# DEVELOPING A GENETIC EVALUATION SYSTEM FOR FOREST TREE IMPROVEMENT – THE MAKING OF TREEPLAN®

**R.J. Kerr[1], G.W. Dutkowski[2], L.A. Apiolaza[2], T.A. McRae[3] and B. Tier[1]**

[1]Animal Genetics and Breeding Unit, University of New England, Armidale NSW 2350, Australia
[2]Cooperative Research Centre for Sustainable Production Forestry and School of Plant Science, University of Tasmania, GPO Box 252-12, Hobart, Tasmania 7001, Australia
[3]Southern Tree Breeding Association, PO Box 1811, Mount Gambier, South Australia 5290, Australia

## INTRODUCTION

Some of the features of tree breeding have made the application of best linear unbiased prediction (BLUP) less common than in animal breeding. Breeding programs are in their early generations, with simple shallow pedigrees. Families often result from open pollination, where the maternal parent is known but pollen donors are not. Data are collected from designed trials and trees are often not subject to culling ; thus there are equivalent amounts of data for all individuals in any trial. The magnitude of genotype by environment interactions (GxE) is commonly unknown, except in a large environmental range. The number and type of traits measured is rapidly evolving, as wood quality traits assume greater importance.

Historically there has been an emphasis on experimental design features (e.g. replicate, incomplete block and plot effects) over modelling genetic components and single-generation, single-site and single-trait mixed models have been the norm. Simple pedigrees and genetic material balanced across sites has allowed the use of straightforward methods of analysis, mainly BLP without a numerator relationship matrix. Models corresponding to sire models have been used with a second stage to predict within-family values.

There are several aspects of modern tree improvement prompting greater use of BLUP to predict genetic values. Firstly, breeding programs are progressing and now contain data spanning several generations of selection, individual programs with different samples of the same base population are being consolidated and it is important to account for the effects of selection. Secondly, many programs are now making the transition to overlapping generations, where a proportion of all breeding activities is performed each year, and all families are not tested at all test sites at the same time. Finally, there is a need for integrating all data between trees and between traits, making it easier to monitor the genetic progress of the breeding programs.

The Southern Tree Breeding Association (STBA) is the national breeding cooperative for *Pinus radiata* and *Eucalyptus globulus* in Australia. The two species, with a total estate of 1,027,900 ha, are mostly used for solid wood products and pulp and paper production. Currently, the STBA is collecting performance data on third generation progeny in *P. radiata* and second generation progeny in *E. globulus*. The STBA has worked with the Animal

Genetics and Breeding Unit (AGBU) and the Cooperative Research Centre for Sustainable Production Forestry (CRC-SPF) to develop a genetic evaluation system for forest tree improvement (dubbed TREEPLAN®). This paper discusses the problems and concerns with BLUP applications in forestry and how TREEPLAN® deals with the issues.

## APPLYING BLUP TO FORESTRY

**Heterogeneous variances**. The STBA collects data across a diverse range of site types and age classes. Some traits are or have been assessed in a variety of ways. For example, growth has been measured as total tree height, stem diameter at breast height (DBH) and tree volume ; stem form has been assessed using several scales with different precision. The variance of performance traits such as growth (using any of the surrogate measures) is not independent of the site mean. Sites with larger trees (either older or growing in a more productive site) have larger variances. A linear transformation of the data such that the phenotypic variance is unity is an approach often used in animal breeding to make variances homogenous. A disadvantage of this approach for tree breeding is that a constant heritability would need to be assumed across all sites, despite some sites being more homogeneous than others. Tree breeders have the benefit of large designed trials that provide good estimates of design, genetic and residual variances and correlations specific to each site. TREEPLAN® takes advantage of the availability of these estimates by : (i) transforming the data and variances on a site by site basis prior to joint analysis such that the additive genetic variance equals 1 ; and (ii) using the within site error and design factors (e.g. plot and incomplete block) variances in the BLUP analysis.

**Trait mapping**. The heterogeneity of forestry data can, potentially, lead to an extremely large number of traits. The mapping of multiple selection criteria to a small number of traits is a feature of TREEPLAN®, allowing a reduction in the number of traits (e.g. multiple assessments at different ages or different scales), to cope with GxE, and to provide flexibility to the system for adding new traits. For example, a STBA member organisation may want genotypes specific for low fertility soils in Tasmania. These genotypes may differ to elite genotypes specifically adapted to high fertility soils in Western Australia. Mapping of the traits in TREEPLAN® considers members' requirements regarding specific geographical and environmental subclasses, either creating environmental subclasses or merging environments and ignoring GxE.

**Genetic groups.** In forestry parents of first generation progeny are typically trees from native (or plantations) stands sampled from many different geographical regions that represent different provenances or races. Because the provenances are quite genetically distinct it is important to assume that $E(\boldsymbol{g}) \neq 0$, where $\boldsymbol{g}$ is the vector of genetic values. TREEPLAN® relates all foundation parents on the basis of their original provenance to genetic groups. The modified mixed model equations of Quaas (1988) are used to derive solutions to $\boldsymbol{g}$.

**Clonal data.** Individual trees can be replicated using vegetative propagation (cuttings, tissue culture or somatic embryogenesis). Clonal tests are quite common in *P. radiata* and are also used in some Eucalypt breeding programs. TREEPLAN® currently treats clones as the same individual with a genetic relationship of 1 and matches unique clone identities to a single genotype. Future versions of TREEPLAN® may consider predicting genetic values, including additive and non-additive genetic effects, for individual clones, recognising the potential for somaclonal variation.

**Open pollinated seed.** The collection of seed from known female parents with unknown pollen donors creates half-sibs analogous to a sire mated to many females in a single mating paddock without "mothering up". Trees can also be self-fertile, generating pedigrees where two progeny maybe selfed sibs (i.e. both progeny are the result of selfing), a selfed sib and an outcrossed sib, full sibs and half-sibs. In the *E. globulus* breeding program the majority of the progeny are derived from such open-pollinated seedlots collected from native forest stands. Until many more second-generation progeny (from controlled pollinations) are included in the analysis, the accuracy of breeding value prediction is dependent on how well the relationship coefficients between sibs of open-pollinated trees can be defined. Dutkowski (2001) has outlined simple rules to modify the NRM when a selfing rate in native stands is assumed. These rules can be further extended to account for the equilibrium level of inbreeding in the stand and the level of coancestry in the trees local to the female parent from which the seeds were collected. Sparse stands of trees are expected to have a higher level of inbreeding among the progeny than dense ones. TREEPLAN® will implement these rules in the future.

**Table 1. Description of test data set for current *E. globulus* evaluation**

| | | | |
|---|---|---|---|
| Trials | 20 | Genotypes | 88678 |
| Measured criteria | 35 | Stems | 87629 |
| Analysis traits | 7 | Parents | 1065 |
| Genetic groups | 27 | Non-parents | 87613 |
| Families | 1111 | | |

## AN EXAMPLE EVALUATION

TREEPLAN® has been used to analyse a sample dataset consisting of diameter at breast height over bark (DBH) from 20 test sites (18 first generation and two second generation) ; core basic density (BD) from a 10 % sample in two second generation trials and a 20 % sample in two first generation trials; and pilodyn penetration (PILO, a proxy for wood density) from a 20 % sample of trees in 11 first generation trials. As BD and PILO show little evidence of either GxE or genotype by age interaction, the measured traits were mapped to a single trait. However, DBH measurements were mapped into six growth traits based on three age classes (< 5, 5 - 10, and > 10 years) and four site types according to geographic region. These are preliminary mappings until more detailed analyses demonstrate the presence or absence of interactions with either age or site that could justify further partitioning. Given the flexibility of TREEPLAN®, any recommended changes to the national evaluation system could be easily implemented. For each trait on each site, non-significant design effects are removed, and the data transformed to a unit additive variance. Inter-trait correlations for random design effects are also estimated. While inter-trait genetic correlations are estimated they are not used directly, but only used to estimate correlations between the mapped traits. After transformation all data are combined and breeding values predicted on a unit scale for each mapped trait in a multi-trait analysis.

The STBA have adopted rolling front schemes for its *E. globulus* and *P. radiata* breeding populations. Trial data accumulate regularly throughout the year and need to be analysed quickly and efficiently. The TREEPLAN® system has been integrated with the STBA data management system and is operated via a web based interface. This facilitates the regular update of breeding and genetic values as information is collected. STBA members can make selection decisions for their deployment programs with the knowledge that the latest trial information has been included.

## CONCLUSION

The implementation of large scale individual model BLUP analyses in forest tree breeding involves a combination of both solving new problems and adaptation of features routinely used in animal breeding. Tree breeding programs are evolving in a way that will necessitate the adoption of BLUP analyses to maximize the return of investment in breeding. Trait mapping is essential to reduce data heterogeneity to a manageable problem. The development and adoption of tools like TREEPLAN® will facilitate the use of rolling front strategies with overlapping generations.

## REFERENCES

Borralho, N.M.G. and Dutkowksi, G.W. (1996) *Canadian J. Forest Res.* **28** : 987-993.
Dutkowski, G. (2001) IUFRO 2001.
Quaas, R.L. and Pollak, E.J. (1980) *J. Anim. Sci.* **51** : 1277-1287.
Quaas, R.L. (1988) *J. Dairy Sci.* **71** : 1338-1345.