FILIPE MANOEL FERREIRA

PREDICTING CLONAL COMPOSITES PERFORMANCE AND ENHANCING EUCALYPTUS PRODUCTIVITY BY ACCOUNTING FOR INDIRECT GENOTYPIC EFFECTS

Thesis submitted to the Genetics and Breeding Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Leonardo Lopes Bhering

Co-adviser: Kaio Olimpio da Graças Dias

VIÇOSA - MINAS GERAIS
2023

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
F383p
2023

Ferreira, Filipe Manoel, 1993-
Predicting clonal composites performance and enhancing eucalyptus productivity by accounting for indirect genotypic effects / Filipe Manoel Ferreira. – Viçosa, MG, 2023.
1 tese eletrônica (53 f.): il. (algumas color.).

Orientador: Leonardo Lopes Bhering.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Biologia Geral, 2023.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2023.526
Modo de acesso: World Wide Web.

1. Eucalipto - Melhoramento genético. 2. Genética quantitativa. 3. Modelos multiníveis (Estatísticas). 4. Competição (Biologia). I. Bhering, Leonardo Lopes, 1980-. II. Universidade Federal de Viçosa. Departamento de Biologia Geral. Programa de Pós-Graduação em Genética e Melhoramento. III. Título.

CDD 22. ed. 634.973766

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB-6/2523

FILIPE MANOEL FERREIRA

PREDICTING CLONAL COMPOSITES PERFORMANCE AND ENHANCING EUCALYPTUS PRODUCTIVITY BY ACCOUNTING FOR INDIRECT GENOTYPIC EFFECTS

Thesis submitted to the Genetics and Breeding Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: August 9, 2023.

Assent:

Documento assinado digitalmente
FILIPE MANOEL FERREIRA
Data: 30/08/2023 16:17:49-0300
Verifique em https://validar.iti.gov.br

Filipe Manoel Ferreira
Author

Documento assinado digitalmente
LEONARDO LOPES BHERING
Data: 30/08/2023 16:23:17-0300
Verifique em https://validar.iti.gov.br

Leonardo Lopes Bhering
Adviser

To my mother, Claudinéia, for her love and hard work that moved financial and demographic mountains, making it possible for me to walk the road of science.

DEDICATION

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my mother Claudinéia, for her love and support that made this day possible. I also, thank my father Afonso, for the teachings, and my sister Fernanda for all the laughs and tears we have shared and for my niece Helena. Also, I want to thank my grandparents for all the love and care, and Luiza, my love, for teaching me how to be a better person.

I am truly grateful to Professor Leonardo Bhering for his invaluable guidance and support throughout the entire duration of my master and doctorate. His expertise, patience, and encouragement have been instrumental in awakening my interest to this wonderful area of knowledge that is quantitative genetics. Also, I extend my heartfelt appreciation to Professor Kaio Dias, for sharing his expertise, patience, and encouragement to me and many others and help me to become a better scientist.

I also extend my appreciation to the members of my thesis committee – Dr. Cosme Damião Cruz, Dr. Evandro Vagner Tambarussi, and Dr. Rodrigo Silva Alves. Their insightful comments and suggestions have significantly contributed to the improvement of my research. I am sincerely grateful for the time and effort they have invested in reviewing and evaluating my work. Additionally, I would like to thank the company Celulose Nipo-Brasileira S.A. (CENIBRA) for providing the data that made this study possible.

I am incredibly fortunate to have had the opportunity to collaborate with many amazing fellow researchers and colleagues, and to be able to call some of them friends. Their intellect, enthusiasm, and willingness to engage in stimulating discussions have enriched my research experience. I am immensely grateful that our paths crossed, and we could learn together. Thank you very much to all members of the UFV biometrics laboratory, specially Saulo Chaves who contributed a lot to the development of this thesis.

Furthermore, I am indebted to the staff and faculty of Federal University of Viçosa (UFV) for their ongoing support and provision of necessary resources. I would also like to thank Dr. Juarez Campolina and professor Evandro Tambarussi for the partnership established in several works during my doctorate.

I would like to acknowledge the financial assistance provided by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Without their support, conducting this research would have been significantly more challenging.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Finally, I would like to express my profound gratitude to all the professors who have graced my life with their wisdom, guidance, and unwavering commitment to education. "Education is the most powerful weapon you can use to change the world." (Nelson Mandela).

“Time is really the only capital that any human being has, and the only thing he can’t afford to lose.” (Thomas Edison

ABSTRACT

FERREIRA, Filipe Manoel, D.Sc., Universidade Federal de Viçosa, August, 2023. **Predicting clonal composites performance and enhancing eucalyptus productivity by accounting for indirect genotypic effects.** Adviser: Leonardo Lopes Bhering. Co-adiviser: Kaio Olimpio das Graças Dias.

Biotic and abiotic factors have been making it difficult to maintain high rates of realized genetic gain for tree species, especially those cultivated in monocultures. Planting a mixture of genotypes or clonal composites (CC) can be an alternative to increase phytosanitary security and even the productivity of forest plantations. Clones grown in CC may present residual and genetic competition. Competition effects can affect the heritable portion of the total variability and impact the genetic progress of the population under selection. We aim to jointly model the spatial and genetic competition using a linear mixed model at the spatial and genetic level (SCM) to estimate genetic parameters and study the impacts of intergenotypic competition. In addition, we propose a strategy to predict the best combination of clones to compose a CC that has not yet been planted. To the best of our knowledge, no previous study has explored the prediction of CC accounting for competition effects. The main advantage of our methodology consists in modeling the competition at the genetic and residual level to predict the total genotypic value (TGV) of clones and the phenotypic performance of any CC combination. The proposed approach was illustrated in a dataset from clonal trials of eucalyptus in a randomized block design with 24 replications, containing a single tree per plot evaluated for mean annual increment (MAI – $m^3ha^{-1}ano^{-1}$) at ages 3 and 6. The fitted model was efficient in partitioning genetic variation into variations due to direct genotypic effects (DGE) and indirect or competition genotypic effects (IGE). Additionally, we proposed a way to classify clones as aggressive, homeostatic, and sensitive based on the magnitude of the IGE. The SCM was the most suitable according to the Akaike Information Criterion. By accounting for indirect genotypic effects, for MAI, the total heritability decreased from 0.25 to 0.10 for 3 years and from 0.30 to 0.14 for 6 years, compared to a reduced model for IGE. Therefore, heritability was overestimated when IGE was not considered. Based on the TGV, we were able to identify CC with a high expected average performance for MAI, considering the trade-off between DGE and IGE. Therefore, predicting CC by capitalizing on the IGE can provide a strategic advantage in recommending the best combination of clones to be planted.

Keywords: Tree Breeding. Quantitative Genetics. Linear Mixed Models. Associative Effects. Competition.

RESUMO

FERREIRA, Filipe Manoel, D.Sc., Universidade Federal de Viçosa, agosto de 2023. **Predição da performance de compostos clonais para melhorar a produtividade de eucalyptus considerando os efeitos genotípicos indiretos**. Orientador: Leonardo Lopes Bhering. Coorientador: Kaio Olimpio das Graças Dias.

Fatores bióticos e abiótico vem dificultando a manutenção das altas taxas de ganho genético para espécies arbóreas, principalmente as cultivadas em monoculturas. O plantio de mistura de genótipos ou compostos clonais (CC) pode ser uma alternativa para aumentar a segurança fitossanitária e até mesmo a produtividade dos plantios florestais. Os clones cultivados em CC podem apresentar competição a nível residual e genético. Os efeitos de competição podem afetar a porção herdável da variabilidade total e impactar no progresso genético da população sob seleção. O presente estudo tem como objetivos ajustar um modelo linear misto de competição a nível espacial e genético (MCEG) para estimar parâmetros genéticos e estudar os impactos da competição intergenotípica. Além disso, visa propor uma estratégia para prever a melhor combinação de clones para compor um CC que ainda não foram plantados. Aparentemente, a exploração dos efeitos de competição na recomendação da melhor combinação de clones para a composição de CC ainda não foi explorada em estudos prévios. A maioria destes estudos restringe as inferências apenas aos tradicionais modelos lineares mistos de competição. A principal vantagem da metodologia apresentada consiste em modelar a competição a nível genético e residual para prever o valor genotípico total (VGT) de um CC que ainda não foi plantado. A abordagem proposta foi ilustrada em um conjunto de dados de experimentos de testes clonais de eucalyptus em um delineamento de blocos ao acaso com 24 repetições, contendo uma única árvore por parcela avaliada para incremento médio anual (IMA – $m^3ha^{-1}ano^{-1}$) aos 3 e 6 anos. O modelo ajustado foi eficiente em particionar a variação genética em variação devido à efeitos genéticos diretos (EGD) e efeitos genéticos indiretos ou de competição (EGC). Adicionalmente, foi proposta uma forma de classificar os clones como agressivos, homeostáticos e sensíveis com base na magnitude dos EGC. O MCEG foi o mais adequado de acordo com o Critério de Informação de Akaike. A herdabilidade total diminuiu de 0,25 para 0,10 aos 3 anos e de 0,30 para 0,14 aos 6 anos em comparação com um modelo reduzido para os EGC. Portanto, a herdabilidade foi superestimada quando EGC não foram considerados. Com base nos VGT foi possível a identificação de CC com alto desempenho médio esperado para IMA levando em conta o equilíbrio entre EGD e EGC. Em resumo, a utilização dos EGC na determinação das melhores combinações de clones para a composição de CC fornece vantagens estratégicas nas etapas de recomendação de misturas de clones.

Palavras-chave: Melhoramento Florestal. Genética Quantitativa. Modelos Lineares Mistos. Efeitos Associativos. Competição.

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

IGE	Indirect Genotypic Effects
DGE	Direct Genotypic Effects
TGV	Total Genotypic Values
SM	Spatial Models
SCM	Spatial Competition Models
MET	Multiple Environmental Trials
CC	Clonal Composities
CIF	Competition Intensity Factor
MAI	Mean Annual Increment
AR1	First-order Autoregressive Process

LIST OF SYMBOLS

λ	Lower case greek letter lambda
ζ	Lower case greek letter zeta
σ	Lower case greek letter sigma
ρ	Lower case greek letter rho
$\otimes$	Math symbol kronecker product
$\sim$	Math symbol tilde

SUMMARY

1. INTRODUCTION

1.1. Associative effects

In cases of social interactions, an individual's phenotype is not only influenced by its own genes (direct genetic effects - DGE) but also by the genes expressed in its social partners (indirect genetic effects- IGE). Social or associative effects play a crucial role in unraveling the dynamics of interactions within a group and their impact on the artificial and natural selection of genotypes. These interactions can result in positive or negative changes in the mean of traits influenced by interactions among individuals of the same species (BIJMA, 2014).

Cooperation or altruism, and competition are examples of how the social environment can influence on individual and group performance and fitness. Altruistic behaviors, such as alarm sounds in some bird species, exemplify how certain traits, that may decrease individual fitness, have been evolutionarily preserved due to their advantages for the group (WALSH; LYNCH, 2018). In turn, competition effects can reduce or increase the trait mean of a given individual depending on the covariance between the DGE and IGE (COSTA E SILVA et al., 2017) .This covariance plays a crucial role in determining the influence of IGE on the magnitude of the heritable variation that exists at the population level for traits influenced by associative effects (BIJMA, 2011a; GRIFFING, 1967).

In some situations, like plantations or confined spaces, individuals often have conspecific neighbors with whom they compete for resources. The Costa e Silva and Kerr (2013) proposal derived from Cappa and Cantet (2008) original idea can help estimate the competition variance component and the covariance between DGE and IGE. This methodology calculates a mean competition intensity factors (f_{ij}) that a focal individual i exerts over its j neighbors $j = j_1, j_2, \ldots j_m$, being m the maximum number of first neighbors for a focal individual i (Figure 1 - extracted from Ferreira et al., (2023) The f_{ij} will vary according to the distance between these individuals in the field and the number of missing neighbors (Figure 1a). The magnitude of f_{ij} is related to the inverse of the distance between individual i and its m neighbors. In this context, IGE can be regarded as the sum of effects that the m neighbors exert over the focal individual i (Figure 1b) or as the mean IGE that the focal individual i exert over its neighbors (Figure 1c).

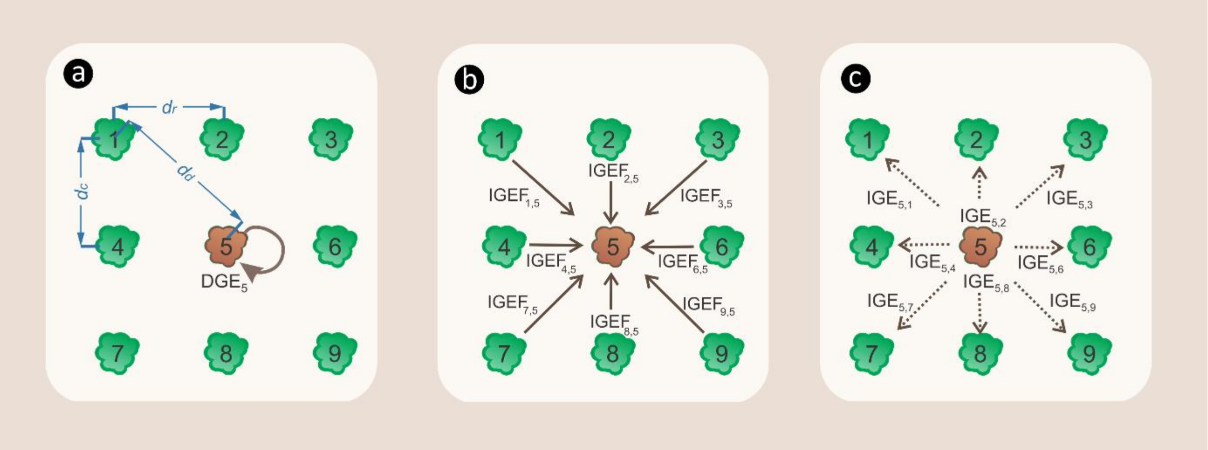

Figure 1: Illustration of (a) direct genetic effect (DGE) and (b) indirect genetic effects (IGE) of the j neighbors on the focal individual *i* and (c) IGE of a focal individual i on its *j* neighbors. The focal individual *i* stands out by a different color and is represented by the number 5; its *j* neighbors are represented by the remaining numbers (the number of neighbors can vary from 0 to *m* depending on the position of focal individual *i*; d_r, d_c, and d_d are the distance between the focal tree and its neighbors in the row, column and diagonal, respectively.

The phenotypic performance of individuals under competition can be explained by the expression of its genes (DGE_i) combined with the influence received from its *m* neighboring individuals' gene expression (IGE_j). Therefore, the competitive capacity of individual *i* will only be perceived in the phenotype of its *m* neighbors, similar to that observed in maternal effect in mammals, which are observed only in the offspring (WALSH; LYNCH, 2018). The total breeding value (TBV) combines the DGE_i and the mean IGE that the focal individual *i* exerts over its *m* neighbors. Therefore, only the phenotype of the neighbors provides information about the TBV_i.

When the covariance between DGE and IGE is significant, selecting solely based on DGE may lead to undesirable changes in the population average. Therefore, it is recommended to select based on total breeding value (TBV) to ensure that the gains are in the desired direction (BIJMA, 2014). Additionally, considering kinship is important since a higher degree of relatedness allows for the exploitation of a greater fraction of the total heritable variation (COSTA E SILVA et al., 2017). Highly competitive individuals can potentially reduce the average production of a productive unit due to their DGE being overshadowed by negative IGE. Thus, considering IGE in breeding can enhance the success of selecting desirable genotypes for quantitative traits, since it allows the exploration of previously unexplained variability.

1.2. Spatial and genetic competition

In certain breeding situations, the presence of correlations between plots can be attributed to spatial or genetic competition. Spatial competition can arise from factors such as management practices, soil fertility or moisture differences, disease incidence, and inadequate experimental design, among others (BURGUEÑO, 2018; FAVERI et al., 2015).

These factors can reduce the efficiency of the selection. Traditionally, blocking has been proposed to address spatial variation in field experiments (FISHER; OTHERS, 1925). However, blocking does not control small-scale variation, since it assumes complete homogeneity within the block and independence among plots (SELLE et al., 2019). To account for spatial heterogeneity and improve the estimation of genetic parameters, alternative approaches have been suggested. Papadakis (1937) and Wilkinson (1983) proposed the use of covariates to adjust plot results, while Cullis and Gleeson (1991) recommended modeling the correlation between plot errors using autoregressive-integrated-moving average models (ARIMA). This modeling approach was further extended to a two-dimensional scheme of rows and columns by using a first-order autoregressive (AR1 X AR1) residual structure (GILMOUR; CULLIS; VERBYLA, 1997). More recently, spline-based models such as spatial analysis of field trials with splines (SpATS) have been explored (RODRIGUEZ-ALVAREZ et al., 2018). However, the results from SpATS models tend to be similar to those obtained using AR1 X AR1 residual models (VELAZCO et al., 2017). Consequently, the AR1 X AR1 residual models have become common in plant breeding studies (FERREIRA COELHO et al., 2021; FERREIRA et al., 2022; HUNT et al., 2013; STRINGER; CULLIS, 2002).

Aside from spatial competition, genetic competition among individuals within a plot can significantly bias the assessment of individual performance and reduce the accuracy of genetic prediction (COSTA E SILVA; KERR, 2013; HUNT et al., 2013; STRINGER; CULLIS; THOMPSON, 2011). Recent developments in quantitative genetics emphasize the significance of interactions among individuals in shaping the inheritance and response to selection for quantitative traits (BIJMA, 2011b). Indirect genetic effects (IGEs) can arise from multiple genetic causes and vary with sex, age, and environment (BAILEY; HOSKINS, 2014; CAMERLINK et al., 2015; MOORAD; NUSSEY, 2016; WOLF; MUTIC; KOVER, 2011).

Griffing (1967, 1977) introduced models that incorporate indirect genetic effects (IGE) and investigated their implications for response to selection.

\citeonline{wright1986individual} derived expressions for the covariance between relatives in Griffing's model, considering additive, dominance, and epistasis effects for both DGE and IGE. In forest tree breeding, Cannell (1978) proposed selecting non-competitive genotypes to increase yield per unit area. This strategy is most effective when the DGE on growth are positively correlated with the IGE, resulting in less competition induced by high-growth individuals and allowing for increased tree density and enhanced yield per unit area. However, if the correlation between DGE and IGE is negative, selecting for higher growth can lead to increased competition among individuals and reduced overall production (CAPPA; CANTET, 2008).

Several studies have delved into incorporating IGE into genetic models. Muir et al (2002) described an individual model that accounts for both direct and competitive effects, while Van Vleck and Cassady (2005) utilized simulation to unravel the covariance structure of DGE and IGE. Similarly, Arango et al. (2005) attempted to estimate the covariance components for a swine population during a growth test, highlighting the importance of considering additive relationships among competing individuals. Muir (2005) developed mixed model equations considering competition effects to analyze a simulated forest tree data but did not account for the consequences of varying numbers of competitors, such as mortality or thinning, on estimating the additive genetic variance for competition.

The study by Resende et al. (2005) was the first to use spatial competition models in real data from forestry field trials. However, it did not consider the presence of missing neighbors resulting from mortality or account for the varying distances between a focal tree and its neighbors in different spatial directions. Cappa and Cantet (2008) discussed the estimation of DGE and IGE in tree breeding using a Bayesian approach with an individual tree mixed model that incorporates the number and position of competitor trees, revealing the biased estimation of breeding values when competition effects are neglected. Further advancements have been made by Costa e Silva and Kerr (2013), who extended the model proposed by Cappa and Cantet (2008) to account for competition effects in multiple directions (rows, columns, and diagonals) and investigated the impacts of genetic relatedness within the neighborhood and overall survival on the magnitude of IGE. Stringer et al. (2011) introduced the Random Treatment Interference Model (R-TIM), which jointly models spatial and genetic competition at the plot level, revealing a negative residual correlation between neighboring plots in different rows due to competition effects.

In tree breeding, various spatial competition mixed models, such as univariate single-environment, multivariate, and multi-environment competition mixed models, have been employed to investigate the relationships between DGE and IGE (BELABER et al., 2021; CAPPA et al., 2015; COSTA E SILVA et al., 2013, 2017; HERNANDEZ; LOPEZ; CAPPA, 2019). The Cappa and Cantet`s model was expanded by incorporating a two-dimensional smoothing surface accounting for the additive relationships among individuals in a Bayesian framework (CAPPA et al., 2015).

Overall, classical approaches in genetic selection often overlook the effects of genetic competition, leading to biased estimates of genetic and residual effects. Incorporating competition effects poses modeling challenges, including the interaction between direct and indirect genetic effects, and the potential biases introduced when ignoring or misinterpreting these effects in selection models. A comprehensive review on this subject for crop and animal breeding can be found in Chapter 22 of Walsh and Lynch(2018) and for a cross-disciplinary perspective, please see Baud et al. (2022).

1.3. References

ARANGO, J. et al. Estimation of variance components including competitive effects of Large White growing gilts. **Journal of animal science**, v. 83, n. 6, p. 1241–1246, 2005.

BAILEY, N. W.; HOSKINS, J. L. Detecting cryptic indirect genetic effects. **Evolution**, v. 68, n. 7, p. 1871–1882, 2014.

BAUD, A. et al. Indirect genetic effects: A cross-disciplinary perspective on empirical studies. **Journal of Heredity**, v. 113, n. 1, p. 1–15, 2022.

BELABER, E. C. et al. Accounting for competition in multi-environment tree genetic evaluations: a case study with hybrid pines. **Annals of Forest Science**, v. 78, n. 1, p. 1–15, mar. 2021.

BIJMA, P. A General Definition of the Heritable Variation That Determines the Potential of a Population to Respond to Selection. **Genetics**, v. 189, n. 4, p. 1347–1359, 1 dez. 2011a.

BIJMA, P. A general definition of the heritable variation that determines the potential of a population to respond to selection. **Genetics**, v. 189, n. 4, p. 1347–1359, 1 dez. 2011b.

BIJMA, P. The quantitative genetics of indirect genetic effects: a selective review of modelling issues. **Heredity**, v. 112, n. 1, p. 61–69, jan. 2014.

BURGUEÑO, J. Spatial analysis of field experiments. **Applied statistics in agricultural, biological, and environmental sciences**, p. 319–344, 2018.

CAMERLINK, I. et al. Indirect genetic effects for growth rate in domestic pigs alter aggressive and manipulative biting behaviour. **Behavior genetics**, v. 45, p. 117–126, 2015.

CANNELL, M. G. **Improving per hectare forest productivity**. Proceedings of the Fifth North American Forest Biology Workshop. **Anais**...1978.

CAPPA, E. P. et al. A novel individual-tree mixed model to account for competition and environmental heterogeneity: a Bayesian approach. **Tree genetics & genomes**, v. 11, p. 1–15, 2015.

CAPPA, E. P.; CANTET, R. J. C. Direct and competition additive effects in tree breeding: bayesian estimation from an individual tree mixed model. **Silvae Genetica**, v. 57, n. 1–6, p. 45–56, 1 dez. 2008.

COSTA E SILVA, J. et al. Genetic control of interactions among individuals: contrasting outcomes of indirect genetic effects arising from neighbour disease infection and competition in a forest tree. **New Phytologist**, v. 197, n. 2, p. 631–641, 2013.

COSTA E SILVA, J. et al. Genetic-based interactions among tree neighbors: identification of the most influential neighbors, and estimation of correlations among direct and indirect genetic effects for leaf disease and growth in Eucalyptus globulus. **Heredity**, v. 119, n. 3, p. 125–135, set. 2017.

COSTA E SILVA, J.; KERR, R. J. Accounting for competition in genetic analysis, with particular emphasis on forest genetic trials. **Tree Genetics & Genomes**, v. 9, n. 1, p. 1–17, 2013.

CULLIS, B.; GLEESON, A. Spatial analysis of field experiments-an extension to two dimensions. **Biometrics**, p. 1449–1460, 1991.

FAVERI, J. et al. Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, v. 66, n. 9, p. 947, 2015.

FERREIRA COELHO, I. et al. Accounting for spatial trends in multi-environment diallel analysis in maize breeding. **PloS one**, v. 16, n. 10, p. e0258473, 2021.

FERREIRA, F. M. et al. Bioenergy elephant grass genotype selection leveraged by spatial modeling of conventional and high-throughput phenotyping data. **Journal of Cleaner Production**, v. 363, p. 132286, 2022.

FERREIRA, F. M. et al. A data-driven approach for enhancing forest productivity by accounting for indirect genetic effects. **bioRxiv**, 2023.

FISHER, R. A.; OTHERS. Statistical methods for research workers. **Statistical methods for research workers.**, n. 1th Ed, 1925.

GILMOUR, A. R.; CULLIS, B. R.; VERBYLA, A. P. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 2, n. 3, p. 269–293, 1997.

GRIFFING, B. Selection in reference to biological groups I. Individual and group selection applied to populations of unordered groups. **Australian Journal of Biological Sciences**, v. 20, n. 1, p. 127–140, 1967.

GRIFFING, B. **Selection for populations of interacting genotypes**. Proceedings of the international conference on quantitative genetics. **Anais**...Iowa State University Press Ames, 1977.

HERNANDEZ, M. A.; LOPEZ, J. A.; CAPPA, E. P. Improving Genetic Analysis of Corymbia citriodora subsp. variegata with Single- and Multiple-Trait Spatial-Competition Models. **Forest Science**, v. 65, n. 5, p. 570–580, 2019.

HUNT, C. H. et al. Predicting additive and non-additive genetic effects from trials where traits are affected by interplot competition. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 18, p. 53–63, 2013.

MOORAD, J. A.; NUSSEY, D. H. Evolution of maternal effect senescence. **Proceedings of the National Academy of Sciences**, v. 113, n. 2, p. 362–367, 2016.

MUIR, W. M. Incorporation of competitive effects in forest tree or animal breeding programs. **Genetics**, v. 170, n. 3, p. 1247–1259, jul. 2005.

MUIR, W.; SCHINCKEL, A.; OTHERS. **Incorporation of competitive effects in breeding programs to improve productivity and animal well being**. 7th world congress on genetics applied to livestock production, montpellier, france. **Anais**...2002.

PAPADAKIS, J. Méthode statistique pour des expériences sur champ. **Bull. Inst. Amel. Plantes a Salonique**, v. 23, p. 13–29, 1937.

RESENDE, M. D. V. et al. Joint modelling of competition and spatial variability in forest field trials. **Revista de Matemática e Estatística, São Paulo,** v. 23, n. 2, p. 7–22, 2005.

RODRIGUEZ-ALVAREZ, M. X. et al. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. **Spatial Statistics**, v. 23, p. 52–71, 2018.

SELLE, M. L. et al. Flexible modelling of spatial variation in agricultural field trials with the R package INLA. **Theoretical and Applied Genetics**, v. 132, p. 3277–3293, 2019.

STRINGER, J.; CULLIS, B. Application of spatial analysis techniques to adjust for fertility trends and identify interplot competition in early stage sugarcane selection trials. **Australian Journal of Agricultural Research**, v. 53, n. 8, p. 911–918, 2002.

STRINGER, J.; CULLIS, B.; THOMPSON, R. Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement. **Journal of Agricultural Biological and Environmental Statistics**, v. 16, p. 269–281, jun. 2011.

VAN VLECK, L. D.; CASSADY, J. Unexpected estimates of variance components with a true model containing genetic competition effects. **Journal of Animal Science**, v. 83, n. 1, p. 68–74, 2005.

VELAZCO, J. G. et al. Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. **Theoretical and Applied Genetics**, v. 130, n. 7, p. 1375–1392, 2017.

WALSH, B.; LYNCH, M. **Evolution and selection of quantitative traits**. [s.l.] Oxford University Press, 2018.

WILKINSON, G. N. et al. Nearest Neighbour Analysis of Field Experiments. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 45, n. 2, p. 151–178, 1983.

WOLF, J. B.; MUTIC, J. J.; KOVER, P. X. Functional genetics of intraspecific ecological interactions in Arabidopsis thaliana. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 366, n. 1569, p. 1358–1367, 2011.

2. ARTICLE

A NOVEL STRATEGY TO PREDICT CLONAL COMPOSITES BY JOINTLY MODELING SPATIAL VARIATION AND GENETIC COMPETITION

Ferreira, F. M.; Chaves, S. F.; Bhering, L. L.; Alves, R. S.; Takahashi, E. K.; Souza, J. E.; Resende, M. D. V.; Leite, F. P.; Gezan, S. A.; Viana, J. M. S.; Fernandes, S. B.; Dias, K. O. G.

2.1. Abstract

Biotic and abiotic stresses have been making it difficult to maintain the rates of genetic gains in forest tree breeding. Planting a mixture of genotypes or clonal composites can be an alternative to increase the security of forest plantations. The formation of clonal composites is complex due to inter-genotypic competition and indirect genotypic effects that can affect the efficiency of genetic selection. This study aims to understand how indirect genotypic effects can impact the response to selection and how the stand composition can be used to explore these effects and enhance forest yield. We used two clonal trials of eucalyptus hybrids implanted in a randomized complete block design with 24 replications, containing a single tree per plot evaluated for mean annual increment at 3 and 6 years. We focused on partitioning the genotypic variation into direct and indirect genotypic effects based on competition intensity factors. We identified aggressive, homeostatic, and sensitive clones based on the magnitude of indirect genotypic effects. Considering indirect genotypic effects, for mean annual increment, the total heritability decreased from 0.25 to 0.10 for 3 years and from 0.30 to 0.14 for 6 years, respectively. We proposed a workflow that uses the direct and indirect genotypic effects to predict the mean values for clonal composite combinations and to select the one with the highest yield. Our methodology considered spatial variation and interplot competition that can affect the total heritable estimates and response to selection in forest trials. The clones were classified as aggressive, homeostatic, and sensitive, according to their deviation from the indirect genotypic effects means. Our methodology enabled the extraction of useful information and the prediction of many clonal composites' performance, which can be used in recommending genotypes to be planted in mixtures.

2.2. Introduction

In the last decades, there was a great increase in *Eucalyptus* spp. forest productivity due to genetic improvement. However, maintaining these rates is not a trivial task. Projections for the coming years suggest that changes in temperature and precipitation patterns will increase the vulnerability of forests to drought (CHOAT et al., 2012), and pests and diseases outbreaks (TRUMBORE; BRANDO; HARTMANN, 2015). These aggravating productivity-reducer factors are forcing changes in breeding strategies. Forest tree breeders are concerned with keeping populations with sufficient genetic variability, capable of sustaining the harmful influence of biotic and abiotic stresses (REZENDE et al., 2019). Forests with greater genetic variability are expected to be more resilient to climate change (MORIN et al., 2018).

Recent evidence has shown the potential of cultivating mixtures of improved clones, also called clonal composites (CC), to increase productivity (REZENDE et al., 2019) and to mitigate the effects of genotypes-by-environments interaction (OLIVEIRA et al., 2023). Comparing CC with monoclonal planting, Rezende (2019) found a 9.8 and 6.3 % increase in mean annual increment (MAI) in trials and commercial forests of eucalyptus, respectively. Similar results were found for a mixture of genotypes of *Pinus* (CARTER et al., 2020), where individual trees planted in mixed rows had approximately 7% greater volume compared to the ones that were planted in pure rows; and *Populus* (FOSTER; ROUSSEAU; NANCE, 1998) where the mixture presented 27% increase in volume over the best clone. The good performance of the mixture of genotypes can be due to better exploitation of environmental resources for the allo-competitor genotypes. Theoretically, there is a complementarity between different genotypes when allocated in the same area as suggested by the niche partitioning hypotheses (YOUNG, 1981). However, many questions remain unexplored in this topic, such as the number of genotypes that should be planted together, i.e., the size of the CC, and the impacts of indirect genotypic effects (IGE) on the performance of different CC combinations.

Linear mixed models (LMM) are routinely used to predict breeding values or genotypic values in tree breeding (CHAVES et al., 2022; GEZAN; DE CARVALHO; SHERRILL, 2017). The LMM was extended to include IGE as a random effect for animal or plant breeding. The IGE can be described as the influence that an individual's genes pool exerts over the phenotype of neighboring conspecifics, which can be studied by variance-component competition models (MUIR, 2005). Some evidence suggests that IGE, also called associative or competition effects, is a quantitative genetic trait (GRIFFING, 1967), thus it is capable of affecting the heritable variation of a population. More information can be found in Chapter 22 of Walsh and Lynch (2018) and in Chapter 11 of Resende et al. (2014). A spatial LMM accounting for IGE was applied in a few studies in forest genetics and tree breeding (COSTA E SILVA; KERR, 2013; HERNANDEZ; LOPEZ; CAPPA, 2019; RESENDE et al., 2005). The classical quantitative genetic model $P_{il} = G_i + E_{il}$, in which P_{il} is the phenotypic value of individual i into the environment l, G_i is the genotypic effects of individual i and E_{il} is the environmental effects for individual i into the environment l will be expanded in the competition philosophy. Now, the genotype's total genotypic value (TGV) comprises its own genetic merit or direct genotypic effects (DGE) summed to the weighted IGE that it exerts over its neighbors, therefore the genetic expression of individual i will contribute to the phenotype of its neighbors and its phenotypic value will have the neighbor`s contribution (BIJMA, 2011b).

Since IGE are expected to be common in nature (SAKAI, 1955), their impacts in the magnitude and direction of response to selection have been investigated in plant and animal breeding. Muir (2005) suggested that incorporating IGE into breeding programs can increase the accuracy of selection and improve breeding outcomes. The total heritable variation that determines the potential of a population to respond to selection depends on the DGE and IGE covariance (BIJMA, 2011a). Costa e Silva and Kerr (2013) simulated data to explore the impacts of different levels of genetic relatedness within the neighborhood and overall survival on the ability to estimate IGE. Silva et al. (2017) and Nunes et al. (2018) showed the implications of considering direct and indirect genotypic effects on forest management and breeding programs. The complete reviews of Bijma (2011, 2014) provide more examples and details.

Based on the aforementioned, we aimed to fit a spatial competition linear mixed model for estimating genetic parameters and studying the impacts of IGE. We also propose a strategy to predict the CC with the highest potential to form highly productive commercial stands. To the best of our knowledge, no previous study has explored the prediction of CC accounting for IGE. In fact, most of them restrict inferences only to traditional linear mixed models. The main advantage of our methodology is combining genetic competition and autoregressive residual models to predict the total genotypic value (TGV) of a CC and the competition class of the evaluated clones.

2.3. Material and methods

2.3.1. Trials description

We evaluated two trials of eucalyptus hybrids belonging to Celulose Nipo Brasileira S.A. (CENIBRA). These trials were side-by-side and were connected by three checks. The first trial was composed of 55 clonal hybrids generated from crosses between *E. urophylla* and *E. grandis*. The second trial was composed of 15 clonal hybrids generated from crosses between *E. urophylla* and *E. pellita* and *E. urophylla* and *E. globulus*. Both trials were implanted in a randomized complete block design with 24 replications and one tree per plot. Both trials had replications installed in two different areas. Replications from 1 to 12 were placed in an area distinct from replications from 13 to 24. The inter-rows and inter-columns spacing was 2.5 x 3 m, respectively, which gives a plantation density of 1333 trees ha^{-1}. The trials were in the municipality of Ipaba, Minas Gerais state, Brazil, at 19°41' S 42°41' W, and 213 m of altitude. This region has two well-defined seasons: one hot and humid (October to May) and the other cold and dry (April to September) with a minimum and maximum temperature of 13.5 and 34.6 °C, respectively, and monthly rainfall ranging from 0 to 112 mm.

The mean annual increment (MAI, m^{3}ha^{-1}year^{-1}) was calculated as the quotient of the volume of individual trees (VOL_{ind} - m^3) collected at 3 and 6 years after planting. This volume was extrapolated to ha (VOL_{ha} - m^3) as shown below:

$$VOL_{ind} = exp\big(V1 + V2 \times ln(DBH) + V3 \times ln(H)\big),$$

$$VOL_{ha} = VOL_{ind} \times (10000/AP) \times (SRV/100),$$

$$MAI = VOL_{ha}/A$$

where DBH is the diameter at breast height in cm, H is the total height in m, V1 = -10.300612, V2 = 1.689752, V3 = 1.292432, AP is the occupied area per plant, SRV is the percentage of survival, and A is the age (3 and 6) in years.

2.3.2. Statistical analyses

We fitted three linear mixed models. The traditional linear mixed model (TM) was used as a benchmark, and it was given by the following equation:

$$\boldsymbol{y} = \boldsymbol{X}\mathbf{b} + \boldsymbol{Z}_g\boldsymbol{g} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is a (n x 1) vector containing the phenotypes records, which n the number of records; $\boldsymbol{X}$ is the (n x r) incidence matrix identifying which of the r fixed effects are associated with each observation; $\mathbf{b}$ is the (r x 1) vector of fixed effects (intercept, 24 complete blocks, and 2 trials); $\boldsymbol{Z_g}$ is the (n x m) incidence matrices relating phenotypes records to their genotypic effects contained in the (m x 1) random vector $\boldsymbol{g}$ with $g \sim N(0,\ G\ =\ \mathrm{I}_{g\sigma_g^2})$ and $\boldsymbol{e}$ is the (n x 1) vector of random residuals with $e \sim N(0,\ R\ =\ \mathrm{I}_{e\sigma_{e^2}})$. **I** is an identity matrix of proper order, σ_g^2 is the genotypic variance and σ_e^2 is the residual variance.

The spatial model (SM), uses a separable first-order autoregressive process in two directions (AR1 x AR1) for modeling the covariance matrix of residual effects considering both rows and columns directions (GILMOUR; CULLIS; VERBYLA, 1997):

$$\mathbf{y}\ =\ \mathbf{X}\,\mathbf{b}\ +\ \mathbf{Z_g}\,\mathbf{g}\ +\boldsymbol{\xi}$$

where X is a vector that includes random effects that represent spatial trends revealed by the autoregressive correlated residual structure in the rows and columns, $\xi \sim N(0, R), with\, R = \sigma_\xi^2[AR1(\rho_c) \otimes AR1(\rho_r)] = H\sigma_\xi^2$. The separability property of the stochastic process ensures $H^{-1} = \sigma_\xi^2[AR1(\rho_c)^{-1} \otimes AR1(\rho_r)^{-1}]$ enabling the computation of H^{-1}, which is needed for estimation of variance components and for prediction of genotypic values. The quantity σ_ξ^2 is the variance of spatially dependent residuals, $AR1(\rho_r)\ and\, AR1(\rho_c)$ are the first-order autoregressive correlation matrices for rows and columns, respectively (GILMOUR; CULLIS; VERBYLA, 1997). First-order autoregressive correlation matrices were separately adjusted for each area. The operator ⊗ represents the Kronecker product between columns and rows auto-regressive processes.

We jointly modeled the spatial variation and genotypic competition in the plot level (STRINGER; CULLIS; THOMPSON, 2011). We used the proposal of Costa e Silva and Kerr (2013) derived from the original idea of Cappa and Cantet (2008). This methodology considers the mean competition intensity factors (f_{ij}) that a focal tree i exerts over its j neighbors $j = j_1, j_2,... j_m$, being m the maximum number of first neighbors for a focal tree i (Figure 1). The f_{ij} will vary according to the distance between these individuals in the field and the number of missing neighbors (Figure 2a). The magnitude of f_{ij} is related to the inverse of the distance between individual i and its m neighbors. The f_{ij} derivations were demonstrated by Cappa and Cantet (2008) for equal inter-row and inter-column spacing and Costa e Silva and Kerr (2013) for different inter-row and inter-column spacing. In this context, IGE can be regarded as the sum of effects that the m neighbors exert over the focal individual i $\left(\sum_{i\neq j}^{m} IGE_j\right)$ or the indirect genotypic effects on the phenotype of the focal tree (IGEF) (Figure 2b) or as the mean IGE that the focal individual i exert over its neighbors affecting their phenotype $(\overline{IGE_i})$ - (Figure 2c).

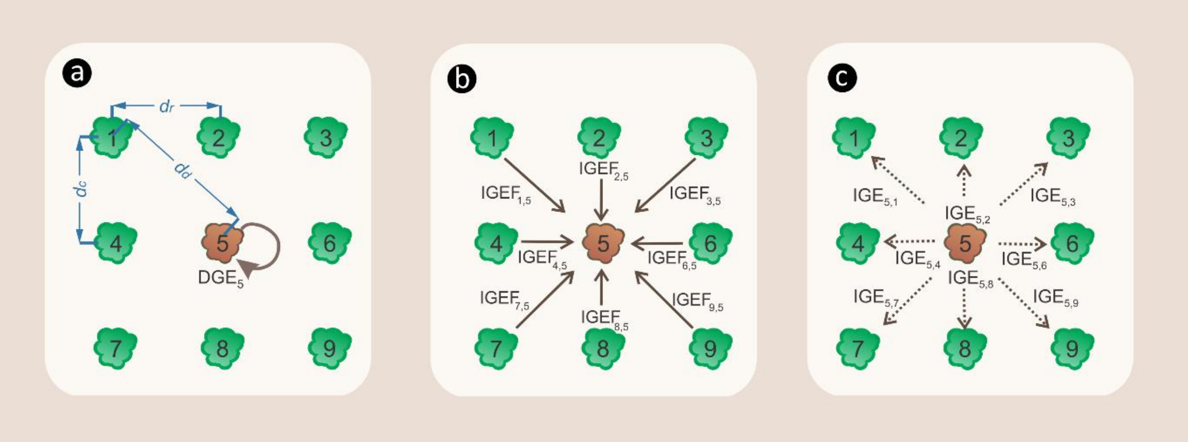

Figure 2: Illustration of (a) direct genotypic effects (DGE); (b) indirect genotypic effects on the phenotype of the focal tree (IGEF) of the j neighbors on the focal individual i, which will affect the phenotype of individual i; and (c) mean IGE of a focal individual i on its neighbors, affecting their phenotype. The focal individual i stands out by a different color and is represented by the number 5; its j neighbors are represented by the remaining numbers (the number of neighbors can vary from 0 to m depending on the position of focal individual i); d_r, d_c, and d_d are the distance between the focal individual and its neighbors in the row, column and diagonal, respectively.

In matrix notation, the SCM was given by the following equation (STRINGER; CULLIS; THOMPSON, 2011):

$$\mathbf{y} = \mathbf{X}\boldsymbol{b} + \mathbf{Z_g}\,\mathbf{g} + \mathbf{Z_c}\,\mathbf{c} + \boldsymbol{\xi}$$

where $\mathbf{Z_g}$ and $\mathbf{Z_c}$ are (n x t) incidence matrices linking phenotypes records to their DGE and IGE contained in the (t x 1) random vectors $\mathbf{g}$ and $\mathbf{c}$ respectively, being t the total number of treatments. Matrices $\mathbf{Z_c}$ were constructed following Costa Silva and Kerr (2013). When building the $\mathbf{Z_c}$ matrix, we considered the geographical separation of each area, i.e., clones in the 12th block do not neighbor clones in the 13th block. More details on how to construct these matrices can be found in the appendix.

The variance-covariance matrix of the DGE and IGE (Σ_g) given by:

$$\Sigma_g = \begin{bmatrix} \sigma_g^2 & \sigma_{gc} \\ \sigma_{gc} & \sigma_c^2 \end{bmatrix} \otimes I_n$$

where Σ_g denote the estimates of the covariance between DGE and IGE, and I_n is an identity matrix of order n.

We used the total genotypic value of each focal individual TGVi to rank and select individuals:

$$TGV_i = g_i + \left(\overline{n}_R\overline{f}_R + \overline{n}_C\overline{f}_C + \overline{n}_D\overline{f}_D\right) \times c_i$$

where g_i and c_i are the direct and indirect genotypic effects, respectively; $\overline{n}_R$, $\overline{n}_C$, and $\overline{n}_D$ are the average number of neighbors of individual i in row, column, and diagonal directions, respectively; and $\overline{f}_R$, $\overline{f}_C$, and $\overline{f}_D$ are the average fij for row, column, and diagonal directions, which are represented in Costa e Silva e Kerr (2013).

For clarity, the term $\left(\overline{n}_R\overline{f}_R + \overline{n}_C\overline{f}_C + \overline{n}_D\overline{f}_D\right)$ will be called competition intensity factor, or simply CIF.

Model selection and genetic parameters

We calculated the likelihood ratio test (LRT), considering an alpha = 0.05, by contrasting the likelihood of the full (full) and reduced model (red):

$$LRT = -2\left(LogL_{red} - LogL_{full}\right)$$

The goodness-of-fit of the evaluated models was assessed by the Akaike information criterion (AIC) – (AKAIKE, 1974).

$$AIC = -2LogL + 2p$$

where p is the number of estimated parameters. The most suitable model is the one that presents the lowest AIC.

Estimation of genetic and non-genetic parameters

We estimated the broad sense heritabilities (H^2) for the TM, SM, and SCM:

$$H^2_{TM} = \frac{\sigma^2_g}{\sigma^2_g+\sigma^2_e};\ H^2_{SM} = \frac{\sigma^2_g}{\sigma^2_g + \sigma^2_{\xi_1} + \sigma^2_{\xi_2}};\text{ and }H^2_{SCM} = \frac{\sigma^2_g}{\sigma^2_g+\sigma^2_c+\sigma^2_{\xi_1}+\sigma^2_{\xi_2}}$$

The broad sense total heritability (T^2) was estimated for the SCM, based on the total heritable variance (σ^2_t) as suggested by Bijma (2011) and Silva et al. (2013).

$$T^2 = \frac{\sigma^2_t}{\sigma^2_p}$$

where $\sigma_t^2 = \sigma_g^2 + 2 \times CIF \times \sigma_{gc} + CIF^2 \times \sigma_c^2$ and $\sigma_p^2 = \sigma_g^2 + \sigma_c^2 + \sigma_{\xi_1}^2 + \sigma_{\xi_2}^2$.

We also estimated the accuracy of both DGE (r_g) and IGE (r_c) following (MRODE, 2014).

$$r_g = \sqrt{1 - \frac{PEV}{\sigma_g^2}}; r_c = \sqrt{1 - \frac{PEV}{\sigma_c^2}}$$

where PEV is the prediction error variance.

The individual reliability of DGE and IGE is the square root of r_g and r_c, respectively. For comparison, we estimated these parameters by the traditional models and in the spatial-competition model. We compared the ranking of genotypes provided by each model using the Spearman correlation. We also predicted the response to selection (R), in percentage, for each model, using the genotypic values for TM and SM and TGV for SCM, as given by Piephoi and Mohring (2007):

$$R = \frac{\bar{g}}{\bar{y}} * 100$$

where $\bar{g}$ is the mean best linear unbiased prediction (BLUP) of the selected clones, for TM and SM, or the mean TGV of the selected clones, for SCM, and $\bar{y}$ is the phenotypic mean for the evaluated trait. We selected 26% of the clones (20 out of 76).

All analyses were performed within the R software environment, version 4.2.1 (R CORE TEAM, 2022). The linear mixed models were fitted using ASReml-R (version 4.1) - (BUTLER et al., 2018). Graphs and figures were made using the ggplot2 package.

2.3.3. Prediction of clonal composites

Leveraging the methods mentioned above, we propose the following workflow: i) fit a CSM in real datasets; ii) estimate genetic parameters and predict the DGE, IGE, and TGV for each clone and select the best-ranked ones based on their TGV; iii) investigate CC with different compositions and sizes by calculating their mean values, considering the DGE and IGE of all possible combinations of the selected clones for a given CC size. Here, we defined that the CC should be composed of 5 clones.

Furthermore, we randomized the clones in the CC 2 times and take a mean of the combinations across each randomization) (at this time we define the size of the CC as 5); iv) choose the CC composition that enhances the forest productivity based on MAI and randomize the clones 100 times to calculate an expected mean and its standard deviation for the evaluated trait (Figure 3).

Based on the TGV, for mean annual increment (IMA), the 20 best-ranked clones were selected to predict all possible combinations of clones and find the CC with the highest potential to form highly productive commercial forests. Since IGE was statistically significant, we followed our proposed workflow (Figure 3). The data from the selected clones were used to create a list with all possible CC combinations. Therefore, based on the 20 selected clones, we predicted 15,504 CC combinations considering the DGE and IGE.

Figure 3: Workflow for clonal composites (CC) prediction. y is a (n x 1) vector that contains the phenotype records, with n being the number of records; X is the (n x r) incidence matrix identifying which of the r fixed effects are associated with each observation; $\boldsymbol{\beta}$ is the r x 1 vector of fixed effects, $\boldsymbol{Z_d}$ and $\boldsymbol{Z_c}$ are (n x m) incidence matrices linking phenotype records to their DGE and IGE contained in the (m x 1) random vectors $\boldsymbol{g}$ and $\boldsymbol{c}$ respectively according to Costa e Silva and Kerr (2013). ξ is an effect (random) that represent spatially trends or correlated residual in the row-column structure. TGV is the total genotypic value, DGE is the direct genotypic effects, IGE is the indirect genotypic effects, BLUP is the best linear unbiased prediction, SM is the spatial model, and SCM is the spatial competition model.

The best CC combination was the one with the highest predicted average $\widehat{y_i}$ for our trait of interest:

$$\widehat{y_i} = \hat{\mu} + \widehat{DGE_i} + \sum_{i \neq j}^{m} \frac{1}{di} \times \widehat{IGE_j}$$

Note that the predicted mean is weighted by the inverse of the distance between rows, columns, or diagonals, which penalizes the effect of competition based on the distance between neighbor j and the focal individual i for the m closest neighbors. $\hat{\mu}$ is the intercept of the model, which can be predetermined based on the mean of the MAI trait.

We propose an innovative classification for competition classes based on the magnitude of the predicted IGE and the impacts that it can have on the performance of some clones. Clones whose predicted IGE falls within one standard deviation of the mean predicted IGE were considered homeostatic, whereas clones below the threshold of one standard deviation were considered aggressive. Finally, clones above the threshold of one standard deviation were considered sensitive.

2.4. Results

2.4.1 Model selection and genetic parameters

The percentages of missing observations in the trial were 13,87% and 21.93% at 3 and 6 years after planting, respectively. According to the LRT, the IGE was statistically significant ($p < 0.05$) for both ages. Thus, IGE should be considered when predicting the genotypic values. Based on the AIC, the best-fitted model, for both ages, was the SCM (Table 1).

Table 1: Model performance considering model complexity via Akaike information criterion (AIC) for a traditional model (TM), a spatial model (SM), and a spatial-competition model (SCM) accounting for the correlation between direct genotypic effects (DGE) and direct genotypic effects (IGE) for eucalyptus at 3 and 6 years.

Age		3	6
Model	DF	AIC	
TM	2	9413.14	9588.06
SM	7	9350.77	9568.31
SCM	9	9325.61	9533.97

where DF is degrees of freedom, and AIC is Akaike information criterion.

The distribution of IGE effects was used to classify clones as aggressive, homeostatic, and sensitive for both ages (Figures 4 a and b). The number of different clones as neighbors that each clone had throughout the experiment (considering all replications) varied from 29 to 73, for 3 years, and from 23 to 69, for 6 years. All clones had neighbors in the 3 classes (Figures 4 c and d). Following our proposed classification method, the checks were aggressive and homeostatic at 3 years and as homeostatic and sensitive at 6 years.

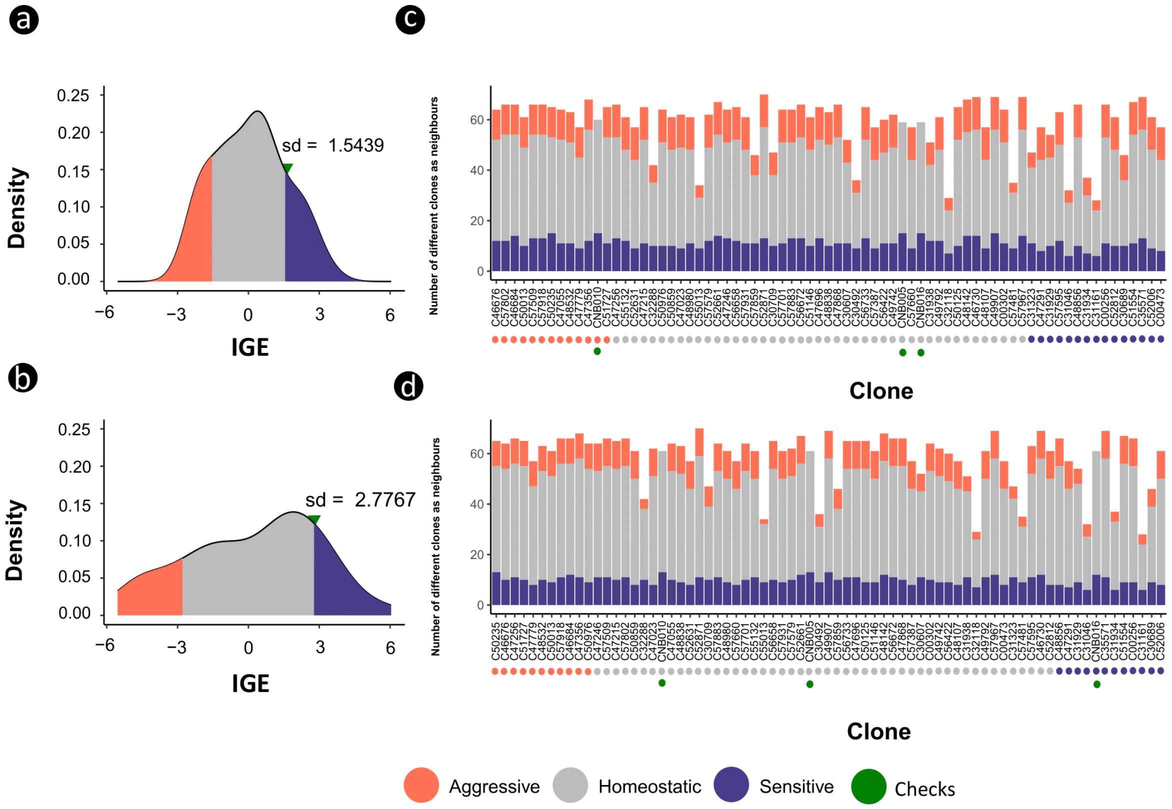

Figure 4: Density plots for IGE for eucalyptus clones belonging to three competition classes based on the magnitude of the IGE at 3 and 6 years, respectively; Homeostatic clones are in the area between plus and minus one standard deviation from the mean IGE, aggressive clones are in the area below one standard deviation from the mean IGE (shaded orange), and sensitive clones are in the area above one standard deviation from the mean (shaded blue); The upside down green triangle indicates the point where values are higher than one standard deviation from the mean IGE; c and d) Competition class of each clone (indicated by the colored circles below the clone names) and the total number of different neighbors and their classes (indicated by the colored bars) at 3 and 6 years, respectively. The green circle indicates the checks.

There were increments in the heritabilities as the age advanced (Tables 2 and 3). The total heritable variation reduced from 0.25 to 0.10 for 3 years and from 0.30 to 0.14 for 6 years when the SCM was employed instead of the TM. Resende (2005) indicated that high (>0.3) positive autocorrelation coefficient estimates reveal that environmental heterogeneity (spatial effects) is predominant over the competition, and negative (< -0.3) autocorrelation coefficient estimates indicate competition effects at the residual level probably together with environmental heterogeneity. Therefore, autocorrelation coefficients smaller than -0.3 were used to identify the predominance of competition effect. Our results showed a high level of competition across rows in both groups (Tables 2 and 3), which can be due to shading in the row direction of the field trial. All models, for both ages, presented very high accuracy (> 0.95), indicating that they can be used to perform selection. The three most aggressive clones, meaning that they exerted the highest competition over their neighbors, were clones C46676, C57802, and C46684 at 3 years and clones C50235, C46676, and C47256 at 6 years (Supplementary material - Tables S1 and S2). These clones were among the ones that had the highest DGE (Figure 5).

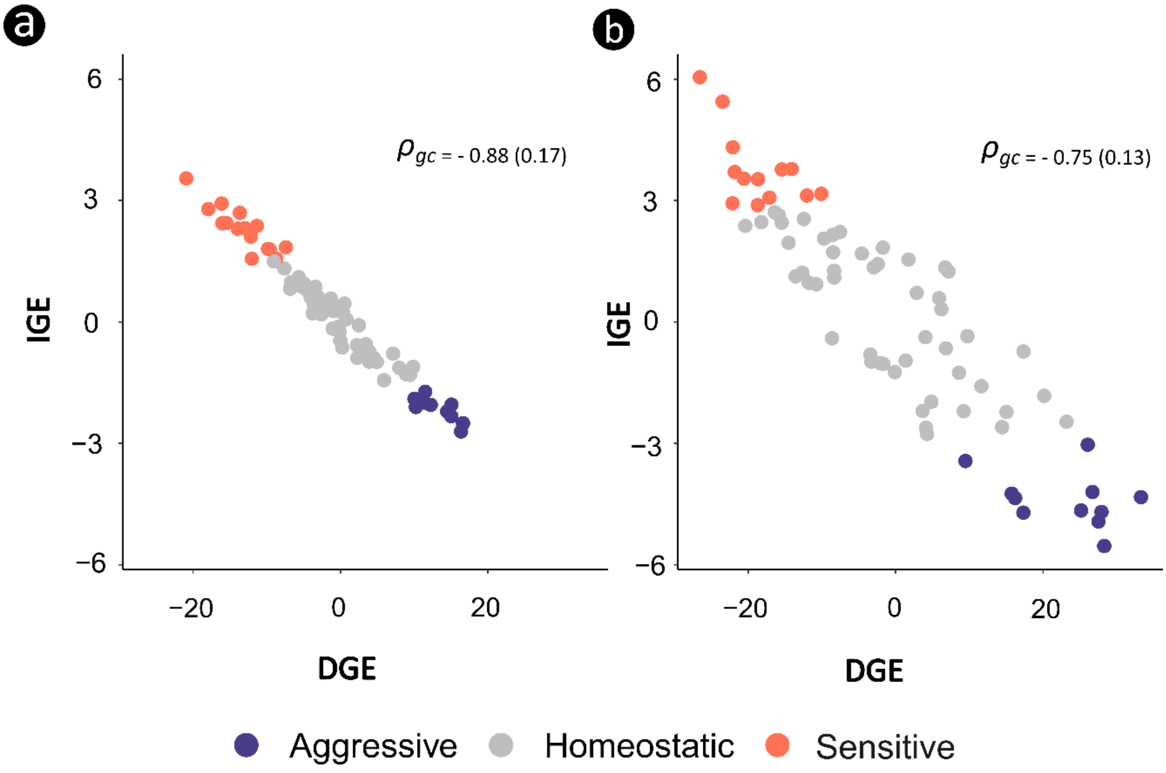

Figure 5: Correlation between direct genotypic effects and indirect genotypic effects (ρ_{gc}) for a eucalyptus clonal trial at ages 3 years (a) and 6 years (b), and their respective competition class. DGE and IGE are direct and indirect genotypic effects, respectively.

Table 2: Estimates (E) of genetic and non-genetic parameters and standard errors (SE) for eucalyptus clonal trial at 3 years old by using a traditional model (TM), a spatial model (SM), and a spatial-competition model (SCM).

	E_{TM}	SE_{TM}	E_{SM}	SE_{SM}	E_{SCM}	SE_{SCM}
ρ_{dc}					-0.88	0.17
σ_g^2	98.92	17.97	89.53	16.31	89.29	16.24
σ_c^2					3.15	1.48
σ_e^2	140.41	5.22				
$\sigma_{\xi_1}^2$			181.99	9.58	178.95	9.51
ρ_{c_1}			0.08	0.04	0.07	0.04
ρ_{r_1}			-0.11	0.04	-0.1	0.04
$\sigma_{\xi_2}^2$			95.86	5.47	92.02	5.37
ρ_{c_2}			-0.08	0.05	-0.1	0.05
ρ_{r_2}			0.03	0.05	0.02	0.05
σ_p^2	239.33		367.38		363.42	
σ_t^2					35.75	
H_d^2	0.41	0.05	0.24	0.03	0.25	0.03
T^2					0.1	0.05
r_d	0.96		0.96		0.96	
r_c					0.86	

ρ_{dc} is the correlation between DGE and IGE; σ_g^2 is the DGE variance; σ_c^2 is the IGE variance; σ_e^2 is the residual variance; $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$ are the residual variances of correlated residuals in the row-column structure for area 1 and area 2; ρ_{c_1} and ρ_{c_2} are the correlations across columns area 1 and area 2, respectively; ρ_{r_1} and ρ_{r_2} are the correlations across rows area 1 and area 2, respectively; σ_p^2 is the phenotypic variance; σ_t^2 is the total heritable variance; H_d^2 is the broad sense heritability; T^2 is the broad sense total heritability; r_d accuracy for DGE; and r_c accuracy for IGE.

Table 3: Estimates (E) of genetic and non-genetic parameters and standard errors (SE) for eucalyptus clonal trial at 6 years old by using a traditional model (TM), a spatial model (SM), and a spatial-competition model (SCM).

	E_{TM}	SE_{TM}	E_{SM}	SE_{SM}	E_{SCM}	SE_{SCM}
ρ_{dc}					-0.75	0.13
σ_g^2	260.62	47.46	243.6	44.41	246.85	44.81
σ_c^2					11.56	4.08
σ_e^2	300.7	11.79				
$\sigma_{\xi_1}^2$			345.4	19.23	333.98	18.91
ρ_{c_1}			0.01	0.04	0	0.04
ρ_{r_1}			-0.16	0.04	-0.16	0.04
$\sigma_{\xi_2}^2$			249.33	15.1	232.14	14.43
ρ_{c_2}			0	0.05	-0.04	0.06
ρ_{r_2}			0.01	0.05	-0.03	0.05
σ_p^2	561.32		838.34		824.53	
σ_t^2					119.54	
H_d^2	0.46	0.05	0.29	0.04	0.3	0.04
T^2					0.14	0.06
r_d	0.96		0.96		0.96	
r_c					0.81	

ρ_{dc} is the correlation between DGE and IGE; σ_g^2 is the DGE variance; σ_c^2 is the IGE variance; σ_e^2 is the residual variance; $\sigma_{\xi_1}^2$ and $\sigma_{\xi_2}^2$ are the residual variances of correlated residuals in the row-column structure for area 1 and area 2; ρ_{c_1} and ρ_{c_2} are the correlations across columns area 1 and area 2, respectively; ρ_{r_1} and ρ_{r_2} are the correlations across rows area 1 and area 2, respectively; σ_p^2 is the phenotypic variance; σ_t^2 is the total heritable variance; H_d^2 is the broad sense heritability; T^2 is the broad sense total heritability; r_d accuracy for DGE; and r_c accuracy for IGE.

The potential of the early selection (correlation between ages 3 and 6) was assessed by the Spearman rank correlation. Lower correlations between genotypes selected at 3 and 6 years for SCM, using the TGV, when compared to the TM and SM were found (Table 4). Also, there were changes in the 20 best-ranked genotypes between 3 and 6 years (Supplementary material - Tables S1 and S2). The proportion of the selected genotypes (PSG) at 6 years that were among the previously selected genotypes at 3 years was higher when the competition models were adopted in cooperation with the TM and equal to the SM (Table 4). For 3 years old, the responses to selection by selecting the 20 best-ranked genotypes were 38.76, 35.95, and 33.74% for TM, SM, and SCM, respectively. For 6 years old, the responses to selection by selecting the 20 best-ranked genotypes were 54.44, 51.98, and 47.94% for TM, SM, and SCM, respectively. In the supplementary material, we present the DGE, IGE, their respective standard errors, and TGV for all evaluated clones at 3 years (Table S1) and 6 years (Table S2).

Table 4: Spearman rank correlation (SRC) between 3 and 6 years in eucalyptus clones for all evaluated genotypes (n = 73) and proportion of the selected Genotypes (PSG) at 6 years old that were among the 10 genotypes previously selected at 3 years old (n = 10).

	TM	SM	SCM_{TGV}	SCM_{DGE}	SCM_{IGE}
SRC	0.92	0.92	0.86	0.92	0.93
PSG	0.8	0.9	0.9	0.8	0.7

TM: traditional linear mixed model; SM: spatial model, SCM: spatial-competition model (SCM)

2.4.2 Clonal composite predictions

To define which ones will be the five selected clones, we used the TGV at 6 years of the 20 best-ranked clones. We predicted 15,504 CC combinations and chose the one that predicted the greatest planted forest production per area considering the competition among clones. The chosen CC composition was the one that had clones ranked in the positions 1, 4, 5, 2, and 3 for TGV; 1, 3, 5, 6, and 8 for DGE; and 7, 4, 9, 11 and 15 for IGE (sort smallest to largest). The chosen CC composition expected MAI was 56.26 m^3 ha^{-1} year $^{-1}$ (Table 5). Despite C50976 and C57802 not being ranked among the top five clones for DGE (Table 5), they were selected prioritizing the trade-off between DGE and IGE. The IGE that C57802 (homeostatic clone) exerts over its neighbors is almost half the IGE that the C50013 (aggressive clone) exerts over its neighbors (Table 5). Based on these results, we can infer that at least 1 out of 5 clones must be homeostatic to compose a CC of size 5. The expected MAI for a monoculture with the best-ranked clone was 61.98 m^3 ha^{-1} year $^{-1}$, which was higher than the best CC (56.26 m^3 ha^{-1} year $^{-1}$). The distribution of the expected mean based on the 15,504 CC combinations of size five is displayed in the Supplementary material (Figure S1).

Table 5: Expected mean $(\hat{Y})$ for mean annual increment at age 6 years for the best clonal composite composition of size 5 for eucalyptus hybrids.

CLONE	$\widehat{DGE}$	$\widehat{IGE}$	$Rank_{TGV}$	$Rank_{DGE}$	$Rank_{IGE}$
C50013	33.3	-4.3	1	1	7
C51727	27.9	-4.7	4	3	4
C46684	26.7	-4.2	5	5	9
C50976	26.1	-3.1	2	6	11
C57802	23.2	-2.5	3	8	15
$\hat{Y}$	56.19 (0.0573)				

where $\widehat{DGE}$ is the predicted direct genotypic effects, $\widehat{IGE}$ is the predicted indirect genotypic effects, $\hat{Y}$ mean annual increment at age 6 years for the best clonal composite composition and its standard error (in parentheses). $Rank_{TGV}$, $Rank_{DGE}$, and $Rank_{IGE}$ are ranking for each of the select genotypes based on the TGV, DGE, and IGE (IGE is sort smallest to largest), respectively.

2.5. Discussion

Our results shed new light on the potential of spatial-competition models to predict the performance of a mixture of genotypes, including CC, as a combination of direct and indirect genetic effects. This section summarizes the findings and contributions made by the present study focusing on our two main goals: fit a spatial-competition model to estimate genotypic parameters and investigate the effects of IGE and propose a new approach to predict CC eucalyptus yield.

2.5.1 Model selection and genetic parameters estimates

We choose to work with MAI (m^3 ha^{-1} year $^{-1}$) because it is directly related to diameter at breast height and plant height. Additionally, MAI is a function of the planted area and the age of the forest. The mean annual increment increases with the age of the stand until competition and physiological maturity slow the growth rate of the cohabitating trees. The chosen trial, established with single-tree plots, was efficient for estimating the IGE since the focal clone is often surrounded by genetically distinct others, favoring the inter-clonal competition experienced in CC plantations.

The plastic response of a given clone under competition is closely related to its ability to capture limited resources and use them efficiently in this shared environment. After canopy closure, the competition for light, water, and nutrients intensifies. Therefore, competition can be considered a force that disproportionately affects cohabiting individuals' growth. Clones present different magnitudes of IGE as age advances, indicating that a differential gene expression - and, consequently, a differential gene interaction between the focal tree and its neighbors - can occur. The

8.06\% increase in mortality observed between 3 and 6 years might be due to competition or ecophysiological effects not considered by our model.

There are many ways to construct the Z_c matrix and calculate the CIF (BIJMA, 2011b; CAPPA; CANTET, 2008; COSTA E SILVA; KERR, 2013). However, the results must make biological sense. The applied methodology provided IGE values much lower than the DGE, which is expected since the clone's performance is mainly affected by the pure expression of their genes due to DGE. Also, we worked with many genetically distinct interacting individuals, which can reduce the model's ability to capture IGE (BIJMA, 2011b). The SCM was efficient in capturing the IGE, allowing the estimation of variance components and predicting TGV on the individual level. The proposed competition classification is a refined way of label clones using a common statistical concept.

All checks were classified as homeostatic or sensitive at 6 years. This suggests that the effect of intragenotypic competition (at more advanced stages) can reduce the performance of competitive genotypes, resulting in their non-selection. The check CNB010 was aggressive at 3 years and homeostatic at 6 years, which may be due to accelerated initial growth that stabilized in later ages. The notion of genetic architecture and the response to selection can be altered by accounting for IGE. IGE can amplify or diminish the response to artificial selection compared to what is estimated via models that do not account for this effect (COSTA E SILVA et al., 2013).

The range of heritability values depends on the correlation between DGE and IGE (ρ_{gc}), which is a biological phenomenon that arises due to the genetic interactions among individuals within a neighborhood. The ρ_{gc} quantifies the extent to which a genotype's genetic makeup influences both its own performance and the performance of neighboring individuals through competition. Positive ρ_{gc} indicates that genotypes with higher individual performance (DGE) tend to positively impact the performance of their neighbors (IGE). Conversely, negative ρ_{gc} suggests that genotypes with higher DGE may have negative effects on their neighbor's performance (COSTA E SILVA et al., 2017). The magnitude and direction of the ρ_{gc} are influenced by factors such as genetic relatedness, plant density, resource availability, and spatial arrangement (BIJMA, 2014).

We observed a strong and negative ρ_{gc}, which impacts the magnitude of the genetic estimates, decreasing the potential response to selection for MAI based on individual performance. Bijma (2011) speculates that this might be due to evolutionary mechanisms that selected related genotypes in natural populations with good growth capacity. As a result, by accounting IGE, there is less heritable variation to be explored. Thus, the predictions in response to selection may be biased if IGE is neglected when it is statistically significant. The opposite can occur when the ρ_{gc} is positive, as evidenced by Costa e Silva et al. (2013) for resistance to *Mycosphaerella* leaf disease. Future studies can help to elucidate the physiological aspects of competition, including how stress signals are perceived and how they trigger genetic and metabolic responses.

2.5.2 Clonal composite predictions

The optimal genotype may not always be the most competitive or productive for scenarios with inter-genotypic competition. In cases where two trees possess identical DGE, opting for the one with the smaller IGE will mitigate adverse effects on the growth of neighboring trees. The TGV is a valuable selection metric as it considers the trade-off between DGE and IGE. Despite the small changes in our study, considering IGE can alter the order of the selected genotypes.

Some authors reported the productive superiority of a mixture of genotypes in relation to monoclonal plantations (CARTER et al., 2020; FOSTER; ROUSSEAU; NANCE, 1998; REZENDE et al., 2019). However, following our methodology, the best-ranked clone at 6 years remained the most productive when randomized with itself (monoclonal) compared to the best CC. The way of our trial was designed (one tree per plot) did not allow the estimation of the intragenotypic competition. Therefore, more research is necessary to explore the advantages of SCM for eucalyptus, particularly using larger datasets from different breeding stages.

Competition models can provide valuable information to predict CC composition, since changes in response to selection and ranking may occur by considering the IGE. Our study was the first to present, methods that use a SCM for predicting CC performance. The magnitude and direction of IGE may vary depending on the size and composition of the CC, the availability of resources, and the number of neighbors. Future studies should aim to validate results in field trials and simulations.

The ideal size of the CC will be based on the researcher's previous experiences, which must consider: the maintenance of a reasonable level of diversity, the availability of seedlings or seeds, and the trade-off between DGE and IGE. The use of homeostatic clones associated with the best-ranked TGV provided a reasonable performance compared to the best CC, suggesting that it can be explored in situations of denser spacing or longer cutting cycles (where a greater magnitude of IGE is expected) without sudden decreases in productivity. The best-ranked genotypes based on the DGE, will not always be the most appropriate to compose a CC due to the trade-off between DGE and IGE. In this study, the proportion of each clone into the CC is not yet optimized as it considers the same contribution proportion (20%) for each of the five clones. More elaborated solutions can be explored by optimization CC via linear programming techniques.

Our strategy should be explored to predict the performance of CC that have been untested in the field. Testing all possible combinations of clones to define a compound is costly, time-consuming, and, in some situations, impractical. We successfully fit parsimonious spatial-competition models, extracting useful information to predict many CC combinations and their expected average performance. We identified genotypes with distinct competition response patterns using the IGE magnitude, which can be a piece of valuable information for assisting with strategic decision-making within forest tree breeding programs or operational plantations. For example, we can predict the expected changes in a trait performance due to modifications in the number of clones composing the CC or their phenotypic plasticity. Clearly, the SCM provided more information to assist the selection process than the TM and SM models. Thus, the SCM can be expanded to different situations of perennial breeding, helping to answer important and under-explored questions regarding the management and mixture of genotypes.

2.6. Conclusion

We showed based on a SCM that overestimated heritability is obtained when IGE are not considered for MAI in eucalyptus. Also, a methodology for predicting high-performance CC based on the trade-off between DGE and IGE was proposed.

Therefore, predicting CC by capitalizing on the IGE can generate a strategic advantage in determining the CC with the highest productive potential.

2.7. Author contribution statement

Ferreira, F.M.; and Dias, K.O.G., designed the research. Ferreira, F.M.; Chaves, S.F.S.; and Dias, K.O.G., performed the statistical analyses and wrote the first draft; Dias, K.O.G., programmed the function to count the neighbors and fit the model; Chaves, S.F.S., programmed the function to simulate the clonal composites; Ferreira, F.M., programmed the general script; Bhering, L.L; Alves, R.S; Resende, M.D.V; Gezan, S.A; Viana, M.S; and Fernandes S.B. revised drafts of the paper. Alves, R.S; Takahashi, E.K; Sousa J.E.; and Leite F.P. provided the eucalyptus dataset. All of the authors read and approved the final manuscript.

2.8. Acknowledgements

This work was financially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001. The authors also acknowledge the Federal University of Viçosa, Society of Forest Investigations, and Celulose Nipo Brasileira S.A. for the partnership.

2.9. References

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716–723, dez. 1974.

BIJMA, P. A general definition of the heritable variation that determines the potential of a population to respond to selection. **Genetics**, v. 189, n. 4, p. 1347–1359, 2011a.

BIJMA, P. A General Definition of the Heritable Variation That Determines the Potential of a Population to Respond to Selection. **Genetics**, v. 189, n. 4, p. 1347–1359, 1 dez. 2011b.

BIJMA, P. The quantitative genetics of indirect genetic effects: a selective review of modelling issues. **Heredity**, v. 112, n. 1, p. 61–69, jan. 2014.

BUTLER, D. G. et al. **ASReml-R reference manual version 4**. [s.l.] VSN International, 2018.

CAPPA, E. P.; CANTET, R. J. C. Direct and competition additive effects in tree breeding: bayesian estimation from an individual tree mixed model. **Silvae Genetica**, v. 57, n. 1–6, p. 45–56, 1 dez. 2008.

CARTER, D. R. et al. Complementarity increases production in genetic mixture of loblolly pine (*Pinus taeda* L.) throughout planted range. **Ecosphere (Washington, D.C)**, v. 11, n. 11, p. 1–15, 2020.

CHAVES, S. F. S. et al. Application of linear mixed models for multiple harvest/site trial analyses in perennial plant breeding. **Tree Genetics & Genomes**, v. 18, n. 6, p. 44–55, dez. 2022.

CHOAT, B. et al. Global convergence in the vulnerability of forests to drought. **Nature**, v. 491, n. 7426, p. 752–755, 2012.

COSTA E SILVA, J. et al. Genetic control of interactions among individuals: contrasting outcomes of indirect genetic effects arising from neighbour disease infection and competition in a forest tree. **New Phytologist**, v. 197, n. 2, p. 631–641, 2013.

COSTA E SILVA, J. et al. Genetic-based interactions among tree neighbors: identification of the most influential neighbors, and estimation of correlations among direct and indirect genetic effects for leaf disease and growth in Eucalyptus globulus. **Heredity**, v. 119, n. 3, p. 125–135, set. 2017.

COSTA E SILVA, J.; KERR, R. J. Accounting for competition in genetic analysis, with particular emphasis on forest genetic trials. **Tree Genetics & Genomes**, v. 9, n. 1, p. 1–17, 2013.

FOSTER, G. S.; ROUSSEAU, R.; NANCE, W. Eastern cottonwood clonal mixing study: intergenotypic competition effects. **Forest Ecology and Management**, v. 112, n. 1–2, p. 9–22, 1998.

GEZAN, S. A.; DE CARVALHO, M. P.; SHERRILL, J. Statistical methods to explore genotype-by-environment interaction for loblolly pine clonal trials. **Tree Genetics & Genomes**, v. 13, p. 1–11, 2017.

GILMOUR, A. R.; CULLIS, B. R.; VERBYLA, A. P. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 2, n. 3, p. 269–293, 1997.

GRIFFING, B. Selection in reference to biological groups I. Individual and group selection applied to populations of unordered groups. **Australian Journal of Biological Sciences**, v. 20, n. 1, p. 127–140, 1967.

HERNANDEZ, M. A.; LOPEZ, J. A.; CAPPA, E. P. Improving Genetic Analysis of Corymbia citriodora subsp. variegata with Single- and Multiple-Trait Spatial-Competition Models. **Forest Science**, v. 65, n. 5, p. 570–580, 2019.

MORIN, X. et al. Long-term response of forest productivity to climate change is mostly driven by change in tree species composition. **Scientific Reports**, v. 8, n. 1, p. 1–12, 2018.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. 3rd. ed. [s.l.] CABI, 2014.

MUIR, W. M. Incorporation of competitive effects in forest tree or animal breeding programs. **Genetics**, v. 170, n. 3, p. 1247–1259, jul. 2005.

NUNES, A. C. P. et al. Statistical strategies design based on competition classes of Eucalyptus clones. **Industrial Crops and Products**, v. 124, p. 66–73, nov. 2018.

OLIVEIRA, A. F. C. F. et al. Clonal composites as a strategy for mitigating the clones × environments interaction in *Eucalyptus*. **Cerne**, v. 29, p. 1–7, 2023.

PIEPHO, H. P.; MOHRING, J. Computing heritability and selection response from unbalanced plant breeding trials. **Genetics**, v. 177, p. 1881–1888, 2007.

R CORE TEAM. **R: A language and environment for statistical computing**. [s.l.] R Foundation for Statistical Computing, 2022.

RESENDE, M. D. V. et al. Joint modelling of competition and spatial variability in forest field trials. **Revista de Matemática e Estatística, São Paulo,** v. 23, n. 2, p. 7–22, 2005.

RESENDE, M.; SILVA, F.; AZEVEDO, C. **Estatística matemática, biométrica e computacional**. [s.l.] Suprema, 2014.

REZENDE, G. D. et al. Clonal Composites: An alternative to improve the sustainability of production in eucalypt forests. **Forest Ecology and Management**, v. 449, p. 1–9, 2019.

SAKAI, K. I. Competition in plants and its relation to selection. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 20, p. 137–157, 1955.

STRINGER, J.; CULLIS, B.; THOMPSON, R. Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement. **Journal of Agricultural Biological and Environmental Statistics**, v. 16, p. 269–281, jun. 2011.

TRUMBORE, S.; BRANDO, P.; HARTMANN, H. Forest health and global change. **Science (New York, N.Y.)**, v. 349, n. 6250, p. 814–818, 2015.

YOUNG, J. P. W. Sib competition can favour sex in two ways. **Journal of Theoretical Biology**, v. 88, n. 4, p. 755–756, 1981.

2.10. Appendix

Below we have a small example of how to construct the incidence matrices X, Z_g, and $\boldsymbol{Z_c}$ for a complete randomized block design with 6 treatments and two blocks and one plant per plot (Figure 4). The first block is represented on the left and the second block on the right. The plots from 1 to 6 are from block 1 and the plots from 7 to 12 are from block 2.

The hypothetical distance between the focal tree and its neighbors in the row, column, and diagonal are 2.5, 3, and 3.90 m, respectively.

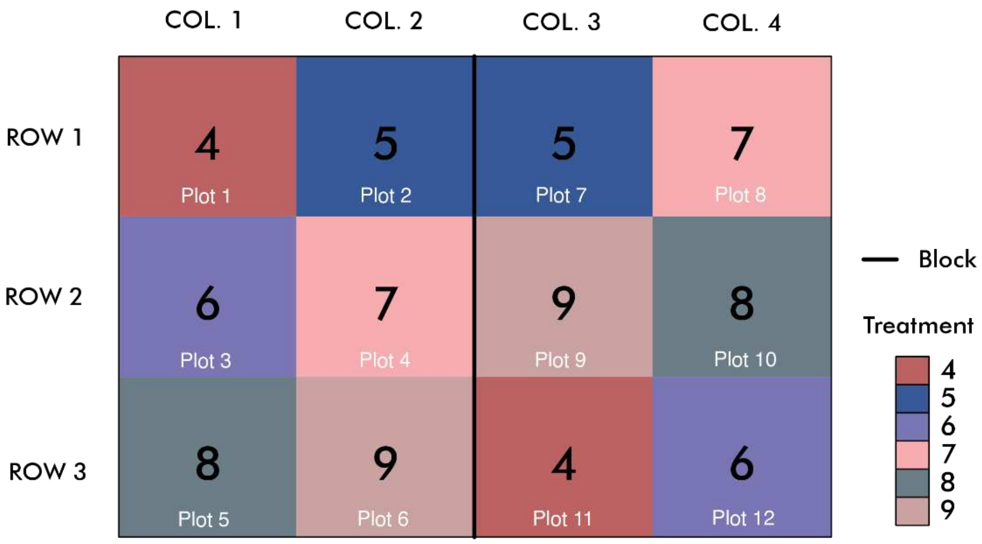

Figure 6: Complete randomized block design for a breeding trial with 6 treatments and two blocks and one plant per plot. The colors represent a production for a hypothetical trait.

We wish to fit the SCM to this data. Therefore, our incidence matrix $\boldsymbol{X}$ can be written as:

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

The Z_g can be written as:

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Finally, the Z_c can be written using the competition intensity factors (CIF) by applying the equations 9, 10, and 11.

$$\boldsymbol{X} = \begin{bmatrix} 0 & 0.69 & 0.57 & 0.44 & 0 & 0 \\ 0.53 & 0.53 & 0.34 & 0.44 & 0 & 0.34 \\ 0 & 0.53 & 0 & 0.53 & 0.34 & 0.44 \\ 0 & 0.69 & 0 & 0 & 0.57 & 0.44 \\ 0.47 & 0.36 & 0 & 0.56 & 0.47 & 0.36 \\ 0.29 & 0.38 & 0.45 & 0 & 0.29 & 0.45 \\ 0.38 & 0.38 & 0.29 & 0.45 & 0.45 & 0.29 \\ 0.36 & 0.36 & 0.47 & 0.47 & 0 & 0.56 \\ 0 & 0 & 0.57 & 0.44 & 0 & 0.69 \\ 0.53 & 0 & 0.34 & 0.44 & 0.53 & 0.34 \\ 0 & 0 & 0.53 & 0.34 & 0.34 & 0.53 \\ 0.69 & 0 & 0 & 0 & 0.57 & 0.44 \end{bmatrix}$$

2.11. Supplementary material

Table S1: Direct genetic effects (DGE), indirect genetic effects (IGE) and their respective standard errors (SE) and reliability (REL), total breeding value (TBV) for 73 clones of eucalyptus at 3 years and competition class (A is aggressive, H is homeostatic and P is passive).

Trat	DGE	SE.DGE	IGE	SE.IGE	IGE_CIF	TBVi	Class
C57802	16.62	2.47	-2.50	0.88	-6.46	10.16	A
C47055	15.05	2.47	-2.04	0.88	-5.27	9.78	A
C46676	16.35	2.49	-2.71	0.88	-6.98	9.37	A
C46684	15.03	2.55	-2.33	0.88	-6.01	9.02	A
C50013	14.47	2.42	-2.21	0.88	-5.70	8.76	A
C51727	11.49	2.46	-1.73	0.88	-4.45	7.04	A
C50976	9.85	2.47	-1.11	0.88	-2.87	6.99	H
C50235	12.24	2.56	-2.05	0.88	-5.29	6.96	A
C47779	11.33	2.42	-1.99	0.88	-5.14	6.19	A
C52631	9.44	2.48	-1.30	0.89	-3.35	6.09	H
C55132	9.41	2.51	-1.31	0.89	-3.37	6.04	H
C47215	8.89	2.62	-1.29	0.89	-3.33	5.56	H
C47356	10.52	2.52	-1.98	0.89	-5.10	5.43	A
C48532	10.51	2.45	-2.04	0.88	-5.25	5.26	A
CNB010	10.03	1.97	-1.90	0.82	-4.90	5.13	A
C57579	7.14	2.42	-0.78	0.88	-2.02	5.11	H
C32288	7.97	2.69	-1.14	0.91	-2.93	5.04	H
C57918	10.30	2.55	-2.06	0.89	-5.32	4.99	A
C57509	10.22	2.47	-2.10	0.88	-5.42	4.80	A
C50859	4.93	2.53	-0.99	0.89	-2.55	2.38	H
C56672	2.50	2.50	-0.08	0.89	-0.22	2.29	H
C55013	4.51	2.66	-0.89	0.90	-2.29	2.22	H
C47256	5.91	2.80	-1.44	0.90	-3.72	2.19	H
C57859	3.45	2.70	-0.55	0.90	-1.43	2.02	H
C52661	3.91	2.56	-0.73	0.89	-1.89	2.01	H
C47246	3.37	2.52	-0.64	0.88	-1.64	1.73	H
CNB005	0.54	2.04	0.45	0.82	1.17	1.71	H
C47023	3.91	2.47	-0.98	0.88	-2.54	1.37	H
C51146	0.81	2.67	0.07	0.89	0.18	0.99	H
C47868	0.38	2.49	0.22	0.88	0.56	0.94	H
C57931	2.24	2.59	-0.58	0.89	-1.49	0.75	H
C57660	-1.30	2.57	0.58	0.88	1.49	0.20	H
C57387	-1.08	2.53	0.42	0.89	1.10	0.02	H
C48980	2.29	2.50	-0.89	0.88	-2.29	-0.01	H
C30607	-0.95	2.67	0.27	0.90	0.69	-0.26	H
C57883	-0.23	2.47	-0.11	0.88	-0.29	-0.52	H
C56733	-1.51	2.50	0.36	0.88	0.94	-0.57	H
C56422	-1.81	2.58	0.43	0.89	1.11	-0.69	H
C30709	-0.14	2.71	-0.24	0.91	-0.61	-0.76	H
C30492	-1.90	2.85	0.35	0.91	0.91	-0.99	H

C48142	-3.38	2.42	0.87	0.88	2.23	-1.15	H
C52871	-0.03	2.46	-0.46	0.88	-1.19	-1.22	H
C56658	0.24	2.63	-0.63	0.90	-1.64	-1.40	H
C31938	-3.07	2.68	0.61	0.90	1.59	-1.48	H
C57701	-1.06	2.62	-0.17	0.89	-0.43	-1.49	H
C49792	-3.68	2.47	0.80	0.88	2.07	-1.61	H
C47696	-2.58	2.52	0.18	0.88	0.48	-2.11	H
CNB016	-4.05	1.97	0.61	0.82	1.58	-2.47	H
C48107	-4.91	2.73	0.94	0.90	2.43	-2.48	H
C49742	-3.66	2.61	0.45	0.89	1.17	-2.49	H
C32118	-4.71	2.71	0.82	0.91	2.10	-2.60	H
C31046	-7.40	2.78	1.84	0.91	4.76	-2.64	P
C00302	-5.66	2.56	1.10	0.89	2.85	-2.81	H
C46730	-5.40	2.52	0.88	0.88	2.28	-3.13	H
C48838	-3.77	2.58	0.21	0.89	0.55	-3.22	H
C49907	-6.72	2.85	0.98	0.90	2.53	-4.19	H
C57481	-7.61	2.88	1.32	0.91	3.41	-4.19	H
C47291	-8.72	2.47	1.56	0.88	4.03	-4.69	P
C50125	-6.83	2.53	0.82	0.88	2.12	-4.71	H
C31929	-9.59	2.73	1.79	0.90	4.63	-4.97	P
C57967	-9.03	2.55	1.50	0.89	3.87	-5.16	H
C57595	-9.83	2.80	1.80	0.90	4.65	-5.17	P
C00256	-11.33	2.85	2.37	0.90	6.12	-5.21	P
C51554	-13.65	3.49	2.69	0.94	6.95	-6.70	P
C48856	-12.15	2.56	2.11	0.89	5.44	-6.72	P
C31161	-12.93	2.87	2.32	0.92	5.98	-6.95	P
C31934	-13.93	2.98	2.30	0.93	5.94	-7.99	P
C31323	-12.02	2.87	1.56	0.91	4.02	-8.00	P
C52006	-16.11	2.72	2.92	0.89	7.54	-8.57	P
C30689	-15.47	3.01	2.44	0.93	6.30	-9.17	P
C52812	-16.02	3.12	2.44	0.92	6.29	-9.74	P
C35571	-17.92	3.33	2.79	0.93	7.19	-10.73	P
C00473	-20.91	4.34	3.55	1.01	9.15	-11.76	P

Table S2: Direct genetic effects (DGE), indirect genetic effects (IGE) and their respective standard errors (SE) and reliability (REL), total breeding value (TBV) for 73 clones of eucalyptus at 6 years and competition class (A is aggressive, H is homeostatic and P is passive).

Trat	DGE	SE.DGE	IGE	SE.IGE	IGE_CIF	TBVi	Class
C50013	33.26	3.89	-4.33	1.96	-10.70	22.56	A
C50976	26.05	3.82	-3.03	1.97	-7.50	18.55	A
C57802	23.18	3.84	-2.47	1.98	-6.09	17.08	H
C51727	27.91	3.83	-4.69	1.98	-11.61	16.31	A
C46684	26.69	3.95	-4.20	1.95	-10.39	16.30	A
C47055	20.15	3.83	-1.83	1.97	-4.51	15.63	H
C55132	17.33	3.90	-0.73	1.98	-1.81	15.53	H

C46676	27.50	3.92	-4.93	1.96	-12.19	15.30	A
C50235	28.28	3.96	-5.54	1.95	-13.69	14.60	A
C47779	25.14	3.75	-4.66	1.97	-11.52	13.62	A
C51146	7.16	4.23	1.25	2.00	3.08	10.25	H
C56672	6.70	4.01	1.34	1.98	3.32	10.02	H
C50859	15.03	4.06	-2.23	1.98	-5.50	9.53	H
C57579	9.73	3.90	-0.35	1.98	-0.87	8.86	H
C47215	14.43	4.03	-2.60	1.95	-6.43	8.00	H
C48838	11.63	4.93	-1.59	1.96	-3.93	7.71	H
CNB005	5.88	3.21	0.59	1.78	1.45	7.33	H
C52661	6.19	4.19	0.31	1.99	0.78	6.97	H
C47256	17.32	4.35	-4.71	1.97	-11.65	5.67	A
C30607	1.73	4.18	1.54	2.05	3.81	5.54	H
C52631	8.58	3.99	-1.26	1.98	-3.10	5.48	H
C48532	16.21	3.82	-4.35	1.97	-10.76	5.45	A
C57918	15.74	3.96	-4.24	1.97	-10.49	5.25	A
C55013	6.80	4.17	-0.65	2.06	-1.61	5.19	H
C30492	2.86	4.36	0.72	2.06	1.77	4.63	H
C32288	9.23	4.20	-2.21	2.07	-5.45	3.78	H
C57931	4.04	4.06	-0.38	1.97	-0.93	3.11	H
C56422	-1.74	4.12	1.84	1.97	4.54	2.79	H
C57387	-2.42	3.97	1.43	1.96	3.53	1.11	H
C47356	9.46	3.89	-3.44	1.96	-8.49	0.97	A
C47868	-2.99	3.91	1.35	1.97	3.33	0.34	H
CNB010	4.84	3.23	-1.97	1.77	-4.88	-0.04	H
C00302	-4.58	4.03	1.69	1.98	4.18	-0.41	H
C57660	1.35	3.95	-0.96	1.96	-2.37	-1.02	H
C47023	3.65	3.89	-2.20	1.98	-5.44	-1.79	H
C49792	-7.54	3.92	2.22	1.96	5.50	-2.04	H
C31046	-10.08	4.47	3.17	2.04	7.83	-2.26	P
C57509	4.11	3.90	-2.61	1.99	-6.46	-2.35	H
C47246	4.25	4.11	-2.77	1.94	-6.85	-2.60	H
C52871	-0.11	3.89	-1.24	1.94	-3.07	-3.18	H
C32118	-8.54	4.32	2.15	2.07	5.31	-3.22	H
C30709	-1.66	4.35	-1.04	2.05	-2.56	-4.22	H
C49742	-8.48	4.21	1.72	1.98	4.25	-4.23	H
C31929	-12.03	4.21	3.12	2.03	7.72	-4.31	P
C57883	-2.09	3.86	-1.02	1.96	-2.51	-4.60	H
C31938	-9.72	4.28	2.06	2.04	5.10	-4.63	H
C00256	-14.10	4.50	3.78	1.98	9.34	-4.77	P
C48142	-8.35	3.84	1.27	1.97	3.13	-5.22	H
C57701	-3.42	4.04	-0.81	1.98	-2.00	-5.42	H
C56733	-8.34	3.89	1.10	1.97	2.71	-5.63	H
C48980	-3.29	4.04	-0.98	1.97	-2.43	-5.72	H
C51554	-15.44	5.21	3.77	1.97	9.31	-6.13	P
C57481	-12.45	4.45	2.54	2.02	6.29	-6.16	H
C49907	-10.77	4.84	0.93	1.95	2.29	-8.48	H
C57595	-15.83	4.47	2.63	1.99	6.50	-9.32	H

C00473	-15.45	9.30	2.45	2.12	6.06	-9.39	H
C57859	-11.79	4.42	0.97	2.04	2.39	-9.40	H
C47291	-17.11	3.98	3.07	1.98	7.59	-9.52	P
C56658	-8.63	4.14	-0.41	1.99	-1.00	-9.64	H
C50125	-12.67	3.90	1.22	1.98	3.01	-9.66	H
C48107	-14.52	4.35	1.96	1.97	4.84	-9.69	H
C46730	-16.40	4.48	2.71	1.96	6.70	-9.70	H
CNB016	-18.67	3.51	3.53	1.75	8.72	-9.95	P
C30689	-23.48	4.90	5.44	2.06	13.46	-10.02	P
C47696	-13.57	4.16	1.12	1.93	2.78	-10.80	H
C31161	-22.07	4.91	4.32	2.06	10.67	-11.40	P
C52812	-18.70	6.19	2.88	1.99	7.13	-11.57	P
C52006	-26.56	5.00	6.05	1.95	14.95	-11.61	P
C35571	-20.58	8.34	3.54	2.06	8.74	-11.83	P
C31323	-18.23	4.69	2.47	2.03	6.10	-12.13	H
C31934	-21.84	4.95	3.70	2.07	9.16	-12.68	P
C57967	-20.40	5.16	2.37	1.97	5.87	-14.53	H
C48856	-22.12	4.30	2.93	1.97	7.24	-14.88	P

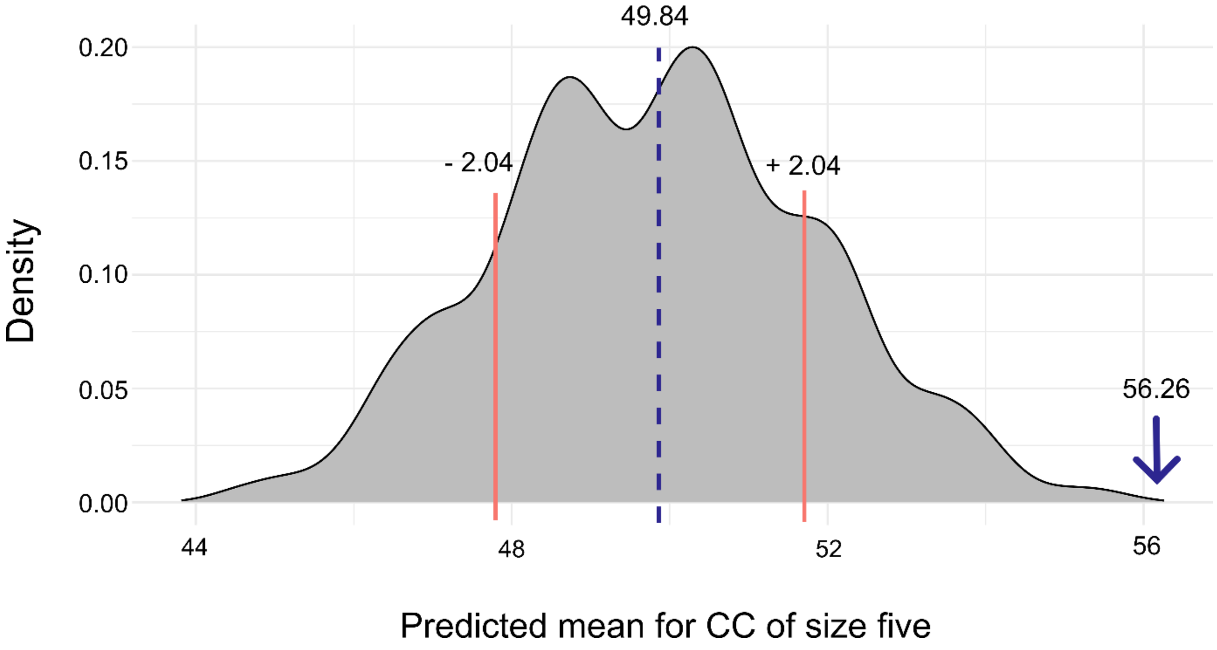

Figure S1: Density plot for expected mean annual increment (MAI) in m³ ha-¹ year-¹ for eucalyptus at 6 years for 15,504 combinations of clonal composites (CC) of size 5. The blue dotted line represents the distribution mean and the coral solid lines represent the standard deviations. The blue narrow indicates the predicted mean for the best