CROP BREEDING AND APPLIED BIOTECHNOLOGY

ARTICLE

Genetic parameter estimation for *Ilex paraguariensis* St. Hill. in Argentina using spatial analysis

Crop Breeding and Applied Biotechnology
23(4): e46082344, 2023
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332023v23n4a39

Vanesa Carolina Schoffen[1*], Eduardo Pablo Cappa[2,3], María Elena Gauchat[4] and Ector Cesar Belaber[4]

Abstract: *Spatial and non-spatial analyses were conducted to estimate genetic parameters for the traits leaf mass weight (LMW), crown height (CH), crown diameter (CD), and crown volume (CV) for ages between 21 and 27 in 10 half-sib progeny trials of Ilex paraguariensis St. Hill. The spatial model gave a better fit than the base model in 87.2% of the analysed dataset, with reductions in residual and plot variances. The narrow-sense heritability estimates ranged from low to moderate for LMW trait (0.01 to 0.43) and from low to high for crown traits (0.08 to 0.74). The additive genetic coefficient of variation for the LMW trait was over 12.4%, while for CH and CD it was below 10%. Generally, the additive genetic correlations ($\hat{r}_g$) between the LMW evaluations and between LMW and crown traits were greater than 0.70.*

Keywords: *Half-sib progeny trials, leaf mass weight, heritability, genetic improvement*

INTRODUCTION

Yerba Mate (*Ilex paraguariensis* A. St. Hil.) is a tree species belonging to the family Aquifoliaceae. Its natural distribution covers southern Brazil, northeastern Argentina, eastern Paraguay, and Uruguay (Coelho et al. 2002). Argentina is the leading producer of Yerba Mate, accounting for 62% of the world's production, followed by Brazil (34%) and Paraguay (4%). In Argentina, Yerba Mate plantations are concentrated in Misiones Province and the northeast of Corrientes province. The plantations cover an approximate area of 209,276 hectares, whose average production exceeds 810 million kg of green leaf and 267 million kg of processed Yerba Mate. Eighty-six percent of processed Yerba Mate is destined for the domestic market and the remaining 14% is for the external market (INYM 2022).

The genetic improvement of this species began in Argentina in the 1970s through phenotypic selection conducted by the National Institute of Agricultural Technology (INTA) in commercial plantations (Prat Kricun 2013). Traits such as green leaf productivity, plant structure, leaf abscission, pest and disease tolerance were used in the selection process. In addition, different breeding programs have been initiated in the private sector, with Pindo S.A. being the only company that has published estimates of genetic parameters for the caffeine content, theobromine content, and yield using one-way analysis of variance (Scherer et

***Corresponding author:**
E-mail: schoffen.vanesa@inta.gob.ar
ORCID: 0009-0000-1992-536X

Received: 14 July 2023
Accepted: 05 September 2023
Published: 10 October 2023

[1] Instituto Nacional de Tecnología Agropecuaria - Estación Experimental Agropecuaria Cerro Azul, Ruta Nacional 14 km, 1085, N3313 Cerro Azul, Misiones, Argentina
[2] Instituto Nacional de Tecnología Agropecuaria - Instituto de Recursos Biológicos, Centro de Investigación en Recursos Naturales, De Los Reseros y Dr. Nicolás Repetto s/n, 1686, Hurlingham, Buenos Aires, Argentina
[3] Consejo Nacional de Investigaciones Científicas y Técnicas, Godoy Cruz 2290, Buenos Aires, Argentina
[4] Instituto Nacional de Tecnología Agropecuaria - Estación Experimental Agropecuaria Montecarlo, Av. El Libertador 2472, 3384, Montecarlo, Misiones, Argentina

al. 2002). In Brazil, during the 1990s, three breeding programs were consolidated: Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI) (Floss 1997), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) (Sturion and Resende 1997), and Federal University of Mato Grosso (UFMT) (Costa et al. 2005). The methodology of Linear mixed models (LMM) (Henderson 1984) was successfully employed to maximize genetic gains in Yerba Mate breeding programs in Brazil (Resende 2000, Simeão et al. 2002).

A variant of LMM used to control the environmental heterogeneity within genetic trials is spatial models with a first-order autoregressive residual covariance structure for rows and columns (Gilmour et al. 1997). These spatial models have been widely used in forest tree species such as *Pinus* and *Eucalyptus* (e.g., Costa Silva et al. 2001, Dutkowski et al. 2002, Belaber et al. 2019); however, few reports are available on spatial analysis in Yerba Mate (Resende 2002). The above-referenced forest tree genetic studies using spatial models showed a consistent reduction in the error variance, as well as increases in both heritabilities and accuracies of predicted breeding values in comparison with the classical model based on block design. Despite the relevance of LMM and the study of spatial variation in genetic testing, these genetic selection techniques have not yet been incorporated into Yerba Mate breeding programs in Argentina, hindering, among other things, the accurate identification and selection of individual genotypes based on their breeding value.

The goals of this research were to evaluate and compare the relative efficiency of the spatial model compared to the standard completely randomized design in terms of goodness of fit and changes in the additive genetic, residual, and plot variances, and to apply the univariate and bivariate spatial models for leaf mass weight and crown traits recorded at ages between 21 and 27 years in ten open-pollinated progeny trials of Yerba Mate, to estimate additive genetic variances, heritabilities, and additive genetic correlations between traits and between ages within trials. These genetic parameters were used for discussing the implications for the genetic improvement of Yerba Mate in Argentina.

MATERIAL AND METHODS

Genetic material, field experiment and traits evaluated

The genetic material evaluated corresponds to 241 open-pollinated families planted in 10 genetic trials. Most of this material (239 families) involved phenotypic selections from commercial plantations of 12 provenances from northeastern Argentina, while two selections were from southern Brazil (Figure 1, Table S1). The number of families per site was 25, except for the YM49 trial with 14 and the YM42 trial with 36, and the only genetic linkage between the ten trials was the open-pollinated progeny CA1/74. Ten trials were established between 1990 and 1996 at the Annex Field of INTA located in San Vicente, Misiones (Figure 1). This region is characterized by soils to the Ultisol order, an average annual rainfall of 1,998 mm, and average annual temperature of 20.7 °C. The field experimental design was the same in all trials: a randomized complete block design with three replications and linear plots with ten plants. More details of the ten trials are summarized in Table 1.

Leaf mass weight (LMW) was evaluated in each plant of the 10 trials during the years 2017, 2018, and 2019 (ages 21 to 27 years according to the planting date). In addition, in 2019, the sex of all plants was recorded, and crown height (CH) and crown diameter (CD) were evaluated in the three trials with the highest survival and number of families (YM37, YM46, and YM48). The LMW trait was recorded in

Figure 1. Approximate location of the 14 sampled provenances of Yerba Mate used in the 10 trials performed in Argentina. The rhombus indicates the sites where the 10 trials were planted. Note: 1: Cerro Azul; 2: Candelaria; 3: Cuartel Río Victoria, 4: Campo Viera; 5: Gobernador López; 6: Montecarlo. 7: Oberá; 8: Puerto Esperanza; 9: Puerto Mineral; 10: San José; 11: Santa Ana; 12: Gobernador Virasoro-Corrientes; 13: Guarapuava; 14: Capão do Leão.

kilograms of green leaf per plant (kg per plant) following the mature branch harvesting system. Before harvesting, CH was measured with a graduated stick and CD with a tape measure, from which the crown volume (CV) was calculated according to the following equation $CV = (\pi\ CD^2\ CH)/12$, as reported by Sturion et al. (1999).

Statistical Analysis

The statistical genetic analysis was performed in two stages. First, a univariate analysis was performed to estimate the genetic parameters of LMW, CH, CD, and CV traits for each trial. Second, covariances were assessed through bivariate analysis for LMW values measured at different ages in the same individual, and for pairs of traits measured at the same age within the same individual. The matrix expression of the univariate individual-tree mixed model (animal model) has the following form:

Table 1. Number of families (#Families), percentage of plant survival (%, Survival), date of planting and spacing (m × m) of the 10 trials evaluated

Trials	#Families	Survival (%)[1]	Date of planting	Spacing (m × m)
YM36	25	94.4	11/07/1990	3 × 1.5
YM37	25	95.5	01/08/1990	3 × 1.5
YM42	36	95.3	07/08/1992	3 × 1.5
YM46	25	96.4	07/06/1993	3 × 1.5
YM47	25	94.8	15/06/1993	3 × 1.5
YM48	25	97.2	21/06/1993	3 × 1.5
YM49	14	98.6	28/06/1993	3 × 1.5
YM59	25	81.4	03/08/1995	3 × 1.5
YM62	25	62.6	20/08/1996	2.5 × 1.5
YM63	25	60.6	20/08/1996	2.5 × 1.5

[1] Survival was calculated at the time of the phenotypic evaluations carried out, in the year 2017. Abbreviations used for the trials are described in the text.

$$y = X\beta + Z_a a + Z_p p + e \qquad [1]$$

where y is the vector of individual-tree observations, β is the vector of fixed effects associated with y by the incidence matrix X, which contains the fixed effects of replication and provenances. The Sex variant was excluded due to non-significance in unreported preliminary analyses (p-value > 0.05). The random vector a contains the additive genetic effects of individual trees and it is related to y by the incidence matrix Z_a with $a \sim N(0, A\sigma^2_a)$, where A is the average numerator relationship matrix (Henderson 1984), with σ^2_a being the additive genetic variance. The random vector p contains the plot effects with $p \sim N(0, I\sigma^2_p)$, related to y by the incidence matrix, Z_p, where I is the identity matrix and σ^2_p is the plot variance. Finally, two parameterizations were performed for the term of the error: (1) the residual vector e includes the residual random effects with $e \sim N(0, I\sigma^2_e)$, where σ^2_e is the residual variance (standard model - *Base*); and (2) the residual vector e is divided into two correlation structures ($\xi + \eta$), where ξ refers to spatially correlated residuals and η to independent random residuals (spatial model - *Spa*). The covariance structure of the spatially correlated residuals (ξ) was specified using a first-order autoregressive process for rows (row) and columns (col) (Gilmour et al. 1997). Therefore, the residual matrix R for the *Spa* model is $R = \sigma^2_\xi\ [AR1(\rho_{col}) \otimes AR1\ (\rho_{row})] + I_\eta\ \sigma^2_\eta$ (Dutkowski et al. 2002), where σ^2_ξ is the spatially dependent residual variance, σ^2_η is the independent residual variance, and AR1 (ρ) is the first-order autoregressive structure, where (ρ) is the spatial correlation coefficients for rows (ρ_{row}) and columns (ρ_{col}).

Genetic covariances between pairs of traits were estimated using the following bivariate individual-tree mixed model:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} Z_{a_1} & 0 \\ 0 & Z_{a_2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} Z_{p_1} & 0 \\ 0 & Z_{p_2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \qquad [2]$$

where y_1 and y_2 are the vectors of individual tree observations on traits or ages 1 and 2, respectively. Matrice $X_1 \oplus X_2$, $Z_{a1} \oplus Z_{a2}$ and $Z_{p1} \oplus Z_{p2}$ relate observations to fixed effects in $[\beta'_1 | \beta'_2]$, breeding values in $[a'_1 | a'_2]$, random effects of plot in $[p'_1 | p'_2]$, respectively, and $[e'_1 | e'_2]$ is the residual vector. Symbols $\oplus$ indicate the direct sum of matrices and ' the transpose operation. Expected value and variance-covariance matrix for breeding values are equal to

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{a_{1,1}} & \sigma_{a_{1,2}} \\ \sigma_{a_{2,1}} & \sigma^2_{a_{2,2}} \end{bmatrix} \otimes A \right) \qquad [3]$$

where $\sigma^2_{a_{1,1}}$ and $\sigma^2_{a_{2,2}}$ are the additive genetic variances for the traits or ages 1 and 2, respectively, $\sigma_{a_{1,2}}$ is the additive covariance between traits or ages 1 and 2. The symbol $\otimes$ indicates the Kronecker products of matrices. Expected value and variance-covariance matrix for the plot effects are equal to

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{p_{1,1}} & 0 \\ 0 & \sigma^2_{p_{2,2}} \end{bmatrix} \otimes I \right) \qquad [4]$$

where $\sigma^2_{p_{1,1}}$ and $\sigma^2_{p_{2,2}}$ are the variances of the plot effects for the traits or ages 1 and 2. Finally, the expected value and covariance matrix of the residuals are equal to

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e_{1,1}} & \sigma_{e_{1,2}} \\ \sigma_{e_{2,1}} & \sigma^2_{e_{2,2}} \end{bmatrix} \otimes I\right) \quad [5]$$

where the residual variances for the traits or ages 1 and 2 are $\sigma^2_{e_{1,1}}$ and $\sigma^2_{e_{2,2}}$, and $\sigma^2_{e_{1,2}}$ is the residual covariance between the two traits or ages measured in the same trial. The spatial bivariate analyses for trait and age were performed following a two-step approach (Belaber et al. 2019). In the first step, the detrended data were obtained by subtracting the estimated spatially dependent residual from the univariate spatial model [1] from the measured phenotype. In the second step, the detrended data was analysed using the bivariate model [2] and assuming a residual covariance structure [5].

Genetic parameters and model comparison

The dispersion parameters of the random effects in the mixed model [1] and its spatial variant with a first-order autoregressive residual structure, along with the additive genetic covariances of the model [2] and their respective standard errors, were estimated by the restricted maximum likelihood method (REML; Patterson and Thompson 1971), using the average information algorithm ("Average information", AI) with R software version 3.4.4 (R Core Team 2022), and the statistical package breedR (Rodriguez and Munoz 2016). The statistical significance of both variances and genetic correlations were assessed by the likelihood ratio test (LRT; Stram and Lee 1994). For additive variance, a one-tailed distribution with one degree of freedom was used. In the case of correlations, a two-tailed test with one degree of freedom was employed. The *Base* and *Spa* models were compared using the LRT test with 3 degrees of freedom corresponding to the difference in the number of parameters estimated by both models.

The narrow-sense individual-tree heritability ($\hat{h}^2$) in the *Base* model was estimated according to the following expression: $\hat{h}^2 = \hat{\sigma}^2_a/(\hat{\sigma}^2_a + \hat{\sigma}^2_e)$, where $\hat{\sigma}^2_a$ is the estimate of the additive genetic variance, and $\hat{\sigma}^2_e$ is the estimate of the residual variance. For the calculation of the heritabilities in the *Spa* model, the estimate of the independent residual $\hat{\sigma}^2_\eta$ was used (i.e., $\hat{h}^2 = \hat{\sigma}^2_a/(\hat{\sigma}^2_a + \hat{\sigma}^2_\eta)$). The additive genetic correlations ($\hat{r}_a$) between traits within a trial and between the same trait measured at different ages were estimated with the following equation: $\hat{r}_a = \hat{\sigma}_{a_{1,2}} / \sqrt{\hat{\sigma}_{a_1} \times \hat{\sigma}_{a_2}}$, where $\hat{\sigma}_{a_{1,2}}$ corresponds to the estimated additive genetic covariance between traits 1 and 2 or ages 1 and 2 for the same trait, and $\hat{\sigma}_{a_1}$ and $\hat{\sigma}_{a_2}$ to the estimates of the additive variances of traits (or ages) 1 and 2. The additive genetic coefficient of variation ($\widehat{CV}_a$) was calculated with the expression $\widehat{CV}_a = (\hat{\sigma}_a/\overline{x}) \times 100$, where $\hat{\sigma}_a$ is the additive genetic standard deviation and $\overline{x}$ is the phenotypic population means. Finally, the theoretical accuracy ($\hat{r}$) of the breeding values obtained from the *Base* and *Spa* models was compared using the following expression: $\hat{r} = \sqrt{1 - (PEV/\hat{\sigma}^2_a)}$, where the acronym *PEV* stands for "prediction error variance" of the predicted breeding values, which was calculated following Henderson (1984). In addition, Spearman correlations between the breeding values were calculated to detect possible changes in the genetic rankings of both models.

RESULTS AND DISCUSSION

Model comparison

In this work, standard (*Base*) and spatial (*Spa*) models were used to analyse 39 datasets generated from three evaluations of LMW in 10 trials and one evaluation of CH, CD, and CV traits in three trials. According to the LRT criterion, the *Spa* model provided a better fit than the *Base* model in 87.2% of the analysed datasets (Table S3). Much of the observed efficiency of the *Spa* model, in comparison with the *Base* model, was due to a decrease in residual variance ($\hat{\sigma}^2_e$) and plot variance ($\hat{\sigma}^2_p$). In general, the spatially correlated error ($\hat{\sigma}^2_\zeta$) absorbed most of the $\hat{\sigma}^2_p$ and part of the $\hat{\sigma}^2_\eta$ (Tables 2 and S3). The *Spa* model decreased the residual variance compared to the *Base* model by more than 10% in 69.2% of the analysed datasets, and in the remaining cases, there was generally no change between models. The *Spa* model reduced the $\hat{\sigma}^2_p$ in comparison with the *Base* model in 64.1% of the analysed dataset. The reduction in the residual and plot variances from the *Base* to the *Spa* model has been reported by several authors for growth, stem quality, and branch characteristics in forest species (e.g., Costa Silva et al. 2001, Dutkowski et al. 2006, Ye and Jayawickrama 2008, Cappa et al. 2015, Dong et al. 2020). In Yerba Mate, the only work we have found was Resende (2002), who evaluated LMW in open-pollinated progenies and reported that the spatial model reduced, on average, 40% of the residual variance and 100% of the plot variance compared to the standard model.

In general, the estimated additive genetic variance ($\hat{\sigma}^2_a$) of the traits evaluated showed significant differences between

the *Base* and *Spa* models. In 35.8% of the cases analysed there were increases of more than 10%, and in 25.6% of them there were decreases of less than 10% (Tables 2 and S4). This inconsistency in the behaviour of $\hat{\sigma}^2_a$ between the models has been reported in several studies on forest trees that indicated an increase or decrease in $\hat{\sigma}^2_a$ when the *Base* and *Spa* models were compared (Costa Silva et al. 2001, Dutkowski et al. 2002, Chen et al. 2018, Belaber et al. 2019, Dong et al. 2020). For example, Dong et al. (2020) reported a decrease in the estimated additive genetic variance for diameter at breast height (23.1%) and total height (27.3%), as revealed by spatial analysis. In contrast, Resende (2002), in Yerba Mate, reported average increases of 33% in additive variance estimates with the spatial model. According to Dutkwoski

Table 2. Estimation of additive variance ($\hat{\sigma}^2_a$), plot variance ($\hat{\sigma}^2_p$), independent residual variance ($\hat{\sigma}^2_\eta$), spatially correlated residuals ($\hat{\sigma}^2_\xi$)**,** additive genetic coefficient of variation ($\widehat{CV}_a$), and narrow-sense individual-tree heritability ($\hat{h}^2$) with its respective approximate standard error, obtained for each trial using the spatial model for the traits LMW17, LMW18, LMW19, CH, CD and CV at different ages

Traits	Trials	Ages	$\hat{\sigma}^2_a$	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_\xi$	$\hat{\sigma}^2_\eta$	$\widehat{CV}_a$ (%)	$\hat{h}^2$
LMW 17	YM36	27	1.78(1.47)**	0.83 (0.44)	1.90 (0.80)	16.67(1.61)	17.88	0.10 (0.08)
	YM37	27	0.81(1.05)	0.51 (0.32)	1.84 (0.67)	12.52(1.16)	13.97	0.06 (0.08)
	YM42	25	0.35(0.87)	1.08 (0.34)	4.72 (1.41)	4.96 (1.52)	9.99	0.07 (0.16)
	YM46	24	3.59 (1.97)*	0.42 (0.37)	0.73 (0.42)	9.77 (1.67)	29.47	0.43 (0.23)
	YM47	24	0.16(1.08)	0.91 (0.43)	1.63 (0.60)	10.3 (1.11)	5.81	0.02 (0.01)
	YM48	24	2.29(1.56)	0.87 (0.41)	5.13 (1.77)	4.55 (2.06)	21.81	0.33 (0.22)
	YM49	24	2.39(1.28)**	0.04 (0.03)	9.93 (1.24)	1E^{-4} (0.08)	19.20	0.19 (0.10)[a]
	YM59	22	2.79(2.66)	1.89 (0.89)	3.26 (1.14)	5.85 (2.10)	25.50	0.32 (0.28)
	YM62	21	0.76(1.36)	0.10 (0.42)	2.84 (1.08)	10.19(1.41)	13.86	0.07 (0.12)
	YM63	21	1.04(1.72)	1.23 (0.61)	1.46 (0.74)	7.57 (1.55)	15.45	0.12 (0.19)
LMW 18	YM36	28	0.12(0.39)	0.45 (0.29)	1.19 (0.60)	15.47(1.59)	4.74	0.01 (0.05)
	YM37	28	3.87(2.09)**	0.86 (0.42)	2.32 (0.82)	10.73(1.82)	29.27	0.27 (0.14)
	YM42	26	0.94(0.57)**	0.57 (0.24)	1.15 (0.35)	6.63 (0.58)	19.43	0.13 (0.07)
	YM46	25	3.63 (2.22)*	1.31 (0.48)	7.35 (1.90)	1.56 (2.44)	29.36	0.36 (0.15)
	YM47	25	0.24(0.89)	0.73 (0.34)	1.39 (0.49)	8.33 (0.91)	8.06	0.03 (0.01)
	YM48	25	2.48(1.12)**	0.90 (0.31)	1.23 (0.44)	5.95 (0.95)	27.49	0.29 (0.12)
	YM49	25	1.14 (0.77)*	0.13 (0.23)	8.82 (0.89)	2E^{-4} (0.05)	14.24	0.11(0.08)[a]
	YM59	23	3.52 (2.05)*	0.32 (0.79)	5.90 (1.39)	6.92 (1.72)	27.75	0.34 (0.18)
	YM62	22	3.31 (2.9)	0.10 (0.73)	5.98 (2.10)	15.65(2.73)	18.47	0.17 (0.15)
	YM63	22	2.92(3.46)	0.10 (0.60)	6.23 (1.99)	13.1 (2.87)	18.88	0.18 (0.29)
LMW 19	YM36	29	2.41 (1.56)*	0.61 (0.42)	1.23 (0.65)	15.46(1.61)	20.64	0.14 (0.09)
	YM37	29	2.07(1.70)	0.10 (0.46)	2.67 (0.88)	15.65(1.71)	19.01	0.12 (0.09)
	YM42	27	3.27(1.46)**	0.44 (0.31)	5.36 (1.89)	6.46 (2.18)	25.98	0.34 (0.15)
	YM46	26	1.92 (1.46)*	0.34 (0.39)	10.39(2.78)	5.68 (2.86)	18.90	0.11 (0.08)
	YM47	26	2.30 (1.39)*	0.23 (0.32)	2.19 (0.72)	10.13(1.35)	23.33	0.19 (0.11)
	YM48	26	2.52(1.33)**	0.50 (0.28)	1.28 (0.51)	8.18 (1.15)	25.56	0.24 (0.10)
	YM49	26	2.05 (1.44)*	0.52 (0.39)	10.90(1.39)	2E^{-4} (0.01)	17.85	0.16(0.11)[a]
	YM59	24	2.35(2.09)	1.08 (0.65)	3.42 (1.03)	6.92 (1.70)	24.81	0.25 (0.21)
	YM62	23	1.48(1.27)	0.10 (0.27)	3.99 (1.02)	7.20 (1.33)	18.86	0.17 (0.14)
	YM63	23	4.60(3.41)	0.12 (0.13)	4.12 (1.58)	10.14(2.80)	28.11	0.31 (0.21)
CH	YM37	29	0.04(0.03)*	0.01 (0.00)	0.03 (0.01)	0.22 (0.03)	6.94	0.16 (0.10)
	YM46	26	0.05(0.03)**	6E^{-3} (7E^{-3})	0.01 (9E^{-3})	0.16 (0.02)	7.87	0.24 (0.14)
	YM48	26	0.18(0.07)**	2E^{-3} (7E^{-3})	0.06 (0.01)	0.06 (0.05)	14.58	0.74 (0.23)
CD	YM37	29	0.04(0.03)	0.02 (0.01)	0.04 (0.01)	0.20 (0.03)	7.97	0.16 (0.13)
	YM46	26	0.02 (0.01)*	2E^{-3} (6E^{-3})	0.02 (9E^{-3})	0.19 (0.01)	5.50	0.08 (0.09)
	YM48	26	0.08(0.04)**	0.01 (0.01)	0.05 (0.01)	0.13 (0.03)	10.17	0.38 (0.18)
CV	YM37	29	0.01(0.01)*	2E^{-3} (1E^{-3})	7E^{-3} (2E^{-3})	0.03 (0.01)	16.39	0.16 (0.13)
	YM46	26	0.01(0.01)**	8E^{-4} (9E^{-4})	3E^{-3} (1E^{-3})	0.02 (0.01)	16.13	0.22 (0.14)
	YM48	26	0.02(0.01)**	1E^{-3} (1E^{-3})	0.01 (2E^{-3})	0.01 (0.01)	20.80	0.72 (0.26)

[a] $\hat{h}^2$ was estimated using the $\hat{\sigma}^2_\xi$, as the $\hat{\sigma}^2_\eta$ approached zero. Significance effects are denoted as follows: * Statistically significant ($0.01<p<0.05$), ** Statistically significant ($p<0.01$).

et al. (2002), such inconsistency in the additive genetic variance could be due to high independent errors. In this study, inconsistencies in $\hat{\sigma}^2_a$ when moving from the *Base* to the *Spa* model may be also due to high independent errors, which accounted for more than 60% of the total variation in 53.3% of cases. The spatially correlated error term absorbed over 20% of the total variation in 43.6% of cases. Regarding the standard error of $\hat{\sigma}^2_a$, for the LMW trait, the *Spa* model was associated with lower estimates compared to the *Base* model in 83.3% of the cases. However, no significant changes were observed for crown traits, except for a 25% reduction in the standard error of $\hat{\sigma}^2_a$ for CH trait in the YM36 trial (Tables 2 and S4).

Spearman correlations between the breeding values obtained with both models were generally high (≥ 0.95); however, 28% of them were lower than 0.95, indicating differences between the genetic rankings (Table S5). The changes in ranking between models indicate that each model will select different individuals, which will affect the expected genetic gains. For example, when selecting the 10 best individuals for trait LMW18 in trial YM59, the proportion of common trees selected by both models was 60% (data not shown). In addition, the average accuracy of breeding values from parents and offspring estimated with the *Spa* model were higher than the corresponding values estimated with the *Base* model (averaging 3% for parents and 5% for offspring, Table S5). Therefore, based on the better fit (LRT test) of the *Spa* model and its generally higher accuracy of breeding values compared to the *Base* model, this article presented and discussed the genetic parameters obtained using the *Spa* model. Furthermore, results from the *Base* model are included as supplementary material.

Table 3. Estimated additive genetic correlations (approximate standard error) for each trial using the bivariate individual-tree mixed model (2), for the traits LMW17, LMW18, LMW19, CH, CD and CV at different ages

Trials	Trait	LMW18	LMW19	CD	CV
YM36	LMW17	0.69 (0.13)**	0.93 (0.06)**		
	LMW18		0.93 (0.37)*		
YM37	LMW17	0.83 (0.08)**	0.64 (1.19)		
	LMW18		0.79 (0.62)**		
	CH		0.96 (0.15)**	0.67 (0.43)	0.93 (0.26)**
	CD		0.88 (0.64)**		0.89 (0.13)**
	CV		0.93 (0.25)**		
YM42	LMW17	0.50 (2.58)	0.93 (1.14)*		
	LMW18		0.82 (0.45)**		
YM46	LMW17	0.74 (0.86)**	0.67 (0.65)**		
	LMW18		0.57 (0.94)*		
	CH		0.96 (1.29)	0.77 (0.81)**	0.90 (0.38)**
	CD		0.18(1.56)		0.76 (0.80)**
	CV		0.66 (0.61)**		
YM47	LMW17	0.28 (0.16)*	0.82 (0.10)*		
	LMW18		0.90 (0.20)*		
YM48	LMW17	0.90 (0.26)**	0.89 (0.26)**		
	LMW18		0.92 (0.11)**		
	CH		0.87 (0.70)**	0.70 (0.23)**	0.91 (0.15)**
	CD		0.90 (0.56)**		0.92 (0.19)**
	CV		0.89 (0.28)**		
YM49	LMW17	0.93 (0.47)*	0.96 (0.27)**		
	LMW18		0.91 (1.03)*		
YM59	LMW17	0.78 (0.28)**	0.97 (0.06)**		
	LMW18		0.96 (0.26)**		
YM62	LMW17	0.32 (2.16)	0.64 (0.33)*		
	LMW18		0.84 (0.75)*		
YM63	LMW17	0.91 (0.08)**	0.91 (0.61)*		
	LMW18		0.89 (0.31)**		

Significance effects were tested with respect to 0 and are denoted as: * Statistically significant ($0.01<p<0.05$), ** Statistically significant ($p<0.01$).

Genetic variance, heritability, and additive genetic coefficient of variation

Overall, estimates of the additive genetic variance ($\hat{\sigma}^2_a$) for LMW and crown traits were significantly different from zero and accounted for between 0.7% and 26.2% of the total variation. Standard errors of $\hat{\sigma}^2_a$ were in general high, but lower than the parameter estimates in 74.4% of the dataset (Table 2). When the $\hat{\sigma}^2_a$ of LMW trait was analysed across ages over the three-year evaluation period (year 2017, LMW17; year 2018, LMW18, year 2019, LMW19), we detected an increased behaviour for trials YM42, YM47, YM48 and YM63, with a greater increment occurring in the YM63 trial, where the $\hat{\sigma}^2_a$ increased with age (2017, $\hat{\sigma}^2_a$ =1.04; 2018, $\hat{\sigma}^2_a$ =2.92; 2019, $\hat{\sigma}^2_a$ =4.60). In contrast, no consistent behaviour of $\hat{\sigma}^2_a$ was observed for the remaining trials. For example, in the YM49 trial, the $\hat{\sigma}^2_a$ decreased by 48% between 2017 and 2018, and then increased by 180% between 2018 and 2019.

The narrow-sense individual-tree heritabilities ($\hat{h}^2$) for the LMW trait were low to moderate, ranging from 0.01 to 0.43, with an average value of $\hat{h}^2$ =0.19 (Table 2). In general, trials YM48 and YM59 displayed the highest and most similar $\hat{h}^2$ values for the LMW trait across years, which is related to higher $\hat{\sigma}^2_a$ and lower $\hat{\sigma}^2_e$ compared to the other trials. The lowest $\hat{h}^2$ was observed in different trials according to the year of evaluation of the LMW trait (LMW17 for trial YM47, $\hat{h}^2$ =0.02; LMW18 for trial YM36, $\hat{h}^2$ =0.01; LMW19 for trial YM46, $\hat{h}^2$ =0.11) resulting from lower $\hat{\sigma}^2_a$ and intermediate $\hat{\sigma}^2_e$ compared to the other trials. The low to moderate $\hat{h}^2$ values obtained for the LMW trait were comparable to those reported by other authors in open-pollinated Yerba Mate progenies between 3 and 18.5 years of age (Sturion et al. 1999, Resende et al. 2000, Rosse and Fernandes 2002, Floss et al. 2003, Sturion and Resende 2005, Sturion et al. 2017). For example, Sturion et al. (2017) estimated an individual heritability of 0.17 for LMW at 18.5 years of age, a value and age similar to those of the present study. However, in contrast to our findings, Wendling et al. (2018) reported higher heritability estimates for the LMW trait at various ages, including 0.59 at age 2.5, 0.79 at age 4.5, 0.88 at age 6.5, and 0.65 at age 18.7. Concerning the crown traits, low to high values of $\hat{h}^2$ were obtained (0.08 and 0.74), with average values of 0.38 for CH, 0.21 for CD, and 0.37 for CV (Table 2). These heritabilities were comparable to those reported by Rosse and Fernandes (2002) at 4 years of age, namely 0.37, 0.41, and 0.21 for the CH, CD, and CV traits, respectively. In contrast, Sturion et al. (1999) obtained lower heritabilities than those reported in this study at the age of 5.8 years (0.05 for CH, 0.02 for CD, and 0.07 for CV).

The additive genetic coefficient of variation ($\widehat{CV}_a$) for the LMW trait showed values between 4.74% and 29.47%, and 86.6% of them were higher than 10% (Table 2). According to Sebbenn et al. (1998), coefficients higher than 10% are considered high, which would indicate high additive genetic variation. In general, the $\widehat{CV}_a$ values obtained in this study for LMW were lower than those reported in other studies (Sturion et al. 1999, Rosse and Fernandes 2002, Simeão et al. 2002, Floss et al. 2003, Sturion et al. 2017, Wendling et al. 2018). For example, Sturion et al. (2017) reported a $\widehat{CV}_a$ of 37.2% for LMW at age of 18.5 years. However, Floss et al. (2003) reported a $\widehat{CV}_a$ of 22.5% and 11.6% when evaluating LMW at 6 years in the trials from two provenances. Overall, the $\widehat{CV}_a$ estimates for CH and CD were below 10%, indicating low levels of additive genetic variation for these traits. In contrast, the $\widehat{CV}_a$ estimate for CV was higher than 16%. However, it should be noted that trait CV was created from multiplicative combinations of traits CH and CD, resulting in $\widehat{CV}_a$ higher than those obtained from the original variables. The $\widehat{CV}_a$ for the crown traits obtained in this study were lower than the values reported by other authors at younger ages (Sturion et al. 1999, Rosse and Fernandes 2002, Costa et al. 2005). For example, Costa et al. (2005) evaluated ages before pruning and obtained higher coefficients of variation (14% and 62.2% for CH and CD, respectively). However, Sturion et al. (1999) obtained values of 24.2% for CH, 28.9% for CD, and 59.2% for CV at the age of 5.8 years. The low values of $\widehat{CV}_a$ reported in this study for crown traits could be related to a high number of harvesting and crown management interventions in each trait (one per year), a situation that generates a greater uniformity among the crowns of the individuals within the trial.

Additive genetic correlations

In general, the additive genetic correlations ($\hat{r}_a$) for the LMW evaluated at different ages during 2017, 2018, and 2019 were statistically significant and moderate to high, with values ranging from 0.50 to 0.97 (Table 3). The $\hat{r}_a$ values were low only in the LMW17-LMW18 evaluations in YM47 ($\hat{r}_a$=0.28) and YM62 trials ($\hat{r}_a$=0.32). In general, lower $\hat{r}_a$ than those found in the present study were reported by Sturion and Resende (2005) for LMW assessed at ages 2, 4, and 6 years (mean $\hat{r}_a$=0.40). Similarly, Wendling et al. (2018) obtained lower estimates than those of this study by correlating ages 2.5 and 18.7 ($\hat{r}_a$=0.09) and 4.5 and 18.7 ($\hat{r}_a$=0.41). However, these authors reported values similar to our correlations at

ages 6.5 and 18.7 ($\hat{r}_a$=0.90). For the crown traits, the $\hat{r}_a$ values were positive, significant, and high ($\hat{r}_a \geq 0.70$); the only non-significant and moderate $\hat{r}_a$ was between CH and CD ($\hat{r}_a$=0.67) in trial YM37 (Table 3). Similar values to those obtained in this study were reported in Yerba Mate plants at 3 years of age by Rosse and Fernandes (2002), who obtained an $\hat{r}_a$= 0.97 between CH and CD traits, and a $\hat{r}_a$= 0.99 between CH and CV and between CD and CV traits. Finally, the $\hat{r}_a$ values between crown traits and LMW19 evaluated in three trials (YM37, YM46, and YM48) were positive, generally significant, and high ($\hat{r}_a \geq 0.87$) (Table 3). Similar results were reported by Rosse and Fernandes (2002) with values of $\hat{r}_a$=0.91 between CH and LMW, $\hat{r}_a$= 0.94 between CD and LMW, and $\hat{r}_a$= 0.99 between CV and LMW. In contrast, Sturion et al. (1999) reported lower correlations at 5.8 years of age than those found in this study (CH-LMW $\hat{r}_a$= 0.31; CD-LMW $\hat{r}_a$= 0.38 and CV-LMW $\hat{r}_a$= 0.58). In summary, the $\hat{r}_a$ values between ages for the same trait and between traits obtained in this study were generally high and significant. This indicates a similar behaviour of the genotypes over the years evaluated and demonstrates that indirect selections for LMW through the crown traits are possible.

Implications for Yerba Mate Breeding in Argentina

The analysis of data from 10 half-sib Yerba Mate progeny trials in this study showed that accounting for environmental heterogeneity (*Spa* model) consistently reduced non-genetic variation and improved breeding value accuracy for LMW and crown traits compared to the *Base* model. The presence of significant additive genetic variation ($\widehat{CV}_a$) suggests that selecting for general combining ability could effectively enhance LMW production. The strong additive genetic correlation ($\hat{r}_a$) between LMW and crown traits indicates that indirect selection of LMW using crown traits, particularly CV, could be effective. However, further evaluations are needed to confirm and obtain more precise information on these trait relationships. The high $\hat{r}_a$ observed over three consecutive years for LMW suggests consistent genotype performance in adulthood. However, determining the juvenile-adult genetic correlation is crucial for accelerated breeding and early selection. Additionally, the genotype by site interaction and the suitability of selected genetic material for different Yerba Mate-growing regions require further investigation. To address these issues, new trials with strong genetic connections will be conducted across diverse site conditions. Furthermore, the increased tree density per hectare in the new Yerba Mate paradigm in Argentina raises questions about inter-tree genetic and environmental competition within the INTA's Yerba Mate breeding program.

ACKNOWLEDGEMENTS

Supplementary Tables are available from the corresponding author.

REFERENCES

Belaber EC, Gauchat ME, Rodríguez GH, Borralho NM and Cappa EP (2019) Estimation of genetic parameters using spatial analysis of *Pinus elliottii* Engelm. var. *elliottii* second-generation progeny trials in Argentina. **New Forests 50**: 605-627.

Cappa EP, Muñoz F, Sanchez L and Cantet RJC (2015) A novel individual-tree mixed model to account for competition and environmental heterogeneity: a Bayesian approach. **Tree Genetics and Genomes 11**: 1-15.

Chen Z, Helmersson A, Westin J, Karlsson B and Wu HX (2018) Efficiency of using spatial analysis for Norway spruce progeny tests in Sweden. **Annals of Forest Science 75**: 2.

Coelho C, Araujo MT and Schenkel E (2002) Populational diversity on leaf morphology of maté (*Ilex paraguariensis* St. Hil., Aquifoliaceae). **Brazilian Archives of Biology and Technology 45**: 47-51.

Costa RB, Resende MDV, Contini AZ, Rego FLH, Roa RAR and Martins WJ (2005) Avaliação genética dentro de indivíduos de erva-mate (*Ilex paraguariensis* St. Hil.), na região de Caarapó, MS, pelo procedimento REML/BLUP. **Ciência Florestal 15**: 371-376.

Costa Silva J, Dutkowski GW and Gilmour AR (2001) Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. **Canadian Journal of Forest Research 31**: 1887-1893.

Dong L, Xie Y, Wu HX and Sun X (2020) Spatial and competition models increase the progeny testing efficiency of Japanese larch. **Canadian Journal of Forest Research 50**: 1373-1382.

Dutkowski GW, Costa e Silva J, Gilmour AR and Lopez GA (2002) Spatial analysis methods for forest genetic trials. **Canadian Journal of Forest Research 32**: 2201-2214.

Dutkowski GW, Costa e Silva J, Gilmour AR, Wellendorf H and Aguiar A (2006) Spatial analysis enhances modeling of a wide variety of traits in forest genetic trials. **Canadian Journal of Forest Research 36**: 1851-1870.

Floss PA (1997) Programa de melhoramento genético da erva-mate na EPAGRI. In **Congresso sul-americano da erva-mate.** EMBRAPA, Colombo, p. 279-284.

Floss PA, Croce DM, Sturion JA (2003) Desenvolvimento de duas procedências de erva-mate na região de Chapecó - SC. In **Congresso**

sul-americano da erva-mate 3. Epagri, Chapecó, p. 1-4.

Gilmour AR, Cullis BR and Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiment. **Journal of Agricultural, Biological and Enviromental Statistics 2**: 269-293.

Henderson CR (1984) Estimation of variances and covariances under multiple trait models. **Journal of Dairy Science 67**: 1581-1589.

INYM - Instituto Nacional de Yerba Mate (2022) Superficie cultivada por departamento. Available at <https://inym.org.ar/>. Accessed on January 4, 2022.

Patterson HD and Thompson R (1971) Recovery of inter-block information when block sizes are unequal. **Biometrika 58**: 545-554.

Prat Kricun SD (2013) Mejoramiento genético de la yerba mate en la estación experimental Agropecuaria Cerro Azul – Período 1974-2011. 56p. (**Circular técnica, 68**).

R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at <http://www. R-project. org/>. Accesed on February 15, 2022.

Resende MDV (2002) **Genética biométrica e estatística no melhoramento de plantas perenes.** Embrapa Informação Tecnológica, Brasilia, 975p.

Resende MDV, Sturion JA, Carvalho AP, Simeao RM and Fernandes JSC (2000) Programa de Melhoramento da Erva-Mate coordenado pela Embrapa: resultados da avaliação genética de populações, progênies, indivíduos e clones. Embrapa Florestas, Colombro, 66p. (**Circular técnica, 43**).

Rodriguez LS and Muñoz F (2016) breedR: statistical methods for forest genetic resources analysts. **IUFRO Genomics and Forest Tree Genetics**. Arcachon, France, p. 1-135.

Rosse LN and Fernandes JC (2002) Escolha de caracteres para o melhoramento genético em erva-mate por meio de técnicas multivariadas. **Ciência Florestal 12**: 21-27.

Scherer R, Urfer P and Mayol M (2002) Inheritance studies of caffeine and theobromine content of Mate (*Ilex paraguariensis*) in Misiones, Argentina. **Euphytica 126**: 203-210.

Sebbenn AM, Siqueira ACM, Kageyama PY and Machado JAR (1998) Parâmetros genéticos na conservação da cabreúva – *Myroxylon peruiferum* L.F. Allemão. **Scientia Forestalis 53**: 31-38.

Simeão RM, Sturion JA, Resende MDV, Fernandes JSC, Neiverth DD and Ulbrich AL (2002) Avaliação genética em erva-mate pelo procedimento BLUP individual multivariado sob interação genótipo x ambiente. **Pesquisa Agropecuária Brasileira 37**: 1589-1596.

Stram D and Lee JW (1994) Variance components testing in the longitudinal mixed effects model. **Biometrics 50**: 1171-1177.

Sturion JÁ and Resende MDV (1997) Programa de melhoramento genético da erva-mate no Centro Nacional de Pesquisa de Florestas da Embrapa. In **Congresso sul-americano da erva-mate 1.** EMBRAPA, Colombo, p. 285-298.

Sturion JÁ and Resende MDV (2005) Seleção de progênies de erva-mate (*Ilex paraguariensis* St. Hill) para produtividade, estabilidade e adaptabilidade temporal de massa foliar. **Boletim de Pesquisa Florestal 50**: 37-51.

Sturion JA, Resende MDV and Carpanezzi AA (1999) Controle genético e estimativa de ganho genético para peso de massa foliar em erva-mate (*Ilex paraguariensis* St.Hil). **Boletim de Pesquisa Florestal 38**: 5-12.

Sturion JA, Stuepp CA and Wendling I (2017) Genetic parameters estimates and visual selection for leaves production in *Ilex paraguariensis*. **Bragantia 76**: 492-500.

Wendling I, Sturion JA, Stuepp CA, Fioravante Reis CA, Ramalho Patto MA and Resende MDV (2018) Early selection and classification of yerba mate progenies. **Pesquisa Agropecuária Brasileira 53**: 279-286.

Ye TZ and Jayawickrama KJS (2008) Efficiency of using spatial analysis in first-generation coastal Douglas-fir progeny tests in the US Pacific Northwest. **Tree Genetic and Genomes 4**: 677-692.