

UNIVERSIDADE FEDERAL DO PAMPA

EZEQUIEL LUÍS VIDAL

**OTIMIZAÇÃO DA SELEÇÃO DE
ACASALAMENTOS PARA
MELHORAMENTO GENÉTICO
ANIMAL**

**Bagé
2021**

EZEQUIEL LUÍS VIDAL

**OTIMIZAÇÃO DA SELEÇÃO DE
ACASALAMENTOS PARA
MELHORAMENTO GENÉTICO
ANIMAL**

Dissertação apresentada ao Programa de Pós-Graduação em Computação Aplicada como requisito parcial para a obtenção do título de Mestre em Computação Aplicada.

Orientador: Sandro da Silva Camargo
Co-orientador: Fernando Flores Cardoso

**Bagé
2021**

—

Vidal, Ezequiel Luís

Otimização da Seleção de Acasalamentos para Melhoramento Genético Animal / Ezequiel Luís Vidal. – agosto, 2021.

130 f.: il.

Dissertação (Mestrado) – Universidade Federal do Pampa, Campus Bagé, Programa de Pós-Graduação em Computação Aplicada, 2021.

“Orientação: Sandro da Silva Camargo; Co-orientação: Fernando Flores Cardoso”.

1. Algoritmo genético. 2. Critério de parada. 3. Diferença esperada na progênie. 4. Pecuária de Precisão.
I. Título.

EZEQUIEL LUÍS VIDAL

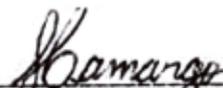
**OTIMIZAÇÃO DA SELEÇÃO DE ACASA-
LAMENTOS PARA MELHORAMENTO
GENÉTICO ANIMAL**

Dissertação apresentada ao Programa de Pós-Graduação em Computação Aplicada como requisito parcial para a obtenção do título de Mestre em Computação Aplicada.

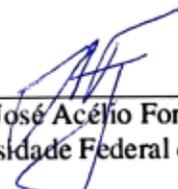
Área de concentração: Tecnologias para a produção agropecuária

Dissertação defendida e aprovada em: 18 de Agosto de 2021.

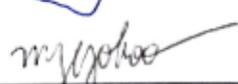
Banca examinadora:



Prof. Dr. Sandro da Silva Camargo
Orientador



Prof. Dr. José Acélio Fontoura Júnior
Universidade Federal do Pampa



Dr. Marcos Jun-Iti Yokoo
EMBRAPA Pecuária Sul



Prof. Dr. Marilton Sanchotene de Aguiar
Universidade Federal de Pelotas

AGRADECIMENTO

Agradeço primeiramente a Deus pelo cuidado e por me dar capacidade. Quero agradecer também aos colegas de mestrado, pela amizade e companhia, e aos professores Sandro da Silva Camargo e Fernando Flores Cardoso pela orientação deste trabalho. Um agradecimento especial à minha esposa Driely, pelo amor, pela força e pelo incentivo.

“O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis.”

— José de Alencar

RESUMO

As transformações demográficas no mundo podem elevar a demanda por produtos da pecuária. Com isso, estratégias que contribuam para uma maior oferta desses produtos, e que agreguem qualidade e sustentabilidade, precisam ser formuladas. O PampaPlus é um programa de melhoramento genético que visa analisar o desempenho genético de bovinos e fornecer ferramentas que permitam o direcionamento dos trabalhos de melhoramento. Uma prática importante neste tipo de programa é a seleção de acasalamentos. Neste sentido, um algoritmo genético foi proposto anteriormente para analisar dados fenotípicos dos animais selecionados e recomendar acasalamentos aos produtores. Experimentos realizados comprovaram a sua eficácia. Porém, a abordagem necessita de melhorias, principalmente de desempenho. O algoritmo utilizava um critério de parada baseado em um número máximo de iterações, mas geralmente atingia a convergência para uma resposta ao problema antes do limite pré-estabelecido. Isso fazia com que executasse por mais tempo e com iterações que não forneciam nenhuma melhoria significativa nos índices genéticos. Neste sentido, foram implementadas duas abordagens de critério de parada encontradas na literatura, sendo criados casos de teste para validação e comparação, cada um com um conjunto de animais. A primeira abordagem não pôde ser validada em diferentes casos de teste. Porém, a segunda abordagem foi validada e apresentou resultados de desempenho com redução de até 50,8% no tempo de processamento em comparação com o critério de parada original, mantendo os mesmos índices de qualificação genética alcançados. Também foram implementadas melhorias na restrição de utilização de touros, pois, em algumas situações, o algoritmo apresentava recomendações que excediam um limite. Após a implementação, os resultados indicaram que a recomendação de acasalamentos dada como resposta pelo algoritmo obedecia a restrição de utilização imposta. Além disso, algumas rotinas e estruturas de dados foram refatoradas para eliminar dependências com as características genéticas medidas e utilizadas no PampaPlus, possibilitando assim a utilização do algoritmo em outros programas de melhoramento. Após a implementação de melhorias, um *benchmark* realizado mostrou que o algoritmo genético obtém recomendações de acasalamentos com melhores índices de qualificação genética quando comparado a outra ferramenta de recomendação de acasalamentos.

Palavras-chave: Algoritmo genético. Critério de parada. Diferença esperada na progênie. Pecuária de Precisão.

ABSTRACT

Demographic changes in the world may increase the demand for livestock products. Thus, strategies that contribute to a greater offer of these products, and that add quality and sustainability, need to be formulated. PampaPlus is a breeding program that aims to analyze the genetic performance of cattle and provide tools that allow the direction of breeding work. An important practice in this type of program is the selection of matings. In this sense, a genetic algorithm was previously proposed to analyze phenotypic data of the selected animals and recommend matings to producers. Experiments performed proved its effectiveness. However, the approach needs improvement, mainly from performance. The algorithm used a stopping criterion based on a maximum number of iterations, but generally reached convergence for an answer to the problem before the pre-established limit. This made it run longer and with iterations that did not provide any significant improvement in genetic indexes. In this sense, two approaches to stopping criteria found in the literature were implemented, and test cases were created for validation and comparison, each with a set of animals. The first approach could not be validated in different test cases. However, the second approach was validated and presented performance results with a reduction of up to 50.8% in processing time compared to the original stopping criterion, maintaining the same genetic qualification indexes achieved. Improvements were also implemented in the restriction of the use of bulls, because, in some situations, the algorithm presented recommendations that exceeded a limit. After implementation, the results indicated that the mating recommendation given as an answer by the algorithm obeyed the usage restriction imposed. In addition, some routines and data structures have been refactored to eliminate dependencies with genetic characteristics measured and used in PampaPlus, thus making it possible to use the algorithm in other breeding programs. After implementing improvements, a benchmark performed showed that the genetic algorithm obtains mating recommendations with better genetic qualification indexes when compared to another mating recommendation tool.

Keywords: Genetic algorithm, Stopping criterion, Expected progeny difference, Precision livestock farming.

LISTA DE FIGURAS

Figura 1	Sumário online do PampaPlusNet.	26
Figura 2	Ferramenta de acasalamento do PampaPlusNet.	27
Figura 3	Matriz de produtos gerada pela ferramenta de acasalamento.	27
Figura 4	Fluxograma de um algoritmo genético.	32
Figura 5	Representação binária de um cromossomo.	33
Figura 6	Exemplo de <i>crossover</i> e mutação.	35
Figura 7	Exemplo de cromossomo do AG.	38
Figura 8	Distribuição de IQG dos animais selecionados no Caso de Teste 1.	49
Figura 9	Gráfico de dispersão das repetições dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) do Caso de Teste 1.	51
Figura 10	Evolução das soluções geradas pelo AG com a abordagem de parada original no Caso de Teste 1.	52
Figura 11	Evolução das soluções geradas pelo AG com a abordagem de parada proposta com limiar de 0,025 no Caso de Teste 1.	52
Figura 12	Distribuição de IQG dos animais selecionados no Caso de Teste 2.	54
Figura 13	Gráfico de dispersão das repetições dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) do Caso de Teste 2.	56
Figura 14	Evolução das soluções geradas pelo AG com a abordagem de parada original no Caso de Teste 2.	56
Figura 15	Evolução das soluções geradas pelo AG com a abordagem de parada proposta com limiar de 0,02 no Caso de Teste 2.	57
Figura 16	Distribuição de IQG dos animais selecionados no Caso de Teste 3.	63
Figura 17	Distribuição de IQG dos animais selecionados no Caso de Teste 4.	66
Figura 18	Distribuição de IQG dos animais selecionados no Caso de Teste 5.	68
Figura 19	Evolução das soluções geradas pelo AG no experimento 1 de utilização de touros.	73
Figura 20	Relação de utilização de touros do experimento 1.	74
Figura 21	Evolução das soluções geradas pelo AG no experimento 2 de utilização de touros.	75
Figura 22	Relação de utilização de touros do experimento 2.	75
Figura 23	Evolução das soluções geradas pelo AG no experimento 3 de utilização de touros.	76
Figura 24	Relação de utilização de touros do experimento 3.	77
Figura 25	Evolução das soluções geradas pelo AG no experimento 4 de utilização de touros.	77
Figura 26	Relação de utilização de touros do experimento 4.	78
Figura 27	Evolução das soluções geradas pelo AG no experimento 5 de utilização de touros.	79
Figura 28	Relação de utilização de touros do experimento 5.	79
Figura 29	Evolução das soluções geradas pelo AG no experimento 6 de utilização de touros.	80
Figura 30	Relação de utilização de touros do experimento 6.	81
Figura 31	Tela de escolha de animais da interface <i>web</i>	82
Figura 32	Tabela de banco de dados dos animais selecionados nos dois casos de teste.	82
Figura 33	Tela de listagem de animais selecionados da interface <i>web</i>	83
Figura 34	Tela de escolha de DEPs e edição do IQG da interface <i>web</i> para o Caso de Teste 1.	83

Figura 35	Tabela de banco de dados das DEPs selecionadas no Caso de Teste 1.	84
Figura 36	Descrição de DEPs cadastradas no banco de dados.	85
Figura 37	Tela de recomendação de acasalamentos da interface <i>web</i>	86
Figura 38	Recomendação de acasalamentos no banco de dados para o Caso de Teste 1.	86
Figura 39	Tela de escolha de DEPs e edição do IQG da interface <i>web</i> para o Caso de Teste 2.	87
Figura 40	Tabela de banco de dados das DEPs selecionadas no Caso de Teste 2.	88
Figura 41	Recomendação de acasalamentos no banco de dados para o Caso de Teste 2.	88
Figura 42	Relação de utilização de touros em uma das execuções do AG no Experimento 1 do <i>Benchmark</i>	91
Figura 43	Relação de utilização de touros em uma das execuções do EVA no Experimento 1 do <i>Benchmark</i>	92
Figura 44	Relação de acasalamentos e consanguinidades em uma das execuções do AG no Experimento 2 do <i>Benchmark</i>	93
Figura 45	Relação de acasalamentos e consanguinidades em uma das execuções do EVA no Experimento 2 do <i>Benchmark</i>	94
Figura 46	Relação de acasalamentos e consanguinidades em uma das execuções do AG no Experimento 3 do <i>Benchmark</i>	94
Figura 47	Relação de acasalamentos e consanguinidades em uma das execuções do EVA no Experimento 3 do <i>Benchmark</i>	95
Figura 48	Modelo conceitual no contexto de um sistema computacional.	104
Figura 49	Protótipo da tela de escolha dos touros.	107
Figura 50	Aba de selecionados do protótipo da tela de escolha dos touros.	108
Figura 51	Protótipo da tela de escolha das vacas.	108
Figura 52	Protótipo da tela de escolha dos parâmetros do AG.	109
Figura 53	Protótipo de tela de andamento da execução do AG.	110
Figura 54	Protótipo de tela de recomendação de acasalamentos.	111

LISTA DE TABELAS

Tabela 1	Parâmetros utilizados nos experimentos do Caso de Teste 1.	48
Tabela 2	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) no Caso de Teste 1.	50
Tabela 3	Parâmetros utilizados nos experimentos do Caso de Teste 2.	53
Tabela 4	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) no Caso de Teste 2.	55
Tabela 5	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 1.	60
Tabela 6	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 2.	62
Tabela 7	Parâmetros utilizados nos experimentos do Caso de Teste 3.	63
Tabela 8	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 3.	64
Tabela 9	Parâmetros utilizados nos experimentos do Caso de Teste 4.	65
Tabela 10	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 4.	67
Tabela 11	Parâmetros utilizados nos experimentos do Caso de Teste 5.	68
Tabela 12	Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 5.	69
Tabela 13	Parâmetros utilizados nos experimentos de utilização de touros.	72
Tabela 14	Parâmetros utilizados em comum pelas ferramentas de acasalamento nos experimentos do <i>Benchmark</i>	90
Tabela 15	Média dos dados resultantes dos experimentos do <i>Benchmark</i>	90

LISTA DE ABREVIATURAS E SIGLAS

ABHB	Associação Brasileira de Hereford e Braford
AG	Algoritmo Genético
BLUP	Melhor Predição Linear Não Viesada
EVA	<i>Evolutionary Algorithm</i>
DEP	Diferença Esperada na Progenie
IQG	Índice de Qualificação Genética
NP	Nível de Problema
PP	Pecuária de Precisão

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Justificativa	16
1.2 Problema de Pesquisa	17
1.3 Objetivos	17
1.4 Organização do trabalho	18
2 REVISÃO DE LITERATURA	19
2.1 Pecuária de Precisão	19
2.2 Melhoramento Genético	20
2.2.1 Abordagens para o Melhoramento Genético	21
2.2.2 Programa PampaPlus	23
2.2.3 Explosão Combinatorial no Melhoramento Genético.....	28
2.3 Problemas de Busca em Espaço de Estados	29
2.3.1 Estratégias de Busca	29
2.3.1.1 Busca Cega.....	30
2.3.1.2 Busca Heurística	31
2.3.2 Algoritmos Genéticos	31
2.3.2.1 Representação dos Indivíduos.....	33
2.3.2.2 População Inicial.....	33
2.3.2.3 Função de <i>Fitness</i>	33
2.3.2.4 Seleção	34
2.3.2.5 Operadores de <i>Crossover</i> e Mutação	34
2.3.2.6 Condição de Parada e Convergência.....	35
2.3.3 Seleção de Acasalamentos Baseada em Algoritmos Genéticos.....	37
2.3.3.1 Critério de Parada	40
2.3.3.2 Convergência.....	40
2.4 Trabalhos Correlatos	40
3 METODOLOGIA	44
3.1 Caracterização da Pesquisa	44
3.2 Material e Métodos	44
4 RESULTADOS E DISCUSSÃO	47
4.1 Critério de Parada	47
4.1.1 Critério de Parada de Yeng, Yoke and Suhaimi (2019)	47
4.1.1.1 Caso de Teste 1	48
4.1.1.2 Caso de Teste 2	53
4.1.1.3 Discussão	58
4.1.2 Critério de Parada de Bhandari, Murthy and Pal (2012).....	58
4.1.2.1 Caso de Teste 1	59
4.1.2.2 Caso de Teste 2	61
4.1.2.3 Caso de Teste 3	63
4.1.2.4 Caso de Teste 4	65
4.1.2.5 Caso de Teste 5	67
4.1.2.6 Discussão	70
4.2 Restrição de Utilização de Touros	71
4.3 Eliminação de <i>Hardcode</i> no AG	80
4.4 Benchmark do Algoritmo Genético	89
5 CONCLUSÕES E TRABALHOS FUTUROS	96
5.1 Trabalhos Futuros	97
REFERÊNCIAS	98

APÊNDICE A — INTERFACE DA FERRAMENTA	102
A.1 Interface Homem-Computador	102
A.1.1 Princípios de Interfaces	103
A.1.2 Ferramenta <i>Web</i> para Seleção de Acasalamentos	105
A.2 Protótipos de Interface	106
A.3 Documento de Requisitos	112
A.3.1 Introdução	112
A.3.1.1 Descrição Geral do Sistema	112
A.3.1.2 Usuários de Requisitos de Sistema	112
A.3.2 Requisitos Funcionais	113
A.3.3 Requisitos Não-Funcionais	114
A.3.4 Restrições	115
A.3.5 Referencial	115
A.4 Plano de Testes	115
A.4.1 Introdução	115
A.4.1.1 Documento de Referência	116
A.4.2 Estratégia de Testes	116
A.4.3 Objetivos e Prioridades do Sistema.....	116
A.4.4 Casos de Teste	117
A.4.5 Referencial	130

1 INTRODUÇÃO

Estima-se que, no ano de 2050, a população mundial seja superior a 9,5 bilhões de pessoas. De 2012 a 2050, esses números representariam um crescimento populacional de 34,90% (UNITED NATIONS DEPARTMENT OF ECONOMICS AND SOCIAL AFFAIRS, 2013). Essa rápida e contínua expansão populacional elevaria a demanda por alimentos em geral. Em 2018, em torno de 11% da população mundial não possuía alimentos suficientes, o que representava cerca de 820 milhões de pessoas (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, 2019).

Aliados à expansão populacional no mundo, outros fatores relevantes na demanda mundial por alimentos tendem a ser: a maior concentração de pessoas nos centros urbanos, pois mais 2,5 bilhões de pessoas poderiam migrar para estes centros até 2050 e; em consequência disso, um crescimento da renda per capita e mudanças no comportamento de consumo (UNITED NATIONS DEPARTMENT OF ECONOMICS AND SOCIAL AFFAIRS, 2015).

As transformações demográficas no mundo podem elevar a demanda por produtos da pecuária (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, 2009). Com isso, estratégias que contribuem para uma maior oferta desses produtos, e que agreguem qualidade e sustentabilidade, precisam ser formuladas (HERRERO; THORNTON, 2013).

Segundo dados da Associação Brasileira das Indústrias Exportadoras de Carnes (ABIEC), em 2018, o Brasil se consolidou como o maior exportador de carne bovina do mundo, quando foram exportadas 1,64 milhões de toneladas, uma diferença de 11% do registrado em 2017 (ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNES, 2020). A introdução de técnicas para a melhoria da genética do rebanho, como fecundação *in vitro*, produção de embriões, clonagem, entre outras, foi uma das maiores contribuições para que a pecuária bovina brasileira se consolidasse no cenário mundial. Atualmente, a aplicação de inovações tecnológicas que monitoram os animais e seus ambientes vem sendo utilizada com o conceito de Pecuária de Precisão e, nesse contexto, programas de melhoramento genético foram implementados no Brasil.

O objetivo de um programa de melhoramento genético é alcançar a melhoria futura dos rebanhos sob a perspectiva das características genéticas desejáveis, como taxa de crescimento e rendimento de produtos como carne ou leite, por exemplo (ROSA; MENEZES; SILVA, 2013). Através da seleção de animais adequados, é possível ter níveis

de produtividade e qualidade em sintonia com o sistema de produção e com aquilo que o mercado exige (ROSA; MENEZES; EGITO, 2013). Técnicas estatísticas de modelo misto e modelo animal que utilizam a metodologia conhecida por Melhor Predição Linear Não Viesada (BLUP) são utilizadas nas avaliações genéticas de muitos programas. Através destas técnicas, é possível expressar os valores genéticos dos animais na forma de Diferença Esperada na Progenie (DEP). Uma DEP prediz a habilidade de transmissão genética de um animal em comparação com a média do rebanho para determinada característica. Seu valor é expresso na unidade da característica avaliada, como Kg de peso ao nascer, por exemplo (LOBO; LÔBO, 2007). Dentre os programas de melhoramento existentes no Brasil que utilizam as DEPs, está o PampaPlus.

O PampaPlus é um programa de melhoramento genético para raças bovinas da Associação Brasileira de Hereford e Braford (ABHB). Em parceria com a Embrapa Pecuária Sul, o programa visa analisar o desempenho genético dos rebanhos e fornecer, aos criadores associados, ferramentas que permitam o direcionamento dos trabalhos de melhoramento. O programa conta com um Índice de Qualificação Genética (IQG), onde as características de maior interesse do programa são definidas e calculadas com base nos valores ponderados das DEPs. Com isso, um único valor de classificação é fornecido para cada animal. Para subsidiar a realização dos cálculos dos valores genéticos dos animais, inspetores técnicos do PampaPlus coletam medidas objetivas e de escores visuais dos animais nas propriedades associadas e, através do portal *web* PampaPlusNet, os usuários cadastrados realizam a inclusão dos dados coletados que podem ser acessados posteriormente. Um sumário *online* dos touros também é disponibilizado. Nele, constam as DEPs de todos os touros avaliados no programa. (CARDOSO; LOPA, 2016).

Com o passar dos anos, o PampaPlus vem aprimorando a genética dos seus plantéis, atingindo maiores níveis de eficiência produtiva e, conseqüentemente, incrementando a rentabilidade da atividade pecuária. Não apenas isso, vem ao encontro do objetivo de aumentar a produção de alimentos, com qualidade, para suprir a demanda cada vez mais crescente.

Uma prática importante para o melhoramento genético dentro de um programa como o PampaPlus é a seleção dos acasalamentos entre os animais. O produtor associado tem o objetivo de receber sugestões de acasalamento pautados por uma abordagem estatística, relacionando suas matrizes com reprodutores específicos, touros ou sêmen, disponíveis no programa (FONTOURA, 2019). Dado um conjunto de animais selecionados, é necessário realizar uma busca por uma combinação de acasalamentos que maximize

a média do IQG utilizado para aquele grupo, e que atenda algumas restrições, como de consanguinidade.

Segundo Abbass, Towsey and Finn (1998), o problema da alocação de parceiros é considerado um problema de otimização combinatória. Um exemplo seria selecionar 10 touros e 100 vacas. O espaço de busca do problema teria o total de 10 elevado a 100 respostas candidatas, evidenciando uma explosão combinatorial. Este tipo de problema pode se tornar intratável do ponto de vista computacional à medida que o número de animais aumenta, sendo impossível resolvê-lo de maneira ótima em tempo polinomial (GAREY; JOHNSON, 1979).

No trabalho de Fontoura, Camargo and Cardoso (2019), foi implementada uma estratégia heurística para resolver o problema no âmbito do programa PampaPlus. A solução é baseada na utilização de um Algoritmo Genético (AG) para analisar o conjunto das variáveis envolvidas no processo de seleção e recomendar acasalamentos ao produtor. Um conjunto de animais pré-selecionados é fornecido como entrada e restrições de consanguinidade e utilização de touros são impostas. A cada iteração do AG são formadas respostas candidatas ao problema, representando diferentes combinações de acasalamentos. Uma função de avaliação é utilizada para calcular o IQG de cada combinação e, ao final de execução do AG, aquela que maximiza a função de avaliação é retornada como solução ao problema no formato de uma lista de acasalamentos.

1.1 Justificativa

De acordo com Fontoura (2019), o AG necessita de melhorias, principalmente de desempenho. O algoritmo atinge a convergência para uma resposta ao problema perto das 800 iterações. Porém, para realizar a parada, o critério utilizado baseia-se em um número máximo de 1000 iterações. Como o AG executa até atingir um número fixo de iterações, há um acréscimo de cerca de 20% no tempo de processamento sem necessidade pois, neste tempo, as demais iterações do algoritmo não fornecem nenhuma melhoria significativa nos resultados. Além disso, o ideal é que o produtor aguarde o menor tempo possível por uma recomendação de acasalamentos. Por isso, estudar e implementar um critério de parada mais adequado poderá melhorar o tempo de execução do AG.

Também, segundo técnicos e pesquisadores da Embrapa, há a necessidade de melhoria na restrição de utilização de touros, pois o algoritmo apresenta recomendações que excedem o limite imposto em alguns casos de teste. Além disso, o AG apresenta um có-

digo restrito à utilização pelo Pampaplus, e não poderia ser utilizado em programas que observam outras DEPs e utilizam índices de qualificação genéticos diferentes. Neste sentido, faz-se necessário ajustes no código do AG, e que também possibilitassem uma futura integração com interface *web*, pois o algoritmo receberia os parâmetros selecionados pelos usuários, independente do programa de melhoramento, para realizar a recomendação de acasalamentos.

Como o AG foi desenvolvido para resolver um problema de otimização combinatória, a implementação de um novo critério de parada poderia trazer ganho de desempenho ao algoritmo, visto que o tempo de processamento do AG aumenta exponencialmente em função do tamanho do conjunto de animais pré-selecionados. Acrescentando a implementação das demais melhorias, o AG poderia ser utilizado para suprir a necessidade por uma abordagem que oriente os criadores quanto aos acasalamentos entre os seus animais, aumentando assim o ganho genético do programa de melhoramento PampaPlus.

1.2 Problema de Pesquisa

O problema de pesquisa deste trabalho consiste em responder a seguinte questão: A implementação de um novo critério de parada e o aperfeiçoamento do código do AG proposto por Fontoura, Camargo and Cardoso (2019) podem otimizar a seleção de acasalamentos no âmbito do programa PampaPlus?

1.3 Objetivos

O objetivo geral deste trabalho é realizar implementações de melhorias no AG desenvolvido por Fontoura, Camargo and Cardoso (2019), de modo que o algoritmo possa ser utilizado para apoiar o melhoramento genético dos plantéis integrantes do programa PampaPlus. Os objetivos específicos deste trabalho são:

- a) Investigar e implementar um novo critério de parada para o AG;
- b) Refatorar o código do AG para eliminar problemas relacionados à restrição de utilização de touros;
- c) Refatorar o código do AG para permitir sua utilização em diferentes programas de melhoramento genético e possibilitar uma integração com interface *web*; e

- d) Realizar um *benchmark* do AG com outra ferramenta de seleção de acasalamentos.

1.4 Organização do trabalho

O presente trabalho está organizado em cinco Capítulos. No Capítulo 2, é realizada uma revisão de literatura sobre os tópicos mais relevantes para este trabalho, sendo que a abordagem de AG desenvolvida por Fontoura, Camargo and Cardoso (2019) é descrita com mais detalhes. O Capítulo 3 expõe o material e os métodos utilizados, além da caracterização da pesquisa e suas fases. No Capítulo 4, os resultados são apresentados e discutidos. O Capítulo 5 traz as conclusões e trabalhos futuros.

2 REVISÃO DE LITERATURA

Neste Capítulo é realizada uma revisão de literatura sobre os tópicos de interesse deste trabalho. Na seção 2.1, a Pecuária de Precisão é apresentada. A seção 2.2 trata sobre o Melhoramento Genético voltado à pecuária, sendo que algumas abordagens são exploradas e o programa de melhoramento PampaPlus é apresentado. Na seção 2.3, o Problema de Busca em Espaço de Estados é abordado juntamente com algumas estratégias que podem ser utilizadas, entre elas os Algoritmos Genéticos. Por fim, os Trabalhos Correlatos são apresentados na seção 2.4, e uma discussão de relevância para este trabalho é realizada.

2.1 Pecuária de Precisão

As exigências atuais e futuras do mercado de alimentos requerem uma adequação dos sistemas de produção. Isso, aliado à busca por eficiência produtiva, demanda uma melhoria da gestão de toda cadeia produtiva. Nesse contexto, a Pecuária de Precisão (PP) surgiu para promover a integração de diversas áreas de conhecimento com o objetivo de aperfeiçoar a captação e interpretação dos dados envolvidos nos processos produtivos e, com isso, otimizar a produção animal (RIBAS et al., 2017).

A PP pode ser definida como: aplicação à pecuária de princípios e técnicas de engenharia de processo no monitoramento, modelagem e gerenciamento automático da produção animal, convertendo os dados em informações relevantes para serem aplicadas em diferentes aspectos do manejo, com foco no animal e no meio ambiente (TULLO et al., 2017). Conhecer o comportamento ingestivo do animal, através do monitoramento dos animais e de características nas pastagens, pode servir como base para a tomada de decisão (CARVALHO et al., 2009). O emprego de tecnologias de informação e comunicação avançadas, o uso direcionado dos recursos e o controle preciso do processo de produção, além de aumentar o bem-estar animal e humano, traz maior rentabilidade à atividade pecuária (BANHAZI et al., 2012).

Através da simplificação dos processos de coleta, processamento e análise dos dados coletados e do gerenciamento eficiente das informações geradas, os produtores podem receber soluções para os seus problemas (BANHAZI et al., 2012). Vários sinais relativos à atividade fisiológica do animal podem ser medidos, tais como o consumo de alimentos, os movimentos, o peso vivo, entre outros (WATHES et al., 2008). Segundo Tullo, Finzi

and Guarino (2019), após a identificação e classificação dos dados coletados dos animais, a etapa a seguir seria o desenvolvimento de um algoritmo para modelar as bio-respostas, o qual deve ser validado para verificar se as informações geradas correspondem à realidade e são úteis ao produtor.

A aplicação da PP se torna cada vez mais frequente para aumentar a eficiência produtiva dos animais. Utilizar tecnologias de precisão na avaliação da eficiência alimentar, por exemplo, é uma oportunidade para estudar a variabilidade individual e acelerar o processo de seleção e melhoramento genético dos animais (CAMPOS et al., 2015). Também, o emprego da PP para conhecer os parâmetros da curva de crescimento e desenvolvimento dos animais é importante para que os programas de melhoramento alterem a forma desta curva conforme seus objetivos (RIBAS et al., 2017). Portanto, de acordo com o que foi visto, o conceito de PP pode ser aplicada no melhoramento genético de animais.

2.2 Melhoramento Genético

O melhoramento genético é o processo de acasalamento seletivo de animais com características genéticas desejáveis, para manter ou melhorar essas características nas gerações futuras. Para o gado, por exemplo, isso envolve a estimativa do valor genético de indivíduos com base em suas características, incluindo a taxa de crescimento e o rendimento de produtos como leite ou carne. Isto tende a mensurar o quão superior um animal é do outro, ou da média dos outros animais da raça, em termos de valor genético para as características desejadas, propiciando a seleção de animais e, com isso, o melhoramento genético do rebanho (ROSA; MENEZES; SILVA, 2013).

Através do melhoramento genético busca-se atingir uma maior produtividade com qualidade agregada, de modo que satisfaça as exigências impostas pelo mercado. Para atingir tal objetivo, características animais como peso corporal, curva de crescimento, qualidade de carcaça e carne, entre outras, devem ser monitoradas (ROSA; MENEZES; EGITO, 2013).

Os autores apresentam a equação básica do melhoramento animal. Esta, observa as características dos animais e é conhecida por fenótipo (P). O fenótipo depende do genótipo do animal (G), que é a composição genética compartilhada pelo pai e pela mãe no momento da fecundação; do ambiente (E), no qual este animal está sendo criado e também da interação entre genótipo e ambiente (GxE), que consiste da expressão dos genótipos afetados pelas diferentes condições ambientais de criação. Considerando estas

relações conjuntamente, temos então a equação: $P = G + E + G \times E$. Com base na equação apresentada, vemos não só a contribuição da transmissão dos genes dos pais para os filhos, mas também de como o ambiente interfere no melhoramento dos animais. Nesse sentido, escolher os animais adaptados ao ambiente de criação pode ser o primeiro passo para implementação do melhoramento genético. A seguir, algumas abordagens utilizadas são descritas.

2.2.1 Abordagens para o Melhoramento Genético

Em Lobo and Lôbo (2007), os autores apontam duas abordagens básicas do melhoramento genético: a seleção de progenitores e a aplicação de métodos de acasalamento entre os selecionados. A seleção consiste em escolher os animais para reprodução que apresentam as características desejadas pelo produtor para transmiti-las para as gerações seguintes. Os animais são selecionados de acordo com seu fenótipo, que representa o seu desempenho quanto às características avaliadas, ou pelo genótipo, de acordo com a medida de potencial genético. Para isto, torna-se necessário formar bancos de dados confiáveis, a partir de um número grande de animais. As informações sobre os animais devem advir de uma rigorosa escrituração zootécnica, onde o máximo possível de dados confiáveis sejam coletados.

A partir dos dados coletados, avaliações genéticas podem ser realizadas através de técnicas estatísticas de modelo misto e modelo animal, utilizando a metodologia BLUP, que reúne toda informação coletada para aquela característica. Um valor genético pode ser expresso por uma DEP, que possui sinal positivo ou negativo de acordo com um valor médio, e permite prever a habilidade de transmissão genética de um animal comparando a sua progênie com a de outros animais.

Outras abordagens que podem ser utilizadas, segundo os autores, são os métodos de acasalamento como cruzamento e endogamia. Estas abordagens alteram as frequências dos genes nos rebanhos, podendo promover o melhoramento genético se forem bem concebidas. O cruzamento é um método que apoia-se no acasalamento de indivíduos de genética ou raças diferentes, tendo como resultado produtos mestiços. Esse método é utilizado para obter uma progênie superior em relação a média das características observadas nos pais, ou também para reunir, em um único animal, as características de duas ou mais raças.

A endogamia, também chamada de consanguinidade, fundamenta-se no acasala-

mento de indivíduos que possuem genética semelhante, com ancestrais em comum. Esse acasalamento promove o incremento da homozigose, que consiste no aumento da tendência genética para a característica desejada. Porém, esse método pode promover não apenas o aumento das características desejadas, mas também das indesejadas, acarretando em problemas como redução da fertilidade, da sobrevivência e do vigor dos animais.

No trabalho de Pires (2011), o objetivo foi planejar e implementar um programa de melhoramento genético de ovinos. Para isso, o autor apontou tópicos que considera primordiais para traçar as estratégias de um programa. Em primeiro lugar, é necessário definir os objetivos, ou seja, quais são os tipos de animais a serem desenvolvidos e o que produzir a partir deles.

Após a primeira etapa, deve-se determinar os caminhos para alcançar os objetivos do programa. Isso compreende definir, não só o conjunto das características que serão observadas nos animais, mas também os métodos utilizados para predizer o mérito genético. A identificação dos animais com mérito superior fornece os dados necessários para as estratégias de seleção e cruzamento, que também precisam ser definidas.

O autor comenta que também é necessário definir a maneira de disseminação do material genético. Para isso, estruturas de núcleo no formato de pirâmide podem ser utilizadas. Estas estruturas fornecem um controle sobre todos os animais e menores custos na coleta de dados. Os animais são divididos em estratos como, o núcleo do rebanho no topo da pirâmide, os multiplicadores do rebanho ao centro e o rebanho comercial na base. Os esforços na coleta de dados são concentrados no topo da estrutura, do qual se adquire a genética superior para ser disseminada às demais camadas.

Existem duas estruturas de disseminação, de núcleo aberto e de núcleo fechado. A estrutura de núcleo aberto possui livre fluxo animal entre os estratos. Por exemplo, se um animal do estrato comercial apresentar genética superior, ele poderá ser levado ao estrato núcleo. Também, esta estrutura permite a importação de material genético entre os estratos. Já na estrutura de núcleo fechado o fluxo de animais se dá somente na direção do estrato do topo da pirâmide para o estrato mais inferior, e não permite importação de material genético.

Por fim, o autor apresenta a última grande decisão a tomar quando se planeja uma estrutura de melhoramento genético. Consiste do uso de tecnologias reprodutivas, ou seja, a escolha de qual tecnologia será utilizada para a disseminação do mérito genético dos animais entre os estratos. Existe uma ampla gama de tecnologias como inseminação artificial, transferência de embrião ou clonagem. Porém, é necessário observar o custo

benefício de para cada adoção.

Já o trabalho de Nobre et al. (2013), diferente do trabalho de Pires (2011), apresenta um programa de melhoramento genético para bovinos de corte já implementado, o GenePlus. O objetivo do programa é aumentar a genética desejável nos indivíduos que serão pais nas futuras gerações, considerando o rebanho, isoladamente, ou a raça. Para alcançar o objetivo, o procedimento empregado possui duas fases. Na primeira fase, uma recomendação de indivíduos é realizada por profissionais que prestam assessoria aos produtores. E na segunda fase, o indivíduo avaliado é mantido ou descartado segundo o poder de decisão e responsabilidade do produtor.

O programa baseia-se nos dados coletados junto aos criadores. Nesse sentido, os autores enfatizam a necessidade de ter escrituração zootécnica e coleta de dados devidamente planejadas. No GenePlus, as mensurações sugeridas podem ser quantitativas, como pesos em determinadas idades, ou qualitativas, através de escores visuais atribuídos a cada animal individualmente para características como, por exemplo, a condição corporal das matrizes ao parto. Comenta-se também que quanto maior for o volume dos dados, atrelado à qualidade, maior será a precisão da identificação de indivíduos geneticamente superiores. Isso aperfeiçoa o processo de escolha dos melhores candidatos a pais das gerações futuras.

Segundo os autores, com os avanços da informática e das metodologias e aplicativos utilizados nas avaliações genéticas, informações cada vez mais precisas são geradas. Os dados coletados possibilitam estimar as DEPs, que serão utilizadas como critério de seleção ou de descarte dos indivíduos. Por fim, o programa Geneplus propõe a classificação dos animais segundo o IQG, agregando em um único valor classificatório as estimativas dos valores genéticos dos indivíduos expressas em DEPs. O IQG, expresso em unidades de desvios-padrão, é ponderado de acordo com os graus de importância das características desejadas no objetivo geral de seleção. Sendo assim, as informações geradas podem ser utilizadas na seleção dos animais com potencial para aumentar o mérito genético do rebanho nas futuras gerações.

2.2.2 Programa PampaPlus

O PampaPlus é um programa de avaliação e melhoramento genético para as raças bovinas Hereford e Braford. O programa fornece, aos associados da ABHB, ferramentas de informática que apoiam o processo de melhoramento genético dos animais e possibili-

tam alcançar os objetivos mais rapidamente e com eficiência . O programa é constituído de Conselho Técnico, Inspectores Técnicos e Comitê de Apoio composto por pesquisadores convidados e criadores do próprio programa. Essa estrutura permite o direcionamento do melhoramento genético dos rebanhos na busca de animais com melhor desempenho e rendimento financeiro. O Laboratório de Bioinformática e Estatística Genômica (LABEGEN) da EMBRAPA Pecuária Sul em convênio com o programa de melhoramento genético GENEPLUS, desenvolvido pela EMBRAPA Gado de Corte, realizam as avaliações genéticas dos animais. (CARDOSO; LOPA, 2016).

Para que as avaliações possam ser realizadas, os Inspectores Técnicos do programa coletam medidas objetivas e de escores visuais dos animais nas propriedades associadas. Algumas medidas são obrigatórias e outras são opcionais. As medidas obrigatórias são fundamentais para predizer a contribuição genética dos animais através das DEPs. As DEPs são utilizadas no cálculo do IQG do PampaPlus, um valor classificatório único para cada animal. Por outro lado, as medidas opcionais dizem respeito a outras características importantes das raças que não entram no cálculo do IQG, e sim no IQG Ampliado, utilizadas por criadores que buscam por melhorias adicionais através de um nível maior de controle (CARDOSO; LOPA, 2016).

As características que fazem parte do IQG, com seus respectivos graus de importância, são apresentadas a seguir.

Características produtivas (85% do IQG):

- Peso Total Materno à Desmama (*kg*) 30%;
- Ganho de Peso Pós Desmama (*kg*) 15%;
- Peso ao Sobreano (*kg*) 15%;
- Escore de Musculatura (1-5) 12,5%;
- Escore de Estatura (1-5) 12,5%.

Características reprodutivas (15% do IQG):

- Circunferência Escrotal (*cm*) 15%.

Demais características aferidas no PampaPlus:

- Acabamento (Espessura de gordura subcutânea em *mm*);
- Área de olho de lombo ao Sobreano (*cm*²);

- Escore de Condição Corporal da Vaca à Desmama (1-5)
- Escore de Estrutura Corporal (1-5);
- Escore de Pigmentação Ocular (1-5);
- Escore Racial (1-5);
- Escore de Umbigo (1-5);
- Gordura Intramuscular (%);
- Peso à Desmama (Kg);
- Peso à Desmama Materno (Kg);
- Peso da Vaca à Desmama (Kg);
- Peso ao Nascer (Kg).

No portal *web* PampaPlusNet é disponibilizado o sumário *online*, um relatório com os valores das avaliações genéticas dos touros. Embora seja gratuito e sem necessidade de criar-se um perfil, algumas funcionalidades estão disponíveis apenas a usuários cadastrados, como a inclusão de dados coletados nas propriedades. As informações obtidas dão subsídio à realização dos cálculos dos valores genéticos que permitem prever a contribuição do animal. A Figura 1 apresenta o sumário disponibilizado no PampaPlusNet.

Outras opções disponíveis no PampaPlusNet são:

- Consultar todos os dados de animais da propriedade;
- Gerar catálogo, em formato PDF ou Excel, de animais pré-selecionados em que o criador escolhe quais DEPs deseja apresentar;
- Cadastrar os Comunicados de “Cobertura das Fêmeas por Estação de Monta”, de sua propriedade, junto a ABHB;
- “Ativar” touros pais ou mães doadoras para torna-los disponíveis no cadastro de coberturas;
- Cadastrar grupos de reprodutores múltiplos junto a ABHB e torná-los disponíveis no cadastro de coberturas;

Figura 1 – Sumário online do PampaPlusNet.

Registro	Tatuagem	Nasc	Nome	Estabelecimento	Raça	Livro	GS	NReb	NF	IQG	%	Tat Pai	Nome Pai	Tat Mãe	Nome Mãe
143119	D1369	2008	SANTA TEREZA 38-D1369 MONTREAL	FAZENDA SANTA TEREZA	BR	CG	38	7	66	4.2	0.1	8188	NELIUS SHAKA 38-81105 (SIDNEY)	X0535	SANTA TEREZA 38-X0535
194602	H2134	2012	SANTA TEREZA 38-H2134 MARQUÊS	FAZENDA SANTA TEREZA	BR	CG	38	3	37	3.6	0.1	D1369	SANTA TEREZA 38-D1369 MONTREAL	C1254	SANTA TEREZA 38-C1254
181526	TE1001	2011	SERENO 38-TE1001	AGROPECUARIA SERENO	BR	CG	38	2	207	3.3	0.1	9309	SAO MIGUEL 38-9309 TALURA	3026	NOVA ESPERANCA 38-3026
170551	L231	2009	RECLUTA L231 PATRIOTA	RECLUTA	PH	PC		9	360	3.3	0.1	F41	RECLUTA F41	G017	RECLUTA G017
282109	4005	2014	SERENO 38-4005	AGROPECUARIA SERENO	BR	CG	38	2	25	3.2	0.1	TE1001	SERENO 38-TE1001	1218	SERENO 38-1218
236983	M059	2013	DO BOLSO M059 ÚNICO	ESTANCIA DO BOLSO	PH	PC		5	132	3.0	0.5	X3605	PALEQUERO DOMINANTE 117	H08	DO BOLSO H08
IA-B043	316	2007	RANCHO 316 TIE MAGNO		BR	ES	38	19	847	2.8	0.5	792	ALBERTA LIGHT FOOT 792	105 TIE	RANCHO 105 TIE
167016	K170	2008	RECLUTA K170	ESTANCIA TAMANCA	PH	PC		4	600	2.8	0.5	E124	RECLUTA E124	X687	RECLUTA X687

Fonte: PampaPlusNet

- Cadastrar receptoras de sua propriedade registradas ou não;
- Cadastrar os “Comunicados de Nascimentos” dos produtos perante a ABHB tornando-os imediatamente disponíveis para cadastro das avaliações fenotípicas à desmama e/ou ao sobreano;
- Cadastrar as avaliações fenotípicas realizadas ao Desmame e ao Sobreano dos animais que possuem registro genealógico, também daqueles que não se enquadram para registro ou mesmo aqueles que ainda não foram registrados;
- Consultar protocolos de cadastramento, consultar ou ativar touros pais ou mães doadoras;
- Consultar Sumários de touros ou touros jovens on-line, onde constam as DEPs de todos os touros avaliados no PampaPlus.

O sistema PampaPlusNet também possui uma ferramenta para seleção de acasaelamentos, ainda em fase de testes. A tela principal da ferramenta é mostrada na Figura 2. Conforme pode ser observado na figura, é permitida a seleção dos animais através de alguns filtros como: gênero, raça, ano/safra e valores de DEPs desejados. Pode-se realizar também uma busca por nome, tatuagem ou registro do animal. O usuário pode ainda editar o IQG, ponderando as características de seu interesse.

Figura 2 – Ferramenta de acasalamento do PampaPlusNet.

Fonte: PampaPlusNet

Após a seleção e confirmação, é exibida na tela uma matriz de simulação dos produtos (filhos) gerados através dos animais selecionados (pais). A matriz traz os produtos gerados de todos com todos, ou seja, se a seleção for de n touros e m vacas, tem-se uma matriz $n \times m$ de produtos. A Figura 3 apresenta a matriz gerada pela ferramenta de acasalamento para quatro touros e dez vacas.

Figura 3 – Matriz de produtos gerada pela ferramenta de acasalamento.

	VACAS / TOUROS			
	M831 174341	M187 174345	A012	M157 176699
A101 2010	IQG: 2.48 CONS: 0.00%	IQG: 1.98 CONS: 0.00%	IQG: 1.55 CONS: 13.00%	IQG: 1.48 CONS: 0.00%
A084 2010	IQG: 2.41 CONS: 0.00%	IQG: 1.91 CONS: 0.00%	IQG: 1.45 CONS: 13.00%	IQG: 1.48 CONS: 0.00%
A018 2010	IQG: 2.29 CONS: 0.00%	IQG: 1.79 CONS: 0.00%	IQG: 1.33 CONS: 13.00%	IQG: 1.38 CONS: 0.00%
A234 2010	IQG: 2.23 CONS: 0.00%	IQG: 1.73 CONS: 0.00%	IQG: 1.28 CONS: 0.00%	IQG: 1.24 CONS: 0.00%
A055 2010	IQG: 2.23 CONS: 0.00%	IQG: 1.73 CONS: 0.00%	IQG: 1.28 CONS: 13.00%	IQG: 1.29 CONS: 0.00%
A094 2010	IQG: 2 CONS: 0.00%	IQG: 1.49 CONS: 0.00%	IQG: 1.06 CONS: 13.00%	IQG: 1.09 CONS: 0.00%
A432 2010	IQG: 2.07 CONS: 0.00%	IQG: 1.65 CONS: 0.00%	IQG: 1.11 CONS: 0.00%	IQG: 1.15 CONS: 0.00%
A121 2010	IQG: 2.05 CONS: 0.00%	IQG: 1.54 CONS: 0.00%	IQG: 1.1 CONS: 0.00%	IQG: 1.13 CONS: 0.00%
A013 2010	IQG: 2.11 CONS: 0.00%	IQG: 1.6 CONS: 0.00%	IQG: 1.17 CONS: 13.00%	IQG: 1.05 CONS: 0.00%
A109 2010	IQG: 1.97 CONS: 0.00%	IQG: 1.46 CONS: 0.00%	IQG: 1.01 CONS: 0.00%	IQG: 1.05 CONS: 0.00%

Fonte: PampaPlusNet

Como pode ser visto na figura, é mostrada uma matriz de resultados com os possíveis produtos gerados, os filhos, observando todos os acasalamentos possíveis entre todos os pais. Porém, quando o usuário seleciona um número grande de animais, a visualização dos resultados fica prejudicada. Além disso, esta ferramenta não provê como resultado uma configuração de acasalamentos que maximiza a média do IQG entre os animais selecionados. Esta não é uma tarefa trivial, pois a medida que o número de animais selecionados aumenta, ocorre uma explosão combinatorial de possíveis acasalamentos entre os animais.

2.2.3 Explosão Combinatorial no Melhoramento Genético

Definida uma abordagem para seleção dos animais conforme os objetivos de melhoramento genético, a etapa seguinte corresponde ao acasalamento dos indivíduos selecionados, ou seja, quais animais acasalarão com quais. Com base nas avaliações individuais dos animais elaboradas a partir de dados coletados, é realizada uma análise dos acasalamentos com maior potencial de mérito genético na progênie. Cada acasalamento possui um valor agregado, que deve ser observado em conjunto com os outros acasalamentos. O objetivo é selecionar os acasalamentos que produzem os melhores resultados em termos da média do grupo, e não individuais (KINGHORN, 1987).

Realizar a seleção dos animais para acasalamento visando atingir o melhoramento genético em um rebanho não é uma tarefa trivial quando muitas variáveis estão envolvidas no processo. O problema de Seleção de Parceiros (*Mate Selection*, no inglês) consiste em escolher pares de animais ou até grupos para acasalamento realizando os processos de seleção e acasalamento simultaneamente. No processo, é realizada uma análise com múltiplas variáveis. As variáveis envolvidas podem ser referentes ao ganho genético, à diversidade, à endogamia, entre outras, além de restrições como a mínima e máxima utilização de um touro, por exemplo. Vários algoritmos foram propostos para resolver o problema, baseados em técnicas heurísticas como algoritmos evolucionários e recozimento simulado (KINGHORN, 2011).

No trabalho de Fontoura (2019), uma heurística baseada em algoritmos genéticos foi aplicada para resolver o problema de Seleção de Parceiros, sendo descrita com mais detalhes na seção 2.3.3. Segundo o autor, selecionar os animais e encontrar a melhor configuração de acasalamentos com o objetivo de aumentar o mérito genético na progênie e, respeitando restrições de utilização de touros e consanguinidade, foi considerado um

problema intratável do ponto de vista computacional. Se 100 vacas e 10 touros fossem selecionados, o espaço de busca do problema teria o total de 100 elevado a 10 respostas candidatas, resultando em um alto número e evidenciando a explosão combinatorial presente neste tipo de problema.

2.3 Problemas de Busca em Espaço de Estados

A Seleção de Parceiros está entre os problemas de busca em espaço de estados. Um estado é uma configuração dos elementos de análise. Neste caso, os elementos são os pares de animais escolhidos para acasalamento. O espaço de estados consiste de todas as configurações possíveis que são respostas candidatas ao problema. O problema é definido como: um conjunto de estados; um estado inicial e um conjunto de estados objetivos e; um conjunto de ações (ou operadores) que mapeiam estados para estados sucessores. O objetivo é encontrar uma sequência de ações que transformam o estado inicial em um estado objetivo otimizando alguma medida de qualidade (HANSEN; ZILBERSTEIN, 2001).

Este tipo de problema é considerado NP-difícil, quando não é conhecido um algoritmo que resolva em tempo polinomial. Na busca da melhor solução, seria preciso uma exploração sistemática de todas as configurações possíveis de seus elementos. Porém, na maioria das aplicações de interesse, o tamanho do problema proíbe a enumeração explícita de todos os estados e operações antes do início do procedimento de pesquisa. Nessas aplicações, normalmente o tempo de solução aumenta de forma exponencial com o tamanho do problema (SUN et al., 2011).

Nesse sentido, estratégias devem ser empregadas na busca pela solução do problema. A seguir, são apresentadas algumas estratégias de busca em espaço de estados.

2.3.1 Estratégias de Busca

Tratando-se de estratégias de busca em espaço de estados, dois grupos principais podem ser citados: Busca Cega e Busca Heurística. De acordo com Poole and Mackworth (2017), as estratégias de Busca Cega (*Uninformed Search*, no inglês) não possuem informação sobre o estado objetivo. Já a Busca Heurística (*Heuristic Search*, no inglês), leva em consideração informações sobre o estado objetivo. Nessas estratégias, o problema é modelado na forma de grafo de estados.

Um grafo é formado por um conjunto de vértices, ou nós, e um conjunto de arestas. Os vértices representam os estados do problema e as arestas os operadores ou ações aplicadas em um estado. Cada aresta está associada a dois vértices, representando a transformação de um estado em outro estado do espaço de busca (HANSEN; ZILBERSTEIN, 1998). A seguir, a Busca Cega é descrita com mais detalhes.

2.3.1.1 Busca Cega

As estratégias de Busca Cega mais utilizadas são os algoritmos de Busca em Largura e Busca em Profundidade. Estas estratégias não levam em consideração o estado objetivo e seguem um padrão de busca até encontrar uma solução que satisfaça o problema.

A Busca em Largura, também conhecida como BFS (do inglês *Breadth-First Search*), começa em um vértice inicial do grafo, ou vértice raiz, e explora todos os vértices vizinhos na profundidade atual antes de passar para os vértices no próximo nível de profundidade. A profundidade de um vértice é a distância dele para o vértice raiz, em número de passos, sendo que o vértice raiz possui profundidade 0. Esta estratégia não é adequada quando todas as soluções possuem muitas arestas ou há alguma heurística disponível. Também não é utilizada com frequência em grandes problemas em que o grafo é gerado dinamicamente devido a um espaço de estados muito grande. (POOLE; MACKWORTH, 2017).

Já a Busca em Profundidade, ou DFS (do inglês *Depth-First Search*) usa a estratégia oposta, uma pesquisa aprofundada, que explora a ramificação do vértice o máximo possível antes de ser forçada a voltar e expandir para outros vértices. Não é uma estratégia adequada se as soluções estiverem em profundidade rasa, pois gasta-se tempo percorrendo todo o resto do grafo. Outro problema é se existirem vários caminhos para um mesmo vértice. Um caminho se dá ao percorrer um grafo sem repetição de vértices. Esse tipo de busca gastaria mais tempo percorrendo a profundidade de todos os caminhos para o mesmo vértice. Esta estratégia também não é adequada quando há ciclos no grafo. Um ciclo acontece quando percorre-se vários vértices e atinge-se um vértice já visitado, ou seja, trata-se de um caminho fechado. Com isso, é possível que a busca fique em um *loop* (POOLE; MACKWORTH, 2017).

2.3.1.2 Busca Heurística

A Busca Heurística é uma estratégia de busca informada, pois leva em consideração informações sobre o estado objetivo. Informações heurísticas sobre quais vértices são mais promissores para atingir um vértice objetivo podem orientar a pesquisa, alterando qual vértice é selecionado. Para isso, uma função heurística é utilizada. Para um vértice do grafo, a função calcula um número real não negativo que é uma estimativa do caminho de menor custo deste vértice para um vértice objetivo. Um custo de um caminho é a soma dos custos de cada vértice percorrido. Trata-se de um número atrelado a cada aresta e depende da modelagem do problema e dos seus objetivos. A função é uma heurística admissível se seu resultado for sempre menor ou igual ao custo real de um caminho de menor custo do vértice investigado para um vértice objetivo (POOLE; MACKWORTH, 2017).

Uma estratégia bastante eficiente de Busca Heurística é o algoritmo A^* , que combina duas estratégias de busca. A primeira estratégia é a busca pelo menor custo, que minimiza o custo do caminho percorrido até o estado corrente. A segunda estratégia é a busca pela melhor estimativa, que tenta minimizar o custo estimado do caminho a ser percorrido do estado corrente até um estado objetivo. O algoritmo A^* minimiza o custo total do caminho percorrido, de um estado inicial até um estado objetivo, através da função $f(s) = g(s) + h(s)$, onde $g(s)$ é uma função de custo, implementando a busca pelo menor custo, e $h(s)$ é uma função heurística, implementando a busca pela melhor estimativa (POOLE; MACKWORTH, 2017).

2.3.2 Algoritmos Genéticos

Na busca de uma solução em problemas que possuem um espaço de estados muito grande, utilizar uma estratégia de técnica heurística é a opção mais adequada frente à complexidade do problema. Nesse sentido, os Algoritmos Genéticos podem ser aplicados no problema de busca em espaço de estados.

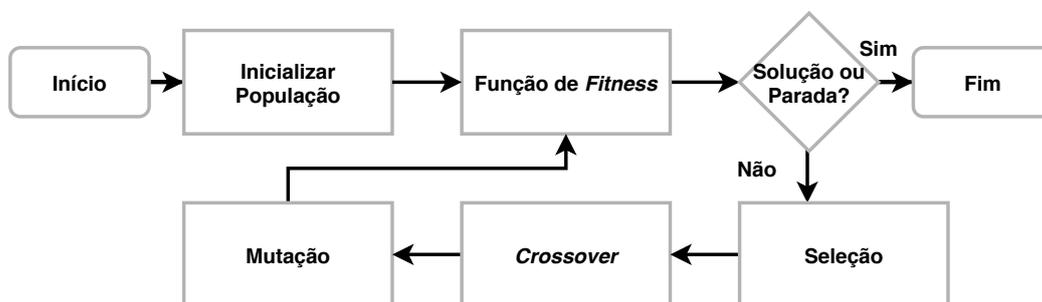
Os AGs utilizam técnicas de otimização estocástica no qual os estados se originam a partir de eventos aleatórios, baseando-se na mecânica da evolução natural e sobrevivência da estratégia mais adequada encontrada em organismos biológicos (TOĞAN; DALOĞLU, 2008). Pautados nos mecanismos de seleção natural e da genética, os AGs foram propostos pela primeira vez por John Holland em meados dos anos 70, quando pes-

quisador na Universidade de Michigan. Desde então, e especialmente a partir da década de 1980, os AGs têm sido aplicados a uma variedade de aplicações (ABBASS; TOWSEY; FINN, 1998).

Os AGs iniciam com um conjunto de soluções formadas aleatoriamente chamado de população inicial. Cada indivíduo da população é chamado de cromossomo, representando uma solução para o problema em questão. Um cromossomo é uma codificação, que pode ser uma série de bits, números inteiros ou números reais por exemplo, utilizados conforme a representação do problema. Os cromossomos evoluem através de iterações sucessivas, chamadas de gerações. Durante cada geração, os cromossomos são avaliados por uma função de *fitness*, que é uma medida de aptidão segundo os parâmetros de avaliação utilizados. Os melhores cromossomos, de acordo com a função de *fitness*, passam por uma seleção para criar a próxima geração (GEN; LIN, 2007). Na seleção, explicada na seção 2.3.2.4, são utilizadas técnicas como a roleta viciada e o torneio.

Novos cromossomos, chamados de filhos, são gerados pela fusão de dois cromossomos pais da geração atual usando um operador de cruzamento (*crossover*) e/ou modificando um cromossomo usando um operador de mutação. Cada par de cromossomos pai gera um novo par de cromossomos filhos, sendo que a nova geração de cromossomos substitui a anterior. A cada nova geração, a tendência é que o *fitness* médio da população aumente e, após várias iterações, o algoritmo converge para uma população onde espera-se que o seu cromossomo melhor avaliado represente a solução ótima ou subótima para o problema (GEN; LIN, 2007). A Figura 4 apresenta o fluxograma de um AG:

Figura 4 – Fluxograma de um algoritmo genético.



Fonte: Autor(2021)

A seguir, é realizada uma descrição mais detalhada da representação dos indivíduos e também de cada etapa envolvida no fluxo de um AG.

2.3.2.1 Representação dos Indivíduos

A representação dos indivíduos da população consiste na modelagem do problema real de modo que ele possa ser tratado computacionalmente. Estruturas de dados do tipo vetor ou lista podem ser utilizadas para representar os indivíduos, chamados de cromossomos. Cada posição de um cromossomo é chamada de gene, como na biologia, e possui um tipo de informação que varia de acordo com o problema tratado. Um gene pode conter binários, números inteiros, reais, entre outros. Com isso, a função de *fitness* e os operadores genéticos são concebidos de acordo com o conteúdo desses genes (LINDEN, 2012). A Figura 5 apresenta um exemplo de representação binária de um cromossomo com oito genes.

Figura 5 – Representação binária de um cromossomo.

1	0	1	1	0	1	1	0
---	---	---	---	---	---	---	---

Fonte: Autor(2021)

2.3.2.2 População Inicial

Para que o processo de evolução do AG aconteça, uma população deve ser gerada. O modo de geração da população inicial afeta criticamente a convergência, o desempenho e a capacidade do AG. Quando não há condição de gerar todos os indivíduos da população inicial a partir do espaço de solução do problema, a geração de indivíduos é aleatória (TOĞAN; DALOĞLU, 2008).

Além da geração aleatória, a inicialização pode ser realizada aplicando-se uma heurística. Com isso, o *fitness* médio da população inicial tende a ser alto, o que ajuda os AGs a encontrar soluções mais rapidamente. Porém, na maioria dos problemas de larga escala apenas uma pequena parte do espaço da solução é explorado, tornando difícil a busca pelas melhores soluções devido à falta de diversidade na população (GEN; LIN, 2007).

2.3.2.3 Função de Fitness

A avaliação de aptidão do cromossomo segundo os parâmetros de otimização observados é realizada ao verificar-se o valor de uma função objetivo conhecida como função

de *fitness*. Esta função fornece um mecanismo de avaliação para cada indivíduo na população. Para manter a uniformidade sobre vários domínios de problemas, pode-se usar a função de *fitness* normalizada para um intervalo de 0 a 1. O valor normalizado é utilizado pelo mecanismo de seleção para avaliar os indivíduos da população (GEN; LIN, 2007).

2.3.2.4 Seleção

Um operador de seleção é utilizado para melhorar a aptidão média da população dando aos cromossomos melhores avaliados maiores chances de serem utilizados para gerar a próxima geração. Em cada iteração do AG, a população passa por uma evolução na sua aptidão, pois soluções relativamente boas se reproduzem e soluções relativamente ruins são eliminadas, de modo que a prole composta pelas boas soluções é reproduzida (GEN; LIN, 2007).

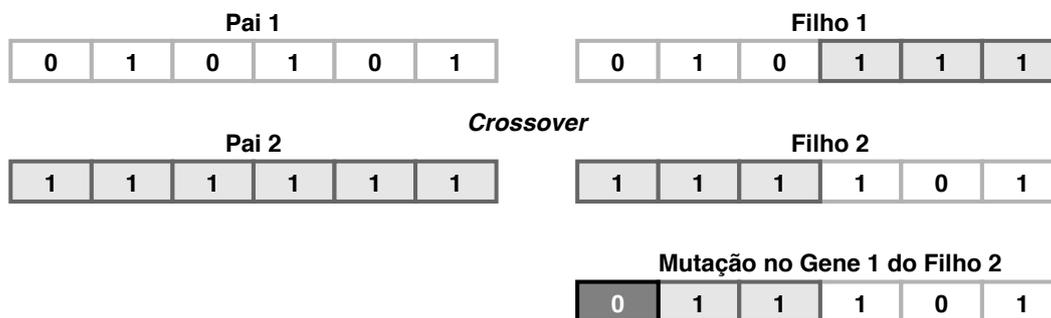
Os tipos de seleção mais comuns são: seleção por roleta viciada, onde cada indivíduo recebe um percentual de uma roleta sendo que o indivíduo melhor avaliado pela função de *fitness* possui maior probabilidade de ser selecionado; seleção por torneio, que utiliza os valores de *fitness* a fim de realizar confrontos entre os indivíduos; e a seleção elitista, utilizada para garantir que uma porção de determinados indivíduos superiores se mantenham nas próximas gerações (LINDEN, 2012).

2.3.2.5 Operadores de Crossover e Mutação

O operador de cruzamento, ou *crossover*, é o principal operador genético. Ele opera em dois cromossomos por vez e gera dois descendentes combinando os genes de ambos os cromossomos. Uma maneira simples de realizar o *crossover* para gerar os descendentes é escolher um ponto de corte aleatório e combinar o segmento de um pai à esquerda do ponto de corte com o segmento do outro pai à direita do ponto de corte (GEN; LIN, 2007).

Após o *crossover*, alguns genes dos cromossomos são selecionados, de acordo com uma taxa de mutação, para terem seus valores alterados. Se o valor do gene for binário igual a 1 e a mutação for determinada, o gene terá seu valor alterado para 0 (LINDEN, 2012). A seguir, a Figura 6 apresenta um exemplo de dois cromossomos pais em um cruzamento para gerarem dois filhos, e também a mutação de um gene em um dos filhos.

Figura 6 – Exemplo de *crossover* e mutação.



Fonte: Adaptado de (FONTOURA, 2019)

2.3.2.6 Condição de Parada e Convergência

Existe um número significativo de trabalhos sobre componentes de algoritmos evolucionários como os operadores de seleção, *crossover* e mutação. Por outro lado, o assunto sobre decidir o melhor momento de parada desses algoritmos é pouco explorado na literatura. O critério de parada mais utilizado é interromper a execução do algoritmo quando um número de iterações/gerações pré-definido é excedido. Porém, essa não é uma tarefa trivial, pois torna-se fundamental estimar o número de iterações necessárias para que o AG encontre uma solução (LIU; ZHOU; ZHANG, 2018).

Aytug and Koehler (2000) modelaram um AG através de cadeias de *Markov* e a partir desse modelo derivaram um limite suficiente de iterações para garantir que o algoritmo iterasse por todas as soluções possíveis e encontrasse uma solução ótima. O mesmo foi feito por Pendharkar and Koehler (2007), porém, os autores comentam que trabalhar com tamanhos grandes de população traz sérios desafios computacionais uma vez que o tamanho da matriz de transição da cadeia de *Markov* aumenta muito rápido. Por isso, a computação de valores de parada se torna problemática.

De acordo com Liu, Zhou and Zhang (2018), definir um limite de iterações depende do problema a ser tratado e é difícil fazê-lo na prática. Se o limite de iterações for muito pequeno, a maioria dos algoritmos podem não encontrar uma boa solução. Se o limite for muito grande, pode haver desperdício computacional e de tempo.

Para Bhandari, Murthy and Pal (2012), um critério que pode ser utilizado para a parada do algoritmo, além do critério de fixar um limite de iterações, é o critério baseado no valor de *fitness*. Nesse critério, o algoritmo é encerrado se nenhuma melhoria adicional no valor de *fitness* do melhor cromossomo avaliado for observada após várias gerações.

Porém, definir o número de gerações onde não há melhorias também é um desafio. Por isso, os autores propuseram um critério de parada baseado na variância dos melhores valores de *fitness* obtidos ao longo de todas as gerações.

O critério de parada utiliza uma função que possui variáveis para o intervalo de gerações (IG) observado, os melhores *fitness* de cada geração (MF_i) e a média dos melhores *fitness* (MF). A função é mostrada a seguir:

$$((1/IG) \cdot \sum_{i=1}^{IG} (MF_i - MF)^2) < \text{limiar}$$

Os autores demonstraram teoricamente que a variância tende a zero quando o número de iterações tende ao infinito. Ou seja, conforme o AG realiza as gerações, ele encontra as melhores soluções até que a população atinja a convergência e nenhuma melhoria adicional ocorra. Nesse sentido, o critério de parada observa um limite de variância pré-fixado, próximo a zero, e encerra a execução do AG quando a variância calculada atinge este limiar. Nenhuma forma automática de escolher esse valor de limiar é dada pelos autores e depende da precisão desejada em cada problema, sendo um problema desafiador e assunto para pesquisas futuras.

No trabalho de Yeng, Yoke and Suhaimi (2019), os autores desenvolveram uma proposta de critério de parada também baseada na convergência da população porém observando a sua saturação de aptidão. Ou seja, a proposta observa quando toda a população converge para uma solução ao problema e não acontecem melhorias adicionais no valor de *fitness* médio da população. A cada iteração do AG, o cálculo da saturação é realizado utilizando uma função que observa o tamanho da população (TP), o *fitness* de cada cromossomo (F_i) e o *fitness* médio (F) da geração atual. O critério de parada observa um limiar para interromper a execução do AG quando a variação de *fitness* nos cromossomos é pequena. Como na proposta de Bhandari, Murthy and Pal (2012), esse limiar é um valor próximo a zero e também não foi definido pelos autores pois depende do problema em cada aplicação. A função utilizada nesta proposta é a mostrada a seguir:

$$((1/TP) \cdot \sum_{i=1}^{TP} (F_i - F)^2) < \text{limiar}$$

Os resultados mostraram que o desempenho do critério de parada proposto é superior quando comparado ao critério de parada convencional baseado em um número pré-fixado de iterações.

2.3.3 Seleção de Acasalamentos Baseada em Algoritmos Genéticos

No trabalho de Fontoura (2019) foi desenvolvida uma solução baseada em AGs voltada à pecuária de corte para selecionar e recomendar acasalamentos no âmbito do programa de melhoramento genético PampaPlus. O autor justifica que a procura por combinações de acasalamentos considerando múltiplas características genéticas na busca pela maximização do índice de qualidade genética tornava o problema inviável de ser tratado em tempo polinomial, sendo necessário então aplicar uma estratégia heurística.

O objetivo principal do AG é melhorar a escolha de acasalamentos com base nas DEPs. Para isso, a solução analisa o conjunto principal das variáveis envolvidas no processo para então recomendar os acasalamentos ao produtor. As variáveis analisadas são medidas objetivas e de escores visuais dos animais descritas com mais detalhes na seção 2.2.2.

Os dados de entrada do algoritmo são os animais escolhidos para acasalamento. Estes são modelados em vetores, os cromossomos, com tamanho igual à quantidade de vacas escolhidas. Ou seja, se o usuário escolher 100 vacas, o cromossomo será um vetor de 100 posições. Cada posição representa uma vaca (gene), porém os cromossomos não guardam informações de vacas e, para isso, o AG utiliza outras estruturas de dados. Estas estruturas mantêm dados de DEPs tanto de vacas como de touros e são utilizadas em funções que as observam em paralelo com os vetores de cromossomos e que são providas pela linguagem de programação utilizada na implementação do AG.

A população do AG é composta por n cromossomos. Para formar a população inicial, n é igual ao dobro do número de vacas selecionadas, ou seja, se forem escolhidas 100 vacas, haverá 200 cromossomos e n se manterá constante até o fim da execução do AG. Segundo Fontoura (2019), o tamanho da população inicial de cromossomos foi definido com base nos trabalhos correlatos.

Para simular um acasalamento, cada posição do cromossomo recebe o identificador de um touro (alelo), sendo que os touros podem ser utilizados mais ou menos vezes de acordo com limites de utilização informados pelo usuário. Cada cromossomo tem um valor de aptidão calculado e representa uma solução candidata dentre os demais indivíduos da população. A abordagem busca como resposta o cromossomo melhor avaliado e que também seja válido a algumas restrições. A Figura 7 ilustra um exemplo de cromossomo onde cada posição representa uma vaca e cada informação contida nela é um identificador para um touro.

Figura 7 – Exemplo de cromossomo do AG.

Cromossomo	454	111	205	102	102	Ids de Touro
	334	341	344	348	351	<i>Ids de Vaca</i>

Fonte: Adaptado de (FONTOURA, 2019)

O *fitness*, ou aptidão, de cada cromossomo consiste na observância de dois componentes com diferentes pesos, o IQG (90%) e o Nível de Problema - NP (10%). O IQG é o índice utilizado no programa de melhoramento genético e consiste em uma equação que leva em conta as DEPs analisadas nos animais. Para cada acasalamento, são calculadas as médias de DEPs entre os animais, sendo que cada média de DEP é multiplicada por seu fator de relevância dentro do IQG. A seguir, é apresentada a equação utilizada por Fontoura (2019) para calcular o IQG de cada acasalamento do cromossomo.

$$IQG_{\text{acasalamento}} = ((DEP_1\text{vaca} + DEP_1\text{touro})/2) \cdot DEP_{1\text{param}} + \dots \\ + ((DEP_n\text{vaca} + DEP_n\text{touro})/2) \cdot DEP_{n\text{param}}$$

Então, a soma destas médias ponderadas de cada característica de cada par vaca-touro (gene x alelo) finalmente é o valor de IQG calculado para aquele cromossomo. Já o NP observa, para cada acasalamento, as DEPs inferiores à média do rebanho e, através de uma padronização que leva em conta a variabilidade das DEPs no grupo, calcula o desvio padrão de cada DEP. Assim, um acasalamento tem 1 NP se possui um desvio padrão para uma DEP, 2 NP para duas, e assim por diante. Se o acasalamento possuir alguma DEP com mais de um desvio padrão da média, o NP é elevado a um fator, de valor 1,2, aumentando ainda mais. A equação utilizada para calcular o NP final de cada acasalamento do cromossomo é mostrada a seguir, onde MNP e DPNP são a média e o desvio padrão de NP para todos os acasalamentos do cromossomo respectivamente.

$$NP_{\text{acasalamento}} = (NP_{\text{acasalamento}} - MNP)/DPNP$$

As restrições impostas que devem ser observadas nos cromossomos candidatos à resposta são: a máxima e mínima utilização dos touros e; a consanguinidade máxima. A restrição quanto a utilização máxima dos touros deve-se ao fato de que um touro pode

acasalar com um número limitado de vacas. Já a utilização mínima é empregada quando o produtor deseja um número mínimo de acasalamentos por touro. Quanto à restrição de consanguinidade, é imposta uma taxa máxima pois o acasalamento entre parentes pode trazer efeitos negativos na progênie, como perda de vigor. Caso não atenda essas restrições, o cromossomo é penalizado com a perda de 50% no seu valor de *fitness*. Isso permite que sejam gerados descendentes de cromossomos que atendem as restrições com mais frequência.

As restrições são implementadas através de estruturas de dados auxiliares. Há um vetor com o número mínimo e máximo de utilização para cada touro e outro vetor contendo a taxa de consanguinidade entre os animais. Isso não garante que aconteça alguma quebra de restrição, sendo então aplicada a penalidade já mencionada.

Após os cromossomos serem avaliados quanto à sua aptidão, o AG parte para a etapa de seleção, que inicialmente se dá por uma roleta viciada, onde os cromossomos melhores avaliados possuem maior probabilidade de serem selecionados. De dois em dois, os cromossomos são pré-selecionados por roleta e em seguida passam pela seleção por torneio, no qual são comparados e o cromossomo que tiver maior aptidão é selecionado para ser um pai na próxima etapa de *crossover*.

No *crossover*, a partir de um ponto de corte escolhido de forma aleatória, os cromossomos filhos são gerados através da contribuição dos seguimentos dos cromossomos pais, similar ao descrito na seção 2.3.2.5. Após isso, a mutação é realizada nos cromossomos filhos conforme um grau de probabilidade (0-100% das posições). Na mutação, ocorre a alteração do touro previamente escolhido em alguma posição do cromossomo filho observando-se a restrição de consanguinidade. Um lista de touros compatíveis é mantida pelo AG para que esta restrição não seja quebrada. O algoritmo executa várias iterações, onde cada iteração corresponde a uma geração de cromossomos, interrompendo sua execução ao atingir uma condição de parada, que é um número máximo de gerações. Após isso, o cromossomo da última geração com maior valor de aptidão é dado como resposta pelo AG.

Simulações foram realizadas para validação do AG, sendo que as configurações de acasalamentos de animais dadas como resposta maximizaram os valores da função objetivo. O autor comenta que esta não é uma abordagem final, pois o AG necessita de implementações adicionais. As próximas seções abordam algumas questões neste aspecto.

2.3.3.1 Critério de Parada

No AG de Fontoura (2019), o critério de parada adotado foi o número máximo de iterações do algoritmo. O AG possui condição de parada inicial de 1000 gerações e, ao final destas, o melhor cromossomo válido é indicado como combinação para acasalamento. A justificativa do autor para esta escolha foi de que os testes iniciais mostravam convergência do AG antes das 800 gerações. Então, para validação, esta condição de parada foi adotada.

2.3.3.2 Convergência

Fontoura (2019) comenta que a população inicial foi implementada de forma não aleatória em sua totalidade e orientada à resolução do problema, observando desde já as regras impostas de máxima e mínima utilização de touros e consanguinidade. Mas apenas isso não foi suficiente para visualizar a convergência do AG em busca da maximização da função objetivo. Foi necessário penalizar os cromossomos que quebravam as regras, fazendo com que cromossomos válidos pudessem ser gerados com maior frequência.

Porém, mesmo penalizando os inválidos, se a forma de seleção para o torneio fosse apenas aleatória, ainda não seria possível atingir os resultados pretendidos. Deste modo, como uma segunda forma de seleção, foi implementada também a roleta viciada, com maior probabilidade de seleção para os cromossomos com maior aptidão. Com a utilização destas formas de seleção conjuntas, mesmo com o AG apresentando uma convergência rápida, os resultados foram satisfatórios.

2.4 Trabalhos Correlatos

No trabalho de Berg, Nielsen and Sørensen (2006) é apresentado o *EVA Program*, uma ferramenta que surgiu da necessidade de monitorar e controlar o risco de consanguinidade em populações de animais no contexto da agropecuária. O *software* utiliza um algoritmo evolucionário (EVA, do inglês, *Evolutionary Algorithm*) para fornecer uma recomendação de acasalamentos que busca manter um equilíbrio entre o avanço genético e a consanguinidade no rebanho. O EVA é disponibilizado de forma gratuita no portal da NordGen ¹, uma organização que possui um centro de estudos de recursos genéticos para os países nórdicos, com sede na Suécia.

¹<https://www.nordgen.org/en/our-work/nordgen-farm-animals/eva-program/>

O *software* descreve o histórico da população de animais selecionada através de coeficientes de consanguinidade e de contribuição genética e busca recomendar acasalamentos de contribuições ótimas, maximizando o valor genético na prole e minimizando a consanguinidade. Para isso, é preciso fornecer ao EVA informações de linhagem confiáveis e consistentes através da edição de arquivos externos que serão lidos pelo *software*, descrevendo informações genéticas dos pais, de utilização máxima e do valor de qualificação genética do animal. Também, é possível realizar ajustes nos parâmetros utilizados pelo algoritmo evolucionário, como número máximo de gerações, tamanho da população de cromossomos, taxa de *crossover* e taxa de mutação. Segundo os autores, a ferramenta foi testada e utilizada com sucesso em populações de animais nos países nórdicos e bálticos.

Carvalho, Queiroz and Kinghorn (2010) utilizaram a técnica de Diferencial Evolutivo (DE) no desenvolvimento de um *software* de aplicação para guiar a seleção de animais para acasalamento através da determinação de contribuições genéticas ótimas. O DE consiste de um método para busca global utilizando a heurística e conceitos amplamente difundidos na computação evolutiva (STORN; PRICE, 1997). A função objetivo a ser otimizada foi composta por DEPs e pela consanguinidade média dos animais em reprodução, com penalização para as soluções que quebravam as restrições de consanguinidade impostas.

Para validar e testar o desempenho do *software* desenvolvido, foram utilizados conjuntos de dados reais e simulados de populações. A solução foi considerada computacionalmente eficiente e viável para ser aplicada na prática. De acordo com os autores, quando comparada a procedimentos empíricos de controle de consanguinidade e/ou com a seleção baseada apenas no valor de mérito genético da progênie, a solução desenvolvida tem o potencial de melhorar o mérito genético limitando mais efetivamente a taxa de consanguinidade. Portanto, os autores concluem que é possível utilizar a evolução diferencial como um método de otimização para executar a seleção ideal de contribuição genética.

Kinghorn (2011) apresenta uma extensão para o algoritmo de seleção de acasalamentos apresentado por Carvalho, Queiroz and Kinghorn (2010). Através da implementação de uma matriz de permissão de acasalamento em grupo, é possível antecipar a aplicação das restrições e não envolver soluções que quebrem as restrições, diferente da abordagem anterior que penalizava as soluções ilegais que surgiam durante a análise.

As restrições impostas neste algoritmo são implementadas por meio das variáveis *Maxuse*, *Minuse* e *AbsMinuse*. A variável *Maxuse* recebe o valor correspondente à utili-

zação máxima de um animal. Por exemplo, essa variável pode receber o valor 1 para as vacas ou 30 para os touros em acasalamento naturais. Para touros de inseminação artificial, essa variável pode receber 1000; *Minuse* recebe o valor de utilização mínima de um animal. Por exemplo, se um touro tiver o valor 15 nessa variável, esse touro terá que ter um grupo feminino para acasalamento composto de no mínimo 15 vacas; Já *AbsMinuse*, corresponde ao valor absoluto de utilização de um animal, tendo valor zero se não utilizada. Essa variável pode receber o número de doses de sêmen disponíveis, por exemplo, quando se deseja utilizar o número exato de doses.

O método utiliza um algoritmo baseado nos algoritmos evolucionários e possui os componentes de representação do problema, de função objetivo e, de otimização. O componente de representação utiliza um vetor de números, semelhante a um cromossomo de AG, para representar o problema. O componente de função objetivo, similar à função de *fitness*, avalia cada representação candidata à solução através do cálculo de sua aptidão. O componente de otimização utiliza o valor da função objetivo para selecionar, modificar e recombinar as representações candidatas gerando outras a partir delas. Ao final de um processo iterativo, o método chega à uma solução, que é uma lista de acasalamentos. Segundo o autor, resolver o problema através de componentes separados é uma vantagem. O problema pode se tornar cada vez mais complexo sem ser preciso aumentar a complexidade do componente de otimização.

O algoritmo proposto permite aplicações com maior velocidade de computação, sendo centenas de vezes mais rápido que o método utilizado anteriormente de penalizar soluções que infringem restrições. Isso viabiliza a implementação da seleção de acasalamento de animais em programas relativamente grandes entre suas unidades de criação. Intitulado Matesel/GroupFix, o algoritmo é executado no portal Matesel ², que é utilizado pelos produtores de bovinos de corte da Austrália participantes de associações com programas de melhoramentos genéticos cadastrados.

Barreto Neto (2014) aplicou a teoria da contribuição genética para o melhoramento das futuras gerações de ovinos dos núcleos de seleção associados ao Programa de Melhoramento Genético da Raça Santa Inês – ASCCO/USP, um programa de pedigree estruturado e com valores genéticos de características de importância econômica estimados através de DEPs.

Através do banco de dados do programa, foram obtidas as informações de ascendência e a DEP para peso aos 60 dias. Com isso, um AG foi utilizado para encontrar a

²www.matesel.com

contribuição genética ótima para a próxima geração dos animais, através da maximização de uma função objetivo que observa o aumento da DEP utilizada enquanto penaliza o aumento de consanguinidade.

Os resultados comprovaram a eficácia da utilização de seleção de animais por meio de DEPs e AGs. Segundo o autor, o AG é uma ferramenta com potencial de ser utilizada para aumentar o intercâmbio de material genético entre os núcleos de animais, levando ao ganho genético e mantendo o controle sobre a consanguinidade em programas de melhoramento que já calculam o valor genético dos animais em DEPs. Porém, um problema a ser resolvido é que a exigência computacional aumenta exponencialmente de acordo com o número de animais inseridos, o que, em bancos de dados grandes, pode ser um sério empecilho.

Todos os trabalhos revisados aplicaram restrições de consanguinidade e a maioria utilizou DEPs, com exceção do trabalho de Berg, Nielsen and Sørensen (2006). Carvalho, Queiroz and Kinghorn (2010), Kinghorn (2011) e Berg, Nielsen and Sørensen (2006), mesmo utilizando as técnicas de computação evolutiva, não fizeram uso de AGs. Já o trabalho de Neto (2014), utilizou AG apenas para encontrar a contribuição genética ótima para a próxima geração de animais, mas não para fornecer uma lista de acasalamentos.

O trabalho de Fontoura (2019), abordado na seção 2.3.3, com o objetivo de recomendar acasalamentos aos criadores associados da ABHB e participantes do programa PampaPlus, utilizou AG, DEPs, restrições de consanguinidade e de utilização máxima e mínima de touros, além da flexibilidade de manutenção pelo usuário da função objetivo utilizando as DEPs. Apenas o trabalho de Fontoura (2019) reuniu todos estes requisitos. Contudo, todos os trabalhos revisados fundamentam utilizar ferramentas, como as heurísticas da computação evolutiva, na busca do melhoramento genético dos animais através de características de maior interesse para os produtores.

3 METODOLOGIA

Este Capítulo aborda a metodologia utilizada neste trabalho, apresentando uma caracterização da pesquisa e uma exposição do material e dos métodos utilizados.

3.1 Caracterização da Pesquisa

De acordo com Silva and Menezes (2001), uma pesquisa científica pode ser classificada do ponto de vista de sua natureza, da forma de abordagem do problema, de seus objetivos e dos procedimentos técnicos. Quanto à natureza, a pesquisa realizada neste trabalho é classificada como aplicada, pois visa gerar conhecimentos científicos para serem aplicados na resolução de um problema concreto.

Quanto à forma de abordagem do problema, é tida como quantitativa pois se baseia em medidas numéricas como instrumento na abordagem proposta e também em sua avaliação, tendo como exemplo os parâmetros que são utilizados pelo AG em sua função objetivo. Quanto aos objetivos, é uma pesquisa explicativa pois, a partir da análise de resultados, explicará os fenômenos envolvidos na utilização da abordagem proposta. Por fim, quanto aos seus procedimentos técnicos, esta pesquisa pode ser classificada como experimental pois será validada no ambiente computacional do PampaPlus e em sua base de dados.

3.2 Material e Métodos

O AG, descrito com mais detalhes na seção 2.3.3, é escrito na linguagem de programação R. Para realizar as implementações adicionais do AG neste trabalho, foi utilizada a versão 3.6.2. Sob licença de *software* livre, a linguagem possui pacotes que foram utilizados para realizar cálculos estatísticos, geração de gráficos e análise dos dados.

O AG faz uso do pacote Rcpp na geração de código em linguagem C++, para a implementação de algumas funções que necessitam percorrer vetores e matrizes elemento a elemento, o que eleva muito o tempo de execução quando implementado em R. Neste trabalho foi utilizada a versão 1.0.2 do Rcpp. Para geração de gráficos, foi utilizado o pacote Plotly na versão 4.9.1. Para conexão direta com o banco de dados, fez-se uso do RMySQL na versão 0.10.18, o que permitiu a utilização da base de dados do PampaPlus.

A linguagem R é rotineiramente utilizada pela EMBRAPA, com total disponibilidade de infraestrutura.

Como ambiente de programação integrado foi utilizado o RStudio ¹ na versão 1.2.1335. O RStudio inclui um console, um editor para execução direta do código, bem como ferramentas de histórico, de depuração interativa, de diagnóstico e correção de erros de forma mais rápida e, de gerenciamento de pacotes e espaço de trabalho.

Os experimentos foram realizados em um Notebook com processador Intel Core i5-7200U de 2.5 GHz, memória principal de 8 GB, disco rígido com capacidade de 1TB e sistema operacional Linux Ubuntu 16.04 LTS de 64 bits. O sistema gerenciador de banco de dados (SGBD) utilizado foi o MYSQL na versão 5.7.28.

Após a configuração do ambiente de programação, foi realizado um estudo do código original do AG e dos aspectos envolvidos na sua execução, como o acesso aos dados dos animais e o processamento das recomendações de acasalamentos. Os dados utilizados nos experimentos fazem parte do banco de dados do programa PampaPlus e correspondem às medidas objetivas e de escores visuais coletadas dos animais descritas na seção 2.2.2.

Uma instância de testes foi disponibilizada pela EMBRAPA para realização das simulações do AG, sendo a mesma utilizada nos experimentos do trabalho de Fontoura (2019). Foi utilizada uma tabela formada pelo identificador do proprietário, o identificador do animal e valores de todas as DEPs do programa para aquele animal.

Foram implementadas duas abordagens investigadas na seção 2.3.2.6 para o critério de parada do AG. A primeira foi a abordagem de Yeng, Yoke and Suhaimi (2019) e a segunda a abordagem de Bhandari, Murthy and Pal (2012). Após a implementação, foram realizados casos de teste, cada um com um conjunto diferente de animais, para validação e análise dos resultados obtidos com as duas propostas de parada antecipada, sendo descritos nas seções 4.1.1 e 4.1.2.

Porém, a utilização do AG com diferentes casos de teste evidenciou que para alguns conjuntos de animais o AG realizava recomendações de acasalamentos que excediam o limite imposto de utilização de touros. Por conta disso, o critério de utilização de touros foi revisado e teve seu código refatorado para dar seguimento aos experimentos.

O código do AG também foi modificado para possibilitar a sua utilização em outros programas de melhoramento. O AG foi concebido para ser utilizado apenas no PampaPlus, sendo codificado de maneira *Hardcode* (Codificação Rígida, em Português),

¹<https://rstudio.com/products/rstudio/>

embutindo muitos dados, principalmente as DEPs, no código fonte. Por isso, algumas rotinas e estruturas de dados do AG foram refatoradas para eliminar suas dependências com as DEPs utilizadas pelo PampaPlus. Por fim, após a implementação das melhorias, foi realizado um *benchmark* do AG com outra ferramenta de recomendação de acasalamentos, o EVA descrito na seção 2.4.

4 RESULTADOS E DISCUSSÃO

Inicialmente, os resultados da implementação de duas abordagens para o novo critério de parada do AG são mostrados na seção 4.1. Em seguida, a seção 4.2 traz os resultados obtidos com a refatoração do critério de utilização de touros. A seção 4.3 mostra os resultados da refatoração de código do AG para que o mesmo possa ser utilizado em outros programas e, por fim, a seção 4.4 apresenta um *benchmark* do AG com outra ferramenta de seleção de acasalamentos.

4.1 Critério de Parada

Esta seção apresenta os resultados da implementação de duas abordagens, sendo a primeira o critério de parada de Yeng, Yoke and Suhaimi (2019), e a segunda o critério de parada de Bhandari, Murthy and Pal (2012), ambas descritas com detalhes na seção 2.3.2.6. Ao final de cada seção é apresentada uma discussão dos resultados.

4.1.1 Critério de Parada de Yeng, Yoke and Suhaimi (2019)

Após a implementação da abordagem de parada, foram realizados dois casos de teste, com bases de dados de animais distintas para cada. Em ambos os casos de teste, foram realizados oito experimentos no total, com dez repetições para cada experimento. Nos dois primeiros experimentos, foi utilizada a abordagem original onde o critério de parada observa um número fixo de gerações/iterações, conforme mostrado na seção 2.3.3.1.

O primeiro experimento perfaz o total de 1500 iterações e o segundo, 1000 iterações, sendo o segundo utilizado para comparação com os demais experimentos da abordagem de parada proposta. O primeiro experimento foi adicionado com o objetivo de verificar se um número maior de iterações faria com que o AG encontrasse recomendações de acasalamentos com melhores valores de *fitness*.

A partir do segundo experimento com critério de parada original nas 1000 iterações, foi calculado o valor de limiar médio alcançado para servir como base para definição dos valores de limiar a serem utilizados nos experimentos da abordagem proposta. Para a realização deste cálculo, utilizou-se a função apresentada por Yeng, Yoke and Suhaimi (2019), obtendo-se um valor médio de 0,0363. Os seis experimentos restantes foram re-

alizados com a abordagem proposta, utilizando valores de limiar que partiram de 0,03 com um decremento de 0,0025 para cada até o limiar 0,02. Foi incluído também o limiar 0,01, para testar uma aproximação maior com a abordagem original. Estes experimentos utilizaram os seguintes valores de limiar: 0,03, 0,0275, 0,025, 0,0225, 0,02 e 0,01.

Os dados coletados nos experimentos realizados foram: o tempo de execução do AG; o *fitness* alcançado, que é o IQG do melhor cromossomo dado como resposta pelo algoritmo e; o número de gerações do AG, que é o número de iterações executadas pelo AG e que no caso da abordagem original estava preestabelecido como 1000. Com estes dados, o objetivo foi observar se, e para quais valores de limiar, a nova abordagem proposta atingia valores de *fitness* próximos aos alcançados pela abordagem original porém com um número de gerações menor e, conseqüentemente, com melhores índices de tempo de execução.

As médias de tempo e de *fitness* passaram por uma análise estatística para verificar a existência de diferença significativa entre os experimentos. Com o apoio do *software* R, foi realizada uma análise de variância e quando verificada a diferença, foi aplicado o Teste de Tukey para comparação múltipla dos experimentos. A seguir, são apresentados os resultados do Caso de Teste 1.

4.1.1.1 Caso de Teste 1

A Tabela 1 apresenta os parâmetros utilizados no AG em todos os experimentos para o Caso de Teste 1. Estes foram definidos de acordo com os mesmos parâmetros utilizados na abordagem original de Fontoura (2019), também utilizados em Fontoura, Camargo and Cardoso (2019).

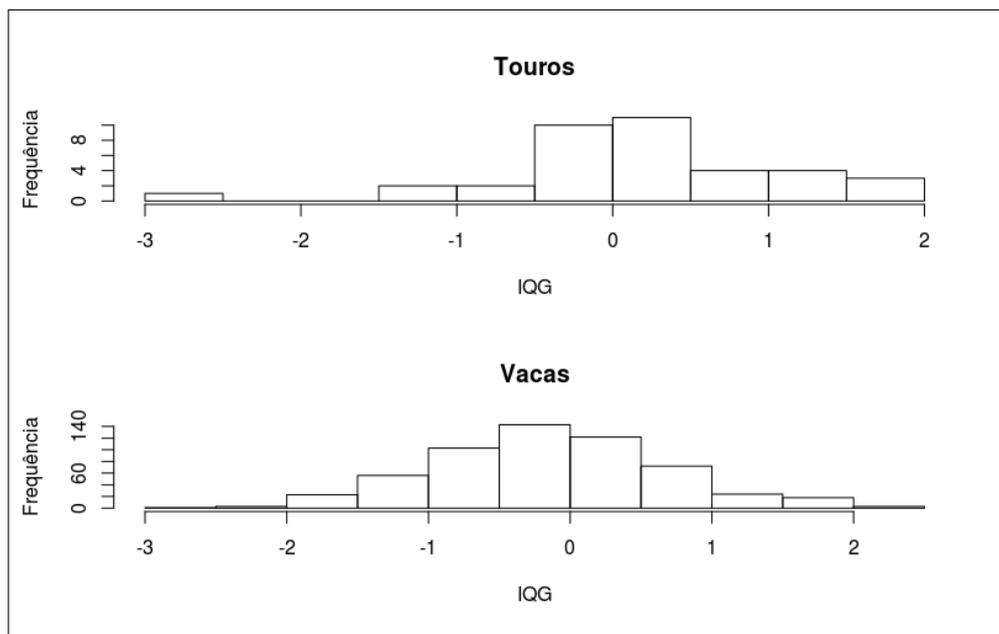
Tabela 1 – Parâmetros utilizados nos experimentos do Caso de Teste 1.

Parâmetro	Valor
Número de Touros	37
Número de Vacas	568
Tamanho da População	1136
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

Fonte: Autor(2021)

As distribuições de valores de IQG de touros e vacas selecionados pode ser visualizada na Figura 8. Para os touros, o valor médio de IQG foi de 0,18, sendo que o menor valor foi -2,61 e o maior 1,88. Para as vacas, o valor médio foi de -0,14, sendo que o menor valor foi -2,95 e o maior 2,48.

Figura 8 – Distribuição de IQG dos animais selecionados no Caso de Teste 1.



Fonte: Autor(2021)

Os valores resultantes dos experimentos podem ser observados na Tabela 2, sendo os seguintes: tempo de execução (em minutos), *fitness* alcançado e gerações, apresentando média, CV (em porcentagem) e a classificação pelo Teste de Tukey das médias. Como o valor de *fitness* inicial independe de experimento, o valor médio observado foi de 75,54.

Como pode ser observado na tabela, a média de tempo de execução dos experimentos diminui conforme incrementa-se os valores de limiar de 0,01 para 0,03. Isso deve-se ao fato de que o valor de limiar imposto, ao ser incrementado, torna-se menos restritivo quanto à convergência da população de cromossomos do AG, fazendo com que o algoritmo interrompa a sua execução tão logo encontre o valor de limiar. Porém, com um tempo de execução menor, o AG realiza menos gerações, tendendo a menores valores de *fitness*.

Também pode-se observar o maior CV de *fitness* para o experimento de limiar 0,03, com os experimentos da abordagem original apresentando valores de CV menores que este. Isso mostra que quando são aplicados valores de limiar menos restritivos, o

Tabela 2 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) no Caso de Teste 1.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (min.)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	23,23	0,42	A	241,61	0,08	A	1500,00	0,00
2. Original: 1000 it.	18,58	0,51	B	241,58	0,07	A	1000,00	0,00
3. Limiar 0,01	18,69	0,22	B	241,59	0,11	A	1000,00	0,00
4. Limiar 0,02	18,43	4,47	B	241,40	0,14	A	969,10	8,89
5. Limiar 0,0225	18,04	7,05	B	241,55	0,05	A	927,40	14,74
6. Limiar 0,025	16,43	8,59	C	241,19	0,32	A	756,30	20,00
7. Limiar 0,0275	14,51	6,40	D	239,94	0,67	B	549,60	17,66
8. Limiar 0,03	14,03	5,59	D	239,34	0,70	B	500,50	17,54

Repetições:10

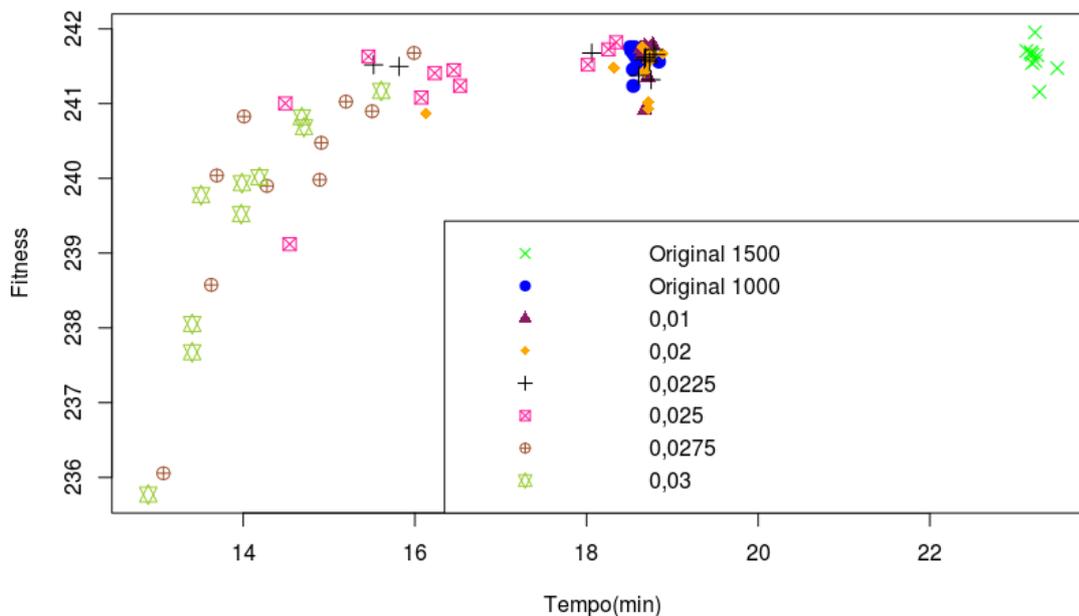
Fonte: Autor(2021)

AG executa menos iterações/gerações e alcança respostas provenientes de uma população de cromossomos com baixa convergência para um valor de *fitness* que representaria uma resposta ótima ao problema.

A Figura 9 apresenta o gráfico de dispersão das repetições dos experimentos. Como pode ser visualizado no gráfico, os experimentos da abordagem proposta com limiar de 0,01, 0,02 e a maior parte dos experimentos com limiar 0,0225, obtiveram tempos de execução e valores de *fitness* muito próximos aos encontrados na abordagem original com 1000 iterações. A partir dos valores de limiar 0,025 até 0,03, é observado que os tempos de execução e os valores de *fitness* tendem a diminuir. É possível observar também que o experimento da abordagem original com 1500 iterações, mesmo com maiores tempos de execução, não elevou o valor médio de *fitness*.

De acordo com Fontoura (2019), testes iniciais mostravam a convergência do AG perto das 800 gerações. Esta informação é confirmada na Tabela 2, quando se observa que o valor de *fitness* alcançado a partir do limiar 0,025, com uma média de 756,3 gerações, se manteve próximo ao alcançado com a abordagem original. Dos limiares que tendem a fornecer soluções com valores de *fitness* próximos a abordagem original, este possui o menor tempo de execução. A partir dos limiares 0,275 e 0,03, mesmo com tempos de execução ainda menores, os valores de *fitness* reduziram quando comparados ao valor

Figura 9 – Gráfico de dispersão das repetições dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) do Caso de Teste 1.



Fonte: Autor(2021)

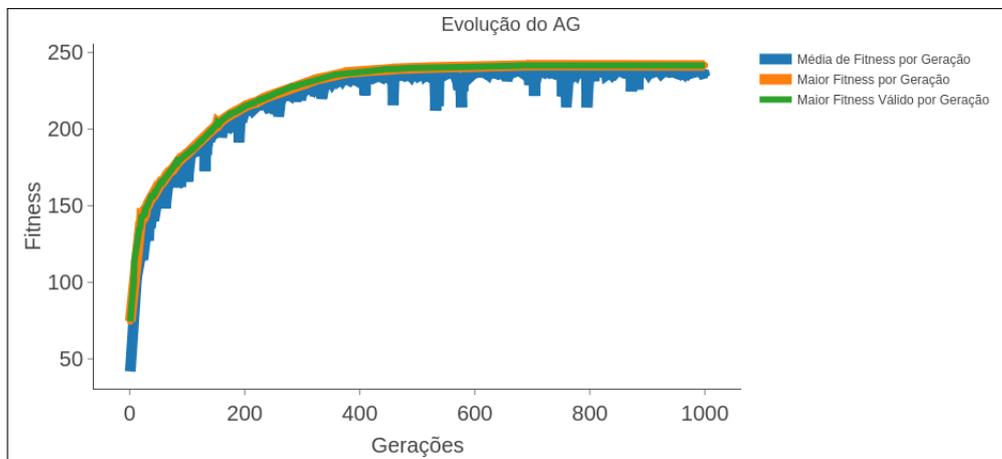
alcançado com a abordagem original.

A Figura 10 apresenta o gráfico da evolução das soluções geradas pelo AG no experimento da abordagem original com parada nas 1000 iterações. Este gráfico é semelhante nas 10 repetições. O gráfico apresenta a evolução do AG para a média de *fitness* por geração, o maior *fitness* por geração e o maior *fitness* válido por geração, sendo este último onde não há quebra das restrições impostas.

O AG executa e sua população evolui até que encerra sua execução na geração 1000, atingindo um *fitness* de $241,58 \pm 0,16$ com um tempo de execução médio de $18,58 \pm 0,09$ minutos. É possível notar que na geração 800 já acontecem iterações do AG sem necessidade, pois não há melhoria de aptidão considerável.

A Figura 11 apresenta o gráfico da evolução das soluções geradas pelo AG na abordagem proposta que tem como base a saturação do *fitness* da população e com limiar de valor 0,025. Este gráfico é semelhante nas 10 repetições. Conforme pode ser visualizado no gráfico e conferido na Tabela 2, o AG executa e sua população evolui até atingir um *fitness* de $241,19 \pm 0,77$, com média de $756,3 \pm 151,28$ iterações. Tendo o AG atin-

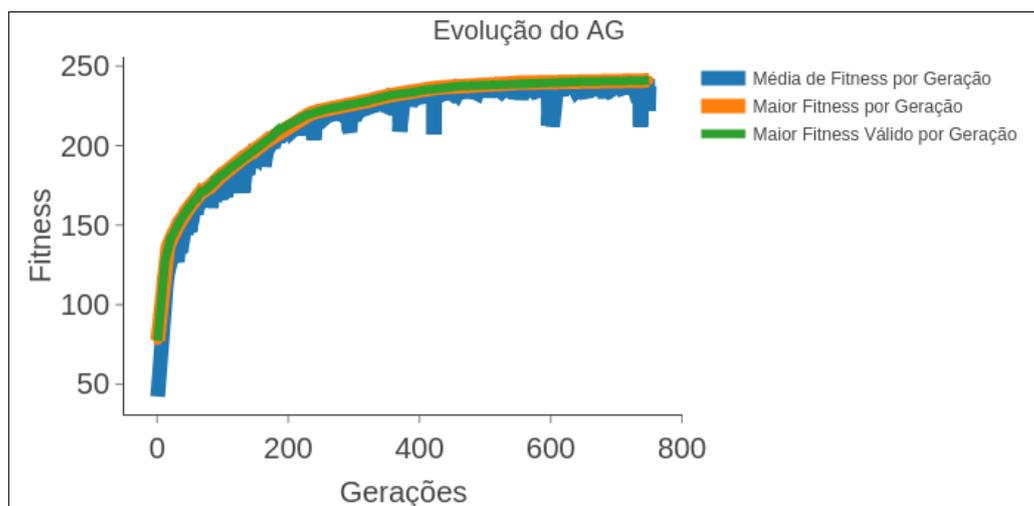
Figura 10 – Evolução das soluções geradas pelo AG com a abordagem de parada original no Caso de Teste 1.



Fonte: Autor(2021)

gido sua convergência e não havendo mais melhoria significativa de *fitness*, a abordagem proposta realiza a parada do AG conforme implementado. Com um tempo de execução médio de 16,43+1,41 minutos, apresentou um ganho de 11,53% comparado à abordagem original com 1000 iterações.

Figura 11 – Evolução das soluções geradas pelo AG com a abordagem de parada proposta com limiar de 0,025 no Caso de Teste 1.



Fonte: Autor(2021)

Como pode ser visualizado na Tabela 2, o Teste de Tukey mostrou que nos experimentos da abordagem proposta com os valores de limiar 0,025, 0,0275 e 0,03 o tempo

médio de execução é significativamente menor em comparação com os demais experimentos analisados, tendo o limiar 0,025 no grupo *C* e os limiares 0,0275 e 0,03 no grupo *D*. O experimento da abordagem original com 1500 iterações obteve significativamente o maior tempo de execução, tendo sido classificado no grupo *A*. Os demais experimentos foram classificados no grupo *B*.

Para as médias de *fitness*, o Teste de Tukey mostrou que os experimentos da abordagem proposta com os valores de limiar 0,0275 e 0,03, no grupo *B*, possuem média de *fitness* significativamente menor que os demais experimentos analisados, no grupo *A*. A partir desta análise também foi possível comparar os experimentos da nova abordagem com o experimento da abordagem original, com 1000 iterações. O experimento com valor de limiar de 0,025 foi o único da abordagem proposta que manteve as médias de *fitness* e diminuiu o tempo médio de execução alcançado pela abordagem original, tal como observado graficamente na Figura 9 e confirmado através da análise estatística realizada. Portanto, a abordagem proposta com o valor de limiar 0,025 apresentou o melhor resultado em comparação com os demais experimentos no Caso de Teste 1.

4.1.1.2 Caso de Teste 2

Em seguida, a Tabela 3 apresenta os parâmetros utilizados no AG em todos os experimentos para o Caso de Teste 2. Os parâmetros do AG como Taxa de Consanguinidade, Taxa de Mutação e Função de Avaliação foram os mesmos que os utilizados no caso de teste anterior. O que difere o Caso de Teste 2 são os animais selecionados, que são distintos dos animais utilizados no Caso de Teste 1. Trata-se de uma base de dados menor e com diferente valor de IQG médio para os animais.

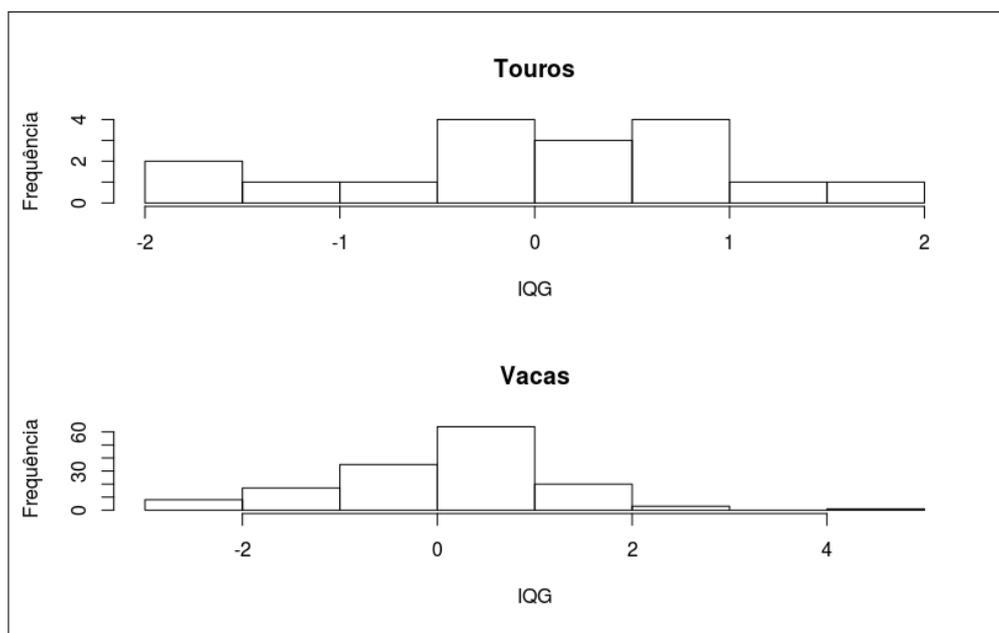
Tabela 3 – Parâmetros utilizados nos experimentos do Caso de Teste 2.

Parâmetro	Valor
Número de Touros	17
Número de Vacas	148
Tamanho da População	296
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

Fonte: Autor(2021)

A Figura 12 apresenta as distribuições de valores de IQG de touros e vacas selecionados. O valor médio de IQG para os touros foi de -0,02, sendo que o menor valor foi -1,79 e o maior 1,82. Para as vacas, o valor médio foi de 0,08, sendo que o menor valor foi -2,75 e o maior 4,31.

Figura 12 – Distribuição de IQG dos animais selecionados no Caso de Teste 2.



Fonte: Autor(2021)

Os valores resultantes dos experimentos do Caso de Teste 2 podem ser observadas na Tabela 4. O valor médio de *fitness* inicial observado foi de 9,86. Semelhante ao ocorrido no Caso de Teste 1, a média de tempo de execução dos experimentos diminui conforme incrementa-se os valores de limiar de 0,01 para 0,03, com o AG realizando menos gerações e com isso alcançando menores valores de *fitness*. Observa-se também, que o CV de *fitness* no experimento com a abordagem proposta de limiar 0,03 é maior que os CV apresentados nos experimentos da abordagem original, da mesma forma como no Caso de Teste 1.

A Figura 13 apresenta o gráfico de dispersão das repetições dos experimentos. Este cenário apresenta algumas diferenças com relação ao encontrado no Caso de Teste 1. O experimento da abordagem proposta com limiar de 0,01 foi o único que obteve tempos de execução e valores de *fitness* muito próximos aos encontrados na abordagem original com 1000 iterações. A partir dos valores de limiar 0,02 até 0,03, é observado que os tempos de execução e os valores de *fitness* tendem a diminuir. De modo semelhante ao

Tabela 4 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) no Caso de Teste 2.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (s)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	77,65	1,16	A	57,34	0,07	A	1500,00	0,00
2. Original: 1000 it.	57,36	0,14	B	57,27	0,23	A	1000,00	0,00
3. Limiar 0,01	54,74	12,12	B	57,18	0,50	A	929,00	17,63
4. Limiar 0,02	37,21	22,41	C	56,83	1,13	AB	429,90	41,48
5. Limiar 0,0225	28,94	12,78	D	55,96	2,21	BC	290,50	30,96
6. Limiar 0,025	28,57	10,32	D	55,88	1,76	BC	280,70	25,92
7. Limiar 0,0275	27,20	9,50	D	55,61	2,18	C	245,60	25,60
8. Limiar 0,03	26,59	8,21	D	55,34	2,22	C	231,90	23,14

Repetições:10

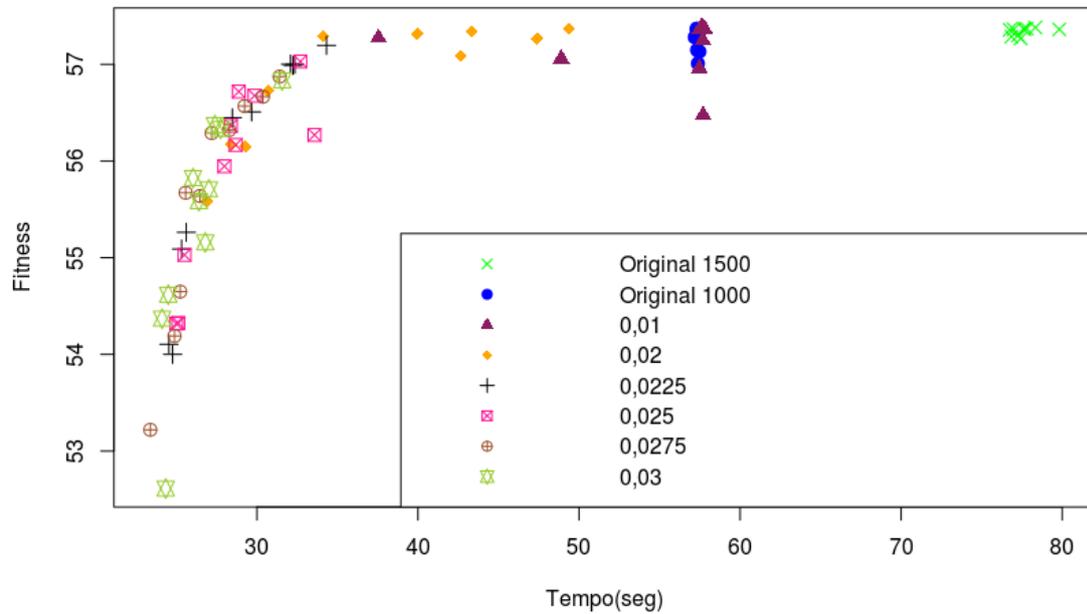
Fonte: Autor(2021)

observado no Caso de Teste 1, o experimento da abordagem original com 1500 iterações, mesmo com maiores tempos de execução, não elevou consideravelmente o valor médio de *fitness*.

Com relação às médias de gerações/iterações executadas pelo AG na abordagem proposta, mais diferenças com o caso de teste anterior podem ser observadas. Como mostrado na Tabela 4, o limiar 0,01 atingiu um valor de 929 gerações, muito próximo à abordagem original com 1000 gerações. Já o limiar 0,02, com 429,9 gerações, apresentou uma pequena perda no valor médio de *fitness* alcançado quando comparado ao limiar 0,01 e à abordagem original com 1000 e 1500 iterações. Os demais limiares apresentaram médias de gerações ainda menores, o que afetou mais os *fitness* alcançados.

A Figura 14 apresenta o gráfico da evolução das soluções geradas pelo AG no experimento da abordagem original com parada nas 1000 iterações. Este gráfico é semelhante nas 10 repetições. O gráfico apresenta a evolução do AG para a média de *fitness* por geração, o maior *fitness* por geração e o maior *fitness* válido por geração. O AG executa e sua população evolui até que encerra sua execução na geração 1000, atingindo um *fitness* de 57,27+-0,13 com um tempo de execução médio de 57,36+-0,08 segundos, conforme a Tabela 4. É possível notar que o gráfico atinge o platô entre a geração 400 e a geração 600, e a partir deste ponto já acontecem iterações do AG sem necessidade, pois não há

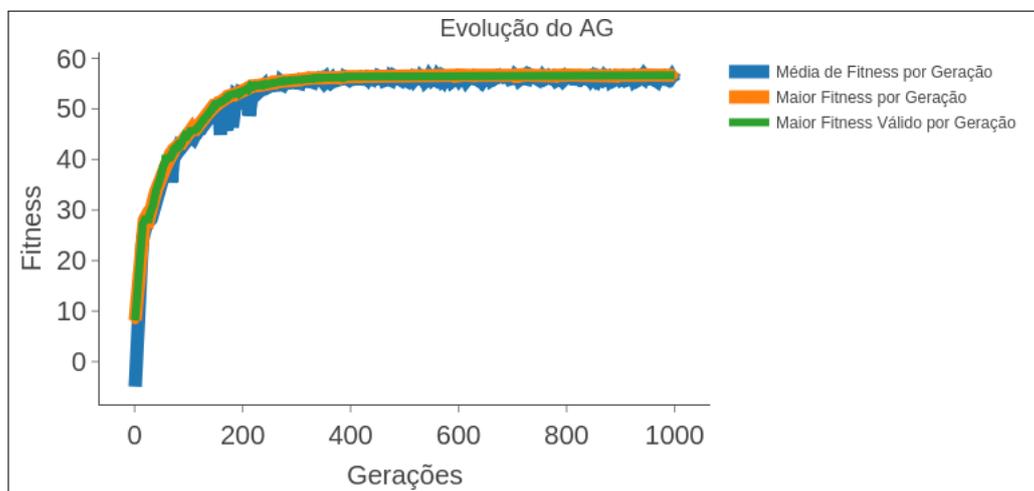
Figura 13 – Gráfico de dispersão das repetições dos experimentos para o critério de parada de Yeng, Yoke and Suhaimi (2019) do Caso de Teste 2.



Fonte: Autor(2021)

melhoria de aptidão considerável.

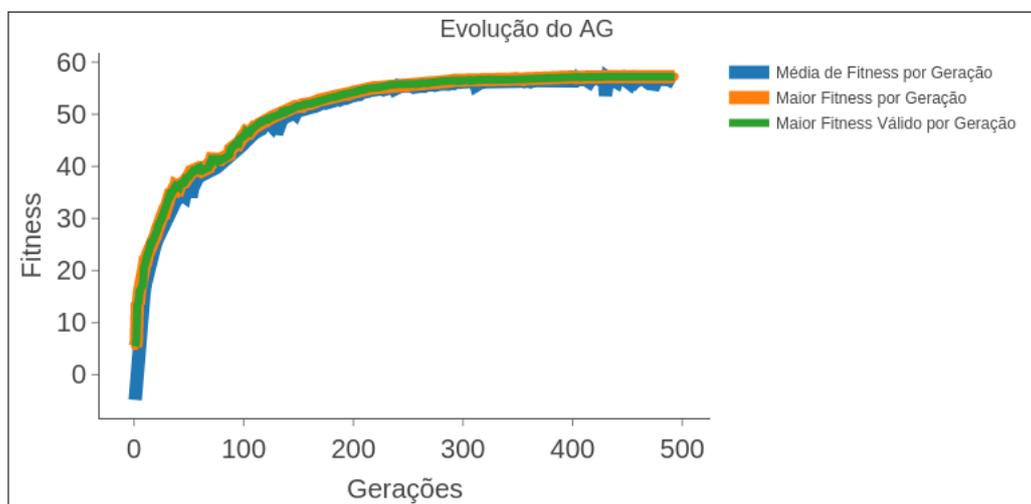
Figura 14 – Evolução das soluções geradas pelo AG com a abordagem de parada original no Caso de Teste 2.



Fonte: Autor(2021)

A Figura 15 apresenta o gráfico da evolução das soluções geradas pelo AG na abordagem proposta que tem como base a saturação do *fitness* da população e com limiar de valor 0,02. Este gráfico é semelhante nas 10 repetições. Conforme pode ser visualizado no gráfico e conferido na Tabela 4, o AG executa e sua população evolui até atingir um *fitness* de 56,83+0,64, com média de 492,9+204,44 iterações. Tendo o AG atingido sua convergência e não havendo mais melhoria significativa de *fitness*, a nova abordagem realiza a parada do AG conforme implementado. Com um tempo de execução médio de 37,2+8,34 segundos, apresentou um ganho de 35,13% comparado à abordagem original com 1000 iterações.

Figura 15 – Evolução das soluções geradas pelo AG com a abordagem de parada proposta com limiar de 0,02 no Caso de Teste 2.



Fonte: Autor(2021)

Com relação à classificação das médias de tempo dada pelo Teste de Tukey realizado, o experimento da abordagem original com 1500 iterações obteve significativamente o maior tempo de execução, tendo sido classificado no grupo A. O experimento da abordagem proposta com limiar de 0,01 foi classificado no grupo B, não tendo redução significativa em comparação com o experimento da abordagem original de 1000 iterações, no mesmo grupo. A partir do experimento com limiar 0,02, no grupo C, houve uma redução significativa do tempo médio de execução. Os demais experimentos com os limiares 0,0225, 0,025, 0,0275 e 0,03 ficaram classificados no grupo D, com uma maior redução.

O Teste de Tukey para as médias de *fitness* mostrou que os experimentos da abordagem proposta com os valores de limiar de 0,0225 até 0,03, nos grupos BC e C, possuem média de *fitness* significativamente menor que os demais experimentos analisados. No

grupo A ficaram os experimentos da abordagem original com 1500 e 1000 iterações, e o experimento da nova abordagem com limiar 0,01.

No grupo AB, ficou o experimento com valor de limiar 0,02, que foi o único da abordagem proposta que manteve as médias de *fitness* enquanto diminuiu o tempo médio de execução alcançado pela abordagem original, tal como observado graficamente na Figura 13 e confirmado através da análise estatística realizada. Portanto, a abordagem proposta com o valor de limiar 0,02 apresentou o melhor resultado em comparação com os demais experimentos no Caso de Teste 2.

4.1.1.3 Discussão

Após a apresentação dos resultados para o critério de parada de Yeng, Yoke and Suhaimi (2019), pode-se comparar os valores de limiar que apresentaram os melhores resultados em cada caso de teste. No Caso de Teste 1, o valor de limiar 0,025 foi o único que reduziu o tempo de execução e manteve as médias de *fitness* alcançadas com a abordagem original. No Caso de Teste 2, o único que reduziu o tempo de execução e manteve as médias de *fitness* foi o limiar 0,02.

Portanto, não foi possível definir um único valor de limiar para efetuar a parada do AG com diferentes seleções de animais. O valor de limiar a ser escolhido mostrou-se dependente das características intrínsecas a cada seleção de animais, onde cada seleção foi um caso de teste diferente. Já no trabalho de Yeng, Yoke and Suhaimi (2019), o algoritmo genético foi desenvolvido para resolver um problema específico, de um único caso de teste, com os autores utilizando um único valor de limiar.

Diante da impossibilidade da adoção de um único valor de limiar para esta abordagem de parada no contexto do problema investigado neste trabalho, uma segunda abordagem de parada para o AG foi implementada e testada. Na seção seguinte, são apresentados os resultados dos experimentos realizados com a segunda abordagem de parada e uma discussão.

4.1.2 Critério de Parada de Bhandari, Murthy and Pal (2012)

Após a implementação da abordagem de parada, foram realizados experimentos com cinco casos de teste. Para os Casos de Teste 1 e 2, foram repetidas as condições dos casos 1 e 2 da seção anterior. Para o Caso de Teste 3, foi utilizada uma nova base

de dados, com animais distintos daqueles já utilizados anteriormente. Já para o Caso de Teste 4, foram utilizados 25% dos animais da base de dados utilizada no Caso de Teste 1. Por fim, para o Caso de Teste 5, foram utilizados 50% dos animais da base de dados utilizada no Caso de Teste 1.

Em todos os casos de teste foram realizados onze experimentos no total. Para os dois primeiros experimentos, foi utilizada a abordagem original desenvolvida por Fontoura (2019), sendo o primeiro experimento de 1500 iterações e o segundo, 1000 iterações. Foram aproveitados os resultados já apresentados na seção anterior e, para fins de comparação com a abordagem proposta, o experimento original com 1000 iterações foi usado como referência.

Os demais experimentos foram realizados com a abordagem proposta com três diferentes valores de limiar: $10E-1$, $10E-2$ e $10E-3$, e três valores de intervalo de gerações: 100, 200 e 300. No trabalho de Bhandari, Murthy and Pal (2012), os autores utilizaram valores de limiar como $10E-5$ e $10E-4$, e comentaram que a utilização destes valores dependem das características da função de *fitness*. Com base nisso, foram realizados testes preliminares no AG a partir destes valores e, foi somente com $10E-3$ que o AG efetuou a parada antecipada. Quanto ao intervalo de gerações, os autores utilizaram o valor de 200 na maioria de seus experimentos, o que guiou a escolha dos intervalos utilizados. Assim, cada um dos três valores de limiar foi testado com cada um dos três intervalos, totalizando os nove experimentos restantes.

Para cada um dos experimentos foram realizadas dez repetições, sendo que os dados coletados foram os mesmos coletados com os casos de teste da abordagem de parada testada anteriormente. A seguir, são apresentados os resultados do Caso de Teste 1.

4.1.2.1 Caso de Teste 1

Os parâmetros utilizados neste caso de teste podem ser visualizados na Tabela 1, e a distribuição de valores de IQG dos animais selecionados na Figura 8. A Tabela 5 apresenta os dados resultantes dos experimentos, sendo os seguintes: tempo de execução (em minutos), *fitness* alcançado e gerações, apresentando média, CV (em porcentagem) e classificação pelo Teste de Tukey. O valor médio de *fitness* inicial foi de 75,54.

Como pode ser visualizado na tabela de resultados, o experimento da abordagem de parada original com 1500 iterações, de identificador 1, obteve o maior tempo de execução, apresentando média de $23,23 \pm 0,09$ minutos. Este experimento apresentou o segundo melhor valor médio de *fitness* de todos os experimentos, com $241,61 \pm 0,20$. O

Tabela 5 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 1.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (min.)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	23,23	0,42	A	241,61	0,08	A	1500,00	0,00
2. Original: 1000 it.	18,58	0,51	B	241,58	0,07	A	1000,00	0,00
3. Limiar 10E-3, 300	18,27	1,98	B	241,66	0,06	A	975,90	3,78
4. Limiar 10E-2, 300	18,06	2,92	B	241,57	0,05	A	951,50	5,73
5. Limiar 10E-1, 300	16,86	2,61	C	241,55	0,07	A	818,00	5,85
6. Limiar 10E-3, 200	17,90	2,91	B	241,53	0,11	A	932,60	6,06
7. Limiar 10E-2, 200	16,88	4,23	C	241,52	0,06	A	820,40	9,58
8. Limiar 10E-1, 200	16,12	2,03	DE	241,53	0,12	A	733,10	5,02
9. Limiar 10E-3, 100	16,67	4,48	CD	241,54	0,12	A	790,60	10,11
10. Limiar 10E-2, 100	15,92	3,10	E	241,49	0,09	A	713,30	7,60
11. Limiar 10E-1, 100	14,74	2,99	F	241,04	0,23	B	583,60	7,47

Repetições:10

Fonte: Autor(2021)

experimento com a abordagem original com 1000 iterações, de identificador 2, obteve o segundo maior tempo de execução médio com 18,58+-0,09 minutos, alcançando o terceiro melhor valor de *fitness* com 241,58 +-0,15.

Para os experimentos da abordagem proposta no quesito tempo médio de execução, o experimento 3 apresentou o terceiro maior valor, 18,27+-0,36 minutos. O experimento 11 apresentou o menor valor, 14,74+-0,44 minutos. Já para os valores de *fitness* médio, o experimento 3 obteve o melhor valor, 241,66+-0,15, com média de 975,9+-36,89 gerações. O experimento 11, por sua vez, apresentou o pior valor, 241,04+-0,56, com média de 583,6+-43,59 gerações.

Os gráficos de dispersão dos resultados dos experimentos com a abordagem de parada de Bhandari, Murthy and Pal (2012) seguem uma distribuição similar àqueles apresentados nos experimentos com a abordagem de Yeng, Yoke and Suhaimi (2019) para o Caso de Teste 1. Os gráficos da evolução das gerações também apresentam o mesmo comportamento.

O Teste de Tukey para a média de tempo mostrou que apenas os experimentos 5,

7, 8, 9, 10 e 11 da abordagem de parada proposta reduziram significativamente o tempo de execução médio quando comparados aos experimentos da abordagem de parada original. Já para as médias de *fitness*, a maioria dos experimentos da abordagem proposta manteve significativamente os valores alcançados com a abordagem original, com exceção do experimento 11.

Portanto, a maior parte dos experimentos da abordagem proposta reduziu o tempo médio de execução e manteve as médias de *fitness* alcançadas pela abordagem original, o que não se concretizou para o experimento 11. Na análise dos melhores pelo tempo de execução, os experimentos 8 e 10 ficaram no grupo *E*, apresentando ganhos de 13,22% e 14,33% respectivamente. Em seguida, são apresentados os resultados dos experimentos do Caso de Teste 2.

4.1.2.2 Caso de Teste 2

A Tabela 6 apresenta os resultados dos experimentos do Caso de Teste 2. Os parâmetros do AG utilizados neste caso de teste podem ser visualizados na Tabela 3, e a distribuição de valores de IQG de touros e vacas na Figura 8. O valor médio de *fitness* inicial foi de 9,86.

De acordo com a Tabela 6, o experimento da abordagem de parada original com 1500 iterações, de identificador 1, obteve o maior tempo de execução, apresentando média de 77,65+-0,01 segundos. Este experimento apresentou também o melhor valor médio de *fitness* de todos os experimentos, com 57,34+-0,04. O experimento com a abordagem original com 1000 iterações, de identificador 2, obteve o segundo maior tempo de execução médio com 57,36+-0,07 segundos, alcançando o segundo melhor valor de *fitness* com 57,27 +-0,13.

Para os experimentos da abordagem proposta no quesito tempo médio de execução, o experimento 3 apresentou o terceiro maior valor, 46,68+-1,02 segundos. O experimento 11 apresentou o menor valor, 30,33+-1,22 segundos. Já para os valores de *fitness* médio, o experimento 5 obteve o terceiro melhor valor, 57,23+-0,12, com média de 550,7+-16,57 gerações. O experimento 11 apresentou o pior valor, 56,89+-0,19, com média de 326,5+-29,76 gerações.

Os gráficos de dispersão dos resultados dos experimentos com a abordagem de parada de Bhandari, Murthy and Pal (2012) seguem uma distribuição similar àqueles apresentados nos experimentos com a abordagem de Yeng, Yoke and Suhaimi (2019) para o Caso de Teste 2. Os gráficos da evolução das gerações também apresentam o mesmo

Tabela 6 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 2.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (s)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	77,65	1,16	A	57,34	0,07	A	1500,00	0,00
2. Original: 1000 it.	57,36	0,14	B	57,27	0,23	A	1000,00	0,00
3. Limiar 10E-3, 300	46,68	2,19	C	57,21	0,28	A	724,10	3,18
4. Limiar 10E-2, 300	43,98	2,60	D	57,16	0,40	AB	658,30	4,04
5. Limiar 10E-1, 300	39,58	2,00	E	57,23	0,22	A	550,70	3,01
6. Limiar 10E-3, 200	43,14	3,94	D	57,18	0,49	A	638,20	6,40
7. Limiar 10E-2, 200	39,63	3,27	E	57,16	0,55	AB	554,30	5,97
8. Limiar 10E-1, 200	35,14	1,06	F	57,19	0,09	A	444,20	2,00
9. Limiar 10E-3, 100	38,10	4,26	E	57,12	0,41	AB	518,00	7,51
10. Limiar 10E-2, 100	34,52	1,96	F	57,10	0,27	AB	429,40	3,83
11. Limiar 10E-1, 100	30,33	4,03	G	56,89	0,34	B	326,50	9,12

Repetições:10

Fonte: Autor(2021)

comportamento.

O Teste de Tukey para a média de tempo mostrou que todos os experimentos da abordagem de parada proposta reduziram significativamente o tempo de execução médio quando comparados aos experimentos da abordagem de parada original. Já para as médias de *fitness*, a maioria dos experimentos da abordagem proposta manteve significativamente os valores alcançados com a abordagem original, com exceção do experimento 11.

Portanto, com exceção de apenas um experimento, os demais experimentos da abordagem proposta reduziram o tempo médio de execução e mantiveram as médias de *fitness* alcançadas pela abordagem original. Entretanto, na análise de tempo de execução, os experimentos 8 e 10 ficaram no grupo *F* e melhor classificados que os demais, apresentando ganhos de 38,74% e 39,82% respectivamente.

4.1.2.3 Caso de Teste 3

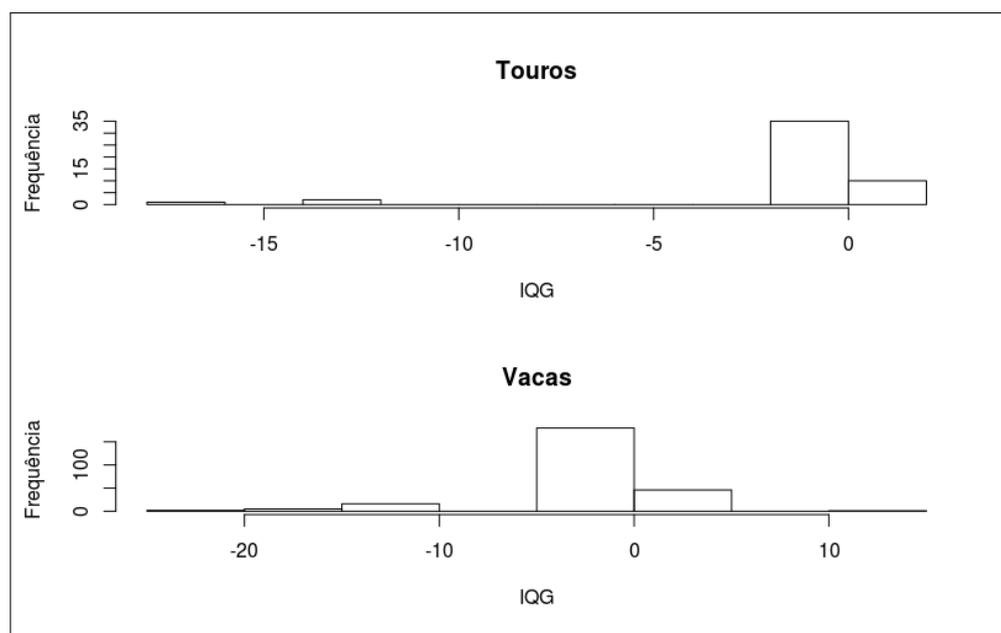
Os parâmetros do AG utilizados no Caso de Teste 3 podem ser visualizados na Tabela 7 e, as distribuições de valores de IQG dos animais selecionados na Figura 16. Para os touros, o valor médio de IQG foi de -1,19, sendo que o menor valor foi -16,79 e o maior 0,61. Para as vacas, o valor médio foi de -1,53, sendo que o menor valor foi -23,72 e o maior 14,37.

Tabela 7 – Parâmetros utilizados nos experimentos do Caso de Teste 3.

Parâmetro	Valor
Número de Touros	48
Número de Vacas	258
Tamanho da População	306
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

Fonte: Autor(2021)

Figura 16 – Distribuição de IQG dos animais selecionados no Caso de Teste 3.



Fonte: Autor(2021)

A Tabela 8 apresenta os resultados dos experimentos. O valor médio de *fitness* inicial foi de -132,47. O valor de *fitness* negativo é resultado do maior número de indivíduos com valores de IQGs negativos no conjunto de animais selecionados.

Tabela 8 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 3.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (min.)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	4,24	0,59	A	-18,65	1,62	A	1500,00	0,00
2. Original: 1000 it.	3,15	0,48	CDE	-18,86	1,68	AB	1000,00	0,00
3. Limiar 10E-3, 300	3,62	5,68	B	-18,74	1,88	A	1205,30	7,70
4. Limiar 10E-2, 300	3,39	5,14	BC	-18,78	1,64	A	1104,10	7,17
5. Limiar 10E-1, 300	3,09	2,92	DEF	-18,76	1,25	A	969,50	4,24
6. Limiar 10E-3, 200	3,59	10,99	B	-18,72	1,13	A	1195,50	15,02
7. Limiar 10E-2, 200	3,29	5,10	CD	-18,78	2,57	A	1062,20	7,35
8. Limiar 10E-1, 200	2,82	4,69	FG	-19,04	1,54	AB	847,70	7,09
9. Limiar 10E-3, 100	2,87	8,38	EFG	-19,05	1,45	AB	870,10	12,64
10. Limiar 10E-2, 100	2,65	8,07	GH	-19,44	3,53	B	767,30	12,69
11. Limiar 10E-1, 100	2,37	4,26	H	-20,40	4,58	C	641,20	7,20

Repetições:10

Fonte: Autor(2021)

O experimento da abordagem de parada original com 1500 iterações, de identificador 1, obteve o maior tempo de execução, apresentando média de 4,24+0,02 minutos. Este experimento apresentou o melhor valor médio de *fitness* de todos os experimentos, com -18,65+0,30, porém muito próximo ao valor obtido pelo experimento da abordagem original de 1000 iterações e aos demais experimentos que realizaram mais iterações. O experimento com a abordagem original com 1000 iterações, de identificador 2, obteve um tempo de execução médio de 3,15+0,01 minutos, com um *fitness* de -18,86 +0,32.

Na abordagem de parada proposta no quesito tempo médio de execução, o experimento 3 apresentou o segundo maior valor, 3,62+0,20 minutos. Já o experimento 11 apresentou o menor valor, 2,37+0,10 minutos. Já para os valores de *fitness* médio, o experimento 6 obteve o segundo melhor valor, -18,72+0,21, com média de 1195,5+179,53 gerações. O pior valor alcançado ficou com o experimento 11, -20,40+0,93, com média

de 641,2+-46,14 gerações.

O Teste de Tukey para a média de tempo mostrou que apenas os experimentos 8, 10 e 11 da abordagem de parada proposta reduziram significativamente o tempo de execução médio quando comparados ao experimento 2 de parada original com 1000 iterações. Já para as médias de *fitness*, a maioria dos experimentos da abordagem proposta mantiveram significativamente os valores alcançados com a abordagem original, com exceção do experimento 11.

Portanto, os experimentos 8 e 10 conseguiram reduzir o tempo médio de execução, em 10,43% e 15,97% respectivamente, e manter as médias de *fitness* alcançadas na abordagem original.

4.1.2.4 Caso de Teste 4

A seguir, são apresentados os resultados dos experimentos para o Caso de Teste 4. Os parâmetros do AG utilizados no Caso de Teste 4 podem ser mostrados na Tabela 9. A Figura 17 apresenta as distribuições de IQG dos animais selecionados. O valor médio de IQG para os touros foi de -0,44, sendo que o menor valor foi -2,61 e o maior 0,41. Para as vacas, o valor médio foi de -0,31, sendo que o menor valor foi -2,44 e o maior 1,34.

Tabela 9 – Parâmetros utilizados nos experimentos do Caso de Teste 4.

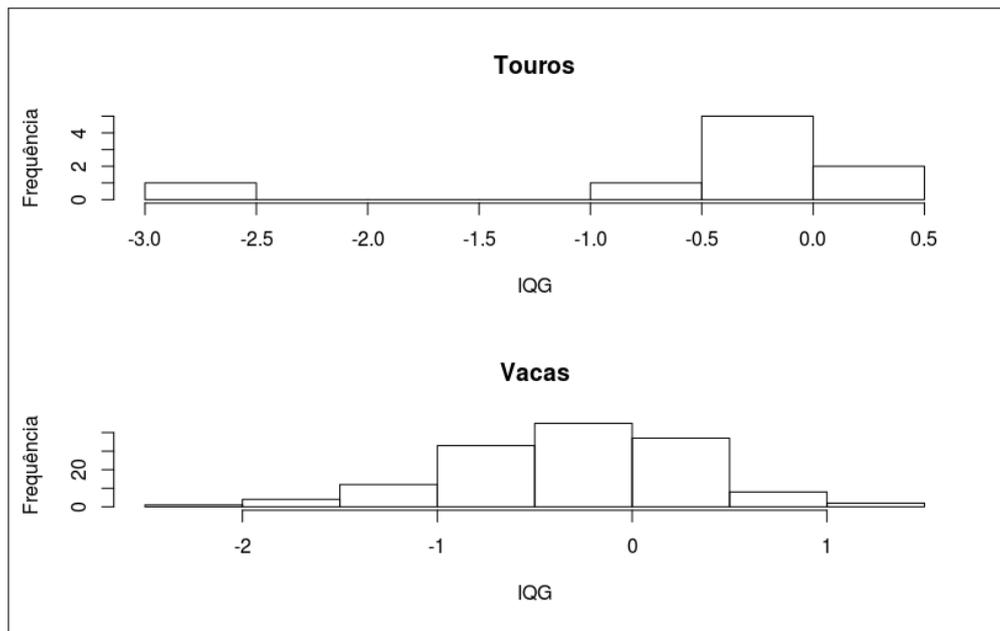
Parâmetro	Valor
Número de Touros	9
Número de Vacas	142
Tamanho da População	284
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

Observação: 25% dos animais utilizados no Caso de Teste 1

Fonte: Autor(2021)

A Tabela 10 apresenta os resultados dos experimentos. O valor médio de *fitness* inicial foi de -63,40. O experimento da abordagem de parada original com 1500 iterações, de identificador 1, obteve o maior tempo de execução apresentando média de 68,24+-0,003 segundos e atingiu o valor de *fitness* de -34,89+-0,13. O experimento com

Figura 17 – Distribuição de IQG dos animais selecionados no Caso de Teste 4.



Fonte: Autor(2021)

a abordagem original com 1000 iterações, de identificador 2, obteve o segundo maior tempo de execução médio, $50,66 \pm 0,09$ segundos, e atingiu o melhor *fitness* de todos os experimentos, com valor de $-34,82 \pm 0,09$.

Analisando os experimentos da abordagem proposta no quesito tempo médio de execução, o experimento 3 apresentou o terceiro maior valor, $36,96 \pm 4,08$ segundos. O experimento 11 apresentou o menor valor, $23,47 \pm 0,40$ segundos. Já para os valores de *fitness* médio, o experimento 4 obteve o segundo melhor valor, $-34,85 \pm 0,07$, com média de $466,4 \pm 22,63$ gerações. O pior valor alcançado ficou com o experimento 11, $-34,96 \pm 0,18$, com média de $215,4 \pm 10,87$ gerações.

Para as média de tempo, o Teste de Tukey mostrou que todos os experimentos da abordagem de parada proposta reduziram significativamente o tempo de execução médio quando comparados com os experimentos 1 e 2 da abordagem de parada original. Já para as médias de *fitness*, não houve diferença significativa nos valores encontrados entre todos os experimentos realizados.

Com isso, todos os experimentos da abordagem proposta reduziram o tempo médio de execução e mantiveram as médias de *fitness* alcançadas na abordagem original. Entretanto, na análise de tempo de execução, os experimentos 10 e 11 ficaram no grupo *H* e melhor classificados que os demais, apresentando ganhos de 50,87% e 53,66% res-

Tabela 10 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 4.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (s)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	68,24	0,30	A	-34,89	0,37	A	1500,00	0,00
2. Original: 1000 it.	50,66	0,19	B	-34,82	0,27	A	1000,00	0,00
3. Limiar 10E-3, 300	36,96	11,05	C	-34,86	0,25	A	548,20	11,24
4. Limiar 10E-2, 300	32,39	2,55	D	-34,85	0,22	A	466,40	4,85
5. Limiar 10E-1, 300	30,30	2,53	DE	-34,89	0,33	A	409,00	5,24
6. Limiar 10E-3, 200	32,51	8,67	D	-34,87	0,37	A	471,50	16,83
7. Limiar 10E-2, 200	28,66	3,56	EF	-34,87	0,21	A	361,20	7,79
8. Limiar 10E-1, 200	26,94	2,91	FG	-34,86	0,67	A	315,60	7,12
9. Limiar 10E-3, 100	27,91	6,12	EF	-34,95	0,30	A	342,70	13,92
10. Limiar 10E-2, 100	24,89	3,39	GH	-34,91	0,39	A	256,80	9,63
11. Limiar 10E-1, 100	23,47	1,72	H	-34,96	0,52	A	215,40	5,05

Repetições:10

Fonte: Autor(2021)

pectivamente.

4.1.2.5 Caso de Teste 5

A seguir, são apresentados os resultados dos experimentos para o Caso de Teste 5. Os parâmetros do AG utilizados podem ser visualizados na Tabela 11. A distribuição de IQG dos animais pode ser visualizada na Figura 18. Para os touros, o valor médio de IQG foi de -0,13, sendo que o menor valor foi -2,61 e o maior 1,86. O valor médio para as vacas foi de -0,31, sendo que o menor valor foi -2,95 e o maior 1,94.

A Tabela 12 apresenta os resultados deste caso de teste. O valor médio de *fitness* inicial foi de -31,83. De acordo com a tabela, o experimento da abordagem de parada original com 1500 iterações, de identificador 1, obteve o maior tempo de execução apresentando média de 4,26+-0,02 segundos e atingiu o valor de *fitness* de 48,48+-0,17. O experimento com a abordagem original com 1000 iterações, de identificador 2, obteve o segundo maior tempo de execução médio, 3,27+-0,01 segundos, e atingiu o segundo

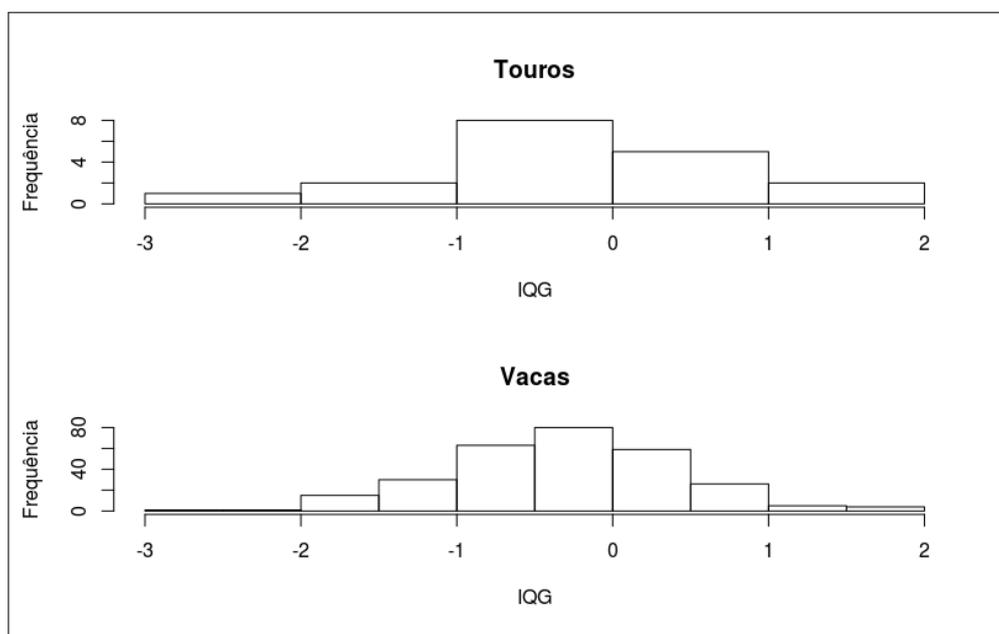
Tabela 11 – Parâmetros utilizados nos experimentos do Caso de Teste 5.

Parâmetro	Valor
Número de Touros	18
Número de Vacas	284
Tamanho da População	568
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

Observação: 50% dos animais utilizados no Caso de Teste 1

Fonte: Autor(2021)

Figura 18 – Distribuição de IQG dos animais selecionados no Caso de Teste 5.



Fonte: Autor(2021)

melhor *fitness* de todos os experimentos, com valor de 48,54 \pm 0,08.

Analisando os experimentos da abordagem proposta no quesito tempo médio de execução, o experimento 3 apresentou o terceiro maior valor, 2,86 \pm 0,22 segundos. O experimento 11 apresentou o menor valor, 1,93 \pm 0,7 segundos. Já para os valores de *fitness* médio, o experimento 9 obteve o melhor valor, 48,55 \pm 0,05, com média de 553,8 \pm 53,27 gerações. O pior valor alcançado ficou com o experimento 11, 47,84 \pm 0,43, com

Tabela 12 – Média, coeficiente de variação e classificação pelo Teste de Tukey dos dados resultantes dos experimentos para o critério de parada de Bhandari, Murthy and Pal (2012) no Caso de Teste 5.

Experimento	Tempo de Execução			Fitness Alcançado			Gerações	
	Média (min.)	CV	Grupo	Média	CV	Grupo	Média	CV
1. Original: 1500 it.	4,26	0,60	A	48,48	0,36	A	1500,00	0,00
2. Original: 1000 it.	3,27	0,22	B	48,54	0,17	A	1000,00	0,00
3. Limiar 10E-3, 300	2,86	7,97	C	48,44	0,45	A	803,70	13,84
4. Limiar 10E-2, 300	2,66	5,25	D	48,52	0,11	A	708,40	9,59
5. Limiar 10E-1, 300	2,50	4,66	DEF	48,48	0,37	A	629,30	8,58
6. Limiar 10E-3, 200	2,57	6,14	DE	48,54	0,35	A	669,80	11,39
7. Limiar 10E-2, 200	2,42	5,96	EFG	48,40	0,84	A	593,40	11,69
8. Limiar 10E-1, 200	2,27	5,47	GH	48,51	0,13	A	521,30	11,71
9. Limiar 10E-3, 100	2,34	4,92	FGH	48,55	0,11	A	553,80	9,62
10. Limiar 10E-2, 100	2,20	6,34	H	48,40	0,65	A	487,50	13,53
11. Limiar 10E-1, 100	1,93	4,01	I	47,85	0,89	B	358,90	9,99

Repetições:10

Fonte: Autor(2021)

média de 358,9+-35,86 gerações.

Para as médias de tempo, o Teste de Tukey mostrou que todos os experimentos da abordagem de parada proposta reduziram significativamente o tempo de execução médio quando comparados com os experimentos 1 e 2 da abordagem de parada original. Já para as médias de *fitness*, com exceção do experimento 11, os demais mantiveram significativamente as médias encontradas nos experimentos da abordagem original.

Com isso, a maior parte dos experimentos da abordagem proposta reduziu o tempo médio de execução e manteve as médias de *fitness* alcançadas na abordagem original, com exceção de apenas um experimento. Entretanto, na análise de tempo de execução, os experimentos 8, 9 e 10 ficaram no grupo *H* e melhor classificados que os demais, apresentando ganhos de 30,56%, 28,24% e 32,67% respectivamente.

4.1.2.6 Discussão

Após a coleta de resultados dos cinco casos de teste foi realizada uma análise de resultados, sendo que uma discussão é apresentada a seguir. De acordo com as médias apresentadas nas Tabelas 5, 6, 8, 10 e 12, pode-se notar um mesmo padrão. Conforme os valores de intervalo de gerações analisadas são decrementados na abordagem de parada proposta, de 300 até 100, a tendência geral é de diminuição na média dos tempos de execução do AG, com valores inferiores aos alcançados com a abordagem original. Analisando, para cada intervalo de gerações, o valor de limiar utilizado, nota-se que as médias dos tempos de execução também diminuem conforme ocorre o incremento nos valores de limiar, de $10E-3$ até $10E-1$.

Como o AG diminui o tempo de execução, menos gerações/iterações são realizadas. Com isso, a tendência é que os valores de médias de *fitness* também diminuam. Isso se mostrou especialmente para o experimento 11 da abordagem de parada proposta, quando para quatro dos cinco casos de teste este experimento obteve uma média de *fitness* significativamente inferior às médias alcançadas com os experimentos da abordagem de parada original.

Também, o experimento da abordagem original com 1500 iterações, mesmo realizando mais iterações que os demais experimentos e com isso aumentando o tempo de execução do AG, não apresentou ganhos significativos em relação ao valor médio de *fitness* alcançado nos demais experimentos em todos os casos de teste.

Com relação aos demais experimentos da abordagem de parada proposta, apenas os experimentos 8 e 10 mantiveram a média de *fitness* alcançada com a abordagem original enquanto diminuíram a média de tempo de execução para todos os casos de teste realizados. O experimento 10, com valor de limiar de $10E-2$ e com análise de variância de 100 gerações, foi o que mais se destacou. De acordo com as análises estatísticas realizadas, este experimento obteve os melhores ganhos com relação ao tempo, chegando a atingir 50,87% em um dos casos de teste.

Portanto, os resultados mostraram que a abordagem de parada proposta por Bhandari, Murthy and Pal (2012) pode ser utilizada para efetuar a parada antecipada e trazer ganhos de desempenho ao AG, sem prejuízo ao valor de *fitness*. Além disso, diferentemente do que foi apresentado com a abordagem de parada de Yeng, Yoke and Suhaimi (2019), é possível utilizar um único valor de limiar para diferentes casos de teste.

4.2 Restrição de Utilização de Touros

A funcionalidade de restringir a utilização de touros, através de variáveis de uso máximo e mínimo no AG, bem como de penalização no caso de quebra dessas restrições, teve sua efetividade avaliada através da realização de seis experimentos, nos quais foram verificados se as soluções geradas estavam de acordo com os limites de utilização pré-estabelecidos.

Todos os experimentos utilizaram uma base de dados distinta daquela utilizada nos experimentos do critério de parada do AG e foram extensivamente repetidos, apresentando sempre o mesmo padrão nos resultados. Quando esta base de dados de animais era executada no AG, antes da implementação de melhorias, apresentava erro na solução gerada pois permitia que a restrição de utilização máxima fosse quebrada.

Através de uma pesquisa minuciosa sobre o código implementado originalmente no AG de Fontoura (2019), identificou-se que alguns cenários de erro não foram considerados. Isso demandou uma refatoração de código para que o AG executasse da forma esperada. A seguir, os experimentos apresentam o problema, como também os resultados após a refatoração de código.

No experimento 1, o objetivo foi justamente demonstrar o erro gerado com a execução do AG antes da implementação das melhorias. A variável de mínima utilização recebeu o valor 0 e a de máxima utilização recebeu o valor 30.

No experimento 2, o objetivo foi demonstrar a solução gerada com a execução do AG após a refatoração de código, com as variáveis de utilização tendo os mesmos valores do primeiro experimento. No experimento 3, a variável de máxima utilização recebeu o valor 15 com o objetivo de verificar se a restrição de utilização máxima era observada pelo AG. A variável de mínima utilização manteve-se com o valor 0.

No experimento 4, a variável de mínima utilização teve seu valor ajustado para 5 com o objetivo de verificar se a restrição de utilização mínima era observada pelo AG. A variável de máxima utilização recebeu o valor 30.

No experimento 5, observando os touros mais relacionados nas soluções do AG dos experimentos anteriores, por terem melhor IQG, foram selecionados dois destes touros que tiveram seus limites individuais de utilização máxima atribuídos para valor zero. O objetivo foi verificar se a solução gerada pelo AG descartaria a utilização daqueles animais. Para os demais animais, o valor de máxima utilização foi de 30 e de mínima utilização foi de 0.

Por fim, no experimento 6, observando os touros não relacionados nas soluções do AG dos experimentos anteriores, por terem pior IQG, foram selecionados dois destes touros que tiveram seus limites individuais de utilização mínima atribuídos para valor 30. O objetivo foi verificar se a solução gerada pelo AG recomendaria a utilização daqueles animais. Para os demais animais, o valor de mínima utilização foi de 0 e o valor de máxima utilização também foi de 30. A Tabela 13 apresenta os parâmetros utilizados em todos os seis experimentos de verificação de utilização de touros.

Tabela 13 – Parâmetros utilizados nos experimentos de utilização de touros.

Parâmetro	Valor
Número de Touros	20
Número de Vacas	200
Tamanho da População	400
Taxa de Consanguinidade	$\leq 3\%$
Taxa de Mutação	10%
Função de Avaliação	90% IQG e 10% NP

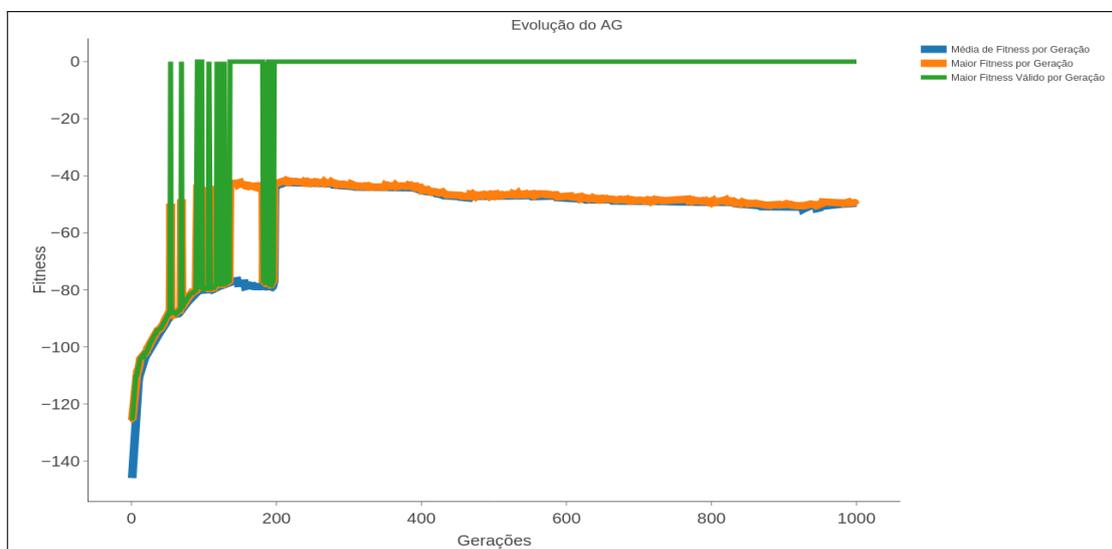
Fonte: Autor(2021)

A seguir, são apresentados os resultados para o experimento 1. A Figura 19 apresenta o gráfico da evolução das soluções geradas pelo AG. Como é possível notar pela linha verde do gráfico, que representa os cromossomos válidos com maior *fitness* por geração, em algumas gerações, o AG não encontra um cromossomo válido. Os valores oscilam entre um valor negativo de *fitness*, que foi o valor computado para o grupo de animais selecionado, e o valor zero absoluto.

O valor de *fitness* negativo resulta do fato que a maioria dos animais possuem valores de IQG negativos, o que impacta no cálculo da aptidão final do cromossomo. Já o valor de *fitness* zero absoluto, indica que o cromossomo foi considerado inválido por ter quebrado alguma restrição, que neste caso foi de utilização máxima de touros. Esta foi a abordagem inicialmente utilizada no AG e que teve que ser refatorada.

O AG criava uma estrutura do tipo matriz para receber os cromossomos e seus valores de *fitness* em cada geração. Os cromossomos considerados inválidos também eram adicionados, porém com o valor de zero absoluto. Após isso, o AG utilizava uma função para buscar o maior valor de *fitness* na estrutura. Como neste conjunto de animais os valores computados eram negativos, sempre que existia um cromossomo com valor

Figura 19 – Evolução das soluções geradas pelo AG no experimento 1 de utilização de touros.



Fonte: Autor(2021)

de *fitness* zero, o AG acabava retornando-o como sendo o de maior valor. No código refatorado, os cromossomos inválidos são adicionados à matriz de cromossomos com valor de *fitness* nulo, ou *Null* na linguagem R, e não são considerados pela função que busca o valor máximo.

Para evitar recomendações de acasalamentos não válidas quanto a utilização de touros, o AG aplica uma penalização no valor de aptidão dos cromossomos que quebram esta restrição a cada geração, evitando que esse problema se propague nas gerações posteriores. Porém, o código original do AG não penalizava corretamente os cromossomos com valor de aptidão negativo, pois previa somente um cenário com valores positivos. A penalização apenas dividia o valor de aptidão por 1,8 e por isso teve que ser refatorada para incluir a multiplicação por este mesmo valor em caso de valores negativos.

A Figura 20 apresenta uma relação de touros, retornada pelo AG, que contém o id de cada animal, o máximo e mínimo de utilização escolhidos antes de executar o AG, a utilização de cada touro retornada como resposta pelo AG e se os acasalamentos para aquele touro foram considerados válidos ou não, além do valor de IQG de cada touro.

Como é possível observar na Figura 20, o touro de id 116146 é relacionado com valor de utilização igual a 32, o que excede o valor de utilização máxima escolhido para aquele experimento, igual a 30. Por causa da quebra da restrição de utilização de um único touro, o cromossomo, que é o conjunto de acasalamentos entre os animais, foi

Figura 20 – Relação de utilização de touros do experimento 1.

Id: 43388	Máximo Uso: 30	Mínimo Uso: 0	Uso: 23	Válido?	SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 30	Mínimo Uso: 0	Uso: 24	Válido?	SIM	IQG: 0.169
Id: 69252	Máximo Uso: 30	Mínimo Uso: 0	Uso: 14	Válido?	SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido?	SIM	IQG: -16.787
Id: 73610	Máximo Uso: 30	Mínimo Uso: 0	Uso: 5	Válido?	SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 30	Mínimo Uso: 0	Uso: 5	Válido?	SIM	IQG: -0.593
Id: 76093	Máximo Uso: 30	Mínimo Uso: 0	Uso: 17	Válido?	SIM	IQG: -0.178
Id: 78866	Máximo Uso: 30	Mínimo Uso: 0	Uso: 12	Válido?	SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 30	Mínimo Uso: 0	Uso: 2	Válido?	SIM	IQG: -13.17
Id: 86180	Máximo Uso: 30	Mínimo Uso: 0	Uso: 9	Válido?	SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 30	Mínimo Uso: 0	Uso: 8	Válido?	SIM	IQG: -0.429
Id: 114845	Máximo Uso: 30	Mínimo Uso: 0	Uso: 13	Válido?	SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 30	Mínimo Uso: 0	Uso: 1	Válido?	SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 30	Mínimo Uso: 0	Uso: 32	Válido?	NÃO	IQG: -0.0997
Id: 116147	Máximo Uso: 30	Mínimo Uso: 0	Uso: 6	Válido?	SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 30	Mínimo Uso: 0	Uso: 8	Válido?	SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 30	Mínimo Uso: 0	Uso: 2	Válido?	SIM	IQG: -0.868
Id: 131384	Máximo Uso: 30	Mínimo Uso: 0	Uso: 5	Válido?	SIM	IQG: -0.912
Id: 137506	Máximo Uso: 30	Mínimo Uso: 0	Uso: 3	Válido?	SIM	IQG: -12.192
Id: 157014	Máximo Uso: 30	Mínimo Uso: 0	Uso: 11	Válido?	SIM	IQG: -0.5337

Fonte: Autor(2021)

considerado inválido.

A Figura 21 apresenta o gráfico da evolução das soluções geradas pelo AG no experimento 2. Diferentemente do que aconteceu no experimento 1, o gráfico gerado mostra, pela linha verde, que o AG evoluiu de uma solução válida inicial com valor de *fitness* de -128,35 para a solução final do problema que atingiu o valor de -63,49.

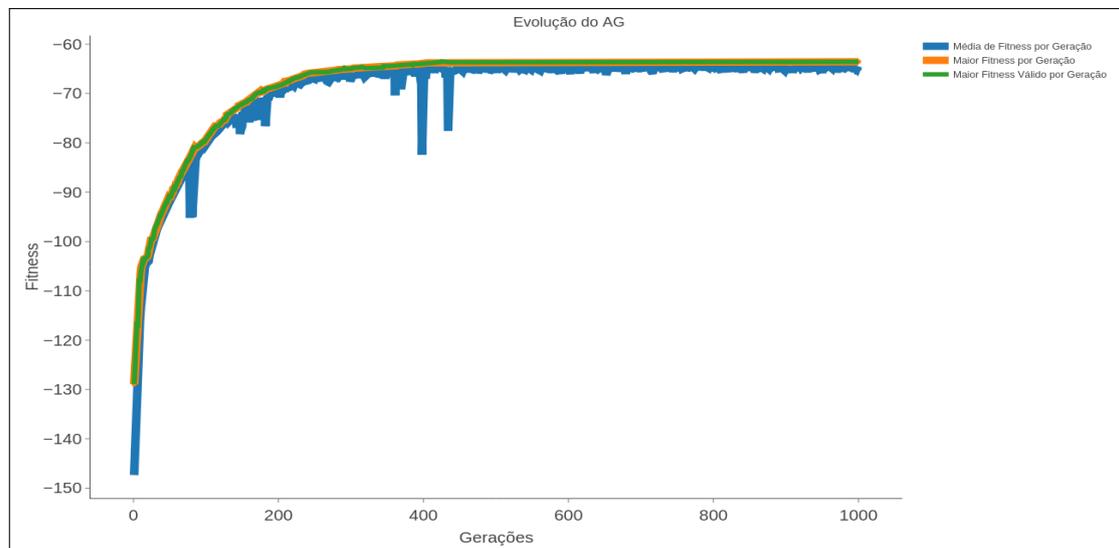
O *fitness* inicial é gerado da primeira geração de cromossomos que são formados por acasalamentos aleatórios e por isso apresenta um baixo valor. Com a evolução nas gerações seguintes, o AG obteve um resultado final 2,02 vezes melhor que o inicial.

A Figura 22 apresenta a relação de touros retornada pelo AG no experimento 2. Como é possível observar na figura, o touro de id 116146, que antes excedia a utilização máxima, apresenta utilização igual a 30, dentro da restrição imposta. Os demais touros listados encontram-se dentro dos limites de utilização pré-estabelecidos.

No experimento 3, a Figura 23 apresenta o gráfico da evolução das soluções geradas com valores de máxima utilização de touros iguais a 15. Os valores de mínima utilização permaneceram com valor 0. O gráfico gerado mostra, pela linha verde, que o AG evoluiu de uma solução válida inicial com valor de *fitness* de -129,32 até atingir a solução final com valor de -97,13, sendo esta 1,33 vezes melhor que a inicial.

A Figura 24 apresenta a relação de touros retornada pelo AG no experimento 3.

Figura 21 – Evolução das soluções geradas pelo AG no experimento 2 de utilização de touros.



Fonte: Autor(2021)

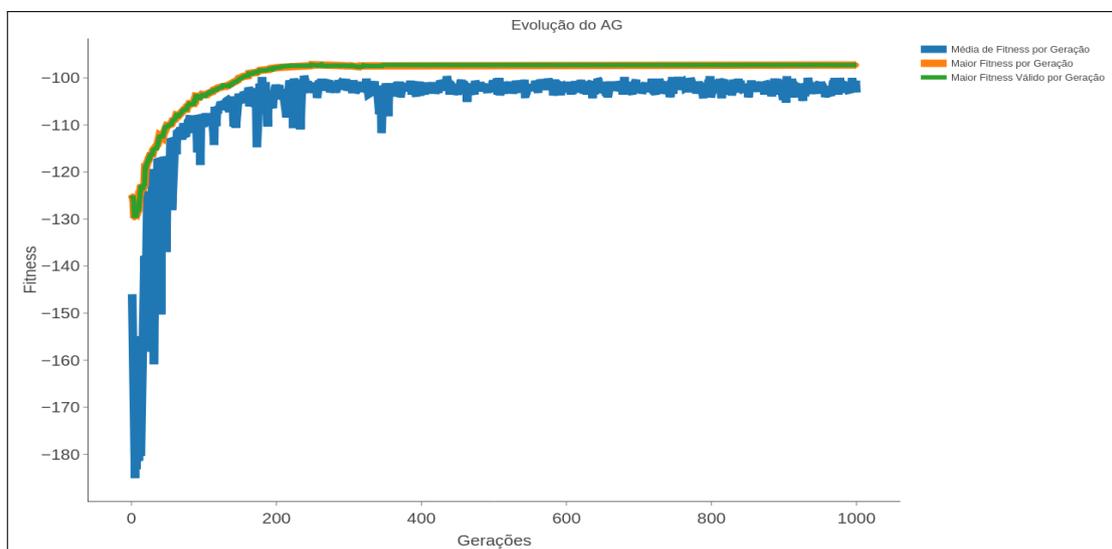
Figura 22 – Relação de utilização de touros do experimento 2.

Id: 43388	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.169
Id: 69252	Máximo Uso: 30	Mínimo Uso: 0	Uso: 28	Válido? SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -16.787
Id: 73610	Máximo Uso: 30	Mínimo Uso: 0	Uso: 2	Válido? SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.593
Id: 76093	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.178
Id: 78866	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -13.17
Id: 86180	Máximo Uso: 30	Mínimo Uso: 0	Uso: 7	Válido? SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.429
Id: 114845	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.0997
Id: 116147	Máximo Uso: 30	Mínimo Uso: 0	Uso: 1	Válido? SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 30	Mínimo Uso: 0	Uso: 11	Válido? SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.868
Id: 131384	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.912
Id: 137506	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -12.192
Id: 157014	Máximo Uso: 30	Mínimo Uso: 0	Uso: 1	Válido? SIM	IQG: -0.5337

Fonte: Autor(2021)

Como é possível observar na figura, todos os touros relacionados apresentam utilização dentro do limite de utilização máxima igual a 15, fazendo com que a solução seja válida. Os touros de id 43388 e 48289, com melhores IQGs e mais utilizados na solução dos

Figura 23 – Evolução das soluções geradas pelo AG no experimento 3 de utilização de touros.



Fonte: Autor(2021)

outros experimentos, não quebraram a restrição de máxima utilização imposta. Alguns touros de IQG intermediário, como é o caso dos touros de id 116187 e 73613, tiveram utilização igual a 14 e 6 respectivamente. Os touros com IQGs mais baixos não foram utilizados na solução.

A limitação de utilização máxima imposta neste experimento fez com que os animais melhor avaliados fossem menos utilizados. Com isso, o valor de *fitness* final foi menor neste experimento 3 quando comparado ao experimento 2.

No experimento 4, a Figura 25 apresenta o gráfico da evolução das soluções geradas com valores de mínima utilização de touros iguais a 5. Os touros receberam valores de máxima utilização iguais a 30. O gráfico gerado mostra, pela linha verde, que o AG evoluiu de uma solução válida inicial com valor de *fitness* de -134,57 até atingir a solução final com valor de -94,94, sendo esta 1,41 vezes melhor que a inicial.

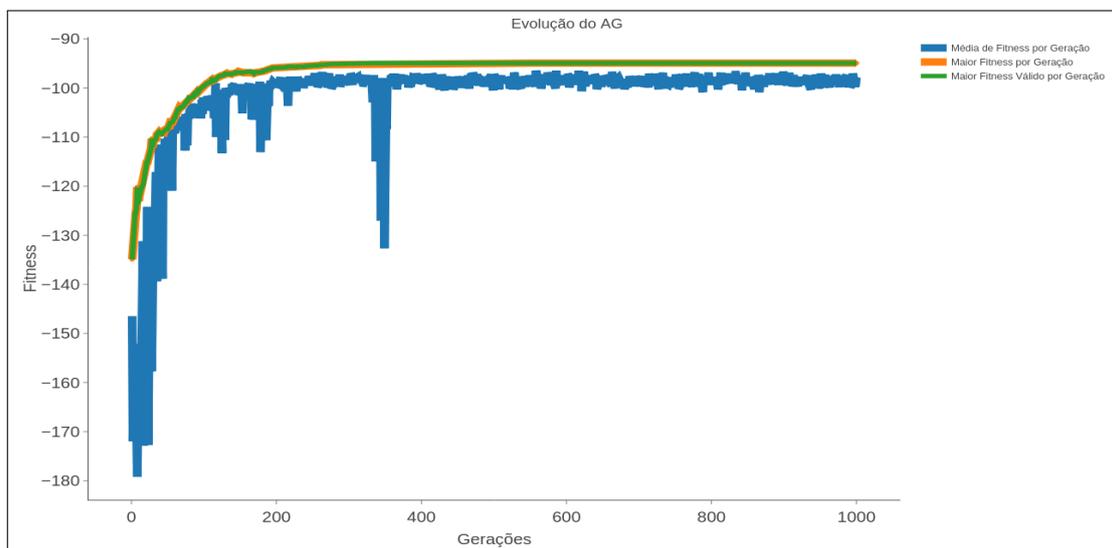
A Figura 26 apresenta a relação de touros retornada pelo AG no experimento 4. Como é possível observar na figura, todos os touros relacionados apresentam utilização dentro do limite de utilização mínima imposto, fazendo com que a solução seja válida. Ou seja, todos os touros foram utilizados com no mínimo 5 vacas. Os touros de id 69344 e 78912, com IQGs mais baixos e não utilizados na solução dos outros experimentos, não quebraram a restrição de mínima utilização imposta. Os touros com IQG mais alto obtiveram utilização igual a 30 na solução.

Figura 24 – Relação de utilização de touros do experimento 3.

Id: 43388	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: 0.169
Id: 69252	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -16.787
Id: 73610	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 15	Mínimo Uso: 0	Uso: 6	Válido? SIM	IQG: -0.593
Id: 76093	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.178
Id: 78866	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -13.17
Id: 86180	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.429
Id: 114845	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.0997
Id: 116147	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 15	Mínimo Uso: 0	Uso: 14	Válido? SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.868
Id: 131384	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.912
Id: 137506	Máximo Uso: 15	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -12.192
Id: 157014	Máximo Uso: 15	Mínimo Uso: 0	Uso: 15	Válido? SIM	IQG: -0.5337

Fonte: Autor(2021)

Figura 25 – Evolução das soluções geradas pelo AG no experimento 4 de utilização de touros.



Fonte: Autor(2021)

A utilização mínima imposta neste experimento fez com que todos os touros de pior IQG fossem recomendados para acasalamento com pelo menos 5 vacas. Por isso, quando comparado com o experimento 2, o valor de *fitness* final foi menor.

Figura 26 – Relação de utilização de touros do experimento 4.

Id: 43388	Máximo Uso: 30	Mínimo Uso: 5	Uso: 30	Válido? SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 30	Mínimo Uso: 5	Uso: 30	Válido? SIM	IQG: 0.169
Id: 69252	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -16.787
Id: 73610	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.593
Id: 76093	Máximo Uso: 30	Mínimo Uso: 5	Uso: 30	Válido? SIM	IQG: -0.178
Id: 78866	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -13.17
Id: 86180	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.429
Id: 114845	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 30	Mínimo Uso: 5	Uso: 30	Válido? SIM	IQG: -0.0997
Id: 116147	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.868
Id: 131384	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.912
Id: 137506	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -12.192
Id: 157014	Máximo Uso: 30	Mínimo Uso: 5	Uso: 5	Válido? SIM	IQG: -0.5337

Fonte: Autor(2021)

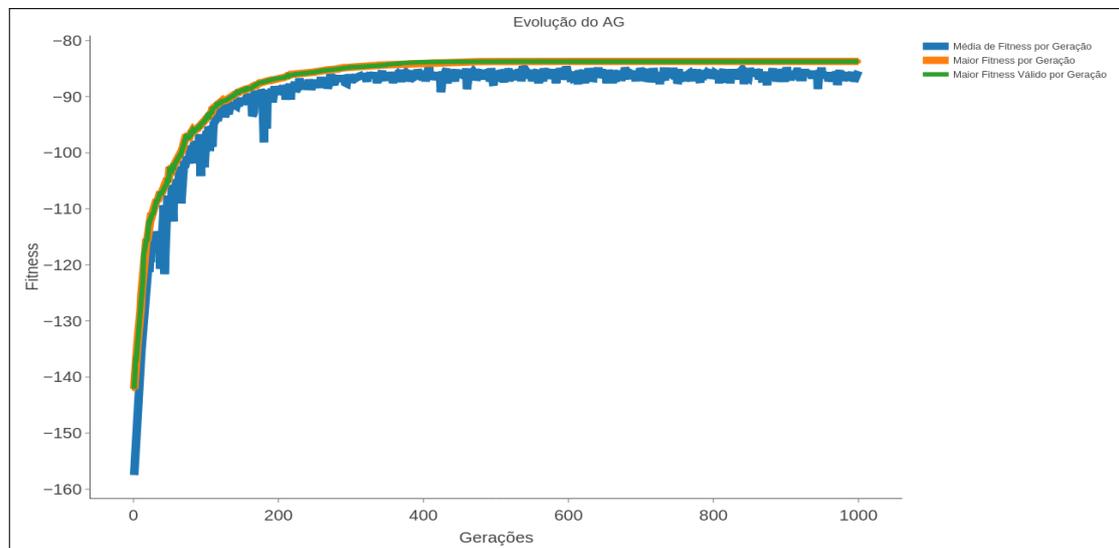
No experimento 5, a Figura 27 apresenta o gráfico da evolução das soluções geradas quando os touros de id 43388 e 48289 receberam valores para o parâmetro de máxima utilização iguais a 0 para analisar se o AG descartaria a utilização daqueles touros na solução dada como resposta. O gráfico gerado mostra, pela linha verde, que o AG evoluiu de uma solução válida inicial com valor de *fitness* de -142,25 até atingir a solução final com valor de -84,44, sendo esta 1,68 vezes melhor que a inicial.

A Figura 28 apresenta a relação de touros retornada pelo AG no experimento 5. Como é possível observar na figura, os touros de id 43388 e 48289 não possuem nenhuma utilização, sendo que os touros restantes com melhor IQG receberam maior utilização. Com isso, a solução dada pelo AG foi válida para todos os touros. Porém, como dois dos melhores touros não foram utilizados, o valor de *fitness* final foi menor neste experimento 5 (-84,44) se comparado ao experimento 2 (-63,49), onde estes touros puderam ser utilizados.

No experimento 6, a Figura 29 apresenta o gráfico da evolução das soluções geradas quando os touros de id 69344 e 78912 receberam valores para o parâmetro de mínima utilização iguais a 30 para analisar se o AG incluiria aqueles touros na solução dada como resposta e com a utilização mínima exigida.

O gráfico gerado mostra, pela linha verde, que o AG evoluiu de uma solução válida

Figura 27 – Evolução das soluções geradas pelo AG no experimento 5 de utilização de touros.



Fonte: Autor(2021)

Figura 28 – Relação de utilização de touros do experimento 5.

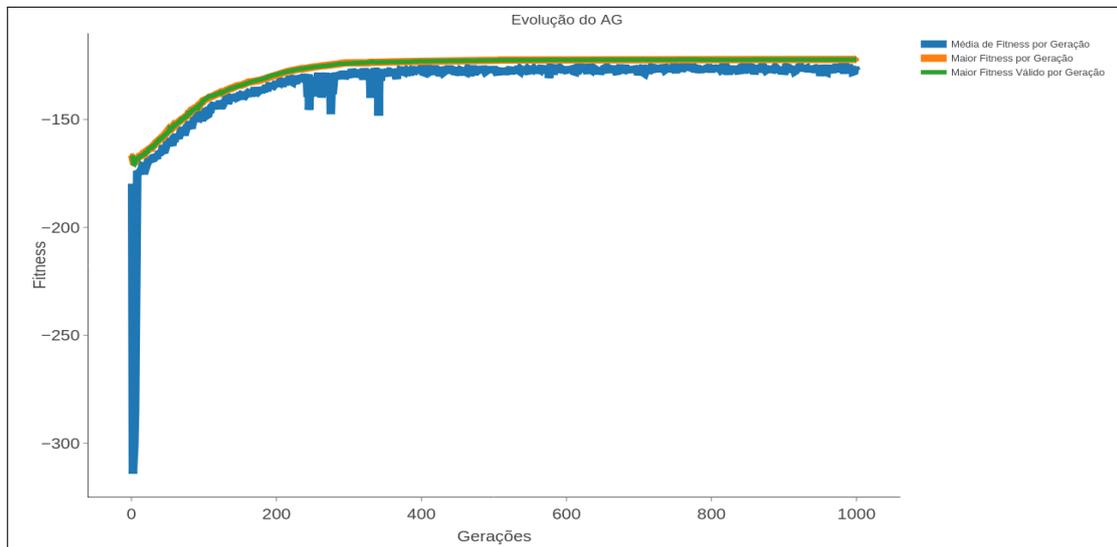
Id: 43388	Máximo Uso: 0	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 0	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.169
Id: 69252	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -16.787
Id: 73610	Máximo Uso: 30	Mínimo Uso: 0	Uso: 4	Válido? SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.593
Id: 76093	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.178
Id: 78866	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -13.17
Id: 86180	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.429
Id: 114845	Máximo Uso: 30	Mínimo Uso: 0	Uso: 16	Válido? SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.0997
Id: 116147	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.868
Id: 131384	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.912
Id: 137506	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -12.192
Id: 157014	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.5337

Fonte: Autor(2021)

inicial com valor de *fitness* de -166,63 até atingir a solução final com valor de -122,11, sendo esta 1,36 vezes melhor que a inicial. O *fitness* alcançado com este experimento, mais negativo que nos demais, resulta-se do fato de que os touros com piores IQGs foram

incluídos na solução do problema.

Figura 29 – Evolução das soluções geradas pelo AG no experimento 6 de utilização de touros.



Fonte: Autor(2021)

A Figura 30 apresenta a relação de touros retornada pelo AG no experimento 6. Como é possível observar na figura, os touros de id 69344 e 78912 possuem utilização igual a 30, respeitando a utilização mínima imposta. A solução dada pelo AG também foi válida para demais touros.

Portanto, os resultados apresentados pelos experimentos validaram as hipóteses levantadas em cada um deles, demonstrando que a refatoração realizada no código sanou o problema da não convergência do AG para uma solução válida em alguns casos de teste, e que também não trouxe nenhum erro adicional ao código já implementado.

4.3 Eliminação de *Hardcode* no AG

Esta seção apresenta os resultados da refatoração do código do AG para eliminação de *hardcode* com DEPs. No código original do AG, havia dependência com as DEPs utilizadas no PampaPlus, o que impossibilitava sua utilização em outros programas de melhoramento genético. Mesmo realizando consultas dos valores de DEPs no banco de dados do programa, fazer uso de DEPs diferentes ou até mesmo mudar a função do cálculo do IQG demandava que um programador fizesse modificações diretamente no código

Figura 30 – Relação de utilização de touros do experimento 6.

Id: 43388	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.1108
Id: 48289	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.169
Id: 69252	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.2395
Id: 69344	Máximo Uso: 30	Mínimo Uso: 30	Uso: 30	Válido? SIM	IQG: -16.787
Id: 73610	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.3813
Id: 73613	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.593
Id: 76093	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.178
Id: 78866	Máximo Uso: 30	Mínimo Uso: 0	Uso: 20	Válido? SIM	IQG: -0.2706
Id: 78912	Máximo Uso: 30	Mínimo Uso: 30	Uso: 30	Válido? SIM	IQG: -13.17
Id: 86180	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.4816
Id: 97836	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.429
Id: 114845	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.5274
Id: 114846	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.8669
Id: 116146	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: -0.0997
Id: 116147	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.4018
Id: 116187	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.3311
Id: 126274	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.868
Id: 131384	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.912
Id: 137506	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -12.192
Id: 157014	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.5337

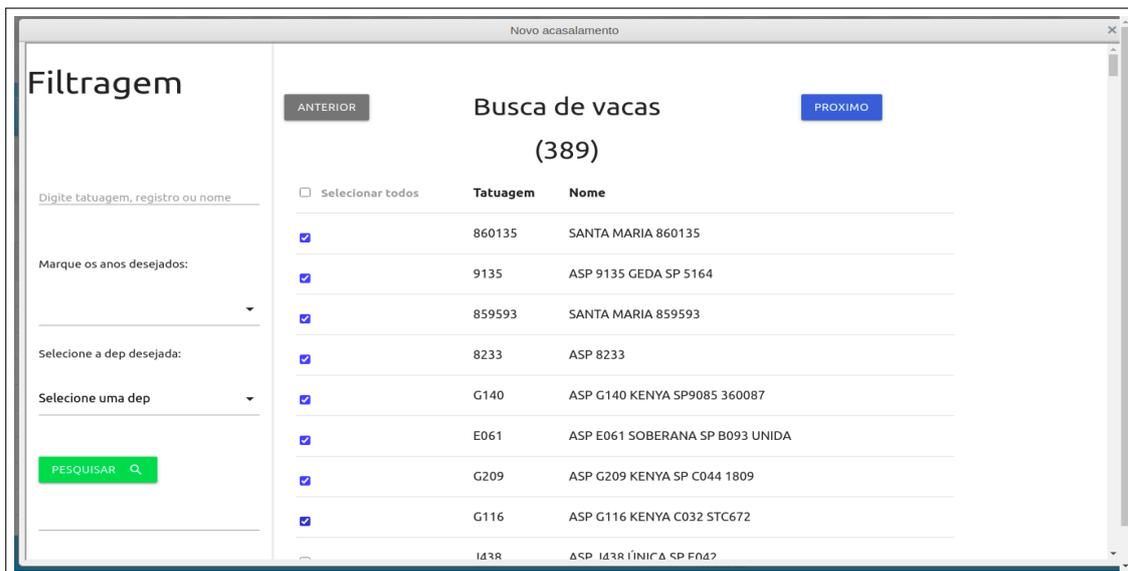
Fonte: Autor(2021)

fonte.

A refatoração de código permitiu que os dados fossem obtidos através de parâmetros externos ao AG, o que também possibilitou a integração do AG com o sistema *web* PampaPlusNet. Foram realizadas implementações de uma interface provisória e de ajustes no banco de dados do programa em parceria com técnicos da Embrapa. Através da interface, os usuários podem selecionar os animais, as DEPs com seus respectivos pesos e também impor limites de consanguinidade. Como se trata de integração inicial, a funcionalidade de impor utilização máxima e mínima de touros não está disponível para seleção pelos usuários ainda, sendo utilizados valores padrão.

Foram desenvolvidos dois casos de teste para demonstrar que o AG recebe os valores das DEPs como parâmetros externos, de acordo com a escolha dos usuários na interface, executa e realiza a recomendação dos animais para acasalamentos. Cada experimento utilizou um conjunto de DEPs diferentes, porém com a mesma seleção de animais.

A Figura 31 apresenta a tela inicial de escolha de animais. Como pode ser visto na figura, inicialmente é realizada a escolha das vacas para acasalamento, sendo que a próxima tela, de escolha de touros, é idêntica a esta. A tela permite ao usuário aplicar filtros para busca de animais por tatuagem, ano de nascimento ou valores de DEPs.

Figura 31 – Tela de escolha de animais da interface *web*.

Fonte: Autor(2021)

A Figura 32 apresenta a tabela de banco de dados desenvolvida para persistência dos dados da seleção já populada com os animais selecionados para acasalamento nos dois experimentos. Como mostrado na figura, para cada animal selecionado, são persistidas as informações de identificador da seleção atual a qual o animal faz parte, de identificador do proprietário, que é o usuário logado que realizou a seleção, e de identificador do animal no sistema junto com sexo e ano de nascimento.

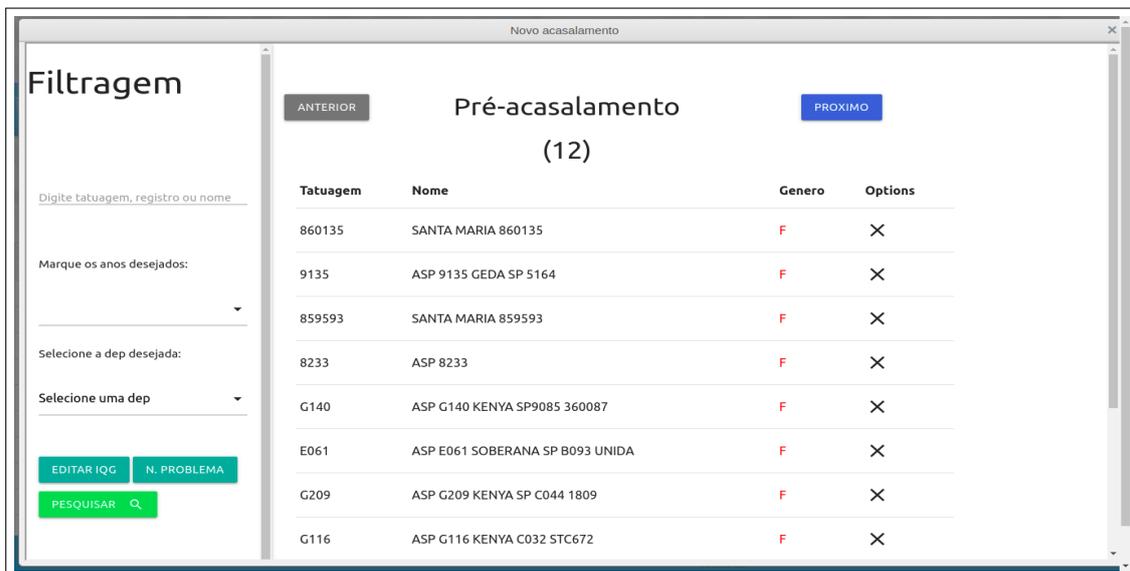
Figura 32 – Tabela de banco de dados dos animais selecionados nos dois casos de teste.

#	selecao_id	proprietarios_id	facadeanimais_id	sexo	ano
1	301	557	1049058	M	2018
2	301	557	944557	F	2016
3	301	557	926544	M	2016
4	301	557	871341	F	2014
5	301	557	363380	F	2008
6	301	557	344803	F	2009
7	301	557	289758	F	2007
8	301	557	289755	F	2007

Fonte: Autor(2021)

Após a seleção de animais para acasalamento, a próxima tela, mostrada na Figura 33, permite ao usuário visualizar todos os animais selecionados, além de editar o IQG. Um clique no botão “Editar IQG” à esquerda da tela, abre uma janela que permite ao usuário definir as DEPs que serão utilizadas e seus pesos no cálculo do IQG.

Figura 33 – Tela de listagem de animais selecionados da interface *web*.



Fonte: Autor(2021)

No Caso de Teste 1, foram utilizados as DEPs e seus pesos do IQG padrão do PampaPlus, de acordo com a seção 2.2.2. A Figura 34 mostra a janela que permite ao usuário realizar esta ação já com as DEPs e seus pesos selecionados.

Figura 34 – Tela de escolha de DEPs e edição do IQG da interface *web* para o Caso de Teste 1.

Editar IQG		
Peso ao Nascer	Peso à Desmama	Peso à Desmama Materno
0,00	0,00	0,00
Total Materno	Peso ao Sobreano	Ganho Pós Desmama
30,00	15,00	15,00
Peso da Vaca à Desmama	Condição Corporal da Vaca à Desmama	Circunferência Escrotal
0,00	0,00	15,00
Musculatura	Estatura	Estrutura Corporal
12,50	12,50	0,00
Umbigo	Pigmentação Ocular	Racial
0,00	0,00	0,00
Área de Olho de Lombo	Gordura Intramuscular	Acabamento
0,00	0,00	0,00

Fonte: Autor(2021)

As informações de DEPs são persistidas no banco de dados como mostrado na Figura 35. Na tabela, são incluídas as informações do identificador da seleção, do tipo de configuração, da sigla da DEP selecionada e o seu valor de ponderação no IQG em porcentagem.

Figura 35 – Tabela de banco de dados das DEPs selecionadas no Caso de Teste 1.

#	selecao_id	tipo_config	dep	valor
1	301	1	CCVDEP	0.00
2	301	1	CEDEP	15.00
3	301	1	dep_AOLd	0.00
4	301	1	dep_EGP8d	0.00
5	301	1	dep_GIMd	0.00
6	301	1	dep_RCSd	0.00
7	301	1	ECDEP	0.00
8	301	1	EDEP	12.50
9	301	1	GPDDEP	15.00
10	301	1	MDEP	12.50
11	301	1	PDDEP	0.00
12	301	1	PDDEPM	0.00
13	301	1	PDTM	30.00
14	301	1	PNDEP	0.00
15	301	1	PODEP	0.00
16	301	1	PSDEP	15.00
17	301	1	PVDEP	0.00
18	301	1	UDEP	0.00
19	301	2		1.00
20	301	3		0.03

Fonte: Autor(2021)

Na coluna do tipo de configuração, o número 1 representa uma configuração de DEP, o número 2 uma configuração do Nível de Problema e o número 3 uma configuração do fator de consanguinidade máxima. A configuração do Nível de Problema será utilizada no futuro, quando será possível aplicar restrições de Nível de Problema máximo nos acasalamentos. Por enquanto, é utilizado o cálculo padrão no AG que penaliza o alto Nível de Problema. As siglas utilizadas nas DEPs e o sua descrição podem ser visualizados na Figura 36, que traz uma consulta em uma tabela do banco de dados que contém o cadastro de DEPs. A esta tabela podem ser incluídas outras DEPs, para outros programas de melhoramento, e serem utilizadas na ferramenta.

Realizada a seleção de animais e de parâmetros de DEPs e consanguinidade, um clique no botão “Próximo” da tela mostrada na Figura 33, inicia a execução no AG no

Figura 36 – Descrição de DEPs cadastradas no banco de dados.

#	sigla	descrição
1	PNDEP	Peso ao Nascer
2	PDDEP	Peso à Desmama
3	PDDEPM	Peso à Desmama Materno
4	PDTM	Total Materno
5	PSDEP	Peso ao Sobreano
6	GPDEP	Ganho Pós Desmama
7	PVDEP	Peso da Vaca à Desmama
8	CCVDEP	Condição Corporal da Vaca à Desmama
9	CEDEP	Circunferência Escrotal
10	MDEP	Musculatura
11	EDEP	Estatuta
12	ECDEP	Estrutura Corporal
13	UDEP	Umbigo
14	PODEP	Pigmentação Ocular
15	dep_RCSD	Racial
16	dep_AOLd	Área de Olho de Lombo
17	dep_GIMd	Gordura Intramuscular
18	dep_EGP8d	Acabamento

Fonte: Autor(2021)

back-end da interface. Neste primeiro caso de teste, o tempo de execução foi de 2,73 segundos. Após isso, a recomendação de animais dada pelo AG é apresentada em tela. A Figura 37 apresenta a tela de resultados com uma matriz de todos os acasalamentos possíveis, com vacas nas linhas e touros nas colunas. Os acasalamentos recomendados pelo AG estão destacados na cor verde, com informações de IQG, consanguinidade e NP para cada acasalamento.

A Figura 38 apresenta a tabela de banco de dados onde são persistidos os dados da recomendação do AG, sendo que cada linha consiste de um acasalamento. A tabela possui colunas para guardar o identificador da seleção de animais, um identificador do produto gerado pelo acasalamento, os identificadores dos touros e das vacas e a consanguinidade para o acasalamento. A última coluna de nome “*ag*” recebe valores binários, com o 1 identificando um acasalamento recomendado e 0 um acasalamento descartado pelo AG.

Como pode ser visualizado na Figura 38, para o conjunto de animais e as DEPs selecionados no Caso de Teste 1, o AG recomendou os acasalamentos das linhas 2, 4, 6, 8, 10 e 11 da tabela. O valor de *fitness* calculado pelo AG foi de 14,82 para esta

Figura 37 – Tela de recomendação de acasalamentos da interface *web*.

VACAS / TOUROS	M060 248138	R082 259452
9135 2009	IQG: 3.03 CONS: 0.00% NP: 0	IQG: 2.92 CONS: 0.00% NP: 0
860135 2007	IQG: 3.03 CONS: 0.00% NP: 0	IQG: 2.91 CONS: 0.00% NP: 0
859593 2007	IQG: 2.98 CONS: 0.00% NP: 0	IQG: 2.88 CONS: 0.00% NP: 0
8233 2008	IQG: 3.12 CONS: 0.00% NP: 0	IQG: 3 CONS: 0.00% NP: 0
G140 2016	IQG: 3.07 CONS: 0.00% NP: 0	IQG: 2.95 CONS: 0.00% NP: 0
E061 2014	IQG: 3.04 CONS: 0.00% NP: 0	IQG: 2.92 CONS: 0.00% NP: 0

Fonte: Autor(2021)

Figura 38 – Recomendação de acasalamentos no banco de dados para o Caso de Teste 1.

#	selecao_id	prod_id	touro_id	vaca_id	consanguinidade	ag
1	301	301000012	926544	289755	0	0
2	301	301000011	1049058	289755	0	1
3	301	301000010	926544	289758	0	0
4	301	301000009	1049058	289758	0	1
5	301	301000008	926544	344803	0	0
6	301	301000007	1049058	344803	0	1
7	301	301000006	926544	363380	0	0
8	301	301000005	1049058	363380	0	1
9	301	301000004	926544	871341	0.00000572205	0
10	301	301000003	1049058	871341	0.00000286102	1
11	301	301000002	926544	944557	0.00000572205	1
12	301	301000001	1049058	944557	0.00000286102	0

Fonte: Autor(2021)

recomendação.

No Caso de Teste 2, foram escolhidas DEPs diferentes daquelas utilizadas no Caso de Teste 1, porém se manteve o mesmo conjunto de animais selecionados. O objetivo foi demonstrar que o AG permite esta opção e também verificar diferenças na recomendação de acasalamentos. A Figura 39 apresenta a janela de seleção de DEPs e seus pesos para o Caso de Teste 2.

A Figura 40 apresenta a tabela do banco de dados com as DEPs selecionadas

Figura 39 – Tela de escolha de DEPs e edição do IQG da interface *web* para o Caso de Teste 2.

Editar IQG		
Peso ao Nascer	Peso à Desmama	Peso à Desmama Materno
15,00	15,00	0,00
Total Materno	Peso ao Sobreano	Ganho Pós Desmama
0,00	0,00	0,00
Peso da Vaca à Desmama	Condição Corporal da Vaca à Desmama	Circunferência Escrotal
0,00	30,00	0,00
Musculatura	Estatura	Estrutura Corporal
0,00	0,00	15,00
Umbigo	Pigmentação Ocular	Racial
0,00	0,00	12,50
Área de Olho de Lombo	Gordura Intramuscular	Acabamento
0,00	0,00	12,50

Fonte: Autor(2021)

juntamente com a ponderação de cada uma. Como mostrado, foram selecionadas DEPs diferentes, mas se manteve o mesmo fator de consanguinidade do Caso De Teste 1.

O AG executou com os novos parâmetros de DEPs em um tempo de 2,78 segundos. A Figura 41 apresenta a tabela de banco de dados onde foram persistidos os dados da recomendação do AG. Como mostrado, para o conjunto de animais e as DEPs selecionadas no Caso de Teste 2, o AG recomendou os acasalamentos das linhas 1, 3, 5, 7, 9 e 12 da tabela. O valor de *fitness* calculado pelo AG foi de 8,95 para esta recomendação.

Portanto, os casos de testes realizados demonstraram que os parâmetros de DEPs e suas ponderações não estão mais embutidos no código fonte do AG e que, através da interface *web*, o usuário pode efetuar a sua escolha. Também, caso o AG seja utilizado em programas de melhoramento genéticos que observam outras DEPs, estas podem ser cadastradas e então utilizadas, desde que seja mantida a mesma estrutura de tabelas de banco de dados.

Figura 40 – Tabela de banco de dados das DEPs selecionadas no Caso de Teste 2.

#	selecao_id	tipo_config	dep	valor
1	302	1	CCVDEP	30.00
2	302	1	CEDEP	0.00
3	302	1	dep_AOLd	0.00
4	302	1	dep_EGP8d	12.50
5	302	1	dep_GIMd	0.00
6	302	1	dep_RCSd	12.50
7	302	1	ECDEP	15.00
8	302	1	EDEP	0.00
9	302	1	GPDEP	0.00
10	302	1	MDEP	0.00
11	302	1	PDDEP	15.00
12	302	1	PDDEPM	0.00
13	302	1	PDTM	0.00
14	302	1	PNDEP	15.00
15	302	1	PODEP	0.00
16	302	1	PSDEP	0.00
17	302	1	PVDEP	0.00
18	302	1	UDEP	0.00
19	302	2		1.00
20	302	3		0.03

Fonte: Autor(2021)

Figura 41 – Recomendação de acasalamentos no banco de dados para o Caso de Teste 2.

#	selecao_id	prod_id	touro_id	vaca_id	consanguinidade	ag
1	302	302000012	926544	289755	0	1
2	302	302000011	1049058	289755	0	0
3	302	302000010	926544	289758	0	1
4	302	302000009	1049058	289758	0	0
5	302	302000008	926544	344803	0	1
6	302	302000007	1049058	344803	0	0
7	302	302000006	926544	363380	0	1
8	302	302000005	1049058	363380	0	0
9	302	302000004	926544	871341	0.00000572205	1
10	302	302000003	1049058	871341	0.00000286102	0
11	302	302000002	926544	944557	0.00000572205	0
12	302	302000001	1049058	944557	0.00000286102	1

Fonte: Autor(2021)

4.4 Benchmark do Algoritmo Genético

Após a efetivação das implementações adicionais no AG, foi realizado um *benchmark* do AG com outra ferramenta de recomendação de acasalamentos. A ferramenta utilizada no comparativo foi o EVA descrito no trabalho de Berg, Nielsen and Sørensen (2006), apresentado como trabalho correlato no Capítulo 2.4.

O objetivo deste *benchmark* foi de comparar as recomendações de acasalamentos geradas pelas duas ferramentas através do IQG apresentado em cada uma delas, sem levar em consideração o nível de problema. O EVA não utiliza qualquer informação de DEP, como o AG, e por isso o cálculo do mérito genético de um acasalamento utiliza somente o valor final de IQG dos animais.

Como o EVA foi desenvolvido com o objetivo primário de controlar os riscos de consanguinidade nos acasalamentos, foram projetados experimentos com bases de dados de diferentes taxas de consanguinidade média entre os animais. O EVA não permite impor restrições quanto ao máximo de consanguinidade entre os animais, ele somente evita recomendar um acasalamento consanguíneo. No AG, a restrição de consanguinidade máxima imposta para um acasalamento recomendado foi de 3%.

Em cada uma das ferramentas foram realizados três experimentos. No Experimento 1, foi utilizada uma base de dados com 165 animais, sendo 17 touros e 148 vacas, com taxa de consanguinidade média de 0%. Esta base de dados foi utilizada nos casos de teste da Seção 4.1. No Experimento 2, foi utilizada uma segunda base de dados, distinta da primeira, que contou com total de 24 animais, sendo 8 touros e 16 vacas e com uma taxa média de consanguinidade entre os animais de 3,88%, com os possíveis acasalamentos apresentando taxas de 0% até 25% de consanguinidade entre os animais.

No Experimento 3, foi utilizado um subconjunto da segunda base de dados com um total de 11 animais, sendo 5 touros e 6 vacas que apresentavam as maiores taxas de consanguinidades entre eles. O objetivo com este experimento foi de verificar o comportamento das duas ferramentas com um conjunto de animais com alta taxa de consanguinidade média. A taxa média de consanguinidade com este subconjunto foi de 9,48% com os potenciais acasalamentos apresentando taxas de 0% até 25% de consanguinidade entre os animais.

Foram utilizados os mesmos parâmetros nas duas ferramentas de acasalamento, sendo que foram ajustados de acordo com o contexto e objetivos deste comparativo. A seguir, a Tabela 14 apresenta os parâmetros em comum entre os três experimentos deste

benchmark.

Tabela 14 – Parâmetros utilizados em comum pelas ferramentas de acasalamento nos experimentos do *Benchmark*.

Parâmetro	Valor
Tamanho da População	Dobro do número de Vacas
Número de Iterações	1000
Taxa de Mutação	10%
Critério de Avaliação	100% do IQG

Fonte: Autor(2021)

Após a execução de cada experimento, foi coletado o tempo de execução e o IQG da solução gerada. Tanto no AG como no EVA foram realizadas 10 repetições para cada um dos três experimentos. A Tabela 15 apresenta a média das repetições para o tempo de execução e o IQG de cada ferramenta de acasalamento para os três experimentos realizados.

Tabela 15 – Média dos dados resultantes dos experimentos do *Benchmark*.

	Ferramenta	Tempo (seg)	IQG
Experimento 1	AG	68,08	0,5438
	EVA	127,50	0,1073
Experimento 2	AG	4,02	0,9912
	EVA	1,07	0,2386
Experimento 3	AG	2,61	1,2372
	EVA	0,37	0,7133

Repetições:10

Fonte: Autor(2021)

No Experimento 1, o AG atingiu um valor de IQG médio 5,07 vezes superior ao atingido com o EVA. Isto se deve a uma diferença de comportamento das ferramentas na recomendação dos acasalamentos. O AG prioriza a utilização dos touros de melhor IQG recomendando-os a um número maior de vacas para acasalamento, em detrimento dos touros de pior IQG, que são recomendados a um número menor de vacas ou a nenhuma.

Quando um touro possui um melhor valor de IQG, a média de IQG calculada com um acasalamento tende a aumentar, a depender também do IQG da vaca. Um exemplo é mostrado na Figura 42, que apresenta a listagem de touros e suas utilizações na recomendação dada pelo AG em uma das execuções do Experimento 1.

Figura 42 – Relação de utilização de touros em uma das execuções do AG no Experimento 1 do *Benchmark*.

Id: 257357	Máximo Uso: 30	Mínimo Uso: 0	Uso: 1	Válido? SIM	IQG: -0.0753
Id: 259451	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -1.6108
Id: 259455	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.3499
Id: 259456	Máximo Uso: 30	Mínimo Uso: 0	Uso: 27	Válido? SIM	IQG: 0.6508
Id: 284385	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.0731
Id: 284407	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -1.2555
Id: 284416	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.4628
Id: 318628	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.6615
Id: 331630	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.4727
Id: 375797	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -1.7902
Id: 531837	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.3522
Id: 535135	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 0.6793
Id: 536853	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: -0.953
Id: 795027	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.3588
Id: 795047	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 1.2078
Id: 795048	Máximo Uso: 30	Mínimo Uso: 0	Uso: 30	Válido? SIM	IQG: 1.8221
Id: 843984	Máximo Uso: 30	Mínimo Uso: 0	Uso: 0	Válido? SIM	IQG: 0.6152

Fonte: Autor(2021)

O EVA procura utilizar a maioria dos touros selecionados. Além disso, atribui os acasalamentos de forma mais distribuída entre os touros. Assim como no AG, os touros com maior valor de IQG também são mais utilizados, porém com pouca diferença para os touros de menor valor de IQG. A utilização dos piores touros faz com que o IQG médio calculado na solução do problema seja menor no EVA quando comparado ao AG. Esses padrões se repetiram em todas as execuções nas duas ferramentas. A Figura 43 apresenta a listagem de touros e suas utilizações na recomendação dada pelo EVA em uma das execuções do Experimento 1. A coluna de nome *sire* apresenta o identificador de cada touro e a coluna *N.matings* o número de acasalamentos associado.

Com relação ao tempo de execução, o EVA obteve um tempo de execução médio 1,87 vezes superior ao atingido pelo AG. Neste caso, o tamanho da população de cromossomos utilizado como parâmetro nas duas ferramentas pode ter sido a causa da diferença de tempo. O EVA possui exemplos prontos de execução e aprendizagem da ferramenta onde o número da população de cromossomos é menor do que o dobro do número de vacas, que é um dos parâmetros utilizados no AG e reproduzido também nos experimentos com o EVA. Quando este parâmetro foi reduzido pela metade no EVA, apenas para verificação, o tempo médio também reduziu pela metade.

Figura 43 – Relação de utilização de touros em uma das execuções do EVA no Experimento 1 do *Benchmark*.

Sire	N. matings
257357	9
259451	7
259455	8
259456	10
284385	9
284407	6
284416	8
318628	10
331630	9
375797	6
531837	8
535135	10
536853	8
795027	9
795047	11
795048	11
843984	9

Fonte: Autor(2021)

No Experimento 2, o AG obteve um valor médio de IQG 4,15 vezes superior ao obtido pelo EVA, apresentando os mesmos padrões das duas ferramentas quanto à distribuição de acasalamentos do Experimento 1. Com relação ao tempo, houve diferença para o Experimento 1, pois o EVA obteve um tempo médio de execução inferior ao tempo médio apresentado pelo AG, 3,76 vezes menor. A explicação para isso está na diminuição do número de cromossomos neste experimento, pois o número de vacas é menor, demonstrando que o EVA é mais rápido que AG para seleções pequenas de animais e que o AG torna-se mais rápido à medida que o número de animais aumenta.

Neste experimento foram incluídos animais consanguíneos no grupo de animais submetidos às ferramentas de acasalamento. No entanto, ambas as ferramentas recomendaram acasalamentos com uma taxa de consanguinidade baixa, nunca atingindo a taxa de 3%, considerada como parâmetro no AG. Esse padrão foi verificado em todas as repetições do experimento.

A Figura 44 apresenta a relação de acasalamentos dada como resposta pelo AG em uma de suas execuções, com os identificadores de touro e vaca e suas respectivas

consanguinidades.

Figura 44 – Relação de acasalamentos e consanguinidades em uma das execuções do AG no Experimento 2 do *Benchmark*.

touro_id	vaca_id	consanguinidade
440199	209754	0.0000000000
986486	209767	0.0000000000
440199	209772	0.0000000000
440199	275232	0.0000000000
986486	370511	0.0000000000
944596	370545	0.0000000000
440199	370546	0.0000000000
986486	370566	0.0000000000
440199	849802	0.0000000000
440199	933124	0.0003547670
986486	944534	0.0000600815
440199	944552	0.0157166000
440199	944558	0.0002403260
258151	944574	0.0000000000
440199	987103	0.0000000000
944596	987104	0.0002174380

Fonte: Autor(2021)

A Figura 45 apresenta a relação de acasalamentos dada como resposta pelo EVA em uma de suas execuções, com identificadores para os touros, vacas e o próprio acasalamento, Nm que corresponde ao número de acasalamentos para um determinado touro, $Nf.m$ que corresponde ao número de acasalamentos para uma determinada vaca, EBV que é o valor médio de IQG e F a consanguinidade para aquele acasalamento expressa em taxa.

No Experimento 3, o AG obteve um valor médio de IQG 1,73 vezes superior ao obtido pelo EVA, apresentando os mesmos padrões das duas ferramentas quanto à distribuição de acasalamentos dos Experimentos 1 e 2. Com relação ao tempo, o EVA obteve um tempo médio de execução inferior ao tempo médio apresentado pelo AG, 7,05 vezes menor, corroborando com o que foi visto no Experimento 2, que o EVA é mais rápido que o AG para seleções pequenas de animais.

Com relação às recomendações de acasalamentos, tanto o AG como o EVA incluíram acasalamentos com taxas de consanguinidade superiores à 3%. Neste Experimento 3, foi utilizado um número menor de animais, sendo um subconjunto dos animais utilizados no Experimento 2 e com uma maior taxa de consanguinidade média, restringindo as pos-

Figura 45 – Relação de acasalamentos e consanguinidades em uma das execuções do EVA no Experimento 2 do *Benchmark*.

Sire	Dam	N.mating	Nm	Nf.m	EBV	F
258151	987103	1	1	1	0.6180	0.00000
275378	209772	2	1	1	-1.020	0.00000
275378	370546	3	2	1	-0.5600E-01	0.00000
275378	944534	4	3	1	0.7650E-02	0.00000
418074	209754	5	1	1	-0.6696	0.00000
418074	944552	6	2	1	0.8930E-01	0.00000
944596	275232	7	1	1	0.4778	0.00000
944596	370511	8	2	1	0.5494	0.00000
974707	370545	9	1	1	-0.4905E-01	0.00000
974707	849802	10	2	1	0.2179	0.00000
974707	933124	11	3	1	0.2460E-01	0.00000
986486	209767	12	1	1	1.554	0.00000
986486	370566	13	2	1	1.576	0.00000
986486	944558	14	3	1	0.5040	0.00000
986773	944574	15	1	1	0.2100E-02	0.00000
986773	987104	16	2	1	-0.8650E-02	0.00000

Fonte: Autor(2021)

sibilidades de recomendação pelas duas ferramentas e fazendo com que algum acasalamento consanguíneo tivesse que ser incluído para completar a listagem de acasalamentos. A Figura 46 apresenta a relação de acasalamentos dada como resposta pelo AG em uma de suas execuções no Experimento 3.

Figura 46 – Relação de acasalamentos e consanguinidades em uma das execuções do AG no Experimento 3 do *Benchmark*.

touro_id	vaca_id	consanguinidade
258151	209767	0.000000000
440199	370545	0.000240326
440199	370546	0.000000000
440199	370566	0.125000000
258151	944534	0.000000000
258151	944574	0.000000000

Fonte: Autor(2021)

A Figura 47 apresenta a relação de acasalamentos dada como resposta pelo EVA em uma de suas execuções no Experimento 3.

Como podemos ver na Figura 46, o AG incluiu um acasalamento com 12,5% de

Figura 47 – Relação de acasalamentos e consanguinidades em uma das execuções do EVA no Experimento 3 do *Benchmark*.

Sire	Dam	N.mating	Nm	Nf.m	EBV	F	vF
258151	209767	1	1	1	1.291	0.00000	0.50000
418074	944574	2	1	1	0.3375	0.06250	0.50000
944596	370545	3	1	1	0.8079	0.00000	0.50000
944596	370546	4	2	1	1.210	0.00000	0.50000
986773	370566	5	1	1	0.3517	0.00000	0.50000
986773	944534	6	2	1	0.2813	0.00000	0.50000

Fonte: Autor(2021)

consanguinidade na relação de acasalamentos dada como resposta. Esse acasalamento utilizou o touro de identificador 440199, que foi o touro que apresentou as maiores taxas de consanguinidade quando relacionado com as vacas disponíveis. Este animal também possuía um dos maiores valores de IQG dentre os demais, e foi isso que o AG priorizou para recomendar a sua utilização.

O EVA por sua vez, incluiu um acasalamento de 6,25%, como mostrado na Figura 47, e não incluiu o touro de identificador 440199. Diferentemente do AG, o EVA priorizou utilizar animais com taxas de consanguinidade menores. Com um conjunto de animais mais restrito, esses padrões de recomendação para AG e para o EVA se repetiram em todas as execuções no Experimento 3.

Através dos resultados dos três experimentos, pode-se observar que o EVA apresentou um melhor controle de consanguinidade na recomendação de acasalamentos. Porém, o AG utilizou os melhores animais e obteve recomendações com melhor valor de IQG médio. Também nota-se que o AG apresenta melhores valores de tempos de execução à medida que o número de animais selecionados aumenta, em contraste com o EVA que é melhor com pequenas seleções.

5 CONCLUSÕES E TRABALHOS FUTUROS

Para atingir o objetivo deste trabalho e realizar as implementações de melhorias no AG proposto por Fontoura, Camargo and Cardoso (2019), inicialmente foi realizada uma revisão na literatura sobre os critérios de parada em algoritmos genéticos. A partir disso, duas abordagens de parada antecipada foram implementadas, sendo criados casos de teste para validação e comparação, cada um com um conjunto de animais. Ambas as abordagens utilizaram uma função para realizar o cálculo da variância do valor de *fitness* na população de cromossomos do AG, porém com algumas diferenças.

A abordagem de Yeng, Yoke and Suhaimi (2019) observa quando toda a população de cromossomos atinge a convergência para a resposta ao problema, interrompendo a execução do AG quando a variância do valor de *fitness* na população atinge um valor de limiar próximo a zero. Como o limiar depende do problema a ser resolvido, a abordagem foi testada com valores de limiar diferentes, tendo seus resultados comparados com a abordagem original do AG. Os resultados mostraram que, mesmo apresentando ganhos de até 35,13% no tempo de execução, não foi possível definir um único valor de limiar para efetuar a parada do AG com diferentes casos de teste. O valor de limiar a ser escolhido mostrou-se dependente das características intrínsecas a cada seleção de animais.

A abordagem de Bhandari, Murthy and Pal (2012) observa a variância dos melhores valores de *fitness* obtidos ao longo das iterações e encerra a execução do AG se nenhuma melhoria adicional for observada após um intervalo fixo de iterações. Esta abordagem foi testada com diferentes valores de limiar combinados com diferentes valores de intervalo de iterações. Diferentemente da abordagem anterior, os resultados mostraram ser possível utilizar um único valor de limiar para diferentes casos de teste, alcançando ganhos de até 50,8% no tempo de execução do AG.

Quanto ao problema na restrição de utilização de touros apresentado pelo AG, após identificada a falha no código e realizada uma refatoração, foram realizados experimentos para validação. Os resultados mostraram que a recomendação de acasalamentos dada como resposta pelo algoritmo obedece a restrição de utilização.

As refatorações de código em rotinas e estruturas de dados realizadas eliminaram dependências com o PampaPlus. Com os resultados, ficou demonstrado que a refatoração de código possibilitou a utilização do AG em programas de melhoramento que observam outras características animais e índices genéticos diferentes. Além disso, implementações adicionais possibilitaram uma integração com interface *web*.

Já os experimentos realizados no *benchmark* do AG com EVA demonstraram algumas diferenças entre as duas ferramentas. O EVA apresentou um melhor controle de consanguinidade e o AG utilizou mais os melhores animais obtendo recomendações com maior valor de IQG médio.

A maior contribuição deste trabalho foi implementar e validar uma abordagem para realizar a parada antecipada de um AG utilizado no processo de melhoramento genético bovino. Foi mostrado que é possível obter ganhos de desempenho sem que ocorra perda no índice de qualificação genética da solução gerada pelo AG. De maneira geral, esta e as demais implementações adicionais realizadas no algoritmo contribuem para a melhoria do processo de seleção de animais para acasalamento no âmbito do PampaPlus, uma vez que serão utilizadas por este programa de melhoramento.

Algumas limitações deste trabalho podem ser citadas. O AG foi desenvolvido para ser utilizado em grandes seleções de animais, sendo que para uma pequena seleção, todo o processo realizado pelo algoritmo poderia consumir mais tempo que uma busca exaustiva por todas as respostas possíveis ao problema. Também, o AG pode ser utilizado em outros tipos de animais, desde que se possua o registro das características fenotípicas medidas e um índice a ser melhorado. Porém, neste trabalho, o algoritmo foi testado apenas com dados de bovinos de corte. Por fim, o funcionamento do AG é totalmente dependente dos dados das avaliações genéticas dos animais. Animais com ausência de dados de DEPs não podem ser utilizados pelo algoritmo.

5.1 Trabalhos Futuros

A seguir, são listadas algumas sugestões de trabalhos futuros:

- Investigação da relação de parâmetros como tamanho da população de cromossomos, taxa de mutação, valor médio de IQG e restrição de utilização na definição do limiar na abordagem de Yeng, Yoke and Suhaimi (2019);
- Refatoração da abordagem de penalização de cromossomos para melhorar o controle de consanguinidade nas recomendações de acasalamentos;
- Investigação da integração do AG com interface *web*.

REFERÊNCIAS

ABBASS, H.; TOWSEY, M.; FINN, G. Or and data mining for intelligent decision support in the australian dairy industry's breeding program. **Proceedings of New Research in OR**, ASOR, Brisbane, p. 1–23, 1998.

ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNES. **Beef Report Brazilian Livestock Profile**. 2020. Disponível em: <<http://www.brazilianbeef.org.br/download/sumarioingles2019.pdf>>. Acesso em: 30 mar. 2020.

AYTUG, H.; KOEHLER, G. J. New stopping criterion for genetic algorithms. **European Journal of Operational Research**, Elsevier, v. 126, n. 3, p. 662–674, 2000.

BANHAZI, T. M. et al. Precision livestock farming: an international review of scientific and commercial aspects. **International Journal of Agricultural and Biological Engineering**, v. 5, n. 3, p. 1–9, 2012.

BARBOSA, S.; SILVA, B. **Interação Humano-Computador**. Rio de Janeiro: Elsevier Brasil, 2010.

BERG, P.; NIELSEN, J.; SØRENSEN, M. K. Eva: realized and predicted optimal genetic contributions. In: **Proceedings of the 8th World Congress on Genetics Applied to Livestock Production**. Minas Gerais, Brazil: Instituto Prociência, 2006. p. 27–09.

BHANDARI, D.; MURTHY, C.; PAL, S. K. Variance as a stopping criterion for genetic algorithms with elitist model. **Fundamenta Informaticae**, v. 120, n. 2, p. 145–164, 2012.

BIAS, R. G.; MAYHEW, D. J. **Cost-justifying usability: An update for the Internet age**. San Francisco CA: Elsevier, 2005.

CAMPOS, M. M. et al. Tecnologias de precisão na avaliação da eficiência alimentar. **Cadernos Técnicos de Veterinária e Zootecnia**, n. 79, p. 73–85, 2015.

CARDOSO, F. F.; LOPA, T. M. B. P. **PampaPlus: Avaliação Genética Hereford e Braford**. Bagé: Embrapa Pecuária Sul, 2016.

CARVALHEIRO, R.; QUEIROZ, S. A. de.; KINGHORN, B. Optimum contribution selection using differential evolution. **Revista Brasileira de Zootecnia**, v. 39, n. 7, p. 1429–1436, 2010.

CARVALHO, P. C. de F. et al. Do bocado ao pastoreio de precisão: compreendendo a interface planta-animal para explorar a multi-funcionalidade das pastagens. **Revista brasileira de zootecnia**, v. 38, p. 109–122, 2009.

FONTOURA, D. de Carvalho Neves da. **Uma solução de recomendações de acasalamentos baseada em algoritmos genéticos**. 97 f. Dissertação (Mestrado em Computação Aplicada) — Universidade Federal do Pampa, Bagé, 2019.

- FONTOURA, D. N. da.; CAMARGO, S. da S.; CARDOSO, F. F. Optimizing mate selection: a genetic algorithms approach. **ICAR Technical Series**, Roma, v. 24, p. 55–62, 2019. Disponível em: <<https://www.icar.org/wp-content/uploads/2019/12/ICAR-Technical-Series-24-Prague-2019-Proceedings.pdf>>. Acesso em: 25 jul. 2020.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. How to feed the world 2050. In: **High-Level Expert Forum**. Roma, 2009.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. Safeguarding against economic slowdowns and downturns. **The State of Food and Nutritional Security in the World**, Roma, 2019.
- GAREY, M. R.; JOHNSON, D. S. **Computers and intractability**. San Francisco: Freeman, 1979.
- GEN, M.; LIN, L. Genetic algorithms. **Wiley Encyclopedia of Computer Science and Engineering**, Wiley Online Library, p. 1–15, 2007.
- HANSEN, E. A.; ZILBERSTEIN, S. Heuristic search in cyclic and/or graphs. **Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)**, Madison, Wisconsin, n. 15, p. 412–418, 1998.
- HANSEN, E. A.; ZILBERSTEIN, S. LAO*: a heuristic search algorithm that finds solutions with loops. **Artificial Intelligence**, v. 129, n. 1-2, p. 35–62, 2001.
- HERRERO, M.; THORNTON, P. K. Livestock and global change: emerging issues for sustainable food systems. **Proceedings of the National Academy of Sciences**, National Academy of Sciences of the United States of America, v. 110, n. 52, p. 20878–20881, 2013.
- HEWETT, T. T. et al. **ACM SIGCHI curricula for human-computer interaction**. Nova York: Association for Computing Machinery, 1992.
- KINGHORN, B. P. On computing strategies for mate allocation. **Journal of Animal Breeding and Genetics**, Wiley Online Library, v. 104, n. 1-5, p. 12–22, 1987.
- KINGHORN, B. P. An algorithm for efficient constrained mate selection. **Genetics Selection Evolution**, BioMed Central, v. 43, n. 1, p. 4, 2011.
- LINDEN, R. **Algoritmos genéticos**. 3. ed. Rio de Janeiro: Editora Ciência Moderna Ltda, 2012.
- LIU, Y.; ZHOU, A.; ZHANG, H. Termination detection strategies in evolutionary algorithms: a survey. In: **Proceedings of the Genetic and Evolutionary Computation Conference**. Kyoto: Association for Computing Machinery, 2018. p. 1063–1070.
- LOBO, R. N. B.; LÔBO, A. M. B. O. Melhoramento genético como ferramenta para o crescimento e o desenvolvimento da ovinocultura de corte. **Revista Brasileira de Reprodução Animal**, v. 31, n. 2, p. 247–253, 2007.
- NETO, A. D. B. **Estrutura populacional e otimização de esquemas de acasalamento em ovinos com uso de algoritmos evolucionários**. 62 f. Dissertação (Mestrado em Zootecnia) — Universidade Federal de Sergipe, São Cristóvão, 2014.

NOBRE, P. R. C. et al. Programa embrapa de melhoramento de gado de corte-geneplus. **Melhoramento genético aplicado em gado de corte: Programa Geneplus-Embrapa**, In: ROSA, A. do N.; MARTINS, E. N.; MENEZES, G. R. de O.; SILVA, L. O. C. da (Ed.), Campo Grande, p. 235–241, 2013.

NORMAN, D. **The psychology of everyday things**. New York: Basic Books, 1988.

NORMAN, D. **The design of everyday things: revised and expanded edition**. New York: Basic books, 2013.

PENDHARKAR, P. C.; KOEHLER, G. J. A general steady state distribution based stopping criteria for finite length genetic algorithms. **European Journal of Operational Research**, v. 176, n. 3, p. 1436–1451, 2007.

PIRES, M. P. **Planejamento e implementação de um programa de melhoramento genético de ovinos no oeste paulista**. 124 f. Dissertação (Mestrado em Genética e Melhoramento Animal (Avaliação Genética)) — Universidade Estadual Paulista (UNESP), Jaboticabal, 2011.

POOLE, D. L.; MACKWORTH, A. K. **Artificial Intelligence: foundations of computational agents. Second Edition**. Cambridge: Cambridge University Press, 2017.

PREECE, J.; ROGERS, Y.; SHARP, H. **Design de interação: além da interação homem-computador**. Porto Alegre: Bookman, 2005.

RIBAS, M. et al. Pecuária de precisão: Uso de tecnologias para apoio à tomada de decisão. In: **XXVII Congresso Brasileiro de Zootecnia**. Santos: Associação Brasileira de Zootecistas, 2017.

ROSA, A. do N.; MENEZES, G. R. de O.; EGITO, A. A. do. Recursos genéticos e estratégias de melhoramento. **Melhoramento genético aplicado em gado de corte: Programa Geneplus-Embrapa**, In: ROSA, A. do N.; MARTINS, E. N.; MENEZES, G. R. de O.; SILVA, L. O. C. da (Ed.), Campo Grande, p. 11–26, 2013.

ROSA, A. do N.; MENEZES, G. R. de O.; SILVA, L. O. C. da. **Melhoramento genético aplicado em gado de corte: Programa Geneplus-Embrapa**. Campo Grande: Embrapa Gado de Corte, 2013.

SILVA, E. L. d.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. Florianópolis: Laboratório de Ensino a Distância da UFSC, 2001.

SOMMERVILLE, I. **Engenharia de Software**. São Paulo: Pearson Prentice Hall, 2011.

STORN, R.; PRICE, K. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. **Journal of global optimization**, v. 11, n. 4, p. 341–359, 1997.

SUN, Y. et al. An adaptive framework for large-scale state space search. In: **2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum**. Anchorage: IEEE, 2011. p. 1798–1805.

TOĞAN, V.; DALOĞLU, A. T. An improved genetic algorithm with initial population strategy and self-adaptive member grouping. **Computers & Structures**, v. 86, n. 11-12, p. 1204–1218, 2008.

TULLO, E.; FINZI, A.; GUARINO, M. Environmental impact of livestock farming and precision livestock farming as a mitigation strategy. **Science of the total environment**, v. 650, p. 2751–2760, 2019.

TULLO, E. et al. Application note: Labelling, a methodology to develop reliable algorithm in plf. **Computers and Electronics in Agriculture**, v. 142, p. 424–428, 2017.

UNITED NATIONS DEPARTMENT OF ECONOMICS AND SOCIAL AFFAIRS.
World population prospects: The 2012 revision. New York, v. 18, 2013.

UNITED NATIONS DEPARTMENT OF ECONOMICS AND SOCIAL AFFAIRS.
World urbanization prospects: The 2014 revision. New York, v. 41, 2015.

WATHES, C. M. et al. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? **Computers and electronics in agriculture**, v. 64, n. 1, p. 2–10, 2008.

YENG, F. F.; YOKE, S. K.; SUHAIMI, A. The saturation of population fitness as a stopping criterion in genetic algorithm. **International Journal of Electrical and Computer Engineering**, v. 9, n. 5, p. 4130, 2019.

APÊNDICE A — INTERFACE DA FERRAMENTA

Inicialmente, este trabalho contava com dois problemas de pesquisa: implementar melhorias no AG e desenvolver uma interface para que o algoritmo pudesse ser utilizado no portal *web* do PampaPlus. As melhorias foram implementadas no AG, como mostrado no capítulo de resultados. Para o desenvolvimento da interface, foram levantados requisitos junto aos pesquisadores e técnicos da Embrapa, e alguns protótipos de interface foram desenvolvidos. Porém, os esforços se concentraram nas melhorias no AG e a implementação da interface não foi completamente efetuada.

Este apêndice apresenta o material que foi desenvolvido até o término deste trabalho. A seção A.1 traz uma revisão de literatura sobre Interface Homem-Computador e princípios de usabilidade de interface são comentados. A seção A.2 apresenta os protótipos que foram desenvolvidos com base nos requisitos elicitados e nos princípios de interface revisados. Na seção A.3, é mostrado o Documento de Requisitos e por fim, a seção A.4.2 apresenta o documento de estratégia de testes.

A.1 Interface Homem-Computador

No trabalho de Fontoura (2019), o autor comenta que a utilização do AG proposto através de uma interface intuitiva para os usuários proporcionaria resultados mais satisfatórios na seleção de animais para acasalamento. Tendo isso em vista, uma busca bibliográfica sobre Interface Homem-Computador (IHC), também chamada de Interação Humano-Computador, foi realizada.

IHC é uma área de estudo relacionada ao desenvolvimento de sistemas de computação interativos com seres humanos e está presente nas fases de projeto, implementação, avaliação e utilização desses sistemas (HEWETT et al., 1992).

Na utilização dos sistemas interativos pelas pessoas, por exemplo, a IHC está interessada no estudo dos fenômenos presentes. Para isso, sendo uma área de estudo multidisciplinar, utiliza não apenas o conhecimento da computação mas também de outras áreas. Os conhecimentos e técnicas de design, ergonomia, linguística e semiótica são utilizados no projeto de interface de usuário. Já a psicologia, a sociologia e a antropologia permitem obter o conhecimento sobre a cultura, os discursos e os comportamentos dos usuários (BARBOSA; SILVA, 2010). Os aspectos associados à IHC são: natureza da interação humano-computador, contexto de uso do computador, características humanas,

arquitetura de sistema e de interface e, processo de desenvolvimento (HEWETT et al., 1992).

A *natureza da interação* diz respeito à utilização de sistemas interativos pelas pessoas em diversas atividades e as suas consequências. A interação com esses sistemas é influenciada pelo *contexto de uso*, pois as pessoas estão inseridas em diferentes culturas e organizações. O processamento de informações e a aprendizagem de utilização em um sistema novo, por exemplo, são *características humanas* que também interferem na interação. A interação das pessoas com os sistemas depende da *arquitetura de sistema e de interface*, pois diferentes dispositivos de entrada e saída e técnicas de diálogo com o usuário podem ser utilizadas. Já o *processo de desenvolvimento* influencia a qualidade do produto final. O conhecimento de métodos, técnicas e ferramentas de construção de interface com usuário e de avaliação é importante, bem como casos de sucesso e de insucesso de interfaces com usuário e seus motivos (BARBOSA; SILVA, 2010).

De acordo com Bias and Mayhew (2005), a IHC ajuda a reduzir erros na operação dos sistemas, custos com desenvolvimento, treinamento de usuários e suporte técnico e, aumenta a produtividade dos usuários e a fidelidade do cliente. Portanto, o estudo e a aplicação dos princípios de IHC são importantes na construção de sistemas interativos com os usuários. Nesse sentido, os princípios de interfaces utilizados em IHC foram revisados e são apresentados a seguir.

A.1.1 Princípios de Interfaces

De acordo com Norman (2013), o princípio mais importante em um sistema é o seu modelo conceitual, que deve fornecer uma compreensão do sistema como um todo. Um modelo conceitual é uma explicação, geralmente simplificada, de como algo funciona.

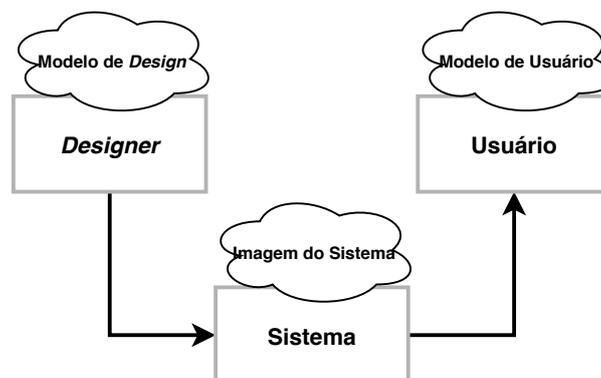
Um exemplo dado por Norman (2013) é o sistema de arquivos do computador, exibido na tela como pastas e ícones. Esse modelo cria o conceito de documentos e pastas e de aplicativos que podem ser clicados na tela. As pastas não existem literalmente nos computadores, porém esse conceito é eficaz em facilitar a organização e utilização dos arquivos por parte dos usuários. Segundo o autor, um bom modelo conceitual permite prever os efeitos das ações das pessoas. Se o modelo não for bom, as pessoas fazem operações sem saber o que acontece nos passos seguintes e, se algo der errado, é difícil saber o motivo, os efeitos dos erros e as ações a serem tomadas. Por isso, quando

uma situação nova é encontrada, torna-se necessário um entendimento mais profundo do funcionamento, fornecido através de um bom modelo conceitual.

A Figura 48 apresenta o modelo conceitual no contexto de um sistema computacional. O modelo de *design* é o modelo conceitual criado pelo *designer*. O modelo de usuário é o modelo que o usuário desenvolve através da sua interação com o sistema. Já a imagem do sistema resulta do sistema em si e sua estrutura física. Ao projetar o sistema, o *designer* espera que o modelo conceitual do usuário seja idêntico ao modelo por ele projetado. Porém, a comunicação não acontece de forma direta entre *designer* e usuário, ela é realizada através da imagem do sistema. Se a imagem do sistema não apresentar o modelo de *design* de forma clara e consistente, o usuário terá um modelo conceitual diferente daquele projetado (NORMAN, 2013).

Baseados no trabalho de Norman (1988), Preece, Rogers and Sharp (2005) destacaram seis princípios de *design*, sendo eles: Visibilidade, *Feedback*, Restrições, Mapeamento, Consistência e *Affordance*. Esses princípios, descritos a seguir, são utilizados no desenvolvimento de sistemas interativos para que o usuário compreenda e utilize o sistema da forma em que foi projetado.

Figura 48 – Modelo conceitual no contexto de um sistema computacional.



Fonte: Adaptado de (NORMAN, 2013)

Visibilidade: Quando as funções de um sistema são visíveis, os usuários conseguem utilizá-las com mais facilidade. Por outro lado, se a visibilidade é afetada, fica mais difícil encontrar as funções e operar o sistema (PREECE; ROGERS; SHARP, 2005). De acordo com um conceito descrito por Norman (1988), a relação entre a maneira em que os elementos são posicionados e o que eles fazem facilita para o usuário encontrar o elemento apropriado para a tarefa em questão.

Feedback: Relacionado com o conceito de visibilidade, o *feedback* diz respeito ao retorno de informações sobre as ações dos usuários em um sistema. As ações realizadas em um sistema devem produzir efeitos para os usuários, sendo que vários tipos de feedback estão disponíveis para *design* de interação como áudio, tátil, verbal, visual e combinações destes (PREECE; ROGERS; SHARP, 2005).

Restrições: Relacionado com maneiras de restringir o tipo de interação do usuário em um determinado momento. Uma prática comum de *design* em interfaces gráficas de usuário é desativar certas opções de menu, sombreando-as, restringindo assim o usuário a apenas ações permitidas naquele estágio da atividade. Uma das vantagens dessa forma de restrição é que ela impede o usuário de selecionar opções incorretas e, assim, reduz a chance de cometer um erro (PREECE; ROGERS; SHARP, 2005).

Mapeamento: Refere-se à relação entre os controles e os seus efeitos. Na reprodução de mídia, por exemplo, os controles devem ser projetados de modo a produzir os efeitos desejados. A configuração deve ser mapeada diretamente na direcionalidade das ações. Uma convenção comum é fornecer uma sequência de botões com o botão *play* ao centro, o botão retroceder à esquerda e o avanço rápido à direita (PREECE; ROGERS; SHARP, 2005).

Consistência: Envolve o *design* de interfaces que ofereçam operações semelhantes ou que utilizem elementos semelhantes na realização de tarefas similares. Interfaces consistentes seguem regras, como utilizar sempre a mesma operação para determinada tarefa, tornando aquela tarefa mais fácil de aprender e usar. Por outro lado, interfaces inconsistentes permitem exceções às regras e trazem dificuldades aos usuários, fazendo com que cometam mais erros (PREECE; ROGERS; SHARP, 2005).

Affordance: Um *affordance* é uma indicação de como operar um objeto sem necessidade de ter instruções ou imagens ilustrativas pois, ao olhar, o usuário já sabe o que fazer (NORMAN, 2013). Elementos gráficos de interface como botões, ícones e barras de rolagem são discutidos e projetados de modo que sua utilização pareça óbvia. Botões são projetados para permitir que sejam pressionados, ícones para serem clicados e barras de rolagem para mover-se para cima e para baixo (PREECE; ROGERS; SHARP, 2005).

A.1.2 Ferramenta *Web* para Seleção de Acasalamentos

O trabalho de Fontoura (2019) apresentou uma solução baseada em AGs que analisa as DEPs, resultantes das medidas objetivas e de escores visuais dos animais, no pro-

cesso de melhoramento genético bovino para recomendar acasalamentos de animais aos produtores. Os resultados mostraram que o AG proposto recomenda uma configuração de acasalamentos que maximiza a média do IQG entre os animais selecionados, respeitando as restrições que são impostas.

O AG foi desenvolvido para ser utilizado no PampaPlus. Através do portal do programa, é disponibilizada uma ferramenta onde o produtor cadastrado pode simular acasalamentos. Esta ferramenta encontra-se em fase de testes, mostrando apenas aos produtores uma matriz de resultados com os possíveis produtos gerados, os filhos, observando todos os acasalamentos possíveis entre todos os pais.

Porém, quando o usuário seleciona um número grande de animais, a visualização dos resultados fica prejudicada. Além disso, esta ferramenta não provê como resultado uma configuração de acasalamentos que maximiza a média do IQG entre os animais selecionados como no AG proposto. Por isso, uma sugestão de trabalho futuro de Fontoura (2019) é integrar o AG na ferramenta de acasalamento do sistema PampaPlusnet e desenvolver uma interface para dar apoio à utilização por parte dos usuários.

Segundo o autor, a utilização do AG através de uma interface intuitiva para o usuário facilitaria a utilização da ferramenta e traria resultados mais satisfatórios no processo de seleção de animais. Através da interface o usuário poderia selecionar os animais, escolher as ponderações da sua função objetivo, editando o IQG, e visualizar graficamente de forma simplificada os resultados gerados pelo AG. Portanto, na busca pela melhoria da ferramenta já existente, além da integração com o AG proposto, seria fundamental utilizar os conceitos de Interação Humano-Computador e seus princípios para melhorar a experiência dos usuários.

A.2 Protótipos de Interface

Com base nos requisitos de funcionalidade definidos para a ferramenta e nos princípios de IHC vistos na seção A.1.1, foram elaborados protótipos de interface utilizando o software Pencil. Estes protótipos apoiaram o processo de especificação e validação dos requisitos e darão suporte à implementação, ainda não realizada.

A interface da ferramenta de acasalamento foi inicialmente projetada para ter quatro ambientes, ou telas, que serão utilizadas em sequência lógica pelo usuário. A primeira tela irá prover ao usuário as funcionalidades para escolha dos touros para acasalamento. A segunda, muito parecida com a primeira, proverá a escolha das vacas. Já a terceira tela

permitirá ao usuário informar os parâmetros que serão utilizados pelo AG para executar. Por fim, a quarta tela informará ao usuário o resultado da execução do AG, que é a recomendação de acasalamentos entre os animais. A seguir, os protótipos elaborados são apresentados. As Figuras 49 e 50 apresentam o protótipo da tela de escolha dos touros.

Figura 49 – Protótipo da tela de escolha dos touros.

USUARIO X - Logout

TOUROS VACAS PARÂMETROS RECOMENDAÇÃO

TATUAGEM/ REG/ NOME

Pesquisa Selecionados (0)

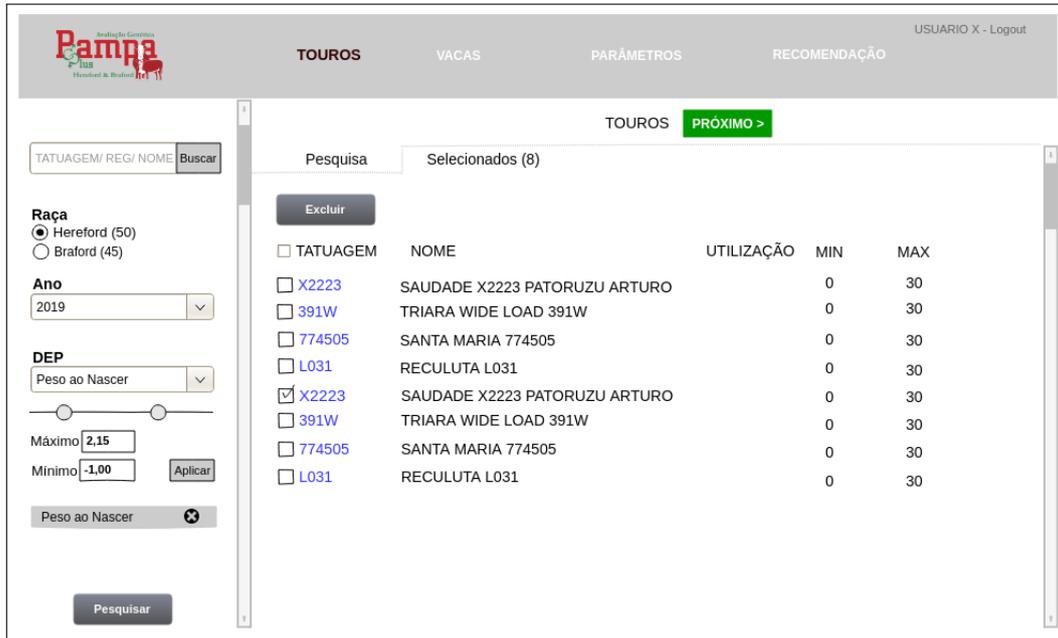
TATUAGEM NOME UTILIZAÇÃO MIN MAX

<input checked="" type="checkbox"/>	X2223	SAUDADE X2223 PATORUZU ARTURO	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	391W	TRIARA WIDE LOAD 391W	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	774505	SANTA MARIA 774505	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	L031	RECULUTA L031	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	X2223	SAUDADE X2223 PATORUZU ARTURO	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	391W	TRIARA WIDE LOAD 391W	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	774505	SANTA MARIA 774505	<input type="text" value="0"/>	<input type="text" value="30"/>
<input checked="" type="checkbox"/>	L031	RECULUTA L031	<input type="text" value="0"/>	<input type="text" value="30"/>
<input type="checkbox"/>	X2223	SAUDADE X2223 PATORUZU ARTURO	<input type="text" value="0"/>	<input type="text" value="30"/>
<input type="checkbox"/>	391W	TRIARA WIDE LOAD 391W	<input type="text" value="0"/>	<input type="text" value="30"/>
<input type="checkbox"/>	774505	SANTA MARIA 774505	<input type="text" value="0"/>	<input type="text" value="30"/>
<input type="checkbox"/>	L031	RECULUTA L031	<input type="text" value="0"/>	<input type="text" value="30"/>

Fonte: Autor(2021)

Como pode ser observado, na coluna da esquerda da tela existem duas maneiras de selecionar os animais. Na primeira, o usuário pode realizar uma busca por um touro específico a partir de sua tatuagem ou nome e confirmar com o botão "Buscar". Na segunda, mais abaixo na coluna, o usuário pode aplicar filtros para escolher por raça, por ano de nascimento e por DEP. Nos filtros de DEP, o usuário pode escolher quantas DEPs desejar e definir os valores de máximo e de mínimo para aquelas características. Nesta segunda maneira de seleção, o usuário deve efetuar a confirmação no botão "Pesquisar". Os animais que correspondem aos parâmetros de busca são listados na coluna da direita da tela na aba "Pesquisa". Nesta aba, o usuário pode então selecionar os touros e informar seus parâmetros de utilização mínima e máxima, confirmando a seleção com o botão "Selecionar". Na aba "Selecionados", que pode ser visualizada na Figura 50, o usuário pode, além de visualizar os animais já selecionados, marcar e excluir uma seleção de animal com o botão "Excluir". Após isso, o clique no botão "Próximo", localizado acima da coluna, conduz para a próxima tela. A Figura 51 apresenta o protótipo da tela de

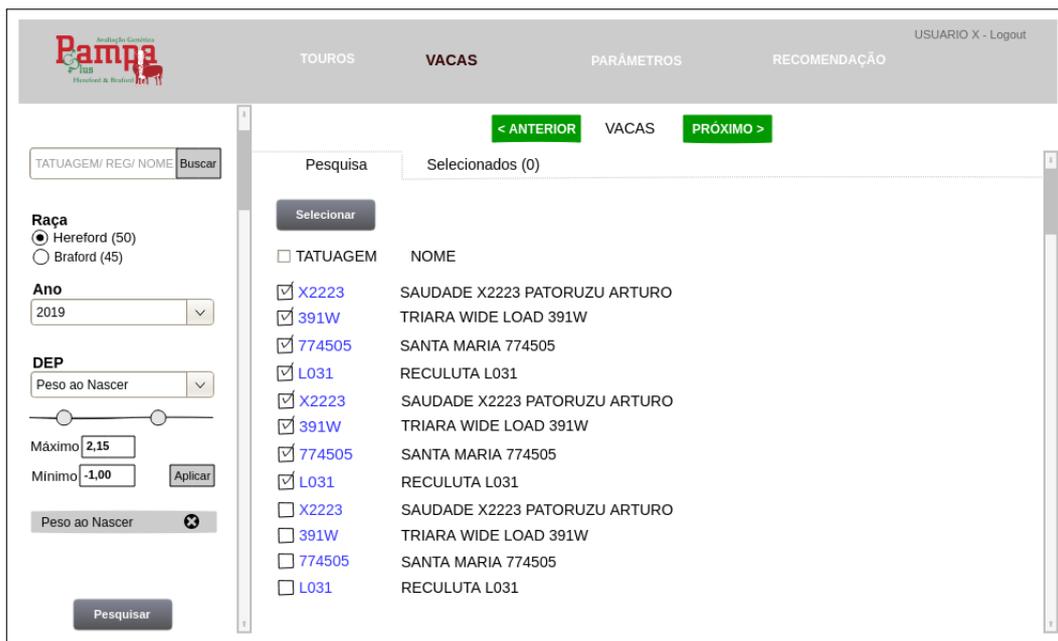
Figura 50 – Aba de selecionados do protótipo da tela de escolha dos touros.



Fonte: Autor(2021)

escolha das vacas.

Figura 51 – Protótipo da tela de escolha das vacas.



Fonte: Autor(2021)

Como pode ser observado, a tela de escolha das vacas é similar a tela de escolha de touros, com a excessão de que o mínimo e o máximo de utilização não é definido para as vacas na aba "Pesquisa", já que o parâmetro de utilização é aplicado apenas para os touros. Ainda, caso o usuário necessite voltar à tela anterior, o botão "Anterior", localizado acima da segunda coluna, permite isso.

A seguir, a Figura 52 apresenta o protótipo da tela de escolha dos parâmetros a serem utilizados pelo AG. Pode-se observar que este protótipo de tela permite ao usuário editar os seguintes parâmetros: a função de avaliação a ser utilizada pelo AG, que é a ponderação entre o IQG e o nível de problema; o IQG com suas ponderações de DEPs; o nível de problema e; a taxa de consanguinidade máxima. Para realizar a edição, o usuário deverá clicar em um dos parâmetros localizados na coluna esquerda da tela para então acessar a sua edição na coluna direita da tela. Como exemplo, é apresentada a função de edição do IQG. Porém, se o usuário não desejar editar nenhum parâmetro, o AG utilizará uma configuração padrão definida nos requisitos do sistema.

Figura 52 – Protótipo da tela de escolha dos parâmetros do AG.

Parâmetro	Valor
DEP	%
Peso Total Materno à Desmama	30
Ganho de Peso Pós Desmama	15
Peso ao Sobreano	15
Escore de Musculatura	12.5
Escore de Estatura	12.5

Fonte: Autor(2021)

Com os parâmetros definidos, o usuário poderá dar início a execução do AG a partir do clique no botão "Simular", localizado acima da segunda coluna. Com isso, as funcionalidades da tela são restringidas e o andamento da execução do AG é informado ao usuário através de uma barra de progresso, como visualizado na Figura 53. Se o

usuário desejar interromper a execução do AG, o botão "Parar" permite isso sem que algum resultado seja gerado.

Figura 53 – Protótipo de tela de andamento da execução do AG.



Fonte: Autor(2021)

O  ltimo prot tipo desenvolvido foi o da tela de recomenda o de acasalamentos, visualizado na Figura 54. Nesta tela, o resultado gerado pelo AG   apresentado ao usu rio. Em consist ncia com a tela anterior de edi o de par metros, esta tela possui as op es na coluna da esquerda e a visualiza o na coluna da direita. No exemplo, podem ser visualizados os pares de acasalamento recomendados com as vacas e seus respectivos touros. Esta tela tamb m permite ao usu rio ver os poss veis produtos gerados a partir dos acasalamentos recomendados e suas estat sticas. Al m disso, documentos podem ser gerados conforme a necessidade do usu rio. Por fim, se o usu rio necessitar, o processo pode ser realizado novamente utilizando o bot o "Anterior" para voltar  s telas anteriores de sele o de animais e edi o de par metros.

A partir da observ ncia do princ pio de *Visibilidade* de IHC, foi poss vel desenvolver uma interface mais minimalista poss vel, a partir de um fundo de tela limpo com bot es de cores sugestivas. Com funcionalidades claras e vis veis, buscou-se facilitar a identifica o e aprendizado destas por parte do usu rio. O princ pio de *Consist ncia* foi utilizado para manter a orienta o das funcionalidades ao usu rio. Todas as telas possuem duas colunas, com op es principais   esquerda e visualiza o   direita com op es

Figura 54 – Protótipo de tela de recomendação de acasalamentos.



Fonte: Autor(2021)

secundárias.

Os botões "Próximo" e "Anterior" fornecem o princípio de *Mapeamento*, facilitando a navegação do usuário pelas telas. Além disso, os títulos que aparecem em destaque na parte superior de todas as telas fornecem a noção de tela ou fase da escolha em que o usuário se encontra. Os princípios de *Restrição* e *Feedback* são utilizados na tela de andamento da execução do AG. Enquanto o AG executa, as funcionalidades são restritas com exceção do botão de "Parar", sendo que uma barra de progresso disponibiliza o *feedback* necessário ao usuário.

O princípio de *feedback* também é utilizado na barra de navegação, localizada no topo de todas as telas, indicando em qual tela ou fase do processo da ferramenta o usuário se encontra. Além disso, conforme as DEPs são selecionadas como filtros na seleção de touros e vacas, elas são colocadas na tela para mostrar a seleção ao usuário, onde também é permitindo a exclusão desse filtro. Também, após a definição dos filtros e o clique no botão "Pesquisar", os animais são listados na coluna da direita da tela, onde o usuário pode clicar na tatuagem do animal listado, um *link* na cor azul, e visualizar uma ficha detalhada com seus valores de DEP.

Por fim, o princípio de *Affordance* é visto nos elementos clássicos de interface utilizados, já amplamente conhecidos e de fácil usabilidade como botões simples, *Radio*

Buttons, Scales, Combo Boxes e ScrollBars.

A.3 Documento de Requisitos

A.3.1 Introdução

Este documento descreve os requisitos especificados em conjunto com pesquisadores e técnicos da EMBRAPA para o sistema de seleção de acasalamentos. Inicialmente é realizada uma descrição geral do sistema. Após isso, os usuários responsáveis pela definição dos requisitos são apresentados. Os requisitos funcionais (RF), os requisitos não-funcionais (RNF) e as restrições do sistema (RST) são descritos logo a seguir. Por fim, as referências para a estruturação deste documento são apresentadas.

A.3.1.1 Descrição Geral do Sistema

O sistema fornecerá um ambiente *web* de seleção de acasalamentos apoiado por AG para os produtores associados ao programa de melhoramento genético PampaPlus. Através da interface, o usuário poderá pesquisar os animais pertencentes ao seu rebanho, selecionar e descartar algum animal selecionado e definir parâmetros a serem utilizados pelo AG. O AG realizará uma simulação com os animais e parâmetros selecionados e fornecerá uma recomendação de acasalamentos entre os animais. O sistema permitirá a visualização dos resultados, dos produtos gerados, de estatísticas e exportação de relatórios.

A.3.1.2 Usuários de Requisitos de Sistema

Os requisitos presentes neste documento foram definidos a partir de uma reunião com pesquisadores e técnicos do Laboratório de Bioinformática e Estatística Genômica (LABEGEN) da EMBRAPA Pecuária Sul. Os usuários do sistema que foram responsáveis pela de definição de todos os requisitos são os seguintes: Fernando Flores Cardoso, Pesquisador e Henry Gomes de Carvalho, Analista.

A.3.2 Requisitos Funcionais

- **RF01: Pesquisar animais através de função de busca.**

O sistema deve permitir ao usuário a pesquisa de touros (ou sêmens) e vacas por tatuagem, registro ou nome.
- **RF02: Pesquisar animais através de aplicação de filtros.**

O sistema deve permitir ao usuário a pesquisa filtrando touros (ou sêmens) e vacas por raça, ano de nascimento e DEPs.
- **RF03: Selecionar animais ou excluir a seleção.**

O sistema deve permitir ao usuário selecionar touros (ou sêmens) e vacas resultantes da pesquisa e também excluir a seleção.
- **RF04: Definir parâmetros de utilização de touros.**

O sistema deve permitir ao usuário definir a utilização mínima e máxima para cada touro (ou sêmem). Para ambos os campos, de utilização mínima e máxima, o sistema deve aceitar valores de 1 (um) até o número de vacas selecionadas. O valor 0 (zero) é utilizado para a não utilização destes parâmetros. Caso o usuário não defina estes parâmetros, o sistema deve utilizar automaticamente 0 (zero) para os dois parâmetros.
- **RF05: Definir a ponderação da função de avaliação do AG.**

O sistema deve permitir ao usuário definir a ponderação entre IQG e Nível de problema para a função de avaliação.
- **RF06: Editar o IQG.**

O sistema deve permitir ao usuário editar a ponderação dos parâmetros de DEP do IQG. Após a edição, o IQG deve ser observado pela função de avaliação do AG.
- **RF07: Informar Nível de Problema.**

O sistema deve permitir ao usuário informar um valor para o Nível de Problema e selecionar DEPs às quais deseja aplicar este valor.
- **RF08: Definir taxa de consanguinidade.**

O sistema deve permitir ao usuário definir uma taxa de 0% a 100% de consanguinidade a ser observada pelo AG. Caso o usuário não defina a taxa, o sistema deve utilizar automaticamente um valor de 3%.

- **RF09: Valores padrão para os parâmetros do AG.**

No caso em que o usuário não informe valores de parâmetros para o AG, devem ser utilizados automaticamente valores padrão nas simulações.

- **RF10: Ação do sistema.**

Deve ser implementada a ação do sistema quando a escolha dos animais resultarem em acasalamentos de consanguinidade superior ao limite definido. Perguntar ao usuário se deseja prosseguir mesmo assim ou realizar a escolha de outros animais.

- **RF11: Iniciar simulação com AG.**

O sistema deve permitir ao usuário iniciar a simulação do AG com os dados selecionados.

- **RF12: Visualizar resultado do AG.**

O sistema deve permitir ao usuário visualizar a recomendação de acasalamentos fornecida pelo AG.

- **RF13: Visualizar potenciais produtos gerados.**

O sistema deve permitir ao usuário visualizar os potenciais produtos gerados pela recomendação de acasalamentos fornecida pelo AG.

- **RF14: Visualizar estatísticas.**

O sistema deve permitir ao usuário visualizar medidas de tendência central e de dispersão do IQG e do Nível de Problema para os produtos.

- **RF15: Exportar relatórios.**

O sistema deve permitir ao usuário exportar relatórios do resultado do AG e dos produtos.

A.3.3 Requisitos Não-Funcionais

- **RNF01: Segurança de Acesso.**

O usuário deverá estar logado no sistema PampaPlusNet para acesso à ferramenta de seleção de acasalamentos.

- **RNF02: Usabilidade de interface.**

O sistema deve possuir interface intuitiva e amigável, de acordo com princípios de IHC validados a partir de teste de usabilidade.

- **RNF03: Critério de parada do AG.**

O AG deve possuir critério de parada com base na função de avaliação. Atualmente, o critério de parada é baseado em uma quantidade pré-definida de gerações.

- **RNF04: Implementação do AG.**

As implementações adicionais no AG deverão ser desenvolvidas na linguagem R.

A.3.4 Restrições

- **RST01: Utilização de touros (ou sêmens).**

O sistema deve recomendar acasalamentos respeitando os limites de utilização máxima e mínima de touros.

- **RST02: Consanguinidade.**

O sistema deve recomendar acasalamentos de acordo com a taxa de consanguinidade definida pelo usuário.

A.3.5 Referencial

Este documento foi baseado nas recomendações do documento IEEE-Std 830-1998 e no modelo apresentado por Sommerville (2011).

A.4 Plano de Testes

A.4.1 Introdução

Este documento descreve o plano de testes do sistema de seleção de acasalamentos. Inicialmente é apresentado o documento de referência para elaboração deste documento. Em seguida, a estratégia de testes a ser utilizada é abordada. Após isso, são apresentados os objetivos e prioridades do sistema e então os casos de testes das funcionalidades do sistema são mostrados. Por fim, a referência para estruturação deste documento é apresentada.

A.4.1.1 Documento de Referência

O documento de referência para elaboração dos casos de teste contidos neste documento é o Documento de Requisitos localizado no *Apêndice A.3*. No Documento de Requisitos encontra-se todos os requisitos elicitados e também uma descrição geral do sistema e seu propósito.

A.4.2 Estratégia de Testes

Os tipos de teste aplicados na verificação e validação de requisitos do sistema serão os seguintes: Teste Unitário, Teste de Integração, Teste de Sistema e Teste de Aceitação. Será aplicada a técnica de regressão, sendo utilizados os casos de teste descritos na seção A.4.4 deste documento.

Os testes serão aplicados de acordo com o progresso de implementação, com o Teste Unitário utilizando apenas o caso de uso correspondente à unidade implementada, o Teste de Integração abrangendo os casos de uso correspondentes às unidades integradas e o Teste de Sistema utilizando a totalidade dos casos de uso. O Teste de Aceitação é o último teste a ser realizado, com a participação dos usuários finais na aplicação de todos os casos de teste.

A.4.3 Objetivos e Prioridades do Sistema

Os objetivos e prioridades do sistema são descritos a seguir:

- O sistema deve possibilitar ao usuário realizar a pesquisa de animais.
- O sistema deve possibilitar ao usuário realizar a seleção de animais provenientes da pesquisa e também excluir uma seleção.
- O sistema deve possibilitar ao usuário definir os parâmetros a serem utilizados pelo AG como: utilização de touros (semêns), IQG, Nível de Problema e taxa de consanguinidade.
- O sistema deve possibilitar ao usuário iniciar a simulação do AG com os animais e parâmetros selecionados.

- O sistema deve possibilitar ao usuário visualizar os resultados gerados pelo AG como os acasalamentos, potenciais produtos gerados e estatísticas.

- O sistema deve possibilitar ao usuário exportar relatórios dos resultados gerados.

A.4.4 Casos de Teste

Caso de Teste 01	ID	Passos	Resultado Esperado
RF01 - Pesquisar animais através de função de busca.	1	Acessar a tela de seleção de touros (sêmens) ou de vacas.	A tela de seleção é exibida.
	2	Digitar uma tatuagem, um registro ou um nome de animal no campo superior de pesquisa da coluna esquerda da tela.	O campo de pesquisa deve receber o dado digitado.
	3	Clicar no botão “Buscar” localizado ao lado do campo de pesquisa.	O sistema deve exibir os dados do animal na aba “Pesquisa” da coluna direita da tela.

Caso de Teste 02	ID	Passos	Resultado Esperado
RF01E - Pesquisar animais através de função de busca - Animal não encontrado.	1	Acessar a tela de seleção de touros (sêmens) ou de vacas.	A tela de seleção é exibida.
	2	Digitar uma tatuagem, um registro ou um nome de animal no campo superior de pesquisa da coluna esquerda da tela.	O campo de pesquisa deve receber o dado digitado.
	3	Clicar no botão “Buscar” localizado ao lado do campo de pesquisa.	O sistema deve exibir a seguinte mensagem na aba “Pesquisa” da coluna direita da tela: ”Animal não encontrado”.

Caso de Teste 03	ID	Passos	Resultado Esperado
RF02 - Pesquisar animais através de aplicação de filtros.	1	Acessar a tela de seleção de touros (sêmens) ou de vacas.	A tela de seleção é exibida.
	2	Selecionar a raça do animal no <i>Input Radio</i> da coluna esquerda da tela.	O <i>Input Radio</i> deve marcar a raça selecionada.
	3	Selecionar o ano de nascimento do animal no <i>Input Select</i> “Ano” da coluna esquerda da tela.	O <i>Input Select</i> deve exibir o ano selecionado.
	4	Selecionar uma característica no <i>Input Select</i> “DEP” da coluna esquerda da tela.	O <i>Input Select</i> deve exibir a DEP selecionada.
	5	Selecionar o valor de máximo e mínimo para a DEP selecionada no <i>Input Range</i> ou nos <i>Inputs</i> “Máximo” e “Mínimo” da coluna esquerda da tela.	Os campos de <i>Input</i> devem exibir os valores selecionados. Quando o <i>Input Range</i> for alterado, os <i>Inputs</i> “Máximo” e “Mínimo” devem ajustar-se automaticamente e vice-versa.
	6	Confirmar a seleção de DEP no botão “Aplicar” da coluna esquerda da tela.	O sistema deve exibir na tela, logo abaixo dos <i>Inputs</i> de DEP, a DEP selecionada e seus valores de máximo e mínimo.
	7	Clicar no botão de “Pesquisar” localizado na parte inferior da coluna esquerda da tela.	O sistema deve exibir os dados do animal (ou dos animais) na aba “Pesquisa” da coluna direita da tela.

Caso de Teste 04	ID	Passos	Resultado Esperado
RF02E - Pesquisar animais através de aplicação de filtros - Nenhum animal encontrado.	1	Acessar a tela de seleção de touros (sêmens) ou de vacas.	A tela de seleção é exibida.
	2	Selecionar a raça do animal no <i>Input Radio</i> da coluna esquerda da tela.	O <i>Input Radio</i> deve marcar a raça selecionada.
	3	Selecionar o ano de nascimento do animal no <i>Input Select</i> “Ano” da coluna esquerda da tela.	O <i>Input Select</i> deve exibir o ano selecionado.
	4	Selecionar uma característica no <i>Input Select</i> “DEP” da coluna esquerda da tela.	O <i>Input Select</i> deve exibir a DEP selecionada.
	5	Selecionar o valor de máximo e mínimo para a DEP selecionada no <i>Input Range</i> ou nos <i>Inputs</i> “Máximo” e “Mínimo” da coluna esquerda da tela.	Os campos de <i>Input</i> devem exibir os valores selecionados. Quando o <i>Input Range</i> for alterado, os <i>Inputs</i> “Máximo” e “Mínimo” devem ajustar-se automaticamente e vice-versa.
	6	Confirmar a seleção de DEP no botão “Aplicar” da coluna esquerda da tela.	O sistema deve exibir na tela, logo abaixo dos <i>Inputs</i> de DEP, a DEP selecionada e seus valores de máximo e mínimo.
	7	Clicar no botão de “Pesquisar” localizado na parte inferior da coluna esquerda da tela.	O sistema deve exibir a seguinte mensagem na aba “Pesquisa” da coluna direita da tela: “Nenhum animal encontrado”.

Caso de Teste 05	ID	Passos	Resultado Esperado
RF03A - Selecionar animais.	1	Acessar a tela de seleção de touros (sêmens) ou de vacas.	A tela de seleção é exibida.
	2	Caso de Teste 01 ou Caso de Teste 02 ou Clicar no botão "Pesquisar", localizado na parte inferior da coluna esquerda da tela.	O sistema deve exibir os dados do animal (ou dos animais) na aba "Pesquisa" da coluna direita da tela.
	3	Clicar na <i>CheckBox</i> ao lado de cada registro de animal desejado.	A <i>CheckBox</i> deve ficar selecionada.
	4	Clicar no botão "Selecionar" localizado na parte superior da aba "Pesquisa" da coluna direita da tela.	O sistema deve persistir os registros dos animais selecionados no banco de dados e exibi-los na aba "Selecionados" da coluna direita da tela.

Caso de Teste 06	ID	Passos	Resultado Esperado
RF03B - Excluir a seleção de animais.	1	Caso de Teste 03.	O sistema deve exibir os registros dos animais selecionados na aba "Selecionados" da coluna direita da tela.
	2	Clicar na <i>CheckBox</i> ao lado de cada registro de animal a ser excluído.	A <i>CheckBox</i> deve ficar selecionada.
	3	Clicar no botão "Excluir" localizado na parte superior da aba "Selecionados" da coluna direita da tela.	O sistema deve deixar de exibir os registros dos animais na tela e excluí-los da tabela de selecionados do banco de dados.

Caso de Teste 07	ID	Passos	Resultado Esperado
RF04 - Definir parâmetros de utilização de touros.	1	Acessar a tela de seleção de touros (sêmens).	A tela de seleção é exibida.
	2	Caso de Teste 03 (passos 2 e 3).	A <i>CheckBox</i> deve ficar selecionada.
	3	Digitar um valor de utilização mínima e máxima nos <i>Inputs</i> junto aos dados dos animais selecionados.	Os <i>Inputs</i> de utilização devem receber os dados digitados.
	4	Clicar no botão "Selecionar" localizado na parte superior da aba "Pesquisa" da coluna direita da tela.	O sistema deve persistir os registros dos touros (sêmens) selecionados e seus parâmetros de utilização no banco de dados.

Caso de Teste 08	ID	Passos	Resultado Esperado
RF05 - Definir a ponderação da função de avaliação do AG.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Clicar na opção "Função de Avaliação" localizado na coluna esquerda da tela.	A coluna direita da tela deve exibir um <i>Input Range</i> com a ponderação da função de avaliação padrão do AG (90% IQG e 10% Nível de Problema).
	3	Clicar e arrastar o <i>Input Range</i> para selecionar a ponderação desejada.	O sistema deve exibir a nova ponderação selecionada.

Caso de Teste 09	ID	Passos	Resultado Esperado
RF06 - Editar o IQG.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Clicar na opção "IQG" localizado na coluna esquerda da tela.	A coluna direita da tela deve exibir um <i>Input Range</i> para cada DEP pertencente ao IQG com sua ponderação padrão e também para as demais DEPs medidas no PampaPlus.
	3	Clicar e arrastar o <i>Input Range</i> para selecionar a ponderação desejada em cada DEP.	O sistema deve exibir a nova ponderação selecionada e a soma dos coeficientes.

Caso de Teste 10	ID	Passos	Resultado Esperado
RF06E - Editar o IQG - Ponderação inválida.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Clicar na opção "IQG" localizado na coluna esquerda da tela.	A coluna direita da tela deve exibir um <i>Input Range</i> para cada DEP pertencente ao IQG com sua ponderação padrão e também para as demais DEPs medidas no PampaPlus.
	3	Clicar e arrastar o <i>Input Range</i> para selecionar a ponderação desejada em cada DEP.	O sistema deve exibir a soma dos coeficientes e a seguinte mensagem: "Ponderação inválida: a soma dos coeficientes deve ser 100%.

Caso de Teste 11	ID	Passos	Resultado Esperado
RF07 - Informar Nível de Problema.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Clicar na opção "Nível de Problema" localizado na coluna esquerda da tela.	A coluna direita da tela deve exibir um <i>Input</i> para receber um valor de Nível de Problema e abaixo dele uma lista com todas as DEPs do PampaPlus com uma <i>CheckBox</i> para cada.
	3	Digitar um valor no <i>Input</i> de Nível de Problema.	O sistema deve exibir o <i>Input</i> com o valor de Nível de Problema selecionado.
	4	Clicar no <i>CheckBox</i> das DEPs desejadas.	O sistema deve exibir o <i>CheckBox</i> marcado para todas as DEPs selecionadas.

Caso de Teste 12	ID	Passos	Resultado Esperado
RF08 - Definir taxa de consanguinidade.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Clicar na opção "Consanguinidade Máxima" localizado na coluna esquerda da tela.	A coluna direita da tela deve exibir um <i>Input</i> e abaixo dele um <i>Input Range</i> para receber um valor de consanguinidade máxima (%).
	3	Escolher a forma desejada para informar o valor. Digitar um valor no <i>Input</i> ou arrastar o <i>Input Range</i> .	Independente do tipo de <i>Input</i> escolhido, o sistema deve exibir o valor informado tanto no <i>Input</i> como no <i>Input Range</i> .

Caso de Teste 13	ID	Passos	Resultado Esperado
RF09 - Iniciar simulação com AG.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Editar os parâmetros desejados ou utilizar os parâmetros padrão do sistema.	Se um ou mais parâmetros forem definidos, a tela deve exibir o último parâmetro definido. Caso contrário não deve exibir nada.
	3	Clicar no botão "Simular" localizado na parte superior da coluna direita da tela.	As opções da barra de navegação superior da tela, da coluna da esquerda e os botões da coluna da direita devem ficar desabilitados. Uma barra de progresso e um botão "Parar" devem ser exibidos na coluna da direita. O AG deve iniciar a sua execução com a lista de animais selecionados e os parâmetros informados.

Caso de Teste 14	ID	Passos	Resultado Esperado
RF09E - Iniciar simulação com AG - Taxa de consanguinidade máxima excedida.	1	Acessar a tela de seleção de parâmetros.	A tela de seleção de parâmetros é exibida.
	2	Editar os parâmetros desejados ou utilizar os parâmetros padrão do sistema.	Se um ou mais parâmetros forem definidos, a tela deve exibir o último parâmetro definido. Caso contrário não deve exibir nada.
	3	Clicar no botão "Simular" localizado na parte superior da coluna direita da tela.	As opções da barra de navegação superior da tela, da coluna da esquerda e os botões da coluna da direita devem ficar desabilitados. Uma barra de progresso e um botão "Parar" devem ser exibidos na coluna da direita. O AG deve iniciar a sua execução com a lista de animais selecionados e os parâmetros informados. Assim que detectar que o limite de consanguinidade foi excedido, deve exibir um botão de "Continuar" e a seguinte mensagem: "Limite de consanguinidade excedido. Continuar mesmo assim?"

Caso de Teste 15	ID	Passos	Resultado Esperado
RF10: Visualizar resultado do AG.	1	Caso de Teste 10.	A execução do AG é completada e a tela de recomendação é exibida.
	2	Clicar na opção "Acasalamentos" localizado na coluna esquerda da tela.	A coluna da direita da tela deve exibir uma lista de todos os acasalamentos fornecidos pelo AG. Cada registro deve exibir a tatuagem e o nome da vaca e do touro para aquele acasalamento.
	3	Clicar na tatuagem de uma vaca ou de um touro (sêmen).	A tela deve exibir um Modal com as informações daquele animal.

Caso de Teste 16	ID	Passos	Resultado Esperado
RF11 - Visualizar potenciais produtos gerados.	1	Caso de Teste 10.	A execução do AG é completada e a tela de recomendação é exibida.
	2	Clicar na opção "Produtos" localizado na coluna esquerda da tela.	A coluna da direita da tela deve exibir uma lista de todos os produtos potenciais (filhos) dos acasalamentos fornecidos pelo AG.
	3	Clicar no registro de um produto gerado.	A tela deve exibir um Modal com as informações daquele produto.

Caso de Teste 17	ID	Passos	Resultado Esperado
RF12 - Visualizar estatísticas.	1	Caso de Teste 10.	A execução do AG é completada e a tela de recomendação é exibida.
	2	Clicar na opção "Estatísticas" localizado na coluna esquerda da tela.	A coluna da direita da tela deve exibir medidas de tendência central e de dispersão do IQG e do Nível de Problema para os produtos gerados.

Caso de Teste 18	ID	Passos	Resultado Esperado
RF13 - Exportar relatórios.	1	Caso de Teste 10.	A execução do AG é completada e a tela de recomendação é exibida.
	2	Clicar na opção "Gerar Documentos" localizado na coluna esquerda da tela.	A coluna da direita da tela deve exibir opções para exportar documentos dos acasalamentos gerados e dos potenciais produtos nos formatos Excel e PDF.
	3	Clicar em uma das opções de exportação de documentos.	O sistema deve efetuar o <i>Download</i> do documento selecionado.

A.4.5 Referencial

Este documento foi baseado nas recomendações do documento IEEE-Std 829-2008 e nos conceitos apresentados por Sommerville (2011).