there are many factors that may impact negatively on the ability to predict performance, e.g. altered gene expression, unequal gene expression and unequal segregation. In addition, it is not known whether current selection methodology can be successfully applied to hybrids, because hybrid populations do not conform to the assumptions of large randomly mating populations which underlie current selection theory. Furthermore, disequilibrium in the hybrid population leads to instability of the genetic structure from one generation to the next (Kain 2003). This presents a problem for selection theory, as it relies on genetic equilibrium to predict the outcome of forward selection and recombination reliably (Kain 2003).

1.5.1 Predicting performance in agricultural and horticultural hybrid crops

There are two stages of selection that relate to hybrid breeding programmes. One is the selection of parents which are to be hybridized, and the other is the selection of individuals within a hybrid population for further breeding or production purposes. The techniques used to select in agricultural and horticultural crops will be discussed only briefly.

1.5.1.1 Selection of parents for the production of hybrids

Research on prediction in hybrid breeding programmes in agricultural and horticultural crops, seems to have been focused on the prediction of hybrid performance based on information of the parents to allow for the selection of the best parents, preferably without the need of prior field testing of the hybrids. Many techniques have been explored, including techniques such as genetic distance models (e.g. Lanza *et al.* 1997), factorial regression models (e.g. Charcosset *et al.* 1998), marker associated studies (e.g. Vuylsteke *et al.* 2000), methods based on best linear unbiased prediction (e.g. dos Santos *et al.*, 2005; Piepho *et al.* 2008), and the support vector machine regression technique (e.g. Maenhout *et al.* 2007). Some methods such as the support vector machine regression and marker-based BLUP have been reasonably successful in maize (Maenhout *et al.* 2007) but will not be discussed here in detail as these prediction methods are generally not related with the problem of prediction in a hybrid population.

There seems to be limited information on predictions of the performance of the progeny of hybrids.

1.5.1.2 Selection of plants for clonal production

Selection in many asexually propagated crops, both hybrids and pure species, including sugarcane, potatoes, forage crops, rubber plants, cassava and several fruit species, is based on mass selection and therefore family performance is not taken into account (de Resende & Barbosa 2006). Better results may be obtained if family and within family selection is practised and it is therefore expected that better results may be obtained through improved selection methodology, such as BLUP.

Best linear unbiased prediction (BLUP see section 1.5.2) is widely used for selections in animal breeding programmes, but it has not gained the same degree of popularity amongst plant breeders working with annual crops as trials in these crops are usually well balanced and only information from one generation is considered in the selection process (Piepho *et al.* 2008). It has, however been used in a number of forestry and perennial crops. One example of BLUP being used in an improvement programme of an annual crop is that of the sunflower (*Helianthus annuus* L.) improvement programme in Argentina. Here BLUP was used to measure genetic progress of the intraspecific hybrid programme over a number of years and a number of sites rather than for selection of hybrids (de la Vega *et al.* 2007).

BLUP has been applied for clonal selection in some asexually propagated crops. BLUP was, for example, used in Brazil to predict genetic and genotypic values of *Panicum maximum* plants obtained from intraspecific hybrid crosses (Resende *et al.* 2004). The aim was also to predict the performance of the progeny from existing and potential hybrid crosses. The analysis needed to be adapted to take into account the autotetraploid nature of the crop. The genotypic values were predicted for the production of apomictic hybrids (or clones), whereas the sexual hybrids were selected based on their additive genetic value. In a cashew improvement programme in Australia, BLUP was used to determine breeding values to identify parents for further crosses (Blaikie *et al.* 2002). At the same time, individuals were selected on the basis of their phenotype for vegetative

deployment. The breeders found a large degree of correspondence between the individuals that were identified by means of the two different selection methods. In a sugarcane improvement programme in Brazil, BLUP was adapted to deal with the fact that in some crop species (such as sugarcane, soybean, rice and some grain crops) it is easier to record performance on a family rather than an individual basis. De Resende and Barbosa (2006) have, therefore proposed a BLUP-based analysis for these crops which they have named 'simulated individual BLUP'. The method is used to determine the number of families that are to contribute to the clonal selections, the number of individuals per family which are to be selected, and the total number of clones.

There are only a limited number of reports on selection of clones in agricultural and horticultural hybrid crops. The examples cited illustrate that, although mass selection seems to be commonly used, some breeders are applying BLUP for the selection of clones. In the above mentioned examples the applicability of BLUP in hybrid populations was not questioned. There are also no indications as to whether BLUP was successful in identifying the best genotypes.

1.5.2 Predicting performance in hybrid forest tree species

In forest trees several strategies are employed for the selection of parents for hybrid production and for the selection of hybrids for production purposes. These strategies require different selection methods such as forward, backward and clonal selection. Different selection techniques are used in forestry to address needs that are unique to forestry. As improvement of horticultural trees has similar constraints, examples from these programmes will also be cited.

1.5.2.1 Selection techniques used in forestry

In tree breeding, breeders are faced with long generation intervals, as well as environmental and to a lesser extent seasonal effects. The data obtained from such forestry trials are often 'messy' (Cotterill & Dean 1990). It is often found that parents are not represented in equal numbers at all sites or have not been planted at all sites. The precision of tests at sites may also differ (Cotterill & Dean 1990). There are a number of

different selection techniques that have been used in forest tree improvement programmes but not all deal with the problem of unbalanced data. It is believed that these selection techniques are applied both in hybrid and pure species improvement programmes.

Individual selection or mass selection is a technique in which the individual trees are assessed and selected on the basis of their phenotype only (Cilas 2003). No information on the performance of relatives is used (Cotterill & Dean 1990). This method is typically applied when selecting trees from unimproved stands where family structure is not known (Cotterill & Dean 1990). The success of this type of selection mainly depends on the heritability of the trait: success is limited if the heritability is low, whereas more success can be expected from traits of high heritability (Cilas 2003).

The Smith-Hazel index (Smith 1936; Hazel 1943) is a selection index which has been widely applied in forestry (White & Hodge 1989; Cotterill & Dean 1990). The index places the same relative weights to family mean and single tree values for all families. This would only be appropriate if all families are tested equally; the same number of individuals per family, the same age and equal precision of different tests (White & Hodge 1989). The index will, therefore, only give the best linear prediction of the breeding values of individuals if the data were balanced (White & Hodge 1989).

Henderson (1984) developed best linear prediction (BLP), a method that is based on mixed models that is appropriate for unbalanced data. BLP and best linear unbiased prediction (BLUP) calculate index weights for each candidate depending on the quantity and quality of the data (White & Hodge 1989). While earlier approaches used in tree breeding treat breeding values as fixed effects, both BLP and BLUP treat genetic values as random effects. With BLP it is assumed that the fixed effects are known constants, whereas with BLUP, the fixed effects are estimated simultaneously with the prediction of the random genetic effects (Henderson 1984). The predictions of BL(U)Ps are unbiased, the error variance of the predictions are minimised and the correlations between the predicted and true genetic values are maximised (Henderson 1984; White & Hodge 1989).

In hybrids these selection methods are applied for the selection of parents which are to be hybridized as well as for selection of hybrids for commercial deployment.

1.5.2.2 Strategies for the selection of hybrid parent trees

There are two basic breeding strategies that are used for the selection of parents in forest hybrid improvement programmes. Where hybrid performance can be successfully predicted from the performance of the parents in pure species combinations, recurrent selection (see Figure 1.1) may be applied. However, where the correlation between pure species parental performance and hybrid performance is poor, a reciprocal recurrent selection strategy (see Figure 1.2) needs to be implemented (Vigneron *et al.* 2000; Verryn 2000; Shelbourne 2000; Dieters & Dungey 2000; Potts & Dungey 2004). There are contradictory reports about these correlations and it seems to be a function of both parental species and the traits (Potts & Dungey 2004; Volker *et al.* 2008). Poor correlations between pure species parental performance and hybrid performance may be ascribed to three reasons (Potts & Dungey 2004; Volker *et al.* 2008): different genes may be involved in trait expression in the hybrids and pure species; chromosomal re-arrangements may affect gene expression in the hybrids; and there may be a larger proportion of non-additive genetic variance in hybrids.

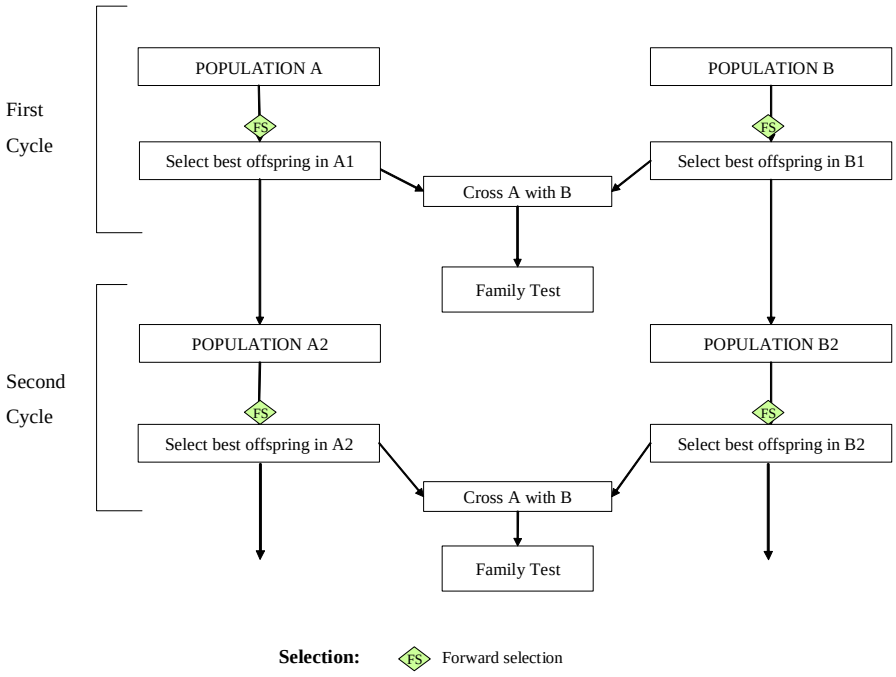


Figure 1.1 Recurrent selection strategy for breeding hybrids.

In a breeding programme that is based on the recurrent selection strategy, the hybrid family tests are aimed at the selection of material for commercial deployment. The parents are selected only on the basis of their performance as pure species, therefore any problems with reliability of the selection in hybrid populations does not affect the selection of the parents in this strategy.

Poor correlations between hybrid and pure species performance may dictate that parents be selected based on hybrid performance. In the reciprocal recurrent selection strategy, parents for hybrid production are selected based on the information on the performance of the hybrid progeny in the family tests. Kain (2003) argues that prediction of the degree of improvement from backward selection is a statistical matter and requires no assumptions on the genetic architecture of the hybrid population.

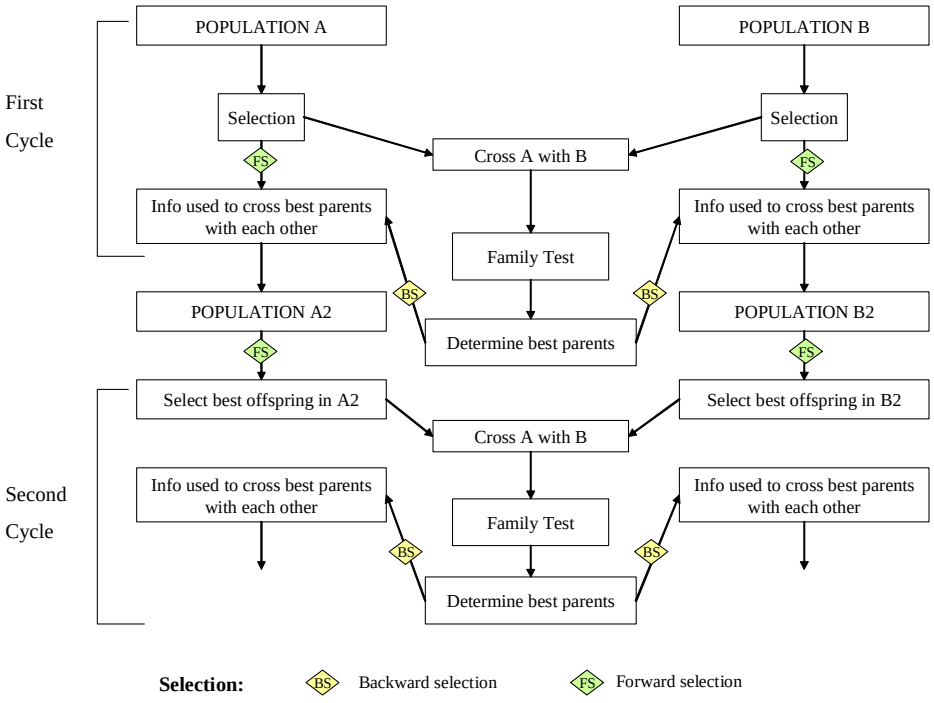


Figure 1.2 Reciprocal recurrent selection strategy for breeding hybrids.

Volker *et al.* (2008) state that there is no reliable method to predict which parents will yield good F_1 hybrids. The authors based this statement on research of *E. globulus* × *E. nitens* hybrids, which tend to have a high incidence of abnormalities and generally produce poor performing hybrids, despite the fact that both species belong to the section *Maidenaria* (Brooker 2000). Their experience may not be valid for other hybrid combinations.

An example of successful prediction of hybrid performance was reported by researchers working on the improvement of oil palm. In both Malaysia and Indonesia research has shown that BLUP was suitable for the prediction of hybrid performance as high correlations were obtained between predicted and observed hybrid performance (Soh 1994; Purba *et al.* 2001). The improvement programme was based on reciprocal recurrent selection, but hybrids were intraspecific hybrids between different groups or provenances.

It has been suggested that the use of molecular markers may assist in the selection of parents for the production of successful hybrid combinations, as has been investigated in crop species (see section 1.5.1). One option is to determine the genetic distance between parents and select only parents that are genetically diverse. This has proved to be successful in the selection of parents in an *E. globulus* improvement programme, where such crosses tended to produce the best full-sib families (MacRae & Cotterill 2000). However, if the improvement of hybrids is a function of complementarity rather than heterosis, this method is likely to fail. Henry and Shepherd (2000) suggested that such a strategy may be of less value in the highly heterozygous tree species as high levels of heterozygosity may not be related to performance.

In the view of the uncertainty around the prediction of performance in hybrids, molecular markers may play a role in more accurate identification of traits through marker-aided selection. However, the reasons that may be responsible for poor correlation between parental and hybrid performance (other genes that may play a role in determination of the trait, chromosomal re-arrangement and large proportion of non-additive genetic variation), will still play a role.

1.5.2.3 Selection for deployment of hybrids in plantations

Once parents have been selected and crosses have been made, hybrids need to be tested and selected for commercial deployment in plantations. There are a number of different options for the deployment of hybrids. Some of these options are illustrated in Figure 1.3. In option 1 hybrid crosses are made by means of controlled pollination and the seedlings are established in a family trial. The family trial is evaluated and genotypes are selected for vegetative propagation. Clones are established in clonal trials for the evaluation of clonal performance. These trials may be repeated at a number of sites. Selected clones may be re-tested at a larger number of sites before clones are selected for deployment in plantations. This strategy is, for example, generally applied for the deployment of eucalypt hybrids.

In option 2 the time to deployment is reduced by cloning the hybrid seedlings and planting clonal trials without prior evaluation in family trials. Family identity is

however, retained in the clonal trials to allow for the selection the evaluation of family and individual performance.

Options 3 and 4 involve the deployment of hybrids as seedlings. The first option is to deploy hybrids as F_1 seed. The best performing parental combinations are selected based on performance in family trials. These controlled crosses are then repeated for deployment in plantations. There is, however, a risk of high within-family variance in the F_1 (Volker 2001). But, as the success of controlled pollinations is relatively low and the associated costs are high, it has been suggested to deploy the hybrids as F_2 'hybrids' (option 4).

In option 4 individuals that are selected from the family trials are grafted in a seed orchard. F_2 seed is produced by means of open pollination. This reduces the cost of mass production of the hybrid significantly, but there is the risk of variability in performance of the F_2s due to segregation.

Options 5 and 6 are primarily designed for genera where vegetative propagation of mature material is not feasible. In option 5, the hybrids are established as seedlings in a family trial for the evaluation of hybrid performance. The crosses that yield the best hybrids are identified and repeated. The resulting seedlings are then established in hedges from which material is harvested for the establishment of clonal trials. Clones are deployed as families rather than individual genotypes. This deployment strategy is also known as family forestry.

In option 6 the seedlings are immediately established in hedges and clonal trials are established from material collected from these hedges. If the hedges can be maintained in a juvenile condition until the trial has been evaluated, selected families may be harvested from the original hedges for deployment purposes. Alternatively, new hedges could be established.

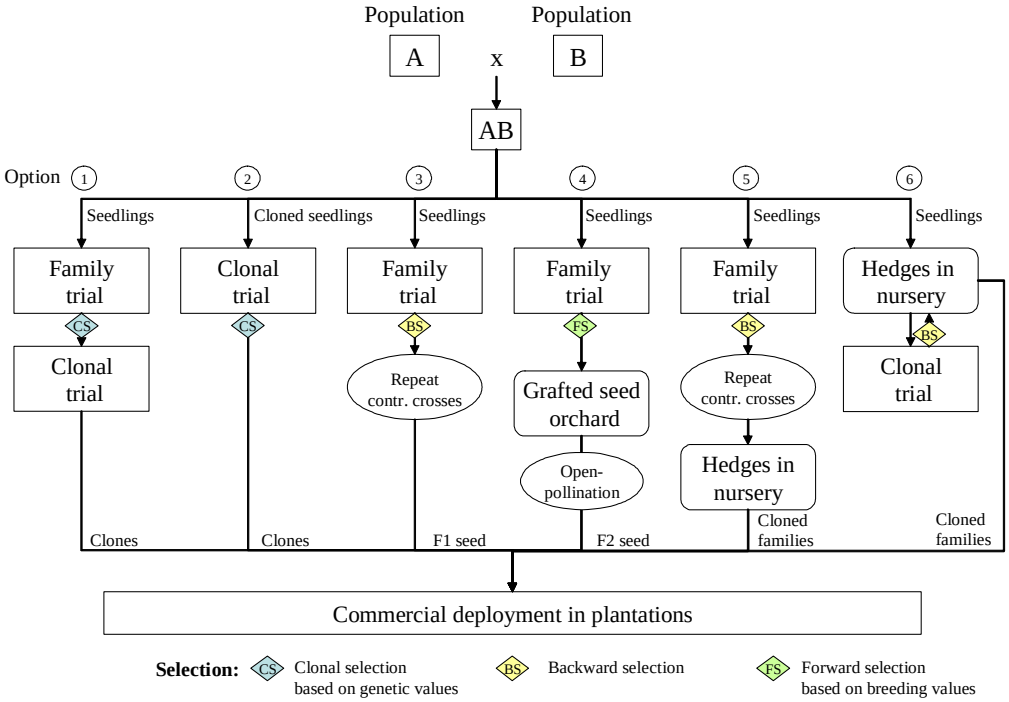


Figure 1.3 Different options available in forestry for the deployment of hybrids.

There are three different types of selection required in the deployment options presented. Clonal selection is based on the genetic value of an individual and is used for the selection of clones in options 1 and 2. In the context of this study, the term ‘clonal selection’ is defined as the forward selection of ortets for the production of clones. Backward selection is used where crosses have to be repeated (options 3 and 5) or where the best families in hedges are selected (option 6). Forward selection is based on the breeding value of an individual whereas backward selection uses breeding values of the parents. Foreward selection is applied for selection of individuals that are to be grafted in a seed orchard.

In hybrid populations it is difficult to successfully predict the performance by means of forward selection, as it depends on additive genetic variation which may not be defined

(see section 1.2.5) and due to linkage disequilibrium (Kain 2003). The efficiency of forward selection in hybrid populations needs to be investigated.

In this study, the focus is, however, on the efficiency of clonal forward selection as many forestry hybrid programmes rely on variations of option 1 for the deployment of hybrids, whereas the seed production route is less commonly used at this stage. Selection of ortets for the purpose of vegetative propagation differs from forward selection for breeding purposes. The genetic gain achieved by the selection of parents for seed production is based almost entirely on the additive component of genetic variation whereas the genetic gain achieved by the selection of ortets for clonal production is based on both additive and non-additive effects, as the clones retain the effect of allelic interaction (Mullin & Park 1992). Therefore, selection of clones (for example in deployment options 1 and 2) should be based on both additive and non-additive genetic components. Because narrow-sense heritability only accounts for the additive genetic variation that is transmitted from one generation to the next, broad-sense heritability is appropriate for clonal selection.

A survey of the literature has provided limited information on the selection of ortets. In a number of improvement programmes where selection was based on individual tree performance only, researchers realised that this strategy could be improved by adding information on family performance. For example, in oil palm improvement based on intraspecific hybrids, ortet selection was based on individual tree or mass selection, but it has been suggested that selection could be improved by first selecting the best families and then the best individuals within those families (Kee *et al.* 2003; Soh *et al.* 2003). In the improvement of cocoa (*Theobroma*) it was also found that individual tree selection was not efficient for the selection of ortets and it was suggested that combined individual-family selection would be more appropriate (Eskes 2004). The clonal selection strategy can be further improved by the use of BLUP. This method was applied for the selection of *E. urophylla* × *E. grandis* and *E. urophylla* × *E. pellita* ortets in the Congo (Combes *et al.* 1997). The efficiency of such selections in hybrid populations was not investigated.

1.6 Objective of the study

Hybrid populations differ in many aspects from pure species. Selection theory has been developed on assumptions of large randomly mating pure species populations. Hybrid populations do not, however, conform to these assumptions. It is therefore largely speculative whether selection in hybrid populations is efficient at identifying superior genotypes. The objective of this study is to investigate the applicability of BLUP in hybrid populations.

As case study selection of ortets was investigated in three *Eucalyptus* hybrids. Hybrids between species that are closely related and species that have diverged to a greater degree were chosen. The three hybrids that were assessed included: *E. grandis* × *E. saligna*, *E. grandis* × *E. urophylla* and *E. grandis* × *E. camaldulensis.*

The objective was addressed by investigating:

- correlations of predictions of clonal performance in eucalypt hybrid family trials with realised performance of the clones in clonal trials,
- comparisons of a number of BLUP selection options to see which method achieves the best predictions,
- comparisons of the correlations of the different hybrids with each other to determine whether there are differences between narrow and wide hybrids, and
- comparisons of the correlations obtained for the *Eucalyptus* hybrids with the corresponding correlations in *E. grandis* in order to determine whether there are differences between pure species and hybrid populations.

CHAPTER 2
MATERIALS

2.1 Introduction

The aim of this study was to use historical data to investigate whether best linear unbiased prediction (BLUP) may be applied for the selection of clones in three *Eucalyptus* hybrid populations. The use of historical data has the advantage of data being immediately available for analysis with no establishment, maintenance and assessment of trials being required. This is of particular importance in forestry as these activities are expensive and it would take many years to collect the appropriate data. The disadvantage of using historical data is that the trials are not optimally designed for the study and the researcher has no control over factors such as trial design, genetic controls, traits assessed or age of assessment.

This study was planned to investigate whether the predicted performance of clones was realised in clonal trials in order to obtain an indication of the efficiency of the best linear predictions in hybrids. An outline of the research strategy is presented in Figure 2.1. Clonal performance was to be predicted from a series of hybrid family trials (step 1 in Figure 2.1) whereas realised clonal performance was to be assessed in clonal trials (step 2). The predicted and realised clonal performances were to be correlated (step 3) to assess the efficiency of predictions. It was proposed to use an array of hybrid parental species to allow for the comparison of prediction in hybrids between closely and distantly related parent species (step 4). In order to detect differences in the efficiency of prediction in hybrids and pure species, it was proposed to predict clonal performance in a pure species case (step 5) and to determine the realised performance of these clones (step 6). The predicted and realised performances were to be correlated (step 7). These correlations were then to be compared with the corresponding correlations obtained in the hybrid populations (step 8).

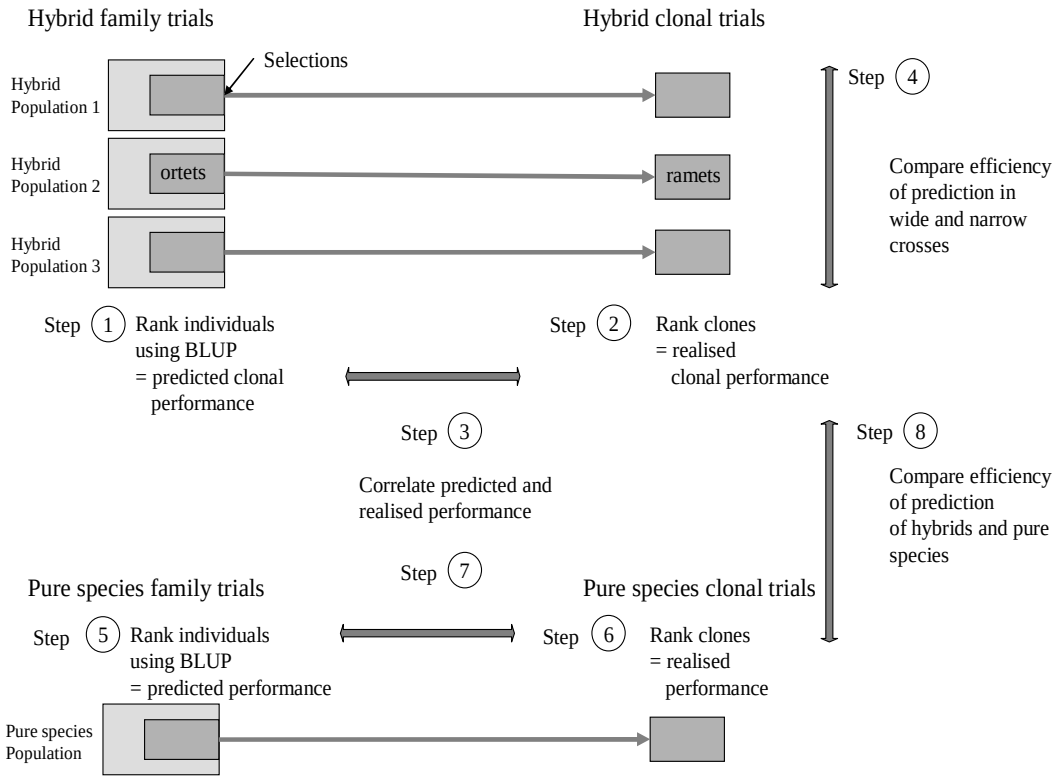


Figure 2.1 Research strategy to investigate the performance of BLUP in hybrids.

This strategy required data on ortet performance in family trials and clonal performance in clonal trials from both hybrids and pure species. The clonal performance needed to be linked to the performance of the ortet or mother tree from which the ramet originated. Appropriate data were selected from hybrid and pure species breeding populations. These data will be described in this chapter together with the criteria that were used to determine the suitability of the data for the purpose of this study.

2.2 Data selection criteria

The integrity of the results of the study relied to a large extent on the integrity and suitability of the data used. The data needed to meet a number of requirements to ensure meaningful genetic analysis and reliable statistical analysis. Genetic and statistical criteria were therefore developed to evaluate the suitability of the data.

2.2.1 Genetic criteria

A number of criteria were drawn up to determine whether the data were suited to the genetic analyses that formed part of this study. The data had to comply with the following set of criteria:

- **Appropriate trials:** Sets of hybrid family trials and corresponding clonal trials were required. Records of the origin of the ramets were required to establish the relationship between the ortets and ramets. As a control, pure species family trials and clonal trials were required for the comparison of the efficiency of prediction in hybrids as opposed to prediction in pure species.
- **Appropriate traits:** Traits that were assessed in both hybrid and clonal trials were a prerequisite for the study. Traits should preferably be of medium to high heritability. Traits which were not selected with a high selection intensity would be preferred as selection would reduce variability between clones.
- **Differences in the degree of relatedness of the parent species of hybrids:** Hybrids between both closely as well as distantly related parent species were required to determine whether there were differences in the efficiency of prediction in narrow and wide hybrids.
- **Appropriate genetic structure in family trials:** Full-sib or half-sib relatedness of individuals in family trials (for hybrid and pure species) and sufficient numbers of families were required for reliable heritability estimates and for the prediction of genetic and breeding values.

In addition to the above mentioned criteria the following properties were considered to be advantageous:

- **Similar environmental conditions:** Limited environmental differences between the hybrid family trial sites and the clonal trial sites would be advantageous as the effect of genotype-by-environment interaction would be minimized and thus reducing the error.
- **Similar assessment ages:** As trait expression varies with age, it would be preferable to compare trait assessments that were done at similar ages.
- **Appropriate controls:** The same genetic controls across trials would allow for the determination of genetic gains.

2.2.2 Statistical criteria

Statistical analyses form the basis of genetic analysis therefore the validity of the statistical analysis was also an important factor in determining the success of the study. In order to ensure the reliability of the statistical analysis, statistical criteria were defined which would assist in determining whether the data were suitable. The criteria are listed below:

- **Size of the trials:** Trials were required to be large enough to allow for meaningful analysis of the data, thus sufficient trees per family in the family trials and sufficient numbers of ramets per clone in the clonal trials were required for reliable estimates of performance.
- **Normality of the data and homogeneity of error variances:** Many statistical tests are based on the assumption that the error variances are distributed normally and that variances are equal (Snedecor & Cochran 1967).
- **Variability of data:** The analysis required statistically significant differences between the clones to draw meaningful conclusions about better or poorer performance.
- **Ideally a continuous trait** was required as such traits are better suited to the analysis of variance.

2.3 Description of data sets

The data that were selected for this study were obtained from the CSIR eucalypt hybrid breeding programme. This programme was initiated by the South African Forestry Research Institute (SAFRI) in 1982. The breeding programme has been managed by the CSIR since 1990. The first *Eucalyptus* hybrid family trials of 350 hybrids were planted in 1984. Initially crosses were made between *E. grandis* and species such as *E. camaldulensis, E. tereticornis* and *E. urophylla* in an attempt to improve *E. grandis* by adding drought tolerance and adaptability on shallow soils to the desirable growth and wood properties of *E. grandis*. Crosses between *E. grandis* and *E. nitens* were made to introduce cold-tolerance while crosses between *E. grandis* and *E. saligna* were made to obtain a red-coloured saw timber resource. In the late 1980s, early 1990s many double crosses and backcrosses were produced including crosses that did not use *E. grandis* as female parent. The programme was structured to allow for deployment of hybrid seed or hybrid clones. This was achieved by the planting of family hybrid trials as well as clonal hybrid trials which were based on the cutting material harvested from superior genotypes within the evaluated family trials.

2.3.1 Hybrid parent species

Three eucalypt hybrids were included in the study to represent an array of hybrid parent species combinations. The hybrids were *E. grandis* × *E. saligna, E. grandis* × *E. urophylla* and *E. grandis* × *E. camaldulensis*. The classification of the four parental species is presented in Table 2.1. From the classification it can be seen that the hybrids represent intra- and intersectional as well as intra- and interseries crosses. It is assumed that the classification reflects the genetic relatedness of parent species.

The distribution of the species may also be an indication of the degree to which populations were isolated. The natural distribution of *E. grandis* extends from northern Queensland (16°S) to New South Wales (32°S) in a relatively narrow band within 160 km from the coast (Jovanovic & Booth 2002). The natural distribution of *E. saligna* ranges from the south coast region of New South Wales to Maryborough in Queensland, and is mainly concentrated within 120 km from the coast (Jovanovic & Booth 2002).

The distribution of *E. grandis* and *E. saligna* therefore overlaps. *E. urophylla* has a totally distinct distribution as it originates from seven islands of eastern Indonesia (Pepe *et al.* 2004). *E. camaldulensis* has the widest distribution which covers most of Australia but it is not found in the coastal margins of south eastern Australia and Queensland (Jovanovic & Booth 2002).

Table 2.1 Classification of four parent species according to Brooker (2000).

Species	Series	Section	Subgenus
E. grandis	*Transversae*	*Transversaria*	*Symphyomyrtus*
E. saligna	*Transversae*	*Transversaria*	*Symphyomyrtus*
E. urophylla	*Annulares*	*Transversaria*	*Symphyomyrtus*
E. camaldulensis	*Rostratae*	*Exsertaria*	*Symphyomyrtus*

The hybrid *E. grandis* × *E. saligna* represents an intraseries, intrasectional cross. The two parent species are thus closely related and were once considered to be the same species. Despite the relatedness of the two parental species, a difference in the DNA content has been reported with 710 Mbp for *E. saligna* and 640 Mbp for *E. grandis* (Grattapaglia & Bradshaw 1994). The hybrid *E. grandis* × *E. urophylla* represents an interseries, intrasectional cross. At 650 Mbp, the size of the nuclear genome of *E. urophylla* is the closest to the 640 Mbp reported for *E. grandis* (Grattapaglia & Bradshaw 1994). The *E. grandis* × *E. camaldulensis* hybrid represents an interseries, intersectional cross. It has, however, been suggested that the hybrids between *Transversaria* and *Exsertaria* are ‘atypically successful’ (Potts & Dungey 2004). Steane *et al.* (2002) has questioned the separation of the sections based on molecular data but Poke *et al.* (2006) found more resolution between the species than previously reported by Steane *et al.* (2002). There is also a difference in the genome size between these two species. With a genome size of 590 Mbp, *E. camaldulensis* has a smaller nuclear genome than *E. grandis* (Grattapaglia & Bradshaw 1994).

It has been reported that the frequency of hybrid inviability is often high in wide crosses (Potts *et al.* 2003) and that hybrid inviability tends to increase with increasing

taxonomic distance (Potts & Dungey 2004). Therefore it may be expected that the presence of inviability and abnormalities in the hybrids will be more prevalent in the *E. grandis* × *E. camaldulensis* than in *E. grandis* × *E. saligna*. A relatively large proportion of weak or abnormal phenotypes in a hybrid population will contribute to the fact that the assumption of normality of distribution within the population will not be met (Griffin *et al.* 2000).

2.3.2 Trial descriptions of the hybrid trials

Three hybrid family trials, an *E. grandis* × *E. saligna* (G x S), an *E. grandis* × *E. urophylla* (G x U) and an *E. grandis* × *E. camaldulensis* (G x C) trial were selected together with their corresponding clonal trials. The trials are presented in Figure 2.2.

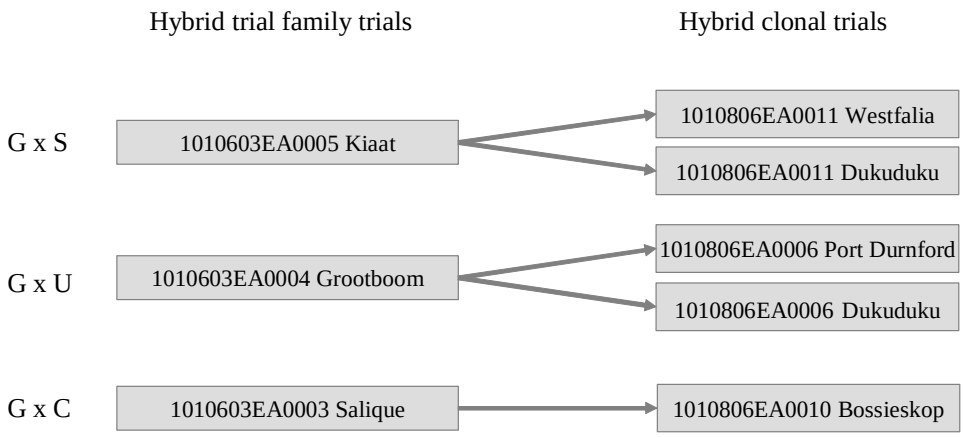


Figure 2.2 Sets of hybrid family and clonal trials that were available for the study.

The hybrids used in this study had been produced by means of controlled pollination between selected pure species parents. The hybrid seedlings were established in hybrid family trials for evaluation of family and individual tree performance. Within the hybrid family trials, the best performing trees were selected for the establishment of clonal trials. The selections were based on individual tree values and no family information was used. The trials were felled and coppice was collected from the selected trees

(ortets) for the propagation of macro-cuttings. The cuttings (ramets) were used for the establishment of clonal trials.

2.3.2.1 E. grandis × E. saligna (G x S) hybrid

The *E. grandis* × *E. saligna* hybrids originated from controlled crosses that were made using *E. saligna* pollen that was collected from a seed orchard at the De Hoek Forestry Research Station in the Limpopo province. The pollen was applied to flowers of *E. grandis* mother trees in the seed orchard at the J. D. M. Keet Forestry Research Station in the Limpopo province (Hoon 1992). A total of 36 *E. grandis* × *E. saligna* full-sib families were included in the hybrid family trial. These families resulted from crosses between 16 *E. grandis* mothers and five *E. saligna* pollen sources. From the mating design, which is presented in Table 2.2, it can be seen that in addition to the within-family full-sib relationship, there was also a half-sib relationship between some families in the trial (e.g. G4 x S2 and G4 x S4). The hybrid family trial was established at the Kiaat (Westfalia) plantation in the Limpopo province. The description of the trial site is presented in Table 2.3 and the trial description is presented in Table 2.4.

Table 2.2 Mating design of the G x S hybrid family trial.

♀ \ ♂		***E. saligna***				
		S1	**S2**	**S4**	**S5**	**S7**
E. grandis	**G4**		x	x		x
	G6					x
	G15			x	x	x
	G17		x	x	x	x
	G19		x	x	x	x
	G22		x	x		x
	G45		x	x		x
	G50		x		x	x
	G58			x		x
	G76	x				
	G77	x				
	G78	x				
	G87	x				
	G88	x				
	G92	x				
	G101		x	x	x	x

Table 2.3 Site description of the G x S hybrid family trial.

Trial site descriptors	Description of hybrid family trial site
Plantation name	Kiaat
Province	Limpopo
Closest town	Politsi
Latitude	23º46'S
Longitude	30º02'E
Altitude	900 m
Mean annual rainfall	1200 mm
Soil	Hutton

Table 2.4 Trial description of the G x S hybrid family trial.

Trial descriptors	Hybrid family trial description
Plantation	Kiaat
Trial number	1010603EA0005
Trial type	Hybrid family
Hybrid	36 G x S hybrids (F)
Design	RCB
Replications	6
Plot size	1 x 6
Spacing	2. 7 m x 2.7 m
Controls	2 G x T hybrids (F) 4 G x R hybrids (F) 2 G x C hybrids (F) 1 G x U hybrid (F) 2 *E. grandis* (F)
Date planted	04/1988
Traits assessed (age)	height, DBH, defects (2 years); height, DBH, stem, defects, disease (6 years)

F = family, G x T = *E. grandis* × *E. tereticornis*; G x R = *E. grandis* × *E. resinifera*

Clonal selections were made from the *E. grandis* × *E. saligna* hybrid trial at Kiaat. These selections were based on rankings of tree stem volumes (m^3) at 74 months and were followed by an independent culling for disease (Pierce 1994 b). One clonal trial was established at Westfalia in the Limpopo province, whereas the second clonal trial was established in KwaZulu-Natal. Site descriptions of the clonal trials are presented in Table 2.5 and the trial descriptions are presented in Table 2.6. The relatedness of the clones of the trial 1010806EA0011 at Westfalia and Dukuduku and the identification of the ortets are presented in Table 2.7. The 24 clones in the clonal trials consisted of full-sibs (e.g. GxS152 and GxS154), half-sibs (e.g. GxS156 and GxS177) and unrelated clones (e.g. GxS160 and GxS154).

Table 2.5 Site description of the G x S clonal trials.

Site descriptors	**Description of clonal trial site 1**	**Description of clonal trial site 2**
Plantation	Dukuduku	Westfalia
Province	KwaZulu-Natal	Limpopo
Closest town	Mtubatuba	Politsi
Latitude	28º13'S	23º44'S
Longitude	32º25'E	30º06'E
Altitude	40 m	950 m
Mean annual rainfall	950 mm	1342 mm
Soil	Sandy/Fernwood	Hutton

Table 2.6 Trial description of the G x S clonal trials.

Trial descriptors	Description of clonal trial 1	Description of clonal trial 2
Plantation	Dukuduku	Westfalia
Trial number	1010806EA0011	1010806EA0011
Trial type	Clonal	Clonal
Hybrid	24 G x S (C)	24 G x S (C)
Design	RCB	RCB
Replications	10	10
Plot size	Single tree	Single tree
Spacing	3 m x 3 m	3.5 m x 3.5 m
Controls	4 *E. grandis* (C) 2 seedling controls (F)	4 *E. grandis* (C) 2 seedling controls (F)
Date planted	08/1997	06/1997
Traits assessed (age)	height, DBH, stem, disease (3 years)	height, DBH, stem, disease (3 years)

C = clone, F = family

Table 2.7 Pedigree of the G x S clones in the clonal trials.

G x S Clone	***E. grandis* Female parent**	***E. saligna* Male parent**	**Family**	**Position of ortet in trial 1010603EA0005 Kiaat**		
				Replication	**Tree**	**Plot**
GXS174	G6	S7	4	3	6	136
GXS153	G15	S4	5	6	1	256
GXS149	G15	S5	6	4	2	158
GXS150	G15	S5	6	6	4	247
GXS152	G15	S7	7	3	4	122
GXS154	G15	S7	7	4	1	148
GXS143	G19	S2	11	2	6	57
GXS142	G19	S7	14	5	2	200
GXS164	G22	S7	17	1	1	19
GXS140	G45	S7	21	2	5	80
GXS144	G50	S2	22	2	2	63
GXS151	G50	S2	22	5	1	231
GXS136	G50	S5	24	2	4	86
GXS132	G50	S7	25	5	2	229
GXS158	G50	S7	25	6	3	240
GXS170	G58	S7	27	2	5	92
GXS171	G77	S1	29	2	4	72
GXS177	G77	S1	29	2	2	72
GXS156	G78	S1	30	5	1	198
GXS157	G78	S1	30	2	3	88
GXS160	G78	S1	30	6	6	272
GXS165	G78	S1	30	5	4	198
GXS168	G78	S1	30	6	1	272
GXS159	G88	S1	32	1	4	43

The locations of the G x S hybrid family and clonal trials are indicated in Figure 2.3. The hybrid trial and one of the clonal trials were established at the same plantation (Kiaat/Westfalia) in the Limpopo Lowveld, whereas the other clonal trial was established in the coastal plains of KwaZulu-Natal. This site has lower rainfall and less fertile soils than the site in the northern Limpopo province and it has higher incidence of disease infestation.

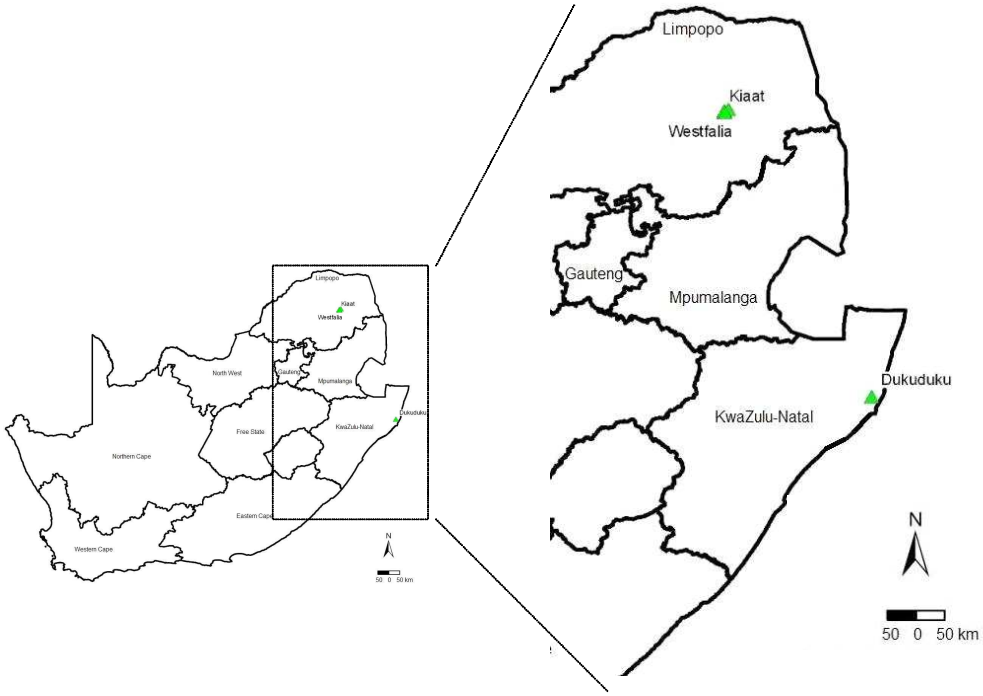


Figure 2.3 Location of the G x S hybrid family and clonal trials.

2.3.2.2 E. grandis × E. urophylla (G x U) hybrid

The *E. grandis* × *E. urophylla* hybrids originated from controlled crosses that were made using *E. urophylla* pollen that was collected from the provenance trials 1010302EA0504 at Frankfort in Mpumalanga and 1010302EA0505 at KwaMbonambi in KwaZulu-Natal. The pollen was applied to flowers of *E. grandis* mother trees in the seed orchard at the J. D. M. Keet Forestry Research Station. The mating design of the trial is presented in Table 2.8. A total of 69 full-sib families were included in the G x U hybrid progeny trial. The families resulted from crosses between 12 *E. grandis* mothers and 10 *E. urophylla* pollen sources. From the mating design it can be seen that, in addition to the within-family full-sib relationships, there were half-sib relationships between some families (e.g. G4 x U1 and G4 x U5) in the trial. The hybrid family trial was established at the Grootboom plantation in the Limpopo province. The trial site is described in Table 2.9 and the trial description is presented in Table 2.10.

Table 2.8 Mating design of the G x U hybrid family trial.

♀ \ ♂		*E. urophylla*									
		U1	**U2**	**U4**	**U5**	**U8**	**U9**	**U11**	**U3-2**	**U3-7**	**24195**
E. grandis	**G4**	x			x			x	x	x	
	G15		x	x		x	x	x		x	x
	G17	x			x						
	G19	x	x		x	x	x	x			
	G35	x	x	x	x	x	x	x			
	G36	x	x	x	x	x	x	x			
	G37		x				x	x		x	
	G38	x	x	x	x	x	x	x		x	
	G39	x	x			x	x	x			x
	G45	x	x		x	x	x	x	x	x	
	G58	x			x						
	G101	x	x	x	x	x	x	x			

Table 2.9 Site description of G x U hybrid family trial.

Site descriptors	**Hybrid family trial description**
Plantation name	Grootboom
Province	Limpopo
Closest town	Politsi
Latitude	23º43'S
Longitude	30º10'E
Altitude	750 m
Mean annual rainfall	1300 m
Soil	Hutton

Table 2.10 Trial description of G x U hybrid family trial.

Trial descriptors	Description of hybrid family trial
Plantation	Grootboom
Trial number	1010603EA0004
Trial type	Hybrid family
Hybrid	69 G x U (F)
Design	RCB
Replications	5
Plot size	1 x 6
Spacing	2. 7 m x 2.7 m
Controls	9 G x T (C) 9 G x C (C) 3 seedling controls (F)
Date planted	02/1986
Traits assessed (age)	height, defects (6 months) height, DBH, stem, defects (2 years); height, DBH, stem, defects, disease (5 years)

C = clone, F = family, G x T = *E. grandis* × *E. tereticornis*

Clonal selections were made from the *E. grandis* × *E. urophylla* hybrid family trial at Grootboom. These selections were based on rankings of volume at 54 months of age (Pierce 1993). The cuttings were planted in clonal trials at Dukuduku and Port Durnford. The trial sites of the two clonal trials are described in Table 2.11 and the trial descriptions are presented in Table 2.12. The relatedness of the clones of the clonal trial 1010806EA0006 at Port Durnford and Dukuduku is indicated in the pedigree presented in Table 2.13. The 65 clones consisted of full-sibs (e.g. GxU69 and GxU88), half-sibs (e.g. GxU88 and GxU96) and unrelated clones (e.g. GxU88 and GxU116).

Table 2.11 Site description of the G x U clonal trials.

Site descriptors	Description of clonal trial site 1	Description of clonal trial site 2
Plantation	Port Durnford	Dukuduku
Province	KwaZulu-Natal	KwaZulu-Natal
Closest town	Empangeni	Mtubatuba
Latitude	28º53'S	28º13'S
Longitude	32º50'E	32º15'E
Altitude	130 m	70 m
Mean annual rainfall	1461 mm	973 mm
Soil	Hutton	Sandy/Fernwood

Table 2.12 Trial description of the G x U clonal trials.

Trial descriptors	Description of clonal trial 1	Description of clonal trial 2
Plantation	Port Durnford	Dukuduku
Trial number	1010806EA0006	1010806EA0006
Trial type	Clonal	Clonal
Hybrid	65 G x U (C)	65 G x U (C)
Design	RCB	RCB
Replications	15	15
Plot size	Single tree	Single tree
Spacing	3.5 m x 3.5 m	2.7 m x 2.7 m
Controls	1 G x T (C) 1 G x C (C) 1 *E. grandis* (F)	1 G x T (C) 3 G x C (C) 1 *E. grandis* (F)
Date planted	09/1993	08/1993
Traits assessed (age)	height, disease (2 years); height, DBH, stem, defects, disease (4 years)	height, disease (2 years); height, DBH, stem, defects, disease (4 years)

C = clone, F = family, G x T = *E. grandis* × *E. tereticornis*

Table 2.13 Pedigree of the G x U clones in the clonal trials.

G x U Clone	*E. grandis* Female parent	*E. urophylla* Male parent	Family	Position of ortet in trial 1010603EA0004 Grootboom		
				Replication	Tree	Plot
GXU53	G4	U1	1	1	1	22
GXU96	G4	U1	1	4	6	272
GXU60	G4	U11	2	1	3	61
GXU118	G4	U11	2	5	6	404
GXU63	G4	U5	3	1	6	75
GXU69	G4	U11	4	2	1	124
GXU88	G4	U11	4	3	1	235
GXU102	G4	U3-2	5	4	1	314
GXU116	G15	U3-7	6	5	5	402
GXU117	G15	U3-7	6	5	6	402
GXU111	G17	U1	7	5	5	371
GXU52	G19	U11	8	1	4	11
GXU86	G19	U9	8	3	4	226
GXU87	G19	U9	8	3	2	226
GXU58	G19	U9	8	1	4	57
GXU72	G19	U1	9	2	6	130
GXU73	G19	U1	9	2	4	130
GXU103	G19	U1	9	4	2	319
GXU89	G19	U5	10	3	3	251
GXU99	G19	U8	11	4	3	295
GXU107	G19	U2	12	4	1	352
GXU51	G35	U2	13	1	5	8
GXU66	G35	U8	14	1	1	90
GXU82	G35	U1	15	3	2	195
GXU95	G35	U1	15	4	3	271
GXU71	G36	U1	15	2	6	129
GXU92	G35	U11	16	3	4	260
GXU61	G36	U11	17	1	3	64
GXU62	G36	U11	17	1	4	64
GXU84	G36	U9	18	3	4	203
GXU110	G36	U9	18	5	2	369
GXU115	G36	U4	19	5	1	401
GXU109	G37	U9	19	4	2	359
GXU50	G37	U9	20	1	4	3
GXU83	G37	U3-7	21	3	1	202
GXU80	G38	U1	22	3	4	181
GXU108	G38	U3	23	4	3	353
GXU54	G39	U11	24	1	1	30
GXU97	G39	U11	24	4	5	289
GXU59	G39	U9	25	1	4	58
GXU105	G39	U9	25	4	3	348
GXU113	G39	U9	25	5	3	398
GXU114	G39	U9	25	5	2	398
GXU81	G39	U2	26	3	2	188
GXU94	G39	U1	27	3	4	269
GXU112	G39	U1	28	5	4	374
GXU57	G39	U24195	29	1	5	47
GXU56	G45	U2	30	1	5	44

G x U Clone	*E. grandis* Female parent	*E. urophylla* Male parent	Family	Position of ortet in trial 1010603EA0004 Grootboom		
				Replication	Tree	Plot
GXU64	G45	U1	31	1	1	89
GXU65	G45	U1	31	1	5	89
GXU77	G45	U1	31	2	3	138
GXU78	G45	U1	31	2	2	138
GXU100	G45	U1	31	4	6	299
GXU67	G45	U9	32	2	1	99
GXU68	G45	U9	32	2	1	119
GXU90	G45	U9	32	3	3	254
GXU91	G45	U9	32	3	5	254
GXU101	G45	U9	32	4	3	303
GXU70	G45	U3-7	33	2	6	125
GXU119	G45	U3-7	33	5	1	437
GXU93	G45	U11	34	3	6	263
GXU55	G58	U1	35	1	1	32
GXU104	G58	U1	35	4	5	320
GXU85	G101	U1	36	3	4	225
GXU79	G101	U2	37	2	2	151

The G x U hybrid family trial was located in the Limpopo province whereas both clonal trials were established in the coastal plain of KwaZulu-Natal (see Figure 2.4). The sites of the clonal trials were chosen for optimal evaluation of disease resistance, as disease infestations are more prevalent in these coastal plantations.

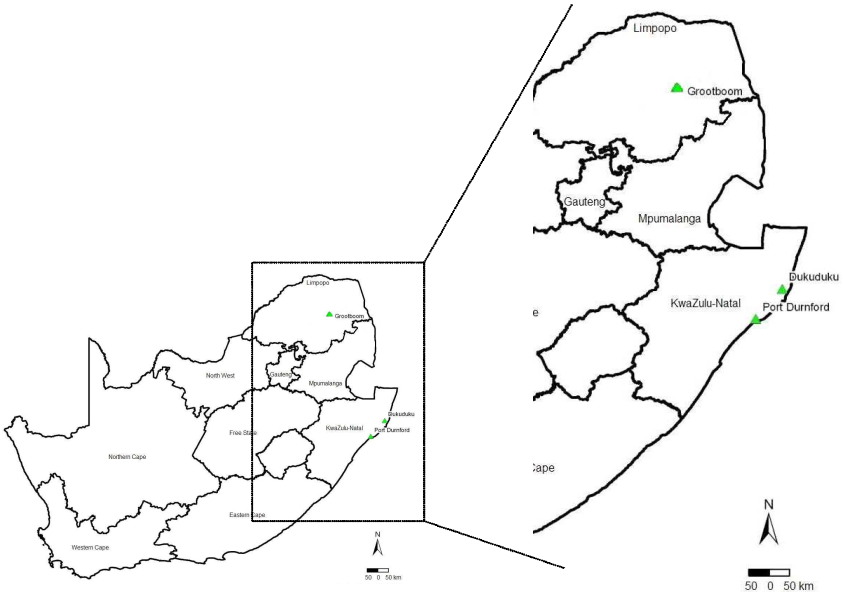


Figure 2.4 Location of the G x U hybrid and clonal trials.

2.3.2.3 E. grandis × E. camaldulensis (G x C) hybrid

The *E. grandis* × *E. camaldulensis* hybrids originated from controlled crosses that were made using *E. camaldulensis* pollen that was collected from the provenance trials 1010302EA2401 at Roodewal (Limpopo province) and 1010302EA2407 at Frankfort (Mpumalanga province). The pollen was applied to flowers of *E. grandis* mother trees in the seed orchard at the J. D. M. Keet Forestry Research Station. The mating design is presented in Table 2.14. A total of 34 full-sib families were included in the trial. These families resulted from crosses between 10 *E. grandis* mothers and seven *E. camaldulensis* pollen sources. In addition to the full-sib relationship within families there was also half-sib relatedness of some families (e.g. G4 x C22694 and G4 x C24223). The hybrid family trial was established at the Salique plantation in the Limpopo province. A description of the site is presented in Table 2.15, whereas the trial description is presented in Table 2.16.

Table 2.14 Mating design of the G x C hybrid trial.

♀ \ ♂		*E. camaldulensis*						
		C22694	**C24216**	**C24223**	**C24225**	**C24226**	**C24229**	**C24230**
E. grandis	**G4**	x		x				
	G17			x		x	x	x
	G19	x		x	x	x		
	G37	x						x
	G38	x	x		x	x		x
	G39					x	x	x
	G45							x
	G50	x		x		x		x
	G58					x	x	x
	G101		x	x	x	x	x	x

Table 2.15 Site description of the G x C hybrid family trial.

Site descriptors	**Hybrid family trial site description**
Plantation name	Salique
Province	Limpopo
Closest town	Klaserie
Latitude	24º45'S
Longitude	30º58'E
Altitude	670 m
Mean annual rainfall	750 mm
Soil	Sandy

Table 2.16 Trial description of the G x C hybrid family trial.

Trial descriptors	Hybrid family trial description
Plantation	Salique
Trial number	1010603EA0003
Trial type	Hybrid family
Hybrid	32 G x C (F)
Design	RCB
Replications	5
Plot size	1 x 6
Spacing	2. 7 m x 2.7 m
Controls	13 G x T + 2 incomplete (F) 1 *E. grandis* (F)
Date planted	03/1985
Traits assessed (age)	height (6 months); height, DBH, defects (1 year) height, DBH, stem, defects (4 years); height, DBH, stem, defects (6 years); DBH, stem, defects (8 years)

F = family, G x T = *E. grandis* × *E. tereticornis*

Clonal selections were made from the *E. grandis* × *E. camaldulensis* hybrid trial at Salique. These selections were based on the 100 month assessments of diameter at breast height (Pierce 1994 a). The cuttings were planted in a clonal trial at Bossieskop in Mpumalanga and at Shallowdrift in KwaZulu-Natal. The trial at Shallowdrift has not been assessed. The trial site at Bossieskop is described in and the trial description is presented in Table 2.18. The relatedness of the clones in the clonal trial 1010806EA0010 at Bossieskop is indicated in Table 2.19. From this table it can be seen that the 37 clones consisted of full-sibs (e.g. GxC250 and GxC265), half-sibs (e.g. GxC264 and GxC245) and unrelated clones (e.g. GxC230 and GxC228).

Table 2.17 Site description of the G x C clonal trial.

Site descriptors	Hybrid family trial description
Plantation	Bossieskop
Province	Mpumalanga
Closest town	Nelspruit
Latitude	25º24'S
Longitude	30º30'E
Altitude	802 m
Mean annual rainfall	850 mm
Soil	Hutton

Table 2.18 Trial description of the G x C clonal trial.

Trial descriptors	Clonal trial description
Plantation	Bossieskop
Trial number	1010806EA0010
Trial type	Clonal
Hybrid	37 G x C (C)
Design	RCB
Replications	6
Plot size	1 x 5
Spacing	2.7 m x 2.7 m
Controls	1 control (F)
Date planted	11/1997
Traits assessed (age)	height, DBH, stem, defects, disease (4 years)

C = clone, F = family

Table 2.19 Pedigree of the G x C clones in the clonal trial.

G x C Clone	*E. grandis* Female parent	*E. camaldulensis* Male parent	Family	Position of ortet in trial 1010603EA0003 Salique		
				Replication	Tree	Plot
GXC230	G4	C24223	424223	2	5	120
GXC228	G17	C24226	1724226	3	6	140
GXC231	G17	C24226	1724226	2	2	67
GXC261	G17	C24226	1724226	4	4	203
GXC264	G17	C24226	1724226	2	6	67
GXC245	G17	C24230	1724230	1	6	52
GXC250	G17	C24230	1724230	3	4	125
GXC265	G17	C24230	1724230	5	1	269
GXC253	G19	C24223	1924223	4	4	226
GXC263	G19	C24223	1924223	4	1	226
GXC211	G19	C24226	1924226	3	1	132
GXC235	G19	C24226	1924226	5	6	289
GXC238	G37	C24230	3724230	3	6	174
GXC232	G38	C24226	3824226	5	4	258
GXC214	G39	C24226	3924226	3	4	168
GXC215	G39	C24226	3924226	3	6	168
GXC218	G39	C24226	3924226	1	6	9
GXC221	G39	C24226	3924226	1	3	9
GXC237	G39	C24226	3924226	4	4	196
GXC246	G39	C24226	3924226	2	1	95
GXC266	G39	C24226	3924226	4	3	223
GXC234	G39	C24229	3924229	3	5	175
GXC272	G39	C24229	3924229	2	2	201
GXC259	G39	C24230	3924230	5	5	280
GXC268	G39	C24230	3924230	3	3	149
GXC256	G45	C24230	4524230	4	3	240
GXC233	G50	C24223	5024223	5	1	267
GXC257	G50	C24226	5024226	4	3	235
GXC300	G50	C24226	5024226	1	2	10
GXC213	G58	C24226	5824226	3	1	128
GXC217	G58	C24226	5824226	3	3	128
GXC225	G58	C24226	5824226	2	1	62
GXC247	G58	C24226	5824226	2	4	62
GXC224	G58	C24229	5824229	5	2	246
GXC226	G58	C24230	5824230	5	1	242
GXC254	G101	C24216	10124216	1	6	5
GXC262	G101	C24225	10124225	5	3	286

The locations of the G x C hybrid and clonal trial are indicated in Figure 2.5. Both trials were located in the Mpumalanga province but the hybrid trial was established at a drier site and on a less fertile, sandy soil.

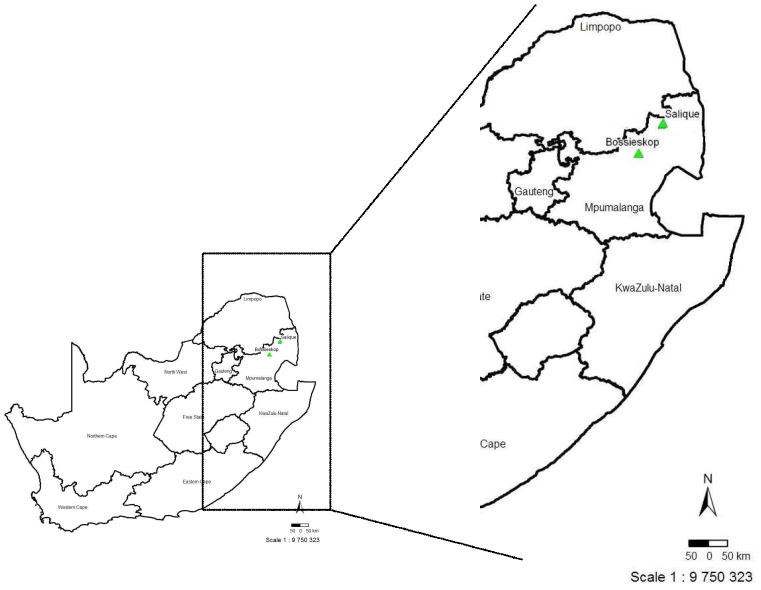


Figure 2.5 Location of the G x C hybrid family and clonal trials.

2.3.3 Trial descriptions of the pure species trials

E. grandis, mother parent in all three hybrids, was chosen as pure species control to compare differences in the efficiency of selection in hybrids with pure species. Three sets of *E. grandis* family trials and clonal trials were chosen. The three sets are presented in see Figure 2.6.

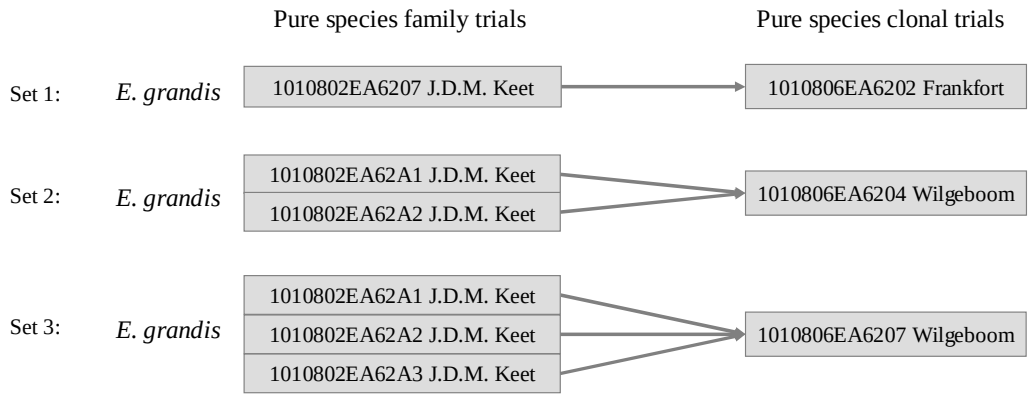


Figure 2.6 Sets of *E. grandis* family and clonal trials that were available for the study.

2.3.3.1 Set 1

The first set of *E. grandis* family and clonal trials consisted of the family trial 1010802EA6207 at J. D. M. Keet and the clonal trial 1010806EA6202 at Frankfort. The F_1 family trial was established from controlled crosses that were made in the seed orchard at J. D. M. Keet Forestry Research Station. The mating design is presented in Table 2.20. The 71 families in the trial consisted of full-sib families, some of which were reciprocal crosses (e.g. G4 x G6 and G6 x G4). Apart from the within-family full-sib relationship, there was also a half-sib relationship between some families (e.g. G4 x G6 and G4 x G15). The trial site of the family trial is described in Table 2.21 and the trial description in Table 2.22.

Table 2.20 Mating design of the *E. grandis* family trial 1010802EA6207.

♀ \ ♂	**4**	**6**	**10**	**15**	**17**	**19**	**22**	**35**	**36**	**37**	**38**	**39**	**44**	**45**	**50**	**58**	**60**	**64**	**66**	**101**
4		x		x								x				x				
6	x					x														
10		x							x							x	x	x	x	x
15			x									x	x				x		x	
17													x			x		x		
19																				
22										x						x	x		x	
35	x			x						x										
36		x				x				x	x	x								
37	x												x							
38																				x
44	x	x		x			x	x		x							x			
45				x			x						x			x		x		x
50	x		x			x	x						x							x
58	x				x															
60						x														
64		x										x				x	x			
66						x						x								
101	x				x	x				x	x			x						

Table 2.21 Site description of the *E. grandis* family trial 1010802EA6207.

Site descriptors	**Family trial description**
Plantation name	J. D. M. Keet
Province	Limpopo
Closest town	Politsi
Latitude	23º47'S
Longitude	30º07'E
Altitude	750 m
Mean annual rainfall	1300 mm
Soil	Hutton

Table 2.22 Trial description of the *E. grandis* family trial 1010802EA6207.

Trial descriptors	Family trial description
Plantation	J. D. M. Keet
Trial number	1010802EA6207
Trial type	Progeny trial
E. grandis families	71 *E. grandis* (F)
Design	8 x 9 Lattice
Replications	3
Plot size	1 x 10
Spacing	2. 7 m x 2.7 m
Controls	1 control (F)
Date planted	03/1976
Traits assessed (age)	height, DBH, stem, crown (5 months) height, DBH, stem (4 years) DBH (5 years) height, DBH, stem, crown (8 years)

F = family

Selections were made in the family trial 1010802EA6207 for the establishment of the clonal trial 1010806EA6202 at the Frankfort plantation. The site description of the clonal trial is presented in Table 2.23 and the trial description in Table 2.24. The pedigree of the clones is presented in Table 2.25. The clones consisted of full-sibs (e.g. SGR429 and SGR428), half-sibs (e.g. SGR451 and SGR452) and unrelated clones (e.g. SGR 411 and SGR417).

Table 2.23 Site description of the *E. grandis* clonal trial 1010806EA6202.

Site descriptors	Clonal trial site description
Plantation	Frankfort
Province	Mpumalanga
Closest town	Sabie
Latitude	25º02'S
Longitude	30º55'E
Altitude	960 m
Mean annual rainfall	1616 mm
Soil	Hutton

Table 2.24 Trial description of the *E. grandis* clonal trial 1010806EA6202.

Trial descriptors	Clonal trial description
Plantation	Frankfort
Trial number	1010806EA6202
Trial type	Clonal
E. grandis clones	22 (C)
Design	9 x 9 Lattice
Replications	4
Plot size	4 x 4
Spacing	2.5 m x 2.5 m
Controls	20 clones from other trials 1 control (F)
Date planted	12/1986
Traits assessed (age)	Height, defects (2 years) Height, DBH (4 years) Height, DBH, stem, defects (5 years) Height, DBH, stem, defect, disease (7 years) Height, DBH (11 years)

C = clone, F = family

Table 2.25 Pedigree of the *E. grandis* clones in the clonal trial 1010806EA6202.

***E. grandis* clone**	**Female parent**	**Male parent**	**Family**	**Position of ortet in trial 1010802EA6207 J. D. M. Keet**		
				Replication	**Plot**	**Tree**
SGR404	G15	G6	1	1	37	2
SGR411	G15	G60	2	2	78	1
SGR417	G17	G44	3	3	166	2
SGR423	G22	G66	4	3	192	3
SGR429	G35	G37	5	2	132	4
SGR428	G35	G37	5	1	71	2
SGR435	G36	G39	6	1	29	3
SGR437	G36	G39	6	2	91	4
SGR438	G37	G4	7	1	17	1
SGR380	G4	G15	8	2	122	1
SGR390	G4	G58	9	3	185	2
SGR451	G44	G4	10	3	159	4
SGR448	G44	G4	10	2	82	2
SGR452	G44	G6	11	1	32	2
SGR467	G45	G58	12	3	203	4
SGR477	G50	G19	13	3	173	2
SGR474	G50	G19	14	1	46	4
SGR478	G50	G22	15	2	126	1
SGR479	G50	G22	15	3	201	2
SGR516	G50	G4	16	3	213	3
SGR470	G50	G4	16	1	48	2
SGR471	G50	G4	16	1	48	4
SGR472	G50	G4	16	2	93	2
SGR488	G58	G17	17	3	163	2
SGR483	G58	G4	18	2	118	3
SGR481	G58	G4	18	1	58	2
SGR394	G6	G4	19	1	62	3
SGR494	G66	G39	20	1	52	2
SGR495	G66	G39	20	1	52	3

The locations of the family trial and the clonal trial are indicated in Figure 2.6. The family trial was established in the Lowveld of the Limpopo province, whereas the clonal trial was established in the Lowveld of Mpumalanga.

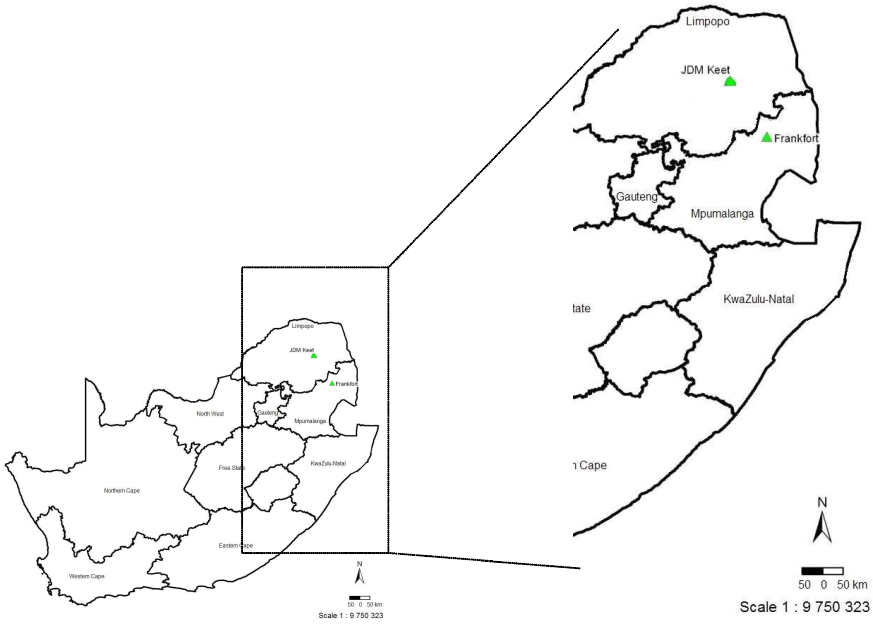


Figure 2.7 Location of the *E. grandis* family trial 1010802EA6207 and the clonal trial 1010806EA6202.

2.3.3.2 Set 2

The second set of *E. grandis* family and clonal trials consisted of two family trials and one clonal trial. The family trials were two trials from the CSIR's F_2 *E. grandis* breeding population. The trials 1010802EA62A1 and 1010802EA62A2 were planted at the J. D. M. Keet plantation (see site description in Table 2.21). The family trials at J. D. M. Keet consisted of half-sib families as they resulted from open pollination, unlike all other family trials used in this study. Selections from these two trials were established in the clonal trial 1010806EA6204 at the Wilgeboom plantation. Because of renumbering of trees after thinning and rogueing of the family trial, the link between the ramets in the clonal trial and the ortets in the family trial 1010802EA62A2 was lost. This trial could thus not be used for this study. Therefore only the trial 1010802EA62A1 will be discussed below. The trial description of this trial is presented in Table 2.26.

Table 2.26 Trial description of *E. grandis* family trial 1010802EA62A1.

Trial descriptors	Family trial description
Plantation	J. D. M. Keet
Trial number	1010802EA62A1
Trial type	Progeny trial
E. grandis families	60 *E. grandis* (F)
Design	8 x 8 Lattice
Replications	9
Plot size	2 x 2
Spacing	2. 7 m x 2.7 m
Controls	4 controls (F)
Date planted	03/1983
Traits assessed (age)	DBH, defect (3 years) height, DBH, stem, defect (4 years)

F = family

Clonal selections from 1010802EA62A1 were made in 1991. Coppice was collected after clearfelling of the trial in 1991 and cuttings were established together with selections from 1010802EA62A2 in the clonal trial 1010806EA6204 at Wilgeboom. Selections from 1010802EA62A1 were based on BLP predictions using information from four sources: parental information from trial 1010802EA6206 (a trial from the previous generation) at 81 months; volume and stem straightness family data from 1010802EA62A1 at 36 and 56 months; and individual tree and family data (94 months) from trial 1010802EA62A1 (Pierce 1992). The clonal trial included both pulp selections, which were selected for volume (70%) and density (30%), as well as selections for saw timber which were based on volume (45%), density (18%), splitting (27%) and stem straightness (10%). It is assumed that forward selection and backward selection for breeding values were used, as the clonal forward selection option was not available at that stage. Very hot weather (>40ºC) was experienced shortly after planting and plants were scorched. The trial file also notes some drought damage in the trial but there is no indication of the timing and the extent of the damage. The Wilgeboom

plantation site is described in Table 2.27 and the trial description in Table 2.28. The pedigree of the clones is presented in Table 2.29. The table shows that only replication and plot information was available for the identity of the ortet and no tree identity.

Table 2.27 Site description of the trial site of the trials 1010806EA6204 and 1010806EA6207.

Site descriptors	Clonal trial site description
Plantation	Wilgeboom
Province	Mpumalanga
Closest town	Graskop
Latitude	24º56'S
Longitude	30º57'E
Altitude	945 m
Mean annual rainfall	1348 mm
Soil	Hutton

Table 2.28 Trial description of *E. grandis* clonal trial 1010806EA6204.

Trial descriptors	Clonal trial description
Plantation	Wilgeboom
Trial number	1010806EA6204
Trial type	Clonal
E. grandis clones	91 (C) from 1010802EA62A1
Design	RCB
Replications	10
Plot size	Single tree
Spacing	3.5 m x 3.5 m
Controls	54 clones from 1010802EA62A2 (C) 2 of unknown origins (C) 2 control (F)
Date planted	02/1992
Traits assessed (age)	height, DBH, stem, defect (4 years) height, DBH, stem, defect, disease (6 years) height, DBH, stem, defect, disease (9 years) splitting (11 years)

C = clone, F = family

Table 2.29 Pedigree of the *E. grandis* clones in the clonal trial 1010806EA6204.

***E. grandis* clone**	**Female parent**	**Male parent**	**Position of ortet in trial 1010802EA62A1 J. D. M. Keet**	
			Replication	**Plot**
SGR1193	AG10	OP	5	277
SGR1202	AG10	OP	8	459
SGR1279	AG12	OP	1	58
SGR1243	AG13	OP	2	86
SGR1264	AG13	OP	6	383
SGR1207	AG15	OP	6	354
SGR1230	AG15	OP	8	480
SGR1284	AG15	OP	1	7
SGR1289	AG15	OP	9	534
SGR1198	AG17	OP	4	249
SGR1229	AG17	OP	8	481
SGR1244	AG17	OP	3	171
SGR1222	AG18	OP	8	449
SGR1227	AG18	OP	1	64
SGR1249	AG18	OP	3	142
SGR1195	AG19	OP	3	169
SGR1217	AG2	OP	7	396
SGR1271	AG2	OP	9	575
SGR1189	AG20	OP	6	329
SGR1197	AG20	OP	4	233
SGR1269	AG20	OP	9	555
SGR1273	AG20	OP	7	424
SGR1247	AG21	OP	2	111
SGR1187	AG23	OP	6	343
SGR1196	AG23	OP	2	70
SGR1206	AG23	OP	1	29
SGR1245	AG23	OP	3	178
SGR1190	AG25	OP	5	275
SGR1220	AG25	OP	3	144
SGR1221	AG25	OP	9	539
SGR1259	AG25	OP	4	200
SGR1263	AG25	OP	8	486
SGR1266	AG25	OP	7	413
SGR1272	AG25	OP	1	92
SGR1203	AG26	OP	8	466
SGR1218	AG26	OP	9	515
SGR1225	AG26	OP	1	42
SGR1253	AG26	OP	5	264
SGR1233	AG27	OP	6	350
SGR1211	AG29	OP	9	550
SGR1261	AG30	OP	7	434
SGR1191	AG31	OP	7	435
SGR1280	AG32	OP	1	26
SGR1204	AG33	OP	6	360
SGR1223	AG33	OP	8	454
SGR1228	AG33	OP	9	568
SGR1192	AG34	OP	7	448
SGR1205	AG35	OP	9	559

E. grandis clone	Female parent	Male parent	Position of ortet in trial 1010802EA62A1 J. D. M. Keet	
			Replication	Plot
SGR1268	AG35	OP	7	429
SGR1237	AG36	OP	8	507
SGR1255	AG36	OP	4	202
SGR1219	AG37	OP	5	263
SGR1252	AG37	OP	4	228
SGR1235	AG38	OP	3	150
SGR1267	AG38	OP	6	372
SGR1194	AG39	OP	5	284
SGR1274	AG39	OP	2	121
SGR1285	AG39	OP	1	16
SGR1209	AG40	OP	9	570
SGR1213	AG40	OP	8	504
SGR1265	AG40	OP	7	406
SGR1282	AG40	OP	2	115
SGR1240	AG42	OP	3	155
SGR1283	AG43	OP	1	28
SGR1238	AG44	OP	8	499
SGR1376	AG44	OP	6	321
SGR1199	AG45	OP	4	211
SGR1250	AG46	OP	4	212
SGR1188	AG47	OP	4	210
SGR1231	AG47	OP	5	271
SGR1234	AG47	OP	6	324
SGR1241	AG49	OP	3	175
SGR1212	AG50	OP	2	98
SGR1377	AG50	OP	3	141
SGR1288	AG51	OP	1	30
SGR1246	AG53	OP	5	270
SGR1256	AG53	OP	3	151
SGR1278	AG54	OP	2	71
SGR1210	AG55	OP	8	491
SGR1260	AG55	OP	6	382
SGR1287	AG55	OP	9	533
SGR1290	AG55	OP	3	134
SGR1375	AG55	OP	5	290
SGR1214	AG57	OP	9	516
SGR1277	AG59	OP	1	40
SGR1292	AG6	OP	3	161
SGR1200	AG60	OP	8	467
SGR1215	AG60	OP	2	83
SGR1258	AG60	OP	7	405
SGR1281	AG60	OP	1	5
SGR1275	AG9	OP	4	243

OP = open pollinated

The location of the family and clonal trials are presented in Figure 2.8. The family trial was established in the northern Limpopo province, whereas the clonal trial was established in the Mpumalanga Lowveld.

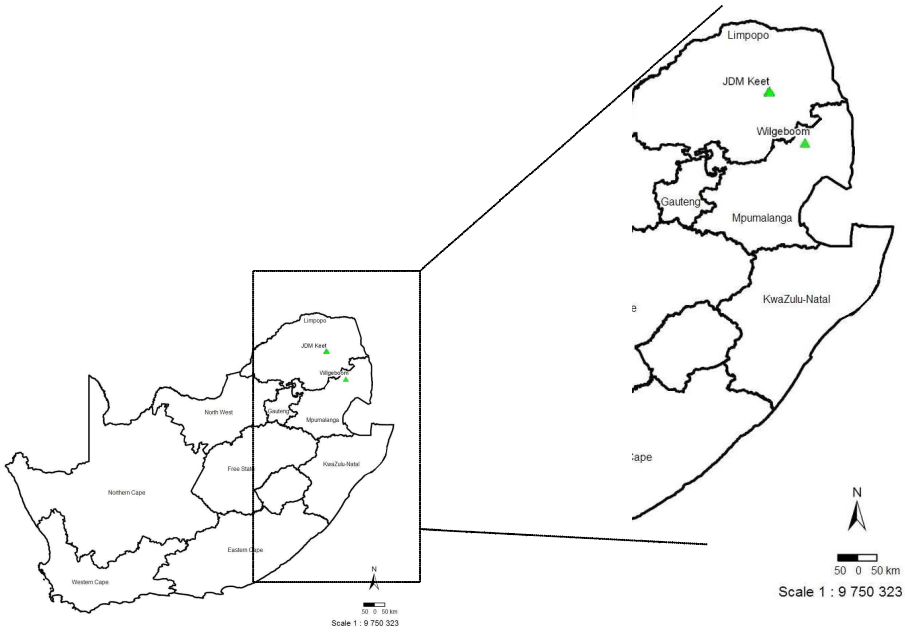


Figure 2.8 Location of *E. grandis* family trial 1010802EA62A1 and clonal trial 1010806EA6204 plantations.

The trial 1010802EA62A1 was allowed to re-coppice after felling and a second round of selections was made in 1996. These selections and together with selections from the trials 1010802EA62A2 and 1010802EA62A3 were used for the establishment of a second clonal trial at the Wilgeboom plantation. These sets of trials formed the third set of *E. grandis* trials.

2.3.3.3 Set 3

The third set of family and clonal trials consisted of the three F_2 family trials 1010802EA62A1, 1010802EA62A2 and 1010802EA62A3 and one clonal trial, the trial 1010806EA6207 at Wilgeboom. The three family trials which were established on the J.

D. M. Keet plantation consisted of half-sib families. As there was no link between the ramets in the clonal trial and the ortets in the trials 1010802EA62A2 and 1010802EA62A3 because of renumbering of trees after thinning, these two trials could not be used for the analysis. Information on the trial site is presented in Table 2.21 and the trial description of 1010802EA62A1 is presented in Table 2.26.

The clonal trial 1010806EA6207 that consisted of low wood splitting selections was established at Wilgeboom. The selections were based on BLP breeding values predictions using clonal forward selection. The analysis was based on the assumption that the total genetic variance consisted of 20% non-additive and 80% additive genetic variance (Verryn *et al.* 1997). The selection was based on 94 month-assessments of volume (40%), stem straightness (10%) and splitting (50%). The trial site has been described in Table 2.27 and the trial description is presented in Table 2.30. The pedigree of the trial is presented in Table 2.31. Twenty of the clones in the clonal trial EA6207 at Wilgeboom were also in the clonal trial EA6204 at Wilgeboom.

Table 2.30 Trial description of *E. grandis* clonal trial 1010806EA6207.

Trial descriptors	**Clonal trial description**
Plantation	Wilgeboom
Trial number	1010806EA6207
Trial type	Clonal
E. grandis clones	25 (C) from 1010802EA62A1
Design	RCB
Replications	6
Plot size	1 x 5 tree
Spacing	3 m x 3 m
Controls	19 from 1010802EA62A2 (C) 12 from 1010802EA62A3 (C) 4 from other trials (C) 1 control (F)
Date planted	12/1996
Traits assessed (age)	height, DBH, stem, defect, disease (4 years)

C = clone, F = family

Table 2.31 Pedigree of the *E. grandis* clones in the clonal trial 1010806EA6207.

***E. grandis* clone**	**Female parent**	**Male parent**	**Position of ortet in trial 1010802EA62A1 J. D. M. Keet**	
			Replication	**Plot**
SGR1187	AG23	OP	6	343
SGR1189	AG20	OP	6	329
SGR1192	AG34	OP	7	448
SGR1212	AG50	OP	2	98
SGR1213	AG40	OP	8	504
SGR1217	AG2	OP	7	396
SGR1228	AG33	OP	9	568
SGR1229	AG17	OP	8	481
SGR1233	AG27	OP	6	350
SGR1235	AG38	OP	3	150
SGR1243	AG13	OP	2	86
SGR1244	AG17	OP	3	171
SGR1249	AG18	OP	3	142
SGR1250	AG46	OP	4	212
SGR1256	AG53	OP	3	151
SGR1261	AG30	OP	7	434
SGR1264	AG13	OP	6	383
SGR1273	AG20	OP	7	424
SGR1274	AG39	OP	2	121
SGR1276	AG15	OP	5	318
SGR1277	AG59	OP	1	40
SGR1559	AG12	OP	4	253
SGR1560	AG59	OP	2	136
SGR1574	AG30	OP	9	560
SGR1576	AG30	OP	4	217

OP = open pollinated

2.3.4 Trait descriptions

Four traits were assessed the trials. Height was measured in metres, using either a height rod or a vertex hypsometer. Diameter at breast height (DBH) was measured in millimetres at a height of 1.3 m above the ground using a measuring tape. Stem straightness was scored on a subjective 8-point scale as indicated in Table 2.32. The scale was used as a fixed scale and was not adjusted for sites. Disease tolerance was scored on a subjective 5-point scale, where 0 represented no visual infestation and 4 represented chronic infestation. Trees showing stunted growth and trees with broken tops were not assessed. Defects were noted.

Table 2.32 Descriptors and scores used to assess stem straightness.

Score	Short description	Quality	Defects
8	Straight	Straight stem - pole quality	No defects
7	Nearly straight	Slight sweep and/or 1 minor bend	1 – 2 minor defects
6	Very slightly crooked	One slight sweep + >1 minor bend OR more than 1 slight sweep + 1 minor bend OR more than 2 minor bends	3 – 4 minor defects
5	Slightly crooked	Moderate sweep + 1 moderate bend OR two moderate sweeps + minor defect OR two moderate bends + minor defect	2 moderate defects OR 2 moderate + 1 minor
4	Moderately crooked	Moderate sweep + major bend OR more than two moderate sweeps OR more than two moderate bends OR two major bends + minor defects	1 moderate + 1 major OR 2 moderate OR 2 major + 2 minor
3	Crooked	Obvious sinuosity or major crooks	2 major OR 2 major + 2 moderate
2	Very crooked	Presence of multiple severe straightness defects	Several major + moderate defects
1	Malformed	Not merchantable as a short log (cork screw)	Major defects

CHAPTER 3
METHODS

3.1 Introduction

The data from three *Eucalyptus* hybrid and one *E. grandis* population were analysed to determine

- the correspondence of predicted and realised clonal performance in hybrid and pure species trials,
- the most successful method to predict clonal performance in the hybrid populations,
- the correspondence of prediction in hybrids between closely and distantly related parent species, and
- the correspondence of prediction in hybrids and pure species.

The methods used are presented in Figure 3.1. Data were first subjected to an exploratory analysis (step 1) to determine which trials were suitable for the study. The data were then edited. Genetic parameters were estimated for the family trials and clonal repeatability was determined in the clonal trials (step 2). Clonal performance was then predicted using different BLUP methods and constructed indices (step 3). Realised clonal performance was determined in the clonal trials (step 4). The predicted performance was then correlated with the realised performance in the clonal trials (step 5). Finally the prediction methods were compared (step 6).

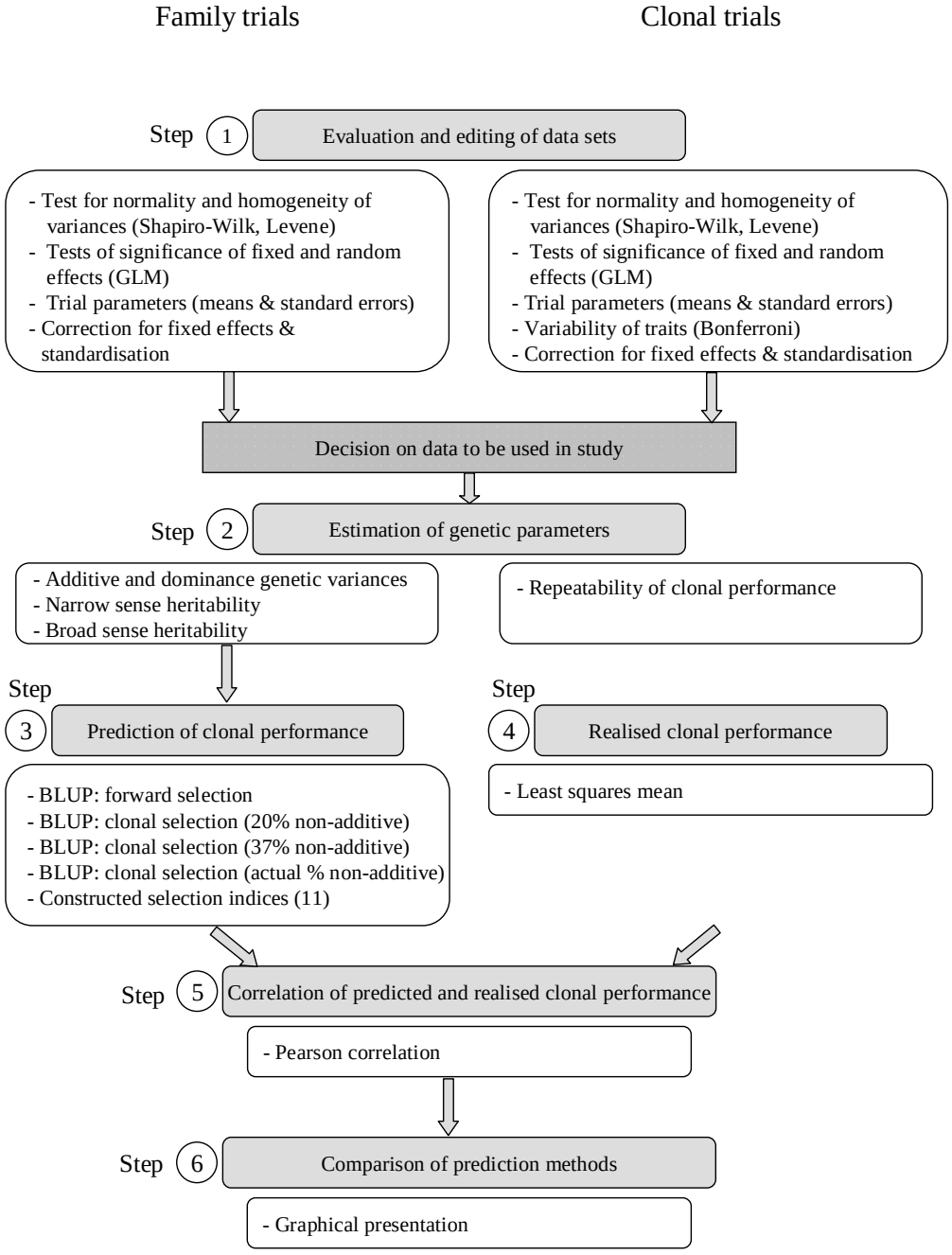


Figure 3.1 Methods used in the study.

3.2 Evaluation and editing of data sets (step 1)

The most suitable trait was first chosen for the study. Then all data sets were evaluated for normality of the distribution of errors, homogeneity of error variances and significance of effects. Data were edited before genetic analyses were performed. All statistical data analyses and data manipulations were performed with SAS/STAT software, of the SAS System for Windows (Copyright © 2002-2003 SAS Institute Inc.).

3.2.1 Trait selection

There were two criteria to establish whether the traits assessed were suitable for the study. The first criterion was that the trait should be assessed in both the family and the clonal trials. The traits height, diameter at breast height (DBH) and stem straightness met this criterion. The second criterion was that the traits needed to show sufficient variability. The variability has two aspects, firstly the variability of the ortets, and secondly the variability of the ramets. Testing of the variability in the ramets will be discussed in section 3.2.4. The variability of the traits in the ortets was expected to be dependent on the degree to which the trait had been subjected to selection historically and in the process of ortet selection.

The ortets represented only a selected subset of the family trials from which they were selected. In the hybrid trials this selection of ortets was based on the trait volume, which is based on height and DBH assessments. It was therefore assumed that the ortets would display a limited variation for these two traits and they would only represent the top-ranked individuals for these traits. Consequently there would not be a range of performances from which the best could be selected in this study. Stem straightness was, however, less strictly selected and a wider range of performances in the ortets could be expected. For this reason the study was focused on stem straightness.

3.2.2 Data evaluation

Analysis of variance (ANOVA) is based on three assumptions: normal distribution of the errors, homogeneity of error variances and independence of observations. The last assumption is difficult to verify and has to be inferred from the trial design. In this case the observations are assumed to be independent. The other two assumptions were tested by means of statistical tests.

The assumption that the errors were normally distributed was tested using a number of tests. Measures of skewness and kurtosis, the Shapiro-Wilk statistical test for normality (Shapiro & Wilk 1965) as well as normality plots (stem-leaf plot, box plot and normal probability plot) were investigated. The PROC UNIVARIATE procedure in SAS was used to perform these tests for normality. The Shapiro-Wilk test has been adapted in SAS to test for normality in sample sizes up to 2000 (Peng 2004). As sample sizes increase, smaller departures from normality can be detected by the Shapiro-Wilk test. This test may, therefore, detect small departures from normality which may not affect the validity of ANOVA tests. It has therefore been recommended to consider these tests together with other measures such as skewness and kurtosis measures and plots (SAS 1990). The SAS normality tests assess the distribution of the data rather than the distribution of the residuals. For this reason the residuals were first calculated using an option in PROC GLM in SAS.

The Levene test (Levene 1960) was used to test for the homogeneity of variances as the Bartlett test which is commonly used may be inaccurate in cases where the distribution is not normal (Box 1953).

An ANOVA procedure was performed to determine whether replications and families or clones differed significantly. The analyses were done using the PROC GLM procedure as it is most suitable for unbalanced data sets. Treatments were considered to be random effects and block and replication effects to be fixed effects. Interaction effects were not considered at this stage as the aim of the analysis was just to determine whether treatment effects were significant and whether there was a need to correct for

block and replication effects. The following model was used for trials with a randomized complete block design:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where

y_{ij} = the j^{th} replication of the i^{th} family or clone
μ = trial mean
α_i = the effect due to the family or clone
β_j = the effect due to replications
ε_{ij} = the random error associated with the j^{th} replication of i^{th} family or clone.

Where trials were designed as lattice or alpha lattice designs, the model contained an additional block variable:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{jk} + \varepsilon_{ijk}$$

where

γ_{jk} = the effect of the k^{th} block in the j^{th} replication

The trial means for stem straightness as well as the corresponding standard errors were determined for all trials by means of the PROC MEANS procedure in SAS.

3.2.3 Data editing

The data were edited before analysis. Missing values, controls and outliers were deleted, data were corrected for significant fixed effects and were standardised.

Data were corrected for replication effect. Replication was considered to be a fixed effect and was corrected for to reduce bias of selection from good performing replications. Correction was done by the following equation:

$$y_{ijk\ corrected} = y_{ijk} - (y_{j.} - \mu)$$

where

y_{ijk} = the kth tree of the j^{th} replication of the i^{th} family or clone

$y_{j.}$ = the mean of the j^{th} replication

μ = trial mean

Similarly, in trials where block effect was significant, a correction was made for block effect within replication and data were then analysed as randomized complete block designs. In cases where data of two sets of clonal trials were concatenated for analysis, the data were also corrected for trial effect. This, however, assumes no clone x site interaction.

All data were standardised by transforming the values to a distribution with a mean of zero and a standard deviation of one using PROC STANDARD in SAS. This standardisation allowed for the comparison of data that were measured at different ages because the standardised data sets were independent of the ranges of values. Another advantage of standardising data is that it becomes easy to compare relative rankings of individuals in a selection index.

3.2.4 Testing for variability in the clonal trials

It was expected that high selection pressure in the family trials would decrease the variability of the trait in the clonal trials. Variability was therefore assessed. The Bonferroni test was chosen as it is one of the few multiple comparison methods that can be applied in the case of unbalanced data. The BON option of the MEANS statement of the PROC GLM procedure was used. The number of statistically different groups was used as a measure of variability.

3.3 Estimation of genetic parameters (step 2)

Additive and dominance genetic variances were determined for all family trials that consisted of full-sibs. Narrow and broad-sense heritability were estimated for these family trials. Narrow-sense heritability was also estimated for the *E. grandis* trial that consisted of half-sibs and clonal repeatabilities were determined for the clonal trials. The genetic parameters were estimated using genetic theory that has been developed for large randomly mating populations and therefore effects such as the coefficient of the HWD were ignored (Wu *et al*. 2004).

The phenotypic variance of a trait can be partitioned into a genetic and an environmental component and interaction between these two components (Falconer 1981):

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2 + \sigma_{ge}^2$$

where

σ_p^2 = the phenotypic variance
σ_g^2 = the genetic variance component
σ_e^2 = the environmental variance component
σ_{ge}^2 = the variance attributed to the interaction between genotype and the environment

The genetic variance has an additive and a non-additive component (Falconer 1981):

$$\sigma_g^2 = \sigma_a^2 + \sigma_{na}^2$$

where

σ_a^2 = the additive genetic variance
σ_{na}^2 = the non-additive genetic variance

The non-additive genetic variance consists of dominance genetic variance (σ_d^2) and an epistatic interaction component which results from the interaction of different loci (Falconer 1981).

3.3.1 Estimation of additive and dominance genetic variance

The additive and dominance genetic variance may be estimated from full-sib family trials. In the absence of inbreeding the genetic variance effects may be estimated by (Becker 1992):

$$4\,\sigma_f^2 = 4\,cov_{HS} = \sigma_a^2 + \frac{1}{4}\sigma_{aa}^2 + \frac{1}{16}\sigma_{aaa}^2 + \ldots + 4\,\sigma_M^2$$

$$4\,\sigma_m^2 = 4\,cov_{HS} = \sigma_a^2 + \frac{1}{4}\sigma_{aa}^2 + \frac{1}{16}\sigma_{aaa}^2 + \ldots$$

$$2\left(\sigma_f^2 + \sigma_m^2\right) = 2\left(cov_{HS(f)} + cov_{HS(m)}\right) = \sigma_a^2 + \frac{1}{4}\sigma_{aa}^2 + \frac{1}{16}\sigma_{aaa}^2 + 2\sigma_M^2$$

$$4\sigma_{fm}^2 = 4\left[cov_{FS} - \left(cov_{HS(f)} + cov_{HS(m)}\right)\right] = \sigma_d^2 + \frac{1}{2}\sigma_{aa}^2 + \frac{1}{2}\sigma_{ad}^2 + \frac{1}{2}\sigma_{ad}^2 + \frac{1}{4}\sigma_{dd}^2 + \frac{3}{8}\sigma_{aaa}^2 + 4\sigma_M^2$$

where

σ_f^2	=	variance due to female parent effects
σ_m^2	=	variance due to male parent effects
σ_{fm}^2	=	variance due to interaction between male and female parents
σ_{aa}^2	=	epistatic genetic variance due to interaction of additive effects at two loci
σ_{aaa}^2	=	epistatic genetic variance due to interaction of additive effects at three loci
σ_{ad}^2	=	epistatic genetic variance due to interaction of additive and dominance effects at two loci
σ_{dd}^2	=	epistatic genetic variance due to interaction of dominance effects at two loci
σ_M^2	=	variance due to maternal effects
cov_{FS}	=	covariance of full-sibs
cov_{HS}	=	covariance of half-sibs with the subscript indicating male and female

If the epistatic variances and maternal variances are assumed to be negligible the additive genetic variance can be estimated as:

$$\sigma_a^2 = 2\left(\sigma_f^2 + \sigma_m^2\right)$$

and the dominance genetic variance as:

$$\sigma_d^2 = 4\sigma_{fm}^2$$

Significance of effects (female, male, female x male, female x replication, male x replication, fem x male x replication) was tested by means of the PROC GLM procedure. As the female x replication and male x replication effects were non-significant, these terms were excluded in the model. The estimation was based on the following model:

$$y_{ijk}=\mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Where

y_{ijk}	=	the observation of the k^{th} full-sib progeny in a plot of the i^{th} pollen parent and the j^{th} maternal parent
μ	=	trial mean
α_i	=	the effect of the i^{th} pollen parent
β_j	=	the effect of the j^{th} maternal parent
γ_k	=	fixed replication effect
$(\alpha\beta)_{ij}$	=	interaction of the paternal and maternal parents
ε_{ijk}	=	environmental effect and remainder of genetic effect between full-sibs on the same plot

Where there were significant female x male x replication effects, an additional term $(\alpha\beta\gamma)_{ijk}$ was included in the model to account for this effect.

In the case of the G x S hybrid family trial at Kiaat, where there were no significant female x male interactions, the term was not included in the model the model was reduced to:

$$y_{ijk}=\mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

The variance components ($\sigma_f^2, \sigma_m^2, \sigma_{fm}^2$ and σ_e^2) were estimated using the REML option in the PROC MIXED procedure in SAS. The standard errors of these variance components were estimated by the programme by specifying the COVTEST option.

The standard error of the additive genetic variance was then calculated as:

$$s.e.\left(\sigma_a^2\right) = 2x\left(s.e.\left(\sigma_f^2\right) + s.e.\left(\sigma_m^2\right)\right)$$

and the standard error of the dominance genetic variance as:

$$s.e.\left(\sigma_d^2\right) = 4x\left(s.e.\left(\sigma_{fm}^2\right)\right).$$

3.3.2 Narrow-sense heritability

Narrow-sense heritability is defined as the ratio of additive genetic variance to phenotypic variance (Falconer 1981):

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

For the family trials consisting of full-sib families estimates of narrow-sense heritability were based on the estimates of additive and dominance genetic variance. Under the assumption of negligible epistatic variance the narrow-sense heritability is estimated as:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2}$$

The standard error of the narrow-sense heritability estimate was calculated as:

$$s.e.(h^2) = \frac{s.e.(\sigma_a^2)}{\sigma_p^2}$$

For the family trial 1010802EA62A1 which consisting of half-sibs, the additive genetic variance was estimated from the family variance (σ_{fam}^2). Although the trial consisted of half-sibs, it was expected that there would be some degree of selfing and mating

between related individuals that would bias the estimate of the additive genetic variance components. It was, therefore, assumed that there was a 20% increased "relatedness" relative to open pollination (Verryn, 1993). A coefficient of relationship of 0.3 was therefore used for the estimation of additive genetic variance. If it is assumed that σ^2_{na} is negligible, the expression $3.33 \times \sigma^2_{fam}$ provides an upper bound estimate for σ^2_a and the heritability for full-sibs is estimated as:

$$h^2 = \frac{3.33 \times \sigma^2_{fam}}{\sigma^2_p}$$

The Mixed Model Least-Squares and Maximum Likelihood program (LSMLMW & MIXMDL PC-2 Version) developed by Harvey (1990) was used to determine the variance components for the estimation of narrow-sense heritability in the half-sib family trial. Model 2 of the program was used as there were no significant family-replication interactions.

3.3.3 Broad-sense heritability

Broad-sense heritability is defined as the ratio of the total genetic variance to the phenotypic variance (Falconer 1981):

$$H^2 = \frac{\sigma^2_g}{\sigma^2_p}$$

Under the assumption of negligible epistatic variance the genetic variance is estimated as:

$$\sigma^2_g = \sigma^2_a + \sigma^2_d$$

The estimates of additive and dominance genetic variance were used to estimate broad-sense heritability. The standard error was calculated as:

$$s.e.(H^2) = \frac{s.e.(\sigma_a^2) + s.e.(\sigma_d^2)}{\sigma_p^2}$$

3.3.4 Repeatability of clonal performance

Repeatability is a measure of the degree of agreement between repeated measurements of the same trait on the same individuals. In this case clonal repeatability was defined as the degree of agreement of different ramets of the same clone. The calculations of the clonal repeatability were based on the following statistical model:

$y_{km} = \mu + \alpha_k + \varepsilon_{km}$

where

y_{km} = the observation of the m^{th} ramet of the k^{th} clone
μ = the common mean
α_k = the effect of the k^{th} clone
ε_{km} = the deviation of the m^{th} ramet of a clone

Table 3.1 ANOVA on which estimation of clonal repeatability was based.

Source	Degrees of freedom	Sum of squares (SS)	Mean sum of squares (MS)	Expected mean sum of squares (EMS)
Among clones	n-1	SS_w	MS_w	$\sigma^2_e + k_1 \sigma^2_w$
Within clones between ramets	n(m-1)	SS_e	MS_e	σ^2_e

where

n = number of trees
m = number of ramets per clone
k_1 = m

The repeatability of clonal performance is given as (Becker 1992):

$$R = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_e^2}$$

where

R = repeatability
σ^2_w = variance between clones
σ^2_e = variance between ramets of the k^{th} clone

The variance components were estimated using the PROC VARCOMP procedure of SAS and the repeatabilities were calculated in Microsoft's Excel (2003).

The standard error of the repeatability estimates was calculated using (Becker 1992):

$$s.e.(R) \cong \sqrt{\frac{2(m.-1)(1-R)^2[1+(k_1-1)R]^2}{k_1^2(m.-n)(n-1)}}$$

where

m. = total number of measurements
k_1 = number of ramets per clone
n = number of clones

3.4 Prediction of clonal performance (step 3)

Two different BLUP selection techniques were used to determine whether these methods were efficient for selecting clones from hybrid trials. In the first approach, breeding values (BV) were predicted. With this method only the additive genetic potential of each individual is estimated by combining both parental and within family breeding value components (Xiang & Li 2001). In the second approach the genetic values (GV) were predicted. This method includes both additive and non-additive

genetic values and accounts for the total genetic effects (Xiang & Li 2003). In addition to the BLUP selection techniques, a series of selection indices were constructed which placed different weights on individual versus family performance.

3.4.1 Choice of programme for BLUP

Two programmes, Matgen and ASReml, were considered for the estimation of breeding and genetic values. ASReml is a programme that was developed by the Biometrics Programme of NSW Agriculture and the Biomathematics Unit of Rothamsted Experiment station (Gilmour *et al.* 2002). This programme has been developed for the analysis of balanced and unbalanced data from agricultural, medical and environmental sciences and is used by a number of tree improvement organisations world-wide (Gilmour *et al.* 2002). ASReml fits linear mixed models using the Residual Maximum Likelihood method. The advantage of this package is that the breeding values are estimated using information from both half-sibs and full-sibs. It would be suited to deal with backward selection of clones. However, for the forward selection of clones the individual tree model would need to be expanded to include the dominance genetic effects and a dominance relationship matrix would need to be constructed that contains the probabilities that the pairs of alleles in each of the two hybrids are identical by descent (Kain, personal communication).

Matgen (Verryn & Geerthsen 2007) is a programme that has been developed by the CSIR. The programme is a multi-trait, multi-site best linear prediction package which has been developed for unbalanced index selection in tree breeding. It has the capacity to deal with forward and backward selection as well as clonal selection. When using least square means for the correction of fixed effects, the predictions that are obtained are considered to be BLUP solutions (White & Hodge 1989). The disadvantage of the programme is that it does not use information of half-sibs when predicting full-sib performance.

It was decided to use Matgen for the prediction of breeding values as it offered several advantages. The main advantage of Matgen is that it can deal with non-additive variance

estimates. Furthermore, this programme is routinely used for the prediction of breeding and genetic values by researchers of the CSIR. Another consideration in the choice of programmes was that it is hoped that this study will in future lead to improvements in the Matgen programme to enhance its capability to deal with hybrid populations.

3.4.2 Estimation of breeding values

The data from the family trials that were corrected for fixed effects and standardised were exported from SAS and saved as a database (.dbf) file which is the format required by Matgen. BLUP predictions of breeding values were obtained using Matgen version 7.0 (Verryn & Geerthsen 2007). The forward selection option of Matgen was used to obtain the breeding values (BV). For the full-sib families a coefficient of relationship of 0.5 was used and for the half-sib families a coefficient of 0.3. Breeding value predictions were saved in a database format.

3.4.3 Estimation of genetic values

The clonal forward selection option of the programme Matgen was used to obtain genetic values (GV). The programme Matgen includes the non-additive genetic variance in predicting clonal performance. It was initially decided to use an estimate of non-additive genetic variance (σ^2_{na}/σ^2_g) in the populations. As epistatic genetic variances have not been estimated, non-additive genetic variances were substituted with dominance genetic variances. In this dissertation the proportion of non-additive genetic variances will therefore refer to the proportion of the total genetic variance that is attributed to dominance genetic variance. Estimates of 37% (GV_{37}) and 20% (GV_{20}) dominance genetic variance were used for the clonal forward predictions. The first is an estimate obtained for a *Eucalyptus grandis* cloned breeding population (Snedden 2001). The second is an approximation of an estimate of the non-additive genetic variation that was reported for volume and DBH in an *E. grandis* × *E. urophylla* hybrid population at three years of age (Bouvet & Vigneron 1996). Estimates of dominance genetic variance that were obtained from analysis of the family trials were also used.

3.4.4 Prediction using constructed selection indices

A series of combined indices were designed which placed different weights on family means versus single tree performance. Eleven indices were constructed. The indices were calculated using the following formulas:

$I_{0\%} = (1.0 \times y_{.j}) + (0 \times y_{ij})$
$I_{10\%} = (0.9 \times y_{.j}) + (0.1 \times y_{ij})$
$I_{20\%} = (0.8 \times y_{.j}) + (0.2 \times y_{ij})$
$I_{30\%} = (0.7 \times y_{.j}) + (0.3 \times y_{ij})$.
$I_{40\%} = (0.6 \times y_{.j}) + (0.4 \times y_{ij})$
$I_{50\%} = (0.5 \times y_{.j}) + (0.5 \times y_{ij})$
$I_{60\%} = (0.4 \times y_{.j}) + (0.6 \times y_{ij})$
$I_{70\%} = (0.3 \times y_{.j}) + (0.7 \times y_{ij})$.
$I_{80\%} = (0.2 \times y_{.j}) + (0.8 \times y_{ij})$
$I_{90\%} = (0.1 \times y_{.j}) + (0.9 \times y_{ij})$
$I_{100\%} = (0 \times y_{.j}) + (1.0 \times y_{ij})$

where

y_{ij} = the trait value of the i^{th} individual in the j^{th} family
$y_{.j}$ = the family mean (least squares mean) of the j^{th} family

The subscript of the index label indicates the percentage weight on individual tree performance. The index $I_{10\%}$, for example, places 90% weight on the family mean and 10% weight on the individual tree performance.

The family least square means were calculated using the LSMEANS option in PROC GLM in SAS.

3.5 Realised clonal performance (step 4)

The least square means were chosen to estimate the realised clonal performance as these means are free of bias due to unequal representation across blocks and plots (White & Hodge 1989). The least square means of the stem straightness scores of the clones were calculated using the LSMEANS option of the GLM procedure of SAS.

3.6 Correlation of predicted and realised performances (step 5)

The correlation of the predicted performance versus the realised clonal performance was used as a measure of efficiency of the predictions. From each family trial at least fourteen sets of predictions of clonal performances were obtained: BLUP breeding value predictions; BLUP genetic value predictions based on the assumption of 20% non-additive genetic variance; genetic value predictions based on an assumption of 37% non-additive genetic variance; and predictions based on the 11 constructed indices. In the case of the full-sib family trials an additional genetic value prediction was added which was based on the proportion of dominance variance (σ_d^2/σ_g^2) estimated for the trial. These predictions were correlated with the realised clonal performances as determined by the least square means in the clonal trials. The correlations are presented in Figure 3.2. The predictions in the family trials were done for all individuals, irrespective of whether they were represented in the clonal trials or not. However, only those trees that were selected as ortets could be used for in the correlations.

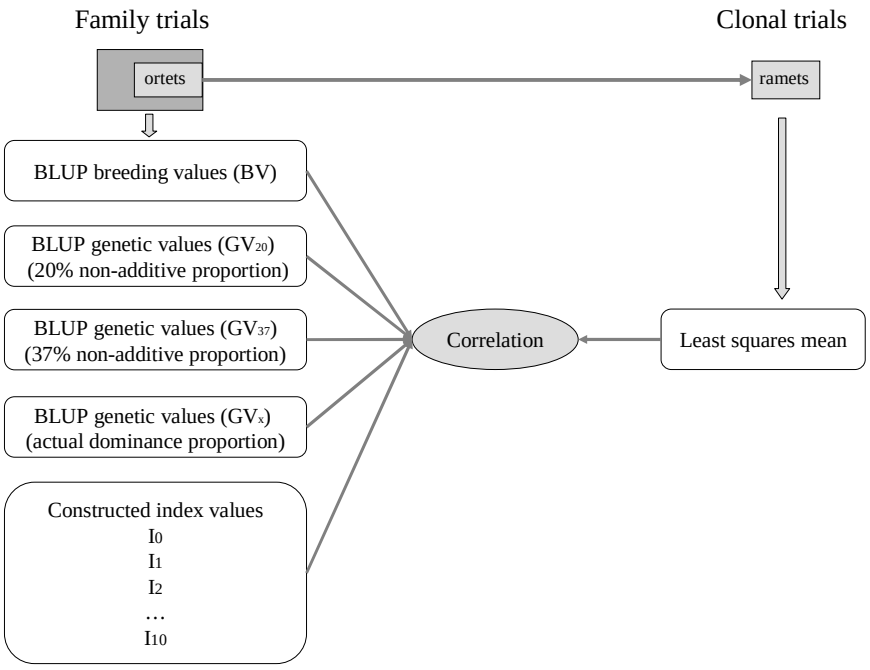


Figure 3.2 Predictions of clonal performance to be correlated with realised clonal performance.

3.6.1 Linkage of ortet and ramet performance

In order to determine the correlations, the predictions which were based on ortet performance and the realised performance of the ramets needed to be linked and joined into a single file. As the data were in a number of different formats, the first step was to convert all data into a single format (comma separated values) that could be read by SAS. Furthermore, a link was required to associate ortet identity with clonal identity to allow for merging of the data sets with the predictions with the data sets containing the realised performance information.

The replication, tree and plot variable of the ortets were merged into a single ‘link’ variable. The tree and replication variables were converted in Microsoft’s Excel (2003) to three-digit variables and the plot to a four-digit variable before merging the three values into a single variable. For the trial 1010802EA62A1 where tree number was not included in the clonal identity, the link variable consisted only of replication and plot.

The link variable was added to all files containing the clonal performance predictions. In the case of the clonal trials, the ortet information had to be added before a link variable could be created. The ortet information was obtained from selection register files that contained information on the clone identity as well as information on the ortet identity.

The files containing the predicted performances and the files containing the corresponding realised performances were imported into SAS and corresponding family and clonal data were merged by the 'link' variable. These merged data sets were used to determine the correlation between the predicted and realised clonal performances.

3.6.2 Correlation

The predicted performances that were based on ortet performance were correlated with the realised clonal performances in the clonal trials to assess the relationship between the breeding values, the genetic values and the 11 constructed indices with the realised clonal performances. As standardised index values were correlated with standardised least square means values of the clones, it was decided to use the Pearson correlation. The values were correlated using the PEARSON option of the PROC CORR procedure in SAS.

3.7 Comparison of prediction methods (step 6)

The correlations of the predicted clonal performances with the realised clonal performance were plotted on a graph to allow for visual assessment of the relative efficiencies of the eleven different predictions. The graphs were plotted using the Microsoft's Excel (2003) graph function.

CHAPTER 4
RESULTS

4.1 Introduction

A large number of trials from the CSIR eucalypt hybrid breeding programme were screened and three hybrid family and five clonal trials were identified as being potentially suitable to investigate the efficiency of BLUP predictions in hybrid populations. In addition, two *E. grandis* family and three clonal trials were identified as potential pure species controls. The data from these trials were evaluated to determine whether they were suitable for the analysis. Once the trials were chosen, genetic parameters were estimated, followed by prediction of clonal performance using different prediction methods and estimation of the realised clonal performance. The predicted and realised performances were then correlated to establish the best prediction method and the efficiency of BLUP prediction methods. The results of this study are presented in this chapter.

4.2 Statistical description of data sets

The data sets from all family and clonal trials that were selected for the study were subjected to an exploratory analysis to determine whether the data met the assumptions of normality and homogeneity which are required for analysis of variance. Trial means and standard errors were determined and significance of effects was investigated.

4.2.1 Trial means

The trial mean values of stem straightness together with their standard errors are presented in Table 4.1. The lowest average score (4.2) was obtained for the hybrid family trial 1010603EA0003 at Salique and the highest average score (7.1) for the *E. grandis* family trial 1010802EA62A1. This high average score could be ascribed to the

fact that this trial consisted of third generation improved material which had been selected for stem straightness.

4.2.2 Normality

All trials were analysed to determine whether the residuals were normally distributed. The skewness and kurtosis measures and the results of the Shapiro-Wilk test for the corrected and standardised data are presented in Table 4.1. The null-hypothesis of the Shapiro-Wilk test is that the residuals are normally distributed, therefore p-values that are larger than 0.05 indicate that values are normally distributed at the 5% level of significance. The Shapiro-Wilk test indicates that the values are normally distributed in the trials 1010802EA6207 at J. D. M. Keet, 1010603EA0004 at Grootboom, 1010806EA0011 at Westfalia and Dukuduku and the trial 1010806EA0006 at Dukuduku. In all other trials the Shapiro-Wilk test indicates a departure from normality.

The data from the sister clonal trials 1010806EA0011 at Dukuduku and Westfalia and 1010806EA0006 at Port Durnford and Dukuduku were concatenated to give larger data sets. The data from clones that were represented in both *E. grandis* clonal trials 1010806EA6204 and EA6207 were also joined to form one data set. Normality tests for the joint data sets are presented in Table 4.2. In the trial 1010806EA0006 a departure from normality was indicated by the Shapiro-Wilk test whereas the other clonal trials showed a normal distribution.

The skewness and kurtosis measures were used as additional measures of normality. In a normal distribution both measures are zero. In some trials high kurtosis values indicated leptokurtic distribution of residuals, whereas large negative values indicated a platykurtic distribution.

A third test for normality that was considered was the normal probability plot, a graph that plots the quantiles of the residuals against the quantiles of the standard normal distribution. A straight line is an indication of a normal distribution. Normal probability plots for the trials are presented in Figure 4.1.

All three tests indicated a departure from normality in some of the trials, but the departures were not considered to be large enough to warrant transformation of the data. Tests that were more robust to departures from normality than the normal ANOVA were therefore used for the analysis of the data where applicable.

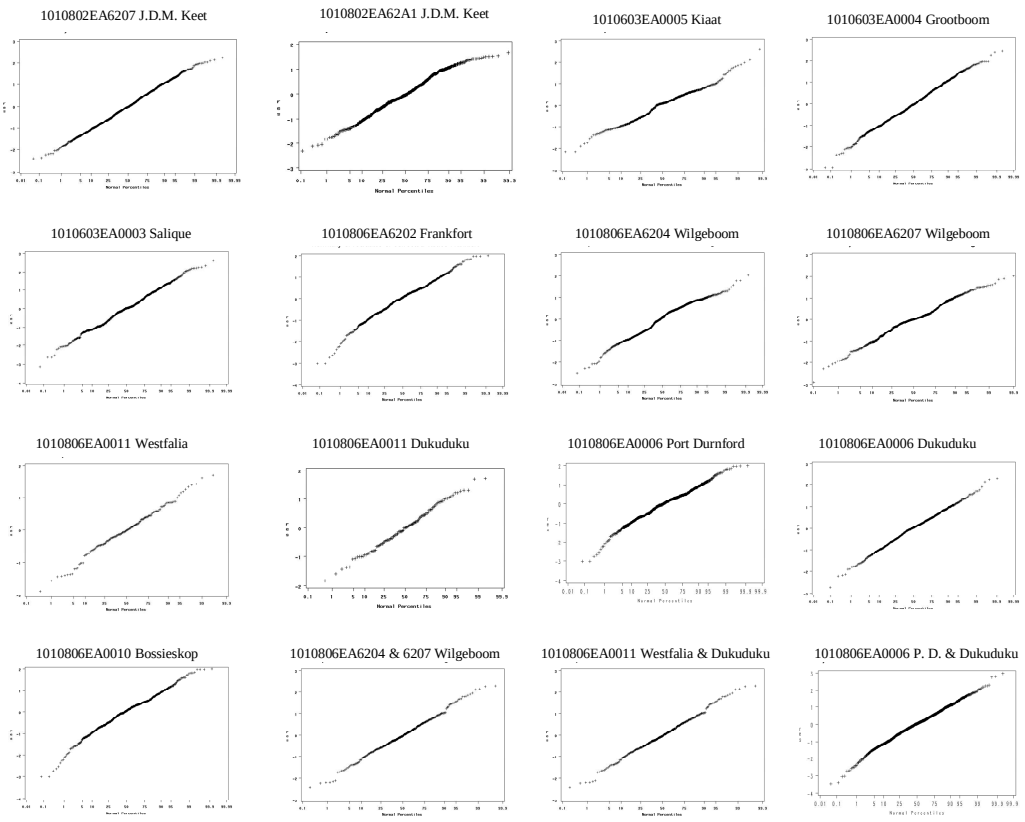


Figure 4.1 Normality probability plots of the trials.

Table 4.1 Trial means for stem straightness, measures of normality, homogeneity and significance of differences between treatments.

Trial	Trial type	Age (months)	n	Mean	Stand. error	Normality of distribution			Homogeneity of variances	Significance of treatments	
						Skewness	Kurtosis	Shapiro-Wilk Pr<W	Levene Pr>F	ANOVA Pr>F	Bonferroni test: no. of groups
1010802EA6207 J. D. M. Keet	family	49	1245	4.557	0.045	-0.028	-0.270	0.2553	0.8329	<0.0001	
1010802EA62A1 J. D. M. Keet	family	56	539	7.128	0.033	-0.177	-0.466	0.0002	0.0167	0.0001	
1010603EA0005 Kiaat	family	74	464	5.506	0.027	-0.015	0.158	0.0004	<0.0001	<0.0001	
1010603EA0004 Grootboom	family	64	1272	5.517	0.023	-0.034	-0.027	0.2238	<0.0001	<0.0001	
1010603EA0003 Salique	family	50	947	4.223	0.032	-0.010	0.023	0.0220	0.6554	<0.0001	
1010806EA6202 Frankfort	clonal	48	3374	5.138	0.091	-0.379	0.859	<0.0001	<0.0001	<0.0001	22
1010806EA6204 Wilgeboom	clonal	52	836	5.492	0.021	-0.380	-0.257	<0.0001	0.2225	<0.0001	1
1010806EA6207 Wilgeboom	clonal	44	609	6.135	0.035	-0.213	0.121	0.0010	0.0013	<0.0001	4
1010806EA0011 Westfalia	clonal	38	163	5.202	0.073	-0.157	0.191	0.4620	0.1049	0.0014	2
1010806EA0011 Dukuduku	clonal	43	119	4.891	0.091	0.018	-0.388	0.8861	0.2922	0.0196	2
1010806EA0006 Port Durnford	clonal	48	629	4.826	0.064	-0.379	0.859	<0.0001	0.5077	<0.0001	4
1010806EA0006 Dukuduku	clonal	48	656	4.664	0.049	-0.187	0.154	0.1756	0.2927	<0.0001	5
1010806EA0010 Bossieskop	clonal	34	818	4.614	0.031	0.379	0.859	<0.0001	0.0111	<0.0001	11

n = number of trees

Table 4.2 Measures of normality, homogeneity and significance of differences treatments in combined clonal data sets.

Trial	n	Normality of distribution			Homogeneity of variances	Significance of treatments	
		Skewness	Kurtosis	Shapiro-Wilk Pr<W	Levene Pr>F	ANOVA Pr>F	Bonferroni test: number of groups
1010806EA6204 & 1010806EA6207 Wilgeboom	612	-0.240	-0.188	0.0649	0.8970	<0.0001	3
1010806EA0011 Westfalia & Dukuduku	297	0.007	0.063	0.5870	0.0950	<0.0001	4
1010806EA0006 Port Durnford & Dukuduku	1312	-0.232	0.568	<0.0001	0.0005	<0.0001	10

n = number of trees

4.2.3 Homogeneity of variances

The results from the Levene test for homogeneity of error variances are presented in Table 4.1. A large p-value (>0.05 for the 5% level of significance) indicates that the data do not reject the hypothesis of homogeneity of variances. The Levene test indicates that in six trials (1010802EA62A1 at J. D. M. Keet, 1010603EA0005 at Kiaat, 1010603EA0004 at Grootboom, 1010806EA6202 Frankfort, 1010806EA6207 at Wilgeboom and 1010806EA0010 at Bossieskop) the stem assessments depart from the assumption of homogeneity of variances.

The merged *E. grandis* trials 1010806EA6204 and EA6207 and the merged G x S trials 1010806EA0011 at Westfalia and Dukuduku showed homogeneous variances (see Table 4.2). The error variances of the trials 1010806EA0006 at Port Durnford and Dukuduku were not homogeneous.

The homogeneity might be improved by taking plot means rather than individual tree measurements for the analysis. This is, however, not a useful approach in the family trials where the aim is to select individual trees rather than families.

4.2.4 Significance of treatment effects

There were significant differences at the 5% level of significance between families in all family trials and significant differences between clones in all clonal trials (see Table 4.1). There were also significant differences between clones in all three sets of merged trials (see Table 4.2).

In the clonal tests variability between clones was also assessed as number of significantly different groups (at the 5% level of significance) identified by means of the Bonferroni test (see Table 4.1). The trial 1010806EA6204 showed few significant differences. This lack of significance was observed despite the fact that the trial was relatively large with 91 clones. The trials 1010806EA0011 at Westfalia and Dukuduku also showed few significant differences between clones.

4.3 Selection of trial data

The trial data were analysed to determine whether they met a number of criteria (see Chapter 2) which had been drawn up to evaluate whether they were suitable for reliable statistical and genetic analysis. The extent to which the data met the other statistical and genetic criteria is summarised in Table 4.3. The data sets and trials that were chosen for the study are presented in Table 4.4.

The statistical criteria were used as guidelines to determine whether the data sets met basic requirements for reliable statistical analysis. These included sufficient numbers of trees per treatment within the trials, normality and homogeneity of residuals and significant differences between treatments. One of the most important of these criteria was the requirement of significant differences between clones. Ranking of clones would not be meaningful if there were no significant differences between the clones. The limited amount of variability between clones eliminated the trial 1010806EA6204 at Wilgeboom and the two clonal trials 1010806EA0011 at Westfalia and Dukuduku. Although there were significant differences between clones, only few significantly groups were identified by the Bonferroni test. However, combining of data sets allowed for the inclusion of these data sets as part of larger data sets that met more of the statistical criteria. Furthermore, a number of trials did not meet the requirements of normality and homogeneity but it was decided that the departure was not sufficiently large to warrant the exclusion of the data sets. It was therefore decided to include them in the study but to use robust methods and to interpret the data with caution. Although traits that were preferred that had been assessed on a continuous scale this criterion could not be met as stem straightness was assessed on a discrete scale.

The genetic criteria were drawn up to assess whether the data were suitable for genetic analysis. The sets of hybrid and clonal trials were selected to meet some of the genetic criteria which included links between ortets and ramets, an appropriate family structure and various degrees of relatedness of parents. One of these genetic criteria that was not met was the requirement for large numbers of families in the family trials for the estimation of heritabilities. In trials such as those of this study approximately 75 families are required for the reliable estimation of heritabilities (Verryn, personal

communication). The *E. grandis* family trials and the G x U family trials had the largest numbers of families (between 60 and 71) and were expected to give more reliable estimates than the G x S and G x C trials. It would have been advantageous to have similar environments and common genetic controls but these criteria were not met (see Table 4.3). Furthermore, although assessments were chosen to be as close as possible to 48 months, there were large deviations from this age with ages varying between 34 and 74 months.

Table 4.3 Assessment of suitability of trials for the study.

Criteria	**Measure**	***E. grandis***		***E. grandis***			**G x S**				**G x U**			**G x C**
		1010802EA6207 J. D. M. Keet	**1010806EA6202 Frankfort**	**1010802EA62A1 J. D. M. Keet**	**1010806EA6204 Wilgeboom**	**1010806EA6207 Wilgeboom**	**1010603EA0005 Kiaat**	**1010806EA0011 Westfalia**	**1010806EA0011 Dukuduku**	**1010603EA0004 Grootboom**	**1010806EA0006 Port Durnford**	**1010806EA0006 Dukuduku**	**1010603EA0003 Salique**	**1010806EA0010 Bossieskop**
Genetic	Type	family	clonal	family	clonal	clonal	family	clonal	clonal	family	clonal	clonal	family	clonal
	Coeff. of relationship	0.5	1	0.3	1	1	0.5	1	1	0.5	1	1	0.5	1
	Similarity of environments		no		no	no		yes	no		no	no		no
	Ages (months)	49	48	56	52	44	74	38	43	54	48	48	50	34
	No. of fam. or clones	71	22	60	91	25	36	24	24	69	67	64	32	37
Statistical	Number of trees	1245	3374	539	836	609	464	163	119	1272	629	656	947	818
	Normality	yes	no	no	no	no	no	yes	yes	yes	no	yes	no	no
	Homogeneity	yes	no	+/-	yes	no	yes	no	yes	no	yes	yes	yes	no
	Variability		yes		no	yes		no	no		yes	yes		yes
	Normality (merged trials)					yes			yes			no		
	Homogeneity (merged trials)					yes			yes			no		
	Variability (merged trials)					+/-			yes			yes		

Table 4.4 Trials included in the study.

Type	Family trial	Clonal trial(s)
E. grandis	1010802EA6207 JDM Keet	1010806EA6202 Frankfort
E. grandis	1010802EA62A1 JDM Keet	1010806EA6204 & 1010806EA6207 Wilgeboom
G x S	1010603EA0005 Kiaat	1010806EA0011 Westfalia & Dukuduku
G x U	1010603EA0004 Grootboom	1010806EA0006 Port Durnford & Dukuduku
G x C	1010603EA0003 Salique	1010806EA0010 Bossieskop

4.4 Genetic parameters

The additive and dominance genetic variances were estimated for the full-sib family trials. These estimates were required for the estimation of narrow and broad-sense heritability and to obtain estimates of the proportion the total genetic variance that is attributed to non-additive genetic effects. In the clonal trials repeatability of clonal performance was estimated.

4.4.1 Additive and dominance genetic variance

Estimates of additive and dominance genetic variances were required for heritability estimates and to obtain an estimate of the proportion of the genetic variance that is attributed to non-additive genetic variance. These estimates for stem straightness are presented in Table 4.5. In the G x S hybrid family trial at Kiaat no significant interaction between female and male parents could be detected, therefore no dominance effects could be determined. This may indicate the absence of dominance genetic variance or alternatively that the number of treatments was too low for an estimate the variance. In this study the estimates of the proportion of total genetic variance that can be ascribed to dominance genetic variance decreased with an increasing distance between the parent species.

Table 4.5 Additive and dominance genetic variances and the proportion of additive and dominance over total genetic variance in family trials*.

Trial	**Type**	**Age (months)**	σ^2_a **(s.e.)**	σ^2_d **(s.e.)**	$\frac{\sigma^2_a}{\sigma^2_g}$	$\frac{\sigma^2_d}{\sigma^2_g}$
1010802EA6207 JDM Keet	*E. grandis*	49	0.84 (0.41)	0.59 (0.26)	0.59	0.41
1010603EA0005 Kiaat	G x S	74	0.12 (0.07)	-	1.00	-
1010603EA0004 Grootboom	G x U	64	0.17 (0.09)	0.08 (0.04)	0.68	0.32
1010603EA0003 Salique	G x C	50	0.34 (0.21)	0.13 (0.08)	0.73	0.27

s.e. = standard error
* note that no estimates could be obtained for 1010802EA62A1 as it consisted of half-sib families

4.4.2 Narrow and broad-sense heritability

In the half-sib *E. grandis* family trial narrow-sense heritability estimates were based on family variance estimates and no estimates of broad-sense heritabilities were obtained. In the full-sib family trials estimates of both narrow and broad-sense heritability were obtained with the exception of the trial 1010603EA0005 at Kiaat where no estimate of dominance genetic variance was obtained. The estimates of broad and narrow-sense heritability are presented in Table 4.6. Estimates of narrow-sense heritability ranged from 0.214 to 0.315 in the hybrids and from 0.263 to 0.298 in *E. grandis*. Three estimates of broad-sense heritability were obtained: 0.455 for *E. grandis*, 0.313 for G x U and 0.344 for G x C.

Table 4.6 Estimates of narrow and broad-sense heritability for stem straightness in family trials.

Trial	Type	Age (months)	Number of		h^2	H^2
			families	trees / family	(s.e.)	(s.e.)
1010802EA6207 JDM Keet	*E. grandis*	49	71	16-18	0.263 (0.129)	0.445 (0.210)
1010802EA62A1 JDM Keet	*E. grandis*	56	60	8-10	0.298 (0.114)	-
1010603EA0005 Kiaat	G x S	74	36	35-36	0.315 (0.183)	-
1010603EA0004 Grootboom	G x U	64	69	8-21	0.214 (0.116)	0.313 (0.165)
1010603EA0003 Salique	G x C	50	34	10-80	0.251 (0.155)	0.344 (0.215)

s.e. = standard error

4.4.3 Repeatability of clonal performance

Repeatability of clonal performance was estimated to obtain an indication of the variability of performances of ramets within clones. A high clonal repeatability value (close to 1) would indicate that there is little variance due to the environment whereas a low clonal repeatability indicates a large error variance.

The repeatability estimates are presented in Table 4.7. Repeatability estimates ranged from 0.062 to 0.200. The trial 1010806EA0010 at Bossieskop gave the highest clonal repeatability estimates. A very low clonal repeatability (0.062) was estimated for the trial 1010806EA0011 at Dukuduku.

Table 4.7 Clonal repeatability of clonal performance in the clonal trials.

Trial	Type	Age (months)	Number of ramets per clone (k_1)	R (s.e.)
1010806EA6202 Frankfort	*E. grandis*	48	41.64	0.246 (0.032)
1010806EA6204 Wilgeboom	*E. grandis*	52	5.68	0.111 (0.031)
1010806EA6207 Wilgeboom	*E. grandis*	44	9.96	0.212 (0.044)
1010806EA0011 Westfalia	G x S	38	6.78	0.243 (0.085)
1010806EA0011 Dukuduku	G x S	43	4.92	0.062 (0.078)
1010806EA0006 Port Durnford	G x U	48	9.67	0.199 (0.042)
1010806EA0006 Dukuduku	G x U	48	10.09	0.259 (0.046)
1010806EA0010 Bossieskop	G x C	34	22.00	0.299 (0.056)

s.e. = standard error

4.5 Predicted clonal performance

Clonal performance was predicted by means of three different selection techniques, the constructed selection indices, the BLUP breeding value and the genetic value predictions. There were two to three predictions for the genetic values based on different assumptions of the proportion of non-additive genetic variance. As predictions were obtained for each tree in the family trials, only extracts from the results will be presented here. The first example of the predictions is the breeding value predictions of the G x S hybrid trial at Kiaat. The 10 best and 10 poorest performing trees are presented in Table 4.8. The second is an example of the best and poorest clonal value predictions for the G x U hybrid trial at Grootboom and is presented in Table 4.9. A third example which is presented in Table 4.10 is the constructed indices of the G x S hybrid trial at Kiaat.

Table 4.8 Breeding value prediction for stem straightness of the 10 best and 10 poorest individuals of the Kiaat hybrid trial 1010603EA0005.

Family	Replication	Plot	Tree	$\hat{g}$	Corr ($\hat{g}_i$, g_i)	var ($\hat{g}$)	Rank
76x1	4	158	2	0.9845665	0.7585950	0.1846671	1
101x5	5	222	2	0.9121858	0.7459578	0.1785658	2
15x4	5	196	4	0.9000666	0.7685619	0.1895516	3
15x4	1	9	3	0.8909704	0.7685619	0.1895516	4
45x4	3	111	5	0.8711099	0.7665333	0.1885523	5
4x2	2	70	3	0.8505773	0.7703670	0.1904430	6
4x2	3	98	3	0.7618901	0.7703670	0.1904430	7
45x7	1	36	6	0.7416471	0.7780576	0.1942644	8
45x7	3	108	1	0.7149272	0.7780576	0.1942644	9
6x7	6	262	6	0.5913543	0.7665333	0.1885523	10
50x2	6	244	1	-0.817827	0.7685619	0.1895516	454
19x7	1	11	4	-0.881305	0.7734394	0.1919651	457
19x7	5	200	5	-0.981078	0.7734394	0.1919651	458
58x7	6	242	1	-1.081779	0.7719834	0.1912431	459
50x2	1	21	3	-1.141308	0.7685619	0.1895516	460
50x7	4	160	1	-1.147367	0.7719834	0.1912431	461
50x7	3	140	4	-1.181762	0.7719834	0.1912431	462
50x2	3	133	4	-1.189631	0.7685619	0.1895516	463
50x7	2	54	6	-1.233212	0.7719834	0.1912431	464

$\hat{g}$ = predicted breeding value
var ($\hat{g}$) = variance of the breeding value prediction
Corr ($\hat{g}_i$, g_i) = the estimated correlation between true and predicted genetic values

Table 4.9 Genetic value prediction for stem straightness of the 10 best and 10 poorest individuals of the Grootboom hybrid trial 1010603EA0004.

Family	Replication	Tree	Plot	$\hat{g}$	Corr ($\hat{g}_i$, g_i)	var ($\hat{g}$)	Rank
045000008	5	4	392	1.5567033	0.8532086	0.1917605	1
045000008	3	4	248	1.5245502	0.8532086	0.1917605	2
045000008	1	2	56	1.4117591	0.8532086	0.1917605	3
045000008	1	3	56	1.4117591	0.8532086	0.1917605	4
045000008	2	1	154	1.3706745	0.8532086	0.1917605	5
045000008	2	2	154	1.3706745	0.8532086	0.1917605	6
045000008	2	4	154	1.3706745	0.8532086	0.1917605	7
045000008	3	1	248	1.2617111	0.8532086	0.1917605	9
045000008	3	5	248	1.2617111	0.8532086	0.1917605	8
045000008	4	3	326	1.25329	0.8532086	0.1917605	10
035000009	2	1	126	-1.075701	0.8512772	0.1908933	1265
039000009	2	2	103	-1.091542	0.8512772	0.1908933	1266
039000009	2	4	103	-1.091542	0.8512772	0.1908933	1267
039000009	1	1	58	-1.115274	0.8512772	0.1908933	1269
039000009	1	2	58	-1.115274	0.8512772	0.1908933	1268
037000011	1	6	72	-1.215211	0.8468805	0.1889266	1270
039000009	5	1	398	-1.216582	0.8512772	0.1908933	1271
039000009	4	6	348	-1.333201	0.8512772	0.1908933	1272
039000009	2	6	103	-1.374541	0.8512772	0.1908933	1273
039000009	3	5	204	-1.440889	0.8512772	0.1908933	1274

$\hat{g}$ = predicted genetic value
var ($\hat{g}$) = variance of the genetic value prediction
Corr ($\hat{g}_i$, g_i) = the estimated correlation between true and predicted genetic values

Table 4.10 Predictions of clonal performance based on eleven constructed indices in the hybrid trial 1010603EA0005 at Kiaat.

Clone	Rep	Tree	Plot	$I_{0\%}$	$I_{10\%}$	$I_{20\%}$	$I_{30\%}$	$I_{40\%}$	$I_{50\%}$	$I_{60\%}$	$I_{70\%}$	$I_{80\%}$	$I_{90\%}$	$I_{100\%}$
6x7	6	6	262	-0.173	0.087	0.346	0.606	0.865	1.125	1.385	1.644	1.904	2.163	2.423
76x1	4	2	158	0.561	0.747	0.932	1.118	1.304	1.490	1.676	1.862	2.047	2.233	2.419
4x2	2	3	70	0.288	0.500	0.711	0.923	1.135	1.346	1.558	1.770	1.982	2.193	2.405
101x5	5	2	222	0.591	0.754	0.917	1.079	1.242	1.405	1.568	1.731	1.893	2.056	2.219
15x4	5	4	196	0.471	0.646	0.820	0.995	1.170	1.345	1.520	1.695	1.869	2.044	2.219
15x4	1	3	9	0.471	0.642	0.814	0.986	1.157	1.329	1.501	1.672	1.844	2.015	2.187
45x7	1	6	36	0.196	0.395	0.594	0.793	0.992	1.191	1.390	1.590	1.789	1.988	2.187
45x4	3	5	111	0.490	0.650	0.811	0.971	1.131	1.292	1.452	1.612	1.772	1.933	2.093
45x7	3	1	108	0.196	0.385	0.575	0.765	0.955	1.144	1.334	1.524	1.714	1.903	2.093
4x2	3	3	98	0.288	0.468	0.649	0.829	1.010	1.190	1.371	1.551	1.732	1.912	2.093
19x7	1	4	11	-0.502	-0.657	-0.812	-0.967	-1.123	-1.278	-1.433	-1.588	-1.744	-1.899	-2.054
50x2	1	3	21	-0.975	-1.083	-1.190	-1.298	-1.406	-1.514	-1.622	-1.730	-1.838	-1.946	-2.054
17x2	4	6	143	0.047	-0.168	-0.383	-0.598	-0.813	-1.028	-1.243	-1.458	-1.673	-1.888	-2.103
17x2	4	1	143	0.047	-0.168	-0.383	-0.598	-0.813	-1.028	-1.243	-1.458	-1.673	-1.888	-2.103
50x7	4	1	160	-0.937	-1.053	-1.170	-1.287	-1.403	-1.520	-1.636	-1.753	-1.870	-1.986	-2.103
50x2	3	4	133	-0.975	-1.100	-1.224	-1.349	-1.474	-1.599	-1.724	-1.849	-1.974	-2.099	-2.224
50x7	3	4	140	-0.937	-1.065	-1.194	-1.323	-1.452	-1.580	-1.709	-1.838	-1.967	-2.095	-2.224
19x7	5	5	200	-0.502	-0.692	-0.882	-1.073	-1.263	-1.453	-1.644	-1.834	-2.024	-2.215	-2.405
50x7	2	6	54	-0.937	-1.084	-1.230	-1.377	-1.524	-1.671	-1.818	-1.965	-2.111	-2.258	-2.405
58x7	6	1	242	-0.670	-0.845	-1.021	-1.196	-1.371	-1.547	-1.722	-1.897	-2.072	-2.248	-2.423

Constructed indices from $I_{0\%}$ to $I_{100\%}$ where the subscript indicates the weight on individual tree performance, the remainder of the weight is attributed to the family mean

4.6 Realised clonal performance

In this study realised clonal performance has been defined as the performance of the clones in the clonal trials. Realised clonal performance was expressed as clone least square means of the standardised stem straightness assessments. As an example of the results, the least square means from one trial are presented here. The least square means of the G x S hybrid are presented in Table 4.11.

Table 4.11 Realised clonal performance of the clones in the combined clonal trials 1010806EA0011 at Westfalia and Dukuduku.

Clone	LSM stem
GxS132	0.142161
GxS136	0.273799
GxS140	-0.05502
GxS142	0.515317
GxS143	-0.1659
GxS144	-0.20506
GxS149	0.882029
GxS150	0.285299
GxS151	-0.63724
GxS152	0.159288
GxS153	-0.51172
GxS154	-1.18932
GxS156	0.112845
GxS157	0.074204
GxS158	-0.71371
GxS159	0.074636
GxS160	0.776182
GxS164	0.105674
GxS165	0.018813
GxS168	0.243695
GxS170	-0.22783
GxS171	-0.1639
GxS174	-0.22838
GxS177	-0.23661

4.7 Correlation between predicted and realised clonal performance

Correlations between predicted and realised clonal performance was obtained for four sets of trials: *E. grandis* was represented by the family trial 1010802EA6207 at J. D. M. Keet and a clonal trial 1010806EA6202 at Frankfort as well as the family trial 1010802EA62A1 at J. D. M. Keet and the clonal trials 1010806EA6204 and 1010806EA6207 at Wilgeboom; the G x S hybrid was represented by the hybrid family trial 1010603EA0005 at Kiaat and the clonal trials 1010806EA0011 at Westfalia and Dukuduku; the G x U hybrid by the hybrid family trial 1010603EA0004 at Grootboom and the clonal trials at Port Durnford and Dukuduku; and the G x C hybrid by the hybrid family trial 1010603EA0003 at Salique and the clonal trial EA010 at Bossieskop.

The correlations for the constructed and BLUP indices together with the significance of the correlation are presented in Table 4.12. Most of the correlations were significant except for the indices placing a high weight on family mean and all correlations of the G x C trial 1010603EA0003 at Salique which were all non-significant at the 5% level of significance.

The values of the highest correlations lay between 0.301 and 0.645. The value of these correlations decreased with increasing genetic distance between the parent species. These correlations were lower than the estimated correlation between the true and predicted genetic values which were given by Matgen (see Corr ($\hat{g}_i$, g_i) in Table 4.8 and Table 4.9). These estimated correlations in *E. grandis* varied between 0.6 (for BV) and 0.88 (for GV_{37}) whereas the highest realised correlation was 0.645. In the G x S hybrid, the estimated values lay between 0.74 and 1.14, whereas the highest realised value was 0.522. For G x U the estimated correlations lay between 0.53 and 1.04 and the highest realised value was 0.335 whereas in the G x C hybrid the estimated correlations were between 0.63 and 1.01 and the highest realised correlation 0.301.

The correlations of the clonal performance predicted in the *E. grandis* trial 1010802EA6207 with the realised clonal performance in trial 1010806EA6202 are not presented. For these sets of trials the predictions based on family means gave better correlations than the predictions based on individual tree performance. Investigation

into the reasons for these results showed that the correlations were not between actual ortets and ramet performance, as the trees of the family trial had been re-numbered after thinning. Through this re-numbering of trees the link between the actual ortet and ramet was lost. The correlations were therefore between trees that were full-sibs of the ortets rather than the ortets themselves and therefore the family means were more informative on the clonal performance than the individual tree performances. For this reason only one *E. grandis* control was available for comparison purposes.

Table 4.12 Correlation between predicted and realised clonal performance for different prediction methods.

Type	**n**	**Correlation between predicted and realised clonal performance (Prob. > |r|)**														
		$I_{0\%}$	**$I_{10\%}$**	**$I_{20\%}$**	**$I_{30\%}$**	**$I_{40\%}$**	**$I_{50\%}$**	**$I_{60\%}$**	**$I_{70\%}$**	**$I_{80\%}$**	**$I_{90\%}$**	**$I_{100\%}$**	**BV**	**GV_{20}**	**GV_{37}**	**GV (other)**	
E. grandis	21	0.201	0.307	0.403	0.482	0.541	0.582	0.609	0.626	0.637	0.642	**0.645**	0.535	0.586	0.621	-	-
		(0.382)	(0.176)	(0.070)	(0.027)	(0.011)	(0.006)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.012)	(0.005)	(0.003)		
G x S	23	0.354	0.405	0.446	0.477	0.498	0.512	0.519	**0.522**	**0.522**	0.521	0.518	0.477	0.491	0.502	-	-
		(0.098)	(0.055)	(0.033)	(0.021)	(0.016)	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.021)	(0.017)	(0.015)		
G x U	64	0.207	0.259	0.297	0.320	0.331	**0.335**	0.334	0.331	0.326	0.322	0.317	0.292	0.303	0.314	GV_{32}	0.311
		(0.101)	(0.039)	(0.017)	(0.010)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.019)	(0.012)	(0.012)		(0.012)
G x C	34	0.221	0.255	0.279	0.293	0.300	**0.301**	0.299	0.295	0.291	0.286	0.281	0.295	0.301	0.305	GV_{27}	0.302
		(0.210)	(0.145)	(0.110)	(0.092)	(0.085)	(0.084)	(0.086)	(0.090)	(0.096)	(0.101)	(0.107)	(0.090)	(0.084)	(0.080)		(0.082)

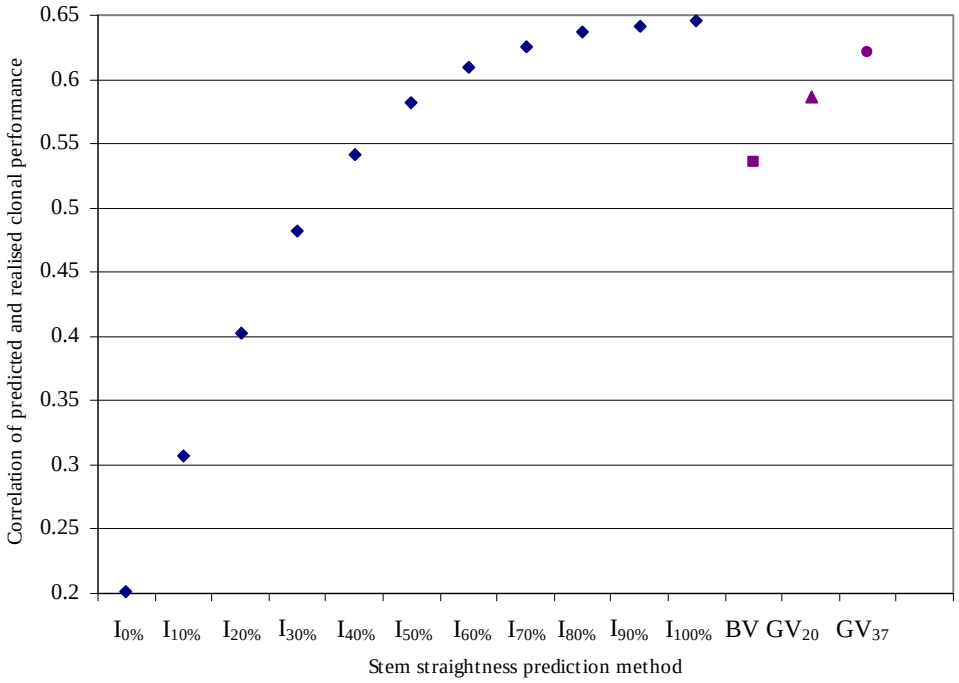
$I_{x\%}$ = Constructed indices where the subscript (x %) indicates the weight on individual tree performance, the remainder of the weight is attributed to family mean
BV = BLUP breeding value prediction
GV = BLUP genetic value prediction with the subscript indicating the percentage non-additive genetic variance
n = number of pairs of values that were correlated
Note: The highest correlations per type are indicated in bold and the highest BLUP predictions are underlined
.

4.8 Comparison of prediction methods

The predicted and realised clonal performances were correlated to assess the effectiveness of the predictions and to determine which of the methods gave the best results. The eleven constructed indices ($I_{0\%}$ to $I_{100\%}$) were included to obtain a measure of the optimal prediction that could be expected for the data sets assuming balanced data. The correlations obtained for the BLUP predictions were therefore compared with the correlations obtained with the constructed indices to assess the performance of BLUP predictions in hybrids. The results are graphically represented in Figure 4.2 (*E. grandis*), Figure 4.3 (G x S), Figure 4.4 (G x U) and Figure 4.5 (G x C).

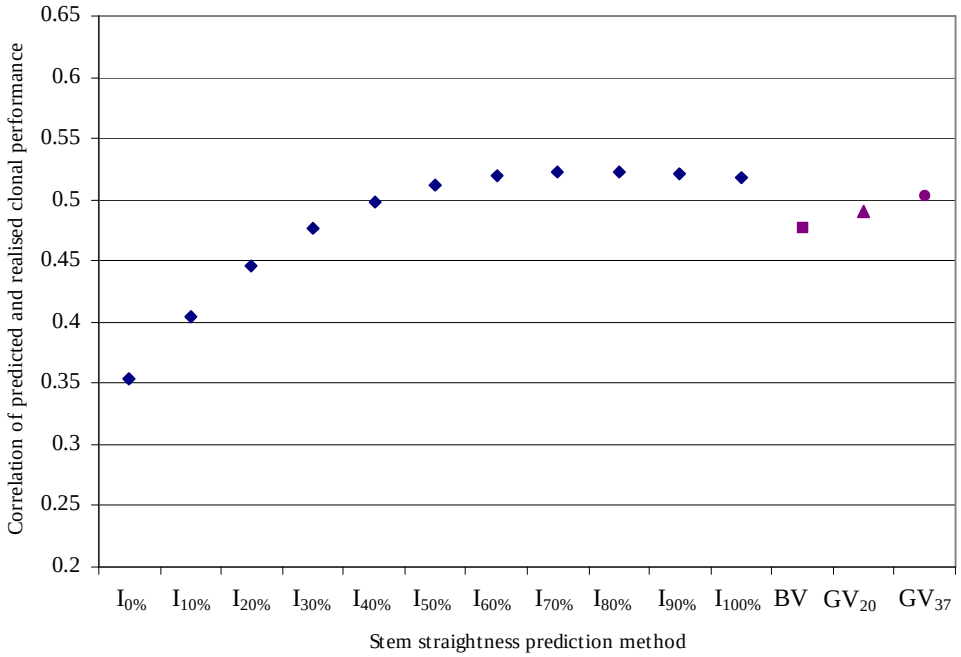
The highest correlations of the BLUP indices with realised clonal performance were obtained with the clonal predictions GV_{37}. These predictions were made based on an assumption that 37% of the total genetic variance was attributable to non-additive genetic variance. The predictions that were based on the actual estimations of non-additive (dominance) genetic variance did not yield higher correlations but the difference was minimal.

In all trials the constructed indices gave the highest correlations but the BLUP genetic value predictions based on the assumption of 37% non-additive genetic variance were very close to these values. This indicates that these clonal predictions were satisfactory.



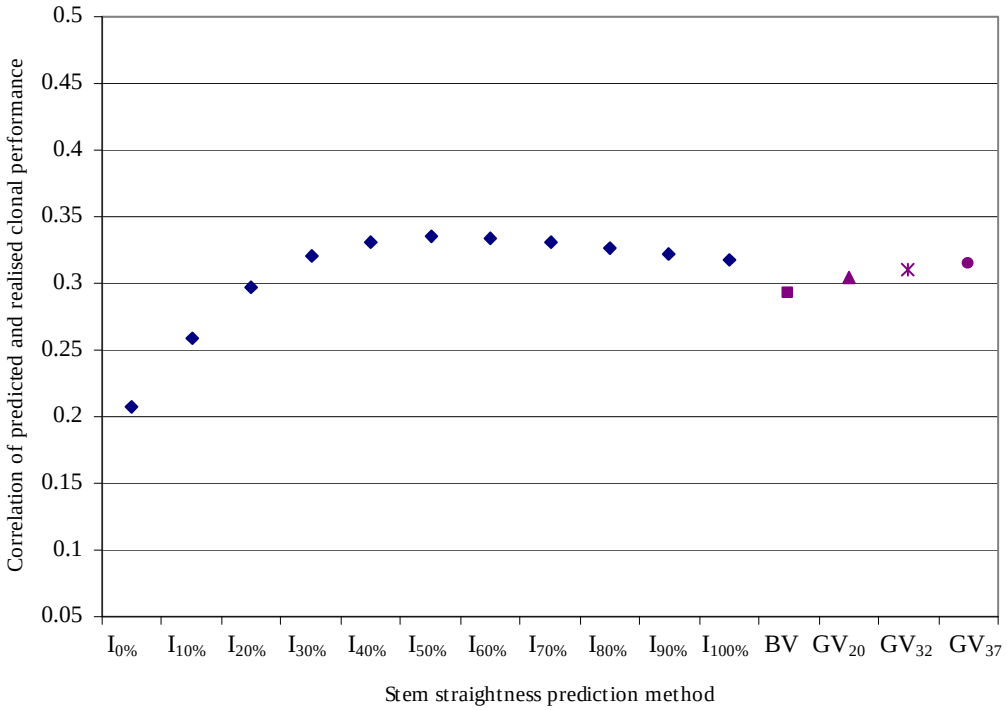
$I_{x\%}$ =	Constructed indices from where the subscript indicates the weight on individual tree performance, the remainder of the weight is attributed to family mean
BV =	BLUP breeding value prediction
GV =	BLUP genetic value prediction with the subscript indicating the proportion non-additive genetic variance of the total genetic variance

Figure 4.2 Correlation of predicted and realised clonal stem straightness performance for different prediction methods in *E. grandis*.



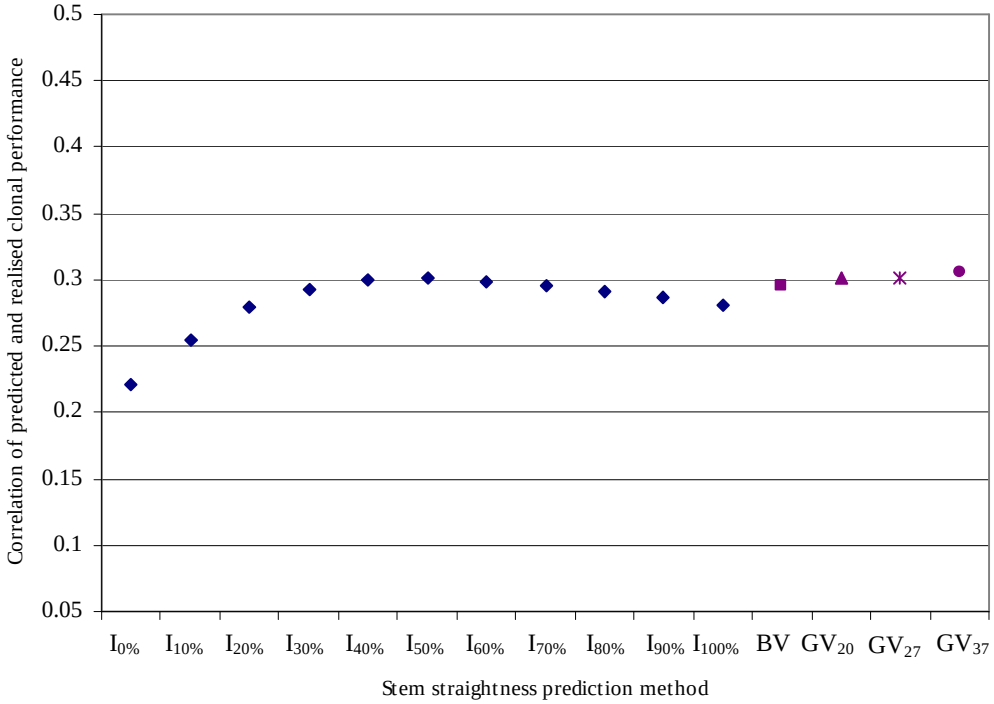
$I_{x\%}$ =	Constructed indices from where the subscript indicates the weight on individual tree performance, the remainder of the weight is attributed to family mean
BV =	BLUP breeding value prediction
GV =	BLUP genetic value prediction with the subscript indicating the proportion non-additive genetic variance of the total genetic variance

Figure 4.3 Correlation of predicted and realised clonal stem straightness performance for different prediction methods in the G x S hybrid.



$I_{x\%}$ =	Constructed indices from where the subscript indicates the weight on individual tree performance, the remainder of the weight is attributed to family mean
BV =	BLUP breeding value prediction
GV =	BLUP genetic value prediction with the subscript indicating the percentage non-additive genetic variance of the total genetic variance

Figure 4.4 Correlation of predicted versus realised clonal stem straightness performance for different prediction methods for the G x U hybrid.



$I_{x\%}$ =	Constructed indices from where the subscript indicates the weight on individual tree performance, the remainder of the weight is attributed to family mean
BV =	BLUP breeding value prediction
GV =	BLUP genetic value prediction with the subscript indicating the percentage non-additive genetic variance of the total genetic variance

Figure 4.5 Correlation of predicted and realised clonal stem straightness performance for different prediction methods in the G x C hybrid.

CHAPTER 5
DISCUSSION AND CONCLUSIONS

Prediction of breeding and genetic values in interspecific hybrid populations is challenging. In hybrids differences in the genetic composition of parental species may for example lead to the production of unbalanced gametes, infertility of certain genotypes, unequal expression of parental genes, transgressive segregation and genomic modifications which will all have an impact on the resulting phenotypic and genotypic composition of the progeny. Transmission of alleles may for example be affected leading to departures from Mendelian inheritance. Furthermore the quantitative genetic theory of inheritance in hybrid populations has not been fully developed. Some models have been developed to explain some aspects of inheritance and some properties of hybrids such as heterosis (e.g. Stuber & Cockerham 1966; Shockley 1973; Wei *et al.* 1991; Wu & Li 2000) but these are very complex and have not reached a stage where they have been implemented to adapt selection practices for hybrid populations. Therefore in practice hybrids are often selected using selection methodology that has been developed for pure species despite large differences in the underlying genetic properties of the populations. It is not known to what extent such selections are able to predict the performance of progeny or clones in hybrid populations. Although the fundamental problem has been identified by several authors including Griffin *et al.* (2000), Potts and Dungey (2004) and Flachenecker *et al.* (2006), it has not been investigated to which extent the use of pure species selection methodology is applicable in hybrid populations.

This study is the first study known to investigate the efficiency of prediction in hybrid populations. The study was based on historical data to establish to which degree the predicted performance corresponded with the realised performance. As a case study the selection of ortets in three hybrid *Eucalyptus* populations was investigated. As this selection deals with the selection of mother trees for the propagation of clonal material, it focuses on selection in a population that is in Hardy-Weinberg and linkage

disequilibrium without addressing the added complexities of segregation and transmission of alleles.

The trait stem straightness was chosen for this study. Stem straightness is an important trait as it can translate into utilisable saw timber volume. Deviations from stem straightness such as lean, crook, sweep or twist all reduce the utilisable saw timber volume and increase the cost of handling and transport (Ehrenberg 1970). This trait was selected for the study as it had been less severely selected and therefore it was expected that there would be a larger variability in the performance of ortets. The trait is, however, difficult to assess and assessments are usually based on a subjective scale. In this study stem straightness was assessed on an 8-point scale. The resulting data were treated as continuous data as the standardised data that were corrected for fixed effects were continuous rather than categorical.

One of the first problems encountered in the analysis of hybrid data is lack of balance of data and the tendency of the errors to depart from normality and homogeneity. Data from forest trials tend to be unbalanced anyway and deviation from balance may be increased in hybrid trials because of inviability of some genotypes. Although many of the runts may be removed in the nursery, inviability may sometimes only be expressed after establishment in the field and there may be very poor performing trees in the trials (Potts *et al.* 2003). Some families may have a better survival rate than others, leading to lack of balance of data. Abnormalities and inviability may lead to poor performance of hybrid individuals (e.g. Edmans 1999; de Assis 2000; Hancock 2005), but heterosis on the other hand may contribute to individuals that outperform most other individuals (e.g. Pierce 1994 a; Dale & Dieters 2007). This may yield outliers on both ends of the spectrum and these may contribute to a non-normal distribution of the data and potentially also non-normality of the distribution of the residuals. Volker *et al.* (2008) reported, for example, that F_1 *E. globulus* × *E. nitens* hybrids showed a more skewed distribution for DBH at two years than the parent species and suggested that this could be ascribed to trees that were growing very slowly. As these abnormal trees died out, the distribution improved at later ages. Factors such as unequal gene frequencies, dominance, non-allelic interactions, genotype-by-environment interactions and linkage disequilibrium may also contribute to non-normality (de Toledo & de Miranda Filho

1985). These authors suggest that directional dominance and unequal gene frequencies may be the biggest cause of distortion whereas the other factors play a lesser role but that this distortion will only be a problem in the case of a small number of segregating loci. Large differences in the performance of hybrids may also result in the departure from homogeneity of variances. Many eucalypt hybrid families contain both abnormal and normal individuals (Potts & Dungey 2004). Such large within-family variability in some families may lead to unequal variances. Although departures from normality and homogeneity were encountered in this study, the problem was not limited to the hybrid trials and similar problems were experienced in the pure species trials.

The departure from normality and homogeneity may impact on the reliability of the analysis of hybrid data. Normality and homogeneity of variances are assumptions on which some statistical methods such as ANOVAs are based. For the estimation of variance components methods are required that are more robust to departure from these assumptions. For unbalanced data REML estimates offer advantages as these estimators are more robust to violations of the assumption of an underlying normal distribution (Robinson 1991; Dieters *et al.* 1995; Wu & Li 2000). For this reason REML was used to estimate the additive and dominance genetic variances.

In plant breeding applications there have been few reports on the robustness of BLUP against departures from assumptions such as incorrect specification of relationships, non-normality or heterogeneity of variances (Piepho *et al.* 2008). Fonseca *et al.* (2001) have done a simulation study to investigate the impact of two violations of assumptions on genetic gains over 10 generations. The assumptions investigated were that variances are known without errors and normal distribution of random errors. The authors found that in the long term the genetic gain was not impacted by violation of these assumptions. There have, however, been no reports on the effect of violation of the assumptions of a Hardy-Weinberg equilibrium and linkage equilibrium on the efficiency of BLUP. These are assumptions that are violated in hybrid populations.

Good estimates of variance components are essential for the reliability of BLUP estimates as poor predictions may be obtained with inaccurate estimates of genetic variance components (Piepho *et al.* 2008). Although BLUP methodology assumes that

the variances and covariances of the base population are known these are in practice not known and need to be estimated (Mrode 1996). The variance components therefore needed to be studied for BLUP predictions.

The estimates of additive and dominance genetic variance that were obtained in this study should be interpreted with caution because of limitations of the trial designs and the questions around the validity of these concepts in hybrids (Gordon 1999). Epistatic genetic variances could not be estimated from the family trials and were therefore ignored. This resulted in epistatic effects being integrated into the additive and dominance genetic effects. This may lead to an overestimation of the variances (Bouvet & Vigneron 1996). A further assumption that was made in the estimation of these genetic parameters was the absence of inbreeding. This assumption may not be true, as the parents which originated from open pollinations may have some degree of relatedness. All these factors may contribute to errors in the estimation of the genetic variance components.

Bouvet and Vigneron (1996) suggested that the relative importance of additive and dominance genetic variances differ between hybrids and proposed that this may be a function of the relatedness of the parental species. The authors reported a high proportion of additive genetic variance for volume and DBH (80%) in *E. urophylla* × *E. grandis* hybrids whereas a lower proportion (40%) was obtained in *E. urophylla* × *E. pellita,* a hybrid between more closely related parents. This trend of increasing additive genetic variance with increasing genetic distance between the parental species was also found in this study if the case of the *E. grandis* × *E. saligna* hybrid is ignored. In the case of the G x S a lack of dominance genetic variance may possibly be ascribed to limitations of the mating design. The reported trend does, however, contradict the expectation of a larger proportion of non-additive genetic variance in hybrids (Potts & Dungey 2004; Volker *et al.* 2008). As each hybrid is represented by a single set of trials, other factors besides relatedness of parents may play a role in the trends observed and no final conclusions can be made regarding the reasons for the observed trends. In *E. grandis* a lower proportion of additive genetic variance may, for example, be ascribed to the tendency of additive genetic variance to decrease in improved populations as this

population had been subjected to selection for stem straightness for a number of generations (Snedden *et al.* 2007).

The narrow-sense heritabilities that were obtained in this study should be also interpreted with caution, as the estimates were based on a relatively low number of families and values and are likely to be overestimated for a number of reasons. Furthermore the definition of narrow-sense heritability in hybrids is disputed (Gordon 1999). The heritability estimates were based on data from single sites and were therefore expected to be upwardly biased, as the interaction between additive genetic effects and environmental effects were confounded with the additive genetic effects (Dieters *et al.* 1995). Selective thinning before assessments may have also increased heritability estimates (Matheson & Raymond 1984). All family trials, except for the G x C hybrid family trial at Salique, had been subjected to selective thinning before the assessments were made. Heritability estimates may furthermore be inflated in hybrids. Volker *et al.* (2008) reported on such inflated estimates relative to estimates that were based on pure species with common parentage.

In this study narrow-sense heritability estimates between 0.214 and 0.315 were obtained for stem straightness of the hybrids and between 0.263 and 0.298 for the *E. grandis* trials. The estimates that were obtained were relatively high in comparison with estimates of heritability of stem straightness in some *Eucalyptus* species such as *E. camaldulensis* (0.12 - 0.15), *E. dunnii* (0.14), and *E. deglupta* (0.09) (Gülbaba 1999; Marcó & White 2002; Mahmood *et al.* 2003; Mesén *et al.* 2007). Similar estimates were, however, recorded for *E. grandis* (0.17 - 0.3), *E. nitens* (0.28), and *E. longirostrata* (0.384 - 0.447) (Marcó & White 2002; Henson *et al.* 2007; Mesén *et al.* 2007; Snedden *et al.* 2007; Hamilton & Potts 2007). As such estimates are dependent on the trials, the age, the stage of improvement and method of assessment, the comparisons are only of limited value. Low heritabilities may indicate that the scale used to assess the trait is not sufficiently refined or that the subjective assessment methods results in high levels of errors (Snedden *et al.* 2007; Henson *et al.* 2007). The heritabilities recorded in this study were relatively high and therefore did not seem to indicate such problems.

Clonal repeatability estimates were obtained to gain a measure of the uniformity of the clones and in order to assess whether clonal repeatability of performance may have an impact on the efficiency of the predictions. The clonal repeatability estimate is not equivalent to broad-sense heritability, as the relatedness of the clones was not considered in the estimation and the phenotypic variation rather than the genetic variation was considered. The clonal repeatability estimates, which varied between 0.062 and 0.299, were generally in the same order as the narrow-sense heritability estimates and somewhat lower than the broad-sense heritability estimates. This may reflect the underlying connection between the heritability and repeatability estimates. It could be expected that the higher the heritability of a trait, the higher the relative genetic contribution and therefore the higher the relative uniformity of the clones.

As this study aimed at assessing the applicability of BLUP predictions in hybrids, there was a need for a measure to determine the efficiency of these predictions. Correlations between predicted and realised clonal performances were used as a measure of the efficiency of the predictions. An additional measure was, however, required to assess the relative size of these correlations. For this purpose a series of constructed indices were created which could give an approximate indication of the range of correlations that could be expected from the data. A series of different weights on family means versus single tree values were used to create a series of simple Smith-Hazel type selection indices. Such selection indices do not adjust the predictions for lack of balance of data and they are therefore generally considered to be inferior to BLUP predictions. In this case the selections were, however, based on a single site and a single trait and there were, for example, no problems with differences in precision at different sites which may lead to bias (Cotterill & Dean 1990).

If the correlations between the predictions based on the constructed indices and the realised clonal performances are considered, a general trend of improved predictions with increased weight on individual tree performance can be observed. The closer the parent species were related, the bigger the improvement of selections based on individual tree performance versus selections based on family mean. In *E. grandis* a steep increase was observed between 0% and 50% weight on individual tree performance, after which the improvement was less pronounced. In G x S the highest

increase in correlations was between 0% and 40%, in G x U and G x C the increase was between 0% and 30%. It was expected that this shift is linked to the heritability in the trials but no clear trend could be observed. It can also be noted that after the initial improvement in correlations with increased weight on individual tree performance, the improvement seemed to reach a plateau after which additional weight on single tree performance had a relatively small impact on improving the prediction. In the G x U and G x C hybrid, there was, however, a slight decrease in the correlations that were only based on single tree performance. These trends indicate that the predictions based on family and single tree values were relatively robust if at least 50% weight was being placed on individual tree performance.

If the trends of the correlations of the different BLUP predictions with the realised clonal performances are considered, a clear trend can be observed of an improvement of the predictions with an increase in the proportion of non-additive genetic variance on which the BLUP prediction is based. The breeding values, where non-additive genetic variances are not considered, give the lowest correlations, whereas the predictions based on the assumption of 37% non-additive genetic variance give the highest correlations. This confirms the expectation that the genetic values are more suitable for the prediction of clonal performances than the breeding values (Mullin & Park 1992). It was, however, also expected that higher correlations may be obtained with actual estimates of non-additive genetic variances than with predictions based on values that have been estimated for other populations or other traits. This expectation was not realised as the correlations were marginally lower than the best correlations. This phenomenon is not fully understood but may indicate shortcomings in the estimation of the non-additive genetic proportions.

There were relatively large differences between the realised correlations and the estimated correlations of the predicted values and the true genetic values. These differences may be attributed to the fact that the realised correlations were between predicted values and realised values, whereas the estimated correlations were between predicted values and actual genetic values. The realised clonal performances include an environmental and error component, whereas the true genetic values exclude these components. The environmental component may be large as the hybrid trials on which

the predictions were based were on different sites than the clonal trials in which the realised clonal performances were assessed.

The correlations of the genetic values (GV_{37}) and the realised clonal performance were only slightly lower than the best correlations of the constructed indices. As the constructed indices are not likely to be used in practice as the optimal weight on family versus individual tree performance can only be determined after evaluation of clonal performance, the genetic value prediction seems to be the most efficient of the methods used for the prediction of stem straightness in the populations used in this study.

No distinct differences could be detected between the validity of predictions in the pure species and hybrids. The results rather indicate a gradual change with increase in genetic distance between the parent species. If the highest correlations that were obtained for *E. grandis* and each of the hybrids are considered, a trend of higher correlations between predicted and realised values with increased relatedness of the parent species can be observed. For *E. grandis* the best correlation between predicted and realised clonal performance was 0.645; in the G x S hybrid this decreased to 0.522; in the G x U hybrid to 0.335; and in the G x C hybrid to 0.301. This may be a function of the heritability or it may indicate that the ability to predict clonal performance decreases with increased genetic distance. The efficiency of the predictions may thus be less affected by the method of prediction than the ability to predict clonal performance accurately in hybrids between distantly related parents as the predictive power of the selections based on individual and family performance seems to decrease with an increasing genetic distance between parent species. This apparent trend needs to be investigated further.

This study was based on historical data and several factors limited the ability to investigate the applicability of BLUP predictions in hybrids fully. Problems experienced included limited numbers for the reliable estimation of genetic parameters and a limited range of performances of the ortets due to prior selection. The study nevertheless gave a clear indication of that BLUP may be sufficiently robust for the selection of clones in the *Eucalyptus* hybrid populations that were investigated. However, in order to confirm the results it is recommended to design an experiment in which the robustness of BLUP

may be investigated more fully. It is recommended that such a study be based on hybrid and clonal trials that are sufficiently large to obtain reliable estimates of genetic parameters. Furthermore it is recommended that they contain common genetic controls, are assessed at the same age and are established in similar environments. It is important that the ortets and clones are unselected or randomly chosen and that they represent a range of performance. This study was limited to single trait and single site selection, and it is therefore recommended to assess two to three traits to investigate multi-trait selections and to investigate multi-site selections. A drawback of such a study is the duration and costs associated. Three suggestions are presented here that aim at addressing some of these problems.

The first option is based on the hybrid deployment option 2 that was presented in Figure 1.3 and involves the cloning of hybrid seedlings. The aim of this option is to establish the clonal trial shortly after the hybrid trial, thereby reducing the duration of the study. For this purpose it is suggested that the hybrid seedlings are established as clonal hedges. The first cuttings from all hedges are established as a family trial. It may be more practical to have all cuttings as mini-hedges first to reduce nursery space and to establish only the ortets as macro-cutting hedges. A number of ortets are selected from the initial hedges for bulking up of the material for the establishment of clonal trials. The selection should be a systematic selection (e.g. every third hedge) or a random selection and should not be based on performance. This option is illustrated in Figure 5.1.

A second option for the study is to plant the seedlings in a hybrid family trial and to collect cutting material for the establishment of hedges from the trial as soon as possible after establishment. This option would be cheaper as fewer hedge plants would be involved but may take longer. This option is illustrated in Figure 5.2.

A third option would be to use an operational hybrid and clonal trial setup in a breeding programme and to assess additional non-selected traits that are known to have a high heritability, for example pilodyn assessments. In such a study there may still be limitations to the design of the experiments, but there would be a better choice of traits. Furthermore it has to be assumed to clonal effects are negligible.

In all options predictions of clonal performance may be correlated with realised clonal performance as was done in this study, or ideally the clonal data could be used for backward selection of the ortets. At this stage this selection option was not available in Matgen, but it may be developed or programmes such as ASReml may be used for backward selection. Backward selection would give a better indication of the genetic potential of the clones and it is expected that the correlation will be closer to the expected correlations between predicted and actual genetic value. Rank changes may also be investigated as well as genetic gains.

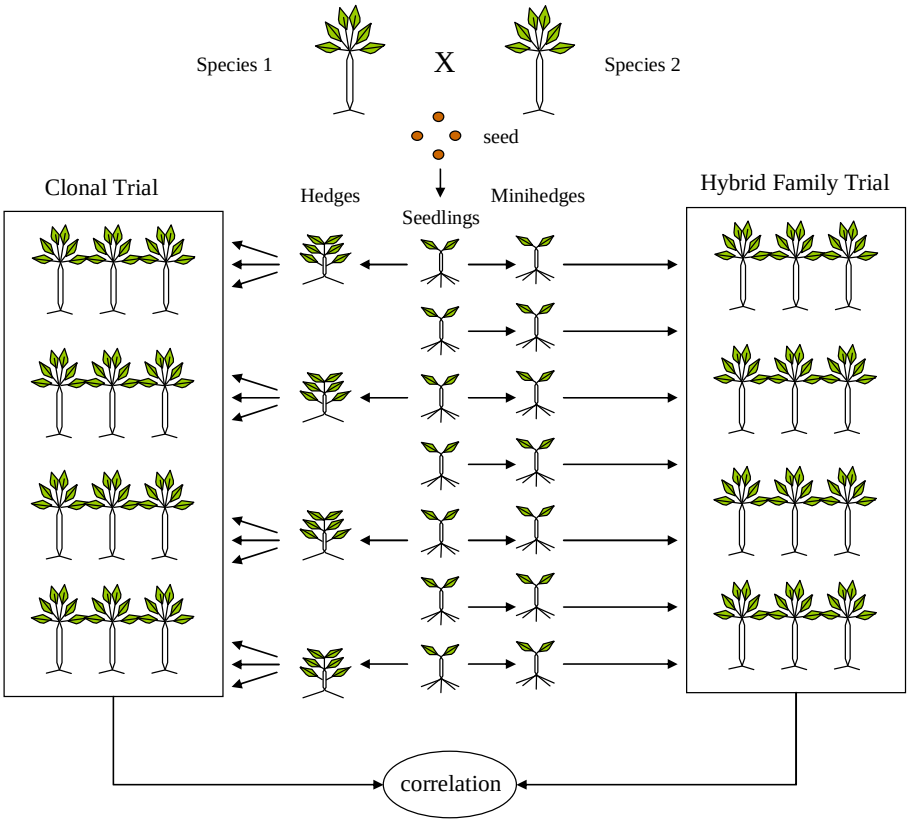


Figure 5.1 Design option 1 for a study investigating BLUP prediction in hybrids.

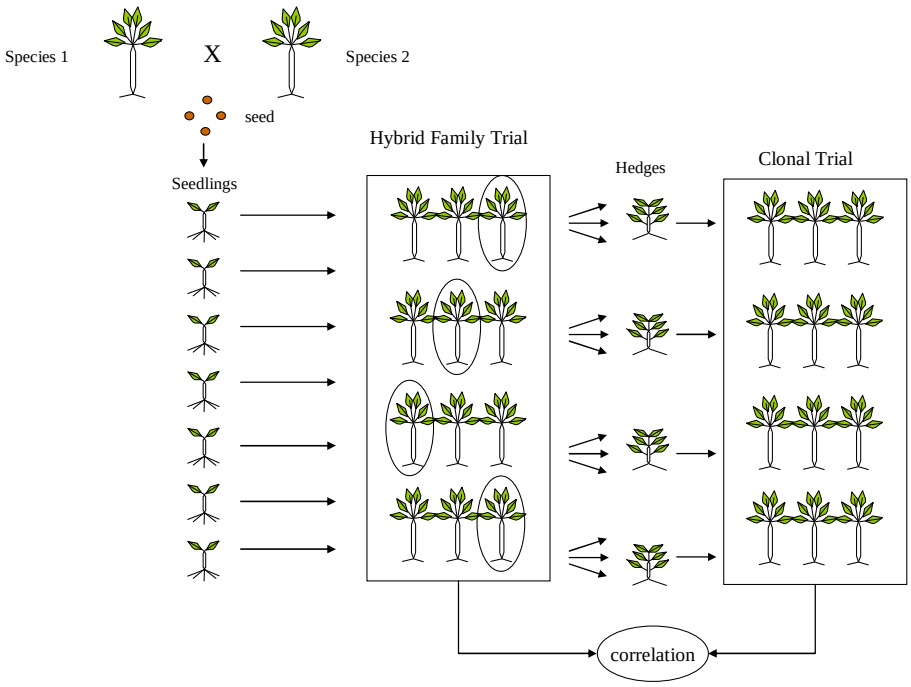


Figure 5.2 Design option 2 for a study investigating BLUP prediction in hybrids.

The currents study has focused on clonal selections. However, with increasing interest in the deployment of F_2 hybrid seed (see option 4 in Figure 1.3) there is a growing need to establish the efficiency of BLUP predictions to select parents for the F_2 generation. Such a study may be based on existing trials, but as in the option 3 discussed above, it would be essential to assess traits that are non-selected. It is expected that the ability to predict the performance of a F_2 hybrid populations will be adversely affected by factors such as departure from Mendelian inheritance.

Forest tree species are not ideally suited as model species due to long generation intervals, large trial sizes and limited numbers that can be tested. Annual crops could be investigated for the study of theoretical questions on the robustness of BLUP. Such results may however only be applicable to forest tree breeding to a limited extent as there may be large differences in the breeding, traits and inheritance. Alternatively simulation of selection in hybrid populations could be considered.

In summary, the principal findings of this study were:

- there were no observable differences in the efficiency of the BLUP predictions in the studied hybrids in comparison to prediction in *E. grandis*,
- there was, however, a decrease in the correlations between predicted and realised clonal performances with increasing genetic distance between the parents,
- genetic values were better predictors of clonal performance than breeding values,
- genetic values which were based on assumptions of higher non-additive genetic variances (37%) were more efficient than those based on assumptions of lower proportions (20%), and
- BLUP methods of predicting genetic values were as efficient as any other methods in predicting the clonal performance in the three hybrid populations that were investigated.

REFERENCES

Baack, EJ & Rieseberg, LH 2007, 'A genomic view of introgression and hybrid speciation', *Current Opinion in Genetics & Development*, vol. 17, pp. 513-518.

Barton, NH 2000, 'Estimating multilocus linkage disequilibria', *Heredity*, vol. 84, pp. 373-389.

Becker, WA 1992, *Manual of Quantitative Genetics*, *5th Edition*, Academic Enterprises, Pullman.

Bison, O, Ramalho, MAP, Rezende, GDSP, Aguiar, AM & de Resende, MDV 2006, 'Comparison between open pollinated progenies and hybrid performance in *Eucalyptus grandis* and *Eucalyptus urophylla'*, *Silvae Genetica*, vol. 55, no. 4-5, pp. 192-196.

Blaikie, S, O'Farrell, P, Müller, W, Wei, X, Scott, N, Sykes, S & Chacko, E 2002, 'Assessment and selection of new cashew hybrids. A report to the Rural Industries Research and Development Corporation', RIRDC Publication No. 01/177, Rural Industries Research and Development Corporation, Barton.

Bouchez, A & Goffinet, B 1990, 'Evaluation of selection index: application to the choice of an indirect multitrait selection index for soybean breeding', *Theoretical and Applied Genetics*, vol. 79, pp. 261-267.

Bouvet, JM & Vigneron, P 1996, 'Variance structure in *Eucalyptus* hybrid populations', *Silvae Genetica*, vol. 45, no. 2-3, pp. 171-177.

Box, GED 1953, 'Non-normality and tests on variance', *Biometrika*, vol. 41, pp. 318-335.

Bradshaw, HD & Grattapaglia, D 1994, 'QTL mapping in interspecific hybrids of forest trees', *Forest Genetics*, vol. 1, no. 4, pp. 191-196.

Brewbaker, JL & Sun, WG 1999, 'Trees and heterosis', in: Coors, JG & Pandey, S (eds.) *The Genetics and Exploitation of Heterosis in Crops*, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Publishers, Madison, pp. 463-478.

Brondani, RPV, Williams, ER, Brondani, C & Grattapaglia, D 2006, 'A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 320 microsatellite markers for the genus', *BMC Plant Biology*, vol. 6, viewed 28 July 2008, http://www.biomedcentral.com/1471-2229/6/20.

Brooker, MIH 2000, 'A new classification of the genus *Eucalyptus* L'Hér (Myrtaceae)', *Australian Systematic Botany*, vol. 13, pp. 79-148.

Burke, JM & Arnold, ML 2001, 'Genetics and fitness of hybrids', *Annual Review of Genetics.*, vol. 35, pp. 31-52.

Charocosset, A, Bonnisseau, IB, Touchebeuf, O, Burstin, J, Barrière, Y, Gallais, A & Denis, J-B 1998, 'Prediction of maize hybrid silage performance using molecular data: comparison of several models for specific combining ability', *Crop Science*, vol. 38, no. 1, pp. 38-44.

Cilas, C 2003, 'How to improve the efficiency of individual cocoa tree selection for quantitative traits in progeny trials?', *Proceedings of the International Workshop on Cocoa breeding for improved production systems*, 19-21 October 2003, Accra, Ghana, pp. 33-40.

Combes, JG, Bouvet, JM & Bailleres, H 1997, 'Analyses of growth, morphological and wood properties traits in ortet populations of *Eucalyptus* in the Congo', in: *Proceedings of the IUFRO conference on Silviculture and Improvement of Eucalypt, volume 3: Silviculture, productivity and utilization of eucalypt*, August 24-29 1997, Salvaldor, Brazil, Colombo: EMBRAPA. Centro Nacional de Pesquisa de Florestas, pp. 37-42.

Cooper, M & Merrill, RE 1999, 'Heterosis: its exploitation in crop breeding', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 316-329.

Cotterill, PA & Dean, CA 1990, *Successful tree breeding with index selection*, CSIRO, Melbourne.

Dale, G & Dieters, M 2007, 'Economic returns from environmental problems: breeding salt- and drought-tolerant eucalypts for salinity abatement and commercial forestry', *Ecological Engineering*, vol. 31, no. 3, pp. 175-182.

De Assis, TF 2000, 'Production and use of *Eucalyptus* hybrids for industrial purposes', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 63-74.

Delaporte, KL, Conran, JC & Sedgley, M 2001, 'Interspecific hybridization within *Eucalyptus* (Myrtaceae): subgenus *Symphyomyrtus*, sections *Bisectae* and *Adnataria'*, *International Journal of Plant Sciences*, vol. 162, no. 6, pp. 1317-1326.

De la Vega, AJ, DeLacy, IH & Chapman, SC 2007, 'Progress over 20 years of sunflower breeding in Central Argentina', *Field Crops Research*, vol. 100, pp. 61-72.

De Resende, MDV & Barbosa, MHP 2006, 'Selection via simulated individual BLUP based on family genotypic effects in sugar cane', *Pesquisa Agropecuária Brasileira*, vol. 41, no. 3, pp. 421-429.

De Toledo, JFF & de Miranda Filho, JB 1985, 'Predicting the potential of open-pollinating populations for the production of superior F1 hybrids', *Theoretical and Applied Genetics*,, vol. 71, no. 3, pp. 563-569.

De Vicente, MC & Tanksley, SD 1993, 'QTL analysis of transgressive segregation in an interspecific tomato cross', *Genetics*, vol. 134, pp. 585-596.

Dieters, MJ, White, TL, Littell, RC & Hodge, GR 1995, 'Application of approximate variances of variance components and their ratios in genetic tests', *Theoretical and Applied Genetics*, vol. 91, pp. 15-24.

Dieters, MJ & Dungey, HS 2000, 'Relationship between the relative importance of non-additive variance and the genetic correlation between hybrid and parental populations in some *Pinus* species', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 87-92.

Dodds, KS 1955, 'Hybrid vigour in plant breeding', *Proceedings of the Royal Society of London, Series B, Biological Sciences*, vol. 144, no. 915, pp. 185-192.

Dos Santos, AJ, Chaves, LJ, Duarte, JB & Brasil, EM 2005, 'Prediction of hybrid means from a partial circulant diallel table using the ordinary least square and the mixed model methods', *Genetics and Molecular Biology*, vol. 28, no. 2, pp. 314-320.

Durel, CE & Kremer, A 1995, 'Hybridization after self-fertilization: a novel perspective for the maritime pine breeding program', *Genetics*, vol. 2, pp. 117-120.

Duvick, DN 1999, 'Heterosis: feeding people and protecting natural resources', in: Coors, JG & Pandey, S (eds.) *The Genetics and Exploitation of Heterosis in*

Crops', American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Publishers, Madison, pp. 19-29.

Edmands, S 1999, 'Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence', *Evolution*, vol. 53, no. 6, pp. 1757-1768.

Ehrenberg, C 1970, 'Breeding for stem quality', *Unasylva*, vol. 97-98, viewed 28 July 2008, http://www.fao.org/docrep/a2173e/a2173e04.htm.

Eldridge, K, Davidson, J, Harwood, C & van Wyk, G 1997, *Eucalypt domestication and breeding*, Oxford Science Publications, Oxford.

Eskes, B 2004, 'Synthesis and conclusions of the INGENIC workshop on "Cocoa breeding for improved production systems" held October 19-21, 2003 at the Miklin Hotel, Accra, Ghana', *INGENIC* Newsletter, vol. 9, pp. 2-7.

Euclid, n.d., *The classification of the genus Eucalyptus*, viewed 26 April 2006, www.anbg.gov.au/cpbr/cd-keys/Euclid/sample/html/classification.htm.

Falconer, DS 1981, *Introduction to quantitative genetics*, Longman, Harlow.

FAO 2001, *Global forest resources assessment 2000: main report.* FAO Forestry Paper No. 140. Rome, viewed 28 July 2008, www.fao.org/forestry/fo/fra/main/index.jsp.

Fehr, WR & Hadley, HM 1980, *Hybridization of crop plants*, American Society of Agronomy and Crop Science Society of America, Publishers, Madison, Wisconsin.

Fishman, L & Willis, JH 2005, 'A novel meiotic drive locus almost completely distorts segregation in Mimulus (Monkeyflower) hybrids', *Genetica*, vol. 169, pp. 347-353.

Flachenecker, C, Frisch, M, Falke, KC &Melchinger, AE 2006, 'Trends in population parameters and best linear unbiased prediction of progeny performance in a European F2 maize population under modified recurrent full-sib selection', *Theoretical and Applied Genetics*, vol. 112, pp. 483-491.

Fonseca, R, Euclydes, RF, Torres, RA, Ribeiro, JI & Silva, MA 2001, 'Efeito da violação de pressuposições da metodologia de modelos mistos na avaliação genética animal (Effect of assumption of violations of the mixed model methodology on the genetic evaluation)', *Arquivo Brasilerio de Medicina Veterinária e Zootecnia*, vol. 53, no. 1, pp. 122-129, viewed 28 July 2008, http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-09352001000100020.

Gilmour, AR, Cullis, BR, Welham, SH & Thompson, R 2002, *ASReml Reference Manual, Second Edition, Release 1.0,* VSN International Ltd., Hemel Hempstead.

Goldman, IL 1999, 'Inbreeding and outbreeding in the development of a modern heterosis concept', in: Coors, JG and Pandey S (eds.) *Genetics and exploitation of heterosis in crops.* American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin, pp. 7-18.

Gordon, IL 1999, 'Quantitative genetics of intraspecies hybrids', *Heredity,* vol. 83, pp. 757-764.

Grattapaglia, D & Bradshaw, HD 1994, 'Nuclear DNA content of commercially important *Eucalyptus* species and hybrids', *Canadian Journal of Forest Research,* vol. 24, pp. 1074-1078.

Grattapaglia, D 2007 'Marker-assisted selection in *Eucalyptus*' in: Guimarães, E, Ruane, J, Scherf, B, Sonnino, A & Dargie, J (eds.) *Marker-assisted selection. Current status and future perspectives in crops, livestock, forestry and fish,* FAO, Rome, pp. 252-281.

Griffin, R, Harbard, J, Centurion, C & Santini, P 2000, 'Breeding *Eucalyptus grandis* × *globulus* and other interspecific hybrids with high inviability – problem analysis and experience at Shell forestry projects in Uruguay and Chile', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 1-13.

Gülbaba, AG 1999, 'Tohum bahçesine dönüştürülecek (*Eucalyptus camaldulensis* Dehn.) döl denemesinde kalitim derecesi ve genetic kazancin belirlenmesi (Heritabilities and expected genetic gain from *Eucalyptus camaldulensis* open pollinated progeny trial)', DOA Dergisi No. 5, Doğu Akdeniz Ormancılık Araştırma Enstitüsü, Yayını, Sayfa: 1-25, Tarsus.

Guo, M, Rupe, MA, Yang, X, Crasta, O, Zinselmeier, C, Smith, OS & Bowen, B 2006, 'Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis', *Theoretical and Applied Genetics*, vol. 113, pp. 831-845.

Gur, A & Zamir, D 2004, 'Unused natural variation can lift yield barriers in plant breeding', *PLoS Biology*, vol. 2, no. 10, pp. 1610-1615.

Gwaze, DP, Bridgwater, FE & Lowe WJ 2000, 'Performance of interspecific F_1 eucalypt hybrids in Zimbabwe', *Forest Genetics*, vol. 7, no. 4, pp. 295-300.

Hajjar, R & Hodgkin, T 2007 'The use of wild relatives in crop improvement: a survey of developments over the last 20 years', *Euphytica*, vol. 156, no. 1-2, pp.1-13.

Hall, MC & Willis, JH 2005, 'Transmission ratio distortion in intraspecific hybrids of *Mimulus guttatus:* implications for genomic divergence', *Genetics*, vol. 170, pp. 375-386.

Hamilton, MG & Potts, BM 2007, 'Review of *Eucalyptus nitens* genetic parameters', *Proceedings of the Australian Forest Genetics Conference: Breeding for wood quality*, 11-14 April 2007, Hobart, Australia, viewed 23 July 2008, http://proceedings.com.au/afgc.

Hancock, JF 2005, 'Contributions of domesticated plant studies to our understanding of plant evolution', *Annals of Botany*, vol. 96: pp. 953-963.

Harvey, W 1990, *User's guide for LSMLMW and MIXMDL PC-2 version.*

Hazel, LN 1943, 'The genetic basis for constructing selection index', *Papers on quantitative genetics and related topics*, Department of Genetics, North Carolina State College, Raleigh, pp. 316-330.

Hegarty, MJ & Hiscock, SJ 2005, 'Hybrid speciation in plants: new insights from molecular studies', *New Phytologist*, vol. 165, pp. 411-423.

Hegarty, MJ, Jones, JM, Wilson, ID, Barker, GL, Coghill, JA, Sanchez-Baracaldo, P, Liu, G, Buggs, RJA, Abott, RJ, Edwards, KJ & Hiscock, SJ 2005, 'Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation', *Molecular Ecology*, vol. 14, pp. 2493-2510.

Henderson, CR 1984, *Applications of Linear Models in Animal Breeding*, University of Guelph.

Henry, RJ & Shepherd, SM 2000, 'Molecular analysis of hybrids', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.) *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 169-176.

Henson, M, Smith, HJ & Boyton, S 2007, '*Eucalyptus longirostrata*: a potential species for Australia's tougher sites?', *Proceedings of the Australian Forest Genetics Conference: Breeding for wood quality*, 11-14 April 2007, Hobart, Australia, viewed 23 July 2008, http://proceedings.com.au/afgc.

Herrmann, RG, Maier, RM & Schmitz-Linneweber, C 2003, 'Eucaryotic genome evolution: rearrangement and coevolution of compartmentalized genetic information', *Philosophical Transactions: Biological Sciences*, vol. 358, pp. 97-97.

Hettasch, MH, Snedden, CL & Verryn, SD 2007, 'The effect of inbreeding in *Eucalyptus grandis*', Poster presented at the IUFRO Working Group 2.08.03 "Eucalypts and Diversity: Balancing Productivity and Sustainability" 22-26 October 2007, Durban, South Africa.

Hill, KD & Johnson, LAS 1995, 'Systematic studies in the eucalypts – 7, a revision of the bloodwoods, genus *Corymbia* (Myrtaceae)', *Telopea*, vol. 6, pp. 173-505.

Hochholdinger, F & Hoecker, N 2007, 'Towards the molecular basis of heterosis', *Trends in Plant Science*, vol. 12, no. 9, pp. 427-432.

Hoon, M 1992, 'Report on the 2 year results of an *Eucalyptus grandis* × *E. saligna* hybrid family trial series established on two sites', *CSIR Report FOR DEA 526*, CSIR, Pretoria.

Jovanocic, T & Booth, TH 2002, 'Improved species climatic profiles', *RIRDC Publication no. 02/095*, Barton.

Junghans, TG, Peters-Robinson, I, Bertolucci, FL & Alfenas, AC 1998, 'The use of self-incompatibility in the production of hybrid eucalyptus seed by Aracruz Cellulose in Brazil', *Genetics and Molecular Biology*, vol. 21, no. 3, viewed 28 July 2008, http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47571998000300015&lng=es&nrm=iso&tlng=es

Kain, DP 2003, 'Genetic parameters and improvement strategies for the *Pinus elliottii* var. *elliottii* × *Pinus caribaea* var. *hondurensis* hybrid in Queensland, Australia', PhD thesis, Australian National University.

Kee, NS, Cheong, TK, Haw, KC, Ooi, HSH, KaYee, L, Kayaroganam, P, von Uexküll, H & Härdter, R 2003, 'Clonal oil palm: production, yield performance and nutritional requirements' in: Fairhurst, T & Härdter, R (eds.), *Oil palm:*

management for large and sustainable yields, PPI/PPIC and IPI, Singapore, pp. 99-114.

Khurana, DK & Khosla, PK 1998, 'Hybrids in forest tree improvement', in: Mandal, AK & Gibson, GL (eds.), *Forest genetics and tree breeding*, CBS Publishers & Distributors, New Delhi.

Kuckuck, H, Kobabe, G & Wenzel, G 1985, *Fundamentals of Plant breeding*, Springer Verlag, Berlin, Heidelberg.

Kumar, S & Singh, PK 2004, 'Mechanisms for hybrid development in vegetables', in: Singh, PK, Dasgupta, SK & Tripathi, SK (eds.), *Hybrid vegetable development*, Food Products Press, Binghamton, pp. 383-409.

Ladiges, PY & Udovicic, F 2000, 'Comment on a new classification of eucalypts', *Australian Systematic Botany*, vol. 13, pp. 149-152.

Lamkey, KR & Edwards, JW 1999, 'The quantitative genetics of heterosis', in: Coors, JG and Pandey, S (eds.) *The genetics and exploitation of heterosis in crops*, Proceedings of an ASA-CSSA-CIMMYT Conference, Mexico City, Mexico. ASA-CSSA-SSSA, Madison, Wisconsin, pp. 31-48.

Landry, CR, Hartl, DL & Ranz, JM 2007, 'Genome clashes in hybrids: insights from gene expression', *Heredity*, vol. 99, pp. 483-493.

Lanza, LLB, de Souza Jr., CL, Ottoboni, LMM, Vieira, MLC & de Souza, AP 1997, 'Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers', *Theoretical and Applied Genetics*, vol. 94, no. 8, pp. 1023-1030.

Larter, EN & Gustafson 1980, JP. 'Triticale', in: Fehr, WR & Hadley, HH (eds.) *Hybridization of crop plants*, American Society of Agronomy and Crop Science Society of America, Publishers, Madison, Wisconsin, pp. 681-694.

Levene, M 1960, 'Robust tests for the equality of variance' in: Olkin, C & Palo Alto, CA (eds.), *Contributions to Probability and Statistics*, Stanford University Press, Stanford.

Li, B & Wu, R 2000, 'Quantitative genetics of heterosis in aspen hybrids', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 184-190.

Lippman, ZB & Zamir, D 2007, 'Heterosis: revisiting the magic', *Trends in Genetics*, vol. 23, no. 2, pp. 60-66.

Lofgren, DL & Stewart, TS 1991, 'Efficacy of alternative multivariate best linear unbiased prediction models for genetic evaluation of swine', *Journal of Animal Science*, vol. 69, pp. 4388-4396.

MacRae, S & Cotterill, PP 2000, 'Application of biotechnology in hybrid breeding', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.) *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 48-52.

Maenhout, S, de Baets, B, Haesaert, G & van Bockstaele, E 2007, 'Support vector machine regression for the prediction of maize hybrid performance', *Theoretical and Applied Genetics*, vol. 115, pp. 1003-1013.

Mahmood, K, Marcar, NE, Naqvi, MH, Arnold, RJ, Crawford, DF, Iqbal, S & Aken, KM 2003, 'Genetic variation in *Eucalyptus camaldulensis* Dehnh. for growth and stem straightness in a provenance-family trial on saltland in Pakistan', *Forest Ecology and Management*, vol. 176, no. 1-3, pp. 405-416.

Marcó, M & White, TL 2002, 'Genetic parameter estimates and genetic gains for *Eucalyptus grandis* and *E. dunnii* in Argentina', *Forest Genetics*, vol. 9, no. 3, pp. 205-215.

Matheson, AC & Raymond, CA 1984, 'Effects of thinning in progeny tests on estimates of genetic parameters', *Silvae Genetica*, vol. 33, pp. 125-128.

McClintock, B 1984, 'The significance of responses of the genome to challenge', *Science*, vol. 226, pp. 792-801.

Meddings, RA, McComb, JA, Calver, MC, Thomas, SR & Mazanec, RA 2003, '*Eucalyptus camaldulensis* × *globulus* hybrids', *Australian Journal of Botany*, vol. 51, pp. 319-331.

Mesén, F, Vásques, W & Víquez, E 2007, 'Ensayos de families F_2 de *Eucalyptus deglupta* Y *E. grandis* con fines de conversion en huertos semilleros', *Agronomía Costarricense*, vol. 31, no. 2, pp. 9-20.

Mráček, J 2005, 'Investigation of interspecific genome-plastome incompatibility in *Oenothera* and *Passiflora*, PhD Thesis, Ludwig-Maximillians Universität, München.

Mrode, RA 1996, *Linear models for the prediction of animal breeding values*, CAB International, Wallingford.

Mullin, TJ & Park, YS 1992, 'Estimating genetic gains from alternative breeding strategies for clonal forestry', *Canadian Journal of Forest Research*, vol. 22, pp. 14-23.

Muro-Abad, JI, Rocha, RB, Cruz, CD, de Araújo, EF 2005, 'Obtainment of *Eucalyptus* spp. hybrids aided by molecular markers –SSR analysis', *Scientia Forestalis*, vol. 67, pp. 53-63.

Myburg, AA, Griffin, AR, Sederoff, RR & Whetten, RW 2003, 'Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F_1 hybrid based on a double pseudo-backcross mapping approach', *Theoretical and Applied Genetics*, vol. 107, pp. 1028-1042.

Myburg, AA, Vogl, C, Griffin AR, Sederoff, RR & Whetten, RW 2004 'Genetics of postzygotic isolation in *Eucalyptus*: whole-genome analysis of barriers to introgression in a wide interspecific cross of *Eucalyptus grandis* and *E. globulus*', *Genetics*, vol. 166, no. 3, pp. 1405–1418.

Nikles, DG & Griffin, AR 1992, 'Breeding hybrids of forest trees: definitions, theory, some practical examples, and guidelines on strategy with tropical acacias', in: Carron, LT & Aken, KM (eds.), *Proceedings of an international workshop held in Tawau, Sabah, Malaysia*, Tawau, Sabah, Malaysia 1-4 July 1991, pp. 101-109.

Orr, HA & Presgraves, DC 2000, 'Speciation by postzygotic isolation: forces, genes and molecules', *BioEssays*, vol. 22, no. 12, pp. 1085-1094.

Peng, G 2004, 'Testing normality of data using SAS ®', paper presented at the PharmaSUG conference, 23-26 May 2004, San Diego, Califorina, USA.

Pepe, B, Surata, K, Suhartono, F, Sipayung, M, Purwanto, A & Dvorak, W 2004, 'Conservation status of natural populations of *Eucalyptus urophylla* in Indonesia and international efforst to protect dwindling gene pools', in: *Forest Genetic Resources no. 31*, FAO, Rome, pp. 62-64.

Piepho, HP, Möhring, J, Melchinger, AE & Büchse, A 2008, 'BLUP for phenotypic selection in plant breeding and variety testing', *Euphytica*, vol. 161, no. 1-2, pp. 209-228.

Pierce, BT 1992, 'Report on the identification of *Eucalyptus grandis* cuttings from 4 seedling seed orchards ex northern Transvaal and Zululand', *CSIR Report FOR-DEA 505*, CSIR, Pretoria.

Pierce, BT 1993, 'Report on the delivery of cuttings from two *Eucalyptus* hybrid family trials at Grootboom and Salique', *CSIR Report FOR-DEA 608*, CSIR, Pretoria.

Pierce, BT 1994 a, 'Growth evaluation plus the delivery of cuttings from a series of *Eucalyptus* hybrid family trials planted at Pan and Salique', *CSIR Report FOR-DEA 751*, CSIR, Pretoria.

Pierce, BT 1994 b, 'The six year results of a *Eucalyptus grandis* × *E. saligna* hybrid family trial series planted at Kiaat and Morgenzon plantations plus selections', *CSIR Report FOR-DEA 839*, CSIR, Pretoria.

Poehlman, JM 1979, *Breeding field crops*, AVI Publishing Company, Westport.

Poke, FS, Martin, DP, Steane, DA, Vaillancourt, RE & Reid, JB 2006, 'The impact of intragenic recombination on phylogenetic reconstruction at the sectional level in *Eucalyptus* when using a single copy nuclear gene (cinnamoyl CoA reductase)', *Molecular Phylogenetics and Evolution*, vol. 39, pp. 160-170.

Potts, BM, Volker, PW, Tilyard, PA & Joyce, K 2000, 'The genetics of hybridisation in the temperate *Eucalyptus*', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 200-211.

Potts, BM, Barbour, RC, Hinston, AB & Vaillancourt, RE 2003, 'Genetic pollution of native eucalypt gene pools- identifying the risks', *Australian Journal of Botany*, vol. 51, pp. 1-25.

Potts, BM 2004, 'Genetic Improvement of eucalypts', in Burley, J & Youngquist, JA (eds.) *Encyclopaedia of Forest Sciences 4 Volume set 1 (1) 200*, Elsevier Science, Oxford, pp. 1480-1490.

Potts, BM & Dungey, HS 2004, 'Interspecific hybridization in *Eucalyptus*: key issues for breeders and geneticists', *New Forests*, vol. 27, pp. 115-138.

Pryor, LD & Johnson, LAS 1971, *A classification of Eucalyptus*, Australian National University Press, Canberra.

Purba, AR, Flori, A, Baudouin, L & Hamon, S 2001, 'Prediction of oil palm (*Elaeis guineensis* Jacq.) agronomic performances using the best linear unbiased predictor (BLUP)', *Theoretical and Applied Genetics*, vol. 102, pp. 787-792.

Resende, RMS, Jank, L, do Valle CB & Bonato, ALV 2004, 'Biometrical analysis and selection of tetraploid progenies of *Panicum maximum* using mixed model methods', *Pesquisa Agropecuária Brasileira*, vol. 39, no. 4, pp. 335-341.

Rieger, RA, Michaelis, A & Green, MM 1991, *Glossary of genetics*, Springer Verlag, Berlin, pp 553.

Rieseberg, LH & Carney, SE 1998, 'Tansley Review No. 102, Plant hybridization', *New Phytologist*, vol. 140, pp. 599-624.

Rieseberg, LH, Archer, MA & Wayne, RK 1999, 'Transgressive segregation, adaptation and speciation', *Heredity*, vol. 83, pp. 363-372.

Rieseberg, LH, Stuart, JEB & Gardner, KA 2000, 'Hybridization, introgression and linkage evolution', *Plant Molecular Biology*, vol. 42, pp. 205-224.

Rieseberg, LH & Willis JH 2007, 'Plant speciation', *Science*, vol. 317, pp. 910-914.

Robinson, GK 1991, 'That BLUP is a good thing: the estimation of random effects', *Statistical Science*, vol. 6, no. 1, pp. 15-32.

Salmon, A, Ainouche, ML & Wendel, JF 2005, 'Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae)', *Molecular Ecology*, vol. 14, pp. 1163-1175.

SAS Institute Inc. 1990, *SAS Procedures Guide, Version 6, 3rd Edition*, Cary, NC:SAS Institute.

Shapiro, SS & Wilk, MB 1965, 'An Analysis of Variance Test for normality (complete samples)', *Biometrika*, vol. 52, pp. 591-611.

Shelbourne, CJA 2000, 'Some insights on hybrids in forest tree improvement', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.) *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 53-62.

Shockley, W 1973, 'Deviations from Hardy-Weinberg frequencies caused by assortative mating in hybrid populations', *Proceedings of the National Academy of Sciences*, vol. 70, no. 3, pp. 732-736.

Shull, GH 1908, 'The composition of a field of maize', *American Breeders Association Annual Report.*, vol. 4, pp. 296-301.

Shull, GH 1952, 'Beginnings of the heterosis concept', in Gowen, JW (ed), *Heterosis*, Iowa State College Press, Ames, pp. 14-48.

Siikamäki, P 1999, 'Developmental instability in hybrids between *Lychnis viscaria* and *Lychnis alpina* (Caryophyllaceae)', *American Journal of Botany*, vol. 86, no. 12, pp. 1683-1686.

Singh, PK, Dasgupta, SK & Tripathi, SK 2004, *Hybrid Vegetable Development*, Food Products Press, Binghamtom, New York.

Smith, HF 1936, 'A discriminant function for plant selection', *Papers on quantitative genetics and related topics*, Department of Genetics, North Carolina State College, Raleigh, pp. 466-476.

Snedden, CL 2001, *Broad and Narrow sense heritabilities in a cloned open pollinated Eucalyptus grandis breeding population*, MSc thesis, University of Pretoria.

Snedden, CL, Roux, CZ & Verryn, SD 2007, 'Broad- and narrow-sense heritabilities in a South African cloned open-pollinated *Eucalyptus grandis* breeding population', *Southern Hemisphere Forestry Journal*, vol. 69, no. 2, pp. 81-90.

Snedecor, GW & Cochran, WG 1967, *Statistical methods*, Iowa State University Press, Ames.

Snyder, LA, Freifelder, D & Hartl, D 1985, *General Genetics*, Jones and Bartlett Publishers, Boston.

Soh, AC 1994, 'Ranking parents by best linear unbiased prediction (BLUP) breeding values in oil palm', *Euphytica*, vol. 76, pp. 13-21.

Soh, AC, Gan, HH, Wong, GG, Ho, TY & Tan, CC 2003, 'Estimates of within family genetic variability for clonal selection in oil palm', *Euphytica*, vol. 133, pp. 147-163.

Springer, NM & Stupar, RM 2007, 'Allelic variation and heterosis in maize: How do two halves make more than a whole?', *Genome Research*, vol. 17, pp. 264-275.

Steane, DA, Nicolle, D, McKinnon, GE, Vaillancourt, RE & Potts, BM 2002, 'Higher-level relationships among the eucalypts are resolved by ITS-sequence data', *Australian Systematic Botany*, vol. 15, no. 1, pp. 49-62.

Stuber, CW & Cockerham, CC 1966, 'Gene effects and variances in hybrid populations', *Genetics*, vol. 54, pp. 1279-1286.

Suzuki, DT, Griffiths, AJF & Lewontin, RC 1981, *An introduction to genetic analysis*, WH Freeman and Company, San Francisco.

Tibbits, WN, Potts, BM & Savva, MH 1991, 'Inheritance of freezing resistance in interspecific F_1 hybrids of *Eucalyptus'*, *Theoretical and Applied Genetics*, vol. 83, pp. 126-135.

Turelli, M & Moyle, LC 2007, 'Asymmetric postmating isolation: Darwin's corollary to Haldane's rule', *Genetics*, vol. 17, pp. 1059-1088.

Turnbull, JW 1999, 'Eucalypt plantations', *New Forests*, vol. 17, pp. 37-52.

van Tuyl, JM & de Jeu, MJ 1997, 'Methods for overcoming interspecific crossing barriers', in: Shivanna, KR & Sawhney, VK (eds.), *Pollen biotechnology for crop production and improvement*, Cambridge University Press, Cambridge and New York, pp. 273-292.

van Tuyl, JM & Lim, K-B 2003, 'Interspecific hybridization and polyploidisation as tools in ornamental plant breeding', *Acta Horticulturae*, vol. 612, pp. 13-22.

Verryn, SD 1993, 'Research and development report compiling estimates of genetic parameters in open-pollination and diallel trials', *CSIR Report For-Dea 601*, CSIR, Pretoria.

Verryn, S, Parfitt, R & Snedden, C 1997, 'A progress report on the programme to provide commercialised improved germplasm and improved raw material to enhance competitiveness in industrial forestry', *CSIR Report ENV/P/I 97024*, CSIR, Pretoria.

Verryn, SD 2000, 'Eucalyptus hybrid breeding in South Africa', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 191-199.

Verryn, SD & Geerthsen, JMP 2007, *Matgen 7.0. A BLP package for index selection in tree breeding*, CSIR, Pretoria.

Vigneron, Ph., Bouvet, J-M, Gouma, R, Saya, A, Gion, J-M & Verhaegen, D 2000, 'Eucalypt hybrids breeding in Congo', in: Dungey, HS, Dieters, MJ and Nikles, DG (eds.), *Hybrid Breeding and Genetics of Forest Trees. Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane, pp. 14-26.

Volker, PW 2002, *Quantitative genetics of Eucalyptus globulus, E. nitens and their F_1 hybrid*, PhD thesis, University of Tasmania.

Volker, PW, Potts, BM & Borralho, NMG 2008, 'Genetic parameters of intra- and inter-specific hybrids of *Eucalyptus globulus* and *E. nitens*', *Tree Genetics & Genomes*, vol. 4, no. 3, pp. 445-460.

Vuylsteke, M, Kuiper, M & Stam, P 2000, 'Chromosomal regions involved in hybrid performance and heterosis: their AFLP ®-based identification and practical use in prediction models', *Heredity*, vol. 85, pp. 208-218.

Waara, S & Glimelius, K 1995, 'The potential of somatic hybridization for crop breeding', *Euphytica*, vol. 85, pp. 217-233.

Walsh, B, 2007, 'Lecture 4, Heritability', Lecture presented at the Summer Institute on Statistical Genetics, Seattle, June 2007, viewed 28 July 2008, http://nitro.biosci.arizona.edu/zdownload/talks/2006/Seattle06-Lecture04.pdf.

Walsh, B & Lynch, M *unpublished*, 'Chapter 12. Selection and Crossbreeding I: basic concepts' in *Evolution and selection on quantitative traits'*, viewed 28 July 2008, http://nitro.biosci.arizona.edu/workshops/Aarhus2006/pdfs/WLChapter12.pdf.

Wang, TL, Hagqvist, R & Tigerstedt, PMA 1996, 'Growth performance of hybrid families by crossing selfed lines of *Betula pendula* Roth', *Theoretical and Applied Genetics*, vol. 92, pp. 471-476.

Warren, M 1884, 'On the etymology of hybrid (lat. Hybrida)', *American Journal of Philology*, vol. 5, no. 4, pp. 501-502.

Wei, M, van der Steen, HAM, van der Werf, JHJ & Brascamp, EW 1991, 'Relationship between purebred and crossbred parameters. I Variances and covariances under the one-locus model', *Journal of Animal Breeding and Genetics*, vol. 108, pp. 253-261.

White, TL & Hodge, GR 1989, *Predicting breeding values with applications in forest tree improvement*, Kluwer Academic Publishers, Dordrecht.

White, T 2001, 'Breeding strategies for forest trees: concepts and challenges', *Southern African Forestry Journal*, vol. 190, pp. 31-42.

Williams, CG & Savolainen, O 1996, 'Inbreeding depression in conifers: implications for breeding strategy', *Forest Science*, vol. 42, no. 1, pp. 102-117.

Wright, JW 1976, *Introduction to forest genetics*, Academic Press, New York.

Wu, R & Li, B 2000, 'A quantitative genetic model for analysing species differences in outcrossing species', *Biometrics*, vol. 56, pp. 1098-1104.

Wu, R, Ma, Ch-X, Wu, J, Fang, W & Casella, G 2004, 'Population genetics of heterosis: effects of Hardy-Weinberg disequilibrium', *Silvae Genetica*, vol. 63, no. 1, pp. 1-6.

Xiang, B and Li, B 2001, 'A new mixed analytical method for genetic analysis of diallel data', *Canadian Journal of Forest Research*, vol. 31, pp. 2252-2259.

Xiang, B and Li, B 2003, 'Best linear unbiased prediction of clonal breeding values and genetic values from full-sib mating', *Canadian Journal of Forest Research*, vol. 33, pp. 2036-2043.

Xu, ZC & Zhu, J 1999, 'An approach for predicting heterosis based on additive, dominance and additive x additive model with environment interaction', *Heredity*, vol. 82, pp. 510-517.

Zamir, D & Tadmor, Y 1986, 'Unequal segregation of nuclear genes in plants', *Botanical Gazette*, vol. 147, no. 3, pp. 355-358.

Zhuang, Y & Adams, KL 2007, 'Extensive allelic variation in gene expression in Populus F_1 hybrids', *Genetics*, vol. 177, pp. 1987-1996.

Zobel, B & Talbert, J 1984, *Applied Forest Tree Improvement*, Waveland Press, Prospect Heights.

PUBLICATION AND PRESENTATION RESULTING FROM THE RESEARCH

Paper presented at an international conference, the IUFRO Working Group 2.08.03 "Eucalypts and Diversity: Balancing Productivity and Sustainability", 22-26 October 2007, Durban, South Africa:

Title: Investigation of the efficiency of selection indices in Non-Hardy-Weinberg eucalypt hybrid populations

Authors: Hettasch, M.H., Verryn, S.D. and Roux, C.Z.

The authors were requested to submit the conference paper to the Southern Hemisphere Forestry Journal as part of a special edition on the conference. The edition will present selected papers from the conference.

Paper submitted to a peer-reviewed journal:

Title: Investigation of the efficiency of selection indices in non-Hardy-Weinberg eucalypt hybrid populations

Authors: Hettasch, M.H., Verryn, S.D. and Roux C.Z.

Paper: Southern Hemisphere Forestry Journal – recently renamed, now: Southern Forests: a Journal of Forest Science

Submission: March 2008

Investigation of the efficiency of selection indices in non-Hardy-Weinberg eucalypt hybrid populations

M H Hettasch[1], S D Verryn[1] and C Z Roux[2]

[1]Natural Resources and the Environment, CSIR, PO Box 395, Pretoria, 0001,
[2]University of Pretoria. Pretoria, 0002, South Africa,
E-mail: mhettasc@csir.co.za

ABSTRACT

The prediction of the performance of hybrid plants as production material is complicated by the genetic structure of the hybrid population. Fundamental quantitative genetic concepts are defined with respect to disomic, randomly fertilised populations which are in linkage equilibrium. Hybrid populations of crosses are, however, in linkage disequilibrium and in Hardy-Weinberg disequilibrium and do not meet the assumptions of quantitative genetic theory. In practice, hybrids are often analysed in the same way as pure species, but the results may be questionable, as the underlying assumptions are not met. There is a need to understand how best to predict the genetic worth or ranking in these non-Hardy-Weinberg hybrid populations. The aim of this study is to investigate the efficiency of selecting F_1 clones from F_1 hybrid family trials and to determine whether the conventional selection techniques can be improved or adapted for such selection. The study is based on a number of different eucalypt hybrid family and corresponding clonal trials. Best linear unbiased prediction rankings of forward selection and clonal forward selection in family trials are compared to the performance in the clonal trials. The results indicate that a best linear unbiased prediction clonal forward selection technique is relatively efficient in predicting clonal performance in the three eucalypt hybrid types that have been investigated.

Keywords: Hybrid, breeding value, genetic value, BLUP, prediction, Eucalyptus hybrids

INTRODUCTION

Hybrids are widely used in forestry as they allow for the exploitation of complementarity, adaptability and better parent heterosis (Kain, 2003). Eucalypt hybrids are planted on a large scale in Brazil and the Congo. There are also reasonably large plantations of eucalypt hybrids in China, Indonesia and South Africa and smaller areas of hybrids in countries in Asia and South America (Potts and Dungey, 2004). The hybrids are usually deployed by means of vegetative propagation, as the production of control-pollinated hybrid seed is usually not feasible in commercial quantities (Shelbourne 2000). Research over more than a century has contributed significantly towards a better understanding of the complexities of hybrids and has contributed to the successful commercial deployment of hybrids but there is still very little known about the quantitative genetic aspects and therefore selection theory for hybrids is not well developed.

Most fundamental quantitative genetic concepts are defined with respect to disomic, randomly mating populations that are in linkage equilibrium (Falconer, 1989; Gordon, 1999). However, hybrid populations are in Hardy-Weinberg disequilibrium as well as in linkage disequilibrium (Falconer, 1989). The Hardy-Weinberg equilibrium is restored after one generation of random mating, but it may take many generations to restore linkage equilibrium after hybridization (Falconer, 1989). The disequilibrium leads to instability of the genetic structure of the population from one generation to the next (Kain, 2003). This presents a problem for selection theory, as it relies on genetic equilibrium to reliably predict the outcome of forward selection and recombination (Kain, 2003). Furthermore it is questionable whether additive genetic variance, dominance genetic variance and therefore narrow-sense heritability can be applied to hybrid populations (Gordon, 1999).

It is believed that researchers generally do not take these limitations into account when performing selections in hybrid populations. Hybrids are often analysed in the same way as pure species. Selection may be based on family mean rankings, individual tree selections or techniques such as best linear unbiased predictions (BLUP) that have been developed for pure species. BLUP is expected to give the best predictions as it gives the

minimum error variance, maximised probability of selecting the better of two candidates and maximised expected genetic gain for a fixed number of selections (Henderson, 1984, White and Hodge, 1989). The reliability and accuracy of these selections in the hybrid population case are unknown. The aim of this study was to obtain a measure of the reliability of BLUP for the selection of F_1 clones from F_1 hybrid family trials using field data.

MATERIAL AND METHODS

Field trials

Three eucalypt hybrids were studied representing intraseries, interseries and intersectional crosses. All parent species belong to the subgenus *Symphyomyrtus*. *E. grandis* × *E. saligna* (G x S) represents a cross between two closely related species, as both belong to the same section (*Transversaria*) and series (*Transversae*) (Brooker, 2000). *E. grandis* × *E. urophylla* (G x U) represent an interseries cross as *E. urophylla* belongs to the series *Annulares* and the section *Transversaria* (Brooker, 2000). *E. grandis* × *E. camaldulensis* (G x C) represents an intersectional cross as *E. camaldulensis* belongs to the section *Exsertaria* (Brooker, 2000).

Three hybrid family trials were selected from the eucalypt hybrid breeding programme of the Council of Scientific and Industrial Research (CSIR) in South Africa. The trials included an *E. grandis* × *E. saligna* trial at the Kiaat plantation, an *E. grandis* × *E. urophylla* trial at the Grootboom plantation and an *E. grandis* × *E. camaldulensis* trial at the Salique plantation. A total of 36 *E. grandis* × *E. saligna* full-sib families were included in the hybrid family trial at Kiaat. The families resulted from crosses between 16 *E. grandis* mothers and five *E. saligna* pollen sources. A total of 69 full-sib families were included in the G x U hybrid progeny trial. The families resulted from crosses between 12 *E. grandis* mothers and 10 *E. urophylla* pollen sources. A total of 34 full-sib families were included in the G x C hybrid trial. The families resulted from crosses between 10 *E. grandis* mothers and seven *E. camaldulensis* pollen sources. In addition to the full-sib relationships there were also a number of half-sibs in all family trials.

Within these hybrid family trials selections had been made in the mid 1990s for the establishment of clonal trials. The selections were based on individual tree values. The trials were felled and coppice was collected for the propagation of macro-cuttings. In the G x S hybrid trial at Kiaat clonal selections were made by ranking the tree stem volumes (m3) at 74 months followed by independent culling for disease (Pierce, 1994a). The clonal trials at Dukuduku and Westfalia were established from these selections. In the G x U trial at Grootboom, selection was based on ranking of volume at 54 months of age (Pierce, 1993). The cuttings were planted in clonal trials at Dukuduku and Port Durnford. In the G x C hybrid trial at Salique selections were made using 100 month assessments of diameter at breast height (Pierce, 1994b). The cuttings were planted in a clonal trial at Bossieskop. All trials consisted of a mixture of half-sibs and unrelated clones. Table 1 presents information on the trial design, site conditions and age of assessment of the trials.

Table 1 Trial design, site description and age of assessment of hybrid family trials and clonal trials.

Number	**1010603 EA0005**	**1010806 EA0011**	**1010806 EA0011**	**1010603 EA0004**	**1010806 EA0006**	**1010806 EA0006**	**1010603 EA0003**	**1010806 EA0010**
Site	**Kiaat**	**Duku-duku**	**West-falia**	**Groot-boom**	**Port Durnford**	**Duku-duku**	**Salique**	**Bossies-kop**
Hybrid	GxS	GxS	GxS	GxU	GxU	GxU	GxC	GxC
Trial type	Hybrid family	Clonal	Clonal	Hybrid family	Clonal	Clonal	Hybrid family	Clonal
Design	RCB	RCB	RCB	RCB	RCB	RCB	RCB	Alpha Lattice
Replications	6	10	10	5	15	15	5	6
Plot size	1x6	Single	Single	1x6	Single	Single	1x6	1x5
Families/ Clones	36	24	24	69	67	67	32	37
Espacement (m)	2.7x2.7	3x3	3.5x3.5	2.7x2.7	3.5x3.5	2.7x2.7	2.7x2.7	2.7x2.7
Latitude	23º46’S	28º13’S	23º44’S	23º43’S	28º53’S	28º21’S	24º45’S	25º24’S
Longitude	30º02’E	32º25’E	30º06’E	30º10’E	32º50’E	32º15’E	30º58’E	30º30’E
Altitude(m)	900	40	950	750	130	70	670	802
Rainfall (mm)	1200	950	1342	1300	1461	973	750	850
Soil	Hutton	Sandy/ Fernwood	Hutton	Hutton	Hutton	Sandy/ Fernwood	Sandy	Hutton
Date planted	04/1988	08/1997	06/1997	02/1986	08/1993	08/1993	03/1995	11/1997
Assessed (months)	74	43	38	64	48	48	50	34

Trait selection and assessment

One of the drawbacks of using historical field data is that the clonal trials represent only ortets that have been selected for their good performance with respect to the selected traits. Therefore it can not be determined whether poor performance of ortets in these traits will also result in the poor performance of clones, i.e. there is an element of bias. There may also be a lack of significant differences between clones. In order to overcome these problems it was decided to focus the study on stem form, a trait that has not been included in the selection criteria which were originally used to select the clones from the hybrid trials. Stem form was assessed on an 8-point subjective scale where one represented poor, crooked, stem form and eight excellent, perfectly straight, stem form.

BLUP methods to predict clonal performance

Two different BLUP selection techniques were performed to determine whether these methods were efficient for selecting clones from hybrid trials. In the first approach, breeding values were predicted. With this method only the additive genetic potential of each individual is estimated by combining both parental and within family breeding value components (Xiang and Li, 2001). In the second approach the genetic values were predicted. This method includes both additive and non-additive genetic values and accounts for the total genetic effects (Xiang and Li, 2003).

BLUP predictions of breeding values and genetic values for stem form were obtained using Matgen Version 7.0 (Verryn and Geerthsen, 2007). The forward selection option of Matgen was used to obtain the breeding values and the clonal forward selection option to obtain genetic values. The programme Matgen includes the non-additive genetic variance in predicting clonal performance.

As the proportion of non-additive genetic variance ($\sigma^2_{na} / \sigma^2_g$) in the populations was unknown, an estimate of 37% was used as well as an estimate of 20%. The first is an estimate obtained for an *Eucalyptus grandis* cloned breeding population at five years of age (Snedden, 2001). The second is an approximation of an estimate of the non-additive genetic variation that was reported for volume and circumference in an *E. grandis* × *E. urophylla* hybrid population at three years of age (Bouvet and Vigneron, 1996)

Both the forward selection and the clonal forward selection options require the input of variance components. The Mixed Model Least-Squares and Maximum Likelihood program (LSMLMW & MIXMDL PC-2 Version) developed by Harvey (1990) was used to determine the variance components.

Unless otherwise stated data was analysed using SAS/STAT software, of the SAS System for Windows (Copyright © 2002-2003 SAS Institute Inc.). The data was corrected for fixed effects where significant and was standardised.

Selection indices

A series of combined indices were designed which placed different weights on family means versus single tree values. The index value for e.g. for 90% weight on the family mean and 10% weight on the individual performance was calculated using the following formula:

Index value= (0.9 x family mean) + (0.1 x single tree value).

Similarly, indices were calculated for 0%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100% weight on single tree values.

Clonal performance

The least square means of the stem scores of the clones were calculated using the GLM procedure of SAS. The least square means were chosen to represent the clonal performance as they are free of bias due to unequal representation across blocks and plots (White and Hodge, 1989). For the G x S and G x U hybrids the information of the two clonal trials was combined after correcting for site effect to obtain one correlation coefficient.

The correlation of clonal versus predicted performance

The correlation of clonal performance versus the predicted performance was used as a measure of reliability of the predictions. The Pearson correlation coefficients were calculated. The breeding values, the genetic values and the constructed indices were correlated with the performances of the clones.

RESULTS AND DISCUSSION

Heritabilities

Narrow-sense heritability only accounts for the additive genetic variation that is transmitted from one generation to the next. Although narrow-sense heritabilities may not be appropriate for hybrids, they were calculated for the purpose of obtaining variance components for the ranking the individual trees of the hybrid trials. Estimates of narrow-sense heritability are presented in Table 2. The estimates are expected to be on the one hand downwards biased due to relatedness of crosses which was ignored, and on the other hand upwards biased due to the wideness of the crosses.

Table 2 Narrow-sense heritabilities of three eucalypt hybrid family trials.

Trial	Site	Type	Number of		h^2	Standard error
			families	trees per family		
1010603EA0005	Kiaat	GxS	36	35-36	0.410	0.106
1010603EA0004	Grootboom	GxU	69	8-21	0.337	0.062
1010603EA0003	Salique	GxC	34	10-80	0.145	0.050

BLUP individual breeding values versus individual genetic values

The correlations between the observed values and breeding values and two different genetic values are presented in Table 3. The correlations obtained for the GxC hybrid were all non-significant. This may be ascribed to the low number of pairs of values that were correlated (34).

Table 3 Correlations between observed and three different types of predicted values.

Hybrid trial (predicted)	Clonal trial (Observed)	Hybrid type	n	Correlations (significance)		
				BV*	GV 20% non-add.**	GV 37% non-add.**
Kiaat	Dukuduku and Westfalia	GxS	46	0.368 (p=0.012)	0.392 (p=0.0071)	0.404 (p=0.0054)
Grootboom	Port Durnford and Dukuduku	GxU	64	0.333 (p=0.0072)	0.346 (p=0.0052)	0.342 (p=0.0056)
Salique	Bossieskop	GxC	34	0.294 (ns.)	0.309 (ns.)	0.322 (ns.)

*. breeding values (BV)

** genetic values with a proportion of 20% (GV 20%) and 37% (GV 37%) non-additive genetic variance

Note: The level of significance is presented in brackets with 'ns.' indicating non-significant correlations

Although the correlations between observed and predicted stem form values were all positive, the values were relatively low with 0.404 being the highest. These are far lower than the theoretical correlations between predicted and actual genetic values that have been calculated by the programme Matgen (which varied between 0.7 and 1.0). The difference needs to be investigated further, but it is believed that the fact that the family trials and the clonal trials were not planted on the same site and were not measured at the same age may play a role.

For all trials the correlation between the observed clonal values and the predicted genetic values (for both a proportion of non-additive genetic variation of 20% and 37%) were higher than the correlation with the predicted breeding values.

These results indicate that the genetic values give better predictions than the breeding values despite the fact that the actual proportion of non-additive genetic variation was assumed for the populations. As this is also true for the hybrid between the closely related species *E. grandis* and *E. saligna*, it is likely that the clonal selection option will also give better results in selecting clones in pure species. The fact that clonal selection gives better results agrees with the theoretical expectations. The genetic gain achieved by the selection of parents for seed production is based almost entirely on the additive component of genetic variation whereas the genetic gain achieved by the selection of ortets for clonal production is based on both additive and non-additive effects, as the

clones retain the effect of allelic interaction (Mullin and Park, 1992). Therefore selection of clones should be based on both additive and non-additive genetic components.

The influence of the proportion of non-additive genetic variance on the individual genetic values

As the actual proportion of non-additive genetic variation of the populations studied was not known, two different assumptions of the proportion were made. These were used as parameters in the clonal forward selection option of the programme Matgen Version 7.0.

The value of 37% that was based on findings for stem form in *E. grandis* (Snedden, 2001) gave slightly higher correlations with the clonal values than the individual genetic values that were based on the 20% value with the exception of the G x U hybrid where a the correlation was slightly lower. This may indicate that 37% is a better reflection of the true proportion of non-additive genetic variation for stem form than 20% for the G x S and G x C hybrids. Despite a relatively large increase (17%) in the proportion of non-additive genetic variance, the improvement in correlation was relatively small.

Comparison of BLUP values and selection constructed index values

A series of index values was calculated which place different weights on family mean versus individual tree values in the hybrid trial. These indices were calculated to explore what may be an optimal ratio between family mean versus individual value and to compare these correlations with the correlations obtained with the BLUP values as measure of reliability of selections.

For each hybrid type a graph was drawn of the correlations obtained for each constructed selection index. For comparison purposes the individual breeding and genetic values were added as horizontal lines. The relationship between the correlations of the clonal performance of the clones with the predicted performance of the hybrid trials are presented in Figure 1 (G x S), Figure 2 (G x U) and Figure 3 (G x C).

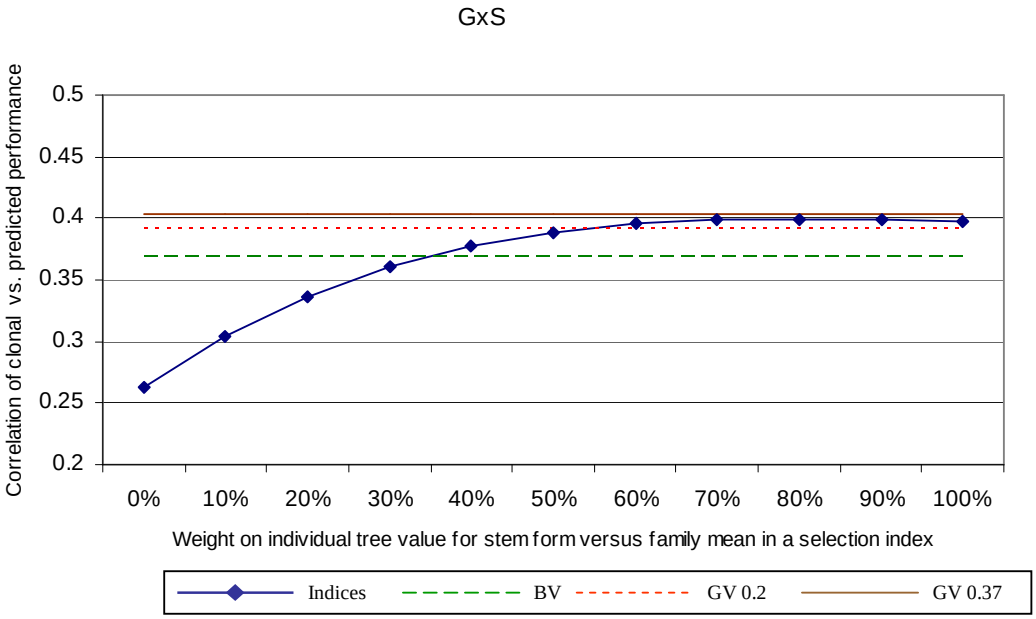


Figure 1 The relationship between the correlations of the performance of the clones in the G x S clonal trials at Dukuduku and Westfalia with the predicted performance from the hybrid trial at Kiaat. The graph shows a number of different selection scenarios which place different weight on individual tree selection versus family selection. The horizontal lines represent the correlation of the clonal performance with three different BLUP selection scenarios: the breeding values (BV), genetic values with a non-additive genetic variance proportion of 20% (GV 0.2) and a non-additive genetic variance proportion of 37% (GV 0.37).

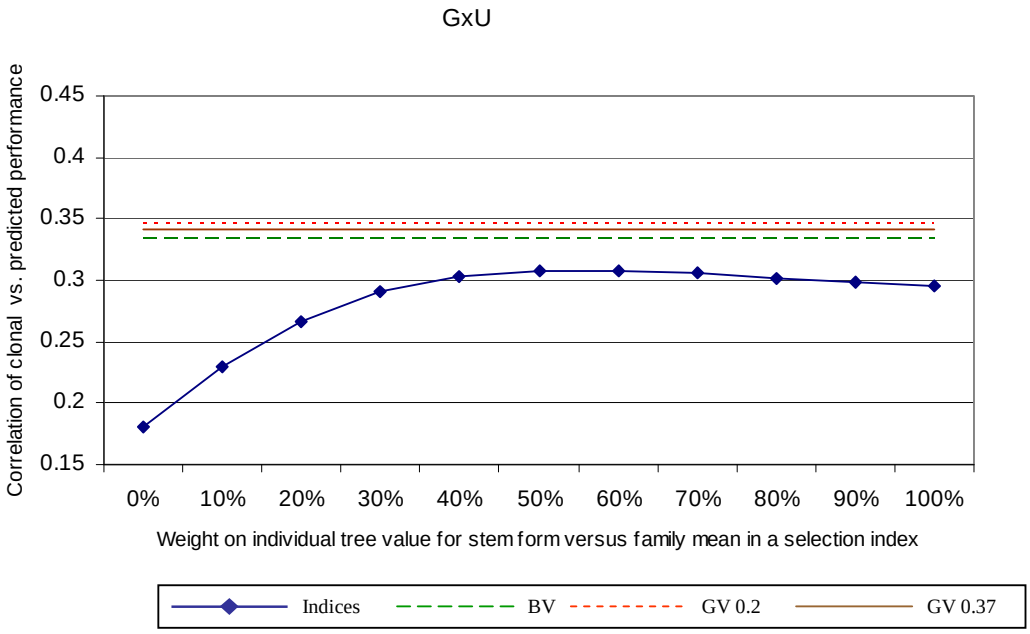


Figure 2 The relationship between the correlations of the performance of the clones in the GxU clonal trials at Dukuduku and Port Durnford with the predicted performance from the hybrid trial at Grootboom. The horizontal lines represent the correlation of the clonal performance with three different BLUP selection scenarios.

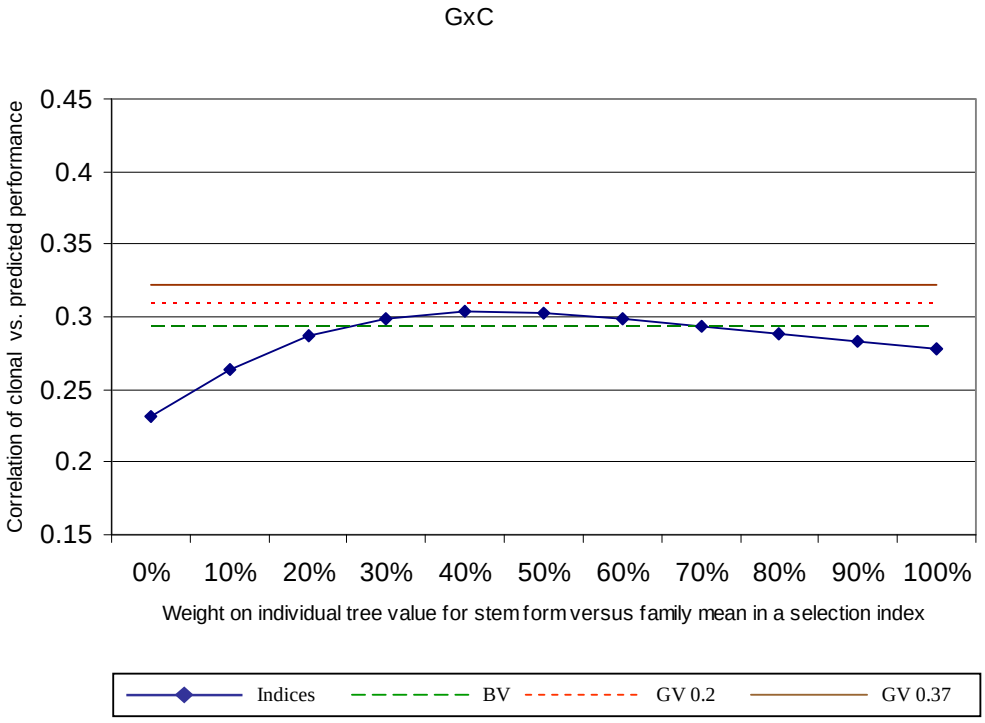


Figure 3 The relationship between the correlations of the performance of the clones in the GxC clonal trial at Bossieskop with the predicted performance from the hybrid trial at Salique. The horizontal lines represent the correlation of the clonal performance with three different BLUP selection scenarios.

The graphs presenting the correlations between the clonal stem measurements and the predictions made using the series of constructed selection indices, show a consistent pattern of relatively low correlations when the predictions are based on the family means only. The correlations increase with increasing weight being placed on individual selection. The correlations decrease slightly when selections are based on single tree values only. The optimal proportion of weight on single tree values versus family means differs for the different hybrid types. In G x S hybrid the correlations increase up until about 70% of the weight placed on the single tree values. In the G x C hybrid the optimal weight lies around 40% on single tree values and for the G x U hybrids on 50%. These differences reflect the differences in the heritabilities of the hybrid trials, with a low heritability for the G x C trial at Salique (0.145), and the highest heritability for the G x S trial at Kiaat (0.410).

The genetic values compare favourably with the highest correlations obtained with the constructed selection indices. This result indicates that the genetic values with a non-additive variation proportion of 37% gives close to optimal correlations. In the GxU and G x S trials the genetic values are higher than the best predictions of the constructed indices. This result is unforeseen, as it is expected that the correlations would not be higher than the best prediction of the constructed indices. In the G x C hybrid this may be due to error as the correlations are not significant in this case. Another possible reason is that the BLUP predictions are better than the selection indices due to the fact that the data was unbalanced, and Matgen accounts for this, whilst the constructed indices did not do so.

CONCLUSION

This study used field data to investigate the reliability of BLUPs applied to hybrid populations. Despite the drawbacks of the data that was available for such a study, the results showed a consistent trend of BLUPs for total genetic value giving near-optimal predictions if the correlation coefficients were used as measure of reliability of clonal selection in hybrids. These results indicate that BLUP may be suitable for clonal selection in the hybrids studied.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr Dominic Kain (Australian National University) and Prof Peter Geerthsen (CSIR) for their support. The research was supported by the Australian Centre for International Agricultural Research (ACIAR), the Innovation Fund of the South African Department of Science and Technology and the CSIR.

REFERENCES

Brooker MIH (2000) A new classification of the genus *Eucalyptus* L'Hér (Myrtaceae). *Australian Systematic Botany* 13: 79-148.

Falconer DS (1989) *Introduction to Quantitative Genetics.* Longman Scientific & Technical , Harlow, England.

Gordon IL (1999) Quantitative genetics of intraspecies hybrids. *Heredity* 83: 757-784.

Harvey W (1990) *User's guide for LSMLMW and MIXMDL PC-2 version.*

Henderson CR (1984) *Applications of Linear Models in Animal Breeding*. University of Guelph.

Kain DP (2003). *Genetic parameters and improvement strategies for the Pinus elliottii var. elliottii x Pinus caribaea var. hondurensis hybrid in Queensland, Australia*. PhD thesis, Australian National University, 460p.

Mulin TJ and Park YS (1992) Estimating genetic gains from alternative breeding strategies for clonal forestry. *Canadian Journal of Forest Research* 22: 14-23.

Pierce BT (1993) Report on the delivery of cuttings from two Eucalyptus hybrid family trials at Grootboom and Salique. *CSIR Report FOR-DEA 608*, CSIR, Pretoria.

Pierce BT (1994a) The six year results of a Eucalyptus grandis x. E.saligna hybrid family trial series planted at Kiaat and Morgenzon plantations plus selections. *CSIR Report FOR-DEA 839*, CSIR, Pretoria.

Pierce BT (1994b) Growth evaluation plus the delivery of cuttings from a series of Eucalyptus hybrid family trials planted at Pan and Salique. *CSIR Report FOR-DEA 751*, CSIR, Pretoria.

Potts BM and Dungey HS (2004) Interspecific hybridization of Eucalyputs: key issues for breeders and geneticists. *New Forests* 27: 115-138.

Shelbourne CJA (2000) Some insights on hybrids in forest tree improvement. In: Dungey, H.S., Dieters, M.J. and Nikles, D.G. (Eds.) Hybrid Breeding and Genetics of Forest Trees. *Proceedings of a QFRI/CRC-SPF Symposium*, 9-14 April 2000, Noosa, Queensland, Australia. Department of Primary Industries, Brisbane. pp. 53-62.

Snedden CL (2001) *Broad and Narrow sense heritabilities in a cloned open pollinated Eucalyptus grandis breeding population*. MSc thesis, University of Pretoria, 135p.

Verryn SD and Geerthsen JMP (2007) *Matgen 7.0. A BLP package for index selection in tree breeding*. CSIR, South Africa.

White TL and Hodge GR (1989) *Predicting breeding values with applications in forest tree improvement*. Kluwer Academic Publishers, Dordrecht.

Xiang B and Li B (2001) A new mixed analytical method for genetic analysis of diallel data. *Canadian Journal of Forest Research* 31: 2252-2259.

Xiang B and Li B (2003) Best linear unbiased prediction of clonal breeding values and genetic values from full-sib mating. *Canadian Journal of Forest Research* 33: 2036-2043.