

**UNIVERSIDADE ESTADUAL PAULISTA (UNESP)
FACULDADE DE CIÊNCIAS AGRÁRIAS E TECNOLÓGICAS
CAMPUS DE DRACENA**

Alexandre de Queiroz

Especialista em Tecnologia da Informação

**SIMULAÇÃO DE DADOS LONGITUDINAIS PARA ESTUDO
DE CARACTERÍSTICAS INDICADORAS DE BEM-ESTAR EM
PROGRAMAS DE MELHORAMENTO GENÉTICO ANIMAL**

Dracena

2023

**UNIVERSIDADE ESTADUAL PAULISTA (UNESP)
FACULDADE DE CIÊNCIAS AGRÁRIAS E TECNOLÓGICAS
CAMPUS DE DRACENA**

Alexandre de Queiroz

Especialista em Tecnologia da Informação

**SIMULAÇÃO DE DADOS LONGITUDINAIS PARA ESTUDO
DE CARACTERÍSTICAS INDICADORAS DE BEM-ESTAR EM
PROGRAMAS DE MELHORAMENTO GENÉTICO ANIMAL**

Dissertação apresentada à Faculdade de Ciências Agrárias e Tecnológicas – Unesp, Câmpus de Dracena como parte das exigências para obtenção do título de Mestre em Ciência e Tecnologia Animal.

Orientador: Prof. Associado Ricardo da Fonseca

Dracena

2023

FICHA CATALOGRÁFICA
Desenvolvida pela Seção Técnica de Biblioteca e Documentação
Campus de Dracena

Q384s

Queiroz, Alexandre de.

Simulação de dados longitudinais para estudo de características indicadoras de bem-estar em programas de melhoramento genético animal / Alexandre de Queiroz. -- Dracena: [s.n.], 2023.

75 f. : il.

Dissertação (Mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Tecnológicas, Dracena. Área do conhecimento: Produção Animal, 2023.

Orientador: Ricardo da Fonseca

1. Programação genética (Computação). 2. Estudos longitudinais. 3. Animais - Proteção. 4. Melhoramento genético de animais. 5. Método de decomposição. I. Título.

Bibliotecário Fábio Sampaio Rosas
CRB 8/6665

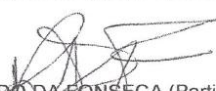
CERTIFICADO DE APROVAÇÃO


TÍTULO DA DISSERTAÇÃO: **Simulação de dados longitudinais para estudo de características indicadoras de bem-estar em programas de melhoramento genético animal**

AUTOR: ALEXANDRE DE QUEIROZ

ORIENTADOR: RICARDO DA FONSECA

Aprovado como parte das exigências para obtenção do Título de Mestre em Ciência e Tecnologia Animal, área: Produção Animal pela Comissão Examinadora:


Prof. Associado RICARDO DA FONSECA (Participação Virtual)
Departamento de Produção Animal - DPA / Faculdade de Ciências Agrárias e Tecnológicas - FCAT
UNESP / Dracena


Prof. Dr. CELSO TADAO MIASAKI (Participação Virtual)
Departamento de Produção Animal - DPA / Faculdade de Ciências Agrárias e Tecnológicas - FCAT
UNESP / Dracena


Dra. MICHELE PORTO PIRES (Participação Virtual)
Laboratório de Sanidade e Genética Animal / Embrapa Suínos e Aves

Dracena, 01 de março de 2023.

DADOS CURRICULARES DO AUTOR

Alexandre de Queiroz, nascido em 17 de maio de 1974, na cidade de Garça/SP. Em 1980, mudou-se com seus pais para a cidade de Araçatuba/SP, onde cresceu e viveu até os 33 anos. Concluiu o ensino médio junto com o técnico em processamento de dados em 1992, iniciou sua carreira como programador de computadores. Até então, era auxiliar administrativo, desde 1987. Em 2000, concluiu a graduação em tecnologia em processamento de dados e, em 2004, iniciou no banco Nossa Caixa como auxiliar administrativo, e professor do curso técnico em informática na Fundec de Dracena/SP. Em 2006, foi bolsista da Fapesp TT-V, no laboratório de Computação Científica Aplicada à Zootecnia – LuCCa-Z, na Unesp de Dracena e, em 2007, concluiu a especialização em Tecnologia da Informação pela Unesp de Bauru/SP. Retornou na carreira acadêmica em 2009, no Curso de Computação e Tecnólogo em Análise de Sistemas da Unifadra e Sistema de Informação na Faca de Tupã/SP. Em 2019, concluiu a Especialização em Formação Pedagógica para Graduados não licenciados pelo CEETEPS, onde também leciona desde 2015.

DEDICATÓRIA

À minha esposa Ana Maria de Souza Queiroz,
aos meus filhos, Arthur Alexandre de Souza
Queiroz e Felipe Alexandre de Souza Queiroz,
que se privaram da minha presença e me
possibilitaram mais essa conquista e ao meu
pai, Aristides de Queiroz, *in memoriam*.

AGRADECIMENTOS

Quero expressar minha profunda gratidão ao Professor Associado Ricardo da Fonseca por ter me aceitado como seu orientado, por sua paciência e companheirismo durante esses dois anos de trabalho conjunto.

Quero também agradecer ao Professor Dr. Celso Tadao Miasaki e à Dra. Michele Porto Pires, que aceitaram o convite para compor a banca deste trabalho e contribuíram significativamente para sua conclusão.

Agradeço a Édna Sayuri Kanno Takaki Lorenzon por sua imensa colaboração e atenção durante o curso, bem como ao departamento de TI da unidade, ao meu amigo e bibliotecário Dr. Fábio Sampaio Rosas e à Biblioteca da FCAT, pelas ajudas valiosas.

Agradeço à minha amiga Fraine Natasha Alessio por sua companhia, ajuda e paciência ao longo desse processo, bem como ao meu amigo Danilo Pereira da Silva, que me incentivou e apoiou muito durante o desenvolvimento deste trabalho.

Agradeço também a Catia Silene Afonso e à Professora Arlene Gonçalez Tenório Etto, minhas atuais colegas de trabalho, por apoiarem e compreenderem a importância deste trabalho para minha vida pessoal e profissional.

Quero agradecer à minha mãe, meu irmão, minha sogra, meus filhos e minha esposa por estarem sempre ao meu lado e me apoiarem durante esta etapa desafiadora.

Agradeço a Deus por ter me permitido conhecer essas pessoas e por ter me guiado na construção e conclusão de mais esta etapa da minha vida.

“As pessoas que são loucas o suficiente para pensar
que podem mudar o mundo são as que o fazem.”
(Steve Jobs, 2005).

RESUMO

Com o crescimento da população mundial, surgem diversos problemas que atingem muitos países, inclusive o Brasil, que produz alimentos em larga escala, para abastecer a necessidade interna e outros países. A preocupação com o meio ambiente e o bem-estar animal vem sendo um assunto muito abordado, nos últimos anos, com a conscientização da população. Os cuidados com os animais, na cadeia produtiva, influenciam os negócios, principalmente, quando se trata de exportação. Este trabalho visou o estudo de indicadores de bem-estar animal e buscou uma implementação, para simuladores de melhoramento genético animal. Esses indicadores devem ser longitudinais, ou seja, os dados se alteram ao longo do tempo, de acordo com a situação em que o animal se encontra e devem ser correlacionados. Foram implementados algoritmos para simular tais características e aplicada a decomposição de Cholesky, que permitiu a correlação de uma matriz, simétrica e positiva definida, gerada aleatoriamente forçando que a correlação fosse diminuindo no tempo, conforme seu distanciamento. Embora essa simulação permita a utilização de apenas uma característica e a complexidade dos algoritmos não sejam lineares, os resultados se mostraram muito promissores com relação aos efeitos longitudinais esperados para cada indivíduo.

Palavras-chave: Programação genética (Computação). Estudos longitudinais. Animais - Proteção. Melhoramento genético de animais. Método de decomposição.

ABSTRACT

With the growth of the world population, various problems arise that affect many countries, including Brazil, which produces food on a large scale to supply internal needs and other countries. Concerns for the environment and animal welfare have been widely discussed topics in recent years, with the public becoming increasingly aware. Care for animals in the production chain influences business, especially when it comes to exportation. This work aimed to study animal welfare indicators and sought to implement them in animal genetic improvement simulators. These indicators must be longitudinal, meaning that data changes over time according to the animal's situation and must be correlated. Algorithms were implemented to simulate such characteristics, and the Cholesky decomposition was applied, which allowed the correlation of a randomly generated symmetric and positive definite matrix, forcing the correlation to decrease over time as it gets more distant. Although this simulation allows for the use of only one characteristic, and the algorithm complexity is not linear, the results were very promising concerning the expected longitudinal effects for each individual.

Keywords: Genetic programming (Computing). Longitudinal studies. Animals - Protection . Genetic improvement of animals. Decomposition method.

LISTA DE ILUSTRAÇÕES

Figura 1 - População mundial atual	17
Figura 2 - Produção de carne bovina em 2022.....	18
Figura 3 - Consumo de carne no Brasil.....	29
Figura 4 – Demonstração da Teoria Coalescente.....	39
Figura 5 - Exemplo de simulação por Gene-dropping.....	42
Figura 6 - Níveis de complexidade.....	53

LISTA DE TABELAS

Tabela 1 - Amostra dos indivíduos simulados com valores genéticos aditivos.....	57
Tabela 2 - Tabela de covariância gerada aleatoriamente	57
Tabela 3 - Matriz LL' resultante da decomposição de Cholesky em $cov(T)$	58
Tabela 4 – Amostra da matriz de valores aleatórias independentes com variância unitária	58
Tabela 5 - Matriz E_t resultante da multiplicação da triangular inferior L por z	59
Tabela 6 - Covariâncias geradas sobre a matriz E_t	59
Tabela 7 - Resultado da simulação do fenótipo para cada indivíduo.....	60

LISTAS DE ABREVIATURAS E SIGLAS

DFD	Dark, firm and dry – escura, dura e seca
DNA	Ácido desoxirribonucleico
EFSA	European Food Safety Authority
FAO	Organização das Nações Unidas para Agricultura e Alimentação
FAWC	Farm Animal Welfare Council
GPS	Global Positioning System
OIE	Organização Mundial de Saúde Animal
PSE	Pale, soft e exudative – pálida, mole e exsudativa
QTL	Quantitative trait locus
SNA	Sistema nervoso autônomo
SNP	Single nucleotide polymorphism
SSH	Secure Socket Shell
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura
VGA	Valor genético aditivo

LISTA DE SÍMBOLOS

A_1	Genótipo do loco A_1
A_2	Genótipo do loco A_2
A_i	Valor genético aditivo do i – <i>ésimo</i> indivíduo
a	Valor genotípico, do loco A_1A_1 , A_1A_2 ou A_2A_2 , multiplicado por 1, 0 e -1, respectivamente
C_i	Efeito dos alelos do i – <i>ésimo</i> indivíduo
$cov(E)$	Matriz onde as covariâncias fornecidas pelo usuário e, as variâncias, são calculadas pela herdabilidade e variância genotípica
$cov(T)$	Covariância da matriz T
E	Matriz de correlações, onde na diagonal tem-se números 1's e fora das diagonais $e_{ij}e'_{ij}$
E_{ij}	Efeito global do ambiente que incide sobre o i – <i>ésimo</i> indivíduo no tempo j
E_p	Efeito de ambiente permanente para a característica
E_{p_i}	Efeito permanente de ambiente sobre o i – <i>ésimo</i> indivíduo
E_t	Matriz resultante da multiplicação da triangular inferior L por z
E_{ti}	Efeito do ambiente temporário sobre a característica no momento da coleta
E_{tij}	Efeito do ambiente temporário (não controlado) sobre o i – <i>ésimo</i> indivíduo no tempo j .
E_{tij}'	Efeito transposto do ambiente temporário (não controlado) sobre o i – <i>ésimo</i> indivíduo no tempo j .
e_{ij}	Valor dos efeitos residuais que incidem sobre o i – <i>ésimo</i> indivíduo no momento j .
h^2	Herdabilidade
L	Triangular inferior da matriz resultante da Decomposição de Cholesky em T
LL'	Matriz resultante da decomposição de Cholesky em T
l	Número de locos para uma determinada característica

O	Ordem de crescimento assintótico do tempo de execução de um algoritmo, ou seja, o número máximo de operações básicas, no pior caso, que ele pode executar
$O(i)$	Complexidade de crescimento assintótico em i vezes
$O(j)$	Complexidade de crescimento assintótico em j vezes
$O(n)$	Complexidade de crescimento assintótico de n vezes
P	Valor fenotípico da característica
P_{ij}	Valor fenotípico da característica do i – <i>ésimo</i> indivíduo no tempo j
p	Frequências dos alelos A_1
q	Frequências dos alelos A_2
T	Matriz gerada aleatoriamente ou definida pelo usuário, simétrica e positiva definida
$T(n)$	Equação da complexidade do algoritmo
t	Tempo que um dado foi coletado para a mesma característica
$varA$	Variância genotípica da característica
$varE_t$	Variância do efeito de ambiente da característica
$varP$	Variância fenotípica da característica
Y_i	Vetor de características repetidas para o i – <i>ésimo</i> indivíduo
Y_{ij}	Matriz de dados longitudinais onde, i é o indivíduo e j , é o número da característica repetida
z	Vetor de variáveis aleatórias independentes e com variância unitária
μ	Constante inerente a média geral da característica
σ_{ir}	covariância entre as i – <i>ésima</i> e r – <i>ésima</i> medidas
σ_{rr}^2	variância da r – <i>ésima</i> medida repetida

SUMÁRIO

1 INTRODUÇÃO	17
1.1 Hipóteses	20
2 OBJETIVOS.....	21
2.1 Objetivo Geral	21
2.2 Objetivos Específicos	21
3 REVISÃO DE LITERATURA.....	22
3.1 Bem-estar Animal e sua Importância.....	22
3.2 Bem-estar vs Produção Animal.....	24
3.3 Coleta de Fenótipos.....	31
3.3.1 Fenótipos como medidas de bem-estar	32
3.4 Simulação de Dados.....	36
3.4.1 Visão geral	36
3.4.2 Simulação de dados na área genética	37
3.4.3 Simulação de dados por teoria Coalescente.....	38
3.4.4 Simulação de dados por <i>Gene-dropping</i>	40
4 MATERIAL E MÉTODOS	43
4.1 Hardware	43
4.2 Sistema Operacional.....	43
4.3 Linguagens de Programação, Compiladores e Softwares Estatísticos.....	43
5 DESENVOLVIMENTO	45
5.1 Simulação de Valores Genéticos Aditivos.....	45
5.2 Simulação de Valores de Ambiente Correlacionados.....	45
5.3 Decomposição Cholesky.....	49
5.3.1 Visão geral	49
5.3.2 Algoritmo da decomposição de Cholesky.....	51

5.3.3 Complexidade de algoritmo.....	52
5.3.4 Complexidade do algoritmo da decomposição de Cholesky.....	53
5.3.5 Complexidade do algoritmo da de cálculo dos efeitos de ambiente temporário correlacionados.....	54
5.3.6 Complexidade do algoritmo para criar a matriz T	55
5.3.7 Algoritmo para simulação de efeitos de ambiente temporários correlacionados.....	55
6 RESULTADOS E DISCUSSÃO	57
7 CONCLUSÃO.....	61
REFERÊNCIAS.....	62
APÊNDICES.....	68
I. Simulador.....	68
II. Rotina de definição de sexo.....	69
III. Rotina de configuração de cada loco.....	69
IV. Rotina de cálculo do valor genético aditivo de cada loco.....	69
V. Rotina para geração de locos poligênicos por cromossomo	70
VI. Rotina para calcular a variância da população	71
VII. Rotina para criação do genoma baseado em parâmetros	71
VIII. Rotina para criar os cromossomos e definição do tipo de cada loco.....	72
IX. Rotina para gerar a matriz de efeitos de ambiente	72
X. Rotina para criar indivíduos baseados no genoma	73
XI. Rotina para geração da população.....	74
XII. Rotina para criação de locos.....	74
XIII. Rotina para criar a matriz de covariâncias aleatórias.....	75

1 INTRODUÇÃO

Com o crescimento da população mundial, surgem os problemas de saneamento, energia elétrica, água potável e alimentos. Somos hoje, segundo o site *WorldOmeter*, que atualiza dados em tempo real, 8.018.838.259 habitantes, conforme ilustrado na figura 1. A demanda por alimento, não de hoje, é, de certa forma, suprida em países que não conseguem produzir o que consomem, por importação, onde países como Brasil, que possuem grande espaço territorial, produzem em larga escala e enviam seus produtos para muitos outros países. Assim, o agronegócio tem crescido e, com o estreitamento das fronteiras da comunicação mundial, surgem algumas preocupações como meio ambiente e bem-estar animal, principalmente em relação aos sistemas que destinam seus produtos à exportação (SILVA, 2022).

Figura 1 - População mundial atual



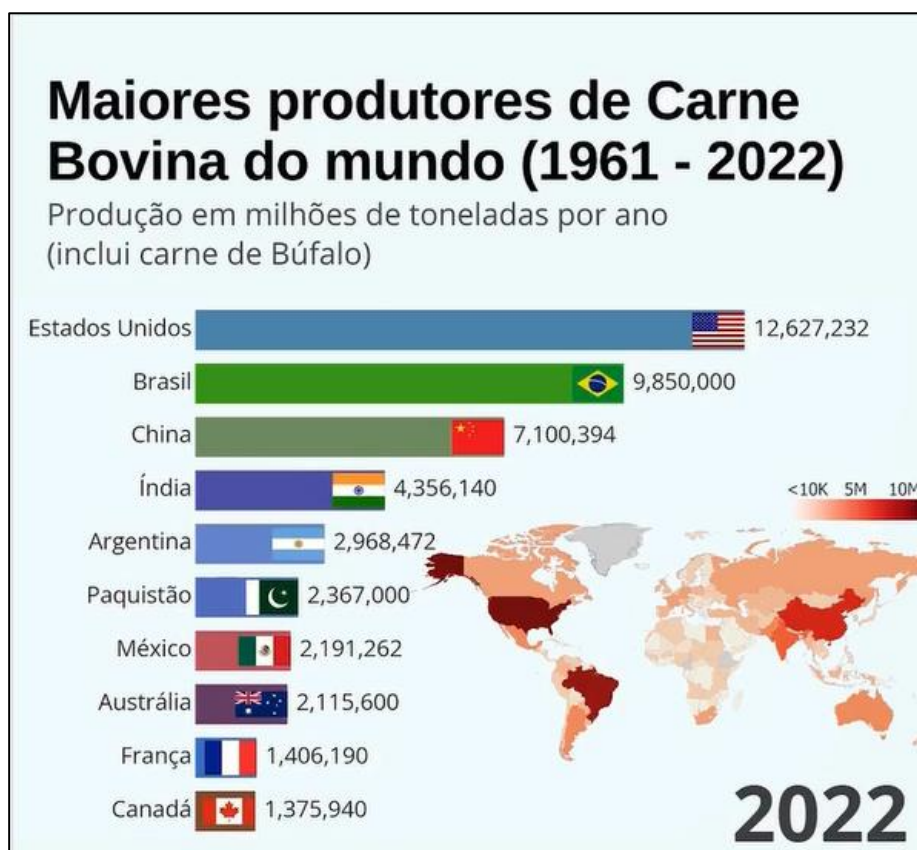
Fonte: WorldOmeter¹

Com a crescente conscientização da população, os sistemas de produção são pressionados a seguirem normas e técnicas para evitar o sofrimento animal. Aqueles que não produzirem de acordo com as exigências de bem-estar terão dificuldades em vender seus produtos, seja pela rejeição do

¹ Disponível em <<https://www.worldometers.info/world-population/>>, Último acesso em 26/02/2023 às 00:00.

consumidor ou pela utilização dessa condição como barreira não-tarifária à exportação. O Brasil, segundo a FAO – Organização das Nações Unidas para Agricultura e Alimentação, foi o segundo maior produtor de carne bovina do mundo em 2022, ficando atrás apenas dos Estados Unidos, conforme ilustrado na figura 2.

Figura 2 - Produção de carne bovina em 2022



Fonte: FAO

O bem-estar é uma característica multifatorial e sua avaliação deve ser realizada pela medição de várias características indicadoras, sejam elas relacionadas à saúde, comportamento, fisiologia ou aspectos emocionais.

Idealmente, a coleta de dados desses indicadores deve ser realizada várias vezes durante a vida produtiva do indivíduo (e talvez, várias vezes durante um dia), para que as avaliações das características envolvidas tenham precisão e qualidade. Dessa forma, o trabalho com a característica bem-estar envolve a coleta e avaliação de dados longitudinais. Entretanto, o processo de coleta é muito difícil se realizado por seres humanos, uma vez que demanda

coleta de dados de maneira frequente e muitas vezes durante a vida dos indivíduos do rebanho (que por sua vez, podem ser compostos de muitos animais). Para que o trabalho de coleta de fenótipos seja viável, deve-se utilizar recursos automatizados. Uma característica da coleta automatizada é a capacidade de gerar um grande volume de dados longitudinais, os quais, nos colocam novos desafios com relação à sua manipulação, avaliação e interpretação. Porém, esses equipamentos ainda possuem alto custo, o que dificulta sua utilização por empresas e instituições de pesquisa.

Uma possível solução para se iniciar os estudos com esses dados, (enquanto os equipamentos não se tornam amplamente disponíveis nos sistemas de produção) que visa a solução de problemas de armazenagem, manipulação, avaliação e interpretação, seria a simulação de dados. Uma vez que os dados longitudinais são correlacionados nas medidas sequenciais nos animais, o desafio é implementar um método que leve em consideração esta estrutura de correlação entre os registros, e simule, com a maior precisão e similaridade possível, os dados, que são fornecidos pela coleta automatizada.

Uma vez que os dados simulados devem se prestar também aos estudos em programas de melhoramento genético animal, além das correlações entre medidas repetidas nos animais, deve se levar também a simulação de outros efeitos de ambiente e valores genéticos aditivos para a composição dos valores fenotípicos dos indivíduos.

Assim, é importante desenvolver algoritmos que sejam eficientes e de qualidade na geração de grande volume de dados longitudinais, para que, posteriormente, possam ser utilizados em estudos de metodologias, avaliação genética, teste de modelos e desenvolvimento de ferramentas de interpretação/avaliação. Todos esses estudos contribuirão para a melhoria da implementação de técnicas de bem-estar geral do rebanho e, conseqüentemente, para a maior adequação da produção pecuária brasileira às exigências crescentes do mercado consumidor interno e externo com relação ao bem-estar nos sistemas produtivos.

1.1 Hipóteses

1. É possível desenvolver um algoritmo para simulação de dados longitudinais que leve em consideração efeitos genéticos e ambientais potencialmente correlacionados;
2. Um algoritmo de complexidade linear pode ser desenvolvido para a simulação de dados longitudinais

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver um algoritmo para a simulação de dados longitudinais que leve em consideração os efeitos genéticos e de ambiente para a formação do fenótipo;

2.2 Objetivos Específicos

- Estudar a complexidade do algoritmo;
- Implementar os algoritmos desenvolvidos no sistema R, o qual se constitui em uma plataforma popular e de fácil acesso à comunidade científica.

3 REVISÃO DE LITERATURA

3.1 Bem-estar Animal e sua Importância

O bem-estar animal vem sendo amplamente discutido nas últimas décadas e é um tema complexo que envolve dimensões científicas, culturais, éticas, sociais, religiosas, econômicas e políticas (CEBALLOS; SANT'ANNA, 2018). É de caráter multidisciplinar e envolve a etologia e questões morais e éticas. Na União Europeia, existe o Protocolo Anexo ao Tratado de Amsterdã Relativo ao Bem-Estar Animal e a nível mundial, desde a década de 1970, a Organização das Nações Unidas para a Educação, a Ciência e a Cultura - UNESCO, emitiu um documento chamado Declaração Universal dos Direitos dos Animais. Na Áustria a Lei Federal sobre o estatuto jurídico dos animais foi aprovada em março de 1988 e no seu parágrafo 285a, traz um conceito muito interessante sobre coisa e animal:

“Os animais não são coisas; estes são protegidos mediante leis especiais. As normas relativas às coisas são aplicáveis aos animais, na medida em que não existam disposições divergentes.”

O conceito de que os animais não são coisas, também são parte do Código Civil da Suíça, Alemanha e França e, além do tema, traz uma série de alterações nos direitos e valores éticos, em favor dos animais, e do direito de sua propriedade.

A ideia do bem-estar animal está longe de ser apenas uma coisa jurídica. Já na década de 1960, o Comitê Brambell, formado por profissionais da área de agricultura e pecuária do Reino Unido, apresentou um conceito que estabelecia um conjunto de estados que foram chamados de “cinco liberdades” para definir melhor a ideia de bem-estar animal. Esse conceito foi aperfeiçoado pelo *Farm Animal Welfare Council* - FAWC e determina a percepção de bem-estar pelo próprio indivíduo independentemente do local onde esteja inserido, seja em ambientes de produção ou na natureza (GUIRRO, 2022). As cinco liberdades são:

1. O animal deve estar livre de fome e sede;
2. O animal deve estar livre de dor e doenças;
3. O animal deve estar livre de desconforto;
4. O animal deve estar livre de medo e de estresse;
5. O animal deve estar livre para expressar seu comportamento natural.

Para alcançar o estado de bem-estar, essas cinco liberdades não podem estar em desacordo. Gentle *et al.* (1982) diz que o animal deve estar em harmonia com a natureza e o meio em que vive e, Broom (1986), definiu que o bem-estar de um animal, é o estado do indivíduo em relação ao seu esforço de sobreviver no ambiente em que está inserido. Se os desafios são vencidos facilmente, o animal está em bem-estar satisfatório. Por outro lado, se o indivíduo encontra muitas dificuldades para conseguir sobreviver, seu bem-estar pode ser pobre, sendo considerados, os fatores negativos como externos, como interferência humana e local onde o animal está inserido ou fatores internos como estresse, doenças entre outros.

Broom e Johnson (2000) destacaram três aspectos importantes a serem observados na questão do bem-estar animal:

“i) o bem-estar é uma característica do animal e é algo que o ser humano não é capaz de fornecer; ii) o bem-estar pode variar de muito bom a muito ruim; iii) o bem-estar pode ser medido cientificamente.”

O bem-estar animal, busca identificar o estado em que se encontram os animais e desenvolver estratégias para amenizar o sofrimento causado pela interferência humana em seu ambiente e manejo (GALHARDO; OLIVEIRA, 2006). Assim, busca analisar o funcionamento biológico, as emoções e experiências mentais e o comportamento natural do indivíduo (MANTECA *et al.*, 2013).

O projeto Welfare Quality, segundo Manteca *et al.* (2013), combina as análises de percepção e atitudes do consumidor com a ciência do bem-estar animal e identifica quatro princípios que são: a boa alimentação, não permitindo que o animal tenha fome e sede prolongados; bom alojamento em relação ao descanso, conforto térmico e facilidade de movimentação; boa saúde,

compreendendo ausências de lesões e enfermidades e dores causadas por manejo; e comportamento adequado ou seja, expressão de comportamento social e de outras condutas adequadas, interação humano animal positiva e estado emocional positivo.

3.2 Bem-estar vs Produção Animal

Pode-se notar que, ao longo dos anos, tem crescido o interesse pelo bem-estar dos animais e nas atividades que os envolvem, sejam elas pesquisas científicas, esportes ou produção e o entendimento que eles também devem ter direitos e os mesmos devem ser respeitados. Organizações internacionais como a Organização Mundial de Saúde Animal - OIE, a *European Food Safety Authority* - EFSA e a Organização das Nações Unidas para Agricultura e Alimentação - FAO, também incluíram suas agendas o bem-estar animal como questão prioritária (CEBALLOS; SANT'ANNA, 2018).

No Brasil, a legislação sobre maus tratos de animais, teve início em 10 de julho de 1934, com o Decreto nº 24.645/1934 que estabeleceu medidas de proteção, a Instrução Normativa 03/2000, que aprovou o regulamento técnico de métodos de insensibilização para o abate humanitário (LEITE *et al.*, 2015), a Instrução Normativa nº 56/2007, que estabeleceu as recomendações de boas práticas de bem-estar animal, para animais de produção e de interesse econômico, que são direcionadas para o manejo, cuidados com as instalações e transporte (LINO *et al.*, 2016), a Portaria nº 185/2008 que instituiu a Comissão Técnica Permanente de Bem-estar animal do MAPA, que tem por objetivo, coordenar as ações de bem-estar animal e fomentar as boas práticas de bem-estar pelos produtores rurais (LEITE *et al.*, 2015). Ainda, a Resolução nº 675 de 21 de junho de 2017, traz normativas sobre o transporte de animais de produção ou interesse econômico, animais para esporte, lazer e de exposição. Essa resolução trata do transporte de animais vivos a fim de minimizar o sofrimento e ferimentos desnecessários e o estresse dos animais, para garantir o bem-estar durante o processo (POLASTRINI *et al.*, 2021).

Um dos maiores desafios relacionados com o bem-estar animal está na produção, justamente, onde ocorrem situações conflitantes entre a produtividade do animal e seu bem-estar. A premissa de que o animal que

expressa melhor a sua característica, em alta produtividade, está em bem-estar é falsa, segundo Molento (2005). Houve, após a segunda guerra mundial, uma crescente demanda por consumo de proteína animal que trouxe um grande aumento na cadeia produtiva, o que trouxe enormes benefícios econômicos (BRAGA *et al.*, 2018).

Já Molento, (2005), afirma que a produção animal pode ter iniciado a cerca de dez mil anos e se intensificou ao longo dos anos, tendo sua maior expansão no início do século XX, o que forçou os produtores a criar os confinamentos de bovinos, suínos e de aves, para atender às pressões comerciais e afetou, drasticamente o bem-estar dos indivíduos, pois houve a necessidade de baixar os custos, aumentar o ganho de produção, às custas do sofrimento do animal.

Hötzel e Machado Filho (2004), também alertam para os maus tratos dos animais nos confinamentos e afirmam que a desinformação ainda é um grande entrave para desenvolver discussões produtivas. Não é porque os animais já estão domesticados que não exigem cuidados e defende as cinco liberdades mínimas descritas no Comitê Brambell. O sofrimento causado pelo descaso, não afeta apenas os animais, mas a falta do bem-estar, pode trazer sérios impactos na cadeia produtiva trazendo insegurança alimentar à população. Podem criar grandes epidemias de doenças, em bovinos, como a febre aftosa, a doença da vaca louca e contaminação de carne de aves por dioxinas (HÖTZEL; MACHADO FILHO, 2004).

Os sistemas produtivos, geralmente, levam em consideração a lucratividade e raramente as dificuldades que os animais possuem em relação ao ambiente em que foram produzidos, não consideram suas necessidades físicas, comportamentais e psicológicas. Entretanto, a qualidade com que o animal é tratado pode trazer diversos benefícios como diminuição da mortalidade e doenças, redução com custos com medicamentos e diminuição de zoonoses (BRAGA *et al.*, 2018). Ainda, segundo Lino, Pinheiro e Ortulho (2016), a forma com que os animais são tratados, pode agregar valores econômicos, independente dos valores produtivos, tendo em vista que, assegurar a melhoria da qualidade de vida dos animais também envolve investimentos mais altos.

Segundo Froehlich (2016), a aplicação das cinco liberdades em um ambiente de produção é utópico e apresenta uma tentativa de proporcionar uma melhor condição de vida aos animais das fazendas, principalmente em confinamentos de gado de corte. A restrição espacial, que um animal confinado sofre, é grande em relação aos criados em pastagens e não são suficientes para garantir o bem-estar necessário, o que causa, principalmente, problemas de saúde devido à umidade, problemas nos cascos e outros, que acabam por diminuir a sua produtividade.

No manejo, Costa *et al.* (2021) relatam que as práticas de bem-estar, evita gritos no manejo, golpes nos animais, os deixam menos reativos e estressados e há uma melhora na resposta reprodutiva. Não são apenas os animais que são beneficiados com isso. Os vaqueiros ficam menos estressados, sofrem menos acidentes, o que significa melhores condições de trabalho a aqueles que não adotaram as práticas de bem-estar no manejo. Macitelli *et al.* (2007), realizaram uma pesquisa com bovinos de corte, confinados a céu aberto, com relação ao espaço disponibilizado por animal, sendo 6, 12 e 24m². Constataram que, os confinamentos com menos espaço por animal, acumula muito mais lama e fezes no período de chuvas e maior concentração de poeira no período de seca, trazendo um risco maior de doenças, mudanças de comportamentos dos animais e estresse. Costa *et al.* (2021), afirmam que um ambiente assim, pode também trazer riscos à saúde das pessoas que trabalham e vivem próximas ao local. Ainda, segundo Macitelli *et al.* (2007), locais com 24m² por animal, foram mais lucrativos, com lucro operacional de R\$ 109,00/cabeça, os com 12m² por animal, R\$ 102,00 e com 6m², R\$ 80,90. Além disso, os confinamentos com área maior por animal apresentaram risco de perda menor que os com área reduzida, o que evidencia que, as boas práticas de bem-estar no manejo, têm potencial para melhorar o desempenho dos animais e qualidade de vida das pessoas envolvidas (COSTA *et al.*, 2021).

Resultados positivos das boas práticas de bem-estar, também foram constatados por Silva-Antunes e Costa (2021), no manejo de bovinos leiteiros em um estudo de dois anos onde, na segunda metade do experimento, as bezerras passaram a receber boas práticas de manejo desde o seu nascimento, com controle de ingestão de colostro, manutenção das bezerras

em grupos, uso de baldes com bicos para aleitamento, estimulação tátil durante a amamentação e desmame progressivo. Os resultados foram que, as bezerras que receberam as boas práticas de bem-estar no manejo tiveram uma redução expressiva da necessidade de uso de antibióticos por mês, de 68,2 para 39,1% e redução de 18,8 para 5,3% na taxa de mortalidade, que representa ganhos econômicos expressivos devido às reduções dos custos.

Nos confinamentos de gado de leite, existem muitas doenças causadas pela baixa ingestão de nutrientes. Isso se dá geralmente pelo espaço que as vacas têm para se alimentarem e pelo tamanho do coxo, em locais com menos de 60 cm disponíveis por animal. Nestes lugares, menos de 70% das vacas conseguem se alimentar simultaneamente, o que aumenta a agressividade e, muitas delas, chegam a parar de comer corretamente (BOND *et al.*, 2012). Ainda, segundo Bond *et al.* (2012), locais que apresentam espaços maiores fazem com que as vacas aumentem a sua frequência de alimentação e diminuam as doenças relacionadas à nutrição. Locais com mais espaço e manejo individualizado, evitam problemas reprodutivos, mastite, a alta contagem de células somáticas e doenças podais, além dos animais ingerirem a quantidade adequada de nutrientes, que evita a fome crônica e doenças metabólicas.

O ambiente onde as vacas se encontram também é um bom indicador de bem-estar. Um local em que as vacas são alojadas em piquetes sem baixa, dificultam a higienização e o manejo dos animais, conseqüentemente resulta em maior número de células somáticas no leite e maior risco de os animais desenvolverem doenças que resultam no empobrecimento do bem-estar (COSTA *et al.*, 2021). Ainda, nos locais onde não foram adotadas as boas práticas no manejo desses animais, não evitou o medo, o estresse e, conseqüentemente a reatividade durante a ordenha, resultou na redução de leite em 1,0 a 1,6 kg por animal. Por outro lado, onde foram adotadas as práticas de bem-estar, não houve queda na produção de leite e muitas vacas ruminaram durante a ordenha, demonstrando tranquilidade durante o manejo (COSTA *et al.*, 2021).

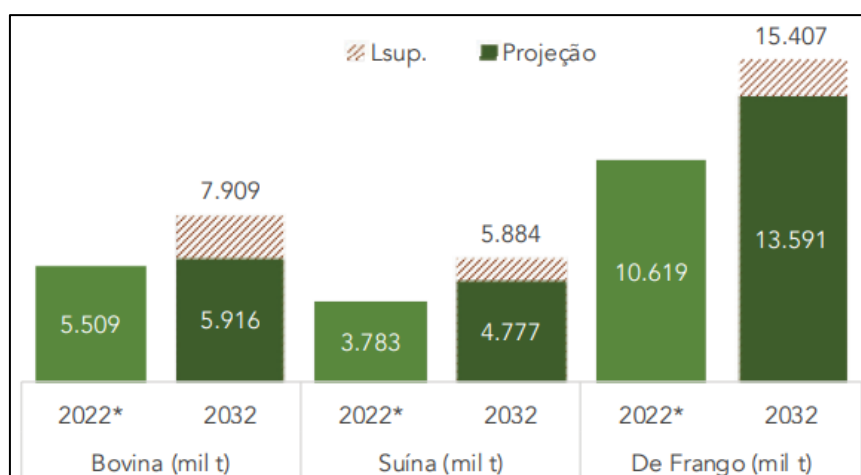
Os ovinos também são muito utilizados na produção animal, uma vez que são animais mais dóceis que os bovinos, necessitam de menos espaço, tem um período de gestação e tempo de abate mais curto e ciclo produtivo

mais rápido. Porém, demandam um manejo sanitário mais adequado pois apresentam diversas doenças que prejudicam o desenvolvimento dos animais e provocam altos índices de mortalidade (LINO *et al.*, 2016). Ainda, o custo do rebanho acaba ficando muito alto devido aos medicamentos e, os endoparasitas, acabam causando perda de peso entre 30 e 40% e isso envolve a produção de carne, leite e lã de qualidades inferiores.

Em vários trabalhos, com ovinos da raça Santa Inês, observando as condições dos ambientes sem sombra, com sombra natural (presença de árvores) e sombra artificial (com sombrite), Lino *et al.* (2016), relatam que os animais que ficaram na sombra, seja natural ou artificial, mantiveram a homeotermia com menor esforço do aparelho termo regulatório e maior consumo de matéria seca, o que resultou em um maior ganho de peso diário (174g/dia) em relação aos animais que estavam sem acesso à sombra (122g/dia). Isso ocorreu porque os animais que ficaram expostos ao sol, constantemente, tiveram que diminuir a ingestão de alimento, reduzindo a fermentação ruminal, para melhor dissipar o calor excedente para o ambiente.

A criação de suínos é a segunda maior do Brasil e a carne de porco é a terceira mais consumida no país, como ilustra a figura 3 (SUEST/SMAE/Embrapa, 2022). O bem-estar na suinocultura também é bem debatido e o ambiente enriquecido com área de 1,75 ou 3,5 m² por animal, baia com palha, área de fuçar e de defecar, é utilizado para minimizar os impactos que os confinamentos impõe como limitação. Os suínos em ambientes enriquecidos possuem comportamento diferente, voltando um quarto do seu tempo para o substrato no piso, enquanto os animais com espaço comum, gastam o seu tempo explorando objetos fixos e comportamentos mais nocivos como fuçar e morder outros porcos. Segundo Machado Filho (2000), o bem-estar dos suínos acaba sendo visto pós abate. Os animais criados em ambiente enriquecidos apresentaram carne com menores perdas na cocção e mais macia além de carcaças mais pesadas e maior espessura de toucinho sendo, esta última, 26,3 ($\pm 1,35$) mm enquanto os de ambiente normal a média foi de 23,85 ($\pm 1,42$) mm. Embora estes resultados não tenham sido estatisticamente significantes, os animais que são expostos ao estresse crônico ou severo, apresentam carne PSE (*pale, soft e exudative* – pálida, mole e exsudativa) ou carne DFD (*dark, firm and dry* – escura, dura e seca).

Figura 3 - Consumo de carne no Brasil



Fonte: CGLAC/DAEP/SPA/MAPA, SUEST/SMAE/Embrapa e Departamento de Estatística/UnB

*Observado

Outra cadeia de produção problemática é a de ovos. As aves de postura são confinadas em gaiolas entre 350 e 700 cm²/ave, dependendo das Leis de cada país, o que proporciona uma maior produção, redução de desperdício de ração, diminuição de mão de obra e maior controle da higiene e de doenças. Porém, há uma grande quantidade de aves no mesmo local, falta de locomoção, debicagem, estresse térmico, problemas com empenamento e nos pés, por conta dos arames das gaiolas e restrição de comportamento, que fere as liberdades individuais de cada indivíduo, comprometendo o seu bem-estar e, conseqüentemente na redução da qualidade dos ovos (SILVA; MIRANDA, 2009).

Em aves de corte, onde a expectativa é a redução do tempo de abate, a criação intensiva permite que os animais não necessitem se locomover em busca de alimentos ou para fugir de predadores. Também em relação aos pintinhos, que são independentes desde o nascimento, mas precisam de aquecimento, o amontoado de animais e aquecimentos artificiais cumprem esse papel. Mas, outras necessidades como a de se empoleirar, seleção de alimentos e expressões de comportamentos que seriam naturais são inibidos nesse cenário (FIGUEIRA *et al.*, 2014).

Os confinamentos não são o único local em que o bem-estar é questionado. No pré-abate, acontecem vários fatores que colaboram para a qualidade de vida dos animais e para a boa qualidade da carne. O final da

cadeia produtiva deve atender a proteção dos animais durante o transporte e vem, principalmente, sendo exigidas no mercado externo. O transporte é uma situação de estresse pois, expõe os animais a novos fatores como dificuldade de embarque e desembarque, barulhos, variações de temperatura, trepidações de estradas e acelerações e desacelerações bruscas dos caminhões com pisos fixos, jogando-os uns contra os outros durante o trajeto (LUDTKE *et al.*, 2012). A intervenção humana inadequada, durante o transporte, na intenção de controlar os animais, também pode contribuir para uma situação de estresse por maus tratos (MIRANDA; CARVALHO; THOMÉ, 2013).

O transporte, então, representa grande riscos de perdas quantitativas e qualitativas da carcaça sendo que, as perdas, ocorrem para os produtores, frigoríficos e consumidores, comprometendo a qualidade, apresentando carne DFD e a mudança do pH, devida à restrição alimentar, desidratação, medo e outros fatores estressantes que ocorrem durante o processo de transporte (POLASTRINI *et al.*, 2021). O Brasil apresenta um problema ainda maior de transporte, por ter rodovias precárias, malhas rodoviárias muito extensas gerando vários prejuízos econômicos e de bem-estar aos bovinos transportados. Polastrini *et al.* (2021), realizaram uma pesquisa para analisar as perdas ocasionadas pelo transporte e observaram que a mentalidade dos envolvidos no processo ainda é de que “o animal já vai ser abatido mesmo, então, tanto faz”. Constataram também uma distância superior a 150km entre a fazenda até os frigoríficos, o que não deveria ser permitido pois aumentam a incidência de estresse e lesões nos animais. Também constataram que as lesões ocorrem com maior frequência nos cortes nobres como contrafilé e até casos de mortes, causando prejuízos éticos e econômicos relevantes. Os valores estimados de perdas referentes ao transporte são alarmantes. Polastrini *et al.* (2021), constataram que em janeiro de 2021 o valor da arroba do boi no estado do Tocantins era de R\$ 262,00 e a quantidade de bois abatidos, por ano, gira em torno de um milhão de cabeças. Estimou-se que o prejuízo econômico pode ter ficado na casa dos R\$ 32.383.200,00, levando em consideração o nível máximo, das lesões, hematomas e contusões das carcaças.

Segundo Rodrigues Leite *et al.* (2015), no Brasil, o manejo pré-abate nos frigoríficos é deixado de lado pelos produtores, transportadores e matadouros,

trazendo consequências negativas no ganho dos pecuaristas e das indústrias, além do estresse sofrido pelos animais. Um estudo realizado por Rodrigues Leite *et al.* (2015), de gado de corte já dentro do frigorífico, levou em consideração o estresse dos animais, contusões nas carcaças e alterações químicas, fisiológicas e biológicas da carne, medindo o seu pH 24 horas. Apontou que, entre o curral de restrição e o boxe de atordoamento, houve um número considerado alto de vocalizações, o que significa estresse por escorregões, quedas, falta de iluminação adequada e uso excessivo de bastão de choque. Também apontou escorregões e vocalização no boxe de atordoamento e problemas com a pistola de atordoamento, tendo que ser realizado mais de um disparo por animal e, ainda assim, havia animais sensíveis durante a sangria, o que é inaceitável para um abate humanitário. A maioria dos lotes de animais apresentaram pH 24 horas post mortem acima de 5,80 (que é o considerado fator máximo) resultando em perda da qualidade da carne trazendo prejuízo para os produtores, frigoríficos e consumidores e é um indicador da falta de bem-estar dos animais no pré-abate (LEITE *et al.*, 2015; POLASTRINI *et al.*, 2021).

Percebe-se, que o bem-estar vem sendo considerado em várias pesquisas na área de produção animal, as quais demonstram que o bem-estar dos animais tem relação direta com o aumento de produtividade e redução de custos e perdas nos sistemas de produção.

3.3 Coleta de Fenótipos

O bem-estar animal é complexo e garantir que os animais estejam dentro das cinco liberdades, que estejam sendo tratados de forma que seja mantida a sua integridade física, mental e comportamental acaba sendo subjetiva. Ela é influenciada pelas diferentes visões humanas e por diversos fatores culturais, o que dificulta, ainda mais, o seu entendimento (HÖTZEL; MACHADO FILHO, 2004). Assim, é importante conhecer e respeitar a biologia do animal de produção para que se possa reconhecer os aspectos que evidenciam o desconforto físico e mental através de seus comportamentos (MIRANDA; CARVALHO; THOMÉ, 2013). De qualquer forma, é necessário que se desenvolvam técnicas de diagnósticos para reconhecer os indicadores de bem-estar animal que utilizem respostas fisiológicas e comportamentais e sua

sanidade (BOND *et al.*, 2012). Além disso, Duncan (2005) considera as emoções dos animais como parte desse diagnóstico, embora seja uma abordagem difícil e subjetiva.

3.3.1 Fenótipos como medidas de bem-estar

Um das formas de se obter um diagnóstico é utilizar medidas fisiológicas diretas, ou seja, aquelas que são ativadas pelo Sistema Nervoso Autônomo - SNA, como frequência respiratória e cardíaca que apresentam respostas de curto prazo e são altamente mensuráveis. Por exemplo, animais que são alojados em sombra apresentam menores frequências cardíacas e respiratórias que animais expostos ao sol. Também é possível verificar a concentração de hormônios adrenocorticais como consequência dos estímulos à prática de manejo. O mais comum é a dosagem de cortisol no sangue que se verifica o estresse agudo dos animais que pode também ser colhido nas fezes como uma ação não invasiva, mas o cortisol atinge o trato intestinal em aproximadamente 12 horas e demanda intenso controle sobre o animal a ser analisado (BOND *et al.*, 2012). Outra forma de medir o estado do animal é através da temperatura retal e frequência respiratória para estimar a tolerância ao calor. Em alguns trabalhos foram encontradas também a frequência cardíaca, taxa de sudorese e temperatura da pele. A medição da temperatura do ambiente também é importante para observar se está, ou não, dentro do tolerável para cada espécie de animal (BARBOSA FILHO, 2005).

No caso dos bovinos, a reatividade é um bom indicativo do seu estado, uma vez que se espera que o animal se expresse naturalmente e isso depende do seu estado mental, expressando medo, sofrimento ou prazer. A ausência de doenças e ferimentos, também são grandes indicadores de bem-estar, uma vez que trazem sofrimento ao animal, comprometem sua reprodução, nível de crescimento e, por muitas vezes, levam à morte (CEBALLOS; SANT'ANNA, 2018).

No caso das aves, é medida a temperatura retal e ofegação, que é um mecanismo de perda de calor por evaporação (que é o mesmo utilizado pelos bovinos). Porém, nas aves, esse processo inicia outro que é a alcalose respiratória que acaba por reduzir a espessura da casca dos ovos (SILVA;

MIRANDA, 2009). Já nas aves de corte, o comportamento social é muito importante. As aves que se isolam, apresentam comportamento de medo e angústia que é um indicativo de que não está confortável com o ambiente (FIGUEIRA *et al.*, 2014).

As características comportamentais são muito importantes para medir o bem-estar dos animais e menos invasivas que a coleta de dados fisiológicos, como exames de sangue. Porém, características comportamentais são mais difíceis de serem coletadas por demandarem observações diretas em tempo real, exigindo grande número de pessoas para as análises e apresentarem baixa precisão dos dados coletados (TURCO *et al.*, 2020). Por outro lado, vem surgindo no mercado equipamentos tecnológicos de precisão, que são menos invasivos, não interferem diretamente com os animais, evitando o estresse e coletando informações de comportamento, fisiológico e de desempenho, mas apresentam um custo elevado por serem importados de países mais desenvolvidos (PEREIRA *et al.*, 2015).

3.3.2 Equipamentos de coleta

Com a popularização dos dispositivos móveis, com a utilização de softwares livres e de código aberto, ficou mais fácil, levar a campo, aparelhos que podem coletar, ao mesmo tempo, dados de coordenadas geográficas, fotos, vídeos, áudios, altitude, além de permitirem maior precisão na coleta de informações manuais e atualização em tempo real com servidores ou armazenamento local, offline, com atualização posterior (TURCO *et al.*, 2020).

Outro importante método de coleta destacado por Turco *et al.* (2020) é a bioacústica. A bioacústica tem sido proposta por apresentar baixo custo e não ser invasiva, para a identificação de atividades ruminantes dos animais, de forma contínua, e não afeta o seu comportamento natural (NELSON *et al.*, 2005). Em ovinos os testes obtiveram 82% de precisão, no comportamento ingestivo, sendo possível observar as espécies forrageiras e sua altura. Com esse equipamento também foi possível observar o tempo diário de pastejo e ruminação, dentre outras atividades de bovinos, pois os sons são bastantes distintos e fáceis de serem diagnosticados por programas de computador (TURCO *et al.*, 2020). Outro fator importante, que vem sendo pesquisado, é para estabelecer alguns padrões a fim de identificar alguma situação como

tosse, doenças respiratórias, comportamentos de dor, injúria, e relacionados à amamentação de leitões, por exemplo (BORGES, 2009).

Existem outros dispositivos de monitoramento que são os sensores. Sensores como GPS (*Global Positioning System*) são ótimos para detectar movimentos e posicionamento, pois captam sinais de satélites em todo o mundo e, conseqüentemente, determinam o seu posicionamento. Este dispositivo é excelente para o monitoramento autônomo dos animais (TURCO *et al.*, 2020).

Em outro trabalho (NASCIMENTO *et al.*, 2018), foram fixados acelerômetros na alça que envolve o maxilar do ruminante para analisar o comportamento de vacas leiteiras, que mostrou-se ser um equipamento interessante para o acompanhamento do comportamento ingestivo de bovinos em pastejo, com maior espaço de tempo. Também foram utilizados os acelerômetros em ovelhas, sendo fixados no pescoço, que forneceu dados como movimento de cabeça e tórax com sinais de aceleração e repouso, pastejo e caminhada. Na perna, que obteve os sinais claros de caminhada do pastejo e locomoção e na orelha, que obteve sinais para todos os estados de movimentação. Os estados de comportamentos, principalmente de pastoreio, são muito importantes para mapear os dados individuais com esses equipamentos de alta precisão (TURCO *et al.*, 2020).

Outro tipo de captura de comportamento é utilizar a visão computacional. Neste processo usa-se um conjunto de câmeras para capturar o comportamento do animal e um software para analisar essas imagens, permitindo um monitoramento contínuo e em número maior que a observação direta (ALVES, 2007; BARBOSA FILHO, 2005). Alves, (2007), cita em seu trabalho, que este processo foi utilizado em aves poedeiras e para monitorar o comportamento digestivo de frangos de corte e, a eficiência desse equipamento, foi confirmada para a observação do comportamento das aves. O mesmo processo foi utilizado para análise comportamental de frangos em gaiolas, através de imagens online e totalmente autônomas, analisando um conjunto de parâmetros mensuráveis e capaz de identificar até 15 tipos diferentes de comportamentos (TURCO *et al.*, 2020). A visão computacional também foi utilizada, juntamente com o algoritmo de Viola-Jones, programa de computador que analisa o comportamento de um animal, para a detecção de

comportamento de vacas leiteiras e foi validado por especialistas, sendo capaz de detectar o comportamento das vacas em aproximadamente 92% (PORTO *et al.*, 2013). Outra medida que é essencialmente importante para os sistemas de criação é o ganho de peso diário, que demanda muito tempo e uma grande quantidade de pessoas. Mortensen *et al.* (2016), criaram de captura de dados com câmera 3D Kinect, onde foi possível estimar o peso de frangos de corte individualmente. O sensor Kinect também foi utilizado para medidas de tórax e do abdômen, altura e área dorsal e comprimento do corpo de bovinos de corte em abatedouro, estimando peso corporal e de carcaça de Angus e Nelore (GOMES *et al.*, 2016). Segundo Barbosa Filho (2005), esses sistemas de coleta de imagens vêm sendo utilizados para a criação de aves, suínos e bovinos e que na criação de leitões, juntamente com a utilização de piso térmico, possibilitou acompanhar a distribuição dos animais, capturar um grande número de dados para a geração das informações para correlação dos perfis.

A temperatura também é um passo importante para identificar dados fisiológicos e avaliar o bem-estar dos indivíduos e pode detectar doenças, estresses e ovulação. São utilizados termômetros de mercúrio no reto do animal que, além de invasivo, o animal precisa ser contido para a realização desse procedimento. Estudos de tecnologia utilizaram a temperatura vaginal transmitida por radiofrequência e observou-se que o cio foi facilmente detectado por esse processo, sendo que as temperaturas foram mensuradas continuamente, porém, não em tempo real. O dispositivo foi acoplado no útero das vacas e comparado com termômetros digitais de alto desempenho verificando que o sistema é confiável. Outros tipos de coleta de temperatura foram as timpânicas e no rúmen do animal. As temperaturas medidas com os três sensores em um único animal obtiveram temperaturas de 39.0, 38.4, e 38.2 °C, que embora apresentem diferenças, essas não foram significativas. A forma não invasiva que permite obter a temperatura corporal do animal é a termografia e esta tecnologia vem sendo usada para mensurar fluxo sanguíneo resultantes de aumento da temperatura corporal e situações e condições estressantes (TURCO *et al.*, 2020).

Um outro equipamento interessante que Brainer *et al.* (2022) relatam, é a ordenha robotizada, A ordenha, além de automatizar todo o trabalho que

seria feito por humanos, livra o animal do estresse, faz a higienização, estimulação e ordenha do úbere da vaca. Além disso, através de uma etiqueta de identificação do animal, o robô faz a adequação do equipamento de acordo com as características morfológicas da vaca, que ajusta a altura, tamanho do úbere e formato e ângulo do teto, sem nenhuma interferência humana. Durante o processo, coleta-se uma série de dados, gera-se relatórios com informações sobre a saúde do animal, saúde do úbere, peso da vaca, temperatura, ordenha, atividades e ruminação, além de indicar o momento exato de se realizar a inseminação. Embora seja um investimento caro, o robô consegue trazer muitos benefícios como aumento da produtividade, condições de saúde e bem-estar dos animais (BRAINER *et al.*, 2022).

3.4 Simulação de Dados

3.4.1 Visão geral

Um simulador de dados, tem por objetivo, otimizar as pesquisas se tornando viável pela economia de tempo e redução dos custos e, visa reproduzir comportamentos de sistemas reais e dar suporte a estratégias e tomadas de decisões (JUNIOR *et al.*, 2007; VERNIZ *et al.*, 2009). É muito utilizado em melhoramento genético animal e vegetal (PÉREZ-ENCISO; LEGARRA, 2016) para previsão de mudanças climáticas (ALMEIDA; OLIVEIRA, 2010), impactos geológicos, infiltração de água no solo (COSTA *et al.*, 2021), simulação de arquiteturas de computadores em fase de desenvolvimento (JUNIOR *et al.*, 2007), na aviação civil para treinamento de pilotos e análises de desastres aéreos (ALMEIDA; CORREA, 2017), na área médica para testes de tratamentos e, mais recentemente, nos Metaversos onde os usuários são imersos em realidades virtuais. Afirmam ainda, (WEN *et al.*, 2018), que sistemas de simulação, são boas alternativas para redução de custos e impactos em modelos consistentes com a realidade.

Ainda, segundo Almeida e Oliveira (2010), os simuladores precisam utilizar dados reais, do que se deseja simular, para que sejam gerados dados sintéticos o mais próximo da vida real. As simulações computacionais são compostas por vários algoritmos e modelos matemáticos e envolve um estudo

detalhado da área desejada que, além de proporcionarem resultados muito próximos dos reais, não ficam presas às suas limitações físicas.

De forma geral, os simuladores de dados científicos, além de gerarem dados confiáveis, suas principais vantagens são a redução do tempo em que os eventos acontecem, custo e evitam limitações éticas e ambientais (BARIONI; MANCINI, 2014). No caso do melhoramento genético animal, por exemplo, para serem analisados dados de uma fazenda, de criação de gado de corte, são necessárias várias gerações de indivíduos, o que demanda muito tempo, altos investimentos, pessoal especializado e grande número de animais para que os dados sejam adequados.

Wen *et al.* (2018) e Bessa Júnior e Setti (2010), criaram simuladores de trânsito, direcionados para a mobilidade urbana. Ambos buscaram uma solução boa, Wen *et al.* (2018), utilizaram algoritmos tradicionais baseados em números aleatórios e Bessa Júnior e Setti (2010), utilizaram parâmetros de calibração baseados em um algoritmo genético, que é capaz de evoluir, em busca de uma solução ótima. É um processo de quatro etapas envolvendo avaliação, reprodução, recombinação e mutação (FRENZEL, 1993).

Seja qual for a área em que se deseja realizar as simulações, é necessário que seja selecionado um bom paradigma de modelagem, para que possa ser programado, e escolher corretamente os métodos matemáticos e algoritmos para a solução do problema proposto (BARIONI; MANCINI, 2014). É necessária a validação dos dados simulados para analisar, o quão preciso foi o método, e determinar se as suposições são confiáveis ou não (PENG *et al.*, 2015).

3.4.2 Simulação de dados na área genética

Simulações também são utilizadas para gerar grandes volumes de dados complexos para modelagem de evolução dos genomas e análises de expansões demográficas, mutações genéticas, seleção natural e recombinações. Através de métodos e associações genótipo-fenótipo, é possível detectar, nos genes, predisposição de doenças associadas às características humanas (PENG *et al.*, 2015). Isso é possível ao simular um grande número de dados com diferentes configurações de parâmetros para a

produção de amostras, que possam corresponder aos dados observados empiricamente. Entretanto, os algoritmos variam de acordo com a sua aplicação e metodologia de dados genéticos da população que se deseja obter. As simulações de história evolutiva de populações são realizadas basicamente por dois métodos: o método *backward-time* é utilizado para criar dados de acordo com a teoria coalescente de genética populacional; e o *forward-time* que cria dados de evolução de populações e extrai amostras dessas populações simuladas. Já as simulações em que as evoluções de populações não são modeladas, são utilizados os métodos teóricos de reamostragem e de *gene-dropping* (PENG *et al.*, 2015).

3.4.3 Simulação de dados por teoria Coalescente

A simulação pela teoria coalescente é uma árvore binária onde cada nó representa um encontro de linhagens ancestrais (KIM; WIEHE, 2008). Ela permite a reconstrução de eventos evolutivos como fluxo gênico, seleção natural e divergência populacional através de fragmentos do DNA (ácido desoxirribonucleico), caminhando do presente para o passado, permitindo a inferência do processo evolutivo através de modelagem probabilísticas, determinando a distância entre uma amostra da população contemporânea e seu ancestral comum mais recente. É um modelo preditivo e faz uso de dados genéticos para aplicações filogenéticas (SIGWART, 2005). O fato de ocorrerem mutações genéticas positivas, o alelo favorável pode se espalhar rapidamente, em toda população, em um curto espaço de tempo, o que sugere que, possa ocorrer um processo coalescente com múltiplas colisões (DURRETT; SCHWEINSBERG, 2005). Na figura 4, temos um exemplo de árvore binária da teoria coalescente onde, em 16 gerações, é encontrado um ancestral em comum para três indivíduos da geração atual.

Liang *et al.* (2007), utilizaram a teoria coalescente para criar um simulador genético, gerando uma população onde, a maioria dos indivíduos, compartilhavam trechos, haplótipos, de uma determinada doença. Para isto, utilizaram uma sequência *backward-time* onde a simulação inicia na geração mais nova e vai criando as gerações mais velhas, uma por vez. Cada aresta,

em que o comprimento da região simulada cresce. A teoria coalescente a ferramenta mais eficaz quando se busca simulações genéticas populacionais que visam inferir a evolução genética histórica com modelos relativamente simples como a seleção direcional em um loco em contextos demográficos não muito complexos, mas podem, com algumas adaptações, gerar padrões com múltiplos locos com uma velocidade razoável (KIM; WIEHE, 2008).

Segundo Carvajal-Rodriguez (2008), nos últimos anos, vem surgindo algoritmos mais eficientes de teoria coalescente, com o uso de estratégias combinadas de *backward-time* e *forward-time*. Porém, os algoritmos que se utilizam dos recursos para “frente no tempo” são muito mais lentos e exigem muito mais memória que as simulações coalescentes tradicionais (KIM; WIEHE, 2008). No entanto, Carvajal-Rodriguez (2008), detalham um processo da teoria coalescente que utiliza *forward-time* em três processos diferentes que são: entrada, evolução e saída, sendo que os dados de saída precisam ser semelhantes aos da entrada, com informações genéticas bialélicas, para a representação dos locos e método de validação eficiente na escala de megabases.

Durret e Schweinsberg (2005), trabalharam com processos de genealogia coalescente quando a população é afetada por uma ou várias mutações, benéficas, repetidas ao longo do tempo e a genealogia de amostragem por processo Markoviano de tempo contínuo. Quando ocorre uma mutação benéfica, em local diferente da área do gene de interesse, se pode rastrear o indivíduo pela mutação o que torna o processo mais fácil, enquanto o que não teve mutação, tem que rastrear seu gene do local de interesse e encontrar um ancestral diferente por causa da recombinação entre os locais do cromossomo.

3.4.4 Simulação de dados por *Gene-dropping*

Outra forma de simular dados é utilizar algoritmos *gene-dropping*. Esse método consiste em gerar a progênie partindo dos pais e, ao contrário do coalescente, são mais adequados para tratar eventos recentes de alta complexidade, pois mantém completa o controle de informações dos ancestrais, dando uma área de aplicação mais ampla, inclusive dos processos

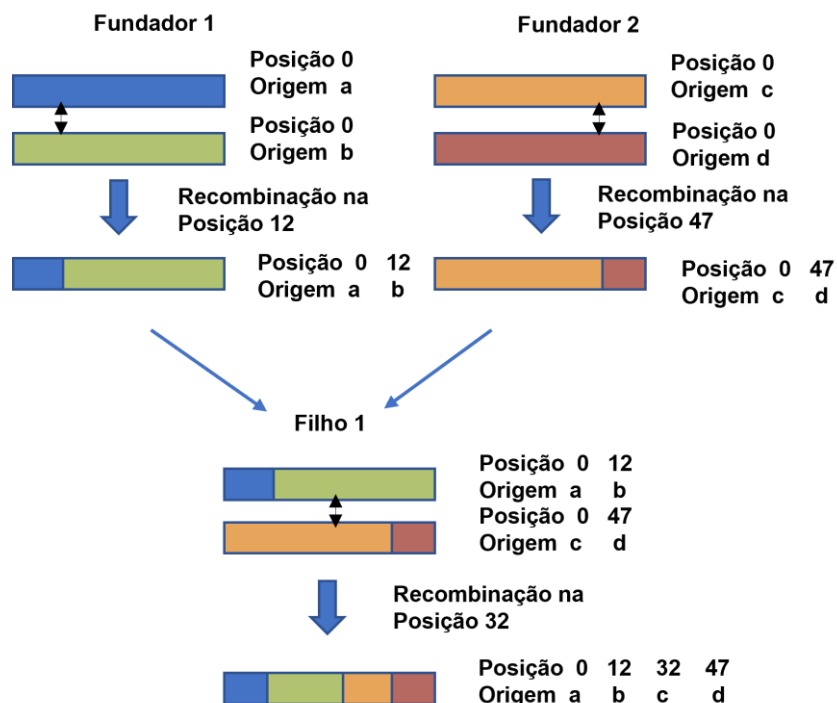
evolutivos. Além disso, as simulações em *forward-time*, permitem simular características genéticas, efeitos de ambiente e que sejam realizados acasalamentos por seleção, o que não são possíveis de serem realizados por coalescência, por serem baseados no modelo de *Wright-Fisher* (PENG *et al.*, 2012).

Embora a simulação genética possa gerar várias características genotípicas e efeitos de ambiente, não consegue contemplar todas as variáveis, pois a arquitetura real não é totalmente conhecida e isso pode influenciar nos resultados finais. Outro fator importante é a história demográfica, pois é ela quem traça a distribuição de polimorfismos, nucleotídeos e desequilíbrio da raça (PÉREZ-ENCISO; LEGARRA, 2016). (Pérez-Enciso e Legarra (2016), desenvolveram um simulador, chamado ms2gs, que combina algoritmos de coalescência e *gene-dropping*, e leva em consideração um curto espaço de tempo quando se tem mutações não significativas sendo muito favorável para programas de melhoramento genético animal e vegetal, além de ser flexível para inserção e definição de QTLs (*Quantitative Trait Locus*) e SNPs (*Single nucleotide polymorphism*), e gerar polimorfismos moleculares por meio de coalescência. Comparado a outros simuladores que combinam os dois métodos, coalescência e *gene-dropping*, o ms2gs apresentou vantagens como simulação de características complexas, flexibilidade para modelar QTLs e uma forma mais realista de verificação de erros, porém ainda apresenta lentidão quando da simulação de genomas completos. Uma abordagem mais detalhada do algoritmos utilizado e parâmetros para simulação pode ser encontrada em (PÉREZ-ENCISO; LEGARRA, 2016). Na figura 5, apresenta-se uma ilustração de como funciona a simulação por *gene-dropping*, onde dois fundadores, são inicialmente cruzados e seus genes sofrem recombinação e são passados para a próxima geração e, assim, sucessivamente.

Manatrinon *et al.* (2009), utilizaram algoritmos de *gene-dropping* para identificação de doenças genéticas autossômicas recessivas letais em uma população com pedigree complexos, de bovinos Pardo Suíços, criados na Áustria, utilizou como população fundadora, dados de 1.830.125 indivíduos e

estimou as frequências alélicas para demonstrar a propagação dos alelos letais.

Figura 5 - Exemplo de simulação por Gene-dropping



Fonte: CHENG *et al.* (2015)

Honda *et al.* (2002), usaram *gene-dropping* com população de bovinos Wagyu, para simular a análise de sobrevivências de determinados alelos presentes na população fundadora. O processo foi replicado 10.000 vezes para estimar a distribuição das frequências alélicas dos animais fundadores e a progênie, simulada, utilizou segregação mendeliana. A partir disso, estimou-se a probabilidade da extinção e da sobrevivência de alelos em uma frequência criticamente baixa.

Ainda, a simulação por *gene-dropping* pode ser utilizada para calcular as proporções dos alelos perdidos, a probabilidade de extinção dos alelos fundadores e verificação de possíveis gargalos na diversidade genética (MELKA *et al.*, 2013).

4 MATERIAL E MÉTODOS

4.1 Hardware

Para o desenvolvimento e implementação das rotinas e simulação dos dados, foram utilizados dois servidores, sendo:

Servidor 1 – Patrick

Intel Core i7 3930K CPU 3.2Ghz

12 núcleos

64GB de memória RAM

HD 2 TB

Servidor 2 – Gary

Intel Core i7 4930K CPU 3.4Ghz

12 núcleos

48GB de memória RAM

HD 2 TB

4.2 Sistema Operacional

Devido à estabilidade, segurança, ser amplamente utilizado no meio acadêmico científico e ser software-livre, o sistema operacional Linux Ubuntu Server versão 20.04.2 LTS, foi utilizado como padrão nos servidores em que ocorreram o desenvolvimento e a simulação dos dados. Entretanto, nas estações do laboratório ou computadores pessoais, foram utilizados outros sistemas operacionais, e as ferramentas Powershell do Windows, Windows Subsystem for Linux (WSL) versão Ubuntu, e acesso os servidores por SSH (*Secure Socket Shell*), interna ou externamente à unidade, e acesso aos arquivos por meio do serviço SAMBA, apenas no laboratório.

4.3 Linguagens de Programação, Compiladores e Softwares Estatísticos

Houve a necessidade de criar as rotinas na linguagem R versão 3.6.3 Release 2020-02-29, que foi instalada no servidor juntamente com o RStudio que provê uma interface amigável que pode ser acessada de qualquer computador através de um browser de internet.

O R é uma linguagem multiplataforma, totalmente voltada para manipulação, análise e visualização de dados, sendo possível a integração de outras linguagens, além de prover recursos para desenvolvimento de funções, scripts e pacotes, que podem ser distribuídos para a comunidade científica.

Optou-se por desenvolver pacotes e utilizar a linguagem do R, uma vez que é amplamente utilizada, gratuitamente, pela comunidade acadêmica para a realização de experimentos e análises genéticas.

Inicialmente a geração e o armazenamento dos dados ficaram em memória RAM, uma vez que as simulações foram realizadas em pequena escala, não sendo necessário o armazenamento em banco de dados.

5 DESENVOLVIMENTO

5.1 Simulação de Valores Genéticos Aditivos

O primeiro passo para simular os dados é a geração dos valores genéticos aditivos para cada indivíduo da população, geralmente por *gene-dropping*, pois parte-se de uma população base e as progênies são simuladas a partir do acasalamento dessa primeira geração. Pires et. al (2007) criaram um simulador que gera o valor genético aditivo para determinada característica com base em locos. O cálculo é feito da seguinte forma:

$$a = \frac{1}{\sqrt{2pql}} \quad (1)$$

em que,

a = valor genotípico, do loco A_1A_1 , para a característica, sendo que, o valor genotípico para A_1A_2 é 0 e para A_2A_2 , é $-a$.

p e q = são respectivamente, as frequências dos alelos A_1 e A_2 para a característica;

l = número de locos para a característica.

Assim, o valor genético aditivo, de cada indivíduo para a característica, é dado pela seguinte equação:

$$A = \sum_{i=1}^l a_i \quad (2)$$

5.2 Simulação de Valores de Ambiente Correlacionados

O segundo passo é simular os valores de ambiente. Porém, a busca de um método eficiente para a simulação de características longitudinais é complexa pois, a literatura nos traz que, o dado coletado de um indivíduo, no tempo t_0 , está correlacionado com os dados coletados nos tempos $t_1, t_2, t_3, \dots, t_n$, do mesmo indivíduo, sendo que, a correlação entre o dado t_0 do dado t_n , vai se aproximando de zero ao longo do tempo. Para tanto, é necessário que o pesquisador forneça uma matriz de covariâncias para que essas correlações sejam calculadas. É mais conveniente para a representação a seguir, considerar as respostas medidas em cada indivíduo durante o tempo

em conjunto, de forma que as relações existentes entre elas possam ser claramente expressas e resumidas. Dessa forma, considera-se as respostas em cada momento como uma variável aleatória, as quais serão armazenadas, para cada indivíduo em um vetor:

Y_{ij} = j – ésimo registro do indivíduo i

em que,

$i = 1 \dots m$, indexa os indivíduos e $j = 1 \dots n_i$, indexa as medidas repetidas no i – ésimo indivíduo.

n_i = número de medidas repetidas do indivíduo i , o qual pode ser diferente para diferentes indivíduos.

Define-se também:

t_{ij} = tempo no qual Y_{ij} é registrado para o i – ésimo indivíduo, o que permite considerar diferentes tempos para diferentes indivíduos.

Pode-se então representar o vetor de medidas repetidas para o i – ésimo indivíduo como:

$$Y_i = [Y_{i1}, \dots, Y_{in}] \quad (3)$$

Para o seguinte desenvolvimento, assume-se que cada Y_{ij} é normalmente distribuído e, portanto, Y_i tem distribuição normal multivariada.

Nos estudos que envolvem medidas repetidas, normalmente assume-se que os vetores aleatórios Y_i são independentes. Entretanto, para os estudos em melhoramento animal essa pressuposição não é razoável, uma vez que existe parentesco entre os indivíduos, gerando uma correlação genética que causará o aparecimento de uma correlação fenotípica entre os indivíduos da população, e conseqüentemente, entre os Y_i .

Naturalmente, espera-se correlações entre as respostas dentro de um mesmo indivíduo, ou seja, os Y_{ij} não são independentes, sendo que quanto mais próximos no tempo são os dados, mais forte é a correlação entre eles.

Considere o modelo:

$$Y_{ij} = \mu + e_{ij} \quad (4)$$

em que,

μ = constante inerente a média geral da característica;

e_{ij} = valor dos efeitos residuais que incidem sobre o i – *ésimo* indivíduo no momento j .

Uma vez que, os efeitos representados e_{ij} não são controlados, considera-se que ocorram aleatoriamente, assumindo valores positivos e negativos e, portanto,

$$E(e_{ij}) = 0 \quad (5)$$

Uma vez que μ é uma constante para cada Y_{ij} , as correlações entre Y_{ij} e Y_{ij}' são determinadas pelas relações entre e_{ij} e e_{ij}' .

A representação dessas relações pode ser vista como:

E = matriz de correlações, onde na diagonal tem-se números 1's e fora das diagonais $e_{ij}e_{ij}'$.

Considere agora o modelo:

$$P_{ij} = \mu + A_i + C_i + E_{ij}, \quad (6)$$

em que,

P_{ij} = valor fenotípico da característica do i – *ésimo* indivíduo no tempo j ;

μ = constante inerente a média geral da característica;

A_i = valor genético aditivo do i – *ésimo* indivíduo;

C_i = efeito dos alelos do i – *ésimo* indivíduo;

E_{ij} = efeito global do ambiente que incide sobre o i – *ésimo* indivíduo no tempo j .

Com medidas repetidas, costuma-se decompor o valor de E_{ij} :

$$E_{ij} = E_{p_i} + E_{t_{ij}}, \quad (7)$$

em que,

E_{p_i} = efeito permanente de ambiente sobre o i – *ésimo* indivíduo;

$E_{t_{ij}}$ = efeito do ambiente temporário (não controlado) sobre o i – *ésimo* indivíduo no tempo j .

O efeito de E_p é o mesmo para o indivíduo em seus vários registros, bem como o efeito de A , C e μ , não sendo, portanto, causa de correlação genética

entre as medidas repetidas. A correlação entre registros consecutivos é causada pelas relações entre $E_{t_{ij}}$ e $E_{t_{ij}}'$.

As relações entre as variáveis de ambiente temporário em diferentes tempos podem ser representadas de forma análoga à equação 8.

Pela equação 6, vemos que os valores de P_{ij} (ou Y_{ij}) serão simulados pela soma dos efeitos, também simulados de A_i , C_i , Ep_i e E_{ij} .

Para que os valores de E_{ij} sejam simulados de acordo com as relações (covariâncias) apresentadas em,

$$cov(E) = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \cdots & \sigma_{rr}^2 \end{bmatrix} \quad (8)$$

em que,

σ_{rr}^2 = variância da r – ésima medida repetida;

σ_{ir} = covariância entre as i – ésima e r – ésima medidas;

As covariâncias $cov(E_{ij}, E_{ij}')$, devem ser fornecidas pelo usuário, enquanto as variâncias, são calculadas a partir da herdabilidade e variância fenotípica fornecidas.

Tendo-se $cov(E_{t_{ij}})$, pode-se, por meio de decomposição de Cholesky, simular os valores de Et que vão reproduzir as correlações demonstradas na tabela 2.

Por questões de simplificação, considera-se que os efeitos de combinações alélicas são desprezíveis, bem como os efeitos de ambiente permanente (E_p) e, portanto:

$$P = \mu + A + E_t \quad (9)$$

Para a simulação dos valores de P para a população, deve-se simular os valores de A e E_t . Para que os valores de A sejam simulados é necessário a entrada de várias informações, como, por exemplo, a herdabilidade, número de genes que determinam a característica e variância fenotípica.

A utilização da informação de herdabilidade nos permite inferir sobre a variância de A e E_t .

Entrando com a herdabilidade h^2 e a variância fenotípica $varP$ (BEIGUELMAN, 2002), tem-se que:

$$varA = h^2 varP \quad (10)$$

mas,

$$varP = varA + varE_t, \quad (11)$$

e, portanto,

$$varE_t = varP - varA \quad (12)$$

em que,

h^2 = herdabilidade;

$varA$ = variância genotípica;

$varP$ = variância fenotípica;

$varE_t$ = variância do efeito de ambiente.

5.3 Decomposição Cholesky

5.3.1 Visão geral

Tendo-se $varE_t$ e as correlações entre registros repetidos, entrada pelo usuário, então pode-se, por meio de decomposição de Cholesky, simular os valores de E_t que vão reproduzir as correlações. Para a simulação de dados correlacionados, a simulação de Cholesky é um método bastante utilizado. Ele permite que a partir da geração de variáveis independentes se obtenha um conjunto de variáveis correlacionadas.

Uma aplicação imediata da decomposição de Cholesky é na simulação de dados longitudinais e medidas repetidas no tempo. Para esse tipo de informação assume-se que existe uma estrutura de correlação entre os registros coletados em diferentes períodos no mesmo indivíduo/unidade experimental, além disso, quanto maior o intervalo de tempo menor deve ser a correlação entre as medidas.

Por exemplo, a simulação de dados para estudos em Melhoramento Genético Animal, considera que medidas repetidas, para uma característica qualquer, podem ser representadas pelo seguinte modelo:

$$P = \mu + A + C + E_p + E_t \quad (13)$$

em que,

P = valor fenotípico do indivíduo para uma medida repetida;

μ = média da característica;

A = valor genético do indivíduo para a característica sendo medida;

C = efeito das combinações alélicas do indivíduo para a característica;

E_p = efeito do ambiente permanente do indivíduo para a característica;

E_t = efeito do ambiente temporário do indivíduo sobre a característica no momento da coleta.

μ, A, C e E_p são efeitos constantes durante toda a vida produtiva do indivíduo, entretanto, os valores de E_t podem se alterar a cada coleta realizada. Portanto, se existe uma estrutura de correlação entre as medidas repetidas do indivíduo, ela é devida a dependência existente entre os efeitos de E_t em diferentes momentos.

Considerando, portanto, a matriz de (co)variâncias entre os efeitos de ambiente temporário como:

$$\text{cov}(T) = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \cdots & \sigma_{rr}^2 \end{bmatrix} \quad (14)$$

em que,

σ_{rr}^2 = variância da r – ésima medida repetida;

σ_{ir} = covariância entre as i – ésima e r – ésima medidas;

O vetor de variáveis aleatórias correlacionadas representando os E_t para as medidas repetidas:

$$E_t = \begin{bmatrix} E_{t_1} \\ E_{t_2} \\ \vdots \\ E_{t_r} \end{bmatrix} \quad (15)$$

em que,

E_{t_r} = efeito da variável aleatória de ambiente temporário para a r – ésima medida no tempo.

Assim, o objetivo é simular valores de E_t para cada medida repetida realizada para cada indivíduo da população de forma a reproduzir a estrutura de covariâncias T .

Uma maneira eficiente de simular esses valores é por meio de:

$$E_t = Lz \quad (16)$$

em que,

L = matriz triangular inferior resultante da Decomposição de Cholesky em T ;

z = vetor de variáveis aleatórias independentes e com variância unitária,

representadas matricialmente conforme:

$$E_t = \begin{cases} et_1 = L_{11}z_1 \\ et_2 = L_{21}z_1 + L_{22}z_2 \\ et_3 = L_{31}z_1 + L_{32}z_2 + L_{33}z_3 \\ et_4 = L_{41}z_1 + L_{42}z_2 + L_{43}z_3 + L_{44}z_4 \\ \vdots \\ et_t = L_{t1}z_1 + L_{t2}z_2 + L_{t3}z_3 + L_{t4}z_4 + \dots + L_{tt}z_t \end{cases} \quad (17)$$

ou seja,

E_t , é a matriz resultante da multiplicação da triangular inferior L por z .

5.3.2 Algoritmo da decomposição de Cholesky

Embora o R já tenha implementado, em suas bibliotecas, um algoritmo da decomposição de Cholesky, foi desenvolvida uma nova função para se realizar o cálculo da complexidade.

Para a decomposição de Cholesky os seguintes fatos devem ser verdadeiros:

Teorema 1: A matriz A é positiva definida $x^T Ax > 0, \forall x \neq 0$

Teorema 2: A matriz A deve ser simétrica, ou seja, existem uma matriz K , triangular superior com diagonal positiva tal que $A = KK^T$

Assume-se que, para a decomposição de Cholesky as seguintes equações:

$$L_{ii} = \sqrt{A_{ii}}, \text{ para todo } i = 1 \quad (18)$$

em que,

T = matriz de entrada simétrica e positiva definida;

L = triangular inferior da matriz resultado da decomposição de Cholesky.

Para a diagonal principal, quando o i for diferente de 1

$$L_{ii} = \sqrt{T_{ii} - \sum_{k=1}^{i-1} L_{ik}^2} \quad (19)$$

Para todo o elemento cujo $j = 1$,

$$L_{ij} = \frac{T_{ij}}{L_{jj}} \quad (20)$$

E, para todo elemento que não for da diagonal principal e não foi contemplado pelos algoritmos anteriores,

$$L_{ij} = \frac{1}{L_{jj}} \left(T_{ij} - \sum_{k=1}^j L_{ik} L_{jk} \right) \quad (21)$$

5.3.3 Complexidade de algoritmo

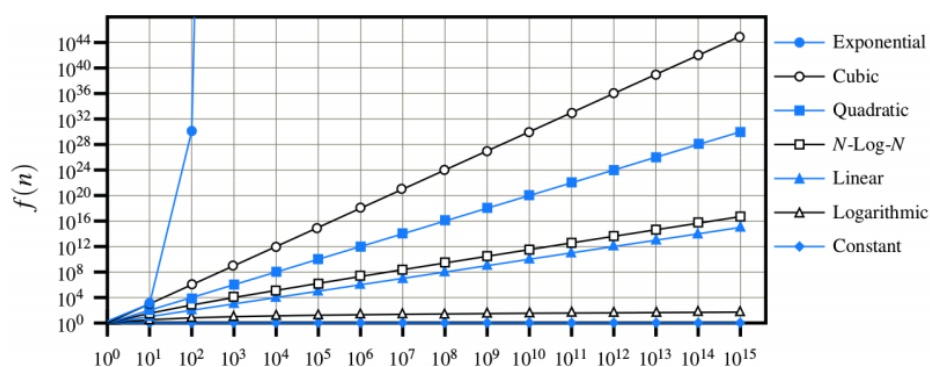
A complexidade de algoritmos é uma medida da eficiência do algoritmo em termos do tempo e espaço necessário para resolver um determinado problema. É uma medida importante para a avaliação de algoritmos e é comumente usada para comparar diferentes soluções para o mesmo problema (BEIGUELMAN, 2002).

Ainda, segundo Beiguelman (2002), complexidade de tempo de um algoritmo refere-se à quantidade de tempo que ele leva para executar em relação ao tamanho da entrada do problema. Geralmente, é medida a complexidade de tempo em termos do número de operações que o algoritmo executa em relação ao tamanho da entrada. Se um algoritmo realiza um *loop* que executa n vezes, dizemos que sua complexidade de tempo é $O(n)$.

Além da complexidade de tempo, também temos a complexidade de espaço, que mede a quantidade de memória necessária para executar o algoritmo em relação ao tamanho da entrada. Essa medida é especialmente importante para algoritmos que manipulam grandes quantidades de dados ou operam em sistemas com recursos limitados e geralmente é representada como $f(n)$ (BEIGUELMAN, 2002; CORMEN *et al.*, 2012), como mostrada na figura 6.

A complexidade de algoritmos pode ser representada por diferentes notações, como $O(n)$, $\Omega(n)$ e $\Theta(n)$, que representam diferentes limites inferiores, superiores e iguais para a complexidade de tempo (CORMEN *et al.*, 2012). Essas notações permitem avaliar a eficiência relativa de diferentes algoritmos para um determinado problema e escolher a melhor opção para a solução do problema, sendo uma medida importante para a avaliação da eficiência e desempenho (BEIGUELMAN, 2002).

Figura 6 - Níveis de complexidade



Fonte: StackOverflow, 2018²

Nas sessões 5.3.4, 5.3.5 e 5.3.6, são apresentadas as complexidades dos três principais algoritmos para a geração dos dados longitudinais correlacionados e suas complexidades.

5.3.4 Complexidade do algoritmo da decomposição de Cholesky

Algoritmo	Custo da linha	Custo total
$L_{ij} \leftarrow 0 \forall i \in 0..n, j \in 0..m$	$O(n * m)$	$O(n^2)$
for $i \leftarrow 1$ to n do	$O(n)$	$O(n)$
for $j \leftarrow 1$ to i do	$O(i)$	$n * O(i)$
$sum \leftarrow 0;$	1	$n * O(i)$
for $k \leftarrow 1$ to j do	$O(j)$	$n * O(i) * O(j)$
$sum \leftarrow sum + L_{ik} * L_{jk}$	1	$n * O(i) * O(j)$

² Disponível em <<https://pt.stackoverflow.com/questions/331506/fun%C3%A7%C3%A3o-de-complexidade>>, Acesso em 20 fev. 2023.

if $i == j$ then	1	$n * O(i)$
$L_{ii} \leftarrow (T_{ii} - sum)^{0.5}$	1	$n * O(i)$
else	1	$n * O(i)$
$L_{ij} \leftarrow 1.0/L_{ij} * (T_{ij} - sum)$	1	$n * O(i)$

Como $O(i)$ e $O(j)$ equivalem a $O(n)$, no pior caso, pode-se substituí-los por $O(n)$, obtendo a seguinte equação para o custo total:

$$T(n) = n + 7n^2 + 2n^3 \quad (22)$$

Como se tem uma função polinomial, segundo (CORMEN *et al.*, 2012), pode-se assumir que o fator dominante da equação é a complexidade deste algoritmo, ou seja, $O(n^3)$.

5.3.5 Complexidade do algoritmo da de cálculo dos efeitos de ambiente temporário correlacionados

Algoritmo	Custo da linha	Custo total
$et_{ij} \leftarrow 0 \forall i \in 0..ncol(L), j \in 0..nrow(L)$	$O(c * r)$	$O(c) * O(r)$
for $i \leftarrow 1$ to $nrow(L)$ do	$O(i)$	$O(i)$
$val \leftarrow 0;$	1	$O(i)$
$j \leftarrow 1;$	1	$O(i)$
while $j \leq i$ do	$O(j)$	$O(i) * O(j)$
$val \leftarrow val + L_{ij} * z_j$	1	$O(i) * O(j)$
$j \leftarrow j + 1$	1	$O(i) * O(j)$
$et_{,j} \leftarrow val$	1	$O(i)$

Como $O(i)$, $O(j)$, $O(c)$ e $O(r)$ equivalem a $O(n)$, no pior caso, pode-se substituí-los por $O(n)$, obtendo a seguinte equação para o custo total:

$$T(n) = 4n + 4n^2 \quad (23)$$

Como se tem uma função polinomial, segundo (CORMEN *et al.*, 2012), pode-se assumir que o fator dominante da equação é a complexidade deste algoritmo, ou seja, $O(n^2)$.

5.3.6 Complexidade do algoritmo para criar a matriz T

Algoritmo	Custo da linha	Custo total
$R_{ij} \leftarrow 0 \forall i \in 0..n, j \in 0..m$	$O(n * m)$	$O(n) * O(m)$
for $i \leftarrow 1$ to n do	$O(n)$	$O(n)$
for $j \leftarrow n - 1$ to n do	$O(n)$	$n * O(n)$
$x \leftarrow rnorm(n - i, mean = 0)$	1	$n * O(n)$
$y \leftarrow pnorm(x, lower.tail = true)$	1	$n * O(n)$
$y \leftarrow sort(y, decreasing = true)$	1	$n * O(n)$
if $size(y) <> 0$ then	1	$n * O(n)$
for $t \leftarrow 1$ to $size(y)$ do	$O(y)$	$y * O(n) * O(n)$
$R_{it+i} \leftarrow y_t$	1	$y * O(n) * O(n)$
$R_{it+i} \leftarrow y_t$	1	$y * O(n) * O(n)$
$S \leftarrow random(n, min, max)$	1	$O(1)$
$diagonal(R) \leftarrow S$	1	$O(1)$

Como $O(i)$, $O(j)$, e $O(y)$ equivalem a $O(n)$, no pior caso, pode-se substituí-los por $O(n)$, obtendo a seguinte equação para o custo total:

$$T(n) = 2 + 1n + 6n^2 + 3n^3 \quad (24)$$

Como se tem uma função polinomial, segundo (CORMEN *et al.*, 2012), pode-se assumir que o fator dominante da equação é a complexidade deste algoritmo, ou seja, $O(n^3)$.

5.3.7 Algoritmo para simulação de efeitos de ambiente temporários correlacionados

A simulação de dados longitudinais envolve processos complexos que necessitam de atenção e conhecimento para que sejam calculados. Dessa forma, o algoritmo completo resultou nos seguintes passos:

1. Para cada indivíduo faça:
 - 1.1. Determine a quantidade de cromossomos;
 - 1.2. Determine a quantidade de SNPs;
 - 1.3. Determine a quantidade de QTLs

- 1.4. Determine o valor genético aditivo de cada loco, de acordo com ilustrado na equação 1;
- 1.5. Armazene os efeitos genotípicos de a , 0 e $-a$, em um vetor;
- 1.6. Para cada loco:
 - 1.6.1. Determine o tipo do loco, se poligênico, SNP ou QTL;
 - 1.6.2. Determine o valor genotípico do loco;
 - 1.6.3. Some os valores genotípicos de cada loco para obter o valor genotípico do indivíduo e armazene em a_i , conforme demonstrado na equação 2;
- 1.7. Crie a matriz T , de covariância dos efeitos, que deverá ser simétrica e positiva definida, gerada pela função *randomCov* no R;
- 1.8. Aplique a decomposição de Cholesky sobre a matriz T , para obter a matriz LL' ;
- 1.9. Gere uma matriz de variáveis aleatórias independentes e armazene em z ;
- 1.10. Multiplique L_{ij} pelo vetor de distribuição normal z_i para obter e_{ij} ;
- 1.11. Aplique o fenótipo para o indivíduo pela equação 9.

O cálculo da complexidade total do algoritmo foi de:

$$T(n) = 8 + 2n + 3n^2 + 3n^3 \quad (25)$$

Da mesma forma que os três principais algoritmos, chegou-se a uma função polinomial, em que, segundo Cormen *et al.* (2012), pode-se dizer que complexidade, para a geração dos dados longitudinais para efeito de ambiente, deste algoritmo, é $O(n^3)$. A complexidade completa dos algoritmos pode ser vista nos anexos deste documento.

Nessa simulação, criou-se uma população de 1000 indivíduos, sendo 512 machos e 488 fêmeas que resultou na variância genotípica da população de 197,68.

6 RESULTADOS E DISCUSSÃO

Uma amostragem dos dados simulados da população de 1000 indivíduos é mostrada na tabela 1.

Tabela 1 - Amostra dos indivíduos simulados com valores genéticos aditivos

id	sexo	vga ³	id	sexo	vga	Id	sexo	vga
1	m	24.526669	9	f	20.937999	17	f	12.980827
2	m	-10.734238	10	f	-2.371265	18	m	11.795195
3	m	-13.917107	11	f	-9.267480	19	f	-39.754086
4	m	-8.612326	12	m	-2.371265	20	m	-9.797959
5	m	5.429457	13	m	9.954407	21	f	-1.310309
6	m	-7.583142	14	m	-11.264716	22	f	14.853387
7	m	23.184588	15	f	-15.757894	:	:	:
8	f	5.180104	16	m	5.304781	1000	f	-1.591344

Fonte: Elaborado pelo autor.

Na tabela 2, é mostrada a matriz de covariância gerada aleatoriamente, pela rotina *randomCov*, em R, levando-se em consideração, os efeitos para cinco medidas de tempos diferentes, de forma que, $cov(T)$, seja simétrica e positiva definida.

Tabela 2 - Tabela de covariância gerada aleatoriamente

$cov(T)$	t_1	t_2	t_3	t_4	t_5
1	6.520180	0.992404	0.774613	0.373809	0.014971
2	0.992404	13.97741	0.873111	0.596431	0.454211
3	0.774613	0.873111	4.366321	0.918250	0.366047
4	0.373809	0.596431	0.918250	10.853417	0.022839
5	0.014971	0.454211	0.366047	0.022839	7.353504

Fonte: Elaborado pelo autor.

Aplicando-se a decomposição de Cholesky em $cov(T)$, se obteve a matriz LL' , apresentada na tabela 3.

³ valor genético aditivo para a característica simulada

Tabela 3 - Matriz LL' resultante da decomposição de Cholesky em $cov(T)$

LL'	t_1	t_2	t_3	t_4	t_5
1	2.5534644	0.000000	0.000000	0.000000	0.000000
2	0.3886503	3.718382	0.000000	0.000000	0.000000
3	0.3033578	0.203102	2.057436	0.000000	0.000000
4	0.1463930	0.145099	0.410399	3.262285	0.000000
5	0.0058632	0.121540	0.165051	-0.019431	2.703901

Fonte: Elaborado pelo autor.

Na tabela 4, são mostrados a amostra dos registros das variáveis aleatórias independentes.

Tabela 4 – Amostra da matriz de valores aleatórias independentes com variância unitária

z_t	z_1	z_2	z_3	z_4	z_5
1	-0.519406	-0.555993	2.420520	1.862141	-0.703461
2	1.120105	0.733699	-0.980111	0.349769	0.628330
3	-0.653483	-1.877910	-0.238688	-1.788556	1.989234
4	-1.797217	-0.660842	-0.261450	0.029847	0.102721
5	0.199735	1.304553	0.654945	1.590221	-2.681655
6	-0.784121	0.099286	-0.159605	0.431968	-0.451481
7	1.758988	-0.629897	0.480301	1.586495	-0.681393
8	1.087878	-0.586768	1.025629	-0.373896	-1.584628
9	0.715033	0.288123	1.064956	1.188198	-0.833089
10	-1.128804	-1.388098	1.026723	1.390796	-0.750200
11	0.174044	-1.759932	-0.791462	-0.663413	-0.809680
12	0.513890	-1.390316	-0.262666	-0.317551	-0.089592
13	-1.105523	0.252432	0.594385	0.391782	2.000706
14	1.457449	0.609799	-1.165474	-0.090609	0.493373
15	0.129120	-2.141884	0.428067	-0.159028	-0.457731
16	0.730492	-1.270351	1.718419	-1.020679	-0.604788
17	-1.031007	0.610975	1.085738	-0.451149	-0.979770
18	0.168551	1.221316	-2.490130	1.420743	-0.092765
19	-0.000108	-1.084181	1.060835	-0.215243	-0.032040
20	0.506680	0.816092	0.152985	-1.266743	0.309740
21	0.130952	0.279120	1.505558	0.859584	0.180288
⋮	⋮	⋮	⋮	⋮	⋮
1000	0.761452	0.420756	0.643192	-0.694382	-0.258484

Fonte: Elaborado pelo autor.

Uma amostra dos efeitos de ambiente temporário correlacionados para a população calculados pela rotina *calcEnvironmentEffect*, em R, que faz a multiplicação das matrizes L e z , pode ser vista na tabela 5.

Tabela 5 - Matriz E_t resultante da multiplicação da triangular inferior L por z

E_t	t_1	t_2	t_3	t_4	t_5
1	-1.326286	-2.269263	4.709579	6.911507	-1.609384
2	2.860149	3.163505	-1.527709	1.009246	1.626118
3	-1.668647	-7.236766	-1.070733	-6.300890	5.141979
4	-4.589132	-3.155754	-1.217336	-0.368917	0.143158
5	0.510017	4.928456	1.673058	5.675077	-7.014005
6	-2.002226	0.064438	-0.546081	1.243320	-1.248030
7	4.491515	-1.658567	1.393859	5.538822	-1.860219
8	2.777859	-1.759027	2.321009	-0.724720	-4.173067
9	1.825813	1.349252	2.466511	4.459784	-2.060695
10	-2.882363	-5.600190	1.488060	4.591879	-2.061358
11	0.444416	-6.476462	-1.933033	-2.718945	-2.519917
12	1.312201	-4.970007	-0.666903	-1.270246	-0.445397
13	-2.822915	0.508979	0.938811	1.396827	5.524403
14	3.721546	2.833906	-1.831910	-0.472061	1.226091
15	0.329706	-7.914163	0.484870	-0.635001	-1.423485
16	1.865288	-4.439747	3.499129	-2.701898	-1.481943
17	-2.632640	1.871139	2.045165	-1.088473	-2.393019
18	0.430390	4.606831	-4.824102	3.814811	-0.540009
19	-0.000277	-4.031444	1.962370	-0.424149	-0.039130
⋮	⋮	⋮	⋮	⋮	⋮
1000	1.944342	1.860471	1.639777	-1.828785	-0.523659

Fonte: Elaborado pelo autor.

A fim de comparação entre a matriz $cov(T)$, mostrada na figura 2, aplicamos a covariância em E_t , obtida pelo cálculo da fórmula 17, e apresentou resultados muito próximos da matriz T , neste caso, criada aleatoriamente pelo R, demonstrando correlação entre os dados simulados.

Tabela 6 - Covariâncias geradas sobre a matriz E_t

$cov(E_t)$	t_1	t_2	t_3	t_4	t_5
1	6.869091	1.302788	0.851793	0.217857	-0.119620
2	1.302788	14.186339	0.696501	1.098017	0.458044
3	0.851793	0.696501	4.744926	0.654165	0.573518
4	0.217857	1.098017	0.654165	11.459969	-0.212509

5 -0.119620 0.458044 0.573518 -0.212509 7.575023

Fonte: Elaborado pelo autor.

Ao final, chegou-se ao resultado, que é ilustrado na tabela 7, com uma amostra dos 1000 indivíduos simulados, com seus respectivos efeitos fenotípicos devidamente calculados como mostrado na equação 9.

Tabela 7 - Resultado da simulação do fenótipo para cada indivíduo

y_i	g	μ	a_i	$\mu + a_i + e_{ij}$				
				j_1	j_2	j_3	j_4	j_5
1	m	197.3051	24.526670	220.5055	219.5625	226.5413	228.7433	220.2224
2	m	197.3051	-10.734239	189.4310	189.7344	185.0432	187.5801	188.1970
3	m	197.3051	-13.917108	181.7193	176.1512	182.3173	177.0871	188.5300
4	m	197.3051	-8.612326	184.1036	185.5370	187.4754	188.3239	188.8359
5	m	197.3051	5.429458	203.2446	207.6630	204.4076	208.4096	195.7206
6	m	197.3051	-7.583143	187.7197	189.7864	189.1759	190.9653	188.4739
7	m	197.3051	23.184588	224.9812	218.8311	221.8835	226.0285	218.6295
8	f	197.3051	5.180105	205.2631	200.7262	204.8062	201.7605	198.3121
9	f	197.3051	20.937999	220.0689	219.5924	220.7096	222.7029	216.1824
10	f	197.3051	-2.371265	192.0515	189.3336	196.4219	199.5257	192.8725
11	f	197.3051	-9.267481	188.4820	181.5612	186.1046	185.3187	185.5177
12	m	197.3051	-2.371265	196.2460	189.9638	194.2669	193.6636	194.4884
13	m	197.3051	9.954408	204.4366	207.7685	208.1983	208.6563	212.7839
14	m	197.3051	-11.264717	189.7619	188.8743	184.2085	185.5683	187.2665
15	f	197.3051	-15.757895	181.8769	173.6330	182.0321	180.9122	180.1237
16	m	197.3051	5.304781	204.4752	198.1701	206.1090	199.9080	201.1279
17	f	197.3051	12.980828	207.6533	212.1571	212.3311	209.1975	207.8929
18	m	197.3051	11.795195	209.5307	213.7071	204.2762	212.9151	208.5603
19	f	197.3051	-39.754086	157.5507	153.5196	159.5134	157.1269	157.5119
20	m	197.3051	-9.797959	188.8009	190.7386	188.1414	183.6300	188.4967
21	f	197.3051	-1.310309	196.3292	197.0836	199.1888	199.4766	196.7488
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1000	f	197.3051	-1.591434	197.6580	197.5741	197.3534	193.8849	195.1900

Fonte: Elaborado pelo autor.

Embora o algoritmo permita a simulação de apenas uma característica parece muito promissor com relação aos efeitos longitudinais esperados para cada indivíduo.

7 CONCLUSÃO

A análise de complexidade mostrou que, algumas rotinas utilizadas para o desenvolvimento, possuem notação assintótica $O(n^3)$, significando que não são lineares, negando a segunda hipótese desse trabalho. Porém, foi possível a simulação de fenótipos, com efeitos de ambientes temporários, o que mostrou que, da forma correta, é possível o desenvolvimento de algoritmos para simular dados longitudinais que leve em consideração os efeitos genéticos e ambientais potencialmente correlacionados. A decomposição de Cholesky, mostrou-se bastante promissora na simulação de dados de forma correlacionada, o que permitirá a identificação de animais potencialmente resilientes para determinadas situações de estresse e análise comportamental.

REFERÊNCIAS

ALMEIDA, L.; CORREA, C. Percepções sobre os jogos de simulação de voo na formação de pilotos privados de avião. **RENOTE**, [s. l.], v. 15, n. 1, 2017.

Disponível em: <http://seer.ufrgs.br/index.php/renote/article/view/75108>. Acesso em: 18 out. 2022.

ALMEIDA, A. T.; OLIVEIRA, A. F. de. Simulação de dados meteorológicos para avaliação de impactos de mudanças climáticas na agricultura. *In*: MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 6., 2010, Campinas. **Resumos** [...]. Campinas: Embrapa Informática Agropecuária: [s. n.], 2010. 2010.electronic resource.

ALVES, S. P. **Uso da zootecnia de precisão na avaliação do bem-estar bioclimático de aves poedeiras em diferentes sistemas de criação**. 2007. 128 f. Tese (Doutorado em Física do Ambiente Agrícola) - Universidade de São Paulo, Piracicaba 2007. Disponível em:

<http://www.teses.usp.br/teses/disponiveis/11/11131/tde-04042007-141113/>. Acesso em: 13 ago. 2022.

BARBOSA FILHO, J. A. D. **Avaliação do bem-estar de aves poedeiras em diferentes sistemas de produção e condições ambientais, utilizando análise de imagens**. 2005. 123 f. Dissertação (Mestrado em Física do Ambiente Agrícola) - Universidade de São Paulo, Piracicaba 2005. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11131/tde-11052005-144156/>. Acesso em: 17 mai. 2022.

BARIONI, L. G.; MANCINI, A. L. Desenvolvimento de simuladores na agropecuária. *In*: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (Ed.). **Tecnologias da informação e comunicação e suas relações com a agricultura**. Brasília, DF: Embrapa, 2014. Cap. 13.

BEIGUELMAN, B. **Curso prático de bioestatística**. 5. ed. Ribeirão Preto: Fundação de Pesquisa Científica de Ribeirão Preto, 2002.

BHARGAVA, A. Y. **Entendendo algoritmos: um guia ilustrado para programadores e outros curiosos**. [S. l.]: Novatec, 2017.

BESSA JÚNIOR, J. E.; SETTI, J. R. Produção de dados de tráfego sintéticos através de algoritmo genético e simulação microscópica. **Transportes**, [S. l.], v. 18, n. 3, 2010. Disponível em:

<http://www.revistatransportes.org.br/anpet/article/view/447>. Acesso em: 25 ago. 2022.

BOND, G. B.; ALMEIDA, R. de; OSTRENSKY, A.; MOLENTO, C. F. M. Métodos de diagnóstico e pontos críticos de bem-estar de bovinos leiteiros. **Ciência Rural**, [S. l.], v. 42, n. 7, p. 1286–1293, 2012.

BORGES, G. **Utilização da pressão sonora (ruído) como indicativo de bem-estar animal na produção industrial de suínos**. 2009. 139 f. Dissertação (Mestrado em Física do Ambiente Agrícola) - Universidade de São Paulo, Piracicaba 2009. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11131/tde-10022009-103022/>. Acesso em: 08 nov. 2022.

BRAGA, J. da S.; MACITELLI, F.; ABREU DE LIMA, V.; DIESEL, T. O modelo dos “Cinco Domínios” do bem-estar animal aplicado em sistemas intensivos de produção de bovinos, suínos e aves. **Revista Brasileira de Zootecias**, [S. l.], v. 19, n. 2, 2018. Disponível em: <https://periodicos.ufjf.br/index.php/zootecias/article/view/24771>. Acesso em: 19 dez. 2022.

BRAINER, M. M. D. A.; SILVA, L. F. F.; MACHADO, A. S.; FARIA JÚNIOR, O. L.; RAMOS, J. S. Pecuária leiteira de precisão. *In*: CIENTÍFICA DIGITAL (org.). **Open Science Research VI**. [S. l.]: Científica Digital, 2022. p. 331–349. Disponível em: <http://www.editoracientifica.com.br/articles/code/220910020>. Acesso em: 22 dez. 2022.

BROOM, D. M. Indicators of poor welfare. **British Veterinary Journal**, [S. l.], v. 142, n. 6, p. 524–526, 1986.

BROOM, D. M.; JOHNSON, K. G.; BROOM, D. M. **Stress and animal welfare**. London: Chapman & hall, 1993.

CARVAJAL-RODRIGUEZ, A. Simulation of genomes: a review. **Current Genomics**, [S. l.], v. 9, n. 3, p. 155–159, 2008.

CEBALLOS, M. C.; SANT’ANNA, A. C. Evolução da ciência do bem-estar animal: Uma breve revisão sobre aspectos conceituais e metodológicos. **Revista Acadêmica Ciência Animal**, [S. l.], v. 16, p. 1, 2018.

CORMEN, T. H.; LEISERSON, C. Eric.; RIVEST, R. L.; STEIN, C.; MARQUES, A. S.; MANDEL, A. **Algoritmos: teoria e prática**. Rio de Janeiro: Campus, 2012.

COSTA, E. J. X.; PIZA, L. V.; LA VEGA, L. T. de. Bem-estar animal durante abate de aves avaliado por meio de eletroencefalograma: um referencial teórico. *In*: COSTA, Ernane José Xavier. **Zootecnia de precisão: desafios e aplicações**. 1. ed. [S. l.]: Científica Digital, 2021. p. 50–58. Disponível em: <http://www.editoracientifica.com.br/articles/code/211106790>. Acesso em: 25 ago. 2022.

DUNCAN, I. J. H. Science-based assessment of animal welfare: farm animals. **Revue Scientifique et Technique Office International des Epizooties**, Paris, v. 2, n. 24, p. 483-492, 2005.

DURRETT, R.; SCHWEINSBERG, J. A coalescent model for the effect of advantageous mutations on the genealogy of a population. **Stochastic Processes and their Applications**, [S. l.], v. 115, n. 10, p. 1628–1657, 2005.

FIGUEIRA, S. V.; NASCIMENTO, G. M.; MOTA, B. de P.; LEONÍDIO, A. R. A.; ANDRADE, M. A. **Bem-estar animal aplicado a frangos de corte**. [S. l.: s.n.], 2014. Disponível em: <http://repositorio.bc.ufg.br/handle/ri/13540>. Acesso em: 11 dez. 2022.

FRENZEL, J. F. Genetic algorithms. **IEEE Potentials**, [S. l.], v. 12, n. 3, p. 21–24, 1993.

FROEHLICH, G. **O bem-estar na carne : um estudo antropológico sobre as relações entre humanos e animais a partir da categoria de “bem-estar animal”**. 2016. 282 f. Tese (Doutorado) - Universidade de Brasília, Brasília, 2016. Disponível em: <https://repositorio.unb.br/handle/10482/21510>. Acesso em: 11 dez. 2022.

GALHARDO, L.; OLIVEIRA, R. Bem-estar animal: um conceito legítimo para peixes? **Revista de Etologia**, [S. l.], v. 8, p. 51–61, 2006.

GENTLE, M. J.; HUGHES, B. O.; HUBRECHT, R. C. The effect of beak trimming on food intake, feeding behaviour and body weight in adult hens. **Applied Animal Ethology**, [S. l.], v. 8, n. 1–2, p. 147–159, 1982.

GOMES, R. A.; MONTEIRO, G. R.; ASSIS, G. J. F.; BUSATO, K. C.; LADEIRA, M. M.; CHIZZOTTI, M. L. Technical note: estimating body weight and body composition of beef cattle through digital image analysis. **Journal of Animal Science**, [S. l.], v. 94, n. 12, p. 5414–5422, 2016.

GUIRRO, E. C. B. do P. Perspectiva bioética sobre o princípio das cinco liberdades e do modelo dos cinco domínios do bem-estar animal. **Revista Inclusiones**, [S. l.], v. 9, n. 3, p. 129–146, 2022.

HEIN, J.; SCHIERUP, M.; WIUF, C. **Gene genealogies, variation and evolution: a primer in coalescent theory**. Oxford: Oxford University Press, 2004.

HONDA, T.; NOMURA, T.; FUKUSHIMA, M.; MUKAI, F. Gene dropping analysis of founder contributions in a closed Japanese black cattle population. **Animal Science Journal**, [S. l.], v. 73, n. 2, p. 105–111, 2002.

HÖTZEL, M. J.; MACHADO FILHO, L. C. P. Bem-estar animal na agricultura do século XXI. **Revista de Etologia**, [S. l.], v. 6, p. 3–15, 2004.

JUNIOR, . G. P.; SILVA, C. L. da; GONÇALVES, R. A. de L.; MARTINI, J. A. R10k: um simulador de arquitetura superescalar. *In*: WORKSHOP SOBRE EDUCAÇÃO EM ARQUITETURA DE COMPUTADORES–WEAC, 2007. **Anais** [...]. [S.l.: s.n.], 2007.

KIM, Y.; WIEHE, T. Simulation of DNA sequence evolution under models of recent directional selection. **Briefings in Bioinformatics**, [S. l.], v. 10, n. 1, p. 84–96, 2008.

LEITE, C. R.; NASCIMENTO, M. R. B. de M.; SANTANA, D. de O.; GUIMARÃES, E. C.; MORAIS, H. R. Influência do manejo pré-abate de bovinos na indústria sobre os parâmetros de bem-estar animal e impactos no pH 24 horas post mortem. **Bioscience Journal**, [S. l.], v. 31, n. 1, p. 194–203, 2015.

LIANG, L.; ZOLLNER, S.; ABECASIS, G. R. GENOME: a rapid coalescent-based whole genome simulator. **Bioinformatics**, [S. l.], v. 23, n. 12, p. 1565–1567, 2007.

LINO, D. M.; PINHEIRO, R. S. B.; ORTUNHO, V. V. Benefícios do bem-estar animal na produtividade e na sanidade de ovinos. **Periódico Eletrônico Fórum Ambiental da Alta Paulista**, [S. l.], v. 12, n. 5, 2016. Disponível em: https://www.amigosdanatureza.org.br/publicacoes/index.php/forum_ambiental/article/view/1506. Acesso em: 25 ago. 2022.

LUDTKE, C. B.; DALLA COSTA, O. A.; ROÇA, R. de O.; SILVEIRA, E. T. F.; ATHAYDE, N. B.; ARAÚJO, A. P. de; MELLO JÚNIOR, A. de; AZAMBUJA, N. C. de. Bem-estar animal no manejo pré-abate e a influência na qualidade da carne suína e nos parâmetros fisiológicos do estresse. **Ciência Rural**, [S. l.], v. 42, n. 3, p. 532–537, 2012.

MACHADO FILHO, L. C. P. M. Bem-estar de suínos e qualidade da carne: uma visão brasileira. *In*: CONFERÊNCIA INTERNACIONAL VIRTUAL SOBRE QUALIDADE DE CARNE SUÍNA, 1., 2000, Concórdia. **Anais [...]**. Concórdia: Embrapa, 2001. p. 34 - 40. Disponível em: http://www.cnpsa.embrapa.br/sgc/sgc_publicacoes/anais00cv_portugues.pdf. Acesso em: 25 ago. 2022.

MACITELLI, F.; BERCHIELLI, T. T.; MORAIS, J. A. da S.; SILVEIRA, R. N. da; CANESIN, R. C. Desempenho e rendimento de carcaça de bovinos mestiços alimentados com diferentes volumosos e fontes protéicas. **Revista Brasileira de Zootecnia**, [S. l.], v. 36, n. 6, p. 1917–1926, 2007.

MANATRINON, S.; EGGER-DANNER, C.; BAUMUNG, R. Estimating lethal allele frequencies in complex pedigrees via gene dropping approach using the example of Brown Swiss cattle. **Archives Animal Breeding**, [S. l.], v. 52, n. 3, p. 230–242, 2009.

MANTECA, X.; SILVA, C. A. da; BRIDI, A. M.; DIAS, C. P. Bem-estar animal: conceitos e formas práticas de avaliação dos sistemas de produção de suínos. **Semina: Ciências Agrárias**, [S. l.], v. 34, n. 6Supl2, p. 4213, 2013.

MARJORAM, P.; WALL, J. D. Fast “coalescent” simulation. **BMC Genetics**, [S. l.], v. 7, n. 1, p. 16, 2006.

- MELKA, M. G.; SARGOLZAEI, M.; MIGLIOR, F.; SCHENKEL, F. Genetic diversity of Guernsey population using pedigree data and gene-dropping simulations. **Animal**, [S. l.], v. 7, n. 2, p. 192–201, 2013.
- MOLENTO, C. F. M. Bem-estar e produção animal: aspectos econômicos - revisão. **Archives of Veterinary Science**, [S. l.], v. 10, n. 1, 2005. Disponível em: <http://revistas.ufpr.br/veterinary/article/view/4078>. Acesso em: 02 mar. 2022.
- MORTENSEN, A. K.; LISOUSKI, P.; AHRENDT, P. Weight prediction of broiler chickens using 3D computer vision. **Computers and Electronics in Agriculture**, [S. l.], v. 123, p. 319–326, 2016.
- NASCIMENTO P. L.; CASANOVA, P. T.; MARIN L. Uso de acelerômetro para identificar padrões de movimentos mandibulares no comportamento ingestivo de bovinos em pastejo. *In*: CONGRESSO BRASILEIRO DE ZOOTECNIA, 28., 2018, Goiania. **Anais** [...]. Goiania: [s.n.], 2018.
- NELSON, D.E.; ALKON, P.U.; KRAUSMAN, P.R. Using acoustic telemetry to monitor foraging by penned mule deer. **Wildlife Society Bulletin**, v. 33, p. 624-632, 2005.
- PENG, B.; CHEN, H.-S.; MECHANIC, L. E.; RACINE, B.; CLARKE, J.; GILLANDERS, E.; FEUER, E. J. Genetic data simulators and their applications: an overview. **Genetic Epidemiology**, [S. l.], v. 39, n. 1, p. 2–10, 2015.
- PENG, B.; KIMMEL, M.; AMOS, C. I. **Forward-time population genetics simulations: Methods, Implementation, and Applications**. Somerset: Wiley, 2012.
- PEREIRA, L. G. R.; PAIVA, C. A. V.; RIBAS, M. N.; FERREIRA, A. L. Pecuária leiteira de precisão: conceitos e tecnologias disponíveis. **Cadernos Técnicos de Veterinária e Zootecnia**, v. 79, p. 7-14, 2015.
- PÉREZ-ENCISO, M.; LEGARRA, A. A combined coalescence gene-dropping tool for evaluating genomic selection in complex scenarios (ms2gs). **Journal of Animal Breeding and Genetics**, [S. l.], v. 133, n. 2, p. 85–91, 2016.
- PIRES, M. P.; ONO, R. K.; UTSUNOMIYA, A. T. H.; QUEIROZ, A.; FONSECA, R. Simulação de características geneticamente correlacionadas. *In*: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRAS DE ZOOTECNIA, 44., 2007, Jaboticabal. **Anais** [...]. Jaboticabal: [s. n.], 2007.
- POLASTRINI, A.; BRACARENSE, L. dos S. F. P.; PEDROZA FILHO, M. X. Perdas econômicas decorrentes de lesões em carcaças bovinas durante o transporte pré-abate: O caso do Estado do Tocantins. **Agri-Environmental Sciences**, [S. l.], v. 7, n. 1, p. 15, 2021.

PORTO, S. M. C.; ARCIDIACONO, C.; ANGUZZA, U.; CASCONI, G. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. **Biosystems Engineering**, [S. l.], v. 115, p. 184–194, 2013.

SIGWART, J. Gene genealogies, variation and evolution: a primer in coalescent theory. Oxford University Press, Oxford. xiii + 276 pp. ISBN 0-19-852996-1; ISBN 0-19-852995-3, **Systematic Biology**, [S. l.], v. 54, n. 6, p. 986–987, 2005.

SILVA, D. P. **Revisão sistemática e metanálise de parâmetros genéticos para indicadores de bem-estar animal em bovinos de leite**. 2022. 95 f. Dissertação (Mestrado em Ciência e Tecnologia Animal) - Universidade Estadual Paulista (Unesp), Dracena, 2022. Disponível em: <https://repositorio.unesp.br/handle/11449/235147>. Acesso em: 18 ago. 2022.

SILVA, I. J. O.; MIRANDA, K. O. S. Impactos do bem-estar na produção de ovos. **Thesis**, [S. l.], v. 6, n. 11, p. 89-115, 2009.

SILVA-ANTUNES, L. C. M.; COSTA, M. J. R. P. da. The adoption of good practices of handling improves dairy calves welfare: Case study. **Acta Scientiarum: Animal Sciences**, [S. l.], v. 43, p. e53327, 2021.

TURCO, S. H. N.; LOURENÇONI, D.; COSTA, D. dos S.; OLIVEIRA, P. T. L. de; ROCHA, D. R. da; SANTOS, I. E. dos A. Ferramentas para o monitoramento de respostas comportamentais, fisiológicas e de desempenho animal a campo. **Revista Científica de Produção Animal**, [S. l.], v. 21, n. 2, p. 69–75, 2020.

VERNIZ, V.; FONSECA, R.; QUEIROZ, A. Desenvolvimento de uma nova interface de texto para o simulador de dados LZ5. *In*: ENCONTRO DE ZOOTECNIA – UNESP DRACENA, 6., 2009, Dracena. **Anais** [...]. Dracena: [s. n.], 2009.

WEN, M.; ROSA, T. D. O.; SOUZA, M. C.; ALEIXO, R. P.; ALVES, C.; SÁ, L.; SANTANA, E. F. Z.; KON, F. Criação de Modelo para Simulação de Movimentação de Ônibus a Partir de Dados Reais. *In*: WORKSHOP BRASILEIRO DE CIDADES INTELIGENTES (WBCI), 2018. **Anais** [...]. [S. l.]: Sociedade Brasileira de Computação - SBC, 2018. Disponível em: <https://sol.sbc.org.br/index.php/wbci/article/view/3225>. Acesso em: 17 nov. 2022.

APÊNDICES

I. Simulador

	Programa	Custo
1	<code>source("tools.R")</code>	1
2	<code>source("genome.R")</code>	1
3	<code>source("cromossome.R")</code>	1
4	<code>source("locus.R")</code>	1
5	<code>source("popgen.R")</code>	1
6	<code>source("individual.R")</code>	1
7	<code>source("randomCov.R")</code>	1
8	<code>source("calcTempEnvEffect.R")</code>	1
9	<code># número de cromossomos</code>	
10	<code># número de locos poligênicos</code>	
11	<code># número de marcadores</code>	
12	<code>cChrom<-getNLoci(29, 1000, 10)</code>	O(n)
13	<code># configuração dos cromossomos</code>	
14	<code># frequência dos alelos A1</code>	
15	<code># variância genotípica</code>	
16	<code># número de qtls</code>	
17	<code># percentual de efeitos dos qtls</code>	
18	<code># número da característica</code>	
19	<code>genome<-newGenome(cChrom, 0.5, 200, 5, 0.3, 1)</code>	O(n ²)
20	<code>pop<-popGen(1000)</code>	O(n ³)
21	<code>T <- randomCov(4,1,20)</code>	O(n ³)
22	<code>L <- t(chol(T))</code>	O(n ³)
23	<code>z<-data.frame(z1 = rnorm(1000), z2 = rnorm(1000), z3 = rnorm(1000), z4 = rnorm(1000))</code>	O(n)
24	<code>et <- calcEnvironmentEffect(L,z)</code>	O(n ²)
25	<code>pop_ <- data.frame(id=unlist(sapply(pop\$pop, function(x) ifelse("indId" %in% names(x), x\$indId, NULL))), gender=unlist(sapply(pop\$pop, function(x) ifelse("gender" %in% names(x), x\$gender, NULL))),mean=197.3051, adGenVal=unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))), et1=(197.3051+unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))))+et[,1], et2=(197.3051+unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))))+et[,2], et3=(197.3051+unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))))+et[,3], et4=(197.3051+unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))))+et[,4], et5=(197.3051+unlist(sapply(pop\$pop, function(x) ifelse("adGenVal" %in% names(x), x\$adGenVal, NULL))))+et[,5])</code>	O(n ²)

$$T(n) = 8 + 2n + 3n^2 + 3n^3 = O(n^3)$$

II. Rotina de definição de sexo

Programa	Custo
1 pickGender<-function(f){	
2 rNumber<-runif(1,0,1)	1
3 if(rNumber<=f)	1
4 return("f")	1
5 else	1
6 return("m")	1
7 }	

$$T(n) = 5 = O(1)$$

III. Rotina de configuração de cada loco

Programa	Custo
1 configLocus<-function(genome, freqAllele){	O(n)
2 for(i in 1:length(genome)){	O(n ²)
3 for(j in 1:length(genome[[i]])){	O(n ²)
4 pAllele<-runif(1,0,1)	O(n ²)
5 mAllele<-runif(1,0,1)	O(n ²)
6 genome[[i]][[j]]\$allelesConfig<-	O(n ²)
c(ifelse(pAllele<=freqAllele,1,0),	
ifelse(mAllele<=freqAllele,1,0))	
7 if((genome[[i]][[j]]\$allelesConfig[1]==0 &&	O(n ²)
genome[[i]][[j]]\$allelesConfig[2]==0)){	
8 genome[[i]][[j]]\$locusEffect <-	O(n ²)
genome[[i]][[j]]\$locusEffect*-1	
9 }else	O(n ²)
10 if((genome[[i]][[j]]\$allelesConfig[1] == 1 &&	O(n ²)
genome[[i]][[j]]\$allelesConfig[2] ==	
0) (genome[[i]][[j]]\$allelesConfig[1] == 0 &&	
genome[[i]][[j]]\$allelesConfig[2] == 1)){	
11 genome[[i]][[j]]\$locusEffect<-0	O(n ²)
12 }	
13 }	
14 }	
15 return(genome)	1
16 }	

$$T(n) = n + 10n^2 = O(n^2)$$

IV. Rotina de cálculo do valor genético aditivo de cada loco

Programa	Custo
1 getValLoci<-function(nVarAdGen, nloci, fp){	
2 fq<-1-fp	1
3 fpq<-2*fp*fq	1
4 a<-(nVarAdGen/(fpq * nloci))^(0.5)	1
5 return(a)	1
6 }	

$$T(n) = 4 = O(1)$$

V. Rotina para geração de locos poligênicos por cromossomo

Programa	Custo
1 # Função para geração de números de locos por cromossomo	
2 # nChrom = número de cromossomos	
3 # nLoci = número de locos por cromossomo	
4 # se nLoci for <=0, gera nChrom números aleatórios de 10 a 10.000	
5 # se nLoci for um número maior que zero, gera nChrom números iguais a nLoci	
6 # se nLoci for um vetor de tamanho igual a nChrom, retorna nLoci	
7 # se nLoci for um vetor de 2 números e nChrom for diferente de dois, gera nChrom números entre nLoci[1] e nLoci[2]	
8 <code>getNLoci<-function(nChrom, nLoci, nMarkers){</code>	
9 <code>chromLoci<-rep(0,nChrom)</code>	0(n)
10 <code>markers<-rep(0,nChrom)</code>	0(n)
11 <code>if(length(nLoci)==1){</code>	1
12 <code>if(nLoci<=0){</code>	1
13 <code>chromLoci<-round(runif(nChrom, 10, 10000))</code>	0(n)
14 <code>}else{</code>	1
15 <code>chromLoci<-rep(0,nChrom)</code>	0(n)
16 <code>for(i in 1:nLoci){</code>	0(n)
17 <code>n<-runif(1, 1, nChrom+1)</code>	0(n)
18 <code>chromLoci[n]<-chromLoci[n]+1</code>	0(n)
19 <code>}</code>	
20 <code>}</code>	
21 <code>}else if(length(nLoci)==2&nChrom!=2){</code>	1
22 <code>chromLoci<-round(runif(nChrom, nLoci[1], nLoci[2]))</code>	0(n)
23 <code>}else{</code>	1
24 <code>chromLoci<-nLoci</code>	1
25 <code>if(nMarkers!=0){</code>	1
26 <code>markers<-rep(0,nChrom)</code>	0(n)
27 <code>for(i in 1:nMarkers){</code>	0(n)
28 <code>n<-runif(1, 1, nChrom+1)</code>	0(n)
29 <code>markers[n]<-markers[n]+1</code>	0(n)
30 <code>}</code>	
31 <code>}</code>	
32 <code>return(list("nChrom"=nChrom, "chromLoci"=chromLoci, "chromMarkers"=markers))</code>	1
33 <code>}</code>	

$$T(n) = 8 + 12n = O(n)$$

VI. Rotina para calcular a variância da população

Programa	Custo
1 <code>popVar<-function(pop){</code>	
2 <code>v<-0</code>	1
3 <code>for(i in 1:length(pop)){</code>	0(n)
4 <code>v<-v+(pop[[i]]\$adGenVal^2)</code>	0(n)
5 <code>}</code>	
6 <code>v<-v/(length(pop)-1)</code>	1
7 <code>return(v)</code>	1
8 <code>}</code>	

$$T(n) = 3 + 2n = O(n)$$

VII. Rotina para criação do genoma baseado em parâmetros

Programa	Custo
1 <code>newGenome<-function(cChrom, freqAllele, polVal, nQtls,</code>	
<code>effectQtls, traitId){</code>	
2 <code>qtlVal<-polVal*effectQtls</code>	1
3 <code>polVal<-polVal-qtlVal</code>	1
4 <code>effectPols<-getValLoci(polVal, sum(cChrom[[2]])-nQtls,</code>	0(1)*0(1)
<code>freqAllele)</code>	
5 <code>effectQtls<-getValLoci(qtlVal, nQtls, freqAllele)</code>	0(1)*0(1)
<code>genome<-NULL</code>	
6 <code>n<-NULL</code>	1
7 <code>for(i in 1:cChrom[[1]]){</code>	0(n)
8 <code>genome[[i]]<-newChromosome(cChrom[[2]][i],</code>	0(n)*0(n)
<code>cChrom[[3]][i], freqAllele, effectPols, traitId)</code>	
9 <code>n[[i]]<-paste("chromosome", i, sep=" ")</code>	0(n)
10 <code>}</code>	
11 <code>while(nQtls>0){</code>	0(n)
12 <code>chrom<-round(runif(1,1,length(genome)))</code>	0(n)
13 <code>locus<-round(runif(1,1,cChrom[[2]][chrom]))</code>	0(n)
14 <code>genome[[chrom]][[locus]]\$locusEffect<- effectQtls</code>	0(n)
15 <code>genome[[chrom]][[locus]]\$type<-"q"</code>	0(n)
16 <code>nQtls<-nQtls-1</code>	0(n)
17 <code>}</code>	
18 <code>for(i in 1:length(genome)){</code>	0(n)
19 <code>genome[[i]]<-sample(genome[[i]], replace=FALSE)</code>	0(n)
20 <code>}</code>	
21 <code>names(genome)<-n</code>	1
22 <code>return(genome)</code>	1
23 <code>}</code>	

$$T(n) = 7 + 12n + n^2 = O(n^2)$$

VIII. Rotina para criar os cromossomos e definição do tipo de cada loco

Programa	Custo
1 <code>newChromossome<-function(chromLoci, chromMarkers,</code>	
2 <code> freqAllele, effectPols, traitId){</code>	
3 <code> chrom<-NULL</code>	1
4 <code> n<-NULL</code>	1
5 <code> for(i in 1:chromLoci){</code>	0(n)
6 <code> chrom[[i]]<-newLocus(freqAllele, effectPols,</code>	0(1)*0(n)
7 <code> traitId, i, "p")</code>	
8 <code> n[[i]]<-paste("locus", i, sep=" ")</code>	0(n)
9 <code> }</code>	
10 <code> if(chromMarkers>0){</code>	1
11 <code> m<-length(chrom)</code>	1
12 <code> for(i in 1:chromMarkers){</code>	0(n)
13 <code> m<-m+1</code>	0(n)
14 <code> chrom[[m]]<-newLocus(freqAllele, 0, traitId, m,</code>	0(1)*0(n)
15 <code> "m")</code>	
16 <code> n[[m]]<-paste("locus", m, sep=" ")</code>	0(n)
17 <code> }</code>	
18 <code> }</code>	
19 <code> names(chrom)<-n</code>	1
20 <code> return(chrom)</code>	1
21 <code>}</code>	

$$T(n) = 6 + 7n = O(n)$$

IX. Rotina para gerar a matriz de efeitos de ambiente

Programa	Custo
1 <code>calcEnvironmentEffect<-function(L, z){</code>	
2 <code> et<-matrix(ncol=ncol(L), nrow=nrow(z))</code>	0(n ²)
3 <code> for(i in 1:nrow(L)){</code>	0(n)
4 <code> val<-0</code>	0(n)
5 <code> j<-1</code>	0(n)
6 <code> while(j<=i){</code>	0(n ²)
7 <code> val<-val+(L[i, j] * z[, j])</code>	0(n ²)
8 <code> j<-j+1</code>	0(n ²)
9 <code> }</code>	
10 <code> et[,i]<-val</code>	0(n)
11 <code> }</code>	
12 <code> e <- data.frame(et)</code>	1
13 <code> return(e)</code>	1
14 <code>}</code>	

$$T(n) = 2 + 4n + 4n^2 = O(n^2)$$

X. Rotina para criar indivíduos baseados no genoma

Programa	Custo
1 # Função para Geração de um Indivíduo	
2 # indId Identificação do Indivíduo	
3 # sireId Identificação do Pai	
4 # damId Identificação da Mãe	
5 # fGender Frequência de Machos na população	
6 # numChrom numChrom + numLocus	
7 # freqAllele Frequência dos Alelos A	
8 # polVal Variância Genética Aditiva	
9 # traitId Identificação da Característica	
10 newIndividual<-function(indId, sireId, damId, fGender, freqAllele, genome){	
11 individual<-NULL	1
12 n<-NULL	1
13 nChromLoci<-0	1
14 individual[[1]]<-indId	1
15 individual[[2]]<-sireId	1
16 individual[[3]]<-damId	1
17 individual[[4]]<-pickGender(fGender)	O(1)*O(1)
18 individual[[5]]<-0	1
19 individual[[6]]<-0	1
20 individual[[7]]<-0	1
21 individual[[8]]<- configLocus(genome, freqAllele)	O(1)*O(n ²)
22 individual[[9]]<-0	1
23 names(individual)<-list("indId", "sireId", "damId", "gender", "phenVal", "adGenVal", "envDevVal", "genome", "locusLength")	1
24 individual\$adGenVal<-0	1
25 for(i in 1:length(individual\$genome)){	O(n)
26 for(j in 1:length(individual\$genome[[i]])){	O(n ²)
27 individual\$adGenVal<-individual\$adGenVal + individual\$genome[[i]][[j]]\$locusEffect	O(n ²)
28 nChromLoci<-nChromLoci+1	O(n ²)
29 }	
30 }	
31 individual\$locusLength<-nChromLoci	1
32 return(individual)	1
33 }	

$$T(n) = 15 + n + 4n^2 = O(n^2)$$

XI. Rotina para geração da população

Programa	Custo
1 <code>popGen<-function(p){</code>	
2 <code> print(paste("Criando a população de ",p," indivíduos"))</code>	1
3 <code> pop<-list()</code>	1
4 <code> n<-list()</code>	1
5 <code> pop[[1]]<-list()</code>	1
6 <code> pop[[2]]<-list()</code>	1
7 <code> pop[[3]]<-NA</code>	1
8 <code> ind<-list()</code>	1
9 <code> for(i in 1:p){</code>	0(n)
10 <code> ind[[i]]<-newIndividual(i, NA, NA, 0.5, 0.5, genome)</code>	0(n*n ²)
11 <code> if(ind[[i]]\$gender=="m"){</code>	0(n)
12 <code> pop[[1]][length(pop[[1]])+1] <-</code>	0(n)
<code>as.integer(ind[[i]]\$indId)</code>	
13 <code> }else{</code>	0(n)
14 <code> pop[[2]][length(pop[[2]])+1] <-</code>	0(n)
<code>as.integer(ind[[i]]\$indId)</code>	
15 <code> }</code>	
16 <code> n[i]<-"individual"</code>	0(n)
17 <code> }</code>	
18 <code> names(ind)<-n</code>	1
19 <code> pop[[4]]<-ind</code>	1
20 <code> names(pop)<-list("males","females","GenAdVar","pop")</code>	1
21 <code> pop[[3]]<-popVar(pop\$pop)</code>	0(1)*0(n)
22 <code> return(pop)</code>	1
23 <code>}</code>	

$$T(n) = 11 + 7n + n^3 = O(n^3)$$

XII. Rotina para criação de locos

Programa	Custo
1 <code>newLocus<-function(freqAllele, locusEffect, traitId,</code>	
<code>locusPos, type){</code>	
2 <code> allelesConfig<-c(0,0)</code>	1
3 <code> locus<-list(allelesConfig, locusEffect, traitId,</code>	1
<code>locusPos, type)</code>	
4 <code> names(locus)<-list("allelesConfig", "locusEffect",</code>	1
<code>"traitId", "locusPos", "type")</code>	
5 <code> return(locus)</code>	1
6 <code>}</code>	

$$T(n) = 4 = O(1)$$

XIII. Rotina para criar a matriz de covariâncias aleatórias

Programa	Custo
1 <code>randomCov<-function(n, dini, dfim){</code>	
2 <code> R<-matrix(data = rep(0,n), nrow = n, ncol = n)</code>	$O(n^2)$
3 <code> for(i in 1:n) {</code>	$O(n)$
4 <code> for(j in n-i:n){</code>	$O(n^2)$
5 <code> x <- rnorm(n-i, mean=0)</code>	$O(n^2)$
6 <code> y <- pnorm(x, lower.tail = TRUE)</code>	$O(n^2)$
7 <code> y <- sort(y, decreasing=TRUE)</code>	$O(n^2)$
8 <code> if(length(y)!=0){</code>	$O(n^2)$
9 <code> R[i,t+i] <- y[t]</code>	$O(n^3)$
10 <code> R[t+i,i] <- y[t]</code>	$O(n^3)$
11 <code> }</code>	
12 <code> }</code>	
13 <code> }</code>	
14 <code> S <- runif(n, min=dini, max=dfim)</code>	1
15 <code> diag(R) <- S</code>	1
16 <code> return(R)</code>	1
17 <code>}</code>	

$$T(n) = 3 + 1n + 6n^2 + 3n^3 = O(n^3)$$