

GABI NUNES SILVA

**REDES NEURAIAS ARTIFICIAIS: NOVO PARADIGMA PARA A
PREDIÇÃO DE VALORES GENÉTICOS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2014

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

S586r
2014

Silva, Gabi Nunes, 1986-

Redes neurais artificiais : novo paradigma para a predição
de valores genéticos / Gabi Nunes Silva. – Viçosa, MG, 2014.
xii,105f. : il. ; 29 cm.

Inclui anexos.

Orientador: Cosme Damião Cruz.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Engenharia genética. 2. Redes neurais (Computação).
3. Inteligência artificial. 4. Simulação por computador.

I. Universidade Federal de Viçosa. Departamento de Estatística.
Programa de Pós-graduação em Estatística Aplicada e Biometria.

II. Título.

CDD 22. ed. 660.65

GABI NUNES SILVA

**REDES NEURAIIS ARTIFICIAIS: NOVO PARADIGMA PARA A
PREDIÇÃO DE VALORES GENÉTICOS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de fevereiro de 2014

Moysés Nascimento
(Coorientador)

Rafael Simões Tomaz

Cosme Damião Cruz
(Orientador)

À DEUS,

Aos meus pais Maria de Fátima Nunes da Silva e Edinilson Santos Silva, à minha irmã Bianca Nunes Silva, família e amigos, que contribuíram incondicionalmente para que mais esta conquista fosse alcançada.

DEDICO.

“A gente pode morar numa casa mais ou menos, numa rua mais ou menos, numa cidade mais ou menos, e até ter um governo mais ou menos”. A gente pode dormir numa cama mais ou menos, comer um feijão mais ou menos, ter um transporte mais ou menos, e até ser obrigado a acreditar mais ou menos no futuro. A gente pode olhar em volta e sentir que tudo está mais ou menos... tudo bem. O que a gente não pode mesmo, nunca, de jeito nenhum... é amar mais ou menos, sonhar mais ou menos, ser amigo mais ou menos, namorar mais ou menos, ter fé mais ou menos, e acreditar mais ou menos.

“Senão a gente corre o risco de se tornar uma pessoa mais ou menos.”

Chico Xavier

“As pessoas são exatamente aquilo que pensam. Tudo o que são surge com seus pensamentos e com esses pensamentos, fazem o seu mundo. Também ao receberem e darem os seus pensamentos, as pessoas se comunicam entre si como nos beijos e abraços; portanto quem recolhe um pensamento não recebe alguma coisa, mas sim recebe alguém.”

Buda

“Os ventos que às vezes tiram algo que amamos, são os mesmos que trazem algo que aprendemos a amar... Por isso não devemos chorar pelo que nos foi tirado e sim, aprender a amar o que nos foi dado. Pois tudo aquilo que é realmente nosso, nunca se vai para sempre...”

Bob Marley

“Quem tem um amigo, mesmo que um só, não importa onde se encontre, jamais sofrerá de solidão; poderá morrer de saudades, mas não estará só.”

Amyr Klink

AGRADECIMENTOS

Agradeço a Deus por me amparar durante toda a minha vida. Obrigada por suas bênçãos, e principalmente, obrigada por ter me agraciado com uma família e amigos mais que especiais.

Aos meus pais, Fátima e Edilson, pelo amor incondicional, dedicação e confiança. Vocês são minha fortaleza e meu refúgio. Sem vocês a realização de mais esse sonho não seria possível.

À Bianca (Bica), que além de irmã é minha melhor amiga. Obrigada pelos conselhos, pelo colo, amizade e paciência.

Aos meus familiares, que mesmo longe, sempre apoiaram minhas escolhas e sonhos.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, pela oportunidade de cursar um mestrado de excelência.

Agradeço aos professores do Departamento de Estatística, pelo saber transmitido e por me ajudarem no meu crescimento pessoal e profissional.

Ao professor Leonardo Lopes Bhering, por ter aceitado o convite para coorientação.

Um especial agradecimento ao professor Cosme Damião, que além das grandes contribuições como orientador, me ensinou muito sobre cooperação, trabalho em grupo e amizade. Você é um exemplo para todos nós.

Ao professor e coorientador Moysés, pelas conversas e ensinamentos. Estou certa de que ganhei um grande amigo.

Aos amigos de Viçosa, amigos do muay thai, futebol, catequese e cursinho. Agradeço à República Sereníssimas (minha segunda casa em Viçosa) e aos meninos da Tratorando (pelos melhores churrascos e festas). Enfim, a todos os amigos pelo carinho e amizade.

Aos amigos do mestrado, Édimo, Eliângela, Fanni, Pâmela e Lucas, por enfrentarmos juntos todas as dificuldades ao longo do curso. Em especial agradeço à Laís, sua amizade foi um presente abençoado por Deus. Obrigada por ser minha segunda irmã aqui em Viçosa.

Aos amigos distantes, pelas ligações e palavras de carinho e incentivo. Pati, obrigada pela amizade que começou no pré-escolar e que não acaba nunca. Carla, sua amizade é uma herança que quero levar por toda a vida. Victor, obrigada pela sua

amizade e sinceridade, você é muito especial. Mayara, obrigada pela honra de ser sua amiga.

A todos os colegas e amigos do laboratório de Bioinformática, em especial, Haroldo, Marciane, João Filipi, Vinícius e Isabela, pelo apoio, incentivo, brincadeiras e zuações, e ao Rafael, pelas enormes contribuições para a realização deste trabalho.

À CAPES, pelo suporte financeiro para o desenvolvimento deste trabalho.

Enfim, a todos que direta ou indiretamente colaboraram de alguma forma, fosse com um abraço ou com uma palavra de incentivo, para o sucesso deste trabalho.

MEU MUITO OBRIGADA!

BIOGRAFIA

GABI NUNES SILVA, filha de Edinilson Santos Silva e Maria de Fátima Nunes da Silva, nascida em 5 de novembro de 1986, em Ipatinga, no estado de Minas Gerais.

Cursou o ensino fundamental e médio na rede pública na cidade de Ipatinga/MG. Em 2005 iniciou o curso de Licenciatura em Matemática, na Universidade Federal de Viçosa – UFV, graduando-se em julho de 2011.

Em fevereiro de 2012 iniciou o Mestrado *Stricto Sensu* no Programa de Pós-Graduação em Estatística Aplicada e Biometria na Universidade Federal de Viçosa, submetendo-se à defesa de dissertação em 27 de fevereiro de 2014.

SUMÁRIO

RESUMO.....	ix
ABSTRACT.....	xi
INTRODUÇÃO GERAL.....	1
REVISÃO DE LITERATURA.....	3
1. Redes Neurais Artificiais.....	3
1.1. Introdução.....	3
1.2. Histórico.....	4
1.3. Neurônios biológicos e neurônios artificiais.....	5
1.4. Funções de ativação.....	7
1.4.1. Função Linear (purelin).....	7
1.4.2. Função Sigmoidal.....	8
1.4.3. Função tangente hiperbólica.....	9
1.5. Arquitetura da Rede.....	10
1.6. O Perceptron Múltiplas Camadas (Multilayer Perceptron).....	10
2. Predição de Valores Genéticos.....	12
2.1. Introdução.....	12
2.2. Modelo para estudo de caracteres quantitativos.....	13
2.2.1. Estimativas da média e da variância fenotípica.....	14
2.2.2. Estimativas dos parâmetros genéticos.....	15
2.3. Uso da experimentação para predição de valores genéticos.....	16
2.3.1. Método de seleção massal.....	19
2.3.2. Método de seleção de progênies.....	20
2.3.3. Teoria de índice de seleção no melhoramento genético.....	21
2.4. Uso da abordagem de modelos mistos na predição de valores genéticos..	22
3. Simulação de dados.....	23
4. Utilização das redes neurais na predição de valores genéticos.....	25
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	25
CAPÍTULO 1.....	32
RESUMO.....	33
INTRODUÇÃO.....	33
MATERIAL E MÉTODOS.....	35
1. Simulação de valores genéticos.....	35

1.1.	Simulação do efeito dos blocos.....	35
1.2.	Simulação do efeito dos genótipos.....	36
1.3.	Simulação dos erros aleatórios.....	36
1.4.	Estabelecimento dos valores fenotípicos.....	37
2.	Ampliação dos dados.....	37
3.	Proposta para estabelecimento do critério de seleção da rede neural artificial.....	38
3.1.1.	Modelo convencional.....	38
3.1.2.	Modelo proposto – Adaptação de MCP.....	39
3.1.3.	Modelo proposto – generalização de MLP.....	40
4.	Construção e avaliação da Rede Neural Artificial.....	40
	RESULTADOS E DISCUSSÃO.....	42
1.	Análise dos dados do arquivo original (teste) e do conjunto de dados ampliados para treinamento da RNA.....	42
2.	Eficiência preditiva e desempenho da Rede Neural Artificial.....	49
	CONCLUSÕES.....	53
	REFERÊNCIAS BIBLIOGRÁFICAS.....	53
	CAPÍTULO 2.....	56
	ABSTRACT.....	57
	RESUMO.....	57
	INTRODUCTION.....	58
	MATERIAL AND METHODS.....	59
	Data simulation.....	59
	Artificial neural network construction and evaluation.....	60
	RESULTS AND DISCUSSION.....	62
	CONCLUSION.....	66
	ACKNOWLEDGEMENTS.....	67
	REFERENCES.....	67
	ANEXOS.....	69

RESUMO

SILVA, Gabi Nunes, M.Sc., Universidade Federal de Viçosa, fevereiro de 2013, **Redes Neurais Artificiais: novo paradigma para a predição de valores genéticos**. Orientador: Cosme Damião Cruz; Co-orientadores: Leonardo Lopes Bhering e Moysés Nascimento.

Para que programas de melhoramento genético tenham êxito é importante levar em consideração alguns fatores, como a variabilidade da população, as condições experimentais, a instalação e condução de ensaios e, fundamentalmente, a acurácia da predição do valor genético a partir de valores fenotípicos. Até então, as formas de aumentar o ganho por seleção tem sido apontadas pelas diferentes estratégias preconizadas nos métodos de melhoramento, ou pela utilização dos princípios básicos da experimentação, ou ainda recorrer a modelos biométricos que buscam parametrizar as influências tanto genóticas quanto ambientais. Assim, no que se refere aos métodos de genética e estatística utilizados para seleção de genótipos superiores, destacam-se os métodos derivados da teoria de seleção direta, seleção indireta e por índice de seleção, o método de seleção combinada e o método REML-BLUP, dentre outros. Apesar das diversas metodologias disponíveis para a seleção de genótipos superiores, ainda é comum e necessário praticar seleção de indivíduos em relação a características com baixa confiabilidade de predição do valor genotípico a partir de um valor fenotípico dado pela média fenotípica corrigida ou ajustada em função da agregação de informações de parentes, de caracteres correlacionados ou de práticas de redução do efeito ambiental. No final, considera-se esta média fenotípica ajustada como a medida mais apropriada para indicar a superioridade genética e predizer o ganho genético. No entanto, tais modelos ou procedimentos não contemplam uma infinidade de outras informações estatísticas de grande relevância, diferentes da média fenotípica que é usualmente adotada, mas que agregam informações importantes acerca do genótipo avaliado e que têm sido deixadas à margem dos estudos envolvendo melhoramento genético e critérios de seleção. Neste contexto, as redes neurais artificiais constituem novo paradigma que tem sido empregado, ainda que de forma tênue, nos programas de melhoramento genético animal e vegetal. Essa abordagem, diferentemente das modelagens estocásticas utilizadas até então, é baseada nos princípios de aprendizado e de inteligência computacional de um conjunto amplo de informação do desempenho do genótipo envolvendo médias, máximos, mínimos, variância e toda ordem de informação possível de ser direta ou indiretamente mensurada. Assim, ao contrário dos métodos

estatísticos que resumem as informações ou realizam a simplificação estrutural dos dados, as redes neurais, à semelhança do cérebro humano, captam toda informação disponível para gerar um critério de tomada de decisão. Assim, este trabalho foi realizado com o intuito de utilizar as redes neurais para melhorar a acurácia na predição de valores e ganhos genéticos, através de uma discussão de seus fundamentos teóricos e utilização de dados simulados, com mesma caracterização em termos de média, herdabilidade e coeficiente de variação dos dados reais, fornecendo um método alternativo para identificação de genótipos superiores.

ABSTRACT

SILVA, Gabi Nunes, M.Sc., Universidade Federal de Viçosa, february, 2013. **Artificial Neural Networks: a new paradigm for predicting genetic values.** Adviser: Cosme Damião Cruz; Co-advisers: Leonardo Lopes Bhering and Moysés Nascimento.

For breeding programs are successful is important to take into account factors such as the variability of the population, the experimental conditions, the installation and conduct of trials and, crucially, the accuracy of prediction of genetic value from phenotypic values. Until then, how to increase the gain by selection has been pointed out by different strategies recommended in breeding methods, or by using the basic principles of experimentation, or even use biometric models that seek to parameterize both genotypic and environmental influences. Thus, as regards the methods of genetics and statistics used for selection of superior genotypes, highlight the methods derived from the theory of direct selection, indirect selection and index selection, combined selection method and REML-BLUP method among others. Despite several methodologies available for the selection of superior genotypes is still common and necessary practice selection of individuals in relation to features with low reliability of predicting genotype value from a phenotypic value given by the corrected or adjusted by the aggregate average phenotypic information from relatives of correlated traits or practices to reduce the environmental effect. In the end, we consider this phenotypic average adjusted as the most appropriate measure to indicate genetic superiority and predict genetic gain. However, such models or procedures do not include a multitude of other statistical information of great relevance, different phenotypic average which is usually adopted, but which provide important information about the reported genotype and who have been sidelined studies involving genetic improvement and selection criteria. In this context, artificial neural networks are a new paradigm that has been employed, albeit tenuously, in animal and plant breeding programs. This approach, unlike the stochastic modeling used so far, is based on principles of learning and computational intelligence on a wide range of performance information of genotype involving averages, maximum, minimum, variance and all sorts of information can be directly or indirectly measured. Thus, unlike the statistical methods that summarize the information or perform the structural simplification of data, neural networks, like the human brain, capture all available information to generate a criterion for decision making. This work was carried out with the intention of using neural networks to improve the accuracy of the predicted values and genetic gains, through a discussion of

its theoretical foundations and use simulated data with the same characteristics in terms of average heritability and coefficient of variation of actual data by providing an alternative method for identification of superior genotypes.

INTRODUÇÃO GERAL

Um dos principais objetivos nos programas de melhoramento vegetal é o aumento da produtividade agrícola associado à melhoria na qualidade nutricional da planta em estudo. Para que esse objetivo seja alcançado, têm sido empregadas estratégias visando a melhoria das condições ambientais e esforços no sentido de identificar o real potencial genético dos indivíduos ou populações, para assim aprimorá-los.

De acordo com Cruz (2010), quando se objetiva melhorar o potencial genético da planta ou do animal, deve-se levar em conta a grande diversidade genética existente e recorrer então a processos de seleção e recombinação das formas genéticas mais adaptadas, de melhor qualidade e mais eficientes no aproveitamento de recursos. Para a identificação de genótipos superiores, é necessária a utilização de métodos de seleção capazes de explorar, de forma eficiente, o material genético disponível, maximizando o ganho genético em relação às características de interesse (ODA *et al.*, 2007). Cruz (2012) afirma que a predição de ganhos realizada por meio de uma estratégia de seleção possibilita orientar, de forma mais efetiva, os programas de melhoramento genético e prever o sucesso do esquema seletivo adotado.

A grande problemática encontrada nos programas de melhoramento genético é obter informações sobre o valor genético do indivíduo, para que se possa praticar a seleção com a máxima acurácia possível (PEIXOTO, 2013). Apesar da grande quantidade de métodos de predição descritos na literatura, é prática comum prever a média genotípica a partir da média fenotípica, agregando valores de outras características correlacionadas ou informações de aparentados.

Para obtenção de informações precisas sobre o comportamento de um dado genótipo, é possível utilizar as estratégias preconizadas nos métodos de melhoramento, utilizar os princípios da experimentação e, ou, recorrer a modelos biométricos que buscam parametrizar as influências tanto genotípicas quanto ambientais. Neste caso, quando o melhorista se depara com análises experimentais complexas, se torna necessária a utilização de um número maior de parâmetros. No entanto, o aumento da complexidade da análise dificulta o processo de tomada de decisão e, conseqüentemente, aumenta a responsabilidade e as incertezas do melhorista na hora de recomendar determinado material genético.

No que se refere aos métodos de genética e estatística utilizados para seleção de genótipos superiores, destacam-se os métodos derivados da teoria de seleção direta,

seleção indireta (PAULA *et al.*, 2002) e por índice de seleção (SMITH, 1936; HAZEL, 1943; MULAMBA & MOCK, 1978; COTTERILL, 1986), o método de seleção combinada (MARTINS *et al.*, 2005) e o método REML-BLUP (GARCIA & NOGUEIRA, 2005). No processo de seleção direta são estabelecidos os materiais genotípicos de teste e as unidades melhoradas resultantes da recombinação de unidades aparentadas com as unidades de teste. Métodos baseados nessa teoria estimam o ganho de seleção por meio dos princípios da regressão linear (CRUZ, 2012). Martins *et al.* (2003) afirmam que, por vezes, a seleção praticada em determinada característica provoca respostas correlacionadas em outras, e, no método de seleção indireta, o melhorista avalia a magnitude dessa resposta correlacionada. A seleção combinada, por sua vez, é aquela em que se identificam genótipos superiores a partir da informação do indivíduo e de seus aparentados (CRUZ & CARNEIRO, 2006).

A grande evolução dos métodos de seleção se deu com a apresentação da metodologia de modelos mistos para obtenção da melhor predição linear não viciada (*best linear unbiased prediction* – BLUP) dos valores genéticos, proposta por Henderson (1973). Segundo Resende (2007), o BLUP é um procedimento superior a outros índices de seleção por se tratar de uma metodologia que maximiza a acurácia seletiva.

Diante das diversas metodologias disponíveis para a seleção de genótipos superiores, uma prática comum é encontrar experimentos nos quais caracteres de importância são avaliados com baixa confiabilidade de predição do valor genotípico, o que força o melhorista a utilizar a média fenotípica das repetições como medida mais apropriada para indicar a superioridade genética e predizer o ganho genético (PETEK *et al.*, 2008). No entanto, se o pesquisador recorrer à literatura, este poderá encontrar uma infinidade de outras informações estatísticas de grande relevância, diferentes da média fenotípica que é usualmente adotada, mas que agregam informações importantes acerca do genótipo avaliado e que têm sido deixadas à margem dos estudos envolvendo melhoramento genético e critérios de seleção.

Neste contexto, as redes neurais artificiais constituem novo paradigma que tem sido empregado, ainda que de forma tênue, nos programas de melhoramento genético animal e vegetal. Essa abordagem, diferentemente das modelagens estocásticas utilizadas até então, é baseada nos princípios de aprendizado e de inteligência computacional. Diante deste cenário, no qual há ainda poucos trabalhos associados à seleção de genótipos superiores utilizando redes neurais, propõe-se neste estudo, uma nova abordagem metodológica por meio da qual utilizamos uma gama de informações

estatísticas para predição do valor genético. Neste estudo preconizamos que as redes neurais, em sua fase de treinamento, seriam capazes de captar informações da avaliação do genótipo que não são contempladas nos procedimentos estocásticos convencionais. Assim, o julgamento da superioridade de um genótipo em relação ao outro poderia ser decidido por meio das redes neurais, considerando como informações de entrada da rede a média, a variância, desvio-padrão, máximo, mínimo dentre outras informações pertinentes.

REVISÃO DE LITERATURA

1. Redes Neurais Artificiais

1.1. Introdução

Uma área de pesquisa que desperta grande interesse é a simulação das capacidades cognitivas de um ser humano devido ao fato de que a inteligência humana é uma das mais avançadas do universo biológico. O cérebro humano é composto por aproximadamente 100 bilhões de neurônios, conectados a aproximadamente 10 mil outros e todas as funções exercidas pelo corpo de um ser humano são controladas diretamente por essas células neurais (SILVA *et al.*, 2010). Ele é responsável por funções cognitivas básicas, associado ainda à sua capacidade de reconhecer padrões e relacioná-los, além de utilizar e armazenar conhecimento obtido por experiência vivida (BRAGA *et al.*, 2007). A estrutura desses neurônios, a topologia de suas conexões e o comportamento conjunto desses elementos constituem a base para o estudo das Redes Neurais Artificiais (RNA's).

As Redes Neurais Artificiais são modelos de processamento de dados que emulam uma rede de neurônios biológicos, capazes de recuperar rapidamente uma grande quantidade de dados e reconhecer padrões baseados na experiência, ou seja, tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento funcional e sua dinâmica (HAYKIN, 2001). De acordo com Braga *et al.* (2007), por não se basear em regras, essa abordagem surge como uma alternativa à metodologia de computação algorítmica convencional.

Braga *et al.* (2007) relatam que estudos envolvendo as Redes Neurais Artificiais despertam cada vez mais o interesse de pesquisadores devido ao fato de que a

solução de problemas apresentada internamente pela rede e o paralelismo natural inerente à arquitetura das RNA's criam a possibilidade de um desempenho superior aos modelos convencionais. Além disso, a capacidade de aprender a partir de exemplos representa um avanço importantíssimo para a utilização da bioinformática em diversas áreas de pesquisa.

A abordagem baseada em Redes Neurais Artificiais tem sido utilizada por alguns pesquisadores e sido descrita como uma ferramenta adicional no processo de tomada de decisão no melhoramento genético animal e vegetal (VENTURA *et al.*, 2012; NASCIMENTO *et al.*, 2013). Segundo Gianola *et al.* (2011), as RNA's oferecem uma alternativa interessante porque essas máquinas de aprendizagem podem atuar como aproximadores universais de funções complexas. Dessa forma, a realização de análises por meio de métodos computacionais que sejam capazes de "aprender" representa um grande avanço para os estudos envolvendo procedimentos estatísticos e para o melhoramento genético. Segundo Bishop (2007), as RNA's são capazes de captar as relações não lineares entre variáveis preditoras e respostas, e ainda podem aprender sobre formas funcionais de forma adaptativa, pois uma série de transformações, chamada genericamente de função de ativação, é conduzida por parâmetros.

Redes Neurais Artificiais são algoritmos de otimização e aprendizagem com o funcionamento inspirado na natureza do cérebro. O primeiro passo da Rede Neural é a etapa de formação, em que um conjunto de dados é utilizado para aprender pela rede (SILVA *et al.*, 2010). Historicamente, para tal fim, o analista tem utilizado um subconjunto de dados, tais como informações históricas dos genótipos, mas que têm a desvantagem de não serem necessariamente contemporâneos do experimento avaliado. Uma alternativa seria utilizar um conjunto de dados virtuais, ou simulados, proporcionados por técnicas estatísticas que sejam capazes de preservar algumas características ou estrutura do experimento original, tais como de média, variâncias e covariâncias.

1.2. Histórico

O primeiro modelo artificial de um neurônio biológico, chamado neurônio artificial, foi criado em 1943 por Warren McCulloch e Walter Pitts. McCulloch, renomado psicólogo e neuro-fisiologista, dedicou grande parte da sua carreira à tentativa de representar e modelar eventos do sistema nervoso e, em 1942, Pitts, matemático recém-graduado na época, juntou-se a ele (BRAGA *et al.*, 2007).

Em 1943, McCulloch e Pitts publicaram o trabalho, “*A Logical Calculus of the ideas Immanent in Nervous Activity*”, que consistiu em um trabalho com fundamentação teórica, cujo principal objetivo foi descrever o modelo de um neurônio artificial, chamado de neurônio MCP, em homenagem a McCulloch e Pitts. O primeiro trabalho científico que relatou o processo de aprendizado das Redes foi apresentado em 1949 por Donald Hebb.

Em 1959, Frank Rosenblatt construiu o primeiro modelo de Rede Neural, que ficou conhecido como *perceptron* simples, pois continha uma única camada. A partir desse modelo, Rosenblatt mostrou que, se fossem acrescentadas de sinapses ajustáveis, as RNA's com neurônios MCP poderiam ser treinadas para classificar certos tipos de padrões (BRAGA *et al.* 2007).

Em 1969, Minsky e Papert relataram que o *perceptron* simples de Rosenblatt se limitava à resolução de problemas linearmente separáveis. Essa crítica fez com que os estudos envolvendo redes neurais ficassem adormecidos por toda a década de 70, salvo alguns poucos pesquisadores, como Igor Aleksander que prosseguiu com os estudos sobre redes sem pesos na Inglaterra; Kunihiko Fukushima que pesquisou sobre *cognitron* e *neocognitron* no Japão; Steven Grossberg, acerca de sistemas auto-adaptativos nos Estados Unidos e Teuvo Kohonen, acerca de memórias associativas e modelos auto-organizáveis, na Finlândia (BRAGA *et al.*, 2007).

O final da década de 1980 marcou o auge dos estudos das Redes Neurais Artificiais. Em 1982, John Hopfield publicou um artigo que falava das propriedades associativas das RNA's. Depois disso, Rumelhart *et al.* (1986), descreveram o funcionamento do algoritmo de treinamento *backpropagation*. A partir de então, inúmeros estudos a respeito das Redes Neurais Artificiais começaram a ser publicados.

1.3. Neurônios biológicos e neurônios artificiais

Como relatado na literatura, a maior inspiração para os estudos envolvendo inteligência artificial fundamenta-se na ideia de ser possível simular as capacidades cognitivas de um neurônio biológico humano.

Esquemáticamente, como exemplificado na Figura 1.1, os neurônios biológicos são divididos em três partes principais: corpo celular, dendritos e axônio, cada uma com sua respectiva função. Os dendritos têm a função de receber todas as informações – chamadas de impulsos nervosos – e conduzi-las ao corpo celular, aonde esses impulsos serão processados e levados a outros neurônios. O axônio tem a função de transmitir

esses impulsos nervosos até os dendritos dos neurônios seguintes. Essa conexão entre o axônio de um neurônio e o dendrito posterior é denominada sinapse. São as sinapses que caracterizam a união funcional entre os neurônios, compondo as redes neurais biológicas.

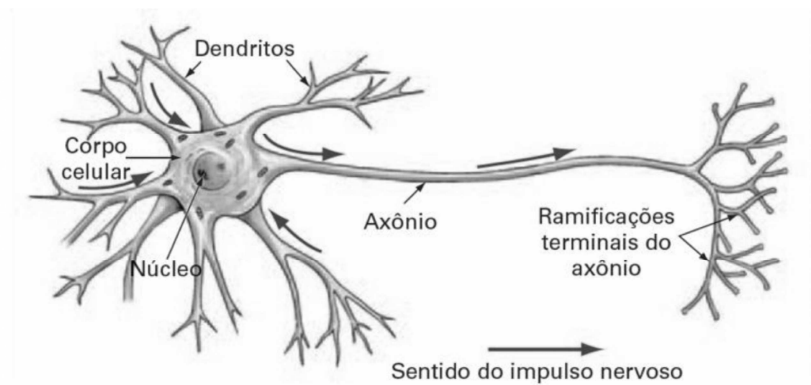


Figura 1.1. Modelo de um neurônio biológico (Fonte: <http://www.borges-solutions.com/algorithmo-genetico-para-otimizacao-de-arquiteturas-de-redes-mlp/>)

No modelo neural artificial podem ser identificados três elementos principais tais como um conjunto de sinapses, um somatório e uma função de ativação (HAYKIN, 2001). Os trabalhos com Redes Neurais Artificiais envolvem processos de treinamento, aprendizado e validação, e, além disso, a escolha de uma arquitetura apropriada que possua funções de ativação eficientes, número de camadas ocultas e número de neurônios por camadas.

O funcionamento do neurônio artificial é inspirado no funcionamento de um neurônio biológico. Para representar o comportamento das sinapses, os m terminais de entrada do neurônio artificial possuem pesos w_1, w_2, \dots, w_m , cujos valores podem ser positivos ou negativos, de acordo com o sinal sináptico correspondente – inibitório ou excitatório.

O neurônio artificial é constituído por m entradas com os pesos ponderados, um limiar de disparo conhecido como bias – representado pelo somatório na Figura 1.2 – associado à função de ativação, e finalmente, um terminal de saída y .

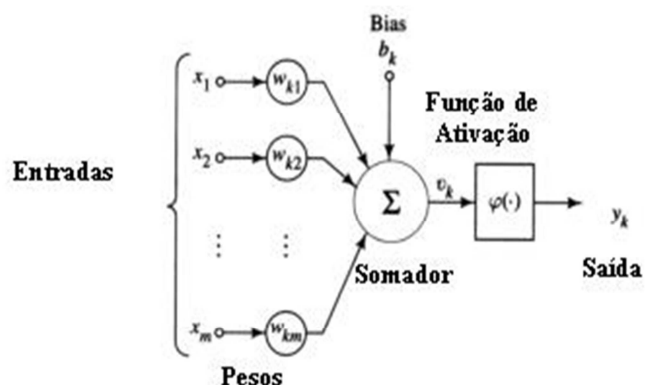


Figura 1.2. Esquema de um neurônio de McCulloch e Pitts (adaptado Haykin 2009).

1.4. Funções de ativação *

De forma geral, a operação de uma célula da rede se resume em sinais emitidos às m entradas. Cada sinal é multiplicado por um peso que indica sua influência na saída da unidade (soma ponderada). A porta de limiar funciona como um mecanismo para o disparo da solução, ou seja, compara a soma ponderada das entradas com um valor limite. Caso a soma exceda o limiar, a saída é ativada, permanecendo desativada em caso contrário (BRAGA *et al.*, 2011).

A função de ativação tem o importante papel de gerar a saída do y neurônio a partir das somas ponderadas recebidas pelo neurônio, e será escolhida de acordo com o problema a ser resolvido. As funções de ativação mais utilizadas na prática são a função linear (purelin), que é recomendada para problemas linearmente separáveis, a função sigmóide logística e a função tangente hiperbólica, dependendo das características dos dados. Outros autores, no entanto, estão recorrendo ao uso de outras funções de ativação. Em 2001, Hornik utilizou funções de ativação não-polinomiais. Singh e Chandra, em 2003, propuseram uma classe de funções sigmóides. Em 2004, Skoundrianos e Tzafestas propuseram uma nova função de ativação sigmoideal e obtiveram bons resultados para modelagem de sistemas dinâmicos de tempo discreto.

1.4.1. Função Linear (purelin)

A função de ativação linear é comumente chamada de purelin, e pode ser vista como uma aproximação de um amplificador não-linear (HAYKIN, 2001).

*As siglas utilizadas para definir as funções de ativação são as mesmas adotadas tal pelo software MATLAB.

A função purelin o gráfico tal como exemplificado na Figura 1.3 abaixo.

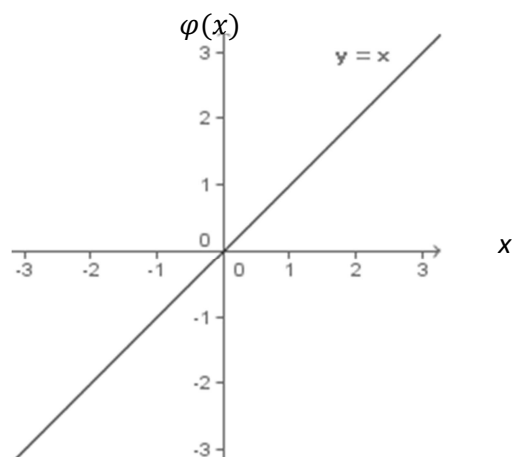


Figura 1.3. Gráfico da função purelin

Essa função é representada pela equação 1.1 abaixo:

$$\varphi(x) = x \tag{1.1}$$

1.4.2. Função Sigmoidal

A função *sigmoide logística* (logsig) é sem dúvida a função de ativação mais utilizada na construção das Redes Neurais Artificiais (HAYKIN, 2001). Por definição, a função sigmoide é monótona crescente com propriedades assintóticas e de suavidade, e como visto na Figura 1.4, apresenta gráfico na forma de “s”.

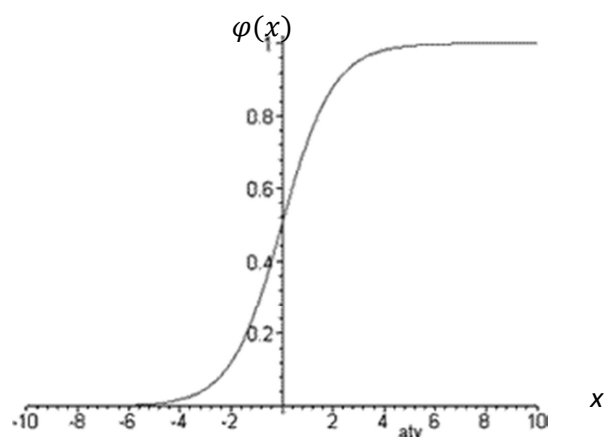


Figura 1.4. Gráfico da função sigmoide.

Um exemplo de função sigmoidal é a função logística, definida pela equação 1.2 dada abaixo:

$$\varphi(x) = \frac{1}{1 + e^{-ax}} \quad (1.2)$$

onde a é o parâmetro de inclinação da função sigmoide. A variação do valor do parâmetro a proporciona funções sigmoides com diferentes inclinações.

1.4.3. Função tangente hiperbólica

Como dito anteriormente, o parâmetro a é quem fornece a inclinação para a função sigmoidal. Quando este parâmetro se aproxima do infinito, a função sigmoide se torna uma função limiar, podendo, no entanto, assumir um intervalo contínuo de valores entre 0 e 1, ao contrário da função limiar que somente assume valor 0 ou 1.

Segundo Haykin (2001), em algumas situações se deseja uma função de ativação que se estenda de -1 a +1, assumindo uma forma anti-simétrica em relação à origem. Neste caso, utiliza-se uma forma correspondente à logsig, denominada de função tangente hiperbólica (tansig) – Figura 1.5 – definida pela equação 1.3:

$$\varphi(x) = \tanh\left(\frac{x}{2}\right) = \frac{1 - e^{-x}}{1 + e^x} \quad (1.3)$$

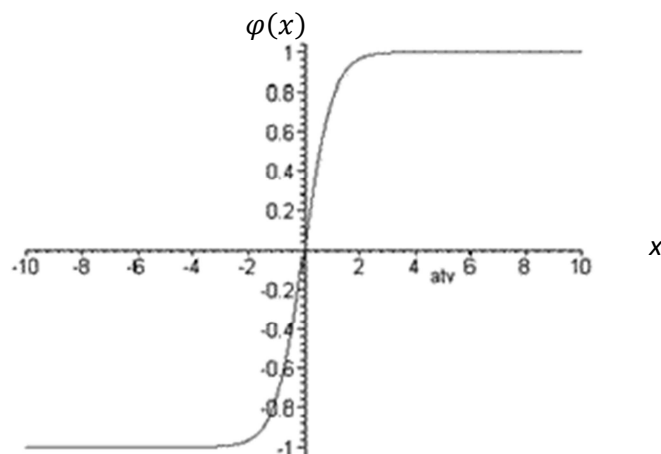


Figura 1.5. Gráfico da função tangente hiperbólica.

1.5. Arquitetura da Rede

A estruturação dos neurônios de uma rede neural está diretamente ligada ao algoritmo de aprendizado adotado para treinar a rede. Um conjunto de neurônios artificiais conectados é capaz de resolver problemas de complexidade elevada (BRAGA *et al.*, 2007).

No contexto mais geral, podem-se considerar duas classes principais de arquiteturas de rede: redes alimentadas por uma única camada de neurônios e redes alimentadas diretamente com múltiplas camadas (HAYKIN, 2001). A rede de camada única apresenta uma estrutura mais simples, na qual se tem uma camada de entrada e uma camada de saída de neurônios que é alimentada para frente – procedimento chamado de *feedforward*.

Segundo Hecht & Nielsen (1989), com apenas uma camada intermediária na rede neural já é possível se calcular uma função arbitrária qualquer a partir de dados fornecidos. Para esses autores, a camada oculta deve ter por volta de $(2i+1)$ neurônios, onde i é o número de variáveis de entrada, no entanto, Haykin (2001) afirma que esse tipo de estruturação é capaz de resolver problemas multivariáveis, mas possui algumas restrições no que se refere a problemas complexos.

Como alternativa para solucionar uma série de problemas de maior complexidade, propõe-se o uso das redes com uma ou mais camadas ocultas – *Multilayer Perceptron* (MLP), nas quais os nós computacionais são chamados de neurônios ocultos ou unidades ocultas (HAYKIN, 2001).

1.6. O Perceptron Múltiplas Camadas (Multilayer Perceptron)

O *Multilayer Perceptron* ou *MLP*, como ilustrado na Figura 1.6, de acordo com Haykin (2001), é uma extensão do perceptron simples, capaz de trabalhar com problemas não linearmente separáveis, avanço este que foi possível graças à inserção das camadas ocultas no modelo.

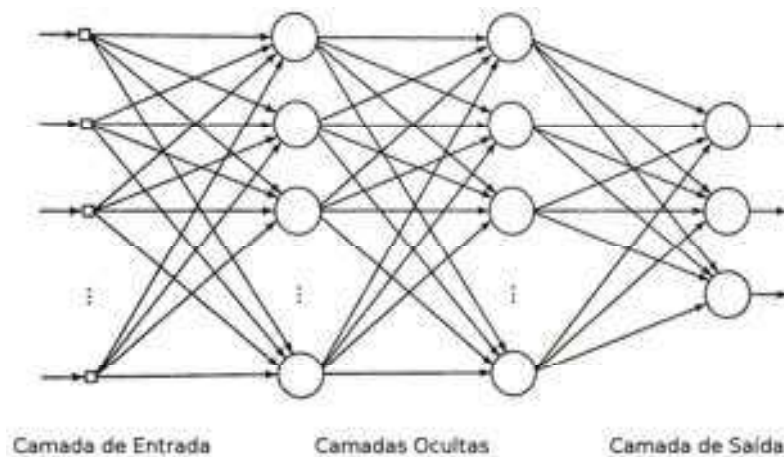


Figura 1.6. Esquema de uma rede MLP com uma camada de entrada, duas camadas intermediárias (camadas ocultas) e uma camada de saída.

Os *perceptrons* de múltiplas camadas têm sido aplicados com sucesso para resolver diversos problemas difíceis, por meio do seu treinamento de forma supervisionada com um algoritmo, conhecido como “algoritmo de retropropagação de erro” (*error backpropagation*) (HAYKIN, 2001).

As principais características de uma rede *MLP* são:

- i. O modelo utiliza uma função de ativação não linear;
- ii. A rede contém uma ou mais camadas de neurônios intermediários (ocultos), que não fazem parte da entrada ou da saída da rede;
- iii. A rede possui um alto grau de conectividade, determinado pelas sinapses da rede.

A atuação conjunta dessas características com a habilidade de aprender por meio da experiência do processo de treinamento é responsável pelo poderio computacional da *MLP* (HAYKIN, 2001).

Diante dos conceitos apresentados, importância, estrutura e funcionamento das Redes Neurais Artificiais, assim como a possibilidade do desenvolvimento de estudos utilizando essa linha de pesquisa aplicada ao melhoramento genético, os tópicos que se seguem referem-se à modelos de predição de valores genéticos e técnicas de melhoramento, cuja utilização potencial das RNA's são imensas.

2. Predição de Valores Genéticos

2.1. Introdução

A genética quantitativa, área da genética que estuda a herança e a variação dos caracteres quantitativos, constitui uma ciência de grande importância e aplicação para o melhoramento genético de plantas e animais.

Características quantitativas ou de herança complexa, apresentam várias peculiaridades, quando comparadas com as características de herança simples ou qualitativas. Geneticamente, os caracteres quantitativos são controlados por um número elevado de genes, os quais "sofrem" grande influência ambiental, enquanto os caracteres qualitativos são controlados por um ou poucos genes e têm pouca ou nenhuma influência do ambiente (FALCONER & MACKAY, 1996). Dessa forma, é mais evidente que nos caracteres quantitativos, o fenótipo individual não constitui um bom preditor para o seu genótipo dos indivíduos. Por este motivo, conjuntos populacionais devem ser considerados, tomando ainda valores médios e medidas apropriadas da variação da característica de interesse. Com relação à variação da característica, os fenótipos dos indivíduos apresentam-se em uma distribuição contínua, em contraste com os fenótipos da característica qualitativa os quais são claramente distinguíveis uns dos outros (classes discretas de fenótipos), podendo assim ser facilmente classificados (FALCONER, 1987).

De forma semelhante, também diferem os métodos de estudo genético. Para o estudo de caracteres de heranças quantitativas, como ressaltado, as informações com base somente no indivíduo têm pouco valor – devido à influência do ambiente –, o que obriga o pesquisador a estender seus estudos a grupos maiores de indivíduos (FALCONER, 1987), adotando-se ainda, um modelo biométrico. Nestes casos, a média de um conjunto de indivíduos constitui parâmetro muito mais confiável para inferir sobre o valor genotípico. Ainda, estimativas de variância fenotípica, genética, e finalmente na fração herdável desta última, tem papel fundamental no estudo da herança quantitativa. Para este estudo, o pesquisador encontra amplo embasamento teórico nas diversas técnicas da Estatística, ciência que segundo Oliveira (1999) e Freund (2006), utilizam as teorias probabilísticas para explicar a frequência de ocorrência de eventos e que ainda envolve um conjunto de técnicas que permitem, de forma sistemática, organizar, descrever, analisar e interpretar dados oriundos de estudos ou experimentos. Quanto aos métodos de estudo das características de herança simples, estes se baseiam

em métodos que avaliam a contagem de indivíduos das diversas classes, não sendo considerados parâmetros de médias e de variância. No estudo deste tipo de característica, o teste de qui-quadrado é comumente utilizado para comprovar a veracidade de hipóteses relativas ao controle gênico do caráter.

2.2. Modelo para estudo de caracteres quantitativos

Estudos genéticos envolvendo caracteres quantitativos adotam o modelo básico descrito na equação (2.1) a seguir:

$$F = G + E \quad (2.1)$$

Em que,

F: é o valor fenotípico, medido dos indivíduos;

G: é a ação do genótipo, ou valor genotípico;

E: desvio causado pela ambiente.

Entende-se como valor fenotípico o valor observado quando uma característica quantitativa é mensurada em um indivíduo, e por esse motivo, observações como média e variância devem ser baseadas nessa medida (FALCONER, 1987).

Como visto, o valor fenotípico do indivíduo pode ser decomposto em componentes devido ao genótipo e na fração influenciada pelo ambiente. O genótipo é o conjunto particular de genes do indivíduo e o ambiente caracteriza toda a influência não genética sobre o fenótipo e, por definição, somente estes dois componentes determinam o valor fenotípico do indivíduo (FALCONER, 1987).

De acordo com Falconer (1987), se o pesquisador ou melhorista pudesse duplicar um particular genótipo em um número de indivíduos e medi-los, sob condições normais de ambiente para a população, o desvio médio ocasionado pelo ambiente, seria zero e seu valor fenotípico médio, por conseguinte, igual ao valor genotípico desse particular genótipo. Esse é o significado do valor genotípico de um indivíduo.

Valor genotípico pode ser definido como sendo a constituição genética do indivíduo e, apesar de não poder ser mensurado diretamente, é de grande importância em processos de seleção e identificação de superioridade genética (CRUZ, 2010). Esse valor genotípico pode ser desdobrado em duas partes. Uma refere-se à sua fração herdável, também denominada valor genético aditivo, e a outra que se refere à uma

fração não herdável por processos sexuais, correspondente aos valores atribuídos aos desvios de dominância (CRUZ, 2010).

Uma vez mensurados os valores fenotípicos, torna-se possível estimar a variância fenotípica (σ_F^2) e, sob determinadas condições, desdobrá-la nos componentes de variância genética ou genotípica (σ_G^2) e em variância ambiental (σ_E^2). Portanto, pela equação (2.2), tem-se:

$$\sigma_F^2 = \sigma_G^2 + \sigma_E^2 + 2cov(G, E) \quad (2.2)$$

Como a avaliação dos genótipos é feita sob os princípios básicos da experimental – repetição e casualização –, a ação do genótipo (G) e os desvios devido aos efeitos ambientais (E) atuam de forma independente, de modo que a covariância entre elas é nula.

De maneira geral, embora o interesse do melhorista seja obter o máximo de informação possível sobre o valor genético dos genótipos sem estudo, a predição destes é feita com base apenas no fenótipo (CRUZ, 2010). O usual, portanto, é fundamentar a predição da média genotípica a partir da média fenotípica e, analogamente, a variância genotípica a partir da variância fenotípica removida a fração da variação ocasionada pelo ambiente, medida por meio de observações da variação em populações homogêneas ou pela experimentação com repetições e casualizações.

2.2.1. Estimativas da média e da variância fenotípica

A média é uma das principais medidas de posição, também chamada de estatística de primeira ordem. A variância, por sua vez, caracteriza uma medida de dispersão das informações acerca dos caracteres quantitativos dos indivíduos, também chamada de estatística de segunda ordem (RESENDE, 2007).

As estimativas da média (\bar{X}) e da variância ($\hat{\sigma}^2$) fenotípica são dadas pelas equações (2.3) e (2.4), respectivamente:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (2.3)$$

$$\hat{\sigma}^2 = \frac{\left[\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n} \right]}{n-1} \quad (2.4)$$

2.2.2. Estimativas dos parâmetros genéticos

Tendo as estimativas da média e da variância, torna-se possível a obtenção de parâmetros genéticos, que são úteis para avaliar o potencial genético do genótipo para a prática do melhoramento e seleção (CRUZ, 2010). Dentre os parâmetros genéticos mais utilizados em programas de melhoramento, podem-se destacar a herdabilidade (h^2), a correlação (r) e a covariância ($cov(F, G)$).

A herdabilidade expressa a proporção da variabilidade observada devido aos efeitos aditivos dos genes, isto é, representam a proporção herdável da variabilidade total (BORÉM, 2001). Segundo Falconer (1987), a melhor estimativa do valor genético de um indivíduo candidato à seleção corresponde ao produto do seu valor fenotípico pela herdabilidade.

Essa fração da variância é expressa pela razão entre a variância genética e a variância fenotípica, de acordo com a equação (2.5) abaixo:

$$h^2 = \frac{\sigma_G^2}{\sigma_F^2} \quad (2.5)$$

Toda a avaliação do potencial genético dos indivíduos ou populações é feita com base nos valores fenotípicos, isso porque o valor genotípico é desconhecido. Se ambos esses valores fossem conhecidos, seria possível o cálculo da correlação entre os mesmos, e esta seria dada pela expressão:

$$r = \frac{cov(F,G)}{\sqrt{\sigma_G^2 \cdot \sigma_F^2}} \quad (2.6)$$

Em que:

$cov(F, G)$: é a covariância entre os valores genotípicos e os valores fenotípicos

σ_G^2 : é a variância genotípica

σ_F^2 : é a variância fenotípica

Ao se admitir que os efeitos do ambiente atuam de forma aleatória e tendo como base a equação (1), $F = G + E$, tem-se:

$$cov(F, G) = cov(G + E, G) = cov(G, G) + cov(G, M) = \sigma_G^2$$

E daí:

$$r = \sqrt{\frac{\sigma_G^2}{\sigma_F^2}} \rightarrow r^2 = h^2 \quad (2.7)$$

O parâmetro $r = h$ obtido na equação (2.7) refere-se à correlação entre o valor genotípico verdadeiro do tratamento genético e aquele estimado ou predito a partir das informações dos experimentos (RESENDE & DUARTE, 2007). Segundo Henderson (1984), este parâmetro estatístico também conhecido como acurácia seletiva (r_{FG}), é o parâmetro mais importante e de maior relevância para a avaliação genotípica. Pela equação (2.7), chega-se que se a herdabilidade for alta, existirá uma alta correlação entre o valor fenotípico e o valor genotípico, ou seja, a acurácia na predição será alta (CRUZ, 2010). Resende *et al.* (1995) compararam 22 diferentes métodos de seleção baseados na acurácia seletiva no melhoramento florestal e concluíram que é necessário considerar simultaneamente o ganho genético e a acurácia ao adotar um método de seleção.

A predição dos valores genéticos e os métodos de seleção dependem, essencialmente, das estimativas de componentes de variância (RESENDE & DUARTE, 2007). Os principais métodos de seleção, assim como as técnicas mais frequentes em programas de melhoramento para fins de estimação dos componentes de variância serão explicitadas no decorrer do trabalho.

2.3. Uso da experimentação para predição de valores genéticos

O estabelecimento de experimentos de campo é essencial nos programas de melhoramento genético e também no processo de recomendação de cultivares melhoradas (RESENDE & DUARTE, 2007). Por esse motivo, no melhoramento, o pesquisador realiza experimentos com o intuito de avaliar a influência de diversos fatores sobre a característica de interesse, uma vez que o valor a ser estimado é fruto de efeitos controláveis – fixos – e de efeitos imprevisíveis – chamados aleatórios (CRUZ, 2010). Esses experimentos de campo são estabelecidos sob a forma de delineamentos experimentais.

Os delineamentos experimentais têm como objetivo principal, estimar a variância genotípica a partir dos dados fenotípicos observados (CRUZ *et al.*, 2012). Um delineamento experimental é dito adequado quando instalado três princípios básicos: repetição, casualização e controle local na eficiência dos experimentos (FISHER, 1926).

Para a obtenção dessa estimativa, o melhorista ou pesquisador deve estabelecer um modelo (equação 2.8) para o caractere de estudo (Y), composto pelos efeitos fixos (E_f) e aleatórios (E_a) avaliados, a média fenotípica da característica estudada (μ) e o erro aleatório (ε) (CRUZ, 2010).

$$Y = \mu + E_f + E_a + \varepsilon \quad (2.8)$$

A partir daí, definem-se os componentes de variância, que correspondem às variâncias associadas aos efeitos aleatórios do modelo estatístico estabelecido (CRUZ, 2010).

Diversos métodos estatísticos para a obtenção dos componentes de variância podem ser encontrados na literatura, e a vasta aplicação dessas metodologias de análise nos estudos dos fenômenos biológicos e a integração da estatística como ciência na essência do método científico são as principais razões para a utilização dessas técnicas na investigação científica (FERREIRA, 2009).

Um método muito utilizado para estimação dos componentes de variância é o Método da Análise de Variância (RESENDE *et al*, 1996). A Análise de Variância (ANOVA) foi desenvolvida por Fisher em 1925, para fins de estimação dos componentes de variância e estabelecimentos de delineamentos experimentais.

Como visto acima, o modelo estatístico para a ANOVA é composto de efeitos fixos e efeitos aleatórios, estes últimos geralmente atribuídos ao erro experimental. Após a modelagem e obtenção dos resultados, o pesquisador deve proceder a uma análise criteriosa da validade dos resultados obtidos na Análise de Variância (REGAZZI, 2008). As pressuposições da Análise de Variância são: aditividade dos efeitos do modelo, normalidade dos erros, independência dos erros e homogeneidade das variâncias dos erros (RESENDE, 2007).

A análise da aditividade consiste em o pesquisador verificar se existe, de fato, uma soma dos efeitos do modelo. Esse pressuposto pode ser verificado através do Teste de aditividade de Tukey (1949) (RESENDE, 2007). O pressuposto de normalidade dos erros verifica se os resíduos associados ao modelo estatístico utilizado aderem a uma distribuição normal, e pode ser avaliado através do teste de hipóteses de Lilliefors (VIEIRA & HOFFMANN, 1989). A suposição de independência entre os erros se refere a não existência de correlação entre os erros, que podem ocasionar erros nas inferências sobre as médias (RESENDE, 2007). Segundo Resende (2007), antes de proceder a um

teste para avaliar a correlação entre os erros – Teste de Durbin-Watson –, o pesquisador deve verificar se os dados foram obtidos de forma independente, isto é, deve verificar se o princípio de casualização foi obedecido. Finalmente, verificar a existência de homogeneidade entre as variâncias consiste em verificar se o efeito do erro experimental atingiu igualmente todos os tratamentos, sendo possível, dessa forma, a obtenção de um estimador comum para a variância (VIEIRA & HOFFMANN, 1989). Testes como o de Bartlett, Hartley ou de Levene podem ser aplicados para inferência sobre a homogeneidade das variâncias (RESENDE, 2007).

Quando são verificados problemas em qualquer uma dessas pressuposições, a transformação de dados se apresenta como uma eficaz alternativa de solução (BUSTOS, 1990; HOAGLIN *et al.*, 1992).

Caracterizada a metodologia de obtenção dos componentes de variância, o próximo passo consiste em escolher um método adequado para praticar a seleção. Em relação aos métodos de genética estatística para seleção, destacam-se os derivados da teoria de seleção direta, indireta e por índice de seleção, o método de seleção combinada e o método REML-BLUP, e todos estes citados são fundamentados no conhecimento ou estimação precisa dos componentes de variância genética e fenotípica (HENDERSON, 1977; HAYES & HILL, 1980).

Como exemplo da utilização dos componentes de variância para a seleção, podem-se citar as situações nas quais o melhorista pratica a seleção com base em uma característica que proporciona alterações em outras, devido à correlação genética entre elas, fenômeno conhecido como “resposta correlacionada à seleção” (MARTINS *et al.*, 2003). Desse modo, a avaliação da resposta correlacionada é interessante para os programas de melhoramento quando o pesquisador tem à sua disposição caracteres de baixa herdabilidade ou de difícil medição ou identificação (GOLDENBERG, 1968). Nestes casos, o pesquisador pode recorrer à análise de variância como método de estimação dos componentes de variância, para tanto, basta avaliar a correlação dessas características com outras características que aparentem ser alternativas para a prática de seleção (PANETO *et al.*, 2002).

A Análise de Variância tem sido, há muito tempo, um dos principais suportes da análise de experimentos em programas de melhoramento e modelagens estatísticas, no entanto, como visto, a ANOVA deve atender a algumas pressuposições básicas, referentes à independência dos erros e entre os efeitos dos tratamentos. Outra deficiência desse método de estimação é que ele se aplica frequentemente a conjuntos de dados balanceados (CRUZ *et al.*, 2012), e quando o pesquisador se depara com

situações complexas, nos quais o experimento é estabelecido sob um delineamento em blocos incompletos – desbalanceado –, por exemplo, a ANOVA não se mostra tão eficiente para controlar a heterogeneidade experimental (RESENDE, 2007).

Com o conhecimento necessário acerca de delineamentos experimentais, análises de variâncias e suas pressuposições, o maior desafio do melhorista é desenvolver variedades melhoradas (BORÉM, 2001), utilizando-se de uma das diversas estratégias de melhoramento. Os tópicos que se seguem trazem ao leitor uma breve revisão acerca de alguns métodos de seleção disponíveis na literatura, e que são pertinentes para a prática seletiva.

2.3.1. Método de seleção massal

A seleção massal tem sido referenciada como uma das primeiras estratégias para obtenção de ganhos por seleção utilizando princípios empíricos de melhoramento genético fundamentados na associação entre o valor genético e o valor fenotípico. Nesse tipo de seleção, uma grande quantidade de indivíduos é selecionada em função de uma avaliação fenotípica, e estes são colhidos em conjunto para constituir a geração seguinte (BORÉM, 2001). Esse tipo de seleção se mostra mais eficaz quando o melhoramento é realizado em um caráter de fácil mensuração, sendo de bastante importância nas etapas iniciais dos programas de melhoramento ou quando há necessidade de produção de sementes em curto prazo (IPEF).

Segundo Borém (2001), a seleção massal constitui um dos mais antigos métodos de melhoramento de plantas e tem como objetivo principal melhorar o nível geral das populações com a reunião de seus fenótipos superiores por meio da escolha dos melhores fenótipos, ou seja, baseia-se, basicamente, nas características qualitativas do indivíduo. Vencovsky (1973) salientou que a seleção de uma planta de milho de grãos amarelos ou brancos, ou ainda, a seleção de animais com ou sem chifre, é em geral simples uma vez que as diferenças genéticas são facilmente detectadas a partir das diferenças na expressão fenotípica. Nestes casos, a prática da seleção massal mostra-se adequada uma vez que as diferenças fenotípicas e genéticas dos indivíduos são de fácil identificação.

Apesar de aparentar ser um método eficaz, a prática da seleção massal apresenta algumas dificuldades. Como visto, a seleção é praticada com base nos caracteres qualitativos do indivíduo, no entanto, indivíduos com fenótipos semelhantes

podem apresentar constituição genética ou características quantitativas distintas, e nesse caso a seleção massal nem sempre é efetiva (BORÉM, 2001).

A maior limitação dessa metodologia se deve ao fato de que os caracteres quantitativos sofrem grande influência do meio no qual se encontram, isto é, muitas vezes o melhorista não consegue saber se o indivíduo selecionado foi superior devido à sua constituição genética ou devido ao ambiente no qual foi avaliado, ou ainda, existem casos nos quais as características não conseguem se expressar no ambiente aonde o experimento foi instalado (BORÉM, 2001).

A heterogeneidade e a baixa eficiência da seleção massal foram, várias vezes, questionadas por alguns autores (GARDNER, 1961; MÉNDEZ, 1971; MOLINA, 1983). Outra deficiência da seleção massal é que seu sucesso depende da variação genética disponível na população (VENCOVSKY, 1978). Paiva *et al.* (1999) conseguiram praticar a seleção massal em acerola para plantio comercial pois a população avaliada apresentava grande variabilidade genética. Por outro lado, Bosco *et al.* (1994) obtiveram baixos percentuais de características desejáveis, resultados estes obtidos devido ao reduzido tamanho da amostra avaliada.

Outra dificuldade existente é que, na grande maioria das vezes, o caráter quantitativo apresenta um controle gênico complexo, ou seja, deseja-se praticar seleção com base em uma característica que é controlada por vários genes (CRUZ, 2010).

Também se encontra dificuldade de praticar o melhoramento quando o caráter estudado apresenta uma pequena expressão gênica. Petek *et al.* (2008) avaliaram a herdabilidade individual para prever valores genéticos aditivos em cultivares de café e concluíram que as condições não eram favoráveis à seleção para os caracteres maturação dos frutos, tamanho e produtividade dos grãos, já que os valores baixos das herdabilidades encontradas eram um indicativo da existência de baixa variabilidade genética.

2.3.2. Método de seleção de progênies

Na busca de maiores ganhos por seleção o pesquisador defronta com um agente perturbador da associação entre o fenótipo e o genótipo, que é o ambiente. Como é sabido, o ambiente exerce uma grande influência sobre o indivíduo, e, por esse motivo, por muitas vezes o melhorista seleciona erroneamente um genótipo, aparentemente superior. Diante dessa dificuldade, diversos programas de melhoramento adotam uma metodologia denominada seleção com base nos testes de progênies (RESENDE &

HIGA, 1994). Neste tipo de seleção, o melhorista possui genótipos que foram classificados como superiores, mas não tem certeza se esta superioridade é de fato devido ao genótipo ou se esta existe devido à influência do ambiente. Para tentar contornar a influência do meio na seleção, o melhorista cruza os genótipos que foram classificados como superiores entre si ou com outro indivíduo também considerado superior, e obtêm a progênie denominada família de irmãos completos ou família de meio-irmãos, respectivamente. Nessas progênies, é possível obter o valor fenotípico, e conseqüentemente, a variância fenotípica.

Vários estudos foram desenvolvidos com o intuito de verificar a eficiência do método de seleção de progênies (RESENDE *et al.*, 1991; PINTO *et al.*, 2011; SANT'ANA *et al.*, 2013). Resende *et al.* (1991) estudaram a variabilidade genética e utilizaram métodos de seleção em testes de progênies em populações de acácia-negra e concluíram que a estratégia adotada manteve variabilidade genética suficiente para produção de sementes melhoradas, pois estas apresentaram coeficientes de herdabilidade de boas magnitudes, indicando excelentes possibilidades para seleção.

2.3.3. Teoria de índice de seleção no melhoramento genético

A seleção baseada em índices é uma alternativa adicional, preconizada pelos melhoristas e biometristas para a obtenção de ganhos adicionais em um conjunto de atributos considerados relevantes. Os métodos de seleção citados até o momento praticam a seleção univariada, ou seja, com base em uma característica quantitativa de interesse. No entanto, por muitas vezes esse tipo de seleção não considera vários caracteres simultaneamente, e como consequência, os indivíduos selecionados não apresentam boa produtividade ou são bem adaptáveis (MARTINS *et al.*, 2003). Uma alternativa para o melhorista é conduzir a seleção com base em vários caracteres simultaneamente, através de uma equação de regressão múltipla cujas variáveis independentes da regressão seriam funções da importância e da herdabilidade dos mesmos (IPEF). Para obtenção de materiais genéticos que o máximo de atributos favoráveis e, portanto, mais produtivos e adaptados, uma alternativa seria usar os índices de seleção (SILVA, 1982; CRUZ & REGAZZI, 1994).

Em suma, a utilização dos índices de seleção consiste em estabelecer um novo caráter – também chamado de super caráter –, que é uma combinação linear dos caracteres envolvidos, em que os coeficientes são ponderados de modo a maximizar a correlação entre o índice e os valores genéticos verdadeiros do indivíduo ou população

(SILVA, 1982; WHITE & HODGE, 1989; VAN VLECK, 1993; CRUZ & REGAZZI, 1994).

Segundo Sampaio *et al.* (2000), seleções praticadas com base na teoria dos índices combinam uma série de informações referentes aos indivíduos candidatos à seleção, e por esse motivo, esperam-se respostas máximas nos valores genéticos desses indivíduos e, conseqüentemente, maior precisão na seleção. A seleção com base nos índices permite que se selecionem genótipos superiores, independentemente da existência ou não de correlações entre as características (SANTOS *et al.*, 2007)

Embora seja aparentemente bastante eficaz Lin (1978) relatou algumas limitações na aplicação dos índices de seleção no que se refere a imprecisões nas matrizes de variâncias e covariâncias, na mudança dos parâmetros genéticos com a seleção e na estimação dos limites de seleção. Alterações nos parâmetros genéticos obrigam que o melhorista reconstrua cada ciclo de seleção (WHITE & HODGE, 1989), e, ao estabelecer os pesos econômicos, existe grande limitação na utilização dos índices porque estes exigem que haja uma proporcionalidade entre os caracteres envolvidos (PESEK & BAKER, 1969).

2.4. Uso da abordagem de modelos mistos na predição de valores genéticos

Como visto anteriormente, a necessidade de se obter maior ganho genético levou os pesquisadores a formularem diferentes estratégias de melhoramento de forma que a observação fenotípica pudesse ser mais acurada por meio de informações de aparentados (seleção de progênie) e, ou, agregando dados de outros caracteres (seleção baseada em índices). Aliado a estes estudos, novas propostas foram apresentadas visando explorar ao máximo os próprios dados experimentais, dando grande ênfase à precisão, às distribuições e aos princípios estocásticos. Assim, inicialmente foram observadas as limitações da análise de variância, as quais forçaram os pesquisadores a desenvolver e/ou aperfeiçoar os métodos desta análise com maior fundamentação estatística. Somente em 1973 Henderson iniciou estudos envolvendo modelos mistos de seleção. Nos modelos mistos, os indivíduos para seleção são variáveis aleatórias não observáveis pertencentes a mais que uma população, e neste caso a seleção também leva em conta os efeitos fixos desconhecidos (RESENDE, 2007).

A partir desses estudos, foram desenvolvidas metodologias alternativas como *Best Linear Unbiased Prediction* (BLUP), por Henderson (1973), ou *Restricted Maximum Likelihood* (REML), por Patterson e Thompson, em 1971.

Em inferência estatística, os procedimentos adequados de estimação se baseiam em quatro propriedades desejáveis: não tendenciosidade, consistência, suficiência e eficiência (RESENDE *et al.*, 1996). O procedimento BLUP determina o melhor preditor não viesado para seleção.

O REML/BLUP permite maior flexibilidade na modelagem e nada mais é do que uma generalização da ANOVA para situações mais complexas, podendo ser ainda derivado para um modelo bayesiano (RESENDE, 2007). Resende *et al.* (1996) compararam as metodologias BLUP e REML para estimar componentes de variância e obtiveram, de fato, estimativas mais precisas quando utilizada a metodologia do REML, embora esta seja uma metodologia computacionalmente complicada.

3. Simulação de dados

Até o momento, toda a teoria envolvida para a prática de seleção, assim como as metodologias mais utilizadas foram apresentadas de forma clara ao leitor. No entanto, na prática, a seleção apresenta algumas dificuldades. Uma problemática encontrada para a execução dessas análises é que comumente os pesquisadores se deparam com situações complexas em que a veracidade de uma hipótese precisa ser comprovada por meio de um número relativamente grande de estudos ou ensaios experimentais, para as quais a obtenção de um banco de dados sob a instalação de delineamentos demanda alto custo e tempo tornando-se, muitas vezes, inviável. Segundo Blasco (1998), mesmo grandes conjuntos de dados de heterogeneidade de variâncias podem apresentar pouca informação.

Uma maneira de contornar esses problemas consiste na utilização da informática como aliada na obtenção de bancos de dados confiáveis para estudos ligados ao melhoramento genético. Uma das grandes contribuições da informática é possibilitar que o estudo de determinadas situações seja realizado por meio da simulação de uma situação complexa, em que são estabelecidos parâmetros e restrições, de forma que os efeitos de certos fatores sejam convenientemente estudados (CRUZ, 2006).

O processo de simulação consiste em emular, por meio de recursos computacionais, o comportamento de um sistema real (DACHS, 1988), englobando certos tipos de modelos lógicos que permitam descrever o sistema natural (NAYLOR, 1971).

Quando se recorre ao uso de modelos simulados, o que se busca é a obtenção de modelos mais simples do que os reais e, portanto, o pesquisador deve estar atento ao desempenho obtido a partir da simulação. A simplicidade do modelo simulado não deve afetar o seu desempenho em comparação ao modelo real; caso esse desempenho seja significativamente afetado, a simulação deve ser refinada ou refeita. Segundo McNitt (1985), a simulação envolve modelos que representam a entidade a ser investigada, porém é mais que um simples modelo: trata-se de uma metodologia para avaliação destes modelos.

Os processos de simulação podem ser constituídos por modelos físicos, matemáticos, biológicos, computacionais, dentre outros e estão distribuídos em duas linhas de pesquisa: a primeira delas envolve a resolução de problemas matemáticos e a outra utiliza simulações de ensaios baseadas em conceitos probabilísticos (CRUZ, 2006).

Um exemplo de grande aplicação é a simulação de ensaios segundo o delineamento de blocos casualizados com o objetivo principal simular planilhas de dados experimentais que possam representar um conjunto de genótipos e acompanhar ou prever a sua diversidade, a precisão experimental, a distribuição de erros, a estrutura das matrizes de covariâncias, dentre outros. De acordo com Cruz (2001), uma forma de validar o sistema simulado é fazer com que a simulação opere sob as mesmas condições do sistema real, e ainda verificar, por meio de testes de hipóteses e outras análises estatísticas relevantes ou comparações com situações já avaliadas anteriormente, se os resultados observados na simulação estão de acordo com os observados no sistema real.

É prudente, em certos estudos da área de melhoramento genético, que cada característica simulada tenha a propriedade de descrever uma variável com a média, herdabilidade e precisão experimental estabelecidas pelo pesquisador (CRUZ, 2006), pois tais informações são rotineiramente trabalhadas pelo pesquisador e conseguem captar o potencial genético do material estudado, sua variabilidade e a influência ambiental.

Coelho (2005) utilizou dados simulados com coeficiente de herdabilidade conhecidos para comparar diferentes métodos de estimação dos componentes de variância e, por consequência, métodos diferentes de estimação do coeficiente de herdabilidade e verificou quais métodos proporcionaram estimativas mais próximas do valor conhecido. A simulação foi de grande utilidade neste contexto, por permitir chegar a conclusões relevantes sem os custos da implantação e condução de experimentos e

sujeita ao tempo demandado pelas particularidades da cultura que se pretende inferir. Com a simulação é possível realizar um número relativamente grande de análises, de maior complexidade, e comprovar uma série de fenômenos ou avaliar a veracidade de diferentes hipóteses. Peixoto (2013) também utilizou dados simulados em ensaios de DBC para fazer estudos acerca do processo de aprendizado de redes neurais artificiais e sua utilização em processo de predição de valores genéticos.

4. Utilização das redes neurais na predição de valores genéticos

Diferentemente de todas as metodologias citadas até então, as Redes Neurais Artificiais (RNA's) trazem a possibilidade de aplicar os princípios da inteligência computacional no melhoramento genético para fins de predição de valores genéticos. A aplicação das Redes Neurais na análise de dados biológicos se encaixa na área de pesquisa denominada Bioinformática, e envolve técnicas e ferramentas computacionais para a resolução de problemas biológicos (BRAGA *et al.*, 2011).

Diversos autores vêm avaliando a utilização das RNA's em problemas que envolvem genética. Ventura *et al.* (2012) investigaram a aplicação das redes na predição de valores genéticos em bovinos e vislumbraram grande potencial nestas para a avaliação genética. Nascimento *et al.* (2013) utilizaram as redes para classificação de genótipos de alfafa.

Em comparação a outras metodologias, as redes apresentam a vantagem de não exigirem pressuposições a respeito da distribuição dos dados para serem utilizadas (BARROSO *et al.*, 2013).

5. Referências Bibliográficas

BARROSO, L. M. A.; NASCIMENTO, M.; NASCIMENTO, A. C. C.; SILVA, F. F.; FERREIRA, R. P. Uso do Método de Eberhart e Russell como informação a priori para aplicação de Redes Neurais Artificiais e Análise Discriminante visando a classificação de genótipos de alfafa quanto à adaptabilidade e estabilidade fenotípica. **Rev. Bras. Biom.** São Paulo, v.31, n.2, p.176-188, 2013.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. Singapore: Springer. p.738. 2007.

BLASCO, A.; SORENSEN, D.; BIDANEL, J. P. Bayesian inference of genetic parameters and selection response for litter size components in pigs. **Genetics**, v.149, n.1, p.301-306. 1998.

BORÉM, A. **Melhoramento de Plantas** – 3. Ed. Viçosa: UFV, p.500. 2001.

BOSCO, J.; AGUIAR, S. P. F.; BARREIRO, M. N. Características fenológicas de plantas de aceroleira. **Congresso Brasileiro de Fruticultura**, 13, 1994, Salvador. Anais. Salvador: SBF, p.87-88. 1994.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais - Teoria e aplicações** – 2. ed. Rio de Janeiro: LTV. 226p. 2011.

BUSTOS, O. Estimadores robustos com regras de detecção de dados surpreendentes no modelo de regressão linear. **Rev. Bras. Estat.** v.51, p.61-79, 1990.

COELHO, A. M. **Simulação de dados visando à estimação de componentes de variância e coeficientes de herdabilidade**. Dissertação (Mestrado em Agronomia). Piracicaba, 131p. 2005.

CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 2. ed. v. 2, Viçosa: UFV, 585p. 2006.

CRUZ, C.D. **Programas Genes: análise multivariada e simulação**. Viçosa, ed. UFV, p.175. 2006.

CRUZ, C. D. **Princípios de Genética Quantitativa**. 2. ed., Viçosa: UFV, p.394. 2010.

CRUZ, C. D.; REGAZZI, A.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 4. ed., v. 1, Viçosa: UFV, p.514. 2012.

DACHS, N. **Estatística computacional**. ed. São Paulo: Livros Técnicos e Científicos Editora. S.A. p.236. 1998.

FALCONER, D. S. **Introdução à genética quantitativa**. Tradução de M.A. SILVA; J.C. SILVA. Viçosa: UFV, p.279.1987.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to quantitative genetics**. Edinburgh: Addison Wesley Longman, p.464. 1996.

FERREIRA, D. F. **Estatística Básica**. 2. ed., Lavras: UFLA, p.664.2009.

FISHER, R. A. The Arrangement of field experiments. **Journal of the Ministry of Agriculture of Great Britain**, v.33, p. 503-513, 1926.

FREUND, J. E. **Estatística Aplicada**. Bookman, p.536. 2006.

GARCIA, C.; NOGUEIRA, M. Utilização da metodologia REML/BLUP na seleção de clones de eucalipto. **Scientia Forestalis**, v.68, p.107-112, 2005.

GARDNER, C. O. An evaluation of effects of mass selection and seed irradiation with thermal neutrons on yield of corn. **Crop Science**, Madison, v.1, n.2, p.124-245. 1961.

GIANOLA, D.; OKUT, H.; WEIGEL, K. A.; ROSA, G. J. M. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC Genetics**. p.12-87, 2011.<<http://www.biomedcentral.com/1471-2156/12/87>>.

GOLDENBERG, J. B. El empleo de La correlación em el mejoramiento genético de las plantas. **Fitotecnia Latino americana**. Costa Rica, v.5, n.2, p.1-8, 1968.

HAYKIN, S. S. **Redes Neurais: princípios e práticas** / Symon Haykin; trad Paulo Martins Engel. 2 ed. – Porto Alegre: Bookman, p.900. 2001.

HAYES, J. F.; HILL, W. G. A reparameterization of a genetic selection index to locate its sampling properties. **Biometrics**. Washington, v.36, p.237-248, 1980.

HAZEL, L. N. The genetic basis for constructing selection indexes. **Genetics**, v.28, p.476-490, 1943.

HENDERSON, C.R. Prediction of future records. **International conference on quantitative genetics**, 1977, Ames. Proceedings. Ames: IOWA Stte University, p.615-638. 1977.

HOAGLIN, D. C. F.; MOSTELLER; TUKEY, J. W. **Análise exploratória de dados.Técnicas robustas**. Lisboa: Salamandra, p.446. 1992.

<http://www.borges-solutions.com/algorithmo-genetico-para-otimizacao-de-arquiteturas-de-redes-mlp/>

IPEF. Melhoramento Genético. Seleção massal e individual. Monte Alegre. Instituto de Pesquisas e Estudos Florestais, p.11. 1976. (**Scientia Florestalis**. Circular Técnica, 21).

MARTINS, I. S.; CRUZ, C. D.; REGAZZI, A. J.; PIRES, I. E. Eficiência da seleção univariada direta e indireta e de índices de seleção em *Eucalyptus grandis*. **Revista Árvore**, Viçosa, v.27, n.3, p.327-333, 2003.

MARTINS, I. S.; CRUZ, C. D.; ROCHA, M. G. B.; REGAZZI, A. J.; PIRES, I. E. Comparação entre os processos de seleção entre e dentro e o de seleção combinada, em progênie de *Eucalyptus grandis*. **Cerne**, v.11, n.1, p.16-24, 2005.

MÉNDEZ, R., I. Refinamiento a la técnica de selección masal moderna. **Agrociência**, Montevidéo, v.6, n.1, p.87-91, 1971.

MOLINA, J. D. **Selección masal visual estratificada em maíz**. Chapingo: Centro de Genética, p.36. 1983. Publicación Especial.

MULAMBA, N. N.; MOCK, J. J. Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. **Egyptian Journal of Genetics and Cytology**, v.7, p.40-57, 1978.

NASCIMENTO, M.; PETERNELLI, L. A.; CRUZ, C. D.; NASCIMENTO, ANA CAROLINA CAMPANA; FERREIRA, R. P.; BHERING, L. L.; SALGADO, C. C.

Artificial neural networks for adaptability and stability evaluation in alfafa genotypes. Crop Breeding and Applied Biotechnology, v.12, p.152-156, 2013.

NAYLOR, T.H.; BALINTFY, J.L.; BURDICK, D.S.; CHU, K. **Técnicas de simulação em computadores.** Ed. Vozes LTDA, São Paulo, p.401. 1971.

ODA, S.; MELLO, E. J.; SILVA, J. F.; SOUZA, I. C. G. Melhoramento florestal. In: BORÉM, A. (Ed.). **Biotecnologia Florestal.** Viçosa: UFV, p. 51-71. 2007

OLIVEIRA, F. E. M. **Estatística e Probabilidade,** Editora Atlas, 1999.

PAIVA, J. R.; ALVES, R. E.; CORREA, M. P. F.; FREIRE, F. C. O.; SOBRINHO, R. B. Seleção massal de acerola em plantio comercial. **Pesq. agropec. bras.,** Brasília, v.34, n.3,1999. p.505-511.

PANETO, J. C. C.; LEMOS, D. C.; BEZERRA, L. A. F.; FILHO, R. M.; LÔBO, R. B. Estudo de Características Quantitativas de Crescimento dos 120 aos 550 Dias de Idade em Gado Nelore. **R. Bras. Zootec.,** v.31, n.2, p.668-674, 2002.

PAULA, R. C.; PIRES, I.E.; BORGES, R. G.; CRUZ, C. D. Predição de ganhos genéticos em melhoramento florestal. **Pesquisa Agropecuária Brasileira,** v.37, p.159-165, 2002.

PEIXOTO, L. A. **Redes Neurais Artificiais na predição do valor genético.** Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, 97p, 2013.

PESEK, J.; BAKER, R.J. Desired improvement in relation to selection indices. **Canadian Journal of Plant Sciences.** Ottawa, v.1, p.215-274, 1969.

PETEK, M. R.; SERA, T.; FONSECA, I. C. B. Prediction of genetic additive values for development of a coffee cultivar with increased rust resistance. **Bragantia,**v. 67, n. 1, p. 133-140, 2008.

PINTO, M. F.; CARVALHO, G. R.; BOTELHO, C. E.; GONÇALVES, F. M. A.; REZENDE, J. C. R.; FERREIRA, A. D. Eficiência na seleção de progênies de cafeeiro avaliadas em Minas Gerais. **Bragantia**, Campinas, v. 71, n. 1, p.1-7, 2012.

RESENDE, M. D. V.; SOUZA, S. M.; HIGA, A. R.; STEIN, P. P. Estudos da variação genética e métodos de seleção em Teste de Progênies de *Acacia mearnsii* no Rio Grande do Sul. **Boletim de Pesquisa Florestal**. Colombo, n. 22/23, p.45-59, 1991.

RESENDE, M. D. V.; ARAUJO, A. J.; SAMPAIO, P. T. B.; WIECHETECK, A.S. Acurácia seletiva, intervalos de confiança e variância de ganhos genéticos associados a 22 métodos de seleção em *Pinus caribaea* var. *hondurensis*. **Revista Floresta**, Curitiba, v.24, n.1/2, p.35-45, 1995.

RESENDE, M. D. V.; PRATES, D. F.; YAMADA, C. K.; JESUS, A. Estimação de componentes de variância e predição de valores genéticos pelo método da máxima verossimilhança restrita (REML) e melhor predição linear não viciada (BLUP) em *Pinus*. **Boletim de Pesquisa Florestal**, Colombo, n.32/33, p.18-45, 1996.

RESENDE, M. D. V. **Matemática e Estatística na Análise de Experimentos e no Melhoramento Genético**. Colombo: Embrapa Florestas, p.362. 2007.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v.37, n.3, p.182-194, 2007.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v.323, n.6088, p.533-536, 1986.

SAMPAIO, P. T. B.; RESENDE, M. D. V.; ARAÚJO, A. J. Estimativas de parâmetros genéticos e métodos de seleção para o melhoramento genético de *Pinus Caribaea* var. *Hondurensis*. **Pesq. agropec. bras., Brasília**, v.35, n.11, p.2243-2253, 2000.

SANT'ANA, V. Z.; Miguel Luiz Menezes FREITAS, M. L. M.; MORAES, M. L. T.; ZANATA, M.; ZANATTO, A. C. S.; MORAES, M. A.; SEBBENN, A. M. Parâmetros

genéticos em progênes de polinização aberta de *Enterolobium contortisiliquum* (Vell.) Morong em Luiz Antonio, SP, Brasil. **Hoehnea**. 40(3): p.515-520, 2013.

SANTOS, F. S.; JÚNIOR, A. T. A.; JÚNIOR, S. P. F.; RANGEL, R. M.; PEREIRA, M. G. Predição de ganhos genéticos por índices de seleção na população de milho-pipoca UNB-2U sob seleção recorrente. **Bragantia**. Campinas, v.66, n.3, p.389-396, 2007.

SILVA, R.G. **Métodos de genética quantitativa aplicados ao melhoramento animal**. Ribeirão Preto, Sociedade Brasileira de Genética, p.162. 1982.

SMITH, H. F. A discriminant function for plant selection. **Annual Eugenics**, v.7, p.240-250, 1936.

VENCOVSKY, R. Genética quantitativa. In: PATERNIANI, E. (Coord.). **Melhoramento do milho no Brasil**. Piracicaba: Fundação Cargill, p.122-201.1978.

VENTURA, R.; SILVA, M.; MEDEIROS, T.; DIONELLO, N.; MADALENA, F.; FRIDRICH, A.; VALENTE, B.; SANTOS, G.; FREITAS, L.; WENCESLAU, R. Use of artificial neural networks in breeding values prediction for weight at 205 days in Tabapuã beef cattle. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**,v.64, n.2, p.411-418, 2012.

VIEIRA, S; HOFFMANN, R. **Estatística experimental**. São Paulo: Atlas S.A., p.179. 1989.

WHITE, T. L.; HODGE, G. R. **Predicting Breeding Values with Applications in Forest Tree Improvement**. Kluwer Academic Publishers. Dordrecht, p.363. 1989.

CAPÍTULO 1

AVALIAÇÃO DA EFICIÊNCIA DAS REDES NEURAIS ARTIFICIAIS NA PREDIÇÃO DE VALORES GENÉTICOS

Avaliação da eficiência das Redes Neurais Artificiais na predição de valores genéticos

RESUMO

Dada a potencialidade das Redes Neurais Artificiais em diversas áreas da ciência, este trabalho foi realizado com o intuito de propor a sua utilização como alternativa viável aos métodos convencionais de predição de valor genético em programas de melhoramento genético. Visando realizar uma avaliação minuciosa da eficiência das RNA's na predição de valores genéticos, foram obtidos e analisados dados de experimentos simulados com o fim de verificar as potencialidades das redes neurais em relação aos métodos estatísticos comumente utilizados no melhoramento genético fundamentados em testes comparativos de médias. As redes neurais foram estabelecidas considerando como entradas, para fins de aprendizado, informações de médias, máximos, mínimos, variâncias e classificação genotípica, dentre outros. Foram considerados oito cenários simulados de experimentação, com coeficiente de variação igual a 15% e herdabilidades de 40, 45, 50, 55, 60, 65, 70 e 80%. Cada cenário consistiu em 100 genótipos dispostos em delineamento de blocos casualizados com seis repetições. Para o processo de simulação, foi utilizado o aplicativo computacional GENES. Foram considerados para predição do valor genético, sete parâmetros estatísticos além da média fenotípica em uma rede com estrutura de perceptron de múltiplas camadas. O algoritmo das redes neurais foi processado pelo aplicativo computacional Matlab por meio de scripts disponibilizado no aplicativo computacional GENES. Após a avaliação das diferentes configurações e arquitetura, os resultados obtidos demonstraram a superioridade das Redes Neurais Artificiais quando comparadas aos processos de estimação baseados em modelos lineares, e indicaram ser viável sua aplicabilidade em processos de predição de valores genéticos. Os resultados indicaram boa capacidade de generalização do modelo ajustado pela rede validado em diversos experimentos.

Palavras-chave: Inteligência artificial; simulação; predição.

INTRODUÇÃO

Um dos grandes desafios do melhorista, nas áreas de genética animal ou vegetal, é selecionar e recomendar genótipos visando o aumento da produtividade da

população de trabalho. Para tanto, a genética biométrica, evocando conceitos da genética quantitativa e de populações, tem fornecido enorme contribuição para o melhoramento genético por meio do desenvolvimento de metodologias que vêm auxiliando o melhorista a identificar os melhores genótipos, predizer ganhos genéticos e recomendar variedades (CRUZ, 2010). De acordo com Resende *et al.* (2007), um programa de melhoramento genético é dito eficiente quando utiliza métodos acurados de seleção. A utilização destes métodos de predição tem possibilitado ao melhorista orientar os programas de melhoramento e decidir, com embasamento científico, por estratégias potencialmente mais eficazes.

Para a recomendação ou seleção de um indivíduo, o pesquisador precisa lançar mão de métodos de experimentação e de análise estatística (RESENDE, 2007) e, em relação aos métodos de genética estatística para seleção, algumas categorias se destacam: os métodos derivados da teoria de melhoramento genético tais como a seleção massal (GARDNER, 1961; MÉNDEZ, 1971; MOLINA, 1983), seleção de progênies (RESENDE & HIGA, 1994) e por índice de seleção (SILVA, 1982; CRUZ & REGAZZI, 1994) e métodos de seleção fundamentados em teorias biométricas englobando a classe de predição de valores genéticos a partir de análise considerando os modelos mistos, como o REML-BLUP (PATTERSON & THOMPSON, 1971).

Apesar da diversidade de princípios e de metodologias para seleção, uma particularidade comum a praticamente todas elas é que o pesquisador busca o máximo de informações sobre o valor genético da característica, mas fundamenta sua predição com base no fenótipo (CRUZ, 2010) e, atualmente, agregando valiosas informações de marcadores moleculares quando possível. Como os efeitos ambientais são pronunciados para grande parte das características de interesse econômico, outra ocorrência comum é encontrar experimentos avaliados com baixa confiabilidade de predição do valor genotípico verdadeiro, o que força os programas de melhoramento a utilizar a média fenotípica das repetições como a medida mais apropriada para indicar a superioridade genética e predizer o ganho genético (PETEK *et al.*, 2008).

Considerando ainda que as Redes Neurais Artificiais têm demonstrado grande potencial em diversas áreas do melhoramento (GIANOLA *et al.* 2011; NASCIMENTO *et al.*, 2013; VENTURA *et al.*, 2013), desenvolvemos este trabalho com o intuito de avaliar uma nova abordagem para predição de valores e ganhos, baseada nos princípios inteligência computacional capaz de captar muito mais informações de um experimento do que valor fenotípico médio e ser livre, em seu processo de predição, da influência das distorções nas distribuições estocásticas dos efeitos considerados. Para tanto, foram

consideradas a construção de RNA's por meio de valores genéticos simulados, nos quais foram consideradas informações comumente negligenciadas como entrada da RNA.

MATERIAL E MÉTODOS

1. Simulação de valores genéticos

O processo de simulação de dados foi realizado por meio do aplicativo computacional GENES, conforme descrito por (CRUZ, 2013). Foram simulados oito cenários com valores de média, coeficiente de variação e herdabilidade bem estabelecidos. O delineamento experimental adotado na simulação foi o delineamento em blocos casualizados (DBC) com 100 genótipos e seis blocos, assumindo-se herdabilidades de 40, 45, 50, 55, 60, 65, 70 e 80%, valores de média igual a 100 e coeficiente de variação de 15%.

Para cada variável devem ser estabelecidos os valores da média, da herdabilidade e do coeficiente de variação experimental, e utilizado o seguinte modelo estatístico descrito na equação (1.1):

$$Y_{ij} = \mu + G_i + B_j + \varepsilon_{ij} \quad (1.1)$$

Em que:

Y_{ij} : observação simulada de uma dada característica;

μ : média geral da característica, cujo valor é especificado pelo pesquisador;

G_i : efeito associado ao i -ésimo genótipo;

B_j : efeito associado ao j -ésimo bloco;

ε_{ij} : erro aleatório, sendo $\varepsilon_{ij} \sim N(0, \sigma^2)$

1.1. Simulação do efeito dos blocos

Para a simulação dos efeitos fixos dos blocos foi considerado um conjunto de dados contendo n valores, em progressão aritmética, de razão r e média \bar{X} , em que o primeiro termo é denotado por X_1 e o último por X_n e variância (s^2) dada pela equação (1.2):

$$s^2 = \frac{n(n+1)}{3(n-1)^2} (X_n - \bar{X})^2 \quad (1.2)$$

Assim, para estimar os efeitos de blocos é admitida a existência de b efeitos fixos, cujos valores configuram uma progressão aritmética de razão r com a particularidade de que $B_1 = -B_b$ e $\bar{B} = 0$. Logo, o valor B_b é estimado por meio de $B_b = \frac{(n-1)\sqrt{3\phi_b}}{\sqrt{n(n+1)}}$ e os demais efeitos estabelecidos considerando a razão da progressão aritmética dada por $r = \frac{B_b - B_1}{b-1}$.

1.2. Simulação do efeito dos genótipos

Para estimar os efeitos de genótipos é necessário conhecer o valor da variância genética, que é obtida a partir das informações sobre a herdabilidade (h^2) e o coeficiente de variação experimental (CVe). Assim, obtêm-se o valor da variância ambiental por meio de $\sigma^2 = \left(\frac{\mu CV_e}{100}\right)^2$.

$$\text{Sabendo que } h^2 = \frac{100\sigma_G^2}{\sigma_G^2 + \frac{1}{b}\sigma^2} \text{ segue que: } \sigma_G^2 = \frac{\sigma^2 h^2}{b(100-h^2)}.$$

A partir dessas informações, podem ser estabelecidos os efeitos fixos ou aleatórios de genótipos. No primeiro caso é admitida a existência de g efeitos fixos, cujos valores configuram uma progressão aritmética de razão r com a particularidade de que $G_1 = -G_g$ e $\bar{G} = 0$. Logo, o valor G_g é estimado por meio de $G_g = \frac{(g-1)\sqrt{3\sigma_G^2}}{\sqrt{g(g+1)}}$ e os demais efeitos estimados considerando a razão da progressão aritmética dada por $r = \frac{G_g - G_1}{g-1}$.

1.3. Simulação dos erros aleatórios

Para simular o efeito dos erros aleatórios, considera-se que $\varepsilon_{ij} \sim \text{NID}(0, \sigma^2)$. Como o pesquisador fornece o valor do coeficiente de variação experimental e da média da característica, o valor de σ^2 torna-se conhecido e, portanto, os erros aleatórios e independentes podem ser estimados usando a função randômica descrita anteriormente.

1.4. Estabelecimento dos valores fenotípicos

Conhecidos o valor da média da característica e dos efeitos envolvidos, os valores fenotípicos de cada variável, em cada ambiente (quando for o caso), são estabelecidos por meio do modelo:

$$Y_{ij} = \mu + G_i + B_j + \varepsilon_{ij} \quad (1.3)$$

2. Ampliação de dados

Para fins de treinamento da rede, procedeu-se a uma ampliação do conjunto de dados inicial por meio do processo descrito a seguir.

Os valores simulados foram tomados como uma variável aleatória $Y \sim N(\phi, \Sigma)$. Os dados foram transformados em uma variável aleatória $Z \sim N(\phi, I)$ por meio da transformação linear $Z = F'Y$, sendo F obtida por meio do processo de decomposição espectral de Σ , tal que $\Sigma^{-1} = FF'$. O processo de ampliação consiste na simulação de novos valores de Y , considerando $Y \sim N(\phi, (F')^{-1}Z)$. Foi considerado um arquivo de dados ampliados, consistindo de 5000 genótipos para treinamento da Rede Neural. O processo de ampliação de dados também foi realizado por meio do módulo integração no aplicativo computacional GENES (CRUZ, 2013).

Em procedimentos de simulação de um conjunto de dados, ou replicação de uma estrutura de dados conhecidos, ou mesmo a ampliação de um conjunto a partir da estrutura de outro, algumas pressuposições devem ser satisfeitas. A primeira delas é que o conjunto de dados deve ter uma distribuição conhecida e, a princípio, média igual a zero e variância igual a V . Para satisfazer a essa exigência, utilizou-se o Teorema de Box Muller, que garante que as variáveis x e y são normalmente distribuídas com média zero e variância V , onde:

$$x = \sqrt{-2\log_e(RND)V} \cos(2\pi RND) \quad (2.1)$$

e

$$y = \sqrt{-2\log_e(RND)V} \sin(2\pi RND) \quad (2.2)$$

Sendo RND um número aleatório.

Para garantir que a covariância do conjunto X de dados é nula, a metodologia recomendada é o uso da técnica de componentes principais e se baseia na simplificação do conjunto de dados para um conjunto reduzido de componentes, os quais apresentam as propriedades de reter o máximo da variação originalmente disponível e ser independentes entre si (CRUZ, 2006).

Considere a variável aleatória $Y \sim N(\phi, \Sigma)$ que desejamos transformar em uma variável aleatória $Z \sim N(\phi, I)$. Por meio do processo de decomposição espectral, temos que $\Sigma^{-1} = FF'$. Então, $(\Sigma^{-1})^{-1} = (FF')^{-1} = (F')^{-1}F^{-1} = \Sigma$

Se $Z = F'Y$, então $E(Z) = E(F'Y) = F'E(Y) = F'\phi = \phi$ e $V(Z) = F'V(Y)F = F'\Sigma F = F'[(F')^{-1}F^{-1}]F = I$

Consideremos que $Z \sim N(\phi, I)$. Se $Z = F'Y$, então $Y = (F')^{-1}Z$

Portanto, $V(Y) = V((F')^{-1}Z) = (F')^{-1}V(Z)[(F')^{-1}]' = (F')^{-1}I(F')^{-1} = \Sigma$

No conjunto de dados ampliados, foram obtidos arquivos com 5000 genótipos para treinamento da Rede Neural Artificial.

3. Proposta para estabelecimento do critério de seleção da rede neural artificial

3.1.1. Modelo convencional

Nos programas de melhoramento, para fins de seleção de genótipos superiores, o modelo convencionalmente adotado baseia-se na decomposição dos valores fenotípicos dos indivíduos na média da característica e nos efeitos envolvidos, tal qual a equação (3.1) descrita a seguir:

$$Y_{ij} = \mu + G_i + B_j + \varepsilon_{ij} \quad (3.1)$$

Em que:

Y_{ij} : observação de uma dada característica;

μ : média da característica;

G_i : efeito do i-ésimo genótipo;

B_j : efeito do j-ésimo bloco;

ε_{ij} : erro aleatório.

Como visto no decorrer do trabalho, a seleção é usualmente praticada com base nos valores das médias fenotípicas obtidos. Assim, dados dois indivíduos i e i' , tem-se que suas médias são dadas por:

$$\bar{Y}_i = \hat{\mu} + \hat{G}_i = \hat{\mu}_i \quad (3.2)$$

$$\bar{Y}_{i'} = \hat{\mu} + \hat{G}_{i'} = \hat{\mu}_{i'} \quad (3.3)$$

Dadas as equações (3.2) e (3.3), o indivíduo recomendado pelo melhorista será aquele que apresentar maior valor para a média fenotípica estimada ($\hat{\mu}_i$).

3.1.2. Modelo proposto – Adaptação de MCP

Na metodologia proposta neste trabalho, acreditamos que a consideração de outras estatísticas adicionais no processo de tomada de decisão aumente a acurácia seletiva, isto é, a ampliação das informações acerca dos genótipos avaliados agrega valor aos genótipos e auxilia o melhorista na seleção.

Desse modo, para um dado genótipo G_i , propõe-se o modelo (3.4):

$$G_i = \sum_k w_k E_{ik}, \varphi(\cdot), \theta \quad (3.4)$$

Em que:

w_k : peso sináptico associado à k -ésima estatística;

E_{ik} : estatística k associada ao genótipo i ;

$\varphi(\cdot)$: a função de ativação utilizada;

θ : o limiar de ativação (bias) da rede.

Estabelecido o modelo, dados dois genótipos G_i e $G_{i'}$, basta comparar os valores obtidos para cada um deles, tal como descrito na equação (3.4). O indivíduo que apresentar maior valor para G_i é considerado superior.

3.1.3. Modelo proposto – generalização de MLP

Esquemáticamente, o modelo de Rede Neural Artificial proposto é representado pela figura (3.2) abaixo.

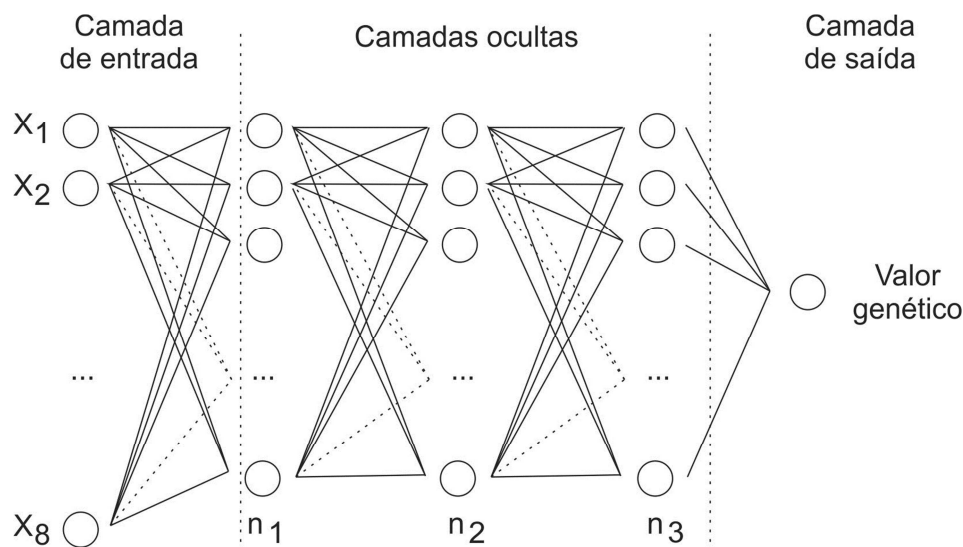


Figura 1. Arquitetura da RNA. Entradas de (x_1) a (x_8) na camada entrada estão relacionadas com as estatísticas estimadas a partir dos dados e são consideradas como entrada. As camadas ocultas foram compostas por n_i (n_i variando de um a sete nós), com funções de ativação *purelin*, *tansig* ou *logsig*. Todas as combinações foram exploradas. Na camada de saída, as RNA's retornaram o valor genético predito.

4. Construção e avaliação da Rede Neural Artificial

A arquitetura de Rede utilizada foi a *Multilayer Perceptron* (MLP), que no presente trabalho foi estabelecida adotando-se três camadas ocultas de neurônios. A Rede MLP foi criada no software Matlab por meio de script disponibilizado no módulo integração no aplicativo computacional GENES.

Para treinamento a rede, adotou-se o algoritmo de treinamento *trainlm*, e as funções de ativação investigadas para estabelecer a melhor arquitetura foram as funções linear – *purelin*, tangente hiperbólica – *tansig*– e a função sigmoial – *logsig*.

O número de neurônios por camada variou de um a sete, adotando-se o número máximo de iterações (ou épocas) igual a 2000. Combinando a variação de número de neurônios por camada e as possibilidades de função de ativação, um total de $7 \times 7 \times 7 \times 3 \times 3 \times 3 = 9261$ arquiteturas de rede foram testadas.

Como alternativa para os métodos usuais de seleção, que baseiam sua seleção apenas nos valores das médias fenotípicas dos genótipos, a rede foi treinada e validada

utilizando como saída desejada os valores genotípicos verdadeiros e como entrada as informações de outras sete estatísticas, além da média, como descrito abaixo:

x_1 : o valor de média fenotípica;

x_2 : o valor máximo de cada genótipo nos b blocos;

x_3 : o valor mínimo de cada genótipo nos b blocos;

x_4 : o valor de desvio padrão de cada genótipo;

x_5 : o valor de coeficiente de variação para cada genótipo;

x_6 : o valor da soma de rank (ou postos) obtidos nos b blocos para cada genótipo;

x_7 : o valor da estatística P_i , adaptada de Lin&Bins (1988) para estudos de adaptabilidade e estabilidade, dada por:

$$\sum_{j=1}^b \frac{(X_{ij} - M_j)^2}{2b}$$

sendo X_{ij} , o valor fenotípico do i -ésimo genótipo no j -ésimo bloco; M_j , a resposta máxima observada entre todos os genótipos no bloco j ; e b , o número de blocos do experimento;

x_8 : os valores discretos, codificados de 1 a 6, estabelecidos a partir da transformação da média fenotípica considerando a sua variação em torno da média nos limites, inferior e superior, tomados como um, dois ou três desvio-padrão.

Para verificar a eficiência da RNA, foram gerados arquivos de validação, também obtidos por meio do processo simulação de ensaios em DBC, e com número de blocos e de genótipos exatamente iguais aos do arquivo original, considerados réplicas do arquivo de teste por manterem o mesmo vetor de médias e a mesma estrutura de variâncias e covariâncias entre os valores obtidos pelos genótipos nos diferentes blocos. Assim sendo, foram utilizados 120 arquivos de validação, tomados como réplicas dos experimentos simulados. O critério adotado pela Rede Neural Artificial para finalizar as análises fundamentadas em inteligência computacional se baseou no número de vezes em que a confiabilidade da predição genotípica dada pela rede se mostrou superior ao da média fenotípica. A confiabilidade foi dada pelo quadrado da correlação entre a resposta da RNA e o valor genotípico verdadeiro em analogia ao quadrado da correlação entre a média fenotípica e o valor fenotípico verdadeiro, que expressa a herdabilidade da característica, aplicada em todas as combinações de neurônios, número camadas ocultas e funções de ativação utilizadas. Assim, foi feita a comparação entre uma medida da confiabilidade da rede e a herdabilidade da característica, que é o critério utilizado de

forma convencional para predição de ganhos, identificação de unidades seletivas e escolha de estratégias de seleção.

RESULTADOS E DISCUSSÃO

1. Análise dos dados do arquivo original (teste) e do conjunto de dados ampliados para treinamento da RNA

A metodologia apresentada neste trabalho consistiu em utilizar um conjunto de dados ampliados, que têm a particularidade de preservar as características do experimento dos quais eles são derivados. Com o intuito de demonstrar que a ampliação de dados não alterou a estrutura do conjunto de dados originais, tal como estabelecido pelo pesquisador, estão apresentadas na Tabela 1.1 as estimativas de parâmetros genéticos estimados para os conjuntos de dados simulados e ampliados, para as herdabilidades de 40, 45, 50, 55, 60, 65, 70 e 80%.

Pela análise de variância dos dados dos diferentes arquivos foi constatada a eficácia do processo de simulação em conservar características desejadas. Os resultados das análises de variância para os oito cenários são apresentados na Tabela 1.1.

Nos procedimentos estatísticos convencionais, utilizados no processo de predição de ganho genético, são considerados somente os valores genotípicos médios assumindo a existência de correlação destes com os valores genéticos reais dos genótipos denotados por herdabilidade da característica. Admite-se que a média seja a principal informação para julgar a superioridade entre genótipos e que, em casos de baixa herdabilidade, o processo seletivo seja de menor confiabilidade.

Tabela 1.1. Estimativas de parâmetros genéticos derivados da análise de variância dos arquivos simulados (S) e dos dados ampliados (A), para os cenários de herdabilidades de 40, 45, 50, 55, 60, 65, 70 e 80%. O processo de expansão de dados preserva a estrutura do arquivo de dados original.

		Estimativas dos parâmetros						
		μ	σ_g^2	σ_e^2	h^2	h_{ind}^2	$CV_e(\%)$	$CV_g(\%)$
$h^2 = 40\%$	S	100.0000	25.0000	225.0000	40.0000	10.0000	15.0000	5.0000
	A	100.0001	25.0019	225.0051	40.0013	10.0005	15.0002	5.0002
$h^2 = 45\%$	S	100.0000	30.6818	225.0000	45.0000	12.0000	15.0000	5.5391
	A	100.0001	25.0019	225.0051	40.0013	10.0005	15.0002	5.0002
$h^2 = 50\%$	S	100.0000	37.5713	224.9946	50.0000	14.2864	14.9998	6.1238
	A	100.0001	37.5013	224.9946	50.0014	14.2864	14.9998	6.1238
$h^2 = 55\%$	S	100.0000	45.8333	225.0000	55.0000	16.9231	15.0001	6.7700
	A	100.0000	45.8282	225.0044	54.9968	16.9212	15.0001	6.7697
$h^2 = 60\%$	S	100.0000	56.2495	225.0030	59.9995	19.9997	15.0001	7.5000
	A	100.0002	56.2495	225.0029	59.9995	19.9997	15.0001	7.4999
$h^2 = 65\%$	S	100.0000	69.6429	225.0000	65.0000	23.6364	15.0000	8.3452
	A	99.9999	69.6454	225.0006	65.0008	23.6370	15.0000	8.3454
$h^2 = 70\%$	S	100.0000	87.5000	225.0000	70.0000	28.0000	15.0000	9.3541
	A	99.9999	87.5005	225.0024	69.9999	27.9999	15.0001	9.3542
$h^2 = 80\%$	S	100.0000	150.0000	225.0000	80.0000	40.0000	15.0000	12.2474
	A	100.0000	150.0029	225.0014	80.0002	40.0003	15.0000	12.2476

μ – média; σ_g^2 – variância genética; σ_e^2 – variância residual; h^2 – herdabilidade em nível de média de genótipos; h_{ind}^2 – herdabilidade em nível de indivíduo; CV_e – coeficiente de variação experimental; CV_g (%) – coeficiente de variação genético.

Uma análise mais criteriosa poderia ser realizada tomando, além da média dos genótipos, outras estatísticas relevantes e, ou, de elevado custo prático, mas que torna a tomada de decisão mais complexa por parte do analista de dados. Assim, dois genótipos poderão ter médias similares, mas com comportamento bastante distinto em termos de variação, ranqueamento, valores máximos e mínimos dentre outros. Na abordagem biométrica estas informações adicionais são negligenciadas, porém quando se utiliza uma abordagem de inteligência computacional tais informações são indispensáveis e constituem entradas relevantes a serem consideradas na Rede Neural Artificial em processos de treinamento.

Ao considerarmos os valores globais de diversas estatísticas de cada genótipo obtidos em um determinado ensaio, observa-se que se obtêm critérios mais confiáveis para prever a superioridade genotípica do que fundamentar-se unicamente no valor médio obtido para o genótipo em ensaios cuja variabilidade genotípica e precisão experimental podem variar enormemente. Com o intuito de verificar se o processo de simulação como o processo de ampliação afetou a estrutura do experimento expandido foi realizada uma análise de correlação entre as estatísticas utilizadas na camada de entrada e os valores genéticos reais (Tabela 1.2). As correlações (r) mostraram ser estatisticamente iguais, considerando um teste t ($p < 0.01$) que avaliou as correlações para o arquivo simulado e para o arquivo expandido. Ainda, embora as estatísticas utilizadas na camada de entrada tenham se mostrado suficientes para predição do valor genético, estudos posteriores são necessários para quantificar a influência de cada uma destes no processo de predição, bem como avaliar a retirada ou a inclusão de estatísticas.

Tabela 1.2. Correlação (r) entre os valores genéticos e os valores das estatísticas estimadas para os dados simulados (S) e para os dados expandidos (A).

r		VG x μ	VG x max	VG x min	VG x dp	VG x CV	VG x Pi	VG x Rank	VG x classe
$h^2 = 40\%$	S	0.4001*	0.2928*	0.2417*	0.0015 ^{ns}	0.011 ^{ns}	0.3623*	0.3848*	0.4290*
	A	0.4001*	0.2253*	0.2552*	0.0000 ^{ns}	0.0323 ^{ns}	0.3923*	0.3772*	0.3624*
$h^2 = 45\%$	S	0.4499*	0.3294*	0.2772*	0.0015 ^{ns}	0.0148 ^{ns}	0.411*	0.4372*	0.4582*
	A	0.4499*	0.2432*	0.2787*	0.0000 ^{ns}	0.0406*	0.4541*	0.4404*	0.3982*
$h^2 = 50\%$	S	0.5000**	0.3681*	0.3155*	0.0015 ^{ns}	0.0195 ^{ns}	0.4608*	0.4823*	0.4719*
	A	0.5000**	0.3262*	0.3439*	0.0001 ^{ns}	0.0565 ^{ns}	0.491*	0.4823*	0.3925*
$h^2 = 55\%$	S	0.5500**	0.4092*	0.3567*	0.0015 ^{ns}	0.0255 ^{ns}	0.5115**	0.530**	0.4748*
	A	0.5500**	0.381*	0.3736*	0.0031 ^{ns}	0.0328 ^{ns}	0.5471**	0.540**	0.4625*
$h^2 = 60\%$	S	0.6000**	0.4532*	0.4014*	0.0015 ^{ns}	0.0332 ^{ns}	0.5634**	0.5796**	0.5168**
	A	0.6000**	0.4427*	0.4498*	.0002 ^{ns}	0.0793*	0.5969**	0.5839**	0.5602**
$h^2 = 65\%$	S	0.650**	0.5002**	0.4503*	0.0015 ^{ns}	0.0433 ^{ns}	0.6161**	0.6302**	0.5256**
	A	0.650**	0.4193*	0.4684*	0.0017 ^{ns}	0.089**	0.6507**	0.6341**	0.6022**
$h^2 = 70\%$	S	0.7001**	0.5512**	0.5039**	0.0015 ^{ns}	0.0570 ^{ns}	0.6696**	0.6871**	0.6395**
	A	0.7001**	0.5253**	0.4982**	0.0022 ^{ns}	0.0978**	0.691**	0.6752**	0.6410**
$h^2 = 80\%$	S	0.800**	0.6676**	0.6291**	0.0015 ^{ns}	0.1046*	0.7804**	0.7949**	0.6994**
	A	0.800**	0.6003**	0.6736**	0.0009 ^{ns}	0.1996*	0.7998**	0.7828**	0.7245**

VG – Valor genético; max – Valor máximo; Min – Valor mínimo, dp – desvio padrão; Rank – soma de ranks; classe – classificação;
 *Significativo a 5%; **Significativo a 1%

A mesma avaliação foi realizada considerando os 120 arquivos replicados utilizados no processo de validação. Para facilitar a avaliação das correlações entre estatísticas adicionais e os valores genéticos reais, para cada herdabilidade, gerou-se um gráfico em boxplot (Figuras 1.1 e 1.2). As correlações para cada réplica encontram-se em anexo.

Foram encontrados resultados semelhantes aos obtidos a partir da simulação e da expansão do arquivo original. O raciocínio subjacente a esta etapa consiste no fato de que o pesquisador poderia questionar sobre a validade do processo validação, uma vez que este pode algumas vezes ser realizado com o arquivo de treinamento ou em parte deste. Os arquivos replicados servem para emular situações na qual a RNA criada é utilizada na predição de experimentos contemporâneos dos genótipos avaliados.

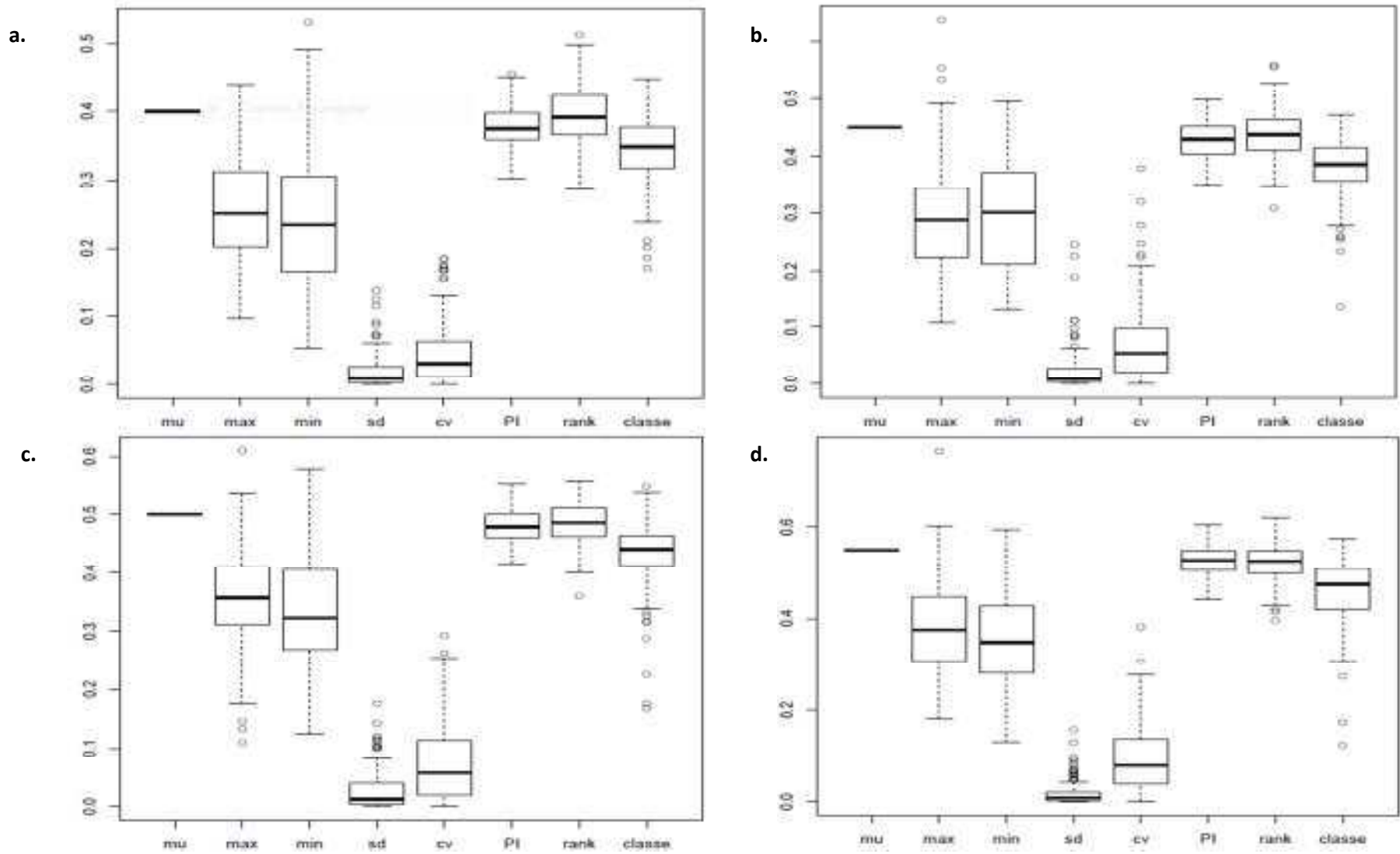


Figura 1.1.Boxplot das estatísticas para as herdabilidades $h^2 = 40\%$ (a.), $h^2 = 45\%$ (b.), $h^2 = 50\%$ (c.), $h^2 = 55\%$ (d.).

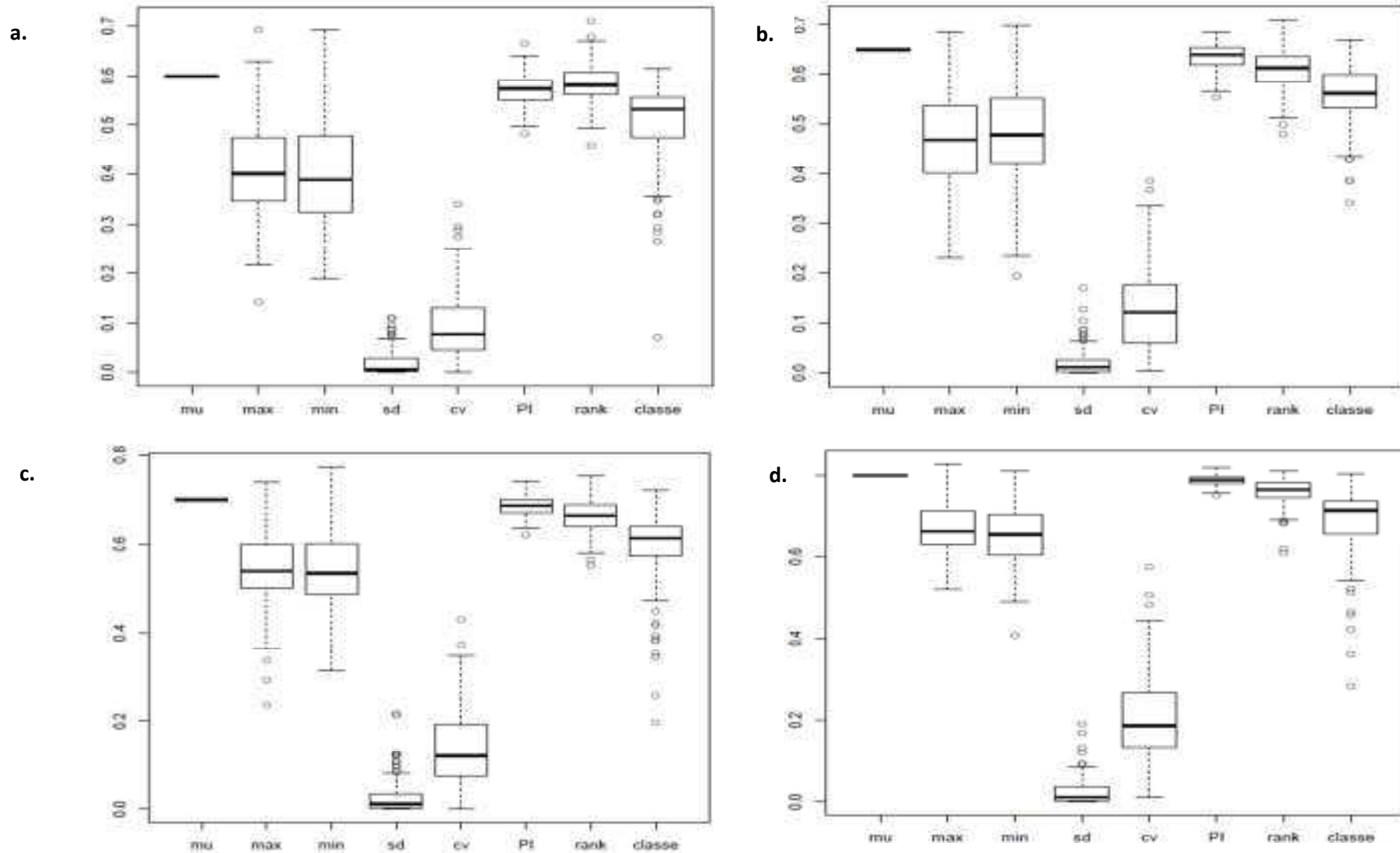


Figura 1.2.Boxplot das estatísticas para as herdabilidades $h^2 = 60\%$ (a.), $h^2 = 65\%$ (b.), $h^2 = 70\%$ (c.), $h^2 = 80\%$ (d.).

2. Eficiência preditiva e desempenho da Rede Neural Artificial

O procedimento usual para determinar o critério de otimização da arquitetura da rede, também denominado de critério de parada do processo iterativo da Rede Neural Artificial, é o erro quadrático médio de previsão. Uma vez que o erro quadrático médio chegou a um nível ideal, o treinamento é interrompido, e os valores obtidos até então são as melhores estimativas dos coeficientes da rede (HAYKIN, 2001; BRAGA *et al.*, 2011). Este critério de parada é normalmente adotado quando o processo de validação da RNA é realizado com apenas um único conjunto de dados de validação, tendo sido utilizado por diversos autores (GIANOLA *et al.*, 2011; VENTURA *et al.*, 2012). Como neste trabalho foram utilizados 120 conjuntos de dados, representativos de réplicas de um experimento base, para fins de validação da rede, uma alternativa foi adotar como critério de parada o número de acertos máximos obtidos pela RNA considerando as 120 réplicas. Neste, considerou-se como acerto o número de vezes em que o quadrado da correlação entre a resposta da rede neural e o valor genotípico verdadeiro foi superior ao quadrado da correlação entre a média fenotípica e o valor fenotípico considerando-se todas as combinações de neurônios, número camadas ocultas e funções de ativação utilizadas. Dessa forma, foi realizada uma comparação entre uma medida da confiabilidade da rede e a herdabilidade da característica, que é o critério utilizado de forma convencional para predição de ganhos e definição de estratégias de seleção.

Uma vez que a melhoria da eficiência de seleção ainda é um desafio compartilhado por melhoristas e biometristas, foi avaliada a eficiência da seleção por meio da capacidade da RNA para descartar um genótipo inferior e/ou selecionar um genótipo superior. Os resultados obtidos por meio das RNA's foram, pelo menos, tão bons quanto aqueles obtidos por meio das metodologias tradicionais, como as estimativas obtidas pelo método dos mínimos quadrados. Em 75% dos cenários a RNA mostrou-se mais eficaz que o método de quadrados mínimos na recuperação do valor genético (Tabela 2.1). As estimativas das acurácias de predição por meio da rede neural, considerando os 120 experimentos de validação, foram em média, 0.6916% superiores considerando as herdabilidades de 40, 45, 50, 55, 65 e de 70%. Para as herdabilidades de 60 e 80%, a rede neural propiciou acurácias inferiores, no entanto, muito próximas das obtidas pela metodologia de mínimos quadrados utilizada em modelos lineares. As porcentagens da superioridade da Rede Neural com relação aos Modelos Lineares encontram-se na Tabela 2.1.

Tabela 2.1. Médias das estimativas das acurácias obtidas por meio das redes (RNA) versus modelos lineares (ML), assim como a média da superioridade da rede (Sup.) para cada cenário.

Acurácias (%)	$h^2 = 40$	$h^2 = 45$	$h^2 = 50$	$h^2 = 55$	$h^2 = 60$	$h^2 = 65$	$h^2 = 70$	$h^2 = 80$
ML	0.6325	0.6708	0.7071	0.7416	0.7746	0.8062	0.8366	0.8944
RNA	0.6411	0.6711	0.7160	0.7432	0.7738	0.8128	0.8403	0.8911
Sup.(%)	1.3697	0.0447	1.2587	0.2157	--	0.8187	0.4423	--

Nota-se, ainda, uma relação praticamente linear entre os valores de herdabilidade avaliados neste trabalho e os valores de acurácia obtidos (Figura 2.1).

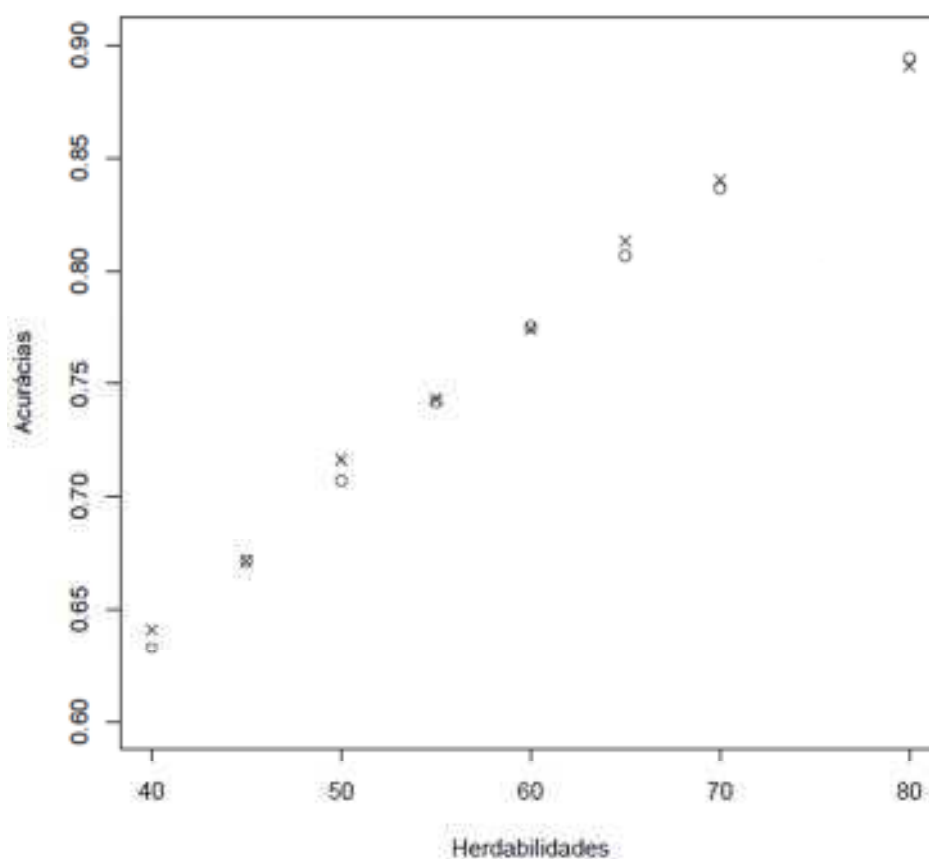


Figura 2.1. Gráfico da acurácia em função das herdabilidades

Pelos resultados obtidos, é notório que as Redes Neurais Artificiais apresentam grande potencial para sua utilização em estudos que envolvam procedimentos de predição. No entanto, os valores de acurácia encontrados não apresentaram, em nenhum dos oitos cenários avaliados, superioridade acima 1.5%, o que demonstra que os

resultados obtidos ainda se aproximam bastante dos resultados encontrados quando o pesquisador lança mão de procedimentos baseados no Método dos Mínimos Quadrados.

Em contrapartida, o paradigma vigente foi construído por meio da utilização de metodologias baseadas em Modelos lineares e os resultados obtidos quase que se equiparam aos resultados dessas metodologias. Assim, ainda é difícil avaliar a eficiência da Rede a partir de um modelo originalmente concebido admitindo ser a média do genótipo o melhor valor de sua superioridade perturbado apenas por ruídos ocasionados pelo efeito ambiental sobre os quais a biometria tenta estabelecer os melhores modelos de predição considerando sua distribuição estocástica.

Apesar da dificuldade em avaliar as Redes Neurais Artificiais, os estudos a partir destas representam um grande avanço acerca da utilização de processos de simulação de ensaios para a predição de genótipos superiores. Em programas de melhoramento, que envolvem estatística experimental, a simulação pode ser utilizada como uma alternativa para gerar ou replicar informações representativas das avaliações de um conjunto de genótipos em experimentos em delineamento apropriado, tais como o de blocos ao acaso em que são observados os princípios da repetição, casualização e controle local (PEIXOTO, 2013).

Processos de simulação têm sido cada vez mais utilizados nas diversas áreas da ciência. Coelho (2005) utilizou dados simulados para comparar diferentes métodos de estimação dos componentes de variância. Junior *et al.* (2010) utilizaram dados simulados para comparar as metodologias Eblup e Bayesiana, em dados com homogeneidade de variâncias, heterogeneidade de variância genética e heterogeneidade de variância genética e ambiental.

A comparação neste trabalho foi estabelecida em condições artificiais tendo em vista o uso de dados simulados. A utilização deste tipo de dados foi importante para o desenvolvimento do trabalho, pois agilizam o processo de comparação entre metodologias e ainda, o processo de treinamento da rede pode ser refinado quando o pesquisador tem em mãos um grande histórico de informações.

No presente trabalho, o paradigma do processo seletivo foi baseado na estimativa da média genotípica verdadeira que não está disponível em situações reais de experimentação. Portanto, outra vez foi de suma importância o uso de dados simulados, uma vez que este procedimento permite que o pesquisador obtenha, a partir das informações de coeficiente de variação, média e herdabilidade, os valores genéticos verdadeiros dos genótipos.

Após predição e ordenamento dos valores genéticos percebeu-se variação na classificação dos genótipos, o que poderia segundo Ventura (2012) acarretar em riscos no uso das redes para avaliação genética da característica. Na Tabela 2.2 estão apresentados estimativas de coincidência na classificação de genótipos selecionados – intensidades de 5 e 10% – e rejeitados – intensidade de 5 e 10% – com base nos reais valores genéticos e nos valores genéticos obtidos por meio da RNA e por meio de ML.

Tabela 2.2. Média percentual de coincidência de genótipos selecionados (Selec.) e rejeitados (Rej.) por meio de previsão através da RNA e dos modelos lineares (ML).

		Coincidência (%)							
		h ² = 40	h ² = 45	h ² = 50	h ² = 55	h ² = 60	h ² = 65	h ² = 70	h ² = 80
Selec. (5%)	ML	30.17	37.33	40.83	42.67	48	52	50.67	59.17
	RNA	34.17	36.67	44	42.67	47.17	52.33	52.33	59
	ML-RNA	69.67	93.33	85.5	89.67	90.67	93	91.83	91.67
Selec. (10%)	ML	38.58	43	44.58	46.08	51.92	53.42	57.92	67.58
	RNA	39.83	42.75	47.33	45.92	52.25	53.67	58.17	65.67
	ML-RNA	80.92	95.25	89.33	88.33	93.5	95.58	94.92	93.33
Rej. (20%)	ML	51.29	52.17	56.29	59	61.54	63.63	67.58	71.29
	RNA	52.33	52.92	56.46	59.29	60.92	64.38	67.67	71.08
	ML-RNA	89.63	89.13	97.04	93.58	95.21	96.04	98.25	94.25
Rej. (10%)	ML	42.5	46.42	48.92	52.42	54.58	59	62.33	67
	RNA	46	47.75	49.5	51.92	53.75	60	62.92	67.33
	ML-RNA	81.87	83.75	94.17	92.5	93.5	95.17	97.17	93.08

Em todos os casos, a RNA proporcionou coincidências médias superiores à proporcionadas pelo processo seletivo por meio de modelos lineares, indicando que as predições permitiram uma melhor discriminação dos genótipos com base no seu valor genético. Apesar de valores baixos, as coincidências de genótipos selecionados refletem a dificuldade em selecionar genótipos em uma característica quantitativa, sendo o efeito mais pronunciado quanto menor a herdabilidade da característica. Em contrapartida, constataram-se coincidências superiores a 69% quando comparados os genótipos selecionados pela RNA e pelo método tradicional evidenciando que, em sua maioria, os mesmos genótipos foram selecionados por ambas as abordagens.

CONCLUSÕES

As Redes Neurais Artificiais demonstraram eficiência em prever valores e ganhos genéticos em ensaios simulados sob o delineamento de blocos casualizados. Vale ressaltar, no entanto que, como todo o trabalho foi desenvolvido a partir de dados simulados, estudos subsequentes em populações reais de melhoramento são necessários para comprovação da superioridade das previsões da Rede Neural Artificial.

A estratégia para treinamento das Redes Neurais Artificiais apresentada consiste em uma nova abordagem de análise que pode ser utilizada para aumentar a eficiência do processo de seleção em programas de melhoramento genético.

A rede neural artificial demonstra superioridade em relação ao método tradicional de discriminar os genótipos com base no seu valor genético real, e ainda, o modelo ajustado para a rede neural artificial obriga o criador a considerar estatísticas adicionais, com alto teor prático, no processo de previsão do valor genético.

Em suma, a rede neural artificial apresenta grande potencial para ser utilizada como um método alternativo para prever valores genéticos e como ferramenta para a seleção genotípica.

REFERÊNCIAS BIBLIOGRÁFICAS

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais - Teoria e aplicações** – 2. ed. Rio de Janeiro: LTV, 226p, 2011

COELHO, A. M. **Simulação de dados visando à estimação de componentes de variância e coeficientes de herdabilidade**. Dissertação (Mestrado em Agronomia). Piracicaba, 131p, 2005.

CRUZ, C.D. GENES - A software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum**. v.35, n.3, p.271-276, 2013.

CRUZ, C. D. **Princípios de Genética Quantitativa**. 2. ed., Viçosa: UFV, 394p, 2010.

CRUZ, C.D. **Programas Genes: análise multivariada e simulação**. Viçosa, ed. UFV, 175p, 2006.

CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, Imprensa Universitária, 390p, 1994.

GARDNER, C. O. An evaluation of effects of mass selection and seed irradiation with thermal neutrons on yield of corn. **Crop Science**, Madison, v.1, n.2, p.124-245, 1961.

GIANOLA, D.; OKUT, H.; WEIGEL, K. A.; ROSA, G. J. M. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC Genetics**. p.12-87, 2011. <<http://www.biomedcentral.com/1471-2156/12/87>>

HAYKIN, S. S. **Redes Neurais: princípios e práticas**. Symon Haykin; trad Paulo Martins Engel. 2 ed. – Porto Alegre : Bookman, 900p, 2001.

JÚNIOR, J. M. C.; ASSIS, G. M. L.; EUCLYDES, R. F.; MARTINS, W. M. O.; WOLTER, P. F. Predição de valores genéticos utilizando inferência bayesiana e frequentista em dados simulados. **Acta Scientiarum**. Maringá, v.32, n.3, p.337-344, 2010.

MÉNDEZ, R., I. Refinamiento a la técnica de selección masal moderna. **Agrociência**, Montevideu, v.6, n.1, p.87-91, 1971

MOLINA, J. D. **Selección masal visual estratificada em maíz**. Chapingo: Centro de Genética, 36p, 1983. Publicación Especial.

NASCIMENTO, M.; PETERNELLI, L. A.; CRUZ, C. D.; NASCIMENTO, ANA CAROLINA CAMPANA; FERREIRA, R. P.; BHERING, L. L.; SALGADO, C. C. **Artificial neural networks for adaptability and stability evaluation in alfafa genotypes**. *Crop Breeding and Applied Biotechnology*, v.12, p.152-156, 2013.

PATTERSON, H.D.; THOMPSON, R. Recovery of inter block information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971.

PEIXOTO, L. A. **Redes Neurais Artificiais na predição do valor genético.** Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, 97p, 2013.

PETEK, M. R.; SERA, T.; FONSECA, I. C. B. Prediction of genetic additive values for development of a coffee cultivar with increased rust resistance. **Bragantia**, v.67, n.1, p.133-140, 2008.

RESENDE, M. D. V. **Matemática e Estatística na Análise de Experimentos e no Melhoramento Genético.** Colombo: Embrapa Florestas, 362p, 2007.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v.37, n.3, p.182-194, 2007.

SILVA, R.G. **Métodos de genética quantitativa aplicados ao melhoramento animal.** Ribeirão Preto, Sociedade Brasileira de Genética, p.162, 1982.

RESENDE, M. D. V.; HIGA, A. R. Maximização da eficiência da seleção em testes de progênies de *Eucalyptus* através da utilização de todos os efeitos do modelo matemático. **Boletim de Pesquisas Florestais**, Curitiba, n.28/29, p.37-56, 1994.

CAPÍTULO 2

THE USE OF NEURAL NETWORKS FOR PREDICTING BREEDING VALUES AND GENETIC GAINS

The use of neural networks for predicting breeding values and genetic gains

Abstract

Analysis with Artificial Neural Networks has been described as an approach in the decision-making process that, although incipient, has been reported as presenting high potential for use in animal and plant breeding. In this paper, we introduce the procedure of using the expanded data set for training the network, and we also proposed utilization of statistical parameters in addition to the mean phenotypic value in a feed-forward back propagation multilayer perceptron network in order to estimate the breeding value of genotypes in simulated scenarios. After evaluating different artificial neural network configurations, our results demonstrated its superiority in relation to estimation procedures based on linear models, as well as its applicability in the genetic value prediction process. The results further indicated the good generalization performance of the neural network model in several additional validation experiments.

Index terms: artificial neural network, simulation, training strategy.

Resumo

Redes Neurais artificiais têm sido descritas como uma abordagem no processo de tomada de decisão que, embora de forma incipiente, têm sido relatadas como de grande potencial para utilização no melhoramento genético animal e vegetal. Neste trabalho, apresentamos o procedimento de expansão de dados para ser utilizado no processo de treinamento da rede, e propomos a utilização parâmetros estatísticos adicionais à média fenotípica em uma rede neural perceptron multicamadas com o intuito de estimar o valor genético em experimentos simulados. Após avaliação de diferentes configurações de redes, os resultados demonstraram a superioridade desta em relação a processos de estimação baseados em modelos lineares, e sua aplicabilidade no processo de predição de valor genético. Os resultados indicaram boa capacidade de generalização do modelo ajustado pela rede validado em diversos experimentos adicionais.

Termos para indexação: rede neural artificial, simulação, estratégia de treinamento.

Introduction

Genetic breeding is a science with the objective of increasing the frequency of alleles and/or obtaining favorable genotypic combinations, in order to increase the production efficiency of an individual or population. According to Cruz *et al.* (2012), the possibility of predicting the gain obtained by traditional selection methods is one of the major contributions to quantitative genetics and biometrics for breeding, since this enables more effective guidance of the breeding program to predict the success of the selective scheme adopted and, based on science, decide on alternative techniques that may be more effective. Regarding the statistical genetic methods for selection, some methodology categories are noteworthy, been widely used: selection index (Martins *et al.*, 2005), the combined selection method (Costa *et al.*, 2000) and the REML-BLUP method (Ferreira *et al.*, 2012).

Nonetheless, a new paradigm can be employed in genetic breeding for selection purposes that does not involve stochastic modeling, but instead the principles of learning in a computational intelligence approach. In this context, approaches based on Artificial Neural Networks (ANN) have been described as an additional tool in the decision making process in various fields of science with great potential in animal and plant, in genetic breeding and for predicting quantitative traits (Castro *et al.*, 2013; Gianola *et al.*, 2011; Ventura *et al.*, 2012; Ventura *et al.*, 2013). Still, it is usual to train the ANN with a data subset, such as historical information of the genotypes, but this presents the disadvantage that they are not necessarily contemporaries of the evaluated experiment. An alternative would be to use a virtual data set, provided by statistical techniques that are able to preserve some features of the original experiment such as the mean, variances and covariances. We will call this the expanded data set, as a concept to be presented and discussed.

Therefore, the goal of this work was to evaluate the ability of the ANN as an alternative method for statistical studies and prediction of breeding values, using simulated data from other relevant statistics as well as the mean phenotypic value. A simulation procedure was used to provide information on the true genetic values for comparison of the ANN technique which may, in practical situations, be suitably replaced by historical values or by performance of advanced generations of these same genotypes. We investigated the performance of several ANN architectures for prediction of breeding values, and introduced the process of expanded data sets to be used in the ANN training process for subsequent prediction of the genetic value.

Material and methods

For clarity of presentation the methodology is presented first, where the main objective of this study was to evaluate the procedure of data expansion. This is followed by a description of two scenarios applied to an ANN for genetic prediction.

Data simulation

Two scenarios of simulated characteristics were considered, with heritabilities of 40 and 70%. The experiments consisted of a randomized block designs with 100 genotypes and six blocks, assuming mean values equal to 100 and coefficient of variation of 15%. One auxiliary file was obtained through a process known as data expansion for the purpose of training the network, which preserves the structure of means and covariance matrix of the test file. This expansion process was carried out as described below. The simulated values were considered a random variable $Y \sim N(\phi, \Sigma)$. The data was transformed into a random variable $Z \sim N(\phi, I)$ by means of the linear transformation $Z = F'Y$, where F was obtained by means of the spectral decomposition process of Σ , such that $\Sigma^{-1} = FF'$. The amplification process consists of simulating new values of Y , considering $Y \sim N(\phi, (F')^{-1}Z)$. A file of amplified data, consisting of 5000 genotypes, was considered for training of the ANN.

For simulation of a data set, or replication of a known data structure or even the expansion of a data set based on the structure of the other, it is required to have a data set with known distribution, and in principle a mean equal to zero and variance equal to V . In order for such requirements to be obeyed, the usual procedure is to resort to the Box Muller transform. For data simulation, the variables considered are: $x = \sqrt{-2\ln(RND)}V \cos(2\pi RND)$ and $y = \sqrt{-2\ln(RND)}V \sin(2\pi RND)$; in which RND is a random number.

To ensure that the data set X in fact has zero covariance, we recommend using the principal components technique. This technique is very useful for breeding programs because it permits the simplification of data sets to a limited set of components, which present the properties to retain as much of the originally available variation and mutually independent (Cruz, 2006).

The random term $Y \sim N(\phi, \Sigma)$ is that which we wish to transport into a random variable $Z \sim N(\phi, I)$. By means of the spectral decomposition process, we have $\Sigma^{-1} = FF'$. Thus, $(\Sigma^{-1})^{-1} = (FF')^{-1} = (F')^{-1}F^{-1} = \Sigma$.

If $Z = F'Y$, then $E(Z) = E(F'Y) = F'E(Y) = F'\phi = \phi$ and $V(Z) = F'V(Y)F = F'\Sigma F = F'[(F')^{-1}F^{-1}]F = I$.

Considering that $Z \sim N(\phi, I)$. If $Z = F'Y$, then $Y = (F')^{-1}Z$.

Therefore, $V(Y) = V((F')^{-1}Z) = (F')^{-1}V(Z)[(F')^{-1}]' = (F')^{-1}I(F')^{-1} = \Sigma$.

The data expansion process was performed by means of the integration module in the computer application GENES (Cruz, 2013).

Artificial neural network construction and evaluation

Feed-forward backpropagation multilayer perceptron networks were created in the Matlab software using the integration module in the computer application GENES (Cruz, 2013). The training algorithm *trainlm* was used, along with network architecture consisting of three hidden layers, activation functions *tansig* or *logsig*, number of neurons varying from one to seven and maximum number of iterations (or epochs) equal to 2000. All combinations of neuron numbers and activation functions in the hidden layers were verified. For training and validation of the ANN the desired outputs of the true genotype values were used and inputs presented the following information:

x_1 : mean phenotypic value;

x_2 : maximum value of each genotype in b blocks;

x_3 : minimum value of each genotype in b blocks;

x_4 : standard deviation of each genotype;

x_5 : coefficient of variation for each genotype;

x_6 : sum of ranks obtained in b blocks for each of the genotypes;

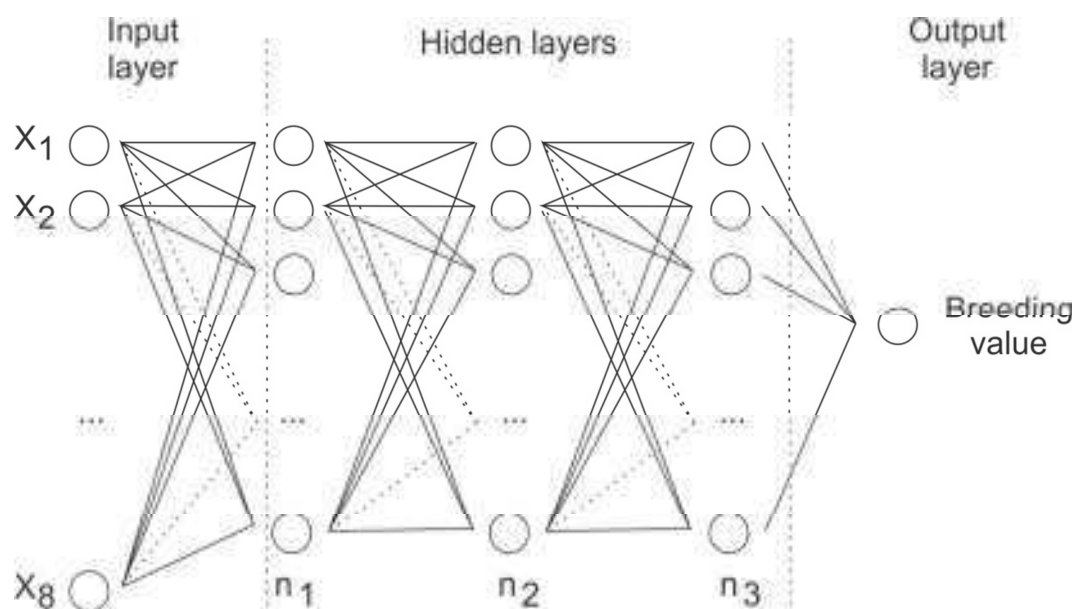
x_7 : statistic value P_i , adapted from Lin & Bins (1988) for studies of adaptability and stability, given by $\frac{\sum_{j=1}^b (X_{ij} - M_j)^2}{2b}$, where X_{ij} is the phenotypic value of the i -

thgenotype in the j -th block; M_j is the maximum response observed in all genotypes in block j ; and b us the number of blocks in the experiment;

x_8 : discrete values, coded 1 to 6, established from transformation of the mean phenotypic value considering its variation around the average within the limits, lower and upper, considered as one, two or three standard deviations.

Exemplified in Figure 1 is the architecture of the ANN.

Figure 1. Architecture of the ANN. Inputs (x_1) to (x_8) in the Input layer are related to the statistics estimated from data and considered as inputs. Three hidden layers consisting of n_i nodes ($i = 1, 2, \dots, 7$) were considered, with the activation functions tansig or logsig. All combinations were explored. The ANN returns the predicted breeding value in the output layer.



To verify the performance of the ANN, replicates of the test file were considered as validation files, also obtained by means of data simulation generated with the same number of blocks and genotypes, preserving the means and structure of variance and covariance between each pair of blocks. Thus, 120 validation files were used, considered as replicates of the simulated experiments. Considered as a criterion of optimization of the architecture of the ANN was the number of times in which reliability of the genotypic prediction by the network was higher than the mean phenotypic value. Reliability was given by the square of correlation between the response of the ANN and the true genotypic value compared with the square of the correlation between the mean phenotypic and true phenotypic values, which expresses the heritability of the trait, applied in all combinations of neurons, hidden layers and activation functions used. Thus, a direct comparison was made between the measure of network reliability and heritability of the trait, which is the criterion conventionally used for predicting genetic value (maximum likelihood estimator – MLE), genetic gain, identification of selective units and choosing the selection strategies.

Results and discussion

The neural networks were fitted to data in the training set, with connection strengths and biases modified iteratively. Oftentimes a data subset, such as historical information of genotypes, is used in such a procedure. However, some alternatives to this process have been presented. One of these is to use the validation data itself for the training process, so that the predictive ability of the network is gauged in parallel with this since they are similar in structure. In this study, we propose the use of an extended data set, which allows for preserving the characteristics of the experiment from which they are derived. The expansion process was performed using statistical methods which allowed the preservation of traits such as the mean, variance and covariance among information of the genotypes which were considered pairs of blocks from the original data. Additional procedures may be performed in order to preserve other traits of the original data set, however, the traits emphasized and preserved for the purposes of this study showed to be sufficient to test the network in its training phase, information that allowed for accurate prediction of genetic values and classification of genotypes.

In order to demonstrate that data expansion did not alter the structure of the original data set, Table 1 presents estimates of genetic parameters estimated for sets of simulated and amplified data sets, for heritability values of 40 and 70% ($\mu = 100$, $CV_e = 15\%$). The expansion process was effective in replicating genetic information from the original dataset, since important features such as genotypic variance, coefficient of genetic variation and heritability of the trait were preserved.

In conventional statistical procedures used in the genetic gain prediction process, only mean genotypic values are considered assuming the existence of their correlation with the actual genetic values of the genotypes, denoted by heritability of the trait. It is acknowledged that the mean is the main parameter needed to judge the superiority between genotypes, and in cases of low heritability, the selection process is of lower reliability. However, a more detailed analysis may be performed by taking into consideration not only the mean genotype, but also other relevant statistics and those of high practical aspect, although these make the decision making process more complex for the data analyst. Thus, two genotypes may have similar averages but significantly different behavior in terms of variation, ranking, maximum and minimum values and others. In the biometrics approach this additional information is neglected, but when utilizing the computational intelligence approach such information is indispensable and make up relevant inputs to be considered in the training process of the ANN.

Table 2. Estimates of genetic parameters derived from the analysis of variance of the simulated files and expanded data, for heritability scenarios of 40 and 70%. The data expansion process preserves the structure of the original data file.

Parameter estimates	$h^2 = 40\%$		$h^2 = 70\%$	
	Simulated	Expanded	Simulated	Expanded
μ	100.0000	100.0001	100.0000	99.9999
σ_g^2	25.0000	25.0019	87.5000	87.5004
σ_e^2	225.0000	225.0051	225.0000	225.0024
h^2	40.0000	40.0013	70.0000	69.9998
h_{ind}^2	10.0000	10.0005	28.0000	27.9998
$CV_e(\%)$	15	15	15	15
$CV_g(\%)$	5	5	9	9

* μ – mean; σ_g^2 – genetic variance; σ_e^2 – residual variance; h^2 – heritability at the level of the genotype mean; h_{ind}^2 – heritability at the level of the individual; CV_e – coefficient of experimental variation; CV_g – coefficient of genetic variation.

When considering the overall values of diverse statistics of each genotype obtained in a determined assay, it is observed that more reliable criteria are obtained to predict genotypic superiority than be based solely on the mean value obtained for genotype assays whose genotypic variability and experimental precision can vary tremendously. In order to verify how the amplification process affected the structure of the expanded experiment, a correlation analysis was performed between the statistics used in the input layer and the actual breeding values (Table 2). The correlations were statistically equal, considering a t-test ($p = 0.01$). Moreover, although the statistics used in the input layer showed to be sufficient to predict the genetic value, additional studies are needed to quantify the influence of each in the prediction process, as well as evaluate the removal or inclusion of statistics.

The same evaluation was performed considering the simulated data file and the 120 replicated files used in the validation process, where a similar result was obtained (data not show). The reasoning behind this step is that the researcher could question the veracity of the validation process, since sometimes this can be done with the training file or only part of it. The replicated files serve to emulate situations in which the created ANN is used in contemporary prediction experiments of the genotypes evaluated.

Table 2.Correlation between the genetic values (GV) and statistic values estimated for the simulated and expanded data sets.

Square of the Correlation coefficient	$h^2 = 40\%$		$h^2 = 70\%$	
	Simulated	Expanded	Simulated	Expanded
GV x mean	0.4001**	0.4001**	0.7001**	0.7001**
GV x max	0.2928*	0.2253**	0.5512**	0.5253**
GV x min	0.2417*	0.2552**	0.5039**	0.4982**
GV x sd	0.0015 ^{ns}	0.0000 ^{ns}	0.0015 ^{ns}	0.0022 ^{ns}
GV x CV	0.0110 ^{ns}	0.0323 ^{ns}	0.0570 ^{ns}	0.0978**
GV x Pi	0.3623**	0.3923**	0.6696**	0.691**
GV x Rank	0.3848**	0.3772**	0.6871**	0.6752**
GV x class	0.4290**	0.3624**	0.6395**	0.641**

* GV– Genetic value; max – Maximum value; Min – minimum value, sd – standard deviation; Rank – sum of ranks; class – classification.

The usual procedure for determining the stop criterion of the ANN is the mean squared error of prediction. Once the mean squared error of prediction reaches an optimal level, training stopped and this led to the best estimates of the network coefficients (Braga *et al.*, 2011; Haykin, 1998). This stop criterion is usually applied when the validation process of the ANN is obtained with one validation file, and has been used by several authors (Gianola *et al.* 2011;Timm *et al.*, 2006; Ventura *et al.* 2012). Because in this work 120 validation files were used, an alternative was adopted in which the stop criterion was the maximum number of hits obtained by the ANN considering the 120 replicates. Herein, we assessed the number of times in which the square of the correlation between the response of the ANN and the true genotypic value was higher than the square of the correlation between the mean phenotypic value and the phenotypic value considering all combinations of neurons, number of hidden layers and activation functions used. Thus, a comparison between the measure of network reliability and heritability of the trait was performed, which is the criterion used in the conventional method for predicting gain, identifying the selective units and choosing the selection strategies.

Because improving selection efficiency still is a challenge shared to breeders and biometricians, selection efficiency was evaluated by means of the ability of the ANN to discard a genotype that would eventually have a low genetic value and/or select a genotype that would have a high genetic value. Our results with the ANN are as good as or better than those obtained with traditional methodologies, based on least squares

estimates. However, this comparison is unfair since the current paradigm of the selection process is based on an estimate of the mean genotype. For this reason, subsequent studies on real breeding populations are needed to prove the superiority of ANN predictions. In both cases, recovery of the genetic value as performed by the ANN was more effective than that based on the mean genotypic value, estimated by maximum likelihood (MLE) in both simulated scenarios (Table 3). Estimates of prediction accuracies by the ANN, considering the 120 validation experiments, were on average 1% and 0.5% higher considering the heritabilities of 40 and 70%, respectively.

Table 3. Mean percentage of coincidence for selected and rejected genotypes by means of prediction via the ANN and linear models, and accuracy of prediction.

		Coincidence (%)	
		$h^2 = 40\%$	$h^2 = 70\%$
Selected (5%)	MLE	30	51
	ANN	34	52
	MLE-ANN	70	92
Selected (10%)	MLE	39	58
	ANN	40	58
	MLE-ANN	81	95
Rejected (20%)	MLE	51	68
	ANN	52	68
	MLE-ANN	90	98
Rejected (10%)	MLE	43	62
	ANN	46	63
	MLE-ANN	82	97
Accuracy	MLE	63	84
	ANN	64	84

* MLE – selection by maximum likelihood estimator compared with the true genetic value, calculated by the software GENES; ANN – Artificial Neural Network selection compared with the true genetic value; MLE-ANN – selection by maximum likelihood estimator compared with Artificial Neural Network selection.

After prediction and ordering of the genetic values a variation in classification of the genotypes was perceived, which according to Ventura (2012) may result in risks for use of the ANN for genetic evaluation of the trait. Table 3 presents estimates of coincidence in classification of the selected– intensities of 5 to 10% – and rejected

genotypes – intensity of 10 and 20% – based on real breeding values and genetic values obtained by the ANN and by MLE.

In all cases, the ANN presented coincidences greater than those obtained by the selection process from linear models, indicating that the predictions allow a better discrimination of genotypes based on their genetic value. Despite the low values, coincidences of selected genotypes reflect the difficulty in selecting genotypes for a quantitative trait, where this effect is more pronounced for lower heritability values. In contrast, coincidences greater than 69% when comparing the genotypes selected by the ANN and the traditional method indicated that, for the most part, the same genotypes were selected by both approaches.

According to Gianola *et al.* (2011), an important question in animal and plant breeding is how an estimated “breeding value”, i.e., an estimate of the total additive genetic effect of an individual, can be arrived at from an ANN output. In their study involving prediction of the genetic value of a quantitative trait, Gianola *et al.* (2011) suggested two approaches for modeling such effects. It is clear, however, that such propositions were made considering a model based on Bayesian neural networks using information of molecular markers – genomic selection. Applied to our context, although of extreme importance to the breeder, it is necessary to understand and use information concerning the relationship between the individuals, a scenario which although relevant, was not assessed herein. Therefore, it is of great importance that more studies be conducted to permit the development and application of ANN in the field of genetic breeding.

Conclusions

The ANN demonstrated superiority over the traditional method to discriminate genotypes based on their real genetic value. Also, the biometrician possesses a range of relevant statistical information, but these are underused in the decision making process. The ANN adjusted model was able to take into account these additional statistics, with high practical content, in the genetic value prediction process. Results obtained using simulated data show that the ANN has great potential for use as an alternative method to predict genetic values and as a tool for genotypic selection.

We demonstrated the potential of the ANN model for use as a decision support tool for selection in situations which historical data is not available by using data

expansion procedures. The ANN consisted of a new analytical approach that can be used to increase selection efficiency in breeding programs.

Acknowledgements

To CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – , FAPEMIG – Fundação de Amparo a Pesquisa do Estado de Minas Gerais – and CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, for financial support and scholarships granted.

References

- Braga, A.P.; Carvalho, A.P.L.F.; Ludermir, T.B. 2011. Artificial Neural Networks. Theory and applications. 2ed. LTV, Rio de Janeiro, RJ, Brazil (in Portuguese).
- Castro, R.V.O.; Soares, C.P.B.; Martins, F.B.; Leite, H.G. 2013. Growth and yield of commercial plantations of eucalyptus estimated by two categories of models. *Pesquisa Agropecuária Brasileira* 48: 287-295 (in Portuguese, with abstract in English).
- Costa, R. B.; Resende, M. D. V., Araujo, A. J.; Gonçalves, P. S.; Bortoletto, N. 2000. Univariate and multivariate combined selection applied to hevea breeding. *Pesquisa Agropecuária Brasileira* 35: 381-388 (in Portuguese, with abstract in English).
- Cruz, C.D. 2013. GENES: a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum* 35: 271-276.
- Cruz, C.D.; Carneiro, P.C.S. 2006. Biometric models applied to genetic improvement. 2ed. UFV, Viçosa, MG, Brazil (in Portuguese).
- Cruz, C.D. Genes program: multivariate analysis and simulation. 2006. UFV, Viçosa, MG, Brazil (in Portuguese).
- Cruz, C.D.; Regazzi, A.; Carneiro, P.C.S. 2012. Biometric models applied to genetic improvement. 4ed. UFV, Viçosa, MG, Brazil (in Portuguese).
- Ferreira, R.T.; Viana, A.P.; Barroso, D.G.; Resende, M.D.V.; Júnior, A.T.A. 2012. *Toonaciliata* genotype selection with the use of individual BLUP with repeated measures. *Scientia Agricola* 69: 210-216.
- Gianola, D.; Okut, H.; Weigel, K.A.; Rosa, G.J.M. 2011. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. *BMC Genetics* 12: 87.

- Haykin, S. 1998. Neural Networks: A Comprehensive Foundation. 2ed. BOOKMAN Upper Saddle River, NJ, United States of America.
- Lin, C.S.; Binns, M.R.A. 1988. Superiority measure of cultivar performance for cultivar x location data. Canadian Journal of Plant Science 68: 193-198.
- Martins, I.S.; Cruz, C.D.; Rocha, M.G.B.; Regazzi, A. J.; Pires, I.E. Comparison between and within family selections with the combined selection in progeny tests of *Eucalyptus grandis*. 2005. Cerne 11: 16-24 (in Portuguese, with abstract in English).
- MATLAB version 7.10.0. 2010. Natick, Massachusetts: The Math Works Inc.
- Timm, L.C.; Gomes, D.T.; Barbosa, E.P.; Reichardt, K.; Souza, M.D.; Dynia, J.F. 2006. Neural network and state-space models for studying relationships among soil properties. Scientia Agricola 63: 386-395.
- Ventura, R.V.; Silva, M.A.; Medeiros, T.H.; Dionello, N.L.; Madalena, F.E.; Fridrich, A.B.; Valente, B.D.; Santos, G.G.; Freitas, L.S.; Wenceslau, R.R.; Felipe, V.P.S.; Corrêa, G.S.S. 2012. Use of artificial neural networks in breeding values prediction for weight at 205 days in Tabapuã beef cattle. Arquivo Brasileiro de Medicina Veterinária e Zootecnia 64: 411-418 (in Portuguese, with abstract in English).

ANEXO 1: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 40\%$

Arquivo	μ	max	min	Sd	cv	PI	rank	classe
1	0.4000	0.3954	0.3300	0.0726	0.0023	0.3694	0.4560	0.3033
2	0.4000	0.2168	0.1370	0.0002	0.0399	0.3755	0.3960	0.3548
3	0.4000	0.3048	0.1817	0.0001	0.0163	0.3334	0.3691	0.3728
4	0.4000	0.1232	0.3109	0.0259	0.0989	0.4561	0.3650	0.3893
5	0.4000	0.2234	0.1940	0.0020	0.0108	0.3584	0.3569	0.3786
6	0.4000	0.4368	0.1639	0.0260	0.0025	0.3101	0.3822	0.3811
7	0.4000	0.3927	0.2489	0.0064	0.0185	0.3478	0.4049	0.2806
8	0.4000	0.4137	0.1681	0.0331	0.0002	0.3523	0.3617	0.3374
9	0.4000	0.3568	0.1834	0.0045	0.0116	0.3516	0.3452	0.3271
10	0.4000	0.1167	0.3786	0.0588	0.1678	0.4157	0.4180	0.3489
11	0.4000	0.1864	0.1054	0.0188	0.0000	0.3851	0.3960	0.3826
12	0.4000	0.2440	0.2202	0.0010	0.0306	0.4127	0.4280	0.3228
13	0.4000	0.4068	0.2420	0.0116	0.0028	0.3562	0.2857	0.3177
14	0.4000	0.3260	0.1068	0.0189	0.0000	0.3440	0.4248	0.2727
15	0.4000	0.3334	0.5300	0.0441	0.1678	0.4009	0.3468	0.2997
16	0.4000	0.2564	0.1135	0.0122	0.0008	0.3471	0.4368	0.3936
17	0.4000	0.1935	0.2542	0.0112	0.0913	0.3971	0.3896	0.3596
18	0.4000	0.2360	0.2243	0.0272	0.1111	0.3686	0.4642	0.3740
19	0.4000	0.4224	0.1404	0.1154	0.0242	0.3593	0.4334	0.3194
20	0.4000	0.3145	0.2341	0.0100	0.0155	0.3599	0.4330	0.3339
21	0.4000	0.2129	0.1288	0.0173	0.0019	0.3396	0.3952	0.3722
22	0.4000	0.3818	0.2611	0.0381	0.0113	0.3914	0.4152	0.3270
23	0.4000	0.1479	0.4407	0.0076	0.0676	0.4399	0.4124	0.3554
24	0.4000	0.2640	0.1460	0.0004	0.0221	0.3753	0.3687	0.2937
25	0.4000	0.0989	0.2341	0.0418	0.1291	0.4090	0.4515	0.3786
26	0.4000	0.2756	0.1563	0.0605	0.0029	0.3607	0.4170	0.3200
27	0.4000	0.2784	0.2146	0.0001	0.0454	0.3544	0.3049	0.3191
28	0.4000	0.2244	0.2102	0.0092	0.0104	0.3618	0.3477	0.3753
29	0.4000	0.3043	0.1892	0.0116	0.0056	0.3523	0.3735	0.2890
30	0.4000	0.2586	0.2465	0.0026	0.0644	0.3749	0.3720	0.3564
31	0.4000	0.2595	0.2657	0.0034	0.0382	0.3720	0.3907	0.3328
32	0.4000	0.1156	0.2524	0.0000	0.0654	0.4022	0.4088	0.3697
33	0.4000	0.3725	0.1544	0.0526	0.0022	0.3613	0.3975	0.3181
34	0.4000	0.2798	0.3978	0.0029	0.1027	0.3906	0.3708	0.2382
35	0.4000	0.3006	0.1228	0.1364	0.0344	0.3530	0.4249	0.3581
36	0.4000	0.4152	0.2626	0.0056	0.0174	0.3616	0.3575	0.3229
37	0.4000	0.3673	0.1569	0.0001	0.0191	0.3413	0.4869	0.3986
38	0.4000	0.2414	0.1379	0.0006	0.0338	0.3673	0.4845	0.3671
39	0.4000	0.3839	0.2232	0.0000	0.0458	0.3507	0.3208	0.3644
40	0.4000	0.2750	0.1945	0.0021	0.0199	0.3604	0.3920	0.3631
41	0.4000	0.2068	0.3091	0.0045	0.0180	0.4298	0.4066	0.3019

42	0.4000	0.1729	0.3341	0.0034	0.0477	0.4111	0.3769	0.3379
43	0.4000	0.1982	0.2405	0.0123	0.0116	0.3788	0.3651	0.3217
44	0.4000	0.3124	0.3181	0.0089	0.0344	0.3975	0.3420	0.3470
45	0.4000	0.3295	0.1790	0.0076	0.0132	0.3371	0.3394	0.3556
46	0.4000	0.1300	0.1527	0.0053	0.0457	0.3728	0.4723	0.3837
47	0.4000	0.3768	0.3161	0.0090	0.0169	0.3744	0.3826	0.3637
48	0.4000	0.1611	0.1710	0.0071	0.0188	0.3696	0.4507	0.3895
49	0.4000	0.2176	0.3103	0.0001	0.0305	0.3776	0.4215	0.3273
50	0.4000	0.3029	0.4052	0.0063	0.0313	0.3826	0.3921	0.3695
51	0.4000	0.4083	0.2128	0.0058	0.0655	0.3434	0.3466	0.3248
52	0.4000	0.2209	0.2592	0.0093	0.0632	0.3919	0.4142	0.3911
53	0.4000	0.2480	0.1730	0.0116	0.0219	0.3633	0.4512	0.3842
54	0.4000	0.3004	0.4030	0.0065	0.0977	0.4145	0.3362	0.3325
55	0.4000	0.3711	0.2735	0.0396	0.0062	0.3674	0.3548	0.3473
56	0.4000	0.1993	0.1357	0.0001	0.0185	0.3627	0.4037	0.3485
57	0.4000	0.3746	0.0791	0.0184	0.0010	0.3230	0.4630	0.3340
58	0.4000	0.1364	0.1994	0.0025	0.0090	0.4058	0.4112	0.4084
59	0.4000	0.2872	0.0536	0.1248	0.0547	0.3516	0.4067	0.3373
60	0.4000	0.1942	0.3745	0.0005	0.0677	0.3991	0.4103	0.3713
61	0.4000	0.1670	0.1985	0.0092	0.0838	0.3665	0.3486	0.3101
62	0.4000	0.2716	0.2344	0.0353	0.0004	0.3712	0.3975	0.2882
63	0.4000	0.1702	0.2647	0.0004	0.0220	0.3986	0.4982	0.3700
64	0.4000	0.1122	0.3806	0.0527	0.1580	0.4465	0.3570	0.3464
65	0.4000	0.2432	0.3019	0.0086	0.0108	0.3850	0.3496	0.3341
66	0.4000	0.2038	0.1824	0.0013	0.0232	0.3580	0.4252	0.3770
67	0.4000	0.1280	0.3144	0.0868	0.1847	0.4053	0.3883	0.3714
68	0.4000	0.1994	0.2051	0.0142	0.0038	0.3494	0.3895	0.4100
69	0.4000	0.2085	0.2514	0.0000	0.0190	0.3882	0.4891	0.3189
70	0.4000	0.3642	0.1008	0.0411	0.0025	0.3399	0.4185	0.3942
71	0.4000	0.2123	0.2953	0.0004	0.0307	0.4139	0.3918	0.3168
72	0.4000	0.0964	0.3290	0.0458	0.1545	0.4310	0.4236	0.3833
73	0.4000	0.2005	0.3438	0.0148	0.0912	0.4063	0.3854	0.1858
74	0.4000	0.1790	0.2102	0.0073	0.0514	0.3882	0.4681	0.3349
75	0.4000	0.2353	0.3385	0.0124	0.0910	0.3842	0.3848	0.2864
76	0.4000	0.2215	0.2855	0.0011	0.0371	0.3865	0.4172	0.4067
77	0.4000	0.2331	0.0902	0.0586	0.0085	0.3630	0.5130	0.2113
78	0.4000	0.4388	0.2484	0.0044	0.0657	0.3698	0.3682	0.3714
79	0.4000	0.1673	0.3630	0.0704	0.1758	0.4129	0.3303	0.3866
80	0.4000	0.1618	0.2398	0.0056	0.0549	0.3724	0.3623	0.1707
81	0.4000	0.1741	0.2738	0.0030	0.0258	0.3983	0.4476	0.3481
82	0.4000	0.1092	0.3631	0.0459	0.1744	0.4508	0.3979	0.3104
83	0.4000	0.2981	0.3991	0.0045	0.0581	0.3906	0.3717	0.3951
84	0.4000	0.3207	0.0731	0.0750	0.0110	0.3021	0.4519	0.2678
85	0.4000	0.4404	0.1679	0.0429	0.0001	0.3256	0.3645	0.3183
86	0.4000	0.2307	0.2653	0.0000	0.0405	0.3923	0.4618	0.3853
87	0.4000	0.2271	0.2287	0.0023	0.0168	0.3758	0.4367	0.3833
88	0.4000	0.2208	0.0925	0.0262	0.0000	0.3608	0.3885	0.4363
89	0.4000	0.1563	0.3287	0.0159	0.0885	0.3995	0.3718	0.3240

90	0.4000	0.3114	0.1799	0.0106	0.0620	0.3616	0.3554	0.3640
91	0.4000	0.2540	0.2166	0.0019	0.0372	0.3787	0.3797	0.2745
92	0.4000	0.3777	0.2687	0.0300	0.0034	0.3595	0.3819	0.3926
93	0.4000	0.2166	0.2473	0.0031	0.0584	0.3689	0.4134	0.3542
94	0.4000	0.2226	0.2030	0.0041	0.0648	0.3949	0.3808	0.4485
95	0.4000	0.2987	0.2153	0.0001	0.0377	0.3780	0.3591	0.2950
96	0.4000	0.2750	0.3481	0.0507	0.1694	0.4011	0.3370	0.2952
97	0.4000	0.2823	0.1032	0.0450	0.0000	0.3553	0.4532	0.3680
98	0.4000	0.2321	0.4016	0.0013	0.0500	0.4181	0.3870	0.3449
99	0.4000	0.2666	0.4918	0.0190	0.1844	0.4110	0.3776	0.3926
100	0.4000	0.1638	0.1592	0.0078	0.0499	0.3732	0.3903	0.3455
101	0.4000	0.2999	0.3746	0.0167	0.1122	0.3870	0.3711	0.4065
102	0.4000	0.1908	0.3145	0.0127	0.1251	0.4197	0.3884	0.3099
103	0.4000	0.3100	0.2212	0.0010	0.0496	0.3876	0.3487	0.3739
104	0.4000	0.2067	0.2938	0.0000	0.0377	0.3863	0.4068	0.3155
105	0.4000	0.1403	0.1464	0.0046	0.0078	0.3954	0.4596	0.3688
106	0.4000	0.2055	0.2429	0.0005	0.0453	0.3835	0.4461	0.2866
107	0.4000	0.3918	0.1704	0.0467	0.0039	0.3804	0.4032	0.2401
108	0.4000	0.2606	0.1428	0.0084	0.0012	0.3766	0.4441	0.2013
109	0.4000	0.2553	0.1698	0.0028	0.0622	0.3456	0.3272	0.2777
110	0.4000	0.2760	0.1651	0.0003	0.0246	0.3627	0.3511	0.3866
111	0.4000	0.4202	0.0975	0.0915	0.0103	0.3039	0.3693	0.3869
112	0.4000	0.2085	0.2469	0.0036	0.0718	0.4052	0.4309	0.3656
113	0.4000	0.2514	0.2713	0.0000	0.0382	0.4004	0.3717	0.3528
114	0.4000	0.3475	0.2744	0.0006	0.0508	0.3833	0.3580	0.3412
115	0.4000	0.2491	0.4477	0.0186	0.1121	0.4175	0.3699	0.3714
116	0.4000	0.3349	0.1324	0.0119	0.0035	0.3386	0.4227	0.4135
117	0.4000	0.2086	0.1611	0.0270	0.0001	0.3631	0.3642	0.3375
118	0.4000	0.2086	0.1223	0.0094	0.0112	0.3550	0.4452	0.2826
119	0.4000	0.2763	0.3357	0.0000	0.0598	0.4031	0.3948	0.3966
120	0.4000	0.2841	0.2380	0.0076	0.0106	0.3942	0.4149	0.3849

ANEXO 2: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 45\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.4500	0.1565	0.2284	0.0137	0.0978	0.4215	0.4777	0.4137
2	0.4500	0.2561	0.2856	0.0099	0.0917	0.4216	0.4756	0.3772
3	0.4500	0.1074	0.2841	0.0228	0.1595	0.4577	0.4949	0.4461
4	0.4500	0.4904	0.1300	0.0853	0.0066	0.3705	0.4267	0.3436
5	0.4500	0.2642	0.1732	0.0001	0.0281	0.3950	0.4149	0.3719
6	0.4500	0.3818	0.3004	0.0054	0.0164	0.4029	0.4502	0.4330
7	0.4500	0.3793	0.2250	0.0001	0.0367	0.4006	0.3931	0.2898
8	0.4500	0.1933	0.2656	0.0409	0.1377	0.4360	0.4693	0.2579
9	0.4500	0.1186	0.3412	0.0257	0.1227	0.4290	0.4325	0.3768
10	0.4500	0.3822	0.3408	0.0063	0.0863	0.4221	0.4549	0.4272
11	0.4500	0.3572	0.4107	0.0027	0.0469	0.4530	0.4330	0.3444
12	0.4500	0.2117	0.2014	0.0002	0.0388	0.4261	0.4156	0.3891
13	0.4500	0.2159	0.3223	0.0344	0.1419	0.4420	0.4519	0.4189
14	0.4500	0.2137	0.3667	0.0673	0.1975	0.4566	0.4296	0.3900
15	0.4500	0.1407	0.4207	0.0351	0.1753	0.4911	0.5562	0.4362
16	0.4500	0.2707	0.4053	0.0140	0.1248	0.4560	0.3688	0.3436
17	0.4500	0.2356	0.2418	0.0147	0.0880	0.4203	0.4414	0.1347
18	0.4500	0.2830	0.2457	0.0329	0.1039	0.4132	0.4250	0.4147
19	0.4500	0.2874	0.3243	0.0088	0.0895	0.4232	0.4107	0.3475
20	0.4500	0.4348	0.3200	0.0293	0.0003	0.4253	0.4382	0.3601
21	0.4500	0.2215	0.4059	0.0056	0.0611	0.4364	0.4038	0.2534
22	0.4500	0.2017	0.3596	0.0228	0.0991	0.4557	0.3977	0.2582
23	0.4500	0.3176	0.3338	0.0002	0.0821	0.4431	0.4181	0.3118
24	0.4500	0.1137	0.4909	0.0988	0.2778	0.4927	0.4671	0.3851
25	0.4500	0.3878	0.2057	0.0001	0.0405	0.4041	0.4141	0.4113
26	0.4500	0.1163	0.3408	0.0410	0.1143	0.4510	0.4168	0.3801
27	0.4500	0.1687	0.3621	0.0123	0.0925	0.4635	0.3849	0.2788
28	0.4500	0.1822	0.3061	0.0293	0.1259	0.4636	0.5276	0.4022
29	0.4500	0.3080	0.3311	0.0025	0.0605	0.4328	0.4274	0.3943
30	0.4500	0.1596	0.2507	0.0120	0.0989	0.4509	0.4559	0.3526
31	0.4500	0.2815	0.2981	0.0000	0.0413	0.4402	0.4502	0.4455
32	0.4500	0.3564	0.2073	0.0383	0.0000	0.3942	0.4795	0.3929
33	0.4500	0.2212	0.3037	0.0027	0.0851	0.4652	0.4516	0.4160
34	0.4500	0.2652	0.2724	0.0045	0.0889	0.4427	0.4249	0.3768
35	0.4500	0.2027	0.4050	0.0185	0.1169	0.4672	0.4366	0.3792
36	0.4500	0.4702	0.1354	0.0201	0.0010	0.3810	0.4920	0.4267
37	0.4500	0.1655	0.4093	0.0108	0.0733	0.4776	0.4912	0.3604
38	0.4500	0.3141	0.3313	0.0002	0.0486	0.4161	0.4519	0.4398
39	0.4500	0.3756	0.3139	0.0000	0.0568	0.4101	0.3950	0.4273
40	0.4500	0.1487	0.2515	0.0599	0.2233	0.4589	0.3826	0.4345
41	0.4500	0.3411	0.3610	0.0080	0.0782	0.4351	0.4567	0.4056
42	0.4500	0.6368	0.1714	0.0282	0.0003	0.3662	0.4554	0.4114

43	0.4500	0.3018	0.1814	0.0024	0.0813	0.4129	0.4616	0.4068
44	0.4500	0.3162	0.4037	0.0051	0.0282	0.4310	0.4055	0.4360
45	0.4500	0.1980	0.3395	0.1119	0.2473	0.4456	0.4328	0.3336
46	0.4500	0.3632	0.1379	0.1108	0.0205	0.3535	0.4190	0.4048
47	0.4500	0.2385	0.2155	0.0003	0.0396	0.4309	0.4142	0.4390
48	0.4500	0.3177	0.1395	0.0070	0.0140	0.3760	0.4705	0.4016
49	0.4500	0.4808	0.3170	0.0069	0.0643	0.4378	0.4408	0.4366
50	0.4500	0.3182	0.1895	0.0000	0.0508	0.4099	0.4845	0.3871
51	0.4500	0.2704	0.2974	0.0016	0.0435	0.4170	0.4316	0.3737
52	0.4500	0.3284	0.2840	0.0008	0.0518	0.4194	0.4741	0.4004
53	0.4500	0.4496	0.2566	0.0018	0.0598	0.3973	0.4490	0.4376
54	0.4500	0.1794	0.4452	0.0160	0.0986	0.4853	0.4414	0.4048
55	0.4500	0.3428	0.3519	0.0000	0.0462	0.4247	0.4619	0.4029
56	0.4500	0.2658	0.3722	0.0261	0.1540	0.4395	0.3913	0.3336
57	0.4500	0.1997	0.2872	0.0123	0.0743	0.4299	0.4899	0.3970
58	0.4500	0.3054	0.4013	0.0199	0.0026	0.4541	0.4355	0.3632
59	0.4500	0.3399	0.3971	0.0008	0.0887	0.4658	0.4421	0.3399
60	0.4500	0.2823	0.3760	0.0162	0.1106	0.4722	0.4185	0.4008
61	0.4500	0.3505	0.2892	0.0096	0.0916	0.4310	0.3083	0.3822
62	0.4500	0.4693	0.1375	0.2454	0.1067	0.3498	0.4123	0.4222
63	0.4500	0.2873	0.4010	0.0216	0.1723	0.4335	0.4411	0.3423
64	0.4500	0.2752	0.2045	0.0036	0.0610	0.4109	0.4557	0.3226
65	0.4500	0.2345	0.3769	0.0011	0.0301	0.4453	0.4656	0.3823
66	0.4500	0.2856	0.2011	0.0502	0.0001	0.4285	0.4172	0.2994
67	0.4500	0.2948	0.1762	0.0459	0.0016	0.4160	0.5236	0.3595
68	0.4500	0.3466	0.4157	0.0293	0.0099	0.4458	0.4032	0.3104
69	0.4500	0.3033	0.2260	0.0036	0.0191	0.4005	0.4156	0.3024
70	0.4500	0.2898	0.2640	0.0004	0.0213	0.4575	0.3741	0.3932
71	0.4500	0.2373	0.2894	0.0037	0.0498	0.4270	0.3729	0.2781
72	0.4500	0.2404	0.2041	0.0001	0.0407	0.4263	0.4152	0.4205
73	0.4500	0.2048	0.2080	0.0057	0.0147	0.3850	0.4159	0.4716
74	0.4500	0.1444	0.4597	0.2253	0.3205	0.4647	0.4082	0.4159
75	0.4500	0.4357	0.1839	0.0040	0.0117	0.3961	0.4787	0.4139
76	0.4500	0.1728	0.3714	0.1892	0.3796	0.4641	0.4651	0.4008
77	0.4500	0.4697	0.2861	0.0572	0.0004	0.3877	0.3977	0.3729
78	0.4500	0.3189	0.2943	0.0006	0.0456	0.4283	0.4650	0.4253
79	0.4500	0.2395	0.3795	0.0041	0.0573	0.4640	0.4574	0.3703
80	0.4500	0.4199	0.1840	0.0004	0.0399	0.3817	0.5020	0.4220
81	0.4500	0.5355	0.2305	0.0019	0.0272	0.3949	0.4713	0.3537
82	0.4500	0.2252	0.2984	0.0009	0.0528	0.4409	0.4767	0.3561
83	0.4500	0.3062	0.1421	0.0001	0.0398	0.3816	0.4986	0.3855
84	0.4500	0.2527	0.1853	0.0081	0.0453	0.3921	0.4052	0.3203
85	0.4500	0.3280	0.4010	0.0190	0.0059	0.3984	0.3489	0.3870
86	0.4500	0.3217	0.3465	0.0085	0.0130	0.4152	0.4584	0.3851
87	0.4500	0.2813	0.4942	0.0012	0.0214	0.4663	0.4024	0.3915
88	0.4500	0.1252	0.3622	0.0307	0.1036	0.4555	0.4379	0.2935
89	0.4500	0.2112	0.4386	0.0866	0.2286	0.4841	0.4270	0.4118
90	0.4500	0.2720	0.2536	0.0049	0.0352	0.4327	0.4546	0.4473

91	0.4500	0.3817	0.1992	0.0357	0.0004	0.3721	0.3614	0.3941
92	0.4500	0.3665	0.2204	0.0510	0.0007	0.4001	0.4478	0.4295
93	0.4500	0.3762	0.4837	0.0030	0.1053	0.4346	0.4112	0.2838
94	0.4500	0.5551	0.1861	0.0012	0.0129	0.3696	0.4742	0.3819
95	0.4500	0.2957	0.3662	0.0079	0.0244	0.4008	0.3659	0.3680
96	0.4500	0.2368	0.1865	0.0008	0.0271	0.4142	0.4795	0.2341
97	0.4500	0.3168	0.3118	0.0055	0.0248	0.4225	0.4643	0.4107
98	0.4500	0.2401	0.3946	0.0096	0.0821	0.4724	0.4035	0.4010
99	0.4500	0.1804	0.4136	0.0080	0.0977	0.4972	0.4584	0.3835
100	0.4500	0.3044	0.1734	0.0026	0.0195	0.4020	0.5591	0.3704
101	0.4500	0.3625	0.2496	0.0004	0.0737	0.4322	0.4085	0.4186
102	0.4500	0.3926	0.3746	0.0625	0.0005	0.4188	0.4541	0.4001
103	0.4500	0.3292	0.2604	0.0139	0.0953	0.3937	0.4005	0.4146
104	0.4500	0.3044	0.4408	0.0273	0.1826	0.4633	0.4211	0.3857
105	0.4500	0.4317	0.2031	0.0203	0.0056	0.3960	0.4591	0.3978
106	0.4500	0.2712	0.1606	0.0025	0.0841	0.4092	0.4842	0.4394
107	0.4500	0.3015	0.4110	0.0006	0.0595	0.4450	0.4480	0.3831
108	0.4500	0.3777	0.1576	0.0822	0.0055	0.3891	0.4093	0.4227
109	0.4500	0.1406	0.3108	0.0438	0.1255	0.4320	0.4750	0.3725
110	0.4500	0.3335	0.3766	0.0088	0.0104	0.4421	0.4299	0.4226
111	0.4500	0.3105	0.3020	0.0146	0.0074	0.4586	0.4791	0.2712
112	0.4500	0.2928	0.3386	0.0029	0.0870	0.4334	0.3898	0.4160
113	0.4500	0.1487	0.4877	0.0490	0.2088	0.4894	0.3850	0.3664
114	0.4500	0.2698	0.3944	0.0082	0.0147	0.4718	0.3769	0.4353
115	0.4500	0.2455	0.3597	0.0165	0.0853	0.4290	0.4387	0.3726
116	0.4500	0.2799	0.3676	0.0059	0.0369	0.4450	0.3664	0.3849
117	0.4500	0.3071	0.1541	0.0215	0.0063	0.3966	0.4194	0.3653
118	0.4500	0.3739	0.2647	0.0024	0.0526	0.4013	0.3557	0.3526
119	0.4500	0.2583	0.1904	0.0078	0.0052	0.4132	0.4019	0.3533
120	0.4500	0.2657	0.3447	0.0026	0.0786	0.4289	0.4383	0.3608

ANEXO 3: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 50\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.5000	0.3362	0.4597	0.0138	0.1540	0.4933	0.4168	0.3880
2	0.5000	0.3866	0.3467	0.0441	0.0135	0.4754	0.4586	0.4791
3	0.5000	0.2659	0.3342	0.0090	0.1132	0.4876	0.4881	0.4406
4	0.5000	0.2879	0.3924	0.0012	0.0620	0.5085	0.4980	0.4839
5	0.5000	0.4184	0.4689	0.0056	0.1339	0.5037	0.4615	0.4622
6	0.5000	0.4731	0.2488	0.0004	0.0952	0.4680	0.4779	0.4193
7	0.5000	0.1104	0.2430	0.0829	0.2272	0.4872	0.4955	0.3391
8	0.5000	0.2824	0.3165	0.0258	0.1281	0.4885	0.5203	0.4394
9	0.5000	0.2857	0.3784	0.0167	0.1377	0.4899	0.5038	0.3542
10	0.5000	0.3584	0.4208	0.0010	0.0362	0.4900	0.5235	0.4225
11	0.5000	0.3542	0.3382	0.0192	0.0358	0.4935	0.4795	0.4550
12	0.5000	0.3205	0.3319	0.0000	0.0484	0.4735	0.4553	0.4717
13	0.5000	0.4053	0.2767	0.1000	0.0006	0.4582	0.5060	0.4421
14	0.5000	0.4266	0.2648	0.0064	0.0138	0.4466	0.4966	0.4560
15	0.5000	0.3116	0.4088	0.0462	0.1575	0.4868	0.5371	0.4476
16	0.5000	0.3490	0.3020	0.0326	0.0079	0.4902	0.4912	0.3257
17	0.5000	0.2696	0.2724	0.0088	0.0052	0.4849	0.4841	0.4602
18	0.5000	0.4044	0.2947	0.0150	0.1070	0.4353	0.4485	0.4661
19	0.5000	0.3683	0.2919	0.0102	0.0283	0.5068	0.4743	0.4518
20	0.5000	0.3120	0.2664	0.0096	0.0783	0.4502	0.4615	0.5465
21	0.5000	0.3000	0.2943	0.0132	0.0877	0.4684	0.5026	0.4182
22	0.5000	0.3945	0.4366	0.0007	0.1598	0.4930	0.4636	0.4555
23	0.5000	0.4892	0.3051	0.0548	0.0064	0.4499	0.5096	0.4573
24	0.5000	0.1340	0.5790	0.1151	0.2627	0.5499	0.5211	0.4311
25	0.5000	0.3627	0.2286	0.1016	0.0123	0.4556	0.5556	0.4828
26	0.5000	0.3162	0.4050	0.0002	0.0518	0.5424	0.5303	0.4739
27	0.5000	0.4798	0.3332	0.1170	0.0090	0.4674	0.4708	0.3961
28	0.5000	0.3764	0.4530	0.0105	0.0382	0.4999	0.4796	0.4458
29	0.5000	0.3523	0.4520	0.0167	0.0309	0.5063	0.4796	0.4324
30	0.5000	0.3293	0.2376	0.0001	0.0909	0.4749	0.5205	0.4825
31	0.5000	0.3881	0.2416	0.0136	0.0067	0.4407	0.4948	0.3830
32	0.5000	0.3594	0.3778	0.0614	0.2436	0.4978	0.4143	0.4204
33	0.5000	0.4858	0.2178	0.0059	0.0222	0.4464	0.5284	0.4646
34	0.5000	0.1453	0.2972	0.0219	0.0996	0.5046	0.5324	0.4224
35	0.5000	0.3025	0.4583	0.0278	0.1319	0.4895	0.4766	0.4148
36	0.5000	0.3802	0.3982	0.0007	0.0846	0.4938	0.4576	0.4577
37	0.5000	0.4351	0.2045	0.0181	0.0092	0.4292	0.4785	0.4751
38	0.5000	0.4169	0.2482	0.0318	0.0027	0.4819	0.5526	0.4864
39	0.5000	0.4099	0.3726	0.0058	0.0625	0.5006	0.4972	0.3982
40	0.5000	0.4459	0.2419	0.0777	0.0025	0.4371	0.4842	0.4641
41	0.5000	0.1908	0.3689	0.1129	0.2554	0.5078	0.4377	0.4197
42	0.5000	0.3502	0.3913	0.0022	0.0256	0.5106	0.4475	0.4123
43	0.5000	0.3845	0.2438	0.0060	0.0333	0.4606	0.5032	0.4474

44	0.5000	0.3831	0.1251	0.0516	0.0006	0.4161	0.5170	0.4740
45	0.5000	0.2728	0.4916	0.0043	0.0633	0.5246	0.4777	0.4252
46	0.5000	0.3193	0.3011	0.0049	0.0875	0.4826	0.4586	0.3555
47	0.5000	0.3348	0.4989	0.0090	0.0127	0.5100	0.4523	0.4586
48	0.5000	0.3836	0.2323	0.0001	0.0275	0.4517	0.5358	0.4459
49	0.5000	0.4625	0.2312	0.0008	0.0442	0.4622	0.4761	0.4259
50	0.5000	0.4772	0.3306	0.0530	0.0060	0.4679	0.4804	0.4286
51	0.5000	0.4894	0.4091	0.0000	0.0673	0.4704	0.5043	0.4220
52	0.5000	0.3652	0.4744	0.0001	0.0547	0.4903	0.4493	0.3866
53	0.5000	0.4827	0.2954	0.0003	0.0332	0.4531	0.5129	0.4655
54	0.5000	0.4342	0.2588	0.0153	0.0124	0.4674	0.4736	0.4447
55	0.5000	0.4477	0.2558	0.1205	0.0061	0.4611	0.5157	0.4641
56	0.5000	0.3951	0.4576	0.0406	0.1741	0.4913	0.4758	0.3322
57	0.5000	0.3254	0.3778	0.0544	0.1641	0.4895	0.4245	0.1767
58	0.5000	0.3576	0.4288	0.0027	0.0988	0.4936	0.4856	0.4702
59	0.5000	0.5180	0.4967	0.0396	0.1578	0.4763	0.4743	0.4553
60	0.5000	0.3961	0.2888	0.0000	0.0469	0.4590	0.4640	0.4395
61	0.5000	0.3126	0.5538	0.1437	0.2244	0.4993	0.4215	0.5007
62	0.5000	0.3482	0.2768	0.0419	0.1482	0.4674	0.5152	0.4944
63	0.5000	0.3498	0.4444	0.0073	0.0997	0.4996	0.4009	0.3763
64	0.5000	0.3713	0.3138	0.0391	0.1141	0.4963	0.4790	0.4090
65	0.5000	0.3702	0.1482	0.0703	0.0000	0.4162	0.5114	0.4653
66	0.5000	0.3798	0.3535	0.0312	0.1270	0.4848	0.5348	0.3957
67	0.5000	0.2280	0.4385	0.0610	0.2423	0.5025	0.4770	0.3726
68	0.5000	0.2633	0.2545	0.0091	0.0791	0.4658	0.5068	0.4365
69	0.5000	0.2033	0.3940	0.0605	0.2308	0.5065	0.4845	0.4440
70	0.5000	0.3216	0.1863	0.0505	0.0018	0.4372	0.4851	0.3887
71	0.5000	0.3957	0.3438	0.0007	0.0476	0.4636	0.4009	0.4220
72	0.5000	0.2870	0.3698	0.0005	0.0720	0.4953	0.4945	0.4272
73	0.5000	0.4280	0.2153	0.1054	0.0057	0.4366	0.4770	0.4411
74	0.5000	0.3224	0.3147	0.0086	0.1182	0.4880	0.4658	0.4396
75	0.5000	0.5343	0.4138	0.0493	0.0130	0.4560	0.4996	0.4453
76	0.5000	0.3176	0.3746	0.0236	0.1237	0.4971	0.4515	0.4391
77	0.5000	0.3722	0.2490	0.0336	0.0051	0.4716	0.4576	0.4769
78	0.5000	0.3221	0.4525	0.0035	0.1927	0.5129	0.3596	0.1676
79	0.5000	0.4559	0.3361	0.0315	0.0124	0.4617	0.5120	0.4707
80	0.5000	0.4508	0.1683	0.1755	0.0508	0.4399	0.5401	0.4292
81	0.5000	0.3707	0.3628	0.0043	0.0749	0.4686	0.5133	0.4755
82	0.5000	0.3333	0.4870	0.0000	0.0534	0.5015	0.5181	0.4523
83	0.5000	0.2187	0.3672	0.0418	0.1557	0.5156	0.4931	0.4251
84	0.5000	0.4671	0.3287	0.0241	0.0059	0.4707	0.4881	0.4649
85	0.5000	0.1752	0.2888	0.0102	0.0743	0.4849	0.5038	0.3173
86	0.5000	0.2790	0.3547	0.0002	0.0556	0.5106	0.4307	0.3559
87	0.5000	0.3321	0.3952	0.0297	0.2277	0.5097	0.4707	0.4956
88	0.5000	0.3581	0.3358	0.0010	0.0804	0.4674	0.4164	0.2257
89	0.5000	0.6091	0.3278	0.0053	0.0632	0.4646	0.4818	0.2882
90	0.5000	0.3777	0.2171	0.0141	0.0091	0.4538	0.4937	0.4458
91	0.5000	0.3576	0.4048	0.0507	0.1632	0.5051	0.4282	0.3146

92	0.5000	0.2275	0.4237	0.0028	0.0228	0.4903	0.4931	0.4965
93	0.5000	0.4675	0.2664	0.0084	0.1027	0.4523	0.4875	0.4750
94	0.5000	0.4101	0.4196	0.0218	0.1154	0.4535	0.4999	0.4461
95	0.5000	0.2789	0.3133	0.0025	0.1124	0.5443	0.5242	0.3659
96	0.5000	0.4350	0.3105	0.0027	0.0477	0.4563	0.4776	0.4149
97	0.5000	0.4914	0.3012	0.0050	0.0588	0.4762	0.4982	0.4514
98	0.5000	0.4768	0.2724	0.0396	0.0173	0.4504	0.5323	0.4048
99	0.5000	0.2871	0.4227	0.0005	0.0361	0.4678	0.4092	0.4381
100	0.5000	0.3565	0.2518	0.0041	0.0411	0.4595	0.5261	0.4405
101	0.5000	0.3375	0.4417	0.0021	0.0647	0.5001	0.4501	0.4690
102	0.5000	0.2964	0.2895	0.0085	0.1399	0.5149	0.5219	0.4447
103	0.5000	0.2975	0.2826	0.0021	0.0318	0.5025	0.4606	0.4535
104	0.5000	0.2299	0.2892	0.0025	0.0425	0.4733	0.4699	0.3626
105	0.5000	0.3434	0.4668	0.0031	0.1005	0.5067	0.5405	0.4790
106	0.5000	0.3191	0.4268	0.0130	0.1065	0.5020	0.4577	0.5370
107	0.5000	0.3482	0.3516	0.0062	0.0231	0.4806	0.4860	0.3643
108	0.5000	0.3287	0.3153	0.0181	0.1133	0.4643	0.4299	0.4609
109	0.5000	0.3700	0.2652	0.0212	0.0019	0.4677	0.4370	0.3497
110	0.5000	0.3350	0.2977	0.0292	0.0060	0.4419	0.4913	0.4501
111	0.5000	0.3120	0.3130	0.0041	0.0624	0.5071	0.5290	0.4390
112	0.5000	0.2557	0.4521	0.0016	0.0189	0.4991	0.5152	0.4394
113	0.5000	0.2208	0.2321	0.0058	0.0617	0.4770	0.5038	0.4607
114	0.5000	0.4846	0.2710	0.1053	0.0135	0.4723	0.4340	0.4056
115	0.5000	0.2051	0.3882	0.1020	0.2932	0.5304	0.4879	0.4800
116	0.5000	0.3213	0.2232	0.0024	0.0263	0.4487	0.5194	0.4195
117	0.5000	0.3238	0.2983	0.0008	0.0329	0.4731	0.4412	0.4325
118	0.5000	0.4727	0.2700	0.0045	0.0204	0.4429	0.4894	0.4718
119	0.5000	0.4495	0.2334	0.0340	0.0071	0.4499	0.4888	0.4627
120	0.5000	0.3687	0.3180	0.0146	0.0852	0.4668	0.5539	0.3982

ANEXO 4: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 55\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.5500	0.5374	0.3415	0.0669	0.2217	0.5272	0.5640	0.5046
2	0.5500	0.2307	0.2165	0.0082	0.0707	0.5120	0.5994	0.4532
3	0.5500	0.2530	0.2807	0.0006	0.0874	0.5337	0.4756	0.3295
4	0.5500	0.2223	0.4790	0.0594	0.2606	0.5787	0.4207	0.5134
5	0.5500	0.5593	0.3088	0.0013	0.0327	0.4979	0.5074	0.4016
6	0.5500	0.4446	0.3083	0.0174	0.1117	0.5150	0.5097	0.4689
7	0.5500	0.3842	0.2391	0.0002	0.0466	0.4915	0.5678	0.5443
8	0.5500	0.3924	0.3096	0.0025	0.0260	0.5197	0.5852	0.4560
9	0.5500	0.3514	0.5251	0.0035	0.0814	0.5829	0.4993	0.4521
10	0.5500	0.2767	0.4161	0.0460	0.1549	0.5505	0.5936	0.5261
11	0.5500	0.4202	0.2784	0.0053	0.0387	0.5047	0.5367	0.4497
12	0.5500	0.2668	0.3489	0.0014	0.0996	0.5284	0.4612	0.4261
13	0.5500	0.4058	0.2083	0.0102	0.0053	0.4843	0.5411	0.5095
14	0.5500	0.3970	0.2539	0.0139	0.0163	0.5322	0.5138	0.4344
15	0.5500	0.3304	0.4449	0.0087	0.1384	0.5530	0.5302	0.5264
16	0.5500	0.5751	0.4768	0.0026	0.1028	0.5143	0.5172	0.5588
17	0.5500	0.4094	0.2283	0.0147	0.0215	0.4815	0.5686	0.4965
18	0.5500	0.4376	0.4536	0.0560	0.1598	0.5312	0.5431	0.5014
19	0.5500	0.2982	0.4107	0.0691	0.2177	0.5512	0.5302	0.5104
20	0.5500	0.4598	0.3936	0.0070	0.0513	0.5439	0.5505	0.4574
21	0.5500	0.4662	0.2523	0.0193	0.0212	0.5029	0.5206	0.3630
22	0.5500	0.5967	0.1941	0.0048	0.0215	0.4488	0.5575	0.5213
23	0.5500	0.4961	0.2051	0.0376	0.0072	0.4937	0.5794	0.5094
24	0.5500	0.3465	0.2751	0.0003	0.0625	0.4945	0.5186	0.4729
25	0.5500	0.4133	0.3971	0.0006	0.0690	0.5200	0.5454	0.4837
26	0.5500	0.2445	0.3140	0.0001	0.0569	0.5346	0.4333	0.3727
27	0.5500	0.3672	0.4761	0.0636	0.2381	0.5539	0.5001	0.4795
28	0.5500	0.3807	0.3370	0.0045	0.0527	0.5372	0.5221	0.4691
29	0.5500	0.2294	0.5124	0.0464	0.2499	0.6039	0.5335	0.4348
30	0.5500	0.3012	0.3884	0.0028	0.0602	0.5096	0.5183	0.4067
31	0.5500	0.3849	0.3794	0.0005	0.0485	0.5471	0.5619	0.5174
32	0.5500	0.4488	0.2163	0.0007	0.0372	0.4921	0.5324	0.5092
33	0.5500	0.3723	0.3165	0.0047	0.0364	0.5057	0.5751	0.3738
34	0.5500	0.2451	0.4168	0.0058	0.0966	0.5281	0.5121	0.3891
35	0.5500	0.4371	0.3587	0.0008	0.0524	0.5090	0.5492	0.4763
36	0.5500	0.2880	0.3307	0.0028	0.0753	0.5550	0.5434	0.5061
37	0.5500	0.4993	0.4390	0.0000	0.0580	0.5107	0.5299	0.4722
38	0.5500	0.3736	0.2365	0.0167	0.0053	0.5003	0.5367	0.4338
39	0.5500	0.6018	0.2369	0.0522	0.0180	0.4690	0.5223	0.5155
40	0.5500	0.3749	0.3636	0.0009	0.0900	0.5521	0.4579	0.3151
41	0.5500	0.2749	0.4591	0.0245	0.2215	0.5662	0.4804	0.3376
42	0.5500	0.4542	0.5251	0.0052	0.0404	0.5464	0.5242	0.4716
43	0.5500	0.4604	0.2002	0.0065	0.0216	0.4680	0.5423	0.4408

44	0.5500	0.4923	0.3172	0.0303	0.0741	0.4983	0.4715	0.3935
45	0.5500	0.2256	0.4161	0.0897	0.2108	0.5584	0.5233	0.4379
46	0.5500	0.5527	0.1300	0.1289	0.0055	0.4417	0.5327	0.5639
47	0.5500	0.2975	0.3888	0.0078	0.1491	0.5348	0.5109	0.4846
48	0.5500	0.4416	0.4543	0.0493	0.1600	0.5365	0.5241	0.5303
49	0.5500	0.5352	0.3304	0.0929	0.0000	0.5006	0.5137	0.4043
50	0.5500	0.4439	0.3159	0.0000	0.0841	0.5332	0.5418	0.5460
51	0.5500	0.2664	0.3730	0.0118	0.1165	0.5279	0.5365	0.4954
52	0.5500	0.3109	0.4385	0.0001	0.0580	0.5497	0.5205	0.4690
53	0.5500	0.2459	0.4050	0.0122	0.1425	0.5339	0.4573	0.4316
54	0.5500	0.3175	0.4384	0.0033	0.1105	0.5719	0.5433	0.5172
55	0.5500	0.4359	0.1893	0.0002	0.0194	0.4633	0.6085	0.5535
56	0.5500	0.4474	0.2773	0.0033	0.0521	0.5042	0.5115	0.4658
57	0.5500	0.4493	0.5313	0.0279	0.1546	0.5377	0.5332	0.4322
58	0.5500	0.4935	0.4297	0.0026	0.0787	0.5031	0.3968	0.4147
59	0.5500	0.7649	0.3585	0.0001	0.0764	0.4724	0.4284	0.4500
60	0.5500	0.5134	0.2841	0.0460	0.0111	0.4573	0.4948	0.4828
61	0.5500	0.5041	0.2677	0.0176	0.0129	0.4911	0.6097	0.1735
62	0.5500	0.4352	0.5083	0.0025	0.1749	0.5487	0.5140	0.4743
63	0.5500	0.4120	0.5009	0.0018	0.1319	0.5643	0.5583	0.3948
64	0.5500	0.2748	0.4586	0.0137	0.3816	0.5731	0.4374	0.3640
65	0.5500	0.4833	0.2813	0.0191	0.0127	0.5133	0.5860	0.4819
66	0.5500	0.4669	0.3204	0.0000	0.0750	0.5281	0.4355	0.3737
67	0.5500	0.3637	0.3590	0.0021	0.0288	0.5288	0.4905	0.3464
68	0.5500	0.4298	0.2313	0.0098	0.0164	0.4783	0.5342	0.5574
69	0.5500	0.3752	0.3275	0.0129	0.1116	0.5046	0.5596	0.4617
70	0.5500	0.4406	0.3650	0.0054	0.1035	0.5185	0.4783	0.3910
71	0.5500	0.3178	0.5493	0.0976	0.3069	0.5711	0.4936	0.3661
72	0.5500	0.4847	0.2942	0.0075	0.0462	0.5150	0.5843	0.5355
73	0.5500	0.3559	0.3743	0.0000	0.1443	0.5605	0.4998	0.3419
74	0.5500	0.3149	0.4151	0.1584	0.2766	0.5238	0.5265	0.4468
75	0.5500	0.3034	0.5124	0.0289	0.1371	0.5586	0.5318	0.4338
76	0.5500	0.2712	0.4627	0.0062	0.1166	0.5606	0.5293	0.4968
77	0.5500	0.3080	0.3491	0.0224	0.1125	0.5521	0.5216	0.4171
78	0.5500	0.3342	0.5692	0.0021	0.0833	0.5907	0.5274	0.4959
79	0.5500	0.2310	0.4944	0.0112	0.1465	0.5786	0.5487	0.4736
80	0.5500	0.3757	0.2562	0.0132	0.0248	0.5213	0.6029	0.5171
81	0.5500	0.3776	0.2749	0.0150	0.0029	0.5268	0.5881	0.4967
82	0.5500	0.2379	0.2166	0.0030	0.0581	0.5006	0.5909	0.5154
83	0.5500	0.4758	0.3321	0.0002	0.0701	0.5123	0.4838	0.4565
84	0.5500	0.2793	0.4926	0.0829	0.2564	0.5385	0.4858	0.4955
85	0.5500	0.3075	0.3944	0.0006	0.0514	0.5296	0.5198	0.5016
86	0.5500	0.3002	0.3112	0.0661	0.1834	0.4831	0.5353	0.4196
87	0.5500	0.3997	0.2708	0.0172	0.0889	0.5033	0.5230	0.5133
88	0.5500	0.3329	0.3174	0.0004	0.0989	0.5565	0.5718	0.4948
89	0.5500	0.3919	0.3583	0.0053	0.0976	0.4939	0.5299	0.5111
90	0.5500	0.5021	0.2809	0.0010	0.0474	0.4940	0.5964	0.5735
91	0.5500	0.3427	0.3435	0.0023	0.1416	0.5459	0.4882	0.5383

92	0.5500	0.3658	0.2786	0.0098	0.0258	0.5228	0.5174	0.4755
93	0.5500	0.3072	0.3944	0.0094	0.1039	0.5535	0.5561	0.4806
94	0.5500	0.4581	0.3178	0.0011	0.0487	0.5063	0.5259	0.1237
95	0.5500	0.5843	0.2108	0.0066	0.0218	0.4708	0.4583	0.4492
96	0.5500	0.3098	0.2758	0.0002	0.0890	0.5358	0.5586	0.4955
97	0.5500	0.4427	0.4270	0.0338	0.1946	0.5487	0.5569	0.4886
98	0.5500	0.1808	0.3306	0.0099	0.1092	0.5592	0.6200	0.5148
99	0.5500	0.2116	0.3325	0.0105	0.1370	0.5300	0.4871	0.4038
100	0.5500	0.4568	0.3924	0.0140	0.1615	0.5203	0.5268	0.4953
101	0.5500	0.3218	0.3613	0.0132	0.0924	0.5033	0.5335	0.5622
102	0.5500	0.3048	0.4593	0.0418	0.2046	0.5608	0.4855	0.5230
103	0.5500	0.3922	0.3517	0.0090	0.0295	0.5457	0.4147	0.3704
104	0.5500	0.3860	0.3263	0.0072	0.0289	0.5203	0.5623	0.4792
105	0.5500	0.4086	0.2464	0.0035	0.0797	0.4898	0.4395	0.3558
106	0.5500	0.4237	0.2978	0.0140	0.0970	0.5102	0.5436	0.4828
107	0.5500	0.4251	0.3340	0.0030	0.0413	0.5110	0.4371	0.3084
108	0.5500	0.3672	0.4591	0.0102	0.1709	0.5614	0.4990	0.4906
109	0.5500	0.3263	0.3111	0.0022	0.0423	0.5427	0.4889	0.4592
110	0.5500	0.3731	0.3140	0.0010	0.0597	0.4909	0.5341	0.5347
111	0.5500	0.4431	0.4086	0.0764	0.1649	0.5272	0.5224	0.5360
112	0.5500	0.3146	0.4435	0.0010	0.2349	0.5456	0.4446	0.3254
113	0.5500	0.3622	0.3929	0.0050	0.0168	0.5471	0.5311	0.3071
114	0.5500	0.4764	0.4281	0.0332	0.0125	0.5187	0.5469	0.5634
115	0.5500	0.2408	0.3875	0.0153	0.1116	0.5427	0.5492	0.4949
116	0.5500	0.3108	0.2735	0.0091	0.0163	0.5331	0.4978	0.4600
117	0.5500	0.2811	0.5936	0.0554	0.2492	0.5601	0.4683	0.4061
118	0.5500	0.3590	0.5139	0.0081	0.0438	0.5517	0.5141	0.5285
119	0.5500	0.3298	0.3834	0.0003	0.0872	0.5337	0.4409	0.2724
120	0.5500	0.3288	0.2803	0.0326	0.0254	0.5339	0.5063	0.4376

ANEXO 5: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 60\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.6000	0.3931	0.3873	0.0001	0.0634	0.5748	0.6044	0.5312
2	0.6000	0.4554	0.2436	0.0694	0.0185	0.5850	0.6014	0.5463
3	0.6000	0.2641	0.3447	0.0128	0.0671	0.5574	0.6403	0.5574
4	0.6000	0.3987	0.4236	0.0167	0.1299	0.5977	0.6127	0.4804
5	0.6000	0.3476	0.4222	0.0336	0.0302	0.5789	0.5907	0.5552
6	0.6000	0.3252	0.3751	0.0329	0.1371	0.5816	0.6104	0.5086
7	0.6000	0.2521	0.4211	0.0000	0.0478	0.6382	0.7087	0.0683
8	0.6000	0.3147	0.4177	0.0282	0.1751	0.5794	0.5907	0.5394
9	0.6000	0.2889	0.6919	0.1090	0.3402	0.6398	0.5035	0.4700
10	0.6000	0.4013	0.4875	0.0602	0.2019	0.5755	0.5838	0.2665
11	0.6000	0.3330	0.3668	0.0035	0.1090	0.5579	0.5791	0.5320
12	0.6000	0.2724	0.3706	0.0136	0.1034	0.5852	0.5785	0.3206
13	0.6000	0.5361	0.4086	0.0010	0.0890	0.5662	0.5545	0.5725
14	0.6000	0.2337	0.2983	0.0001	0.0736	0.5641	0.6364	0.4204
15	0.6000	0.4618	0.5662	0.0193	0.1933	0.6325	0.5615	0.5074
16	0.6000	0.3522	0.4292	0.0024	0.1285	0.5963	0.5001	0.3498
17	0.6000	0.3243	0.5223	0.0051	0.1019	0.6063	0.5847	0.5008
18	0.6000	0.3460	0.6054	0.0002	0.0879	0.6269	0.5221	0.4473
19	0.6000	0.2452	0.3380	0.0292	0.1124	0.5646	0.6574	0.5014
20	0.6000	0.3673	0.5546	0.0885	0.2946	0.6101	0.5058	0.5626
21	0.6000	0.4656	0.2264	0.0005	0.0447	0.5345	0.5696	0.5223
22	0.6000	0.4879	0.2866	0.0662	0.0000	0.5392	0.5337	0.5515
23	0.6000	0.4092	0.2592	0.0268	0.0028	0.5419	0.5810	0.4532
24	0.6000	0.3004	0.3053	0.0034	0.0587	0.5699	0.6539	0.4671
25	0.6000	0.4593	0.3598	0.0013	0.0704	0.5409	0.5793	0.4938
26	0.6000	0.4694	0.2872	0.0064	0.1029	0.5416	0.5619	0.4736
27	0.6000	0.3645	0.3243	0.0002	0.0899	0.5948	0.6021	0.5669
28	0.6000	0.3213	0.3993	0.0019	0.0291	0.5983	0.6156	0.4539
29	0.6000	0.5625	0.3631	0.0013	0.0464	0.5263	0.5598	0.5090
30	0.6000	0.4915	0.3735	0.0138	0.0956	0.5623	0.6359	0.5831
31	0.6000	0.3995	0.3807	0.0001	0.0745	0.5497	0.5673	0.2934
32	0.6000	0.3513	0.5113	0.0073	0.0610	0.6106	0.5744	0.4463
33	0.6000	0.4740	0.5301	0.0003	0.0671	0.5691	0.5182	0.5868
34	0.6000	0.3887	0.3287	0.0026	0.0362	0.5715	0.6464	0.5372
35	0.6000	0.6298	0.2054	0.0876	0.0001	0.4966	0.5711	0.5401
36	0.6000	0.4584	0.3812	0.0030	0.0629	0.5587	0.5593	0.2839
37	0.6000	0.5602	0.4045	0.0153	0.1227	0.5437	0.5996	0.5629
38	0.6000	0.2559	0.5143	0.0045	0.1220	0.6110	0.5089	0.5591
39	0.6000	0.4178	0.4013	0.0521	0.2274	0.5819	0.5689	0.5179
40	0.6000	0.3637	0.4200	0.0754	0.1881	0.5730	0.6223	0.5538
41	0.6000	0.3135	0.4619	0.0199	0.2049	0.6242	0.5535	0.4228
42	0.6000	0.2922	0.5025	0.0284	0.2159	0.6024	0.5944	0.5639
43	0.6000	0.3803	0.3167	0.0003	0.0688	0.5561	0.5664	0.4818

44	0.6000	0.4213	0.3830	0.0006	0.0313	0.5782	0.6061	0.4245
45	0.6000	0.4710	0.3237	0.0084	0.0998	0.5653	0.6100	0.5512
46	0.6000	0.4565	0.2883	0.0065	0.0262	0.5916	0.5949	0.4625
47	0.6000	0.3853	0.6191	0.0585	0.2744	0.6204	0.5770	0.5955
48	0.6000	0.4156	0.4605	0.0001	0.0414	0.5706	0.5735	0.5800
49	0.6000	0.3277	0.4820	0.0023	0.1277	0.5720	0.5082	0.4969
50	0.6000	0.3615	0.3323	0.0001	0.0657	0.5545	0.4998	0.4536
51	0.6000	0.2859	0.3783	0.0017	0.0696	0.5832	0.6313	0.5580
52	0.6000	0.5181	0.3761	0.0001	0.0563	0.5464	0.5636	0.5484
53	0.6000	0.5218	0.4205	0.0013	0.0854	0.5900	0.6099	0.5124
54	0.6000	0.4808	0.2974	0.0009	0.0779	0.5506	0.6190	0.5568
55	0.6000	0.3931	0.3000	0.0611	0.1701	0.5477	0.5894	0.5074
56	0.6000	0.2441	0.4799	0.0259	0.1760	0.6160	0.4583	0.5824
57	0.6000	0.2421	0.4846	0.0788	0.2487	0.6203	0.5964	0.6141
58	0.6000	0.4047	0.5173	0.0011	0.0871	0.5886	0.5683	0.5005
59	0.6000	0.5132	0.3118	0.0115	0.0183	0.5328	0.6026	0.5479
60	0.6000	0.4095	0.5691	0.0012	0.1273	0.6164	0.5690	0.5168
61	0.6000	0.2176	0.3236	0.0397	0.2013	0.5853	0.5153	0.5084
62	0.6000	0.4312	0.3422	0.0047	0.0535	0.5880	0.5473	0.4775
63	0.6000	0.3882	0.3398	0.0007	0.0520	0.5498	0.5742	0.3194
64	0.6000	0.3656	0.3308	0.0075	0.0212	0.5419	0.6069	0.5833
65	0.6000	0.2392	0.4966	0.0397	0.1475	0.5837	0.5842	0.5618
66	0.6000	0.4045	0.5238	0.0316	0.2065	0.5878	0.5464	0.3550
67	0.6000	0.2545	0.3426	0.0048	0.0670	0.5975	0.5295	0.4690
68	0.6000	0.3020	0.4833	0.0238	0.1145	0.5772	0.5757	0.3469
69	0.6000	0.3595	0.2901	0.0157	0.0046	0.5468	0.4928	0.5103
70	0.6000	0.3733	0.4300	0.0017	0.0357	0.5835	0.5391	0.5416
71	0.6000	0.4389	0.4920	0.0026	0.0835	0.6016	0.6088	0.4548
72	0.6000	0.4965	0.2724	0.0010	0.0625	0.5381	0.6185	0.5522
73	0.6000	0.5311	0.3967	0.0003	0.0774	0.5620	0.6093	0.5406
74	0.6000	0.2849	0.2832	0.0000	0.0471	0.5559	0.6583	0.5615
75	0.6000	0.5673	0.3303	0.0012	0.0430	0.5273	0.5890	0.5837
76	0.6000	0.5579	0.1900	0.0134	0.0011	0.5183	0.6610	0.5449
77	0.6000	0.3627	0.3330	0.0181	0.0125	0.5585	0.6045	0.5535
78	0.6000	0.5163	0.4201	0.0001	0.1010	0.5731	0.5700	0.5952
79	0.6000	0.4633	0.4008	0.0026	0.0454	0.5353	0.5628	0.5842
80	0.6000	0.4935	0.5084	0.0506	0.2081	0.5888	0.6294	0.5230
81	0.6000	0.4140	0.5058	0.0769	0.2277	0.6025	0.5801	0.5963
82	0.6000	0.6935	0.2614	0.0187	0.0117	0.4822	0.6706	0.3692
83	0.6000	0.4772	0.2896	0.0025	0.0218	0.5572	0.5654	0.5539
84	0.6000	0.4176	0.5077	0.0329	0.1637	0.5884	0.6013	0.4973
85	0.6000	0.3456	0.4438	0.0136	0.1329	0.5983	0.5722	0.5956
86	0.6000	0.4794	0.2140	0.0014	0.0444	0.5342	0.5640	0.5701
87	0.6000	0.4404	0.4268	0.0004	0.0326	0.5504	0.5574	0.5193
88	0.6000	0.5217	0.3233	0.0193	0.0236	0.5407	0.5622	0.4918
89	0.6000	0.4464	0.2140	0.0420	0.0000	0.5181	0.5832	0.4264
90	0.6000	0.3585	0.3859	0.0005	0.0823	0.5817	0.6233	0.4988
91	0.6000	0.3268	0.2154	0.0016	0.0594	0.5358	0.6626	0.5930

92	0.6000	0.3642	0.4305	0.0059	0.0321	0.5697	0.5498	0.5563
93	0.6000	0.1420	0.4762	0.0291	0.1013	0.6299	0.6207	0.5421
94	0.6000	0.4769	0.4911	0.0020	0.0541	0.5642	0.5486	0.4795
95	0.6000	0.4842	0.4727	0.0251	0.2103	0.5785	0.5691	0.5436
96	0.6000	0.4113	0.3547	0.0000	0.0501	0.5684	0.5725	0.5179
97	0.6000	0.4619	0.4316	0.0008	0.0843	0.5919	0.5440	0.5067
98	0.6000	0.4037	0.3007	0.0221	0.0060	0.5484	0.6079	0.5398
99	0.6000	0.3692	0.5399	0.0052	0.1470	0.5827	0.5637	0.5941
100	0.6000	0.4252	0.4358	0.0029	0.1016	0.5786	0.5245	0.5344
101	0.6000	0.5245	0.3276	0.0002	0.0493	0.5496	0.6483	0.5317
102	0.6000	0.3578	0.4595	0.0140	0.1324	0.5948	0.5229	0.5713
103	0.6000	0.3637	0.2590	0.0000	0.0489	0.5165	0.6077	0.5559
104	0.6000	0.3605	0.2862	0.0015	0.0674	0.5785	0.6032	0.5743
105	0.6000	0.4012	0.3896	0.0001	0.0852	0.5810	0.5583	0.5302
106	0.6000	0.5305	0.3149	0.0004	0.0435	0.5463	0.5919	0.5354
107	0.6000	0.5793	0.4911	0.1111	0.2881	0.5695	0.5551	0.5665
108	0.6000	0.3632	0.4246	0.0003	0.0451	0.5545	0.6148	0.5360
109	0.6000	0.4642	0.3172	0.0002	0.0427	0.5716	0.5280	0.5507
110	0.6000	0.3132	0.5705	0.0397	0.1834	0.6151	0.5685	0.3761
111	0.6000	0.5302	0.3441	0.0063	0.0270	0.5439	0.6057	0.5548
112	0.6000	0.5020	0.2966	0.0062	0.0565	0.5517	0.5854	0.5455
113	0.6000	0.4981	0.6001	0.0001	0.0776	0.5823	0.5756	0.4581
114	0.6000	0.3286	0.6166	0.0029	0.0343	0.6646	0.5360	0.4781
115	0.6000	0.3926	0.3617	0.0418	0.1424	0.5527	0.5931	0.5958
116	0.6000	0.5174	0.4875	0.0354	0.1525	0.5812	0.5894	0.4437
117	0.6000	0.4789	0.4405	0.0043	0.1080	0.6057	0.5895	0.5414
118	0.6000	0.3877	0.4328	0.0071	0.1741	0.5816	0.5842	0.4412
119	0.6000	0.3615	0.2132	0.0508	0.0019	0.5539	0.6790	0.4276
120	0.6000	0.2665	0.4489	0.0955	0.2500	0.5944	0.6040	0.5364

ANEXO 6: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 65\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.6500	0.6500	0.5119	0.0071	0.0468	0.6247	0.5602	0.5244
2	0.6500	0.5075	0.4186	0.0002	0.0998	0.6122	0.6182	0.6039
3	0.6500	0.5539	0.5502	0.0030	0.1356	0.6338	0.5934	0.6297
4	0.6500	0.4509	0.5137	0.0005	0.1464	0.6400	0.5999	0.4306
5	0.6500	0.5720	0.5587	0.0073	0.0501	0.6193	0.6203	0.6160
6	0.6500	0.4304	0.6195	0.1276	0.3672	0.6596	0.5465	0.5922
7	0.6500	0.5311	0.5132	0.0028	0.2029	0.6683	0.6176	0.5625
8	0.6500	0.4279	0.3760	0.0398	0.0306	0.6304	0.5726	0.5592
9	0.6500	0.4613	0.2332	0.0008	0.0726	0.5928	0.6537	0.5714
10	0.6500	0.2944	0.5350	0.0478	0.2086	0.6676	0.5829	0.6258
11	0.6500	0.4650	0.5879	0.0135	0.2435	0.6531	0.5620	0.5995
12	0.6500	0.2338	0.4622	0.0424	0.2199	0.6668	0.6901	0.5932
13	0.6500	0.5386	0.5617	0.0008	0.0959	0.6487	0.5797	0.5329
14	0.6500	0.6862	0.5567	0.0140	0.0835	0.6389	0.6070	0.6107
15	0.6500	0.3962	0.4214	0.0115	0.0606	0.6375	0.5845	0.5592
16	0.6500	0.5059	0.4077	0.0080	0.0691	0.6093	0.6060	0.5394
17	0.6500	0.4187	0.5584	0.0250	0.2104	0.6444	0.6281	0.5200
18	0.6500	0.3574	0.4876	0.0017	0.0989	0.6287	0.6256	0.5487
19	0.6500	0.4167	0.3325	0.0184	0.0050	0.6246	0.6067	0.5887
20	0.6500	0.6211	0.4834	0.0867	0.2347	0.6175	0.6355	0.6166
21	0.6500	0.3620	0.5224	0.0002	0.0690	0.6544	0.6369	0.5865
22	0.6500	0.6204	0.4005	0.0005	0.1616	0.6246	0.5549	0.4623
23	0.6500	0.5691	0.4472	0.0433	0.0313	0.6171	0.4976	0.4810
24	0.6500	0.5531	0.4594	0.0080	0.1239	0.6162	0.6109	0.4796
25	0.6500	0.5399	0.4111	0.0438	0.0363	0.6161	0.6252	0.5609
26	0.6500	0.3613	0.5548	0.0246	0.1666	0.6599	0.5965	0.5636
27	0.6500	0.2298	0.4616	0.0144	0.1161	0.6626	0.6462	0.5739
28	0.6500	0.3027	0.5655	0.0863	0.3345	0.6578	0.5707	0.4634
29	0.6500	0.4341	0.4588	0.0042	0.1407	0.6364	0.5494	0.4895
30	0.6500	0.4524	0.4183	0.0117	0.1290	0.6383	0.4789	0.4804
31	0.6500	0.4606	0.4997	0.0062	0.0907	0.6415	0.5127	0.5697
32	0.6500	0.4885	0.5512	0.0112	0.2648	0.6435	0.6702	0.5695
33	0.6500	0.3876	0.5134	0.0240	0.1214	0.6674	0.6365	0.5976
34	0.6500	0.5673	0.4281	0.0150	0.0140	0.6229	0.6548	0.5458
35	0.6500	0.5129	0.4927	0.0000	0.0564	0.6494	0.5426	0.5143
36	0.6500	0.5558	0.4651	0.0788	0.0580	0.6389	0.5172	0.4400
37	0.6500	0.5887	0.3285	0.0628	0.0041	0.6123	0.6471	0.6162
38	0.6500	0.3401	0.5901	0.0001	0.0799	0.6572	0.6407	0.5937
39	0.6500	0.4450	0.5492	0.0050	0.1264	0.6487	0.6690	0.6041
40	0.6500	0.4521	0.5897	0.0000	0.1009	0.6590	0.6073	0.5187
41	0.6500	0.5006	0.5280	0.0000	0.1354	0.6508	0.6003	0.5998
42	0.6500	0.5289	0.3394	0.0003	0.0853	0.5713	0.6117	0.5534
43	0.6500	0.4653	0.1932	0.0261	0.0327	0.5538	0.6304	0.5528

44	0.6500	0.4511	0.6913	0.0001	0.1173	0.6830	0.5540	0.3409
45	0.6500	0.4709	0.5903	0.0466	0.2799	0.6481	0.5881	0.5433
46	0.6500	0.5000	0.5965	0.0061	0.1359	0.6658	0.5953	0.6063
47	0.6500	0.4896	0.4725	0.0004	0.0726	0.6238	0.5861	0.5498
48	0.6500	0.3534	0.6137	0.0482	0.2169	0.6556	0.5856	0.5735
49	0.6500	0.5902	0.4901	0.0116	0.0922	0.6200	0.6144	0.6168
50	0.6500	0.3743	0.5755	0.0641	0.2363	0.6552	0.7095	0.5612
51	0.6500	0.5150	0.3497	0.0000	0.0622	0.6318	0.6327	0.6179
52	0.6500	0.4522	0.6981	0.0045	0.1860	0.6861	0.6440	0.5331
53	0.6500	0.2945	0.4340	0.0002	0.1589	0.6468	0.5390	0.4292
54	0.6500	0.4992	0.4213	0.0094	0.0358	0.6181	0.6381	0.5574
55	0.6500	0.5607	0.4283	0.0275	0.0208	0.6091	0.6382	0.5973
56	0.6500	0.5060	0.3860	0.0076	0.0312	0.6126	0.6491	0.6185
57	0.6500	0.3910	0.4578	0.0112	0.1844	0.6458	0.6778	0.5530
58	0.6500	0.5483	0.4509	0.0365	0.0479	0.5999	0.6309	0.5489
59	0.6500	0.3668	0.5825	0.0316	0.1714	0.6627	0.6447	0.5674
60	0.6500	0.4436	0.5514	0.0261	0.3857	0.6673	0.5787	0.4776
61	0.6500	0.4874	0.3609	0.0000	0.0859	0.6298	0.5930	0.5637
62	0.6500	0.4089	0.5161	0.0187	0.1543	0.6603	0.6333	0.6694
63	0.6500	0.4345	0.3951	0.0019	0.0649	0.6345	0.6349	0.5387
64	0.6500	0.2665	0.6189	0.0429	0.1610	0.6739	0.6379	0.6176
65	0.6500	0.4884	0.4464	0.0004	0.0995	0.6174	0.6063	0.5886
66	0.6500	0.5496	0.5531	0.0067	0.1806	0.6442	0.6194	0.5468
67	0.6500	0.4685	0.3460	0.0000	0.0677	0.6126	0.6200	0.6094
68	0.6500	0.3182	0.5317	0.0455	0.1958	0.6740	0.6809	0.5632
69	0.6500	0.5170	0.3777	0.0203	0.0465	0.6144	0.6125	0.5478
70	0.6500	0.5598	0.3748	0.0009	0.0771	0.5895	0.5856	0.6163
71	0.6500	0.4790	0.4740	0.0224	0.0186	0.6137	0.5890	0.5889
72	0.6500	0.4545	0.4231	0.0400	0.0352	0.6493	0.5840	0.3861
73	0.6500	0.3424	0.4477	0.0004	0.1291	0.6407	0.6138	0.5969
74	0.6500	0.4489	0.5873	0.0055	0.1493	0.6511	0.5397	0.6128
75	0.6500	0.4994	0.4603	0.0010	0.1203	0.6112	0.6312	0.5562
76	0.6500	0.4337	0.6102	0.0132	0.1971	0.6467	0.6013	0.5807
77	0.6500	0.5687	0.3217	0.0441	0.0104	0.5957	0.6547	0.6221
78	0.6500	0.3128	0.5083	0.0163	0.2061	0.6521	0.5670	0.5532
79	0.6500	0.6836	0.2575	0.1698	0.0098	0.5658	0.6087	0.5614
80	0.6500	0.6201	0.4613	0.0133	0.2005	0.6252	0.6147	0.5980
81	0.6500	0.3452	0.3945	0.0086	0.1561	0.6278	0.6722	0.4819
82	0.6500	0.4747	0.6445	0.0376	0.2858	0.6605	0.5592	0.5616
83	0.6500	0.4688	0.4684	0.0147	0.0142	0.6632	0.5824	0.5470
84	0.6500	0.5636	0.4202	0.0591	0.0033	0.6452	0.6244	0.6138
85	0.6500	0.5804	0.4003	0.0169	0.0376	0.6074	0.5915	0.5245
86	0.6500	0.3923	0.5405	0.0056	0.1475	0.6440	0.6208	0.3850
87	0.6500	0.5054	0.5675	0.0000	0.1624	0.6455	0.6188	0.6108
88	0.6500	0.5076	0.3323	0.0265	0.0356	0.6338	0.5694	0.5051
89	0.6500	0.6259	0.4721	0.0038	0.1522	0.6247	0.5762	0.5802
90	0.6500	0.5391	0.4441	0.0260	0.0164	0.6235	0.5451	0.5361
91	0.6500	0.4537	0.4587	0.0031	0.1310	0.6541	0.6464	0.5175

92	0.6500	0.2757	0.5166	0.0759	0.2419	0.6525	0.6647	0.4356
93	0.6500	0.3932	0.5875	0.0024	0.2433	0.6559	0.5810	0.5393
94	0.6500	0.4955	0.5650	0.0010	0.1268	0.6564	0.6434	0.5635
95	0.6500	0.5820	0.4233	0.0019	0.1421	0.5941	0.6078	0.5997
96	0.6500	0.4355	0.6364	0.0204	0.1807	0.6456	0.6129	0.6320
97	0.6500	0.4512	0.5290	0.0161	0.1579	0.6526	0.5515	0.4725
98	0.6500	0.6146	0.4404	0.0077	0.0642	0.6012	0.6335	0.5321
99	0.6500	0.5326	0.4481	0.0347	0.0378	0.6401	0.6208	0.6337
100	0.6500	0.3445	0.4944	0.0116	0.0258	0.6386	0.6155	0.5611
101	0.6500	0.4778	0.3596	0.0106	0.0747	0.6123	0.6104	0.5103
102	0.6500	0.5783	0.5357	0.0012	0.1527	0.6415	0.6388	0.5863
103	0.6500	0.4549	0.6722	0.0304	0.1909	0.6456	0.6231	0.5862
104	0.6500	0.5050	0.3303	0.0000	0.0729	0.6066	0.6595	0.6274
105	0.6500	0.3278	0.5080	0.0001	0.1251	0.6634	0.5710	0.5489
106	0.6500	0.2998	0.5471	0.0698	0.3286	0.6639	0.6197	0.5466
107	0.6500	0.4682	0.5422	0.0047	0.2016	0.6571	0.5865	0.5956
108	0.6500	0.4733	0.4983	0.0073	0.0730	0.6160	0.6167	0.6010
109	0.6500	0.4460	0.4615	0.0002	0.1596	0.6095	0.5548	0.5642
110	0.6500	0.5491	0.4815	0.0076	0.0974	0.6281	0.6242	0.6165
111	0.6500	0.4431	0.4401	0.0217	0.0358	0.6723	0.6362	0.5256
112	0.6500	0.2923	0.4278	0.0103	0.1345	0.6394	0.6130	0.4550
113	0.6500	0.4379	0.4583	0.0014	0.1351	0.6346	0.6322	0.6243
114	0.6500	0.3784	0.4006	0.0009	0.1126	0.6389	0.6802	0.5705
115	0.6500	0.4557	0.5333	0.0159	0.1929	0.6414	0.6352	0.6096
116	0.6500	0.5192	0.6365	0.0106	0.2220	0.6852	0.5906	0.5445
117	0.6500	0.5277	0.6648	0.0013	0.1217	0.6384	0.5989	0.5598
118	0.6500	0.3512	0.3386	0.0131	0.0493	0.6187	0.6290	0.5572
119	0.6500	0.3464	0.5181	0.0405	0.2537	0.6613	0.5933	0.4813
120	0.6500	0.6427	0.4079	0.1053	0.0065	0.5909	0.5789	0.4997

ANEXO 7: Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 70\%$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.7000	0.5958	0.5312	0.0007	0.1331	0.6755	0.6721	0.6669
2	0.7000	0.5557	0.3254	0.0341	0.0148	0.6477	0.6508	0.6402
3	0.7000	0.5694	0.5763	0.0077	0.0910	0.6960	0.6763	0.5616
4	0.7000	0.7023	0.4990	0.0001	0.1056	0.6532	0.6825	0.6506
5	0.7000	0.5589	0.4863	0.0111	0.0713	0.6847	0.7129	0.6313
6	0.7000	0.5963	0.6195	0.0177	0.2059	0.6882	0.6646	0.6587
7	0.7000	0.5755	0.4364	0.0003	0.0978	0.6617	0.6924	0.6446
8	0.7000	0.5040	0.5003	0.0040	0.0986	0.6904	0.5634	0.5070
9	0.7000	0.4734	0.5880	0.0068	0.0567	0.7053	0.6228	0.5898
10	0.7000	0.4814	0.5388	0.0053	0.0573	0.6890	0.6155	0.5474
11	0.7000	0.5663	0.6234	0.0794	0.4297	0.7017	0.6983	0.3441
12	0.7000	0.6265	0.5565	0.0949	0.0048	0.6798	0.6427	0.6297
13	0.7000	0.4229	0.5646	0.0000	0.0771	0.6867	0.6591	0.6725
14	0.7000	0.5747	0.7088	0.0390	0.0142	0.7097	0.6374	0.5883
15	0.7000	0.5028	0.5260	0.0093	0.0854	0.7053	0.7021	0.6434
16	0.7000	0.4818	0.4373	0.0347	0.1904	0.6649	0.7526	0.5726
17	0.7000	0.5026	0.5146	0.0831	0.3033	0.6817	0.6032	0.5558
18	0.7000	0.5294	0.5346	0.0141	0.0493	0.6803	0.6786	0.6214
19	0.7000	0.3356	0.5743	0.0492	0.2454	0.7176	0.6437	0.1977
20	0.7000	0.7036	0.4918	0.0001	0.1196	0.6499	0.6839	0.6616
21	0.7000	0.6021	0.5291	0.0269	0.1054	0.6756	0.6781	0.6202
22	0.7000	0.6077	0.3817	0.0226	0.0281	0.6561	0.6998	0.6269
23	0.7000	0.5956	0.5694	0.0027	0.2284	0.6705	0.6440	0.2581
24	0.7000	0.5003	0.4819	0.0648	0.3374	0.6769	0.5900	0.5916
25	0.7000	0.6513	0.4526	0.1227	0.0010	0.6603	0.6787	0.7220
26	0.7000	0.5110	0.5413	0.0102	0.1957	0.6866	0.6446	0.6298
27	0.7000	0.5270	0.5255	0.0033	0.0847	0.6852	0.6725	0.6111
28	0.7000	0.5137	0.3569	0.0241	0.0843	0.6640	0.6590	0.6064
29	0.7000	0.5507	0.5419	0.0070	0.1438	0.6874	0.7476	0.6659
30	0.7000	0.5095	0.6275	0.0000	0.1504	0.7113	0.6585	0.5840
31	0.7000	0.5331	0.6873	0.0015	0.1755	0.7095	0.6327	0.5353
32	0.7000	0.2362	0.6552	0.1193	0.3737	0.7240	0.7081	0.6646
33	0.7000	0.5290	0.4973	0.0060	0.2402	0.6827	0.6385	0.6415
34	0.7000	0.5449	0.6167	0.0262	0.0582	0.7042	0.6983	0.7109
35	0.7000	0.5880	0.4882	0.0000	0.1449	0.6595	0.5919	0.6139
36	0.7000	0.7281	0.5022	0.1071	0.0066	0.6580	0.6721	0.6285
37	0.7000	0.5981	0.4603	0.0036	0.2075	0.6738	0.6357	0.5958
38	0.7000	0.5194	0.6462	0.0006	0.1157	0.7256	0.6565	0.6054
39	0.7000	0.6664	0.5045	0.0164	0.0915	0.6630	0.6569	0.6287
40	0.7000	0.6143	0.7417	0.0225	0.3077	0.6946	0.6367	0.6394
41	0.7000	0.5688	0.4954	0.0156	0.2051	0.6749	0.6904	0.4490
42	0.7000	0.4651	0.5049	0.0502	0.2168	0.6880	0.6803	0.5712
43	0.7000	0.5414	0.4998	0.0279	0.0192	0.6962	0.7053	0.5866

44	0.7000	0.5910	0.3283	0.2186	0.0305	0.6621	0.6864	0.6341
45	0.7000	0.4102	0.5564	0.0202	0.1674	0.6999	0.6761	0.6547
46	0.7000	0.4401	0.6679	0.0007	0.1265	0.7070	0.6410	0.6126
47	0.7000	0.6005	0.5054	0.0206	0.0504	0.6715	0.5801	0.4721
48	0.7000	0.6022	0.5848	0.0000	0.1458	0.6984	0.6847	0.5721
49	0.7000	0.4696	0.5857	0.0031	0.1601	0.7030	0.6448	0.5234
50	0.7000	0.7088	0.5049	0.0006	0.1027	0.6641	0.6945	0.5698
51	0.7000	0.5066	0.4875	0.0045	0.1312	0.7019	0.5931	0.3938
52	0.7000	0.5733	0.6117	0.0003	0.1102	0.6923	0.6913	0.6178
53	0.7000	0.5325	0.4974	0.0456	0.2190	0.6743	0.6786	0.6588
54	0.7000	0.4483	0.5978	0.0850	0.3174	0.7081	0.6746	0.5431
55	0.7000	0.5965	0.5325	0.0257	0.2546	0.6848	0.6785	0.6151
56	0.7000	0.5996	0.6289	0.0119	0.1526	0.7116	0.7173	0.6331
57	0.7000	0.5138	0.4862	0.0008	0.0981	0.6854	0.6894	0.6315
58	0.7000	0.4921	0.5351	0.0054	0.0439	0.6955	0.6278	0.6134
59	0.7000	0.5398	0.5262	0.0235	0.1860	0.6884	0.7255	0.4170
60	0.7000	0.6202	0.5812	0.1058	0.0079	0.6835	0.6917	0.6792
61	0.7000	0.6517	0.3647	0.0470	0.0178	0.6607	0.6747	0.6215
62	0.7000	0.5896	0.6120	0.0082	0.1026	0.6985	0.7008	0.6980
63	0.7000	0.5298	0.7191	0.0130	0.0893	0.7117	0.7001	0.5985
64	0.7000	0.5263	0.4620	0.0004	0.0698	0.6540	0.7027	0.6398
65	0.7000	0.6985	0.5271	0.0000	0.1676	0.6828	0.6472	0.3858
66	0.7000	0.5176	0.5542	0.0266	0.1666	0.6886	0.6576	0.5726
67	0.7000	0.4890	0.4709	0.0003	0.0986	0.6838	0.7011	0.6594
68	0.7000	0.5294	0.4191	0.0071	0.0854	0.6648	0.6542	0.5552
69	0.7000	0.5003	0.5478	0.0152	0.2322	0.7086	0.6612	0.6574
70	0.7000	0.6829	0.4550	0.0023	0.1307	0.6596	0.6783	0.6397
71	0.7000	0.4904	0.6201	0.0156	0.0789	0.7091	0.6737	0.5818
72	0.7000	0.5372	0.4808	0.0000	0.0776	0.6742	0.6621	0.5667
73	0.7000	0.4663	0.5137	0.0284	0.1423	0.6785	0.6747	0.6511
74	0.7000	0.6429	0.4296	0.0007	0.0735	0.6571	0.7007	0.6464
75	0.7000	0.4788	0.4635	0.0002	0.0714	0.6774	0.6606	0.6362
76	0.7000	0.4638	0.4413	0.0196	0.1569	0.6725	0.6577	0.6032
77	0.7000	0.6487	0.5842	0.0143	0.0942	0.6710	0.6739	0.6299
78	0.7000	0.6588	0.4863	0.0040	0.1021	0.6620	0.6873	0.5913
79	0.7000	0.5228	0.4796	0.0029	0.0812	0.6836	0.6718	0.6127
80	0.7000	0.3632	0.6655	0.1215	0.3050	0.7415	0.6598	0.6164
81	0.7000	0.7392	0.5775	0.0000	0.1195	0.6653	0.6443	0.6426
82	0.7000	0.6877	0.5047	0.0319	0.0508	0.6363	0.6333	0.6130
83	0.7000	0.5285	0.7734	0.0000	0.2443	0.7340	0.5776	0.5178
84	0.7000	0.6384	0.5332	0.0023	0.1403	0.6744	0.6608	0.6187
85	0.7000	0.4828	0.7041	0.0009	0.1568	0.6953	0.6490	0.6424
86	0.7000	0.5045	0.6637	0.0640	0.3100	0.6891	0.6822	0.6337
87	0.7000	0.2929	0.5351	0.0683	0.2058	0.7075	0.6972	0.6516
88	0.7000	0.3873	0.6029	0.0379	0.2943	0.7068	0.6562	0.6024
89	0.7000	0.6028	0.3138	0.1264	0.0013	0.6217	0.6182	0.6354
90	0.7000	0.5334	0.6332	0.0637	0.0261	0.6936	0.6297	0.6024
91	0.7000	0.5151	0.5950	0.0157	0.2259	0.6923	0.6791	0.6534

92	0.7000	0.4234	0.6211	0.0003	0.0973	0.7010	0.6909	0.6191
93	0.7000	0.4817	0.5048	0.0361	0.2115	0.6623	0.6292	0.3518
94	0.7000	0.6015	0.5657	0.0361	0.2705	0.7009	0.5879	0.5580
95	0.7000	0.4707	0.6406	0.0111	0.0759	0.7202	0.6254	0.5382
96	0.7000	0.4062	0.4644	0.0007	0.1229	0.6765	0.6732	0.6555
97	0.7000	0.5502	0.5626	0.0057	0.1542	0.6939	0.6381	0.5519
98	0.7000	0.5740	0.5097	0.0031	0.1486	0.6918	0.5508	0.5635
99	0.7000	0.5061	0.6491	0.0030	0.2072	0.7073	0.6416	0.6543
100	0.7000	0.5137	0.5128	0.0012	0.1901	0.7010	0.6011	0.7121
101	0.7000	0.5673	0.3869	0.0017	0.1383	0.6443	0.6887	0.5961
102	0.7000	0.6534	0.4143	0.0010	0.1199	0.6635	0.6411	0.5733
103	0.7000	0.5721	0.6384	0.0974	0.0046	0.6988	0.6281	0.6173
104	0.7000	0.5847	0.6097	0.0033	0.2354	0.7058	0.6576	0.5340
105	0.7000	0.5551	0.5868	0.0179	0.0513	0.6997	0.6492	0.5987
106	0.7000	0.4909	0.5482	0.0446	0.3467	0.7034	0.6016	0.6297
107	0.7000	0.4705	0.6871	0.0087	0.1112	0.6903	0.5787	0.5819
108	0.7000	0.6640	0.5183	0.0056	0.1700	0.6710	0.6643	0.5651
109	0.7000	0.6008	0.4352	0.0154	0.0317	0.6737	0.7166	0.6684
110	0.7000	0.6233	0.5250	0.0105	0.1462	0.6946	0.5823	0.4226
111	0.7000	0.5788	0.3928	0.0352	0.0171	0.6664	0.7012	0.6162
112	0.7000	0.6455	0.4129	0.0001	0.0682	0.6501	0.7008	0.6373
113	0.7000	0.4670	0.5425	0.0116	0.1121	0.6904	0.7102	0.3828
114	0.7000	0.6114	0.4936	0.0609	0.2748	0.6772	0.6986	0.6250
115	0.7000	0.7170	0.3580	0.2140	0.0002	0.6458	0.7231	0.6107
116	0.7000	0.5291	0.6030	0.0012	0.1233	0.6986	0.6808	0.6626
117	0.7000	0.6962	0.6865	0.0060	0.2321	0.6795	0.6380	0.6179
118	0.7000	0.4267	0.6022	0.0155	0.0563	0.6907	0.7025	0.5905
119	0.7000	0.5142	0.5673	0.0103	0.1826	0.7098	0.6655	0.5942
120	0.7000	0.5271	0.6924	0.0069	0.1190	0.6897	0.5902	0.4845

ANEXO 8: (80%) Correlações entre o valor genético e os valores das estatísticas adicionais para $h^2 = 80$

arquivo	μ	max	min	sd	cv	PI	rank	classe
1	0.8000	0.8173	0.6283	0.0007	0.1307	0.7574	0.7819	0.7054
2	0.8000	0.6390	0.5723	0.0308	0.0277	0.7767	0.7718	0.7256
3	0.8000	0.6412	0.7369	0.0044	0.2124	0.7902	0.7775	0.6552
4	0.8000	0.7723	0.5079	0.0023	0.1008	0.7621	0.7463	0.7771
5	0.8000	0.6034	0.5897	0.0070	0.2961	0.7818	0.7624	0.2813
6	0.8000	0.6020	0.6972	0.0155	0.2652	0.7854	0.7644	0.6938
7	0.8000	0.6143	0.8109	0.0032	0.1817	0.8192	0.6902	0.6684
8	0.8000	0.6245	0.6637	0.0118	0.1672	0.7950	0.7821	0.7248
9	0.8000	0.5194	0.7041	0.0932	0.3087	0.7892	0.8000	0.7017
10	0.8000	0.6651	0.7686	0.0029	0.3968	0.8078	0.7393	0.4642
11	0.8000	0.6607	0.5898	0.0052	0.2468	0.7820	0.7669	0.7859
12	0.8000	0.7114	0.6246	0.0515	0.1920	0.7652	0.6969	0.6695
13	0.8000	0.6759	0.6927	0.0598	0.1390	0.7857	0.7980	0.7513
14	0.8000	0.5225	0.6876	0.0091	0.1895	0.8094	0.7737	0.7126
15	0.8000	0.6932	0.5839	0.0704	0.0563	0.7650	0.7790	0.7285
16	0.8000	0.6672	0.6520	0.0010	0.1375	0.7788	0.7828	0.7388
17	0.8000	0.7013	0.7049	0.0564	0.4268	0.7960	0.6191	0.5573
18	0.8000	0.6418	0.7148	0.0415	0.3684	0.7929	0.7869	0.4223
19	0.8000	0.7175	0.6608	0.0222	0.1484	0.7799	0.7363	0.7070
20	0.8000	0.5377	0.6897	0.0009	0.1832	0.7950	0.8078	0.7285
21	0.8000	0.6437	0.5706	0.0023	0.0917	0.7855	0.7613	0.6400
22	0.8000	0.6930	0.6632	0.0029	0.2123	0.7954	0.8059	0.7183
23	0.8000	0.5724	0.7671	0.0041	0.2463	0.7995	0.7593	0.7209
24	0.8000	0.7135	0.4079	0.0434	0.1236	0.7588	0.7920	0.7565
25	0.8000	0.7627	0.6340	0.0304	0.3204	0.7855	0.7483	0.7508
26	0.8000	0.7252	0.7109	0.0252	0.1369	0.7911	0.7697	0.7241
27	0.8000	0.5529	0.7096	0.1896	0.4437	0.7947	0.7796	0.7435
28	0.8000	0.5357	0.7291	0.0000	0.1810	0.7867	0.7320	0.6166
29	0.8000	0.6040	0.6730	0.0134	0.2874	0.7918	0.7314	0.5205
30	0.8000	0.6827	0.7829	0.0115	0.2393	0.8040	0.7752	0.7874
31	0.8000	0.6895	0.5287	0.0337	0.0798	0.7708	0.7213	0.6957
32	0.8000	0.8255	0.6098	0.1676	0.0137	0.7502	0.7705	0.7563
33	0.8000	0.7845	0.5974	0.0054	0.1194	0.7700	0.7903	0.6025
34	0.8000	0.5964	0.6387	0.0017	0.4096	0.8022	0.6840	0.6615
35	0.8000	0.6221	0.6283	0.0036	0.1300	0.7904	0.7972	0.6960
36	0.8000	0.6805	0.6344	0.0669	0.0663	0.7854	0.7903	0.7257
37	0.8000	0.6481	0.6785	0.0006	0.1322	0.8001	0.7716	0.7527
38	0.8000	0.6674	0.6455	0.0082	0.1976	0.8057	0.7705	0.7513
39	0.8000	0.7229	0.5544	0.0063	0.1360	0.7826	0.7093	0.6925
40	0.8000	0.7352	0.6475	0.0162	0.1807	0.7866	0.7425	0.7717
41	0.8000	0.7147	0.6610	0.0367	0.2544	0.7876	0.7327	0.7389
42	0.8000	0.7211	0.5333	0.0067	0.1957	0.7604	0.7539	0.7586
43	0.8000	0.6337	0.6103	0.0034	0.2482	0.7878	0.7590	0.6160

44	0.8000	0.5664	0.7345	0.0003	0.1159	0.8123	0.7567	0.6927
45	0.8000	0.6571	0.6408	0.0013	0.2028	0.7805	0.6981	0.6986
46	0.8000	0.6691	0.6884	0.0039	0.2237	0.8022	0.7196	0.6817
47	0.8000	0.6349	0.6899	0.0011	0.2405	0.8038	0.7828	0.7377
48	0.8000	0.7289	0.5718	0.0053	0.0766	0.7921	0.6840	0.5124
49	0.8000	0.8272	0.5350	0.0252	0.1127	0.7580	0.7451	0.4560
50	0.8000	0.6634	0.7277	0.0003	0.1668	0.8002	0.7836	0.7350
51	0.8000	0.6780	0.6659	0.0396	0.0959	0.7940	0.7480	0.7288
52	0.8000	0.6977	0.6406	0.0884	0.0113	0.7910	0.7930	0.6165
53	0.8000	0.7182	0.7447	0.0615	0.3489	0.8019	0.7577	0.6937
54	0.8000	0.6160	0.6376	0.0012	0.2160	0.8061	0.7229	0.6213
55	0.8000	0.6730	0.7133	0.0148	0.0780	0.7796	0.7608	0.6818
56	0.8000	0.8277	0.6550	0.0021	0.1264	0.7829	0.7653	0.7319
57	0.8000	0.6041	0.5391	0.0008	0.0909	0.7763	0.7771	0.7479
58	0.8000	0.7370	0.5568	0.0009	0.1842	0.7594	0.7621	0.7266
59	0.8000	0.6797	0.6093	0.0195	0.1010	0.7798	0.7864	0.7492
60	0.8000	0.5352	0.6741	0.0096	0.2218	0.7947	0.7679	0.7259
61	0.8000	0.6788	0.6517	0.0830	0.0247	0.7859	0.7693	0.7244
62	0.8000	0.7172	0.5932	0.0143	0.0706	0.7791	0.7764	0.7564
63	0.8000	0.7321	0.6747	0.0342	0.5742	0.7925	0.6098	0.6154
64	0.8000	0.7576	0.7486	0.0032	0.2321	0.7896	0.7100	0.6247
65	0.8000	0.6584	0.6443	0.0193	0.1866	0.7890	0.7528	0.6170
66	0.8000	0.6605	0.7250	0.0357	0.1080	0.8104	0.7536	0.7185
67	0.8000	0.7343	0.5115	0.0472	0.0214	0.7900	0.7737	0.7380
68	0.8000	0.6176	0.5928	0.0015	0.1712	0.7908	0.7173	0.7153
69	0.8000	0.7141	0.6037	0.0028	0.2742	0.7962	0.7514	0.6886
70	0.8000	0.7471	0.7484	0.0005	0.3515	0.7838	0.6916	0.6766
71	0.8000	0.6419	0.5486	0.0389	0.0301	0.7747	0.7456	0.6560
72	0.8000	0.6657	0.6927	0.0051	0.1467	0.7902	0.7788	0.6886
73	0.8000	0.7124	0.6053	0.0114	0.1782	0.7829	0.7778	0.7791
74	0.8000	0.6352	0.7668	0.0413	0.4820	0.8020	0.7581	0.6229
75	0.8000	0.7040	0.5622	0.0004	0.2075	0.7808	0.7316	0.5731
76	0.8000	0.6563	0.7028	0.0084	0.2787	0.7835	0.7483	0.7425
77	0.8000	0.6520	0.6334	0.0001	0.2454	0.7874	0.7378	0.5413
78	0.8000	0.6337	0.6605	0.0357	0.3417	0.7907	0.7777	0.6671
79	0.8000	0.6577	0.6267	0.0009	0.1614	0.7963	0.8032	0.7655
80	0.8000	0.7380	0.6246	0.0003	0.2656	0.7863	0.7522	0.7211
81	0.8000	0.8220	0.6718	0.0024	0.1676	0.7735	0.7457	0.7460
82	0.8000	0.6213	0.7370	0.0010	0.2118	0.8079	0.7712	0.7244
83	0.8000	0.6456	0.6447	0.0003	0.1477	0.7795	0.8020	0.8035
84	0.8000	0.7057	0.6266	0.0145	0.1651	0.7845	0.7936	0.6415
85	0.8000	0.5564	0.7554	0.0088	0.3214	0.8117	0.7808	0.6371
86	0.8000	0.6571	0.7418	0.0613	0.0505	0.8030	0.7848	0.6473
87	0.8000	0.6169	0.7921	0.0693	0.5055	0.8094	0.7821	0.5958
88	0.8000	0.6466	0.6298	0.0005	0.2637	0.7768	0.7908	0.7337
89	0.8000	0.6392	0.7572	0.0040	0.2197	0.7993	0.7582	0.6486
90	0.8000	0.6282	0.6915	0.0168	0.2817	0.7983	0.7266	0.6455
91	0.8000	0.6620	0.6402	0.0003	0.1880	0.7929	0.7740	0.7225

92	0.8000	0.7507	0.5922	0.0027	0.1493	0.7753	0.7547	0.7214
93	0.8000	0.6387	0.7301	0.0211	0.0887	0.7859	0.7233	0.7393
94	0.8000	0.6201	0.7934	0.0125	0.3329	0.8114	0.7014	0.6936
95	0.8000	0.6952	0.5403	0.0813	0.0768	0.7790	0.7587	0.7041
96	0.8000	0.6521	0.6866	0.0201	0.1302	0.7914	0.7310	0.7289
97	0.8000	0.6847	0.5327	0.0004	0.1658	0.7652	0.7250	0.3594
98	0.8000	0.7223	0.6602	0.0793	0.3486	0.7816	0.7920	0.7202
99	0.8000	0.6229	0.6675	0.0672	0.4405	0.7989	0.7957	0.7553
100	0.8000	0.6403	0.6103	0.0326	0.2711	0.7901	0.7604	0.7922
101	0.8000	0.6267	0.6586	0.0017	0.2363	0.8025	0.7258	0.5710
102	0.8000	0.6860	0.7357	0.0817	0.2798	0.7886	0.7703	0.7150
103	0.8000	0.7444	0.5596	0.0048	0.1900	0.7605	0.7735	0.6967
104	0.8000	0.6038	0.7135	0.0023	0.1415	0.8011	0.7650	0.7043
105	0.8000	0.5262	0.6525	0.0092	0.2049	0.7890	0.7703	0.6579
106	0.8000	0.7714	0.5934	0.0372	0.0592	0.7780	0.7971	0.7424
107	0.8000	0.6974	0.7030	0.1300	0.3948	0.7798	0.7657	0.6939
108	0.8000	0.6579	0.6593	0.0064	0.1582	0.7999	0.8101	0.7290
109	0.8000	0.6433	0.7931	0.0379	0.2765	0.8022	0.7541	0.6837
110	0.8000	0.6893	0.5927	0.0106	0.2177	0.7811	0.7586	0.6364
111	0.8000	0.7803	0.6964	0.1190	0.4307	0.7897	0.7694	0.7283
112	0.8000	0.5970	0.4903	0.0005	0.1541	0.7676	0.7908	0.7775
113	0.8000	0.5784	0.6663	0.0000	0.2550	0.7951	0.7703	0.7146
114	0.8000	0.5749	0.7046	0.0280	0.3038	0.7938	0.7857	0.6887
115	0.8000	0.6323	0.7446	0.0662	0.3609	0.7961	0.7952	0.7291
116	0.8000	0.7044	0.6962	0.0117	0.1245	0.7891	0.7419	0.7338
117	0.8000	0.6877	0.5420	0.0119	0.1679	0.7664	0.7489	0.7674
118	0.8000	0.6379	0.6454	0.0233	0.1706	0.7692	0.7905	0.7161
119	0.8000	0.6958	0.5512	0.0200	0.1356	0.7724	0.7680	0.7122
120	0.8000	0.6974	0.6260	0.0503	0.0358	0.7773	0.7947	0.7686