CrossMark

ORIGINAL PAPER

Improving prediction accuracy and selection of open-pollinated seed-lots in *Eucalyptus dunnii* Maiden using a multivariate mixed model approach

Craig M. Hardner[1] · Adam L. Healey[1,2] · Geoff Downes[3] · Mónica Herberling[4] ·
Peter L. Gore[5]

Abstract

· *Key message* **Multivariate mixed models can be used to combine complex data into a single-step analysis to improve prediction accuracy of open-pollinated seed lots for all attributes and candidates, and identify elite seed lots.**

Handling Editor: Ricardo Alia

Contribution of the co-authors Craig M. Hardner: undertook the analysis, led the writing.
Adam L. Healey: assisted with writing.
Geoff Downe: assessed wood properties with NIR, reviewed manuscript.
Mónica Herbeling: managed Uruguay trials, reviewed manuscript.
Peter L. Gore: initiated project, managed Australian trials, reviewed manuscript.

✉ Craig M. Hardner
craig.hardner@uq.edu.au

Adam L. Healey
a.healey1@uq.edu.au

Geoff Downes
trooper1325@gmail.com

Mónica Herberling
monica.heberling@montesdelplata.com.uy

Peter L. Gore
pgore@seedenergy.com.au

[1] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St Lucia, QLD 4072, Australia

[2] Present address: HudsonAlpha Institute of Biotechnology, 601 Genome Way, 35806 Huntsville, AL, USA

[3] Forest Quality Pty. Ltd., P.O. Box 293, Huonville, TAS 7109, Australia

[4] Montes del Plata, Luis A. de Herrera 1248,WTC, Torre 3 Piso 9, Montevideo, Uruguay

[5] SeedEnergy, 2 Derwent Avenue, Margate, TAS 7054, Australia

· *Context* Data available for genetic selection may be complex and unbalanced; however, utilisation of all available information for prediction of genetic value may improve prediction accuracy to better identify elite candidates for selection.
· *Aims* This study aimed to develop, implement and evaluate a single-step multivariate mixed model for complex and unbalanced data and use the results to identify elite candidates.
· *Methods* Multivariate mixed models were developed and applied to a case study of seed-orchard open-pollinated *Eucalyptus dunnii* families grown in progeny trials in Australia and Uruguay to identify elite seed lots for biofuel utilisation. This approach combined all available data across trials and ages and included models of spatial variation to predict OP seed lot values for growth and wood quality attributes. Predictions were used to estimate response to selection and correlations between breeding values of parents predicted from their own performance and seed lot values predicted from the progeny trials.
· *Results* Prediction accuracy was highest for a single-step multivariate model. Prediction of seed lot values using this model indicated that selection of the best 12.5 % resulted in a gain of 30 % in cellulose content, and breeding value of parents predicted from own performance was only weakly correlated with seed lot performance in progeny trials.
· *Conclusion* Single-step multivariate approaches provide the most accurate prediction of genetic value for all attributes, for all candidates, and hence leads to greater selection response. In eucalypts, gain from selection using seed lot values predicted from progeny trials will be greater than from selection using breeding values of parents predicted from their own performance.

Keywords Wood properties · Generalised heritability · Repeatability · G × E · Genetic correlation

INRA SCIENCE & IMPACT   Springer

1 Introduction

Genetic improvement is a proven strategy for increasing forest profitability (Borralho et al. 1993; Greaves and Borralho 1996). Response to selection is directly proportional to accuracy of predicting the genetic value of selection candidates (Falconer and Mackay 1996). Data available in forestry for prediction of genetic value may be complex, including multiple traits and locations (Burdon 1977), repeated measures (Greaves et al. 2003) and possibly local spatial variation (Dutkowski et al. 2006; Gilmour et al. 1997). In addition, data are likely to be unbalanced with not all candidates assessed for all traits in all trials. Correlated traits, such as pilodyn penetration (e.g. Greaves et al. 1996) or NIR (e.g. Downes et al. 2010), which are relatively inexpensive to assess, may also be used to predict genetic value for objective traits that are otherwise expensive to assess.

Mixed linear models that treat genetic effects as random provide a flexible framework for combining all available information from complex data structures to support best (i.e. lowest variance), unbiased prediction (BLUP) of genetic value (Henderson 1984; Thompson and Meyer 1986; White et al. 2007). In addition, prediction of accurate genetic values has been demonstrated when terms for global, systematic and local stationary trends in non-genetic variation have been included in the model (Dutkowski et al. 2006; Gilmour et al. 1997). Prediction of genetic effects with these models requires knowledge of the genetic architecture of traits (i.e. variances, heritabilities and correlations). These parameters are often estimated from the available data, with the estimated parameters used to predict empirical best linear unbiased predictions (E-BLUPs) (Kackar and Harville 1981).

Here, we report on the development and implementation of a multivariate linear mixed model with spatial variation modelling to combine growth and wood quality trait data assessed across international progeny trials of open-pollinated seed orchard-derived *Eucalyptus dunnii* families for the prediction of seed lot value and response to selection for pulp and biofuel utilisation.

E. dunnii Maiden is a subtropical tree with limited distribution along the central coastal ranges of Australia (Benson and Hagar 1993). Potential for pulp production in subtropical climates is promising due to frost resistance, pulpwood productivity, growth rate and good form (Thomas et al. 2009). *E. dunnii* plantations have been established in Australia and in drier and frost-affected areas, such as southern Brazil, central China, South Africa, Uruguay, and Argentina, where *E. dunnii* outperforms *E. grandis*, a widely used hardwood species, particularity for pulp and biofuel utilisation where breeding objectives for both uses are similar (Arnold et al. 2004; Geary 2001; Jovanovic et al. 2000; Sheperd et al. 2011).

When selection seeks to improve multiple traits, methods are required to identify the best overall candidate. Selection indices have been used to identify elite germplasm for pulp production using volume, density and pulp yield (Borralho et al. 1993; Greaves et al. 1997; Wei and Borralho 1999). Alternatively, others (Brawner et al. 2012) have used standing pulp productivity, which is the product of volume, density and pulp yield, as the selection objective.

Open-pollination (OP) is a common mating strategy for genetic evaluation and deployment in eucalypts. A half-sib relationship among OP siblings is often assumed (White et al. 2007); however, mixed mating is general in eucalypts (Byrne 2008; Eldridge et al. 1993; Hardner and Potts 1995), suggesting the assumption of half-sib relationships may not be valid in these species. Inbreeding inflates additive genetic variation, and, for traits influenced by dominance, inbreeding depression and additional components of genetic variance arise (Cockerham and Weir 1984). Previous studies have attempted to adjust for inbreeding by inflating the additive relationship coefficients (Squillace 1974), but this does not account for outcrossing rate heterogeneity and dominance effects (Borralho 1994; Hardner et al. 1996). Hence, breeding values for OP individuals predicted using their own performance data may not reflect the value of seed lots collected from these individuals.

2 Methods

2.1 Stockdale seed orchard

Seed lots included in the progeny trials described below were originally selected based on breeding value of seed orchard individuals predicted using their own performance. In 1989, a trial was established near Stockdale in eastern Victoria, Australia (Table 1), with 57 OP families collected from individual native forest *E. dunnii* trees at Boomi Creek NSW, Dead Horse Track, South Yabbra State Forest and unknown origin planted in 15 completely replicated blocks of single tree plots per OP family (total of 870 individuals). At 14 years, diameter at breast height (DBH, cm) was assessed for all individuals and wood basic density (BD kg/m^3 from cores) for individuals in 4 blocks (Table 2). Basal area (BA, cm^2) was calculated from DBH. Prior to the current study, breeding values for these selection traits for individuals within the seed orchard were predicted using only the seed orchard data and a simple (i.e. non-spatial models) individual (Mrode 2005) univariate model assuming a coefficient of relationship of 2.5 (i.e. selfing rate of 30 %, Griffin and

Table 1 Overview of the experimental design of the trials included in this study

Trial	Lat.	Long.	nRow	nSp	nB	nP/B	nT/P	nSL	nT/SL
Stockdale (SD)	37° 50′ S	147° 11′ E	34	48	15	64	1	64	15
Blackwood (BW)	37° 54′ S	140° 50′ E	34	80	16	170	1	168	16.1
FEA (FEA)	28° 43′ S	152° 27′ E	40	73	5	161	2–3	183	12.6
Uruguay—143 (U143)	32° 49′ S	57° 54′ W	90	60	6	52	6	52	34.8
Uruguay—144 (U144)	32° 28′ S	57° 24′ W	76	94	6	50	6	50	36

Row planting row, *Sp* planting space, *B* block, *P* plot, *T* tree, *SL* seed lot

Cotterill [1988]). A selection index of breeding values for volume, wood density, stem form and pulp yield was constructed from the additive genetic correlation matrix between selection and objective traits and economic weights derived from Greaves et al. ([1997]). Selection index values were used to identify the 50 % poorest performing individuals for removal in 2003 to convert the trial into a seed orchard and identify elite seed lots for deployment in subsequent progeny trials.

2.2 Progeny trials

2.2.1 Blackwood progeny trial

The 'Blackwood' progeny trial was planted in September 2007 at Yahl, South Australia, with OP seed lots collected in January 2007 from 77 highly ranked individuals at the Stockdale seed orchard (Table 1). Ninety native-forest OP seed lots collected from mother trees across 6 provenances (Acacia Creek, Beaury Creek, Boomi Creek, Lindesay

Creek, Teviot Brook and Wallaby Creek) and a single bulk of seed from orchards in NSW were also included. Single tree plots of each seed lot were established in 16 completely replicated blocks using a resolvable row-column design. Failed trees were replanted in the first year. DBH was assessed at 3 years on the 3 largest stems of all non-replant individuals across all 16 blocks (2326 individuals) and used to estimate BA as the sum of the BA of the two largest stems.

Only a subset of blocks were assessed for subsequent traits due to resource constraints. BA was again assessed at 4 years and 7 months (5 years) after planting for 1200 trees in blocks 9 to 16. Height (HT, m) was also assessed at this age for 759 trees in blocks 9, 12, 13, 15 and 16. Pilodyn penetration (PP) (mm) was assessed at 3 years on the largest stem for individuals in all 16 blocks except 4, 8 and 14 (1866 trees), and wood basic density (BD, kg/m^3) was assessed at 5 years for the 642 trees in blocks 9, 10, 11, 12 and 13. Separate air-dried drill swarfs from 585 trees in blocks 10, 11, 12 and 13 were used for NIR prediction of basic density (DN kg/m^3), pulp yield (PY, % kg/kg) and cellulose

Table 2 Summary of data collected at different ages from the five trials (*SD* Stockdale seed orchard, *BW* Blackwood progeny trial, *FEA* Banalbo Forest NSW progeny trial, *U143* Uruguay progeny trial 143, *U144* Uruguay progeny trial 144) included in this study

Trial	Trait	Age (yrs)	Unit	nobs	mean	var
SD	Basal Area (BA)	14	cm^2	734	224	6920
	Basic wood density (BD)	14	kg/m^3	202	493	1340
BW	Basal Area (BA)	3	cm^2	2328	85	1400
	Basal Area (BA)	5 (4.6)	cm^2	1199	195	5270
	Height (HT)	5 (4.6)	m	757	9.2	1.72
	Pilodyn penetration (PP)	3	mm	1841	16.1	3.41
	Basic wood density (BD)	5 (4.6)	kg/m^3	645	455	1190
	Wood density- NIR (DN)	5 (4.6)	kg/m^3	585	577	925
	Pulp Yield – NIR (PY)	5 (4.6)	%kg/kg	584	44.4	8.25
	Cellulose content – NIR (CC)	5 (4.6)	%kg/kg	584	38.2	4.34
FEA	Height (HT)	2 (1.7)	m	2186	3.1	0.71
	Basal Area (BA)	3 (2.7)	cm^2	2153	16	97.7
	Basal Area (BA)	5 (5.3)	cm^2	1871	27	301
U143	Basal Area (BA)	2	cm^2	1655	30	166
	Height (HT)	2	m	1666	5.8	1.14
U144	Basal Area (BA)	2	cm^2	1483	29.4	201
	Height (HT)	2	m	1492	5.2	1.18

Values in the parentheses alongside age refer to the precise assessment age

INRA SCIENCE & IMPACT   Springer

content (CC, % kg/kg) using the multi-site and species calibration (Downes et al. 2010).

2.2.2 FEA progeny trial

Eighty-eight native forest OP seed lots and 94 Stockdale seed orchard seed lots were planted at Banalbo in northern NSW (FEA) in 5 completely replicated blocks of 3 tree plots per seed lot, with 147 seed lots in common with Blackwood. HT was assessed at 20 months (2 years), and DBH at 32 (3 years) and 64 (5 years) months for all plants, and BA calculated from DBH.

2.2.3 Uruguay progeny trials

Two trials (U143 and U144) were established in Uruguay with 50 Stockdale seed orchard seed lots (32 in common with Blackwood and 40 with FEA) and 2 seed orchard commercial seed lots in 6 completely replicated blocks of 6 tree plots. Data from unrelated seed lots also planted in this trial were not included in this study. HT and DBH were assessed at 24 months (2 years after planting) and BA calculated from DBH.

2.3 Statistical methods

2.3.1 Linear model

A multivariate linear mixed model was developed and implemented for the analysis of the progeny trial data. For m individuals assessed for w attributes (where the kth attribute was a unique trait-by-age-by-trial combination, i.e. $w = \sum_{l=1}^{l=t} w_l$, where t was the number of trials and w_l was the number of attributes assessed at the lth trial), the linear mixed model used to estimate parameters for the random and residual effects, test the significance and estimate means of fixed effects and predict values for seed lot effects/values (g_i) was

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{r}$$

where $\mathbf{y}$ was a vector of observations, $\mathbf{b}$ was vector of unknown fixed effects, $\mathbf{X}$ was the design matrix for fixed effects, $\mathbf{u}$ was a vector of unknown random effects, $\mathbf{Z}$ was the design matrix for random effects and $\mathbf{r}$ was a vector of unknown residual effects, and

$$\text{var}(\mathbf{y}) = \mathbf{ZGZ}^T + \mathbf{R}$$

where $\mathbf{G}$ was the variance-covariance among genetic and non-genetic random effects and $\mathbf{R}$ was the variance-covariance among residual effects. As planting row-by-planting space coordinates for each trial were available, global, systematic and local stationary spatial trends were modelled after Gilmour et al. (1997).

Fixed effects included the general mean for each attribute, linear regressions on planting row and planting space, and seed lot origin for attributes assessed at Blackwood and FEA (i.e. native forest, Stockdale seed orchard or bulk production area). Random effects included seed lot effects by attribute and independent block, plot within block (for the FEA, U143 and U144 progeny trials), planting row and planting space for each attribute. The covariance structure of genetic effects by attribute was modelled as an unstructured variance-correlation structure for models that included less than 4 attributes and as a factor-analytic structure (Hardner et al. 2010; Smith et al. 2001; Thompson et al. 2003) of order f for models that included greater than 3 attributes.

The variance-covariance matrix of residual effects ($\mathbf{R}$) was modelled as an anisotropic stationary separable first-order autoregressive correlation structure in the two spatial dimensions (planting row and space) with an independent residual (nugget) (Gilmour et al. 1997):

$$\mathbf{R} = \mathbf{R}_S + \mathbf{R}_I$$

where $\mathbf{R}_S$ was the variance-covariance among spatially dependent residual effects and $\mathbf{R}_I$ was the variance-covariance among spatially independent residual effects. In general, $\mathbf{R}_S$ was block diagonal, and with each block, the correlated residuals for the rth section, $\mathbf{R}_{S,r}$. The general structure of $\mathbf{R}_{S,r}$ was

$$\mathbf{R}_{S,r} = \mathbf{E}_r \left[\mathbf{C}_{\text{Row},r}(\rho_{\text{Row},r}) \otimes \mathbf{C}_{Sp,r}(\rho_{\text{Sp},r}) \right]$$

where $\mathbf{E}_r$ was the covariance matrix of the spatially dependent residuals among the k_r attributes in the rth section, and $\mathbf{C}_{Row,r}(\rho_{Row,r})$ was the first-order autoregressive correlation matrix (with common correlation $\rho_{Row,r}$ among attributes) in the planting row dimension for the rth section, with $\mathbf{C}_{Sp,r}(\rho_{Sp,r})$ for the planting space dimension. Different structures of $\mathbf{R}_S$ are possible. For example, each section may be defined by each unique attribute so that spatially dependent residuals among attributes within the lth trial were independent and $\mathbf{E}_r$ was defined as a single variance component. Alternatively, each section may correspond to more than one attribute at a trial, in which case the same spatial processes were assumed for all these attributes. The variance-covariance of spatially independent effects, $\mathbf{R}_I$, was modelled with similar structures with $\rho_{Row,r}$ and $\rho_{Sp,r}$ constrained to 0. Where appropriate, both $\mathbf{E}_{S,r}$ and $\mathbf{E}_{I,r}$ were modelled as structured

covariance matrices using either an unstructured variance-correlation matrix or a factor-analytic parameterisation.

Parameters of the random and residual models were estimated using average information approaches implemented in the statistical software ASReml (Gilmour et al. 2009). Initial single-attribute (i.e. single trait, single trial) univariate analyses were undertaken to check the assumption of normal distribution of residuals and identify significant sources of non-genetic variation for inclusion in multi-attribute analyses. For nested models with common fixed effects, significance of random effects were tested using the likelihood ratio test (Wilks 1938) and adjusted where the estimated parameters were at the boundary of the estimation space (Stram and Lee 1994). Akaike information criterion (Akaike 1974) (AIC) was used to more generally compare the goodness-of-fit among models with common fixed effects. Significance of fixed effects was tested using Wald tests (Kenward and Roger 1997). Standard errors of variances and correlations estimated as functions of factor-analytic parameters in the analysis were obtained using the variance-covariance of model parameters and a first-order approximation of a Taylor series expansion (Kendall et al. 1987; White 2016). Best linear unbiased estimates (BLUEs) of fixed effects and best linear unbiased predictions (BLUPs) of random effects (seed lot values) were produced following Henderson's (1975) mixed model equations using estimated **G** and **R**.

2.3.2 Analysis of genetic architecture

Individual seed lot repeatability for the kth attribute was estimated as

$$\hat{H}^2_k = \frac{\hat{\sigma}^2_{\text{SL},k}}{\hat{\sigma}^2_{\text{SL},k} + \hat{\sigma}^2_{R_I,k}}$$

where $\hat{\sigma}^2_{\text{SL},k}$ was the estimated variance among seed lots and $\hat{\sigma}^2_{R_I,k}$ was the estimated variance of the spatially independent residuals. Individual narrow-sense heritability of the kth attribute was estimated as

$$\hat{h}^2_k = \frac{\hat{\sigma}^2_{A,k}}{\hat{\sigma}^2_{A,k} + \hat{\sigma}^2_{R_I^{\#},k}}$$

where $\hat{\sigma}^2_{A,k}$ was the estimated additive genetic variance and $\hat{\sigma}^2_{R_I^{\#},k}$ was the estimated variance of the spatially independent residuals from a genetic model that assumed 0.3 selfing rate

(Griffin and Cotterill 1988).

Classical repeatability of seed lot predictions for the kth attribute was estimated as

$$\hat{H}^2_{SL,k} = \frac{\hat{\sigma}^2_{SL,k}}{\left(\hat{\sigma}^2_{SL,k} + \dfrac{\hat{\sigma}^2_{R_I,k}}{n_k} \right)}$$

where n_k was the average number of individuals with observations for the kth attribute. Generalised repeatability (Cullis et al. 2006; Piepho and Mohring 2007) of seed lot predictions for the kth attribute was estimated as

$$\hat{H}^{2*}_{SL,k} = 1 - \frac{\bar{\sigma}^2_{\Delta SL,k}}{2 \times \hat{\sigma}^2_{SL,k}}$$

where $\hat{\sigma}^2_{\Delta SL,k}$ was the mean variance of the difference of seed lot predictions, estimated from the prediction error variance matrix (PEV). Generalised repeatabilities of seed lot predictions were estimated using the estimated seed lot variance from the full multi-attribute analysis and corresponding PEV for (i) only seed lots for the attribute ($pH^{2*}_{SL,k}$) or (ii) all seed lots ($aH^{2*}_{SL,k}$). In addition, generalised repeatabilities of seed lot predictions were estimated using the variances and PEV from a univariate analyses of each attribute ($uH^{2*}_{SL,k}$) with the same fixed and random factors as in the multivariate analysis. Standard errors of individual seed lot repeatability, individual narrow-sense heritability and classical seed lot repeatability were estimated using a first-order approximation of the Taylor expansion as described above. Standard errors for generalised repeatability estimates were not estimated, as these parameters are derived from a function of model parameters and the variance-covariance matrix of predicted effects, for which the method of estimating standard errors has not yet been defined.

2.3.3 Response to selection

To evaluate the response to selection of seed lots based on breeding values of parent predicted from own performance in the Stockdale seed orchard, simple correlation coefficients were estimated between these predicted breeding values and seed lot values predicted from the analysis of the progeny trials.

Following Brawner et al. (2012), values for total cellulose mass per tree (TC, kg) of the ith seed lot at the lth trial were predicted as

$$\hat{g}_{TC_{l,i}} = 0.1 \left[\frac{1}{3} \hat{g}_{BA_{l,i}} \times \hat{g}_{HT_{l,i}} \right] \times \hat{g}_{BD_i} \times \frac{\hat{g}_{PY_i}}{100} \times \frac{\hat{g}_{CC_i}}{100}$$

where $\hat{g}_{BA_{l,i}}$ was the predicted mean of the ith seed lot for BA (cm^2/tree) at the lth trial; $\hat{g}_{HT_{l,i}}$ was the predicted mean of the ith seed lot for HT (m/tree) at l^{th} trial; $\hat{g}_{BD_i}$ was the predicted mean of basic density (kg/m^3) of the ith seed lot; $\hat{g}_{PY_i}$ was the predicted mean of pulp yield (%kg/kg) of the ith seed lot at 5 years and $\hat{g}_{CC_i}$ was the predicted mean of cellulose content (%kg/kg) of the ith seed lot at 5 years. Predicted seed lot means were estimated as the sum of the estimated attribute mean and the predicted seed lot values. For attributes transformed prior to analysis, predicted seed lot means were back-transformed prior to prediction of TC. Correlations of seed lot effects among trials for DB, PY and CC were assumed to be 1. Seed lot values for TC were predicted at 5 years for Blackwood and at age 2 years for the Uruguay trials. Predictions for individual attributes and TC at Uruguay were averaged by seed lot across the two trials. Predictions for performance at the FEA trial location were excluded due to poor productivity (see results).

The difference between the seed lots from native forest or the Stockdale seed-orchard were estimated as the average of predicted seed lot means for each seed lot origin plus the effect of the origin, where significant. Averages were also predicted for the top 50 %, 25 % and 12.5 % of Stockdale seed lots based on TC. Difference in TC between origins was not tested as back-transformed predicted seed lot means were used.

3 Results

3.1 Implementation of multivariate mixed model analysis of progeny trials

Residuals from the initial univariate single attribute analyses of progeny trial data were normally distributed, except for BA at 5 years at FEA (log transformed) and 2 years at U144 (square root) and HT at 5 years at Blackwood (squared) and 2 years at the U143 (square root). Single attribute models that included either heterogeneous seed lot or residual variances among origins were not significantly ($p = 0.05$) different from homogeneous models, expect for BA at 5 years at Blackwood where the estimated variation among the Stockdale seed lots was 0.

The most parsimonious multivariate mixed model (MVMM) of 15 attributes across four progeny trials employed a second-order factor analytic structure (i.e. FA2) for the variance-covariance of seed lot effects. Heterogeneous seed lot variances among origins for BA at 5 years Blackwood were not included to avoid over-parameterisation. An FA2 structure was also used to model the covariance among spatially independent residuals for the eight attributes at Blackwood. Unstructured covariance matrices were applied at the other trials. Spatially dependent residuals for attributes at Blackwood were modelled as independent (Table 3). However, the model that included an unstructured variance-covariance among spatially dependent residuals with common row and space autoregressive correlations at the FEA and Uruguay trials was more parsimonious (lower AIC) than the model that allowed independent spatial processes.

3.2 Parameter estimates from most parsimonious model

Spatial effects were detected for most growth traits (BA and HT) (Table 3). Significant spatial effects were not detected for wood quality traits except PP. Where significant, variance of spatially dependent residual was 8–40 % of that for spatially independent (nugget) residual. Autocorrelation of residuals was high (0.73–0.98) and was stronger along planting rows [ar1(Sp)] than among planting rows [ar1(Row)].

Estimated individual seed lot repeatability (H^2) were relatively very low (<0.05) for BA at FEA and DN (0.02) and PY (0.05) at Blackwood, low (0.07–0.14) for all other growth attributes and PP and CC at Blackwood and moderate (0.23) for basic density (BD) at Blackwood (Table 4). Estimates of individual narrow-sense heritability (h^2, assuming a selfing rate of 0.3) varied from 0.17 to 0.25.

Correlation of seed lot values for growth (BA and HT) within and across trials was greater than 0.85, except for HT at FEA with other attributes (Table 5). Wood density attributes (BD and PP) were also highly correlated (−0.98), as were PY and CC (0.93), although correlations with DN were moderate (0.60, −0.44). Seed lot correlations among growth (BA and HT) and wood density (BD and PP) were lower (−0.46 < r^2_{SL} < 0.29), although correlations with DN were stronger (−0.99 < r^2_{SL} < −0.60). Seed lot values for PY and CC were moderately correlated with growth (0.50 < r^2_{SL} < 0.72) and wood density (PP and BD).

3.3 Repeatability of seed lot predictions

Classical estimates of the repeatability of seed lot predictions (H^2_{SL}, Table 5) were highest for HT and BA at the Uruguay

Table 3 Estimated variance components for random effects (*B* block, *P* plot within block, *SL* seed lot, R_I variance of spatially independent residuals, R_S variance of spatially dependent residuals), first-order auto-correlation coefficients (*ar1(Row)* spatially dependent residuals along Row, *ar1(Sp)* spatially dependent residuals perpendicular to Row) for 15 trial-by-trait-by-age attributes (*BA* basal area, *HT* stem height, *PP* pilodyn penetration, *BD* wood basic density determined by volumetric method, *DN* wood density determined by NIR, *PY* pulp value determined by NIR, *CC* cellulose content determined by NIR,)

Trial	Trait	Age	B	B/P	SL	R_I	R_S	ar1(Row)	ar1(Sp)
BW	BA	3	13		88 (15.4)	1000 (30.7)	208 (36.0)	0.82 (0.036)	0.94 (0.013)
BW	BA	5	67		486 (82.3)	4304 (143.2)			
BW	HT[a]	5	116		39.4 (8.93)	248.9 (14.81)	32.3 (9.38)		0.93 (0.032)
BW	PP	3	0.00		0.38 (0.068)	2.80 (0.101)	0.25 (0.087)	0.86 (0.073)	0.96 (0.021)
BW	BD	5	3		277 (55.7)	916 (55.6)			
BW	DN	5	29		20 (11.7)	885 (51.1)			
BW	PY	5	0.55		0.34 (0.124)	5.95 (0.331)			
BW	CC	5	0.06		0.33 (0.095)	3.53 (0.202)			
FEA	HT	2		0.007	0.040 (0.0071)	0.496 (0.0169)	0.214 (0.0579)	0.85 (0.038)	0.98 (0.007)
FEA	BA[b]	3		0.3	3.0 (0.78)	76.9 (2.56)	23.8 (6.00)	0.85 (0.038)	0.98 (0.007)
FEA	BA	5		0.005	0.030 (0.0068)	0.514 (0.0177)	0.063 (0.0195)	0.85 (0.038)	0.98 (0.007)
U143	BA	2		1	10 (2.9)	139 (5.8)	44 (8.4)	0.80 (0.041)	0.96 (0.011)
U143	HT[c]	2		0.0004	0.00274 (0.000836)	0.0373 (0.00164)	0.0221 (0.00391)	0.80 (0.041)	0.96 (0.011)
U144	BA[c]	2		0.00	0.15 (0.039)	1.72 (0.071)	0.36 (0.092)	0.73 (0.053)	0.98 (0.005)
U144	HT	2		0.010	0.070 (0.0194)	0.766 (0.0327)	0.290 (0.0629)	0.73 (0.053)	0.98 (0.005)

Standard errors shown in parentheses

[a] Square transformed

[b] Log transformed

[c] Square root transformed

trials (U143 and U144) but low for NIR assessed attributes (i.e. DN, PY and CC). Generalised seed lot prediction repeatability estimated from individual attribute analyses ($uH_{SL,k}^{2*}$) were correlated (0.74) with H_{SL}^2. Generalised seed lot prediction repeatability estimated from the multi-attribute analysis was consistently greater than the H_{SL}^2 and $uH_{SL,k}^{2*}$. This was true for both generalised repeatability considering only those individuals for which the attribute was assessed

Table 4 Estimates of individual seed lot repeatability (H^2), narrow-sense heritability assuming a selfing rate of 0.3 (h^2), classical repeatability of seed lot predictions (H_{SL}^2) and generalised seed lot prediction repeatability using prediction error variance matrix (PEV) from single attribute analysis ($uH_{SL,k}^{2*}$); PEV from multi-attribute analysis considering only seed lots for which attributes were assessed (pH_{SL}^{2*}) and PEV from multi-attribute analysis considering all seed lots (aH_{SL}^{2*}) for 15 trial-by-trait-by-age attributes (*BA* Basal Area, *HT* stem height, *PP* pilodyn penetration, *BD* wood basic density determined by volumetric method, *DN* wood density determined by NIR, *PY* pulp value determined by NIR, *CC* cellulose content determined by NIR,). Also shown is the average number of observations per seed lot (n_k)

Trial	Trait	Age	H^2	h^2	n_k	H_{SL}^2	$uH_{SL,k}^{2*}$	pH_{SL}^{2*}	aH_{SL}^{2*}
BW	BA	3	0.08 (0.013)	0.20 (0.034)	13.9	0.55 (0.045)	0.58	0.72	0.69
BW	BA	5	0.10 (0.016)	0.25 (0.040)	7.1	0.45 (0.043)	0.52	0.72	0.68
BW	HT[a]	5	0.14 (0.029)	0.34 (0.072)	4.5	0.42 (0.059)	0.38	0.69	0.66
BW	PP	3	0.12 (0.019)	0.30 (0.049)	11.0	0.60 (0.044)	0.59	0.74	0.63
BW	BD	5	0.23 (0.039)	0.58 (0.098)	3.9	0.54 (0.055)	0.48	0.75	0.64
BW	DN	5	0.02 (0.013)	0.06 (0.032)	3.5	0.07 (0.040)	0.26	0.73	0.69
BW	PY	5	0.05 (0.019)	0.14 (0.048)	3.5	0.17 (0.052)	0.43	0.72	0.64
BW	CC	5	0.09 (0.024)	0.21 (0.059)	3.5	0.25 (0.056)	0.44	0.68	0.61
FEA	HT	2	0.07 (0.013)	0.19 (0.032)	12.0	0.49 (0.046)	0.51	0.64	0.60
FEA	BA	3	0.04 (0.010)	0.09 (0.024)	11.8	0.32 (0.057)	0.38	0.69	0.69
FEA	BA[b]	5	0.05 (0.012)	0.14 (0.030)	10.7	0.38 (0.055)	0.35	0.69	0.69
U143	BA	2	0.07 (0.018)	0.17 (0.045)	31.8	0.70 (0.061)	0.46	0.84	0.69
U143	HT[c]	2	0.07 (0.020)	0.17 (0.049)	32.0	0.70 (0.065)	0.51	0.79	0.61
U144	BA[c]	2	0.08 (0.020)	0.20 (0.049)	29.7	0.72 (0.053)	0.58	0.85	0.69
U144	HT	2	0.08 (0.022)	0.21 (0.054)	29.8	0.73 (0.055)	0.52	0.83	0.68

Standard errors are shown in parentheses.

[a] Square transformed

[b] Log transformed

[c] Square root transformed

Table 5 Estimated seed lot correlations among 15 trial (*BW* Blackwood, FEA, *U143* Uruguay 143, *U144* Uruguay 144)-by-trait (*BA* basal area, *HT* height, *BD* wood basic density, *DN* wood density predicted from NIR, *PP* pilodyn penetration, *PY* pulp yield predicted from NIR, *CC* cellulose content predicted from NIR)-by-age attributes

Trait	Trial	Age	BA					HT				DN	PP	DN.N	PY.N	CC.N
			BW	FEA	FEA	U143	U144	BW	FEA	U143	U144	BW	BW	BW	BW	BW
			5	3	5	2	2	5	2	2	2	5	3	5	5	5
BA	BW	3	0.99 (0.014)	1.00 (0.009)	1.00 (0.007)	0.97 (0.039)	1.00 (0.002)	0.97 (0.062)	0.68 (0.069)	0.93 (0.049)	0.96 (0.045)	−0.23 (0.132)	0.04 (0.126)	−0.91 (0.126)	0.68 (0.114)	0.70 (0.083)
	BW	5		0.98 (0.020)	0.99 (0.018)	0.95 (0.046)	0.99 (0.014)	0.96 (0.062)	0.67 (0.069)	0.92 (0.050)	0.96 (0.041)	−0.19 (0.131)	0.00 (0.124)	−0.89 (0.136)	0.70 (0.107)	0.72 (0.078)
	FEA	3			1.00 (0.000)	0.98 (0.034)	1.00 (0.014)	0.96 (0.064)	0.67 (0.070)	0.92 (0.055)	0.94 (0.065)	−0.29 (0.164)	0.11 (0.162)	−0.94 (0.110)	0.63 (0.141)	0.66 (0.111)
	FEA	5				0.98 (0.033)	1.00 (0.011)	0.96 (0.063)	0.68 (0.070)	0.92 (0.053)	0.95 (0.059)	−0.28 (0.154)	0.10 (0.150)	−0.93 (0.113)	0.64 (0.133)	0.67 (0.102)
	U143	2					0.97 (0.045)	0.92 (0.078)	0.64 (0.080)	0.87 (0.060)	0.87 (0.091)	−0.46 (0.158)	0.29 (0.163)	−0.99 (0.054)	0.47 (0.171)	0.53 (0.137)
	U144	2						0.97 (0.062)	0.68 (0.069)	0.93 (0.049)	0.97 (0.020)	−0.21 (0.171)	0.03 (0.166)	−0.91 (0.134)	0.69 (0.134)	0.71 (0.103)
HT	BW	5							0.66 (0.079)	0.90 (0.074)	0.94 (0.071)	−0.16 (0.146)	−0.02 (0.139)	−0.86 (0.150)	0.70 (0.119)	0.71 (0.096)
	FEA	2								0.64 (0.073)	0.67 (0.075)	−0.10 (0.136)	−0.03 (0.132)	−0.60 (0.125)	0.50 (0.111)	0.51 (0.094)
	U143	2									0.92 (0.062)	−0.11 (0.192)	−0.07 (0.186)	−0.80 (0.168)	0.70 (0.143)	0.71 (0.116)
	U144	2										0.05 (0.188)	−0.23 (0.172)	−0.77 (0.210)	0.85 (0.102)	0.84 (0.078)
DN	BW	5											−0.98 (0.022)	0.60 (0.257)	0.56 (0.140)	0.45 (0.126)
PP	BW	3												−0.44 (0.280)	−0.71 (0.114)	−0.59 (0.105)
DN.N	BW	5													−0.32 (0.302)	−0.39 (0.263)
PY.N	BW	5														0.93 (0.034)

Standard errors are presented in parentheses

(pH_{SL}^{2*}) or considering all individuals (aH_{SL}^{2*}). Estimates of pH_{SL}^{2*} were larger and more variable than aH_{SL}^{2*} estimates.

3.4 Response to selection

The correlation between breeding values for BA and BD (Fig. 1) across all individuals in Stockdale seed orchard was near zero (−0.09) but was moderately negative (−0.62) for the subset of individuals selected as sources of seed lots for the progeny trials was considered (Fig. 1). Breeding values for BA for individuals at Stockdale were only weakly correlated with seed lot values for BA predicted from the progeny trials for OP families collected from the same individuals (0.2, Table 6), as were breeding values for BD at Stockdale with seed lot values for BD, DN and PP at Blackwood (−0.16 to 0.17). Selection of all the Stockdale seed lots evaluated in progeny trials increased TC by 12–13 %, relative to native seed lots, which was primarily due to an increase in BA (Table 7). While gain in TC increased with selection intensity, there was little gain in wood-quality attributes.

4 Discussion

4.1 Implementation of multivariate, spatial modelling methods

The comprehensive multivariate modelling approach developed in this study is one of the few in any agricultural crop literature to combine a structured variance-covariance at treatment level, with structured variance-covariance matrices at the spatially dependent and independent residual levels in a single-step analysis. While the most parsimonious model included independent spatial process among attributes at the Blackwood trial, this model included common auto-regressive processes and a structured variance-covariance matrix of spatially dependent residuals among attributes at the FEA and U144 trials. This is likely because only growth attributes at a young age were assessed at these latter trials, compared to the more comprehensive assessment undertaken at Blackwood. In separate studies, both Smith et al. (2007) and De Faveri et al. (2015) assumed a common process for spatially dependent residuals when modelling a single trait at a single trial over repeated years, whereas others (Ivkovic et al. 2015) adjusted raw data for spatial trends detected in single trial analyses, prior to combining data into a multi-environment model.

Fig. 1 Plot of predicted breeding values for basal area (BA, cm^2) by basic density (BD, kg/m^3) of individuals in Stockdale seed orchard selected (*black circle*) and unselected (*white circle*) for establishment of progeny trials
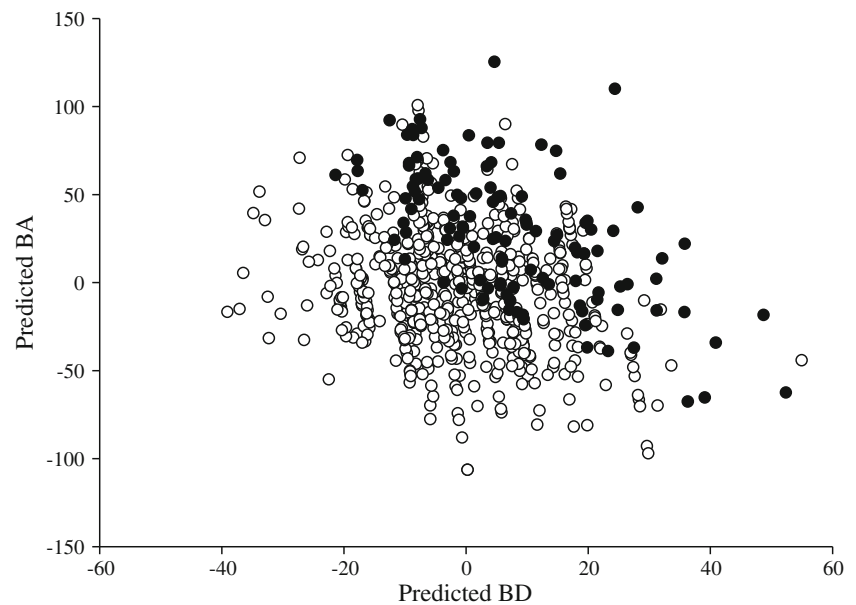
4.2 Spatial effects

Detection of spatial variation for growth attributes in this study is consistent with results published for forest genetic field trials (Dutkowski et al. 2006). While this study reports slightly larger autocorrelation along planting rows, Dutkowski et al. (2006) reports mostly isotropic variation. Not surprisingly, we found no evidence of competition (i.e. negative autocorrelation, Dutkowski et al. (2006)), as most of the attributes in this study were assessed prior to age 4 after which strong competition appears to develop (Hardner and Potts 1997). The absence of significant spatial variation in most wood quality attributes reported here may be due to the true absence of these effects for these attributes (although no literature for these traits in eucalypts was found) or because these attributes were only assessed on a subset of material at Blackwood, as others (Dutkowski et al. 2006) report spatial variation is more commonly detected in larger trials.

4.3 Comparison of genetic architecture with published estimates

Our estimates of individual narrow-sense heritability for early age growth and basic density, and seed lot correlations among these traits, are similar in magnitude to other estimates for volume and density for *E. dunnii* (Arnold et al. 2004; Luo et al. 2012) and other eucalypt species (Costa e Silva et al. 2009; Hamilton and Potts 2008; Hung et al. 2015; Kien et al. 2009a; Kien et al. 2009b; Raymond 2002; Wei and Borralho 1999). Lower heritability of the growth attributes at FEA may be a consequence of the poorer productivity of this trial.

This is the first report of genetic parameters for wood quality traits for *E. dunnii*. However, while our higher estimate for basic density compared to pilodyn penetration agrees with studies in other similar eucalypt species, our heritablity estimates for wood density, pulp yield and cellulose content assessed by NIR are lower. This lower heritability is possibly a consequence of low accuracy of estimation due to small number of individual assessed for these traits. Our results of high age-age correlation for basal area, and moderate correlation among growth traits, are consistent with other studies. However, high seed lot correlation among trials for growth traits observed here contrasts with significant G × E reported for *E. dunnii* in China (Arnold et al. 2004) but is consistent with studies in other eucalypt species. Clearly, the extent of G × E is dependent on

Table 6 Correlation between predicted breeding value for basal area (BA$_{S,14}$) and basic density (BD$_{S,14}$) of individuals at Stockdale seed orchard predicted from own data, and seed lot values of OP families collected from these individuals for similar traits predicted from progeny trials

Trait	Trial	Years	BA$_{S,14}$	BD$_{S,14}$
BA	Blackwood	3	0.20	−0.10
BA	Blackwood	5	0.20	−0.09
BA	FEA	3	0.20	−0.11
BA	FEA	5	0.20	−0.11
BA	U143	2	0.20	−0.11
BA	U144	2	0.20	−0.09
BD	Blackwood	5	−0.09	0.17
PP	Blackwood	3	0.06	−0.16
DN	Blackwood	5	−0.21	0.15

Table 7 Percent gain in 12 attributes (6 traits × 2 locations-by-ages) for different selection intensities of Stockdale seed orchard seed lots (SDs) compared to the average of unselected native forest seed lots (NF)

Trial	Trait	Unit	Age	NF	SD			
					100 %	50 %	25 %	12.5 %
BW	BA	cm^2	5	185.6	9 %	15 %	19 %	20 %
BW	HT	m	5	9.1	2 %	4 %	5 %	6 %
BW	BD	kg/m^3	5	456.9	−1 %	0 %	0 %	1 %
BW	PY	%kg/kg	5	43.8	1 %	1 %	2 %	2 %
BW	CC	%kg/kg	5	37.7	1 %	1 %	2 %	2 %
BW	TC	kg	5	4.3	12 %	22 %	29 %	33 %
U	BA	cm^2	2	24.8	9 %	16 %	19 %	22 %
U	HT	m	2	5.2	3 %	6 %	7 %	8 %
U	BD	kg/m^3	2	456.9	−1 %	0 %	0 %	0 %
U	PY	%kg/kg	2	43.8	1 %	1 %	2 %	2 %
U	CC	%kg/kg	2	37.7	1 %	1 %	2 %	2 %
U	TC	kg	2	0.33	13 %	24 %	31 %	36 %

the environments examined. Published estimates for genetic correlations among growth and wood quality traits also vary in other species, suggesting further studies are required. The low correlation between basic density and wood density assessed by NIR reported here may be due to the low heritability of NIR wood density.

4.4 Repeatability of seed lot predictions

The higher classical seed lot prediction repeatability estimates for growth at the Uruguay trials is due to higher level of replication at these trials. Similarly, the low repeatability estimates for NIR wood quality traits are a consequence of the very low individual seed lot repeatability and the small number of replications. Although the number of replicates assessed for basic density was similar to that for the NIR attributes, the moderate estimate for repeatability of seed lot predictions for basic density is due to the higher individual repeatability of this attribute.

The correlation between classically estimated seed lot prediction repeatability and generalised seed lot prediction repeatability estimated from the individual attribute analyses is expected as they are equal for completely balanced data (Cullis et al. 2006). However, for unbalanced designs (as here), a covariance is induced between predictions which will inflate the variance of the difference between predictions. This does not appear to be accounted for in some other studies (e.g. Oakey et al. 2006; Welham et al. 2010).

The higher generalised repeatability of seed lot predictions from the multivariate combined attribute analysis, compared to those from single-attribute analyses demonstrates the improvement in prediction accuracy from a multivariate approach due to incorporation of correlated information and improved data structure through the residual covariance (Sales and Hill 1976; Thompson and Meyer 1986). The greatest gain in accuracy is expected when the genetic and residual correlations are opposite in sign and when the repeatability of the correlated attribute is greater than that of target attribute (Thompson and Meyer 1986). However, this assumes genetic parameters are known without error. Here, the standard errors of most estimates are small relative to the estimate and our estimates are in general agreement with the published literature.

The higher generalised repeatability when only seed lots for which the attribute was assessed were included, compared to generalised repeatability estimated using the PEV for all seed lots, is as expected. Effects for attributes for which these seed lots have not been directly assessed are predicted from correlated information, and the error variance of these predictions will be greater than those for predictions of attributes for seed lots that were assessed directly. For unbalanced designs, generalised repeatability estimated considering all selection candidates may be the most appropriate measure to assess response to selection (but see Cullis et al. 2006; Welham et al. 2010).

4.5 Response to selection

The high correlation of seed lot means for total cellulose among trials is clearly a consequence of the high environmental correlation of growth attributes and the assumed high (i.e. 1) correlation among locations of seed lot effects for wood quality attributes, as no information was available to estimate this parameter. However, this assumption needs further evaluation as Luo et al. (2012) report significant G × E for wood density for *E. dunnii* across China, and studies in other eucalypt species (Brawner et al. 2012; Hung et al. 2015; Osorio et al. 2003) have reported genetic correlations among trials between 0.63 and 0.88 for this trait and 0.74 and 0.97 for kraft pulp yield. High age-age correlation of seed lot values for growth attributes suggests elite seed lots for total cellulose at 2 years in the Uruguay trials will also be elite at 5 years. A near-zero seed lot correlation between basic density and growth attributes explains the lack of response in wood density from selection on cellulose as no correlated information is available to predict density for seed lots only assessed for growth. Response to selection for total cellulose is likely to be improved if basic density, or the highly correlated attribute pilodyn penetration, was assessed across trials.

The low correlation between breeding values of individuals at Stockdale predicted based on their own information and deployment values of seed lots collected from these individuals may be due to: (i) difference in age at which the Stockdale seed orchard and progeny trials were assessed; (ii) differences in site conditions—although growth at Stockdale, Blackwood and Uruguay

were typical; (iii) the inadequacy of additive genetic models to incorporate confounding effects on family performance of variable outcrossing rates; or (iv) mortality and competition removing inbred individuals from the later age Stockdale population. Our results highlight the importance of backward testing seed orchard families to increase the accuracy for deployment in plantations.

5 Conclusion

Multivariate linear mixed model approaches with spatial modelling can be successfully developed and implemented to combine unbalanced data for multiple traits, from multiple trials, with repeated measures and local spatial variation modelling to improve prediction accuracy. Results from this approach supports selection of elite candidates with substantial gain in total cellulose achievable, and identified opportunities to achieve further gain.

References

Akaike H (1974) New look at model identification. IEEE Trans Autom Control AC19:716–723

Arnold RJ, Johnson IG, Owen JV (2004) Genetic variation in growth, stem straightness and wood properties in *Eucalyptus dunnii* trials in Northern New South Wales. For Genet 11:1–12

Benson J, Hagar T (1993) The distribution, abundance and habitat of *Eucalyptus dunnii* (Myrtaceae) (Dunn's white gum) in New South Wales. Cunninghamia 3:123–145

Borralho NMG (1994) Heterogeneous selfing rates and dominance effects in estimating heritabilities from open-pollinated progeny. Can J For Res-Rev Can Rech For 24:1079–1082

Borralho NMG, Cotterill PP, Kanowski PJ (1993) Breeding objectives for pulp production of *Eucalyptus globulus* under different industrial cost structures. Can J For Res-Rev Can Rech For 23:648–656

Brawner J, Meder R, Dieters M, Lee DJ (2012) Selection of *Corymbia citriodora* for pulp productivity. South Forests 74:121–131

Burdon RD (1977) Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genet 26:168–175

Byrne M (2008) Phylogeny, diversity and evolution of eucalypts. In: Shama AK, Sharma A (eds) Plant genome biodiversity and evolution. Science Publishers, Enfield

Cockerham CC, Weir BS (1984) Covariances of relatives stemming from a population undergoing mixed self and random mating. Biometrics 40:157–164

Costa e Silva J, Borralho NMG, Araujo JA, Vaillancourt RE, Potts BM (2009) Genetic parameters for growth, wood density and pulp yield in *Eucalyptus globulus*. Tree Genet Genomes 5:291–305

Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. J Agric Biol Environ Stat 11:381–393

De Faveri J, Verbyla AP, Pitchford WS, Venkatanagappa S, Cullis BR (2015) Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. Crop Pasture Sci 66:947–962

Downes G, Meder R, Harwood C (2010) A multi-site, multi-species near infrared calibration for the prediction of cellulose content in eucalypt wood meal. J Near Infrared Spec 18:381–387

Dutkowski GW, Silva JCE, Gilmour AR, Wellendorf H, Aguiar A (2006) Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. Can J For Res-Rev Can Rech For 36:1851–1870

Eldridge KG, Davidson J, Harwood CE, van Wyk G (1993) Eucalypt domestication and breeding. Oxford, UK

Falconer DS, Mackay TFC (1996) An introduction to quantitative genetics, 4th edn. Longman

Geary TF (2001) Afforestation in Uruguay—study of a changing landscape. J Forest 99:35–39

Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. J Agric Biol Environ Stat 2:269–273

Gilmour AR, Gogel B, Cullis BR, Thompson R (2009) ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead

Greaves BL, Borralho NMG, Raymond CA, Farrington A (1996) Use of a pilodyn for the indirect selection of basic density in *Eucalyptus nitens*. Can J For Res-Rev Can Rech For 26:1643–1650

Greaves BL, Borralho NMG, Raymond CA (1997) Breeding objective for plantation eucalypts grown for production of kraft pulp. For Sci 43:465–472

Greaves BL, Borralho NMG, Raymond CA (2003) Early selection in eucalypt breeding in Australia—optimum selection age to minimize the total cost of Kraft pulp production. New For 25:201–210

Griffin AR, Cotterill PP (1988) Genetic-variation in growth of outcrossed, selfed and open-pollinated progenies of *Eucalyptus regnans* and some implications for breeding strategy. Silvae Genet 37:124–131

Hamilton MG, Potts BM (2008) *Eucalyptus nitens* genetic parameters. N Z J For Sci 38:102–119

Hardner CM, Potts BM (1995) Inbreeding depression and changes in variation after selfing in *Eucalyptus globulus ssp globulus*. Silvae Genet 44:46–54

Hardner CM, Potts BM (1997) Post dispersal selection following mixed mating in *Eucalyptus regnans*. Evolution 51:103–111

Hardner CM, Borralho NMG, B. Tier, Miller S, Goddard ME (1996) Accounting for dominance and inbreeding in genetic evaluations using individual tree mixed models. In: Dieters MJ, Matheson AC, Nikles DG, Hardwood CE, Walker SM (eds) Tree improvement for sustainable tropical forestry Queensland Forestry Research Institute. Gympie, Caloundra, Queensland, pp. 143–147

Hardner CM, Dieters M, Dale G, DeLacy I, Basford KE (2010) Patterns of genotype-by-environment interaction in diameter at breast height at age 3 for eucalypt hybrid clones grown for reafforestation of lands affected by salinity. Tree Genet Genomes 6:833–851

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423–447

Henderson CR (1984) Estimation of variances and covariances under multiple trait models. J Diary Sci 67:1581–1589

Hung TD, Brawner JT, Meder R, Lee DJ, Southerton S, Thinh HH, Dieters MJ (2015) Estimates of genetic parameters for growth and wood properties in *Eucalyptus pellita* F. Muell. to support tree breeding in Vietnam. Ann For Sci 72:205–217

Ivkovic M, Gapare W, Yang HX, Dutkowski G, Buxton P, Wu H (2015) Pattern of genotype by environment interaction for radiata pine in southern Australia. Ann For Sci 72:11

Jovanovic T, Arnold R, Booth T (2000) Determining the climatic suitability of *Eucalyptus dunnii* for plantations in Australia, China and central and South America. New For 19:215–226

Kackar RN, Harville DA (1981) Unbiasedness of 2-stage estimation and prediction procedures for mixed linear-models. Commun Stat A-Theor 10:1249–1261

Kendall MG, Stuart A, Ord JK (1987) Kendall's advanced theory of statistics, vol 3: design and analysis, and time series. Oxford University Press, New York

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997

Kien ND, Jansson G, Harwood C, Thinh HH (2009a) Genetic control of growth and form in *Eucalyptus urophylla* in northern Vietnam. J Trop For Sci 21:50–65

Kien ND, Quang TH, Jansson G, Harwood C, Clapham D, von Arnold S (2009b) Cellulose content as a selection trait in breeding for kraft pulp yield in *Eucalyptus urophylla*. Ann For Sci 66:711

Luo JZ, Arnold RJ, Cao JG, Lu WH, Ren SQ, Xie YJ, Xu LA (2012) Variation in pulp wood traits between eucalypt clones across sites and implications for deployment strategies. J Trop For Sci 24:70–82

Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H (2006) Joint modeling of additive and non-additive genetic line effects in single field trials. Theor Appl Genet 113:809–819

Osorio LF, White TL, Huber DA (2003) Age-age and trait-trait correlations for *Eucalyptus grandis* hill ex maiden and their implications for optimal selection age and design of clonal trials. Theor Appl Genet 106:735–743

Piepho HP, Mohring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888

Raymond CA (2002) Genetics of *Eucalyptus* wood properties. Ann For Sci 59:525–531

Sales J, Hill WG (1976) Effect of sampling errors on efficiency of selection indexes. 2. Use of information in associated traits for improvement of a single important trait. Anim Prod 23:1–14

Shepherd M, Bartle J, Lee DJ, Brawner J, Bush D, Turnbull P, Macdonell P, Brown TR, Simmons B, Henry R (2011) Eucalypts as a biofuel feedstock. Biofuels 2:639–657

Smith A, Cullis B, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57:1138–1147

Smith AB, Stringer JK, Wei X, Cullis BR (2007) Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. Euphytica 157:253–266

Somerville C, Youngs H, Taylor C, Davis SC, Long SP (2010) Feedstocks for lignocellulosic biofuels. Science 329:790–792

Squillace AE (1974) Average genetic correlations among offspring from open-pollinated forest trees Silvae Genet:23

Stram DO, Lee JW (1994) Variance-components testing in the longitudinal mixed effects model. Biometrics 50:1171–1177

Thomas D, Henson M, Joe B, Boyton S, Dickson R (2009) Review of growth and wood quality of plantation-grown *Eucalyptus dunnii* maiden. Aust For 72:3–11

Thompson R, Meyer K (1986) A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest Prod Sci 15:299–313

Thompson R, Cullis B, Smith A, Gilmour A (2003) A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. Aust N Z J Stat 45:445–459

Wei X, Borralho NMG (1999) Objectives and selection criteria for pulp production of *Eucalyptus urophylla* plantations in south East China. For Genet 6:181–190

Welham SJ, Gogel BJ, Smith AB, Thompson R, Cullis BR (2010) A comparison of analysis methods for late stage variety evaluation trials. Aust N Z J Stat 52:125–149

White IMS (2016) Pin function. Accessed 05/2016 edn, http://www.homepages.ed.ac.uk/iwhite//asreml/

White TL, Adams WT, Neale DB (2007) Forest genetics. CAB International Wallingford

Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann Math Stat 9:60–62