

Rafael Cupertino de Lima Muzzi

**Avaliação da eficiência de modelos baseados
em *pedigree* e modelos genômicos na seleção
de indivíduos de *Eucalyptus***

Viçosa, MG

2026

Rafael Cupertino de Lima Muzzi

**Avaliação da eficiência de modelos baseados em
pedigree e modelos genômicos na seleção de
indivíduos de *Eucalyptus***

Monografia apresentada à Universidade Federal de Viçosa como parte das exigências da disciplina BIO 499 – Trabalho de Conclusão de Curso

Orientador Kaio Olimpio das Graças Dias

Viçosa, MG

2026

Rafael Cupertino de Lima Muzzi

**Avaliação da eficiência de modelos baseados em
pedigree e modelos genômicos na seleção de
indivíduos de *Eucalyptus***

Monografia apresentada à Universidade
Federal de Viçosa como parte das exigên-
cias da disciplina BIO 499 – Trabalho de
Conclusão de Curso

APROVADO: 25 de junho de 2026.

Assentimento:

Documento assinado digitalmente
 **RAFAEL CUPERTINO DE LIMA MUZZI**
Data: 05/07/2026 22:16:32-0300
Verifique em <https://validar.iti.gov.br>

Rafael Cupertino de Lima Muzzi
Autor

Documento assinado digitalmente
 **KAIO OLIMPIO DAS GRACAS DIAS**
Data: 03/07/2026 18:27:28-0300
Verifique em <https://validar.iti.gov.br>

Kaio Olimpio das Graças Dias
Orientador

Agradecimentos

Agradeço a Deus, por me conceder saúde, força, sabedoria, oportunidade e por me guiar durante toda a minha trajetória, tornando possível a realização deste trabalho.

Agradeço ao Prof. Dr. Kaio Olímpio das Graças Dias pela orientação, confiança, disponibilidade e pelos ensinamentos compartilhados ao longo do desenvolvimento deste trabalho.

Agradeço aos colegas do laboratório pela convivência, pelo apoio, pela ajuda, pelas discussões e por todo o conhecimento compartilhado ao longo desta jornada.

À Celulose Nipo-Brasileira S.A. (CENIBRA), agradeço pela parceria com o laboratório, pela disponibilização dos dados utilizados nesta pesquisa e pela autorização para sua utilização neste trabalho.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio às atividades de pesquisa desenvolvidas no laboratório, que contribuíram para a realização deste estudo.

Aos meus pais, expresso minha profunda gratidão pelo amor, apoio incondicional, incentivo e confiança durante toda a minha formação. Sem vocês, esta conquista não seria possível.

Por fim, agradeço à Universidade Federal de Viçosa e a todos aqueles que contribuíram, direta ou indiretamente, para a realização deste trabalho.

Resumo

Com os avanços da genética, os programas de melhoramento passaram a incorporar novas ferramentas e tecnologias que aumentaram a precisão da seleção. No melhoramento florestal, um dos principais desafios está relacionado ao longo ciclo de produção dos indivíduos e à necessidade de maior eficiência seletiva. Entre os avanços mais relevantes, destaca-se o uso de marcadores moleculares, como os SNPs, que possibilitam a obtenção de informações genômicas mais detalhadas sobre os indivíduos. Esses marcadores permitiram estimar relações genéticas com maior precisão do que aquelas baseadas exclusivamente em *pedigree*, ampliando as possibilidades de seleção de progênies. A integração desses dados genômicos com modelos estatísticos resultou no desenvolvimento de modelos genômicos mais eficientes para a predição de valores genéticos quando comparados aos modelos baseados apenas em *pedigree*. Neste contexto, o presente estudo tem como objetivo comparar a eficiência de modelos baseados em *pedigree* e em relacionamento genômico na seleção de indivíduos de *Eucalyptus*. Foram ajustados modelos lineares mistos com estruturas aditiva e aditiva-dominante, utilizando matrizes de relacionamento baseadas em *pedigree* e dados genômicos. A eficiência dos modelos foi avaliada por meio da estimativa de parâmetros genéticos, capacidade preditiva via validação cruzada e coincidência na seleção dos indivíduos superiores. Os modelos genômicos apresentaram melhor ajuste aos dados em relação ao modelo baseado em *pedigree*, sendo o GDBLUP o modelo com melhor ajuste aos dados. Além disso, foi observada elevada concordância entre os modelos genômicos e boa capacidade preditiva para as características avaliadas. Esses resultados evidenciam o potencial da seleção genômica para aumentar a eficiência da seleção e contribuir para a redução do ciclo de melhoramento em *Eucalyptus urograndis*.

Palavras-chaves: Seleção genômica; *E. urograndis*; SNPs; Melhoramento florestal; BLUP genômico.

Lista de figuras

- Figura 1 – Dispersão e concordância entre os valores genéticos preditos obtidos via modelo genômico aditivo-dominante (GDBLUP) e modelo genômico aditivo (GBLUP) para o incremento médio anual em volume (IMAvol) em indivíduos de *Eucalyptus*. Os pontos coloridos destacam os indivíduos que seriam selecionados como superiores por ambos os modelos (laranja), exclusivamente pelo GDBLUP (vinho) ou exclusivamente pelo GBLUP (verde), sob uma mesma intensidade de seleção. Os pontos cinzas representam os indivíduos não selecionados. 18
- Figura 2 – Dispersão e concordância entre os valores genéticos preditos de modelos genômicos em relação ao modelo baseado tradicionalmente em pedigree (ABLUP) para o incremento médio anual em volume (IMAvol) em indivíduos de *Eucalyptus*. Em (a) apresenta-se a comparação entre o modelo genômico aditivo (GBLUP) e o modelo por pedigree (ABLUP). Em (b) apresenta-se a comparação entre o modelo genômico aditivo-dominante (GDBLUP) e o modelo por pedigree (ABLUP). As cores indicam os indivíduos selecionados exclusivamente por um modelo, por ambos simultaneamente ou não selecionados (em cinza). 18

Lista de tabelas

Tabela 1	–	Resumo do processo de filtragem dos marcadores SNP em indivíduos de <i>E. urograndis</i> , considerando o critério de frequência alélica mínima ($MAF \geq 0,05$).	12
Tabela 2	–	Estatísticas descritivas contendo média, desvio-padrão (DP), valores mínimos e máximos para as características diâmetro à altura do peito (DAP), altura total (ALT) e incremento médio anual em volume (IMAvol), avaliadas aos 6 anos de idade em 786 indivíduos sobreviventes de <i>Eucalyptus urograndis</i>	15
Tabela 3	–	Valores do Critério de Informação de Akaike (AIC) obtidos para o ajuste das características diâmetro à altura do peito (DAP), altura total (ALT) e incremento médio anual em volume (IMAvol), considerando os modelos ABLUP, GBLUP e GDBLUP.	16
Tabela 4	–	Coefficientes de correlação de postos de Spearman (r_S) entre os rankings de valores genéticos preditos para o incremento médio anual em volume (IMAvol), avaliando a concordância na classificação dos indivíduos pelos modelos ABLUP, GBLUP e GDBLUP.	17
Tabela 5	–	Capacidade preditiva do modelo GDBLUP, estimada por meio da correlação de Pearson (r) entre os valores genéticos preditos e os valores fenotípicos observados, obtida por validação cruzada do tipo 5-fold. . .	19

Sumário

1	INTRODUÇÃO	8
2	OBJETIVOS	10
2.1	Objetivo geral	10
2.2	Objetivos específicos	10
3	MATERIAIS E MÉTODOS	11
3.1	Dados fenotípicos	11
3.2	Dados moleculares	11
3.3	Modelos estatísticos	12
3.4	Validação cruzada k-fold	14
4	RESULTADOS E DISCUSSÃO	15
4.1	Estatísticas descritivas	15
4.2	Comparação de modelos	15
4.3	Concordância entre modelos	16
4.4	Capacidade preditiva dos modelos	19
5	CONCLUSÃO	20
	REFERÊNCIAS BIBLIOGRÁFICAS	21

1 Introdução

A expansão do setor de florestas plantadas no Brasil intensificou-se a partir de 1967, impulsionada pela implementação de políticas de incentivos fiscais, como a Lei nº 5.106, de 2 de setembro de 1966 (RESENDE, 1999), que visavam reduzir a exploração de florestas nativas e estimular o estabelecimento de florestas plantadas. Desde então, as espécies do gênero *Eucalyptus* consolidaram-se como as principais espécies florestais cultivadas no Brasil, desempenhando papel fundamental no fornecimento de matéria-prima para diferentes cadeias produtivas, como as indústrias de madeira serrada, carvão vegetal e celulose. Em 2024, os plantios de eucalipto totalizaram cerca de 8,1 milhões de hectares no país (Indústria Brasileira de Árvores, 2025). A expansão da cultura para diferentes regiões produtoras, incluindo estados como Mato Grosso do Sul, trouxe novos desafios relacionados à adaptação dos materiais genéticos às distintas condições ambientais e às crescentes demandas por produtividade. Nesse contexto, o melhoramento genético tornou-se uma ferramenta estratégica para o desenvolvimento de genótipos superiores, capazes de atender às necessidades do setor florestal de forma mais eficiente e sustentável.

A crescente demanda por ganhos de produtividade, maior eficiência de seleção e redução do ciclo de melhoramento impulsionou o desenvolvimento de novas estratégias nos programas de melhoramento florestal. Ao longo das últimas décadas, avanços como a hibridação interespecífica, a propagação clonal e a seleção recorrente contribuíram significativamente para o aumento dos ganhos genéticos obtidos em espécies do gênero *Eucalyptus* (RESENDE; ALVES, 2021). Mais recentemente, os avanços da biologia molecular e da genômica possibilitaram a incorporação de novas ferramentas aos programas de melhoramento, destacando-se os marcadores moleculares do tipo SNP (*Single Nucleotide Polymorphism*), que permitem estimativas mais precisas dos relacionamentos genéticos entre indivíduos (MEUWISSEN; HAYES; GODDARD, 2001). Dessa forma, a seleção genômica, que utiliza informações de marcadores moleculares distribuídos ao longo do genoma para prever o valor genético dos indivíduos, tornou-se uma das principais ferramentas para acelerar o progresso genético e aumentar o ganho por unidade de tempo em espécies florestais (SIMIQUELI et al., 2023; GRATTAPAGLIA, 2022).

Em espécies do gênero *Eucalyptus*, características relacionadas ao crescimento e à produtividade, como o diâmetro à altura do peito (DAP), a altura (ALT), o volume e o incremento médio anual em volume (IMAvol), constituem os principais critérios de seleção utilizados nos programas de melhoramento florestal. Essas características são classificadas como quantitativas, sendo controladas por múltiplos genes e influenciadas por fatores ambientais, o que dificulta a identificação precoce dos indivíduos geneticamente superiores. Nesse contexto, a seleção genômica surge como uma alternativa promissora, permitindo

aumentar a precisão das estimativas dos valores genéticos e a eficiência da seleção. Dessa forma, a utilização de informações genômicas tem contribuído para acelerar o progresso genético e reduzir o tempo necessário para a obtenção de materiais superiores (RESENDE, 2008).

Os marcadores moleculares do tipo SNP representam atualmente uma das principais fontes de informação molecular utilizadas em programas de melhoramento genético. Esses marcadores correspondem à forma mais abundante de variação no DNA, caracterizada pela alteração de um único nucleotídeo na sequência genômica. A utilização de SNPs possibilita a estimativa dos relacionamentos genéticos diretamente a partir do DNA dos indivíduos, permitindo a construção de matrizes de relacionamento genômico (MEUWISSEN; HAYES; GODDARD, 2001). Em comparação às matrizes baseadas exclusivamente em informações de pedigree, as matrizes genômicas apresentam maior capacidade de representar a similaridade genética real entre os indivíduos, tornando-se a base da seleção genômica e contribuindo para o aumento da acurácia na predição de valores genéticos e na identificação de indivíduos superiores (RESENDE, 2008).

Nos modelos de seleção genômica, os efeitos genéticos aditivos são tradicionalmente considerados os principais responsáveis pela resposta à seleção, uma vez que são transmitidos de forma previsível aos descendentes por meio da contribuição individual dos alelos para a característica de interesse (RESENDE, 2008). Em contrapartida, os efeitos de dominância decorrem da interação entre alelos, não podendo ser descritos apenas pela soma dos efeitos alélicos. Em espécies do gênero *Eucalyptus*, nas quais a produção de híbridos e a propagação clonal desempenham papel central nos programas de melhoramento, a inclusão dos efeitos de dominância em modelos genômicos pode contribuir para uma representação mais realista da variabilidade genética e para o aumento da capacidade preditiva dos modelos, especialmente para características relacionadas ao crescimento e à produtividade (CANAL et al., 2023; RESENDE, 2008).

Diante desse contexto, parte-se da hipótese de que modelos baseados em relacionamento genômico apresentam desempenho superior aos modelos baseados em pedigree na predição de valores genéticos em híbridos de *Eucalyptus urograndis*. Além disso, espera-se que a inclusão dos efeitos de dominância contribua para melhorar o ajuste e a capacidade preditiva dos modelos. Assim, este trabalho buscou comparar a eficiência dessas abordagens na seleção de indivíduos superiores.

2 Objetivos

2.1 Objetivo geral

Comparar a eficiência de modelos baseados em pedigree e em relacionamento genômico na predição genética e seleção de indivíduos híbridos de *Eucalyptus urograndis* (*Eucalyptus grandis* $\times$ *Eucalyptus urophylla*).

2.2 Objetivos específicos

- Avaliar a eficiência de informações de parentesco baseadas em *pedigree* e em marcadores genômicos na seleção de híbridos de *E. urograndis* para avanço no programa de melhoramento e desenvolvimento de clones comerciais;
- Comparar os parâmetros genéticos estimados para as características DAP, ALT e IMAvol em uma população híbrida de *E. urograndis*;
- Avaliar a concordância na seleção de indivíduos superiores obtida por informações de parentesco baseadas em *pedigree* e em marcadores genômicos em uma população híbrida de *E. urograndis*;
- Comparar o desempenho dos modelos ABLUP, GBLUP e GDBLUP quanto ao ajuste aos dados;
- Avaliar a capacidade preditiva do modelo selecionado com base no Critério de Informação de Akaike (AIC), por meio de validação cruzada.

3 Materiais e métodos

Os dados utilizados no presente trabalho foram fornecidos pelo programa de melhoramento florestal da empresa Celulose Nipo-Brasileira (Cenibra).

3.1 Dados fenotípicos

Os dados fenotípicos são resultantes de experimentos conduzidos em delineamento em blocos completos casualizados (DBCC), com um delineamento dialélico parcial sem recíprocos, utilizando 45 famílias de híbridos de *E. urograndis*, 23 progênies por família, totalizando 1.035 indivíduos. Os indivíduos que apresentaram mortalidade foram representados por valores iguais a zero e tratados como dados ausentes (NA), resultando em 786 observações utilizadas nas análises.

Normalmente, em experimentos florestais são realizadas avaliações em dois períodos, próximos aos 3 anos e aos 6 anos de idade. Neste estudo, foram utilizados os dados referentes à avaliação realizada aos 6 anos. Além disso, o arquivo de dados continha variáveis, como: identificação de bloco, família, indivíduo, código dos pais, DAP, altura, volume e o IMAvol. O IMAvol foi utilizado como variável resposta principal, por representar um indicativo da produtividade baseado no incremento volumétrico em relação ao tempo de plantio. Além disso, as características DAP e altura também foram avaliadas, permitindo verificar o desempenho dos modelos em diferentes características de crescimento de interesse para o melhoramento florestal.

3.2 Dados moleculares

Os dados moleculares foram obtidos para 1035 indivíduos. Esses dados foram importados e processados no software R, onde a matriz original apresentava os marcadores SNP nas linhas e os indivíduos nas colunas, sendo necessário, posteriormente, uma transposição para que se adequasse ao padrão necessário para as análises genômicas, principalmente para a construção da matriz G (matriz de parentesco genômico). Os genótipos foram codificados numericamente segundo o sistema de dosagem alélica, em que os alelos AA correspondem simbolicamente à 0, AB à 1 e BB à 2 (AA=0; AB=1; BB=2).

Os dados genômicos foram submetidos a controle de qualidade utilizando o pacote AGHmatrix (AMADEU et al., 2023) no software R. Foram removidos marcadores com frequência alélica mínima (MAF) inferior a 0,05. Não foi realizada remoção adicional de marcadores redundantes, uma vez que o controle de qualidade adotado baseou-se

exclusivamente no critério de frequência alélica mínima. Em seguida, foram construídas as matrizes de relacionamento genômico aditiva (G) e de dominância (D), utilizadas nas análises subsequentes.

Inicialmente, estavam disponíveis 28.177 marcadores SNP. Após a aplicação dos critérios de controle de qualidade, permaneceram 20.705 marcadores, os quais foram utilizados na construção das matrizes de relacionamento genômico e nas análises subsequentes.

A síntese do processo de filtragem encontra-se apresentada na Tabela 1.

Tabela 1 – Resumo do processo de filtragem dos marcadores SNP em indivíduos de *E. urograndis*, considerando o critério de frequência alélica mínima ($MAF \geq 0,05$).

Etapa	Número de SNPs
SNPs iniciais	28.177
SNPs após controle de qualidade	20.705

3.3 Modelos estatísticos

Inicialmente, foram ajustados modelos mistos utilizando dados de *pedigree* dos indivíduos, levando em consideração efeitos genéticos aditivos baseados na estrutura de parentesco das famílias.

Posteriormente, foram realizadas análises de seleção genômica utilizando matrizes de relacionamento genômico obtidas a partir de marcadores SNP.

Os modelos genômicos avaliados incluíram:

- Modelo baseado em *pedigree*;
- Modelo genômico aditivo;
- Modelo genômico aditivo + dominante.

Os modelos foram ajustados utilizando modelos mistos lineares, considerando efeitos fixos e aleatórios. Os efeitos fixos correspondem aos fatores sistemáticos do experimento, como os blocos, enquanto os efeitos aleatórios representam fontes de variação associadas aos efeitos genéticos dos indivíduos e aos resíduos experimentais. Os modelos estão representados genericamente por:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{e}$$

para o modelo aditivo, e:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \mathbf{e}$$

para o modelo aditivo + dominante, em que:

- $\mathbf{y}$ representa o vetor de observações fenotípicas;
- $\mathbf{b}$ representa o vetor de efeitos fixos;
- $\mathbf{a}$ representa o vetor de efeitos genéticos aditivos;
- $\mathbf{d}$ representa o vetor de efeitos de dominância;
- $\mathbf{e}$ representa o vetor dos erros aleatórios;
- $\mathbf{X}$ é a matriz de incidência dos efeitos fixos;
- $\mathbf{Z}_1$ é a matriz de incidência dos efeitos genéticos aditivos;
- $\mathbf{Z}_2$ é a matriz de incidência dos efeitos de dominância.

O pacote ASReml (VSN International, 2023), empregado no software R (R Core Team, 2025), amplamente utilizado em estudos de melhoramento genético permitiu a estimação de componentes de variância e a predição de valores genéticos por meio de modelos lineares mistos. Inicialmente, foi ajustado o modelo ABLUP, que utiliza uma matriz de relacionamento baseada em *pedigree* (matriz $\mathbf{A}$), construída a partir das informações de genealogia dos indivíduos, para representar o grau de parentesco esperado entre eles.

Em seguida, foi ajustado o modelo GBLUP, no qual a matriz de relacionamento baseada em *pedigree* ($\mathbf{A}$) foi substituída pela matriz de relacionamento genômico ($\mathbf{G}$), construída a partir dos marcadores SNP.

A matriz genômica aditiva foi construída conforme VanRaden (2008):

$$G_{jk} = \frac{\sum_{i=1}^m (x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2 \sum_{i=1}^m p_i(1 - p_i)} \quad (3.1)$$

em que x_{ij} representa o número de alelos da referência no indivíduo j no marcador i , p_i é a frequência alélica do marcador i , e m é o número total de marcadores utilizados.

Por fim, foi ajustado o modelo GDBLUP, que inclui simultaneamente os efeitos genéticos aditivos e de dominância. Esse modelo estende o GBLUP ao incorporar a matriz de dominância ($\mathbf{D}$), permitindo a consideração simultânea dos efeitos genéticos aditivos e de dominância.

A matriz de dominância ($\mathbf{D}$) foi construída conforme o método proposto por Vitezica (VITEZICA; VARONA; LEGARRA, 2013), garantindo a ortogonalidade entre os efeitos aditivos e de dominância.

$$D_{jk} = \frac{\sum_{i=1}^m w_{ij}w_{ik}}{\sum_{i=1}^m 2p_i(1 - p_i) [1 - 2p_i(1 - p_i)]} \quad (3.2)$$

Em que w_{ij} representa o código de dominância do genótipo do indivíduo j no marcador i ; p_i é a frequência alélica do marcador i ; e m é o número total de marcadores utilizados.

3.4 Validação cruzada k-fold

A capacidade preditiva dos modelos foi avaliada por validação cruzada do tipo 5-fold. Inicialmente, os indivíduos com dados fenotípicos disponíveis foram embaralhados aleatoriamente e divididos em cinco grupos de tamanho semelhante. Em cada rodada, um dos grupos teve seus fenótipos mascarados e foi utilizado para validação, enquanto os demais grupos foram utilizados no ajuste do modelo. Esse processo foi repetido até que todos os indivíduos fossem utilizados uma vez na validação. A capacidade preditiva foi obtida pela correlação entre os valores genéticos preditos e os valores fenotípicos observados.

4 Resultados e Discussão

4.1 Estatísticas descritivas

O IMAvol, utilizado como variável resposta, apresentou média de $57,38 \text{ m}^3 \text{ ha}^{-1} \text{ ano}^{-1}$, desvio-padrão de $40,20 \text{ m}^3 \text{ ha}^{-1} \text{ ano}^{-1}$, valor mínimo de $2,04 \text{ m}^3 \text{ ha}^{-1} \text{ ano}^{-1}$ e valor máximo de $216,19 \text{ m}^3 \text{ ha}^{-1} \text{ ano}^{-1}$. A amplitude observada entre os valores mínimo e máximo, associada ao elevado desvio-padrão, indica a existência de variabilidade fenotípica entre os indivíduos avaliados, condição favorável para a identificação de diferenças genéticas e para o avanço dos programas de melhoramento.

Tabela 2 – Estatísticas descritivas contendo média, desvio-padrão (DP), valores mínimos e máximos para as características diâmetro à altura do peito (DAP), altura total (ALT) e incremento médio anual em volume (IMAvol), avaliadas aos 6 anos de idade em 786 indivíduos sobreviventes de *Eucalyptus urograndis*.

Característica	Média	DP	Mínimo	Máximo
IMA	57,38	40,20	2,04	216,19
DAP	14,76	4,36	5,20	27,00
Altura	23,39	5,49	7,30	31,90

Os resultados apresentados na Tabela 2 evidenciam variabilidade fenotípica para todas as características avaliadas. Essa variabilidade constitui requisito fundamental para o progresso genético por seleção, uma vez que permite a identificação de diferenças de desempenho entre os indivíduos da população.

4.2 Comparação de modelos

Os valores do Critério de Informação de Akaike (AIC) obtidos para os modelos avaliados são apresentados na Tabela 3. Para todas as características analisadas, o modelo GDBLUP apresentou os menores valores de AIC, seguido pelo GBLUP e pelo ABLUP. Como menores valores de AIC indicam melhor ajuste aos dados, os resultados sugerem superioridade dos modelos genômicos em relação ao modelo baseado exclusivamente em informações de *pedigree*.

Tabela 3 – Valores do Critério de Informação de Akaike (AIC) obtidos para o ajuste das características diâmetro à altura do peito (DAP), altura total (ALT) e incremento médio anual em volume (IMAvol), considerando os modelos ABLUP, GBLUP e GDBLUP.

Modelo	DAP	Altura	IMAvol
ABLUP	3075,295	3418,637	6461,591
GBLUP	2938,694	3311,525	6318,642
GDBLUP	2863,036	3240,336	6262,760

Os menores valores de AIC observados para os modelos genômicos indicam melhor ajuste aos dados em comparação ao modelo baseado em *pedigree*. Esse comportamento é coerente com os resultados obtidos por Tan et al. (2017), que demonstraram que a utilização da matriz de relacionamento genômico proporciona estimativas mais precisas das relações genéticas entre indivíduos de *Eucalyptus*, resultando em melhor capacidade preditiva.

Esse resultado está de acordo com os achados de Canal et al. (2023), que também observaram melhor ajuste dos modelos genômicos em relação aos modelos baseados em *pedigree*, com redução dos valores de AIC e BIC quando os efeitos de dominância foram incorporados ao modelo.

Dentre os modelos, o GDBLUP apresentou desempenho superior ao GBLUP em todas as características analisadas. O melhor desempenho do GDBLUP pode ser atribuído à inclusão simultânea dos efeitos genéticos aditivos e de dominância, permitindo uma descrição mais completa da variabilidade genética presente na população. A importância dos efeitos de dominância torna-se especialmente relevante em programas de melhoramento de *Eucalyptus*, em que a utilização de híbridos e a propagação clonal são amplamente utilizadas.

Dessa forma, modelos capazes de capturar tanto os efeitos aditivos quanto os não aditivos podem contribuir para aumentar a eficiência da seleção de indivíduos superiores, favorecendo a identificação de materiais com maior potencial produtivo. Em programas de melhoramento florestal, modelos com melhor ajuste podem proporcionar estimativas genéticas mais confiáveis, contribuindo para decisões de seleção mais eficientes e para a redução do ciclo de melhoramento.

4.3 Concordância entre modelos

A concordância entre os modelos foi avaliada por meio da correlação de Spearman entre os valores genéticos preditos pelos modelos ABLUP, GBLUP e GDBLUP. Os resultados indicaram correlação moderada entre os modelos ABLUP e GBLUP ($r_s = 0,612$),

bem como entre os modelos ABLUP e GDBLUP ($r_s = 0,623$). Em contrapartida, foi observada elevada correlação entre os modelos GBLUP e GDBLUP ($r_s = 0,845$).

Tabela 4 – Coeficientes de correlação de postos de Spearman (rS) entre os rankings de valores genéticos preditos para o incremento médio anual em volume (IMAvol), avaliando a concordância na classificação dos indivíduos pelos modelos ABLUP, GBLUP e GDBLUP.

Comparação	Correlação de Spearman
ABLUP $\times$ GBLUP	0,612
ABLUP $\times$ GDBLUP	0,623
GBLUP $\times$ GDBLUP	0,845

Os resultados sugerem que os modelos baseados em informações genômicas apresentaram maior concordância entre si na classificação dos indivíduos, quando comparados ao modelo baseado exclusivamente em informações de *pedigree*. Dessa forma, embora os três modelos identifiquem parcialmente os mesmos indivíduos superiores, a utilização de informações genômicas tende a produzir classificações mais semelhantes entre GBLUP e GDBLUP, refletindo a maior capacidade dessas abordagens em capturar os relacionamentos genéticos reais presentes na população avaliada.

A maior correlação observada entre GBLUP e GDBLUP também foi evidenciada visualmente na Figura 1, que mostra maior concentração dos indivíduos ao longo da diagonal principal, indicando elevada concordância entre os valores genéticos preditos por ambos os modelos.

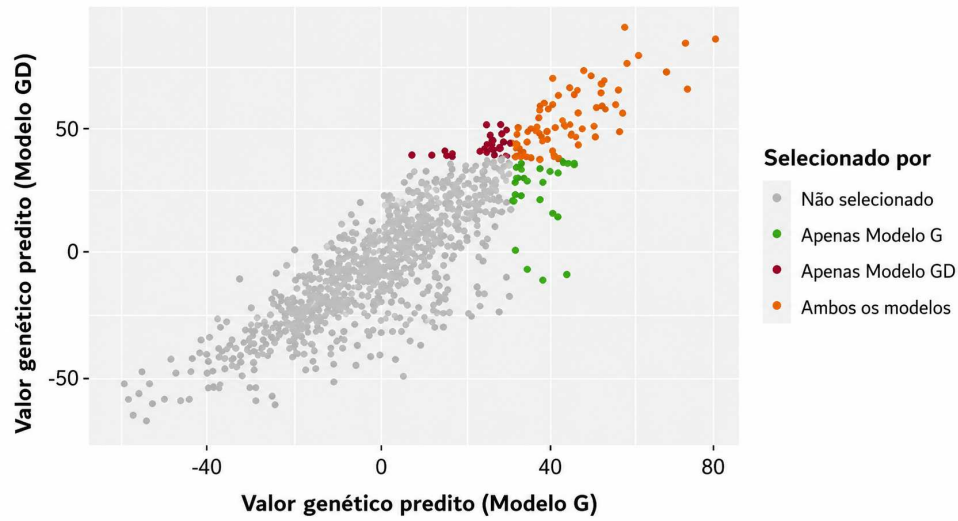


Figura 1 – Dispersão e concordância entre os valores genéticos preditos obtidos via modelo genômico aditivo-dominante (GDBLUP) e modelo genômico aditivo (GBLUP) para o incremento médio anual em volume (IMAvol) em indivíduos de Eucalyptus. Os pontos coloridos destacam os indivíduos que seriam selecionados como superiores por ambos os modelos (laranja), exclusivamente pelo GDBLUP (vinho) ou exclusivamente pelo GBLUP (verde), sob uma mesma intensidade de seleção. Os pontos cinzas representam os indivíduos não selecionados.

Em contrapartida, as correlações entre o modelo ABLUP e os demais modelos foram inferiores.

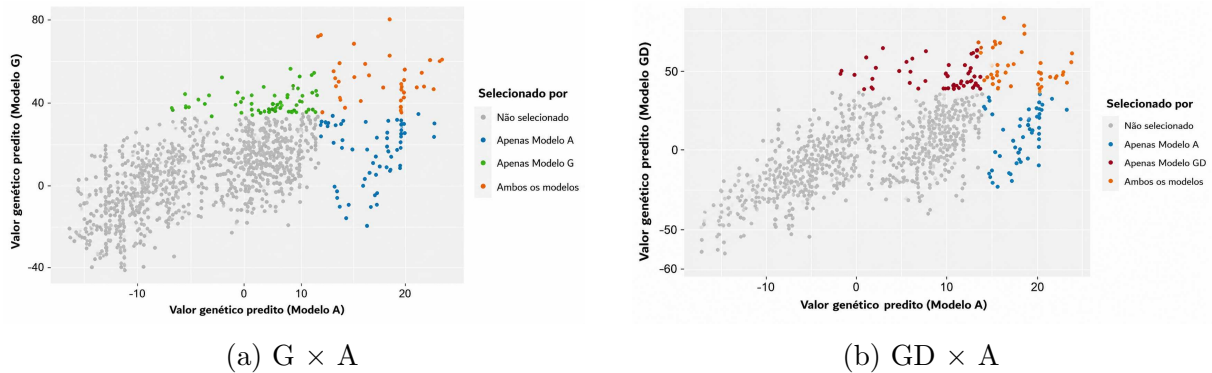


Figura 2 – Dispersão e concordância entre os valores genéticos preditos de modelos genômicos em relação ao modelo baseado tradicionalmente em pedigree (ABLUP) para o incremento médio anual em volume (IMAvol) em indivíduos de Eucalyptus. Em (a) apresenta-se a comparação entre o modelo genômico aditivo (GBLUP) e o modelo por pedigree (ABLUP). Em (b) apresenta-se a comparação entre o modelo genômico aditivo-dominante (GDBLUP) e o modelo por pedigree (ABLUP). As cores indicam os indivíduos selecionados exclusivamente por um modelo, por ambos simultaneamente ou não selecionados (em cinza).

Conforme apresentado na Tabela 4 elevada correlação observada entre GBLUP e GDBLUP sugere que ambos os modelos tendem a selecionar conjuntos semelhantes de

indivíduos superiores, uma vez que compartilham a mesma base genômica para estimativa dos relacionamentos genéticos. As menores correlações observadas entre ABLUP e os modelos genômicos indicam diferenças na classificação dos indivíduos quando apenas informações de *pedigree* são utilizadas.

Esses resultados reforçam a capacidade das matrizes de relacionamento genômico em representar de forma mais precisa a similaridade genética entre os indivíduos avaliados, influenciando diretamente a classificação dos materiais e a tomada de decisão nos programas de melhoramento.

4.4 Capacidade preditiva dos modelos

Após a comparação entre os modelos avaliados, o modelo GDBLUP foi selecionado para a etapa de validação cruzada em função do melhor ajuste aos dados, evidenciado pelos menores valores de AIC observados. A capacidade preditiva foi estimada por meio de validação cruzada do tipo 5-fold, utilizando a correlação entre os valores genéticos preditos e os valores observados para cada característica.

Tabela 5 – Capacidade preditiva do modelo GDBLUP, estimada por meio da correlação de Pearson (r) entre os valores genéticos preditos e os valores fenotípicos observados, obtida por validação cruzada do tipo 5-fold.

Característica	Capacidade preditiva
DAP	0,553
IMAvol	0,544
Altura	0,525

Esses resultados demonstram que o modelo foi capaz de capturar uma parcela importante da variabilidade genética associada às características avaliadas. Capacidades preditivas dessa magnitude indicam potencial para aplicação da seleção genômica na identificação de indivíduos superiores para os programas de melhoramento florestal.

Além disso, a utilização de modelos genômicos com boas capacidades preditivas possibilita a realização de inferências sobre o potencial genético dos indivíduos de forma precoce, possibilitando acelerar os ciclos de seleção e obter um aumento do ganho genético em menor tempo.

5 Conclusão

Os resultados obtidos demonstraram a superioridade dos modelos baseados em relacionamento genômico em relação ao modelo baseado exclusivamente em *pedigree* para a predição genética em híbridos de *E. urograndis*. Entre os modelos avaliados, o GDBLUP apresentou o melhor ajuste aos dados, evidenciado pelos menores valores de AIC para todas as características analisadas.

A consideração simultânea dos efeitos genéticos aditivos e de dominância proporcionou melhor ajuste do modelo aos dados, indicando que a inclusão da variância de dominância pode contribuir para uma representação mais completa da variabilidade genética da população.

Dessa forma, os resultados evidenciam o potencial da seleção genômica como ferramenta para programas de melhoramento de *Eucalyptus*, permitindo maior eficiência na identificação de indivíduos superiores. A utilização de matrizes de relacionamento genômico pode contribuir para aumentar a confiabilidade das estimativas dos valores genéticos e favorecer decisões de seleção mais precisas, com potencial para acelerar o progresso genético e reduzir o ciclo de melhoramento.

Referências Bibliográficas

- AMADEU, R. R. et al. Aghmatrix: Genetic relationship matrices in r. *The R Journal*, v. 15, n. 1, p. 272–284, 2023. Citado na página 11.
- CANAL, G. B. et al. Genomic studies of the additive and dominant genetic control on production traits of euterpe edulis fruits. *Scientific Reports*, v. 13, p. 9795, 2023. Disponível em: <<https://www.nature.com/articles/s41598-023-36970-z>>. Citado 2 vezes nas páginas 9 e 16.
- GRATTAPAGLIA, D. Twelve years into genomic selection in forest trees: Climbing the slope of enlightenment of marker assisted tree breeding. *Forests*, v. 13, n. 10, p. 1554, 2022. Citado na página 8.
- Indústria Brasileira de Árvores. *Relatório Anual 2025*. Brasília, DF: [s.n.], 2025. Disponível em: <<https://iba.org/wp-content/uploads/2025/10/relatorioAnual2025.pdf>>. Citado na página 8.
- MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, v. 157, n. 4, p. 1819–1829, 2001. Disponível em: <<https://academic.oup.com/genetics/article/157/4/1819/6048353>>. Citado 2 vezes nas páginas 8 e 9.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2025. Disponível em: <<https://www.R-project.org/>>. Citado na página 13.
- RESENDE, M. D. V. d. Melhoramento de essências florestais. In: BORÉM, A. (Ed.). *Melhoramento de espécies cultivadas*. Viçosa: UFV, 1999. p. 589–647. Citado na página 8.
- RESENDE, M. D. V. d. Seleção genômica ampla (gws) e maximização da eficiência do melhoramento genético. *Pesquisa Florestal Brasileira*, n. 56, p. 63–77, 2008. Citado na página 9.
- RESENDE, M. D. V. d.; ALVES, R. S. Genética: estratégias de melhoramento e métodos de seleção. In: OLIVEIRA, E. B. d.; JUNIOR, J. E. P. (Ed.). *O eucalipto e a Embrapa: quatro décadas de pesquisa e desenvolvimento*. Brasília, DF: Embrapa, 2021. p. 171–202. Citado na página 8.
- SIMIQUÉLI, G. F. et al. Realized genomic selection across generations in a reciprocal recurrent selection breeding program of eucalyptus hybrids. *Frontiers in Plant Science*, v. 14, p. 1252504, 2023. Citado na página 8.
- TAN, B. et al. Evaluating the accuracy of genomic prediction of growth and wood traits in two eucalyptus species and their f1 hybrids. *BMC Plant Biology*, v. 17, n. 1, p. 110, 2017. Disponível em: <<https://link.springer.com/article/10.1186/s12870-017-1059-6>>. Citado na página 16.
- VANRADEN, P. M. Efficient methods to compute genomic predictions. *Journal of Dairy Science*, v. 91, n. 11, p. 4414–4423, 2008. Citado na página 13.

VITEZICA, Z. G.; VARONA, L.; LEGARRA, A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics*, v. 195, n. 4, p. 1223–1230, 2013. Citado na página 13.

VSN International. *ASReml-R Reference Manual*. [S.l.], 2023. Citado na página 13.