

**UNIVERSIDADE FEDERAL DE PELOTAS**  
**Faculdade de Agronomia Eliseu Maciel**  
**Programa de Pós-Graduação em Zootecnia**



**Tese**

**Uma abordagem computacional para estudo comparativo das estimativas de componentes de variância e parâmetros genéticos sob enfoque bayesiano em relação à abordagem clássica por máxima verossimilhança**

**Gil Carlos Rodrigues Medeiros**

**Pelotas, 2015**

**Gil Carlos Rodrigues Medeiros**

**Uma abordagem computacional para estudo comparativo das estimativas de componentes de variância e parâmetros genéticos sob enfoque bayesiano em relação à abordagem clássica por máxima verossimilhança**

Tese apresentada ao Programa de Pós-Graduação em Zootecnia da Faculdade de Agronomia Eliseu Maciel da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Doutor em Zootecnia.

Orientador: Prof. Dr. Nelson José Laurino Dionello

Co-orientador: Prof. Dr. Willian Silva Barros

Pelotas, 2015

Dados de catalogação na fonte:  
Ubirajara Buddin Cruz – CRB-10/901  
Biblioteca de Ciência & Tecnologia - UFPel

M488a      Medeiros, Gil Carlos Rodrigues  
              Uma abordagem computacional para estudo comparativo das estimativas de componentes de variância e parâmetros genéticos sob enfoque bayesiano em relação à abordagem clássica por máxima verossimilhança / Gil Carlos Rodrigues Medeiros. - 65f. : il. - Tese (Doutorado em Zootecnia). Universidade Federal de Pelotas. Faculdade de Agronomia Eliseu Maciel. Pelotas, 2015. - Orientador Nelson José Laurino Dionello ; coorientador Willian Silva Barros.

1.Zootecnia. 2.Melhoramento genético animal. 3.Análise de dados. 4.Histórico de análises. 5.Interface amigável. 6.Manipulação de dados. 7.Software científico. 8.Usabilidade. 9.Amostragem de Gibbs. 10.Inferência bayesiana. 11.Parâmetros genéticos. 12.REML. I.Dionello, Nelson José Laurino. II.Barros, Willian Silva. III.Título.

CDD: 658.05

Gil Carlos Rodrigues Medeiros

Uma abordagem computacional para estudo comparativo da estimativa de componentes de variância e parâmetros genéticos sob enfoque bayesiano em relação à abordagem clássica por máxima verossimilhança

Tese aprovada, como requisito parcial, para obtenção do grau de Doutor em Zootecnia, Programa de Pós-Graduação em Zootecnia, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas.

Data da Defesa: 30/04/2015

Banca examinadora:

Prof. Dr. Nelson José Laurino Dionello (Orientador)  
Doutor em Biotecnologia Agrícola pela Universidade Federal de Pelotas

Prof<sup>a</sup>. Dr<sup>a</sup>. Arione Augusti Boligon  
Doutora em Genética e Melhoramento Animal pela Universidade Estadual Paulista  
Júlio de Mesquita Filho

Prof. Dr. Fernando Flores Cardoso  
PhD em Animal Science pela Michigan State University

Prof. Dr. Luciano Carlos da Maia  
Doutor em Agronomia - Fitomelhoramento pela Universidade Federal de Pelotas

Prof. Dr. Eduardo Brum Schwengber  
Doutor em Ciências Biológicas (Genética) pela Universidade de São Paulo

***Dedico ...***

*A Deus, fonte inesgotável de toda a energia  
inspiradora que alimenta a ciência.*

*A minha esposa, Sandra Maria, em reconhecimento  
pela perene parceria, dedicação e compreensão.*

*A meus filhos, Daniel, Otávio e Adriano, dons variados  
que completam os meus sonhos.*

*A meus pais, Izabelino e Theresinha, in memoriam,  
exemplos de sabedoria e retidão na simplicidade,  
pelo legado dos valores fundamentais da vida.*

## **Agradecimentos**

Ao Prof. Nelson José Laurino Dionello, pela orientação e principalmente pela cordialidade, generosidade, estímulo e amizade com que assumiu o encargo nesta importante etapa de minha vida profissional.

Ao Prof. Willian Silva Barros, pela orientação, pelo incentivo e pela amizade.

Aos membros da Banca Examinadora, pela honra que me concederam.

Aos demais membros do grupo de pesquisa, bolsistas de iniciação científica ou de pós-graduação, que realizaram todas as tarefas de coleta dos dados e forneceram subsídios para a execução deste trabalho.

A todos os professores do Programa de Pós-Graduação em Zootecnia, em especial, aqueles com quem tive o prazer de conviver em sala de aula, pela acolhida e oportunidade de reavivar os vínculos com a área de minha graduação e integrar a experiência capitalizada da área de tecnologia da informação.

Aos colegas de estudo no curso, em especial, Edenilse, Bruno e Patrícia.

À colega Elisia, parceira de estudo dos mistérios da Estatística, e interlocutora paciente de intermináveis discussões científicas e filosóficas.

Aos meus irmãos, Flávio, Enio, Cláudio e Ricardo, pela confiança e constante incentivo.

Aos meus inúmeros alunos, fontes de alegria e inspiração no trabalho.

Aos colegas professores que, de variadas formas, me incentivaram a persistir e galgar todos os degraus, por mais difíceis que se apresentassem nestes quase 40 anos de docência. Em especial, destaco os Profs. João Baptista da Silva e Fernando Diaz, pela minha introdução na docência, e o Prof. João Carlos Deschamps, pela minha persistente atuação em aplicações da informática na agropecuária.

Aos amigos da ABCS, em especial, Valmir Costa da Rosa, pela oportunidade de conviver vários anos com atividades relacionadas ao melhoramento animal.

À Universidade Federal de Pelotas, pela oportunidade e pelo apoio financeiro para a qualificação, e à Direção do Centro de Engenharias, pelo incentivo e apoio.

A todos aqueles que contribuíram, direta ou indiretamente, para a realização deste trabalho.

## Resumo

MEDEIROS, Gil C. R. **Uma abordagem computacional para estudo comparativo das estimativas de componentes de variância e parâmetros genéticos sob enfoque bayesiano em relação à abordagem clássica por máxima verossimilhança**. 2015. 66f. Tese (Doutorado) - Programa de Pós Graduação em Zootecnia. Universidade Federal de Pelotas, Pelotas.

A avaliação genética de animais é fundamentada em métodos analíticos que associam conhecimentos multidisciplinares, principalmente de genética, estatística e computação. Diversos aplicativos de software implementam esses métodos para obtenção de parâmetros genéticos; muitos somente interagem com o usuário através de arquivos de textos e linha de comando. O primeiro objetivo deste trabalho foi dotar o pesquisador com uma plataforma operacional amigável para integrar múltiplos aplicativos de análise de dados em uma única interface gráfica e facilitar a aplicação e a comparação de diferentes metodologias; o segundo foi comparar a aplicação de dois métodos de análise na avaliação genética de codornas, usando as facilidades dessa plataforma. A ferramenta de software foi desenvolvida e está disponível na forma de dois assistentes de operação. As diversas funções oferecidas incluem: uma interface gráfica de usuário, amigável, genérica e configurável, para integração de diferentes aplicativos científicos de análise de dados, atribuindo-lhes maior usabilidade; recursos para organização, execução, catalogação e recuperação de análises; facilidades para coleta e manipulação preparatória de dados e visualização de dados e relatórios. Esta ferramenta foi utilizada de forma integrada com os aplicativos WOMBAT e INTERGEN para a avaliação genética de 5726 codornas de corte, através das metodologias clássica e bayesiana. Os dados utilizados foram coletados de 10 gerações de um programa de melhoramento e referem-se ao *pedigree* e à característica peso corporal medida em sete idades - um, sete, 14, 21, 28, 35 e 42 dias. Foram realizadas diversas análises unicaracterísticas, relativas às idades avaliadas e a quatro grupos cumulativos de animais formados por 7, 8, 9 e 10 gerações sucessivas. As estimativas resultantes indicam que: a herdabilidade (no sentido restrito) é maior aos 35 dias ( $h^2=0,46$  para o grupo completo); há uma leve redução das herdabilidades, para todas as idades, com a agregação das gerações 8, 9 e 10; os componentes de variância apresentam elevado grau de persistência ao acrescentar as novas gerações. As comparações realizadas com os parâmetros genéticos obtidos pelas duas metodologias evidenciam que as mesmas são equivalentes. A abordagem clássica usando a metodologia REML é suficiente para a obtenção de boas estimativas dos parâmetros.

**Palavras-chave:** amostragem de Gibbs; análise de dados; histórico de análises; inferência bayesiana; interface amigável; manipulação de dados; parâmetros genéticos; REML; software científico; usabilidade.

## Abstract

MEDEIROS, Gil C. R. **Computational approach for comparative study of the estimated variance components and genetic parameters using bayesian methods in relation to the classical maximum likelihood methodology.** 2015. 66f. Thesis (Doctor Degree) - Programa de Pós Graduação em Zootecnia. Universidade Federal de Pelotas, Pelotas.

The genetic evaluation of animals is based on analytical methods that combine multidisciplinary knowledge, especially genetics, statistics and computer science. Genetic parameters can be obtained by these methods via several software applications; many of which interact with the user only through text files and command line. The first aim of this study was to provide the researcher with a user-friendly operating platform to integrate various data analysis applications in a single graphical interface, assisting and easing the use and comparison of different methodologies; the second was to compare the employment of two analytical methods for the genetic evaluation of quails, using the facilities of this platform. The developed computational tool was set up and is available in the form of two operating assistants. Different functions are offered including: a graphical user interface, friendly, generic and configurable, to integrate different scientific applications of data analysis, giving them increased usability; resources for planning, organizing, executing, registering and recovery of analyses; facilities for data manipulation (collects and preparation) and visualization (input data and reports). This tool was used integrated with the WOMBAT and INTERGEN applications for the genetic evaluation of 5726 quails, through the classical and bayesian methods. Data were collected from 10 generations of a breeding program and refer to body weight measured in seven ages - one, seven, 14, 21, 28, 35 and 42 days - and to pedigree. Univariate analysis were applied to several data sets relative to each age, grouping 7, 8, 9 or 10 successive generations. The estimates indicate that: the heritability (narrow sense) is higher at 35 days ( $h^2 = 0.46$  for the full group); there is a slight reduction of heritability for all ages, as the generations 8, 9 and 10 are included; the variance components have a high degree of persistence across generations 7 to 10. It is evidenced by comparisons between the corresponding genetic parameters that both methodologies are equivalent. The classical approach using the REML methodology is satisfactory for obtaining good estimates of the parameters.

**Key words:** Bayesian inference; data analysis; data manipulation; genetic parameters; Gibbs sampling; history of analysis; REML; scientific software; usability; user friendly interface.

## Lista de Figuras

### Artigo 1

- Figura 1 - Duas organizações para o mesmo conjunto de dados de codornas (extrato). .....23
- Figura 2 - Diagrama de relacionamento entre um Aplicativo Alvo e o Assistente de Execução do Guaiaca e suas estruturas. ....30
- Figura 3 - Imagem da janela principal do Assistente de Execução do Guaiaca exibindo exemplo de uso e destacando diversos componentes da interface genérica. ....32
- Figura 4 - Estrutura hierárquica da instalação completa do Guaiaca associado a dois aplicativos alvos a partir de um local escolhido. ....34
- Figura 5 - Exemplos de arquivos de configuração para vinculação entre a interface e dois aplicativos alvos (a) e (b) e trechos dos arquivos de guias de código (c) e de opções (d) da interface para um dos aplicativos vinculados. ....35
- Figura 6 - Janela principal (recorte) durante operação de seleção de uma análise já realizada e catalogada no histórico. ....36
- Figura 7 - Janela do descritor da análise atual e seus três grupos de atributos (recortes sobrepostos).....38

### Artigo 2

- Figura 1 - Gráficos comparativos dos componentes de variância e das herdabilidades, demonstrando similaridade plena entre resultados das duas abordagens. ....54

## Lista de Tabelas

- Tabela 1 - Resultados (estimativas dos parâmetros genéticos) dos testes de dimensionamento da amostragem de Gibbs para as cadeias MCMC via INTERGEN com os dados ponderais dos animais de 10 gerações para a característica PC42 .....52
- Tabela 2 - Estatísticas descritivas dos componentes de variância e de herdabilidades obtidos por análise bayesiana através de MCMC para pesos corporais de 1 a 42 dias em 10 gerações de codornas .....53
- Tabela 3 - Componentes de variância e herdabilidades para os pesos corporais de codornas em 7 idades nas populações crescentes dos 4 grupos cumulativos de gerações obtidos pelos dois métodos comparados (REML e bayesiano).....55

## Sumário

1	Introdução Geral.....	11
2	Artigo 1: Guaiaca - Uma ferramenta de apoio a execução de aplicativos científicos destinados a análises de dados .....	15
	Introdução .....	17
	Funcionalidades.....	21
	Entrada de Dados.....	22
	Manipulações de Dados .....	23
	Visualização dos Dados .....	24
	Tipos de aplicativos alvo .....	25
	Parâmetros configuráveis.....	25
	Detalhes técnicos e funções incorporadas .....	27
	Interface do usuário.....	30
	Aplicações e Exemplos.....	32
	Aplicações configuradas.....	32
	Exemplos de configurações de aplicativos alvos.....	34
	Catálogo histórico de análises.....	36
	Testes.....	38
	Estágio Atual e Disponibilidade.....	39
	Programas.....	40
	Documentação .....	41
	Funcionalidades a serem incorporadas no futuro.....	41
	Agradecimentos .....	42
	Referências.....	42
3	Artigo 2: Comparação de enfoque bayesiano em relação à metodologia clássica para estudo de evolução e tendências dos parâmetros genéticos de codornas de corte ao longo de várias gerações.....	44
	Resumo .....	45
	Abstract.....	46
	Introdução .....	47
	Material e Métodos .....	49
	Resultados e Discussão .....	51

Conclusões .....	58
Agradecimentos .....	59
Referências .....	59
4 Considerações Finais .....	62
Referências .....	63

## 1 Introdução Geral

Uma das bases fundamentais do melhoramento genético de animais é o estudo de parâmetros genéticos e herdabilidades das mais variadas características genéticas, principalmente buscando a evolução da produtividade das espécies de interesse econômico. Esta evolução é dependente de duas componentes primárias: a carga genética dos animais e as condições ambientais a que são submetidos. A primeira pode ser estudada e aprimorada para propiciar evolução por ganho genético e a segunda pode ser trabalhada para produzir ganhos por manejo. Ambas possuem suas limitações, sejam biológicas ou econômicas. Nessa classe de problemas, os objetivos primários incluem necessariamente o reconhecimento, o controle e a obtenção de boas estimativas sobre as componentes que influenciam a variação fenotípica nas populações de animais. Uma boa determinação das proporções destas influências é fundamental para reconhecer e melhorar a participação da carga genética nas manifestações fenotípicas, permitindo selecionar e propagar nas populações as características desejáveis dos genótipos superiores. O valor genético de um indivíduo pode ser estimado através de índices de seleção que representam combinações das informações disponíveis sobre o animal e seus parentes (MRODE, 2005).

A estatística sempre foi uma aliada indispensável para o desenvolvimento dos conhecimentos na área de genética e melhoramento animal. Dessa associação originou-se o desenvolvimento da genética quantitativa, representando a aplicação das leis de Mendel às características quantitativas e econômicas, em particular para a produção animal. As bases da biometria e da genética de populações, estabelecidas por Francis Galton, Ronald Fisher e Sewall Wright, tornaram-se efetivamente aplicadas a partir da década de 1940 com as ações pioneiras de Jay L. Lush, cujos trabalhos foram consolidados em clássico livro texto - *Animal Breeding Plans* (LUSH, 1945). A partir desse período, a difusão das técnicas de inseminação artificial e o advento dos computadores proporcionaram uma massificação dos registros de *pedigree* e de produção dos animais. Outro marco importante foi o desenvolvimento e a aplicação de metodologias analíticas caracterizadas nos

modelos mistos, que Charles Henderson reuniu em outra obra clássica - *Applications of linear models in animal breeding* (HENDERSON, 1984). Na década de 1980, os métodos bayesianos despertaram maior interesse e suscitaram um expressivo crescimento de pesquisas e aplicações para sua utilização, principalmente como decorrência do surgimento dos Métodos de Monte Carlo via Cadeias de Markov (MCMC) (PAULINO et al., 2003). As aplicações na área de melhoramento genético de animais também começaram a ser estudadas (GIANOLA; FERNANDO, 1986) e mais recentemente aparecem com bastante abrangência em diversos trabalhos relacionados (SORENSEN; GIANOLA, 2002; FALCÃO et al., 2004, 2009; WINTER et al., 2006; HIDALGO et al., 2011).

Diversos aplicativos destinados a implementar e tornar viáveis estas metodologias foram surgindo e evoluindo, concomitantemente ao desenvolvimento da computação. Um dos mais reconhecidos é o SAS (*Statistical Analysis System*) (DELWICHE; SLAUGHTER, 2012; SAS INSTITUTE INC, 2010), que representou e ainda representa um paradigma para muitas outras iniciativas. Dentre as implementações focadas na abordagem clássica das aplicações de modelos lineares mistos baseadas na metodologia da máxima verossimilhança restrita (REML), pode ser destacado o software WOMBAT (MEYER, 2007). O programa INTERGEN (CARDOSO, 2008) é um dos produtos oriundos de iniciativas visando uma abordagem centrada nos métodos bayesianos, que implementam os MCMC e outros recursos de simulação relacionados, como a amostragem de Gibbs.

Os dois últimos (e muitos outros exemplos) fazem parte de uma classe de programas que:

- a) representam soluções tecnológicas eficientes para importantes problemas de interesse acadêmico, acessíveis a baixo ou nenhum custo, e
- b) são construídos com foco nos métodos específicos implementados, sem maior atenção a aspectos relacionados à usabilidade (MADAN; DUBEY, 2012) e à disponibilidade de recursos complementares, implementando interfaces muito simples, mas carentes com relação à facilidade de uso, à ergonomia e a outros fatores que influenciam o grau de satisfação do usuário.

A disponibilidade de uma interface gráfica de usuário, definida como uma "central de operações", com desenho uniforme, envolvendo um conjunto expressivo das operações recorrentes necessárias ao pesquisador da área, pode simplificar e

facilitar suas atividades, garantir maior confiabilidade aos dados tratados, valorizar a utilização dessas importantes ferramentas e permitir a navegação entre elas com naturalidade. Os estudos comparativos entre metodologias são exemplos de abordagens operacionais que podem ser bastante beneficiadas por este tipo de recurso.

Uma nova linhagem de codornas de corte vem sendo desenvolvida, há vários anos, pelo Departamento de Zootecnia da Universidade Federal de Pelotas, através de seu Programa de Pós-Graduação em Zootecnia (PPGZ). Nesse programa de melhoramento genético, a seleção das aves para a produção das gerações seguintes é realizada pelo peso corporal aos 42 dias de idade. Parâmetros genéticos para algumas características de interesse econômico vem sendo estimados por abordagem clássica via metodologia REML (DIONELLO et al., 2008; 2009; MEDEIROS et al., 2012). Em uma abordagem inicial, usando os programas MTGSAM (VAN TASSEL; VAN VLECK, 1996), foram aplicados os métodos bayesianos sobre os dados ponderais das primeiras sete gerações, sendo constatadas pequenas diferenças nas estimativas desses parâmetros em relação aos obtidos pela metodologia clássica (DIONELLO et al., 2012). Estudos comparativos mais amplos, que avaliem a evolução dos componentes de variância e das herdabilidades no decorrer de sequências de gerações, podem indicar com mais precisão a influência da metodologia na estimação dos parâmetros.

Os objetivos deste trabalho foram:

- a) desenvolver ferramentas computacionais que viabilizem um ambiente integrado e automatizado para as diversas atividades do pesquisador relacionadas à realização de análises de dados de melhoramento genético de animais, incluindo:
  - interface gráfica de usuário, amigável, genérica e configurável, que agregue maior usabilidade a aplicativos científicos de análise de dados cujos modelos de interação com o usuário são baseados em linha de comando;
  - recursos para organização, catalogação, recuperação e visualização de análises realizadas;
  - facilidades para coleta, manipulação e visualização de dados;
- b) realizar análise comparativa dos parâmetros genéticos resultantes da seleção de codornas de corte, obtidos por duas metodologias - REML e

inferência bayesiana - através do uso de aplicativos conhecidos integrados com as ferramentas desenvolvidas.

Estes dois objetivos são abordados respectivamente nos Capítulos 2 e 3, por meio de dois artigos preparados na forma preconizada pelos periódicos a que se destinam. O Capítulo 2 é dedicado a apresentar a caracterização da ferramenta desenvolvida, descrevendo os recursos modelados e implementados, tanto para apoio à execução de análises como para a manipulação de dados. No artigo, são incluídos exemplos de configurações para integração com os aplicativos utilizados para a execução das análises relativas às metodologias comparadas. No Capítulo 3, são apresentados os resultados de 56 análises realizadas para determinar os componentes de variância e as herdabilidades, que são avaliados comparativamente via análise unicaracterística utilizando os aplicativos WOMBAT e INTERGEN, com o apoio dos recursos da ferramenta desenvolvida. As considerações finais, reunidas no Capítulo 4, encerram este texto com a apresentação de uma síntese das conclusões dos dois artigos.

## **2 Artigo 1: Guaiaca - Uma ferramenta de apoio a execução de aplicativos científicos destinados a análises de dados**

Este artigo está apresentado conforme as normas do periódico a que será submetido. As páginas seguintes deste capítulo (16 a 43) equivalem às páginas do artigo, numeradas junto à margem inferior, independentemente deste contexto. A numeração contínua convencional, utilizada para os índices, está indicada em baixa intensidade junto à margem superior.

# 1       **Guaiaca - Uma ferramenta de apoio a execução de aplicativos** 2       **científicos destinados a análises de dados**<sup>1</sup>

3                   **Gil Carlos Rodrigues Medeiros**<sup>2</sup>, **Willian Silva Barros**<sup>3</sup>,  
4                   **Nelson José Laurino Dionello**<sup>4</sup>

5       <sup>1</sup> *Projeto financiado pelo CNPq - processo nº 484626/2007-8.*

6       <sup>2</sup> *Doutorando do Programa de Pós-Graduação em Zootecnia - UFPel, Brasil. Docente do*  
7       *Centro de Engenharias - UFPel. e-mail: gil.medeiros@ufpel.edu.br.*

8       <sup>3</sup> *Docente do Departamento de Matemática e Estatística - UFPel, Brasil. e-mail:*  
9       *willian.barros@ufpel.edu.br*

10       <sup>4</sup> *Docente do Departamento de Zootecnia - FAEM - UFPel, Brasil. Pesquisador do CNPq.*  
11       *e-mail: dionello@ufpel.edu.br*  
12

13           RESUMO - O objetivo deste trabalho foi desenvolver uma ferramenta de software  
14 com interface gráfica amigável, genérica e configurável, para facilitar pesquisadores na  
15 organização e gerenciamento das atividades de manipulação e análises de dados. A  
16 concepção foi baseada nas características de dois focos principais: a preparação de  
17 dados para análises em pesquisas de melhoramento genético de animais e o incremento  
18 da usabilidade para aplicativos que implementam importantes métodos de análises, mas  
19 interagem com o usuário via arquivos de texto e linha de comando. A ferramenta  
20 Guaiaca foi modelada como um "cinto de utilidades" e desenvolvida na forma de dois  
21 assistentes integrados. Pode ser combinada com múltiplos aplicativos executores de  
22 análises. Oferece facilidades para: coleta, manipulação e reorganização de dados;  
23 organização, catalogação e recuperação de análises via descritores; ativação de  
24 processos de análise. Apresenta-se uma descrição sobre os parâmetros configuráveis  
25 para os assistentes, as estruturas de dados mantidas e a forma de relacionamento entre a  
26 interface genérica e os aplicativos de análise combinados. O Guaiaca foi testado com  
27 dados de 5.726 codornas e 650.000 suínos. A personalização de sua interface foi testada  
28 para peculiaridades dos aplicativos WOMBAT (modelos mistos) e INTERGEN  
29 (modelos bayesianos). Foram realizadas diversas análises para comprovar sua  
30 generalidade e facilidade de instalação, operação e configuração. O uso do conjunto  
31 Guaiaca-WOMBAT é ilustrado com exemplo de análise preparada e visualizada no  
32 Guaiaca e executada pelo WOMBAT. A versão beta do Guaiaca (para Windows) está  
33 disponível de forma gratuita para a comunidade científica e pode ser obtida em  
34 <http://www.ufpel.edu.br/~gil.medeiros/guaiaca> .

35           Palavras-chave: histórico de análises, interface amigável, manipulação de dados,  
36 melhoramento animal, modelos mistos, software científico, usabilidade

37           ABSTRACT - The aim of this work was to develop a software tool that provides a  
38 user friendly graphical interface, with generic and configurable design, to ease the  
39 researcher's tasks in data manipulation and data analysis. Two primary targets have  
40 influenced in the design: data preparation for analysis in animal breeding programs and  
41 usability increase for the applications which perform important methods of analysis, but  
42 interact with user only by text files and command lines. The Guaiaca tool was designed  
43 like a "utility belt" in the form of two integrated assistants; it can be combined with  
44 multiple analysis applications. This tool offers facilities to: collect, manipulation and  
45 reorganization of data; organization, registering and retrieval of analysis through  
46 descriptors; preparation and activation of the analysis processes. The parameters to  
47 configure the assistants, the main data structures maintained, and the form of  
48 relationship between the generic interface and the combined analysis applications are  
49 described in this paper. The Guaiaca's functions for data manipulation were tested and  
50 validated through data of 5,726 quails and 650,000 pigs. Its generic interface has been  
51 customized and validated for peculiarities of two applications, WOMBAT (mixed  
52 models) and INTERGEN (bayesians models). Several analysis were carried out to prove  
53 its generality and ease of installation, operation and configuration. Use of Guaiaca-  
54 WOMBAT integrated is illustrated with an example of analysis, prepared and visualized  
55 on the Guaiaca and executed by WOMBAT. The Guaiaca's beta version for Windows is  
56 available free of charge to the scientific community and can be obtained at  
57 <http://www.ufpel.edu.br/~gil.medeiros/guaiaca> .

58           Key Words: animal breeding, data manipulation, history of analysis, mixed  
59 models, scientific software, usability, user friendly interface

60

61

## **Introdução**

62           Toda a atividade de pesquisa experimental presume a coleta, a manipulação e a  
63 análise de dados. Embora, em alguns casos, a execução manual dessas tarefas possa ser  
64 viável, é inimaginável atualmente a dispensa do apoio de aplicações computacionais  
65 científicas para realizá-las de forma eficaz e eficiente. Em muitos casos, é impossível  
66 realizar tais tarefas sem o uso dessas aplicações. A disponibilidade de aplicações  
67 científicas adequadas a cada necessidade da pesquisa é dependente de vários fatores de

68 ordem econômica, técnica, política, idiomática, ética, inovativa e pessoal, dentre outros.  
69 Frequentemente, o pesquisador é motivado, por alguns desses fatores, a desenvolver as  
70 aplicações de que necessita, agregando ou não colaboradores qualificados das áreas  
71 específicas.

72 Na pesquisa agropecuária, em geral, e na área de melhoramento genético, em  
73 particular, diversos produtos têm sido gerados no escopo desta motivação, tanto para  
74 implementar técnicas relacionadas aos métodos estatísticos e de simulação, como para  
75 particularizar ou especializar estas e outras técnicas (Resende, 2007; Cruz, 2013;  
76 Ferreira, 2013). Destas iniciativas e da agregação de técnicas computacionais avançadas  
77 emergiu, por exemplo, a área de bioinformática, principalmente relacionada à área de  
78 biologia molecular e, mais particularmente, às aplicações em seleção genômica e  
79 seleção genômica ampla (Legarra et al., 2009; Garrick and Saatchi, 2011).

80 A preocupação fundamental ao desenvolver um software de aplicação científica é,  
81 sem dúvida, o provimento imediato dos algoritmos correspondentes às funções objetos  
82 do aplicativo. Aspectos relacionados à usabilidade e disponibilidade de recursos  
83 complementares de apoio às atividades inerentes ao uso desses aplicativos costumam  
84 ser simplificados ao mínimo imprescindível, conforme a necessidade de utilização do  
85 idealizador/desenvolvedor e seus pares imediatos. Muitas vezes esta estratégia, embora  
86 necessária nas fases iniciais, pode determinar o desinteresse de potenciais usuários pela  
87 continuidade de uso de uma ferramenta, após experiência inicial. Dentre as razões desse  
88 desinteresse, está a dificuldade de uso ou simplesmente a dificuldade de reconhecer a  
89 facilidade de uso, mesmo que a ferramenta possa ser considerada de boa qualidade  
90 quanto às suas funcionalidades, mesmo que ela represente uma importante alternativa a  
91 recursos reconhecidos, cujos custos sejam muito elevados. Independentemente da  
92 abordagem dessa abstração, a usabilidade é um importante atributo na apreciação da

93 qualidade de sistemas de software (Madan & Dubey, 2012) e está fortemente associada  
94 à facilidade de uso, à ergonomia e ao grau de satisfação do usuário. Tarefas de apoio,  
95 tais como as inerentes à manipulação de dados, são normalmente remetidas ao uso  
96 prévio ou posterior de ferramentas reconhecidas de uso geral (planilhas eletrônicas e  
97 similares), que também podem gerar custos e demandar outros conhecimentos.

98 O custo de desenvolvimento de interfaces operacionais para ferramentas de  
99 aplicação específica faz com que a tarefa seja frequentemente relegada a uma etapa  
100 posterior que acaba sendo protelada indefinidamente.

101 Para a área de melhoramento genético de animais, são conhecidos vários  
102 aplicativos que definem suas interfaces de usuário de forma simplificada e dependente  
103 do padrão de interação correspondente ao modo linha de comando em janela de texto.  
104 Dentre outros, podem ser citados os seguintes exemplos: DFREML, MTDFREML,  
105 WOMBAT, INTERGEN, ASReML, família BLUPF90 (Meyer, 1988, 2007; Boldman et  
106 al., 1995; Cardoso, 2008; Gilmour et al., 2009; Misztal et al., 2014). As finalidades  
107 destas ferramentas são fundamentais para a área, pois implementam modelos e métodos  
108 de inferência e análise quantitativa de dados, baseados na metodologia da máxima  
109 verossimilhança restrita (REML) com o uso de modelos mistos, permitindo a utilização  
110 simultânea de informações de *pedigree* e parentesco entre os animais. Seus processos  
111 são altamente iterativos e muitas vezes exigem a execução repetitiva de análises para a  
112 busca de uma melhor aproximação e/ou adequação de modelo empregado. Dionello et  
113 al. (2008) destacam a utilização de alguns desses recursos (modelos e aplicativos) em  
114 avaliações genéticas sobre o crescimento de codornas de corte.

115 Recursos complementares normalmente utilizados para a manipulação de dados  
116 incluem: Excel, Word, Notepad++ (no Windows) e gedit, xedit, vi e programas  
117 plotadores de gráficos (no Linux). Uma atividade complementar que pode ser

118 considerada como primária é a coleta de dados. Mesmo com a generalizada utilização  
119 de recursos computacionais móveis, ainda observamos processos de coleta de dados  
120 baseado em planilhas ou fichas de papel e lápis, que posteriormente passam por tarefas  
121 de digitação e validação. Os erros e perdas de dados inerentes a esses processos  
122 poderiam ser evitados ou eliminados na origem com o uso de ferramenta apropriada e  
123 vinculada ao banco de dados desejado. Outro aspecto importante é a possibilidade de  
124 coleta automática de dados via dispositivos de medição que oferecem recursos para  
125 conexão *online* com um computador.

126 Muitos aplicativos científicos desenvolvidos segundo o perfil ora caracterizado  
127 definem as suas interfaces de interação com o usuário baseadas em um padrão  
128 representado por arquivos de parâmetros para definição das tarefas/análises a serem  
129 realizadas e por arquivos de relatórios na forma de textos planos. Em alguns casos, são  
130 definidas portas de interação com outros aplicativos por meio de arquivos específicos  
131 com dados intermediários ou conclusivos produzidos por seus processos. Este tipo de  
132 aplicativos é o alvo primário da interface proposta neste trabalho, visando a execução de  
133 análises de dados de pesquisas.

134 Este documento foi elaborado com o objetivo de descrever uma ferramenta,  
135 denominada Guaiaca, que agrega características de maior usabilidade a aplicativos  
136 científicos de análise de dados cujos modelos de interação com o usuário seguem o  
137 padrão acima identificado. Isto é realizado integrando-os a recursos (facilidades) para  
138 coleta e manipulação de dados e para organização, catalogação e recuperação de  
139 análises realizadas, através de uma interface de usuário amigável, genérica e  
140 configurável.

141

142

## Funcionalidades

143 A versão inicial do Guaiaca foi desenvolvida em um ambiente Delphi® para a  
144 plataforma Windows, na forma de um protótipo que servirá de base para a posterior  
145 implementação de versão multiplataforma. No desenvolvimento, foram utilizados  
146 componentes elementares de bancos de dados, como ponto de partida para a entrada e a  
147 organização das coleções de dados dos indivíduos testados em projetos. As interfaces de  
148 operação (GUI<sup>1</sup>) foram projetadas para proporcionar interação fácil e intuitiva, por meio  
149 de janelas e recursos simples, configuráveis de acordo com cada projeto e com as  
150 necessidades e preferências do usuário. O propósito principal é oferecer ao pesquisador  
151 recursos simples, práticos e de fácil uso para tarefas de: (a) coleta, organização,  
152 visualização e manipulação de dados de pesquisas; (b) execução de análises de dados  
153 com o uso de aplicativos científicos selecionados e (c) manutenção de registro histórico  
154 e de controle sobre as análises realizadas. As tarefas do segundo grupo correspondem à  
155 preparação e ao encaminhamento das análises aos aplicativos executores. O terceiro  
156 grupo, estreitamente relacionado com o segundo, inclui organização, armazenamento  
157 (catalogação) e visualização dos parâmetros e resultados dessas análises. Os aplicativos  
158 executores das análises são genericamente denominados de **aplicativos alvos** no  
159 contexto da interface genérica.

160 Embora concebido sob a ótica das necessidades relacionadas às pesquisas na área  
161 de zootecnia, o Guaiaca pode ser empregado para qualquer outra área cujos dados sejam  
162 organizados de forma semelhante e processados por aplicativos com características  
163 similares às descritas no escopo deste documento. Sua utilização pode ser caracterizada  
164 de duas formas: como um programa único incorporando todas as funcionalidades ou em

---

<sup>1</sup> O acrônimo GUI, do inglês *Graphical User Interface*, é usado para representar a expressão "interface gráfica do usuário" - um tipo de interface de operação que permite a interação do usuário com dispositivos digitais, de forma intuitiva, através de elementos gráficos (ícones e outros elementos visuais, também denominados de *widgets*), contrastando com a interface de linha de comando.

165 dois módulos, o Assistente de Dados e o Assistente de Execução. O primeiro agrupa as  
166 funcionalidades de coleta e manipulação de dados. O segundo reúne todos os recursos  
167 relativos à interface genérica para os aplicativos alvos (análises estatísticas de dados ou  
168 outros tipos de análises) e à catalogação de análises.

#### 169 *Entrada de Dados*

170 A funcionalidade de entrada de dados, diretamente relacionada com as operações  
171 de coleta de dados em laboratório ou a campo, é apresentada ao usuário na forma de  
172 sete funções: criação de novo arquivo de dados, seleção (abertura) de arquivo de dados  
173 existente para operações, cadastramento inicial de novos registros de dados de  
174 indivíduos, alteração de dados cadastrais (dados pré-fixados) de indivíduos, inclusão de  
175 novas medidas (variáveis testadas) de indivíduos, alteração de medidas já catalogadas e  
176 exclusão de registros de dados de indivíduos.

177 A interface é configurável para estabelecer uma filtragem das informações que  
178 serão tratadas pelo operador, conforme a etapa da coleta de dados dos indivíduos,  
179 evitando o manuseio direto de todo o conjunto de dados. O usuário poderá ver os  
180 demais dados do indivíduo em atualização e sua vizinhança, mas a atualização estará  
181 restrita apenas aos campos de dados associados à etapa em questão. Exemplificando,  
182 durante a coleta de dados relativa à terceira pesagem de codornas de um grupo, o  
183 usuário pode configurar a interface para abrir apenas os campos de identificação do  
184 indivíduo e das medidas relativas a essa pesagem (peso, idade, etc), embora os demais  
185 campos sejam visualizados para localização e eventuais avaliações no ato da coleta.

186 A função de criação de novo arquivo de dados representa uma operacionalização  
187 do planejamento do experimento, em que são definidos os campos de dados a serem  
188 coletados e tratados pelo projeto, incluindo a caracterização dos dados, sejam contínuos  
189 ou discretos, numéricos ou alfanuméricos, o nível de precisão etc. Uma facilidade

190 disponível nesta função é a possibilidade de copiar a estrutura de outro conjunto de  
 191 dados similar e fazer os ajustamentos necessários para a composição do novo arquivo.

### 192 *Manipulações de Dados*

193 O aplicativo Guaiaca inclui funções para algumas das mais recorrentes conversões  
 194 de formatos e estruturas dos dados relativas à etapa de preparação para o processamento  
 195 de análises em diversos aplicativos alvos. Dentre essas funções destacam-se:  
 196 (a) ordenação dos registros por campos selecionados, (b) reorganização estrutural dos  
 197 dados por critérios usuais e (c) reorganização formatadora dos dados na mesma  
 198 estrutura original. Um extrato curto de um banco de dados de codornas é usado na  
 199 Figura 1, para ilustrar aspectos de uma reorganização mista (estrutural e formatadora).

(a) Estrutura de um registro por animal com dados de identificação e várias características,  
 vírgula decimal e indicador (#) de dado faltante, sem repetição de campos

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
40834	30442	30138	3	1	9,3	1	28,5	7	82	14	140,5	21	198	28	226	35	267,5	42
40835	30302	30627	3	2	9	1	#	7	89,5	14	149,5	21	222	28	219,5	35	257,5	42
40836	30039	30886	3	2	9,5	1	37	7	97,5	14	164	21	232	28	249,5	35	280,9	42

(b) Estrutura de um registro por característica com ponto decimal, repetição de dados nos registros  
 de um mesmo animal e inclusão de dados de identificação sobre a característica

1	2	3	4	5	6	7	8	9
1	40834	30442	30138	3	1	40834	9.3	1
1	40834	30442	30138	3	1	40834	28.5	7
1	40834	30442	30138	3	1	40834	82	14
1	40834	30442	30138	3	1	40834	140.5	21
1	40834	30442	30138	3	1	40834	198	28
1	40834	30442	30138	3	1	40834	226	35
1	40834	30442	30138	3	1	40834	267.5	42
1	40835	30302	30627	3	2	40835	9	1
1	40835	30302	30627	3	2	40835	89.5	14
1	40835	30302	30627	3	2	40835	149.5	21
1	40835	30302	30627	3	2	40835	219.5	35
1	40835	30302	30627	3	2	40835	222	28
1	40835	30302	30627	3	2	40835	257.5	42
1	40836	30039	30886	3	2	40836	9.5	1
1	40836	30039	30886	3	2	40836	3	7
1	40836	30039	30886	3	2	40836	97.5	14
1	40836	30039	30886	3	2	40836	164	21
1	40836	30039	30886	3	2	40836	232	28
1	40836	30039	30886	3	2	40836	249.5	35
1	40836	30039	30886	3	2	40836	280.9	42

Numeração de  
 ordem dos  
 campos

200

201 Figura 1 - Duas organizações para o mesmo conjunto de dados de codornas (extrato).

202 A primeira tabela (Figura 1a) caracteriza a estrutura usual de um banco de dados  
203 sobre animais, em que o animal é a entidade sobre a qual os registros de dados são  
204 organizados. Na segunda tabela (Figura 1b), os registros de dados são organizados em  
205 função de características específicas e os dados de identificação, incluindo informações  
206 de *pedigree* neste caso, precisam ser repetidos para cada característica (ou grupo)  
207 avaliada. Numa das organizações, é usado o símbolo decimal correspondente ao idioma  
208 e na outra o símbolo exigido pelo aplicativo que processará os dados.

### 209 *Visualização dos Dados*

210 Dados organizados diretamente pelo aplicativo ou importados de outras fontes  
211 podem ser visualizados e destacados na tela de modo integral ou seletivo. Este recurso  
212 permite realizar a seleção (busca) por critérios múltiplos e aplicar dois modelos de  
213 apresentação/navegação: no contexto da coleção integral dos dados (navegação seletiva)  
214 e na forma de extrato dos dados filtrados pelos critérios adotados (visualização seletiva).  
215 O primeiro modelo permite a dupla forma de navegação: entre os registros do arquivo  
216 inteiro e entre os registros selecionados. O segundo modelo oferece, além da  
217 visualização seletiva, a possibilidade de gravação/salvamento do subconjunto dos dados  
218 obtidos pela aplicação dos critérios de seleção, ou seja, a obtenção de um novo arquivo  
219 contendo um extrato seletivo dos dados do arquivo original em operação. A busca de  
220 registros de dados pode ser realizada de duas formas: individual ou coletiva. Na forma  
221 individual, o registro desejado pode ser selecionado por sua ordem no arquivo ou pelo  
222 valor do campo chave de identificação dos indivíduos, independentemente da ordenação  
223 estabelecida para o arquivo. Na forma coletiva, o usuário pode definir os campos a  
224 serem usados como critérios e os valores desses campos para selecionar um grupo de  
225 registros que atenda aos critérios indicados. Os campos deixados em branco (valores

226 não fornecidos) não influenciam no resultado da busca, ou seja, não restringem os  
227 registros selecionados.

### 228 *Tipos de aplicativos alvo*

229 As funcionalidades relacionadas ao Assistente de Dados são destinadas a atender  
230 as necessidades de preparação de dados para o processamento por qualquer tipo de  
231 aplicativo alvo.

232 Embora a generalidade da interface do Assistente de Execução permita uma gama  
233 mais ampla de aplicações, o propósito com que foi definida está diretamente  
234 relacionado às características dos aplicativos que:

235 a) executam cálculos e análises especializados sobre dados fornecidos via  
236 arquivos indicados, tais como análises estatísticas de dados, simulações e avaliações ou  
237 ajustamentos de modelos;

238 b) carecem de uma interface visual de operação (interface gráfica de usuário);

239 c) são parametrizadas por meio de atributos, códigos, declarações, comandos e  
240 opções indicados em um arquivo de texto e/ou argumentos de linha de comando;

241 d) geram relatórios em arquivos de textos e no próprio console de operação (tela  
242 do computador).

### 243 *Parâmetros configuráveis*

244 Na interface do Assistente de Dados, os formulários digitais de entrada de dados  
245 são personalizados, destacando os campos pertinentes, de acordo com as necessidades  
246 do pesquisador em cada etapa da coleta de dados. São mantidas informações sobre a  
247 última atividade realizada, permitindo a retomada de uma operação não concluída nas  
248 mesmas condições anteriores, ou seja, mesmo arquivo, mesmo registro e mesma  
249 personalização do formulário digital.

250 A interface genérica do Guaiaca para o módulo Assistente de Execução deve ser  
251 configurada sob os dois escopos de seu uso: o da própria interface, caracterizado pela  
252 definição de preferências de uso através de opções disponíveis ao usuário, e o da  
253 execução das análises, caracterizado pela vinculação da interface a cada aplicativo alvo  
254 a ser utilizado. Para cada vinculação configurada, que é identificada por um nome  
255 único, um novo item é inserido no menu correspondente e um novo diretório de  
256 configuração é criado no ambiente do Guaiaca. Esse menu permite o usuário ativar a  
257 personalização da interface para o aplicativo alvo, conforme a configuração  
258 selecionada. Além de configurar a vinculação a diferentes aplicativos, é possível  
259 estabelecer diferentes configurações para um mesmo aplicativo alvo. O diretório de  
260 configuração tem a finalidade de manter todos os arquivos referentes a essa  
261 personalização, conforme os parâmetros apresentados a seguir.

262 Os aplicativos alvos são instalados de forma independente, conforme a orientação  
263 específica de sua origem. Para seu uso combinado com a interface genérica do  
264 Assistente de Execução, será necessário definir um local (diretório/pasta) específico  
265 para área de trabalho, onde serão realizadas as suas funções. Esta área é considerada  
266 como temporária e será limpa ao início de cada nova operação do aplicativo alvo, ou  
267 seja, a cada execução de análise. É possível definir uma mesma área de trabalho para  
268 vários ou todos os aplicativos alvos, desde que não haja limitações impostas por seus  
269 requisitos e pela forma de uso da interface. Durante a preparação para execução do  
270 aplicativo alvo e após o seu processamento, o conteúdo da área de trabalho poderá ser  
271 salvo (copiado) para local permanente determinado pelo usuário. Isto é feito ao  
272 caracterizar e registrar a análise desejada, qualquer que seja o estado (preparada ou  
273 realizada).

274 Para vincular (conectar) a interface do Guaiaca a cada aplicativo alvo são  
275 adotados os seguintes parâmetros a serem configurados e mantidos em pasta e arquivos  
276 específicos:

- 277 a) Local (diretório/pasta) de instalação do aplicativo alvo;
- 278 b) Nome do arquivo executável do aplicativo alvo;
- 279 c) Nome do arquivo de ajuda/manual do aplicativo alvo;
- 280 d) Nome da área de trabalho (diretório/pasta);
- 281 e) Nome dos arquivos para representar a entrada e a saída do console (entrada e  
282 saída padrão);
- 283 f) Modo de informação do arquivo de parâmetros (linha de comando ou entrada  
284 do console);
- 285 g) Filtros para reconhecimento de arquivos de saída;
- 286 h) Nomes dos arquivos primários de saída (exibidos automaticamente);
- 287 i) Guia de código (arquivo opcional);
- 288 j) Guia de opções (arquivo opcional) e
- 289 k) Arquivos de exemplo (opcional) para parâmetros e dados.

290 Estes parâmetros são configurados via formulário apresentado pela própria  
291 interface genérica.

#### 292 *Detalhes técnicos e funções incorporadas*

293 Além das funções já descritas (entrada/coleta, visualização e manipulações da  
294 organização de dados e personalização da interface para um aplicativo alvo), o  
295 Assistente de Execução incorpora outros dois grupos de funções: as funções de apoio ao  
296 uso do aplicativo alvo e as funções necessárias para a manutenção de um catálogo  
297 histórico de análises realizadas.

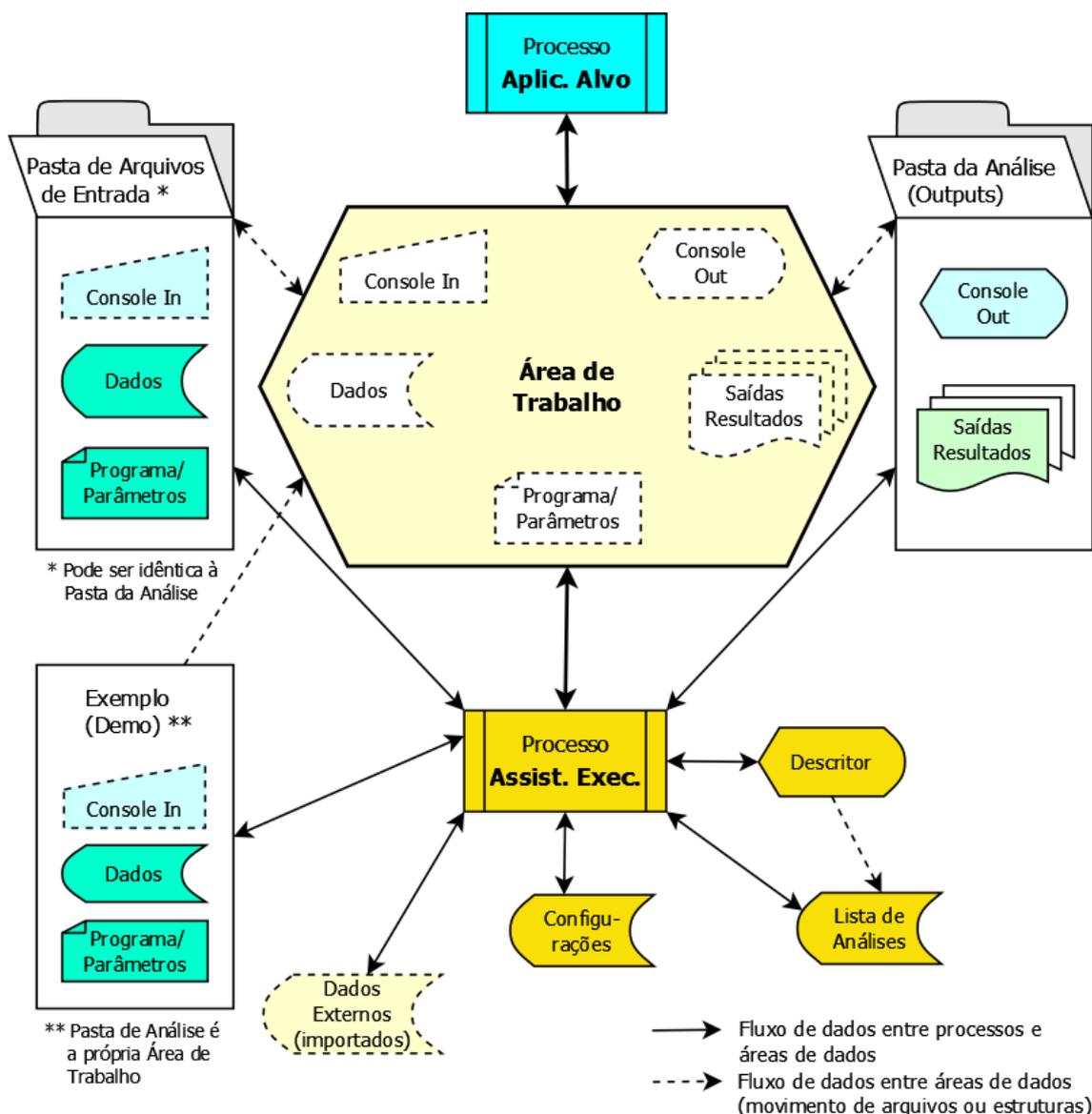
298 As funções de apoio ao uso do aplicativo alvo constituem um grupo de facilidades  
299 para: (a) visualização dos arquivos de entrada e saída, (b) preparação ou visualização de  
300 arquivos de parâmetros (também denominados de programas ou roteiros) e (c) escolha e  
301 indicação de opções (argumentos) de execução. Para que as duas últimas facilidades (b  
302 e c) possam ser usadas em sua plenitude, é necessário criar previamente os arquivos que  
303 representam o Guia de Códigos e o Guia de Opções, de acordo com o interesse do  
304 usuário pesquisador e seguindo as orientações da documentação da interface. Isto pode  
305 ser realizado pelo usuário ou pessoal de apoio da área de tecnologia da informação, a  
306 partir da documentação de recursos do aplicativo alvo, durante a instalação do conjunto  
307 de aplicações. Alternativamente, um padrão recomendado para esses arquivos pode ser  
308 preparado previamente e distribuído como componente deste Assistente de Execução ou  
309 incluído no pacote distribuído pelos fornecedores ou mantenedores de cada aplicativo  
310 alvo. Para a visualização dos arquivos de entrada e de saída do aplicativo alvo (dados,  
311 parâmetros e relatórios) foi adotada estratégia similar à empregada pelo SAS - um  
312 software comercial tradicional de gestão e análise avançada de dados (Delwiche &  
313 Slaughter, 2012; SAS Institute Inc, 2010). O usuário tem a sua disposição, acessíveis  
314 com poucos cliques via indicadores (abas, botões ou menus), janelas integradas com  
315 todos os arquivos importantes para sua apreciação sobre cada análise projetada ou  
316 realizada.

317 No contexto do Guaiaca, cada análise é representada por um Descritor de Análise,  
318 cujo objetivo é manter uma caracterização da tarefa, indicando os recursos envolvidos e  
319 informações relativas aos resultados obtidos. Os descritores podem ser catalogados para  
320 constituir uma Lista de Análises, ou seja, um catálogo histórico de análises realizadas.  
321 Este catálogo é mantido por um grupo de funções que facilitam a organização e o  
322 rastreamento de dados e resultados processados pelos aplicativos alvos vinculados à

323 interface. Este recurso também pode ser usado de forma independente de aplicativo  
324 alvo. É possível organizar e identificar cada análise projetada, reunindo e armazenando  
325 os arquivos associados e mantendo informações adicionais relativas à execução, tais  
326 como, data e tempo de execução, aplicativo processador da análise e anotações  
327 complementares do usuário analista. As funções definidas permitem navegar pela lista  
328 de análises catalogadas, criar novas análises, selecionar e visualizar descritores e/ou  
329 arquivos das análises já realizadas, complementar ou reexecutar análises, salvar  
330 arquivos de dados e resultados das análises no ambiente gerenciado. É recomendável  
331 que esse catálogo seja mantido em uma pasta denominada "Projetos" incorporada ao  
332 ambiente da instalação do Guaiaca.

333         A concepção do relacionamento entre o Assistente de Execução e os aplicativos  
334 alvos baseia-se na coexistência de dois processos independentes no sistema operacional  
335 do computador: o processo Guaiaca e o processo correspondente ao aplicativo alvo  
336 vinculado. Todas as operações do aplicativo alvo são canalizadas a uma área de trabalho  
337 comum gerenciada pelo Assistente. Este prepara e alimenta a área de trabalho com os  
338 arquivos necessários ao processamento de uma análise pelo aplicativo alvo.  
339 Posteriormente, redireciona os resultados obtidos para armazenamento permanente  
340 conforme as especificações do pesquisador indicadas no descritor da análise. Um  
341 diagrama demonstrativo desse relacionamento é apresentado na Figura 2. Nesse  
342 diagrama, são destacados os principais elementos organizacionais definidos pela  
343 interface relativos a uma análise, tais como a pasta de arquivos de entrada, a pasta da  
344 análise, os processos e sua área de trabalho comum, o ambiente de configuração e da  
345 manutenção da lista de análises catalogadas, bem como origens alternativas de dados  
346 para caracterizar uma análise. A caracterização de uma pasta de arquivos de entrada  
347 diferenciada da pasta da análise deve-se ao fato de que um mesmo conjunto de dados

348 pode ser usado em diferentes análises. Isto permite evitar a multiplicação de instâncias  
 349 dos mesmos dados.



350

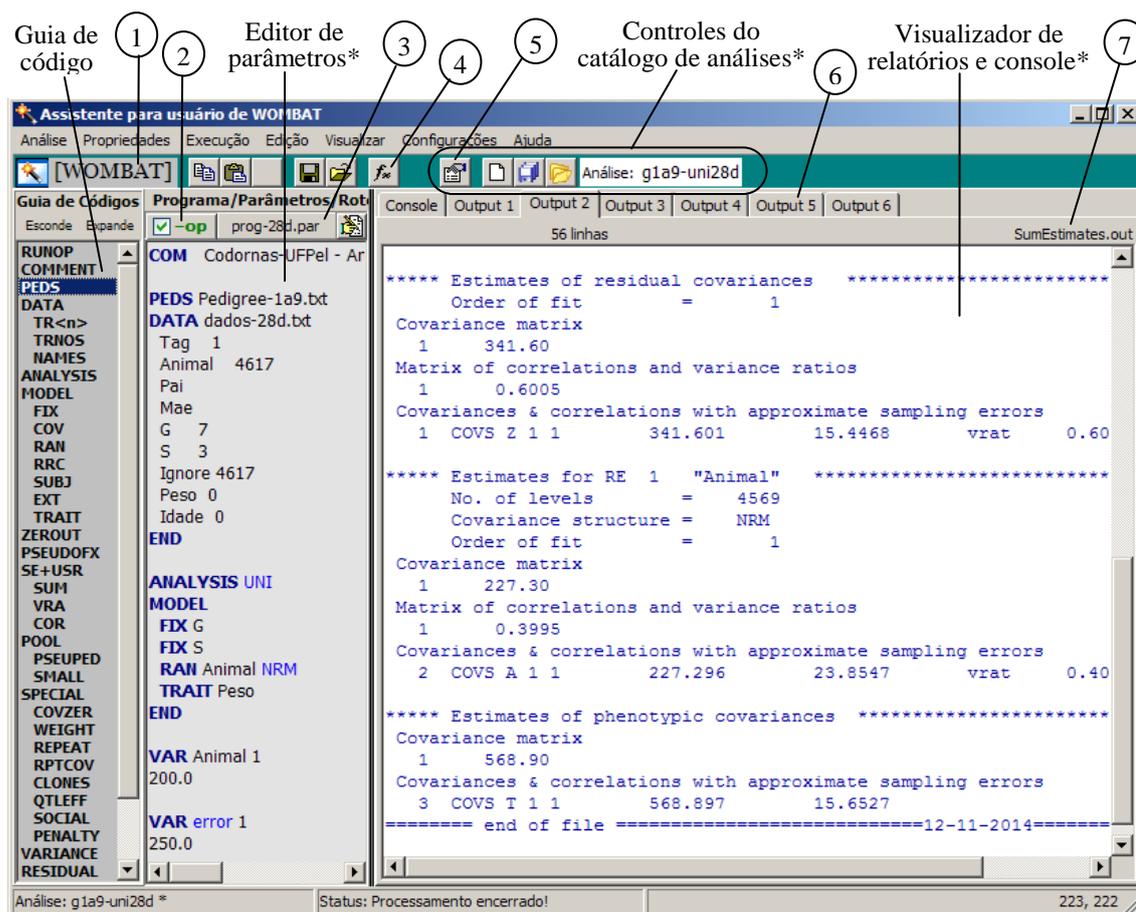
351 Figura 2 - Diagrama de relacionamento entre um Aplicativo Alvo e o Assistente de  
 352 Execução do Guaiaca e suas estruturas.

353

### 354 *Interface do usuário*

355 Toda a interação entre o Guaiaca e o usuário é realizada através de interface  
 356 modelada sobre um ambiente de janelas do sistema operacional e baseada em ações  
 357 apontar e clicar. A interface genérica do Assistente de Execução de análises é formada

358 por um conjunto de janelas que representam os seguintes elementos: o ambiente  
359 principal de operação, o descritor da análise atual, a especificação para busca de análise  
360 catalogada, a seleção de arquivos de dados e parâmetros, a configuração da conexão do  
361 Assistente com o aplicativo alvo em uso, a configuração de opções ou preferências de  
362 uso da interface e a apresentação da ferramenta. A apresentação de documentos de  
363 ajuda (manuais e orientações) é feita em janelas também ativadas pela ferramenta. Essa  
364 ação pode depender da presença de aplicativo complementar padrão, tal como leitor de  
365 PDF ou de outro paradigma adotado para o manual do aplicativo alvo em uso. A  
366 apresentação da janela principal é baseada em componentes dinamicamente ativados  
367 por escolhas definidas em preferências de uso da interface ou durante a operação. Os  
368 principais componentes assim caracterizados são: a lista de análises catalogadas; os  
369 visualizadores de arquivos de dados, de relatórios e do console de operação; o editor de  
370 arquivo de parâmetros e os guias de código e de opções de execução do aplicativo alvo.  
371 O *layout* da janela principal da interface genérica é ilustrado na Figura 3 por intermédio  
372 de imagem instanciada da ferramenta já configurada para uso em conjunto com o  
373 aplicativo WOMBAT. Nessa instância, a janela tem seus componentes carregados com  
374 os arquivos correspondentes a uma análise catalogada com o nome "g1a9-uni28d". Seus  
375 componentes principais visíveis estão identificados através de apontadores.



**Legenda:** 1 - nome do aplicativo alvo conectado; 2 - ativador do seletor de opções de execução; 3 - nome do arquivo de parâmetros; 4 - ativador de execução de análise; 5 - descritor da análise; 6 - abas indicadoras de saídas (resultados) disponíveis; 7 - nome do relatório selecionado. \* O exemplo é uma análise processada pelo aplicativo WOMBAT e catalogada sob o nome "g1a9-uni28d".

376

377 Figura 3 - Imagem da janela principal do Assistente de Execução do Guaiaca exibindo  
378 exemplo de uso e destacando diversos componentes da interface genérica.

379

380

## Aplicações e Exemplos

381 *Aplicações configuradas*

382 O Guaiaca foi configurado e testado visando estabelecer, inicialmente, dois  
383 ambientes integrados, reunindo o módulo Assistente de Execução e dois aplicativos  
384 alvos de análises estatísticas para estimação de componentes de (co)variância e  
385 consequentes parâmetros genéticos. Para esses ambientes, foram escolhidos os  
386 aplicativos WOMBAT e INTERGEN relativos, respectivamente, ao ajustamento de  
387 modelos lineares mistos com o uso da metodologia REML (Meyer, 2007) e à

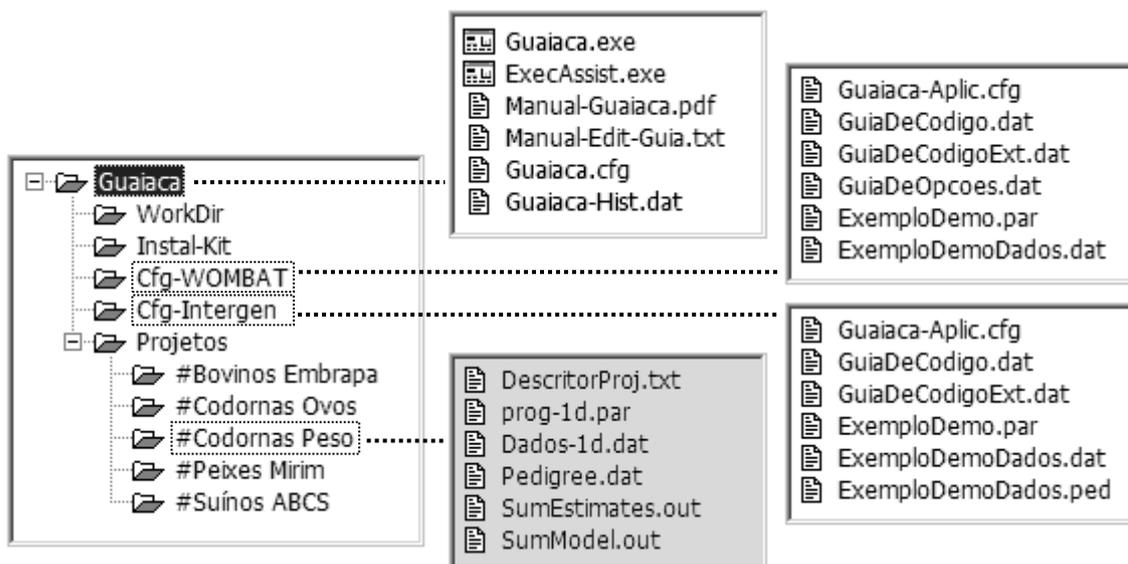
388 implementação de modelos hierárquicos bayesianos com uso de métodos de Monte  
389 Carlo via Cadeias de Markov (Cardoso, 2008).

390 Tanto os aplicativos alvos escolhidos quanto o Guaiaca foram instalados  
391 facilmente pelo simples ato de copiar seus programas e arquivos para os  
392 diretórios/pastas particulares escolhidos. As seguintes estruturas foram estabelecidas  
393 para cada aplicativo:

394 a) WOMBAT - diretório '\WOMBAT' contendo três itens - o arquivo executável  
395 'wombat.exe', o arquivo com o manual de uso 'WombatManual.pdf' e a pasta 'Examples'  
396 com o conteúdo definido de acordo com as recomendações do manual (Meyer, 2014),  
397 baixados do site '<http://didgeridoo.une.edu.au/km/wombat.php>';

398 b) INTERGEN - diretório '\INTERGEN' contendo três itens - o arquivo  
399 executável 'intergen1.2.exe', o arquivo com o manual de uso 'Manual Intergen v1-2.pdf'  
400 e a pasta 'Exemplos' com o conteúdo definido de acordo com as recomendações  
401 indicadas no manual do aplicativo (Cardoso, 2008), baixados do site  
402 '<http://www.cppsul.embrapa.br/estaticos/produtoseseservicos/english.php>';

403 c) GUAIACA - diretório '\GUAIACA' para manter a estrutura e os componentes  
404 representados na Figura 4, ou seja, os arquivos executáveis, de configuração e de  
405 documentação da ferramenta, uma pasta de trabalho (WorkDir), onde ocorre o  
406 processamento do aplicativo alvo em execução (pode ser única ou uma específica para  
407 cada aplicativo), uma pasta (opcional) para manutenção dos arquivos que compõem o  
408 kit de instalação, uma pasta de configuração para definição de cada ambiente integrado  
409 com os aplicativos alvos, uma pasta para catalogação de projetos (coleções de dados e  
410 relatórios de análises realizadas).



**Legenda:** À esquerda, a estrutura de pastas associadas ao diretório da instalação, incluindo a pasta Projetos, como local sugerido para catalogação de dados e relatórios de análises de projetos do pesquisador; ao centro e acima, os arquivos do diretório de instalação; à direita, conteúdos dos diretórios de configuração dos ambientes integrados entre o Guaiaca e os dois aplicativos alvos escolhidos; ao centro e abaixo, um exemplo de conteúdo da pasta associada a um projeto catalogado.

411

412 Figura 4 - Estrutura hierárquica da instalação completa do Guaiaca associado a dois  
413 aplicativos alvos a partir de um local escolhido.

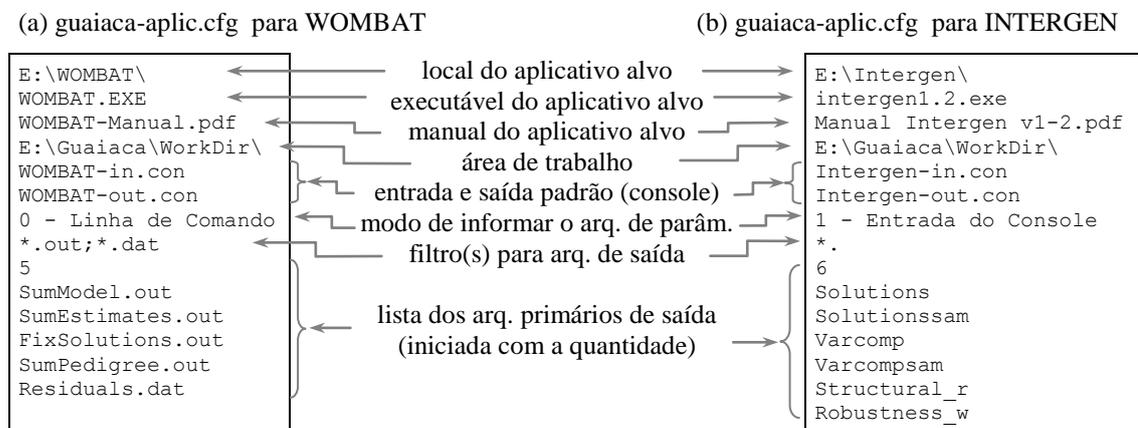
414

415 *Exemplos de configurações de aplicativos alvos*

416 A Figura 5 apresenta uma visualização parcial dos conteúdos dos arquivos de  
417 configuração utilizados para conectar a interface genérica com os programas  
418 WOMBAT e INTERGEN. Está suposto que esses programas foram instalados,  
419 respectivamente, na pasta "WOMBAT" da unidade "e:" e na pasta "INTERGEN" da unidade  
420 "e:", e que o Guaiaca foi instalado em pasta de mesmo nome também na unidade "e:" .

421 Há duas formas de estabelecer os ambientes integrados dos aplicativos alvos com  
422 o Guaiaca: ambientes isolados ou ambientes totalmente integrados. Na primeira, são  
423 constituídas duas instâncias específicas do Guaiaca com bancos de análises realizadas  
424 independentes. A segunda é baseada em uma única instância do Guaiaca, centralizando  
425 o histórico de todas as análises realizadas no único banco de dados e alternando os  
426 ambientes integrados, conforme a necessidade. Esta é a forma preferencial para os casos

427 similares ao experimentado, onde o pesquisador usa diferentes metodologias de análise,  
 428 representadas por diferentes aplicativos (WOMBAT e INTERGEN), em um mesmo  
 429 contexto de pesquisa.



(c) GuiaDeCodigo.dat (trecho) para WOMBAT, alguns com dicas

```

...
PEDS
DATA
TR<n>    DAT
TRNOS   DAT
NAMES   DAT
ANALYSIS Define o tipo de análise.||Complementar com:||UNI, MUV, RR ou MRR
MODEL
FIX      MOD
COV      MOD
RAN      MOD Nome de efeito aleat... ..de variâncias||( <vn>  [{IDE | NRM | SEX |...
RRC      MOD
SUBJ     MOD
EXT      MOD
TRAIT    MOD
ZEROUT
...
VARIANCE Matriz de covar... efeitos aleatórios||...||Obs.: VAR <efeito> <dim.>...
...

```

(d) GuiaDeOpcoes.dat (trecho) para WOMBAT, com descrição

```

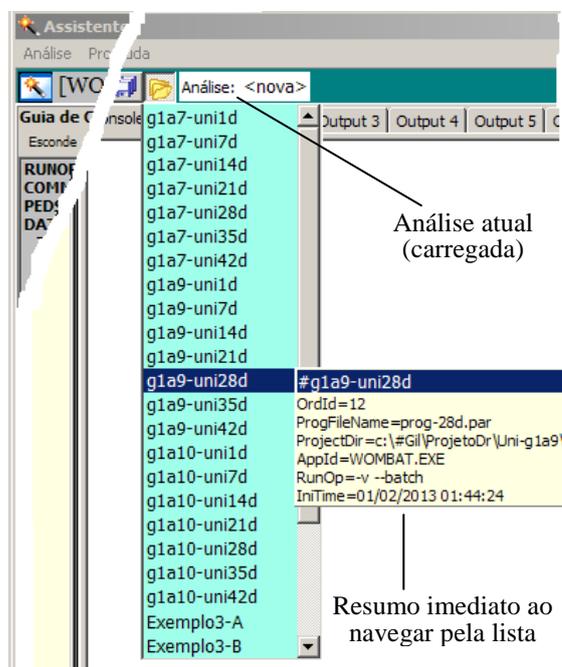
-c      0  Especifica uma rodada de continuação
-v      0  Especifica o modo detalhado para a saída em console (tela)
-d      0  Especifica saída em console muito detalhada
-t      0  Especifica o modo conciso para a saída em console
--good  0  Indica que os valores iniciais são bons (confiáveis)
--bad   0  Indica que os valores iniciais são ruins
--setup 0  Seleciona uma rodada para executar apenas os passos de configuração
--best  0  Executa apenas para imprimir as estimativas do melhor ponto atual
--blup  0  Seleciona uma rodada de predição (BLUP); solução direta
--solvit 2 Seleciona rodada de predição (BLUP); ...; n indica o exp. da precisão
--solvitgs 0 Seleciona rodada de predição (BLUP); iterativamente com Gauss-Seidel
--mmeout 0 Semelhante a --solvit, mas escreve MME em arquivo.
--snap  0  Executa uma análise do tipo GWAS para múltiplos SNPs.
--simul 2 Seleciona uma rodada só de simulação de dadas; n indica o núm. de arq...
...

```

430  
 431 Figura 5 - Exemplos de arquivos de configuração para vinculação entre a interface e  
 432 dois aplicativos alvos (a) e (b) e trechos dos arquivos de guias de código (c)  
 433 e de opções (d) da interface para um dos aplicativos vinculados.

434 *Catálogo histórico de análises*

435 Os descritores das análises realizadas são mantidos em um banco de dados que  
 436 representa o Histórico de Análises. A partir deste histórico, é possível recuperar e  
 437 acessar com facilidade as análises anteriormente planejadas e/ou realizadas e visualizar  
 438 todos os seus arquivos de entrada, saída e parâmetros, reproduzindo o estado da  
 439 interface correspondente ao ato da tarefa. Na própria lista de seleção para abertura de  
 440 uma análise catalogada é possível navegar e obter uma dica (descrição pop-up) com um  
 441 resumo imediato de cada descritor, sem precisar abrir a análise. Para cada instalação da  
 442 ferramenta, um único histórico é criado - arquivo Guaiaca-Hist.dat - independentemente  
 443 do número de aplicativos alvos vinculados (Figura 4). O aplicativo alvo utilizado para  
 444 uma análise é identificado no descritor correspondente. O acesso ao histórico, para  
 445 navegação e seleção de uma análise catalogada, é realizado via controle específico na  
 446 janela principal da interface, conforme mostrado na Figura 6.



447  
 448 Figura 6 - Janela principal (recorte) durante operação de seleção de uma análise já  
 449 realizada e catalogada no histórico.

450

451 Qualquer análise de dados, neste contexto, inicia pela criação de um descritor,  
452 mesmo que seu resultado venha a ser desprezado (não salvo). Inicialmente, o usuário  
453 fornece as informações de preparação para a análise. Na medida em que esta se  
454 desenvolve, as informações complementares são adicionadas automaticamente ao  
455 descritor. É possível também incluir comentários ou observações do analista sobre as  
456 condições e/ou resultados da análise para registro e futura discussão ou apreciação. A  
457 estrutura de dados correspondente ao descritor é formada pelos seguintes atributos:

458 a) informações fornecidas ou confirmadas pelo pesquisador:

459 - nome dado à análise (identificação única);

460 - lista com os nomes dos arquivos de dados;

461 - nome do arquivo de parâmetros ou programa;

462 - indicadores de modo de operação (arquivo de parâmetros protegido contra  
463 alteração, pasta única para dados e resultados, salvamento integral de arquivos  
464 resultantes ou apenas os relevantes);

465 - nome da pasta de projeto (onde a pasta da análise será incluída);

466 - nome da pasta de localização dos arquivos de dados (se indicador de pasta  
467 única estiver desligado);

468 - observações e comentários, podendo incluir identificação do analista;

469 b) informações geradas e atualizadas pela ferramenta:

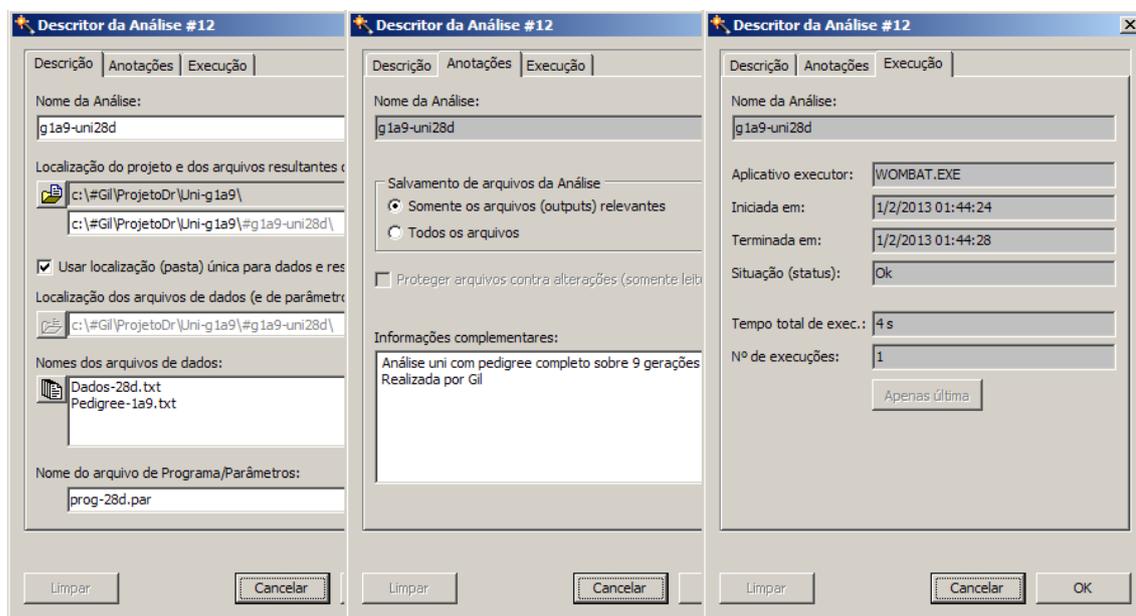
470 - número de ordem da análise no histórico;

471 - nome do aplicativo alvo;

472 - opções de execução do aplicativo alvo (opções de linha de comando);

473 - sumário de execução da análise (número de execuções cumulativas, data e  
474 hora de início e término da última execução cumulativa, tempo total acumulado, estado  
475 da última execução).

476 A visualização do descritor de uma análise na janela correspondente é ativada a  
 477 partir de controles disponíveis na janela principal da interface. Sua composição visual  
 478 (Figura 7) distribui os atributos em três grupos: descrição, anotações e execução.



479  
 480 Figura 7 - Janela do descritor da análise atual e seus três grupos de atributos (recortes  
 481 sobrepostos).

### 483 *Testes*

484 As funções de entrada de dados foram testadas de forma simulada e a campo por  
 485 meio de procedimentos operacionais correspondentes às etapas de coleta de dados dos  
 486 experimentos com codornas realizados no LEEZOO - Laboratório de Experimentação e  
 487 Ensino Dr. Renato Rodrigues Peixoto, do Departamento de Zootecnia/UFPel.

488 Um conjunto de dados de codornas referentes às avaliações de peso corporal,  
 489 constituídos por 5.726 registros compostos por 19 campos, foi utilizado para os testes  
 490 das funções de manipulação de dados, tanto para reorganização de estrutura, quanto  
 491 para reordenações diversas e visualizações seletivas (filtragens). Os dados coletados em  
 492 banco de dados, na forma convencional de registro único de informações sobre cada  
 493 indivíduo, foram reagrupados e convertidos em texto para obter, na sua integralidade ou

494 na forma de extratos, os conjuntos de dados necessários e preparados para análises  
495 através dos aplicativos WOMBAT (Meyer, 2007, 2014) e INTERGEN (Cardoso, 2008).  
496 A eficácia e a eficiência das funções de ordenação e reorganização do módulo  
497 Assistente de Dados foram também comprovadas com um banco de dados mais extenso.  
498 Para isso, foi usado um conjunto de dados genealógicos, reprodutivos e de testagem de  
499 suínos formado por mais de 650.000 registros de animais compostos por mais de 50  
500 campos de dados, gentilmente cedidos pela Associação Brasileira de Criadores de  
501 Suínos.

502 Diversas análises uni e multicaracterísticas foram realizadas com os dados de  
503 codornas e os aplicativos alvos configurados, mediante o uso da interface Assistente de  
504 Execução. Essas análises foram catalogadas e revisitadas posteriormente com o auxílio  
505 da interface para localização, visualização e reexecução com parâmetros ou dados  
506 alterados. Os testes realizados demonstraram que: (a) os resultados das análises e o  
507 tempo para a sua obtenção não sofrem interferência do processamento da interface; (b) a  
508 interface proporciona maior grau de usabilidade dos aplicativos alvos, facilitando ao  
509 usuário pesquisador a manipulação e o uso de suas estruturas de entrada e de saída de  
510 forma imediata e simultânea. Algumas das informações sobre um desses exemplos são  
511 destacadas em ilustrações neste texto (Figura 3 e Figura 7).

512

### 513 **Estágio Atual e Disponibilidade**

514 A ferramenta Guaiaca foi modelada como um "cinto de utilidades" do pesquisador  
515 para a manipulação e análise de dados. Apresenta uma interface gráfica de usuário  
516 genérica e configurável para cada aplicativo específico de análise de dados, segundo as  
517 preferências do usuário. O desenvolvimento da ferramenta continua sendo realizado. A  
518 versão 1.0 Beta deste aplicativo já está disponível de forma gratuita para a comunidade

519 científica e para desenvolvedores que desejarem experimentar ou agregar sua interface  
520 genérica com outros aplicativos científicos. Ele pode ser baixado (*download*) da página  
521 do Guaiaca na internet (<http://www.ufpel.edu.br/~gil.medeiros/guaiaca>). No estágio  
522 atual, o pacote inclui os programas, a documentação, exemplos de arquivos de  
523 configuração para uso combinado com os aplicativos alvos experimentados e o  
524 instalador do mecanismo de software que implementa o acesso ao ambiente de banco de  
525 dados. Os aplicativos alvos desejados devem ser obtidos em seus *Websites* oficiais ou  
526 outros meios previstos. A documentação incluída no pacote contém orientação para a  
527 obtenção dos aplicativos alvos experimentados em conjunto com a interface genérica.

528 O Guaiaca também pode ser útil como ferramenta complementar a aplicativos que  
529 possuam GUI própria. Neste caso, as suas funcionalidades podem ser usadas sem  
530 qualquer vinculação com um aplicativo alvo.

531 Para a instalação do software, basta descompactar o arquivo baixado e seguir as  
532 orientações que acompanham o pacote. As tarefas necessárias são: criar as pastas  
533 indicadas no ambiente de trabalho do usuário, copiar os arquivos obtidos para essas  
534 pastas e instalar o recurso motor do banco de dados.

### 535 *Programas*

536 Nesta versão inicial, exclusivamente para sistema operacional Windows, Guaiaca  
537 é liberado na forma de dois módulos executáveis produzidos para arquitetura de 32 bits.  
538 Esses módulos correspondem ao Assistente de Dados e ao Assistente de Execução, que  
539 poderão constituir ambientes de operação distintos, se desejável, ou um único ambiente  
540 integrado.

541 A ativação do aplicativo alvo pela interface é realizada através da criação de um  
542 processo separado do sistema operacional.

543 A execução da maioria das funções do Assistente de Dados é dependente da  
544 instalação do mecanismo de acesso ao ambiente de banco de dados.

#### 545 *Documentação*

546 Dois tipos de documentos são disponíveis para a orientação do usuário: Manual  
547 do Uso do Guaiaca, em PDF, e documentos de orientação para personalização da  
548 interface genérica. A personalização refere-se à vinculação entre a interface e o  
549 aplicativo alvo e à preparação dos Guias de Código e de Opções. Estes documentos são  
550 acessíveis diretamente na interface gráfica do Guaiaca a partir do menu de Ajuda.

551 O manual do aplicativo alvo, se disponível, também pode ser acessado a partir do  
552 menu de Ajuda do Guaiaca, desde que isto seja informado na configuração da conexão  
553 do Assistente de Execução com o aplicativo alvo. Também é possível selecionar e  
554 conectar na interface gráfica algum exemplo consolidado de uso do aplicativo alvo. Isto  
555 facilita a demonstração exploratória dos recursos da interface e suas vantagens para  
556 visualizar, compreender e explorar os métodos, técnicas e modelos implementados pelo  
557 aplicativo alvo.

#### 558 *Funcionalidades a serem incorporadas no futuro*

559 Pretende-se dar continuidade ao desenvolvimento do Guaiaca para consolidar os  
560 componentes funcionais já definidos e incluir diversas ferramentas ou funcionalidades  
561 complementares. Dentre elas, destaca-se: captura automática de dados a partir de  
562 dispositivos eletrônicos (balanças, por exemplo); transformações sobre dados, gerando  
563 novas estruturas e/ou variáveis; instalação para idiomas alternativos; codificação  
564 alternativa e equivalente para comandos/códigos (sinônimos com abreviações ou outros  
565 idiomas); importação e tratamento implícito de outros formatos de armazenamento de  
566 dados (atualmente só o formato de banco de dados é tratado diretamente e a conversão  
567 de/para texto é feita de forma explícita via função de Manipulações/Reorganizar);

568 geração de gráficos; ampliação da interface genérica com assistentes específicos para  
 569 aplicativos alvos, visando a geração automática de arquivos de parâmetros (programas);  
 570 inclusão de novas informações complementares no descritor de análise; detalhamento  
 571 de critérios para busca de análises catalogadas; funções auxiliares de uso comum em  
 572 análise de dados relativos ao melhoramento genético de animais, tais como o cálculo de  
 573 coeficientes de endogamia em populações alvos de seleção.

574

### 575 **Agradecimentos**

576 Ao CNPq pela concessão de financiamento via Edital Universal, permitindo a  
 577 manutenção da linhagem de codornas de corte, fonte primária de dados que foi  
 578 determinante da motivação para esse projeto.

579

### 580 **Referências**

- 581 BOLDMAN, K.G.; KRIESE, L.A.; Van VLECK, L.D. et al. **A manual for use of**  
 582 **MTDFREML. A set of programs to obtain estimates of variance and**  
 583 **covariance.** Lincoln: USDA/ARS, 1993. 120p.
- 584 CARDOSO, F.F. **Manual de utilização do programa INTERGEN – Versão 1.0 em**  
 585 **estudos de genética quantitativa animal.** Bagé: Embrapa Pecuária Sul, 2008. 74p.  
 586 (Documentos, 74) ISSN 1982-5390
- 587 CRUZ, C.D. GENES - a software package for analysis in experimental statistics and  
 588 quantitative genetics. **Acta Scientiarum. Agronomy**, v. 35, n. 3, p.271-276, 2013.
- 589 DELWICHE, L.D.; SLAUGHTER, S.J. **The Little SAS Book: A Primer.** 5a. ed. Cary,  
 590 NC: SAS Institute Inc, 2012.
- 591 DIONELLO, N.J.L.; CORREA, G.S.S.; SILVA, M.A. et al. Estimativas da trajetória  
 592 genética do crescimento de codornas de corte utilizando modelos de regressão  
 593 aleatória. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.60, n.2, p.  
 594 454-460, 2008.
- 595 FERREIRA, D.F. **Estatística Computacional em Java.** 1. ed. Lavras: Ed. UFLA,  
 596 2013. v. 1. 695p.
- 597 GARRICK, D.J.; SAATCHI, M. Opportunities and challenges for genomic selection of  
 598 beef cattle. **Revista Brasileira de Zootecnia**, v.40, p.310-316, 2011. ISSN 1806-  
 599 9290.
- 600 GILMOUR, A.R.; GOGEL, B.J.; CULLIS, B.R.; THOMPSON, R. **ASReml User**  
 601 **Guide Release 3.0.** VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, 2009.  
 602 Disponível em: <<http://www.vsnl.co.uk>>. Acesso em: 24 nov. 2014.

- 603 LEGARRA, A.; AGUILAR, I.; MISZTAL, I. A relationship matrix including full  
604 pedigree and genomic information. **Journal of Dairy Science**, v.92, n.9, p.4656-  
605 4663, 2009.
- 606 MADAN, A.; DUBEY, S.K. Usability evaluation methods: a literature review.  
607 **International Journal of Engineering Science and Technology**, v.4, n.2, p.590-  
608 599, 2012.
- 609 MEYER, K. DFREML - A set programs to estimate variance components under an  
610 individual animal model. **Journal of Dairy Science**, v.71, suppl. 2, p.33-34, 1988.
- 611 MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by  
612 restricted maximum likelihood (REML). **Journal of Zhejiang University**  
613 **SCIENCE B**, v.8, n.11, p.815–821, 2007.
- 614 MEYER, K. **WOMBAT - a program for mixed model analyses by restricted**  
615 **maximum likelihood**. User Notes. Armidale, Australia: Animal Genetics and  
616 Breeding Unit, University of New England, 2006-2014. 107p. Disponível em:  
617 <<http://didgeridoo.une.edu.au/km/download.php?file=WombatManual.pdf>>.  
618 Acesso em: 18 mar. 2015.
- 619 MISZTAL, I; TSURUTA, S.; LOURENÇO, D. et al. **Manual for BLUPF90 family of**  
620 **programs**. Athens, USA: University of Georgia, 2014. Disponível em:  
621 <<http://nce.ads.uga.edu/wiki/doku.php>>. Acesso em: 24 nov. 2014.
- 622 RESENDE, M.D.V. **Selegen-Reml/Blup : Sistema Estatístico e Seleção Genética**  
623 **Computadorizada via Modelos Lineares Mistos**. 1. ed. Colombo: Embrapa  
624 Florestas, 2007. 360p .
- 625 SAS INSTITUTE INC. **SAS 9.3 Companion for Windows**. Cary, NC: SAS Institute  
626 Inc, 2010.
- 627

### **3 Artigo 2: Comparação de enfoque bayesiano em relação à metodologia clássica para estudo de evolução e tendências dos parâmetros genéticos de codornas de corte ao longo de várias gerações**

Este artigo está apresentado conforme as normas do periódico a que será submetido. As páginas seguintes deste capítulo (45 a 61) equivalem às páginas do artigo, numeradas junto à margem inferior, independentemente deste contexto. A numeração contínua convencional, utilizada para os índices, está indicada em baixa intensidade junto à margem superior.

1 **Comparação de enfoque bayesiano em relação à metodologia clássica para estudo de**  
2 **evolução e tendências dos parâmetros genéticos de codornas de corte ao longo de várias**  
3 **gerações<sup>1</sup>**

4 **Bayesian approach versus classical methodology: comparative study on evolution and**  
5 **trends of genetic parameters of meat quails over several generations**

6 **Gil Carlos Rodrigues Medeiros<sup>2</sup>, Willian Silva Barros<sup>3</sup>, Nelson José Laurino Dionello<sup>4</sup>**

7  
8 **RESUMO**

9 O objetivo deste trabalho foi obter e comparar estimativas de componentes de  
10 variância e herdabilidades usando duas abordagens - clássica e bayesiana - através de análises  
11 unicaracterísticas de dados de peso corporal de 5726 codornas de corte. Os dados referem-se a  
12 pesagens individuais dos animais, realizadas semanalmente do nascimento aos 42 dias,  
13 registradas juntamente com as informações de *pedigree* para 10 gerações em processo de  
14 seleção pelo peso aos 42 dias. Efeitos fixos de geração e sexo foram incluídos em um modelo  
15 animal simples junto aos efeitos aleatórios genético aditivo e residual. Foi analisado o  
16 comportamento dos parâmetros genéticos das populações cumulativas da 7<sup>a</sup> à 10<sup>a</sup> gerações.  
17 Para a abordagem clássica (Máxima Verossimilhança Restrita) foi usado o software  
18 WOMBAT e para a bayesiana (com Métodos de Monte Carlo via Cadeias de Markov e  
19 amostragem de Gibbs) foi adotado o software INTERGEN. A nova ferramenta GUAIACA foi  
20 usada para integrar esses recursos a uma interface amigável e a recursos de manipulação de  
21 dados e análises. As duas abordagens produzem resultados extremamente similares. As

---

<sup>1</sup> Projeto financiado pelo CNPq - processo nº 484626/2007-8.

<sup>2</sup> Doutorando do Programa de Pós-Graduação em Zootecnia - UFPel, Brasil. Docente do Centro de Engenharias - UFPel. e-mail: gil.medeiros@ufpel.edu.br.

<sup>3</sup> Docente do Departamento de Matemática e Estatística - UFPel, Brasil. e-mail: willian.barros@ufpel.edu.br.

<sup>4</sup> Docente do Departamento de Zootecnia - FAEM - UFPel, Brasil. Pesquisador do CNPq. e-mail: dionello@ufpel.edu.br.

22 estimativas de herdabilidades com o grupo completo das 10 gerações, para cada idade, são  
23 respectivamente 0,49, 0,31, 0,28, 0,33, 0,38, 0,46 e 0,39. Estes valores apresentam uma leve  
24 redução entre a 7<sup>a</sup> e a 10<sup>a</sup> gerações, para todas as idades. Fica evidenciado que, após os  
25 primeiros dias de vida, a maior herdabilidade para peso corporal das codornas de corte ocorre  
26 em idade próxima aos 35 dias. A abordagem clássica, mais simples e rápida, é suficiente para  
27 a obtenção de boas estimativas dos parâmetros, nas condições testadas.

28 **Palavras-chave:** herdabilidade, componentes de variância, modelos mistos, amostragem de  
29 Gibbs, REML.

### 30 **ABSTRACT**

31 The aim of the present study was to acquire and to compare the estimates of variance  
32 components and heritability employing two approaches - classic and bayesian - through  
33 univariate analyses for body weight of 5726 meat quails. The data refer to individual weights  
34 accomplished weekly from birth to 42 days, that were recorded with the pedigree info for 10  
35 generations of selection process by the weight at 42 days. Fixed effects for generation and sex  
36 were included into a simple animal model together with the genetic and residual random  
37 effects. It was analyzed the behavior of genetic parameters into cumulative populations from  
38 7<sup>th</sup> to 10<sup>th</sup> generations. For the classic approach (Restricted Maximum Likelihood) it was used  
39 the WOMBAT software and for the bayesian approach (with Markov Chain Monte Carlo  
40 method and Gibbs sampling) it was chosen the INTERGEN software. The new tool  
41 GUAICA was accessed with the aim of combine those applications to a friendly user  
42 interface and to data manipulation features and to data analysis organization feature. Both  
43 approaches produce very similar results. Heritability estimates of complete set with 10  
44 generations, for each age, are respectively 0.49, 0.31, 0.28, 0.33, 0.38, 0.46 and 0.39. These  
45 values have a slight reduction between the 7<sup>th</sup> and 10<sup>th</sup> generation, for each age. It is  
46 demonstrated that after the first few days of life, the best heritability for body weight of quails

47 occurs near to 35 days old. The classic approach, easiest and fastest, is efficient obtaining  
48 good estimates of the parameters, under the conditions tested.

49 **Key words:** heritability, variance components, mixed models, Gibbs sampler, REML.

50

## 51 **INTRODUÇÃO**

52 Os parâmetros genéticos e a herdabilidade dos pesos corporais de animais são  
53 estudados e estimados há vários anos com o uso de metodologias tradicionais, tais como o uso  
54 de modelo animal e a aplicação de modelos de regressão aleatória (MRA) (DIONELLO et al.,  
55 2008). O método de máxima verossimilhança restrita (REML) tem sido largamente adotado  
56 nos processos de estimação relacionados ao melhoramento genético de animais (FALCÃO,  
57 2009), para os quais a metodologia de modelos mistos assume fundamental importância por  
58 permitir a análise conjunta de efeitos fixos e aleatórios (HENDERSON, 1984). As aplicações  
59 dos métodos REML na análise de modelos de regressão aleatória para a seleção de animais  
60 tem seguido esta tendência, em especial, no melhoramento de codornas de corte (DIONELLO  
61 et al., 2008; 2009; BONAFÉ et al., 2008, 2011).

62 O uso de modelos mistos, para a obtenção de estimadores e preditores, é  
63 fundamentado na execução de processos iterativos que se desenvolvem a partir de estimativas  
64 iniciais para os correspondentes parâmetros. O software WOMBAT (MEYER, 2007) tem sido  
65 bastante utilizado com este propósito e é disponível gratuitamente. Embora as vantagens do  
66 método REML - gerar estimativas não viesadas (BLUE / BLUP) e viabilizar a utilização de  
67 toda a informação disponível - a obtenção de componentes de (co)variância para estimação de  
68 herdabilidades, pela abordagem clássica de modelos mistos, apresenta algumas limitações,  
69 como a dificuldade de se estabelecer intervalo de confiança adequado para esse parâmetro.

70 Alternativamente, uma das mais recentes tendências em termos de metodologias para  
71 a avaliação genética de animais é o uso do paradigma bayesiano. Sua concepção viabiliza a

72 valorização de conhecimento anterior sobre os parâmetros, via distribuições *a priori*, e a  
73 obtenção de estimativas pontuais e intervalos de credibilidade para as distribuições *a*  
74 *posteriori* dos parâmetros, independentemente de aproximações ou de normalidade  
75 pressuposta (SORENSEN & GIANOLA, 2002), o que representa uma vantagem para estes  
76 métodos (FALCÃO et al., 2009). A inferência bayesiana associada ao uso de técnicas de  
77 simulação relativamente simples, mas extremamente poderosas, como os métodos de Monte  
78 Carlo via Cadeias de Markov (MCMC) usando a amostragem de Gibbs, tornou-se altamente  
79 estimulada com a forte queda das limitações inerentes a sua aplicabilidade, em decorrência  
80 dos expressivos avanços das capacidades computacionais. HIDALGO et al. (2011) utilizaram  
81 a inferência bayesiana para estimar parâmetros genéticos relacionados à produção de ovos,  
82 considerando quatro linhagens de codornas. WINTER et al. (2006) usaram o mesmo  
83 paradigma na estimação de correlações genéticas e fenotípicas para características de  
84 desempenho ponderal de codornas de corte. Dentre os recursos de software que implementam  
85 esta metodologia, destaca-se o programa INTERGEN, desenvolvido e distribuído  
86 gratuitamente pela Embrapa Pecuária Sul (CARDOSO, 2008).

87 Estes dois aplicativos, WOMBAT e INTERGEN, implementam os métodos  
88 correspondentes e estabelecem as interações com o usuário na forma de uma janela de texto  
89 via linhas de comando. Entretanto, a utilização destes recursos pode ser combinada com a  
90 ferramenta GUAIACA, criada especialmente para suprir uma interface gráfica amigável e  
91 criar um ambiente integrado que facilita a manipulação de dados e a organização de múltiplas  
92 análises, bem como as ações comparativas de diferentes metodologias.

93 O objetivo deste estudo é obter e comparar as estimativas de componentes de  
94 variâncias e herdabilidades para codornas de uma linhagem de corte, selecionadas pelo peso  
95 aos 42 dias, após dez gerações de seleção, através das metodologias REML e inferência  
96 bayesiana, preliminarmente avaliadas por análises unicaracterísticas.

97

98 **MATERIAL E MÉTODOS**

99           Foram analisados os dados de peso corporal de codornas de uma população global de

100 10 gerações sucessivas de uma linhagem de corte em processo de seleção no Departamento de

101 Zootecnia (DZ) da UFPel (DIONELLO et al., 2009). As pesagens individuais das 5726 aves

102 avaliadas foram realizadas para sete idades, semanalmente, do nascimento aos 42 dias, as

103 quais são referenciadas neste contexto por PC1, PC7, PC14, PC21, PC28, PC35 e PC42. Estes

104 dados foram reunidos aos correspondentes registros de *pedigree*, mantidos desde a primeira

105 geração, originada de uma população base constituída por 273 genótipos.

106           A comparação dos parâmetros genéticos foi realizada com quatro grupos

107 correspondentes às populações acumuladas da 1<sup>a</sup> a 7<sup>a</sup>, 1<sup>a</sup> a 8<sup>a</sup>, 1<sup>a</sup> a 9<sup>a</sup> e 1<sup>a</sup> a 10<sup>a</sup> gerações,

108 utilizando-se as estimativas produzidas pelas duas abordagens - REML e bayesiana - para os

109 componentes de variância e herdabilidade direta (no sentido restrito). O primeiro grupo é

110 composto por 4617 animais (DIONELLO et al., 2012). Os demais são compostos pela reunião

111 destes cumulativamente com as populações das gerações seguintes (8<sup>a</sup>, 9<sup>a</sup> e 10<sup>a</sup>),

112 respectivamente, compostas por 335, 410 e 364 animais. Os genitores de cada uma dessas

113 gerações, foram selecionados na geração imediatamente anterior, pelo peso aos 42 dias, sendo

114 62 machos e 93 fêmeas, para a 8<sup>a</sup>, e 54 machos e 88 fêmeas, para a 9<sup>a</sup> e para a 10<sup>a</sup>, mantendo-

115 se o coeficiente de endogamia (F) sob controle (60% dos animais não endogâmicos e

116 coeficiente médio dos animais endogâmicos inferior a 3,2%, até a 10<sup>a</sup> geração). Os objetos

117 comparados foram os resultados das análises unicaracterísticas processadas para peso corporal

118 em cada uma das sete idades estudadas. As análises foram concebidas considerando a

119 aplicação de um modelo animal simples, que incluiu os efeitos fixos de geração e sexo e os

120 efeitos aleatórios genéticos aditivos e residual, cuja expressão matricial é dada por

121           
$$y = X\beta + Za + e ,$$

122 onde  $\mathbf{y}$  representa o vetor de pesos corporais na idade considerada;  $\boldsymbol{\beta}$  representa o vetor de  
 123 efeitos fixos;  $\mathbf{a}$  e  $\mathbf{e}$  representam, respectivamente, os vetores de efeitos genéticos diretos e  
 124 de resíduos (aleatórios);  $\mathbf{X}$  e  $\mathbf{Z}$  são as matrizes de incidência dos efeitos fixos e genéticos  
 125 diretos, respectivamente.

126 Foi admitido que  $\mathbf{y}$ ,  $\mathbf{a}$  e  $\mathbf{e}$  tenham distribuição conjunta normal multivariada, na forma

$$127 \begin{bmatrix} \mathbf{y} \\ \mathbf{a} \\ \mathbf{e} \end{bmatrix} \sim NMV \left\{ \begin{bmatrix} \mathbf{X}\boldsymbol{\beta} \\ \boldsymbol{\phi} \\ \boldsymbol{\phi} \end{bmatrix}, \begin{bmatrix} \mathbf{ZGZ}' + \mathbf{R} & \mathbf{ZG} & \mathbf{R} \\ \mathbf{GZ}' & \mathbf{G} & \boldsymbol{\phi} \\ \mathbf{R} & \boldsymbol{\phi} & \mathbf{R} \end{bmatrix} \right\}$$

128 sendo,  $\mathbf{G} = \mathbf{A}\boldsymbol{\sigma}_a^2$  e  $\mathbf{R} = \mathbf{I}_n\boldsymbol{\sigma}_e^2$ , onde  $\mathbf{A}$  é a matriz de coeficientes de parentesco,  $\mathbf{I}_n$  é uma  
 129 matriz identidade,  $\boldsymbol{\sigma}_a^2$  é a variância genética aditiva direta e  $\boldsymbol{\sigma}_e^2$  é a variância residual.

130 A distribuição normal dos dados foi assumida, tanto para a aplicação do método  
 131 REML, como para a inferência bayesiana. Na análise bayesiana, como densidades *a priori*,  
 132 foram consideradas as pressuposições convencionais de informação vaga sobre os efeitos  
 133 fixos,  $f(\boldsymbol{\beta}) \propto$  constante, de normalidade para a distribuição dos efeitos aleatórios, sendo  
 134  $\mathbf{a} | \boldsymbol{\sigma}_a^2, \mathbf{A} \sim N(0, \mathbf{A}\boldsymbol{\sigma}_a^2)$  e  $\mathbf{e} | \boldsymbol{\sigma}_e^2 \sim N(0, \mathbf{I}_n\boldsymbol{\sigma}_e^2)$ , de normalidade e independência dos erros,  
 135 sendo  $\mathbf{y} | \boldsymbol{\beta}, \mathbf{a}, \boldsymbol{\sigma}_e^2 \sim NMV(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}, \mathbf{I}_n\boldsymbol{\sigma}_e^2)$ , e de distribuição qui-quadrado invertida para as  
 136 variâncias genética aditiva e residual.

137 As referências tecnológicas adotadas nas análises de dados e estimação dos  
 138 componentes de variância foram o software WOMBAT (MEYER, 2007; 2014), para a  
 139 aplicação da metodologia REML, e o software INTERGEN (CARDOSO, 2008), para a  
 140 inferência bayesiana com MCMC e amostragem de Gibbs. A preparação, a organização e a  
 141 visualização dos dados, parâmetros e resultados das análises foram realizadas com o apoio  
 142 dos recursos do Assistente de Dados e do Assistente de Execução (interface genérica) do  
 143 software GUAIACA, desenvolvido junto ao DZ/UFPel. A herdabilidade direta ( $h_d^2$ ) foi obtida

144 da razão de variâncias do WOMBAT (sob REML) e por cálculo via planilha eletrônica com  
145 os componentes de variância do INTERGEN (sob a abordagem bayesiana).

146 Para a amostragem de Gibbs (no INTERGEN), visando a inferência bayesiana, foram  
147 adotadas cadeias geradas com base nos seguintes argumentos: 11000 ciclos (total), 1000  
148 ciclos iniciais descartados (aquecimento) e 10 ciclos de intervalo de amostragem (uso de uma  
149 amostra a cada 10 geradas). Foi considerado que este processo, gerando 1000 amostras dos  
150 componentes de variância para cada avaliação, representava um bom padrão para todas as  
151 comparações e atendia as necessidades, visto que boas estimativas seriam conhecidas *a priori*  
152 através da análise pela metodologia REML previamente executada. Entretanto, foram  
153 realizados diversos testes com variações destes argumentos (número de ciclos) visando  
154 confirmar a adequação do dimensionamento escolhido. Para esses testes foi utilizado o  
155 conjunto de dados da característica de menor cardinalidade - PC42 - referente às 10 gerações.

156 Para estabelecer uma comparação dos tempos de execução, foram adotados os  
157 seguintes critérios: todas as análises executadas no mesmo sistema computacional, mantendo  
158 carga de tarefas do sistema operacional de forma similar durante todos os eventos de  
159 execução das análises; todas as análises ativadas sob controle do software GUAIACA, com  
160 registros de tempo de início e fim de processamento; medidas de tempo tomadas com a  
161 precisão de segundos (usando a aproximação convencional). O sistema computacional  
162 adotado foi um computador pessoal com processador Intel Core i3, memória de 4 GB e  
163 sistema operacional Windows 7, utilizando apenas um dos quatro núcleos do processador para  
164 a execução de cada análise.

165

## 166 **RESULTADOS E DISCUSSÃO**

167 Os testes realizados para a escolha do dimensionamento da amostragem de Gibbs  
168 (Tabela 1) mostraram que a herdabilidade permanece praticamente inalterada, quando se

169 aplica diferentes valores para o número de ciclos, mantendo-se o período de aquecimento e o  
 170 intervalo de aproveitamento de amostras, respectivamente em 1000 e 10 ciclos. Mesmo ao  
 171 aplicar alterações mais expressivas, mudando-se os três argumentos, com maior número de  
 172 ciclos e redução do número de amostras utilizadas, resultados muito semelhantes são obtidos.  
 173 Desta forma, considera-se razoável a adoção dos argumentos escolhidos previamente para  
 174 todas as análises, mesmo que a última alternativa testada apresente resultados mais próximos  
 175 dos determinados pela abordagem clássica (REML), realizada previamente para estes testes.  
 176 A expressiva diferença de custo (tempo computacional) envolvido na amostragem  
 177 hiperdimensionada, da ordem de 35 minutos, comparado com os 8 minutos necessários para  
 178 uma amostragem nos termos adotados, é um fator que empurra aquela abordagem para os  
 179 casos em que uma precisão mais elevada seja fundamental.

180 Tabela 1 - Resultados (estimativas dos parâmetros genéticos) dos testes de dimensionamento  
 181 da amostragem de Gibbs para as cadeias MCMC via INTERGEN com os dados  
 182 ponderais dos animais de 10 gerações para a característica PC42

	REML*			Bayesiana		
Núm. de registros (PC42)				4499		
Núm. de genótipos				4727		
Núm. total de ciclos	-	6000	8500	11000	21000	50000
Período de aquecimento	-			1000		20000
Intervalo de amostragem	-			10		100
$\sigma_a^2$	430,22	443,67	442,64	439,68	444,15	431,56
$\sigma_e^2$	687,08	677,37	677,82	679,29	677,12	684,94
$h_d^2$	0,385	0,396	0,395	0,393	0,396	0,387
Tempo de processamento	2 s	4 min 7 s	5 min 47 s	7 min 56 s	14 min 19 s	34 min 18 s

183 \* Os resultados obtidos por REML (via WOMBAT) são apresentados como referência para comparação.

184

185 Tomando como referência o grupo completo formado pelas 10 gerações, são  
 186 apresentadas na Tabela 2 as estatísticas descritivas relativas aos resultados da análise  
 187 bayesiana. Aqui comprova-se o bom dimensionamento do número de amostras para todas as  
 188 características avaliadas, visto que: as variâncias das amostras ( $\sigma_k^2$ ) das herdabilidades  
 189 estimadas são muito baixas e suas medianas e médias são iguais; as medianas e médias das  
 190 amostras dos componentes de variância são quase iguais. Os correspondentes coeficientes de

191 variação (CV) são considerados baixos, segundo a classificação de PIMENTEL GOMES  
 192 (1990). Para  $\sigma_e^2$ , o CV varia de 3% a 5% e, para  $\sigma_a^2$ , varia de 5% a 13%, o que também  
 193 confirma a maior facilidade de estimar o componente residual em relação aos outros  
 194 componentes, que é considerada como característica comum do método bayesiano (WANG et  
 195 al., 1993; FALCÃO et al., 2009). Outro indicador desta adequação é o intervalo de  
 196 credibilidade da média, determinado pelo método tradicional a 95%, cuja amplitude  
 197 apresenta-se bastante estreita para todos os parâmetros estimados.

198

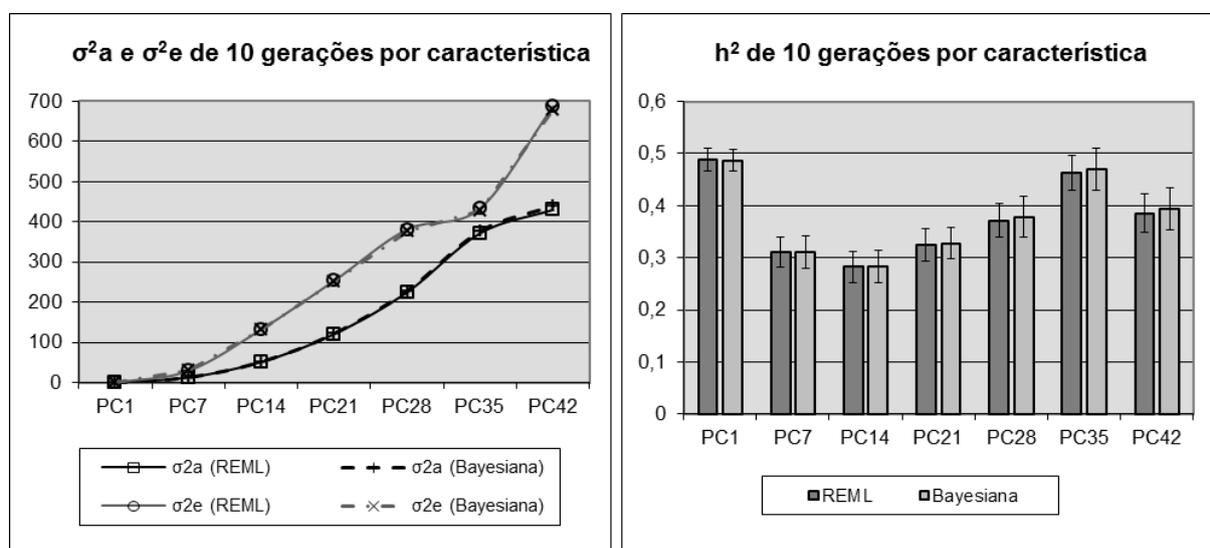
199 Tabela 2 - Estatísticas descritivas dos componentes de variância e de herdabilidades obtidos  
 200 por análise bayesiana através de MCMC para pesos corporais de 1 a 42 dias em 10  
 201 gerações de codornas

Caracte- rística	Parâ- metro*	Média	Mediana	$\sigma_k^2$ *	Desvio Padrão	CV (%)	Mínimo	Máximo	IC* 95%
PC1	$\sigma_a^2$	0,53	0,53	0,0007	0,03	5	0,45	0,62	(0,53 - 0,53)
	$\sigma_e^2$	0,56	0,56	0,0003	0,02	3	0,51	0,62	(0,56 - 0,56)
	$h^2_d$	0,48	0,48	0,0003	0,02	4	0,43	0,55	(0,483 - 0,485)
PC7	$\sigma_a^2$	13,11	13,06	1,8439	1,36	10	9,12	18,38	(13,03 - 13,20)
	$\sigma_e^2$	29,28	29,28	0,9712	0,99	3	26,34	32,93	(29,21 - 29,34)
	$h^2_d$	0,31	0,31	0,0008	0,03	9	0,22	0,40	(0,307 - 0,311)
PC14	$\sigma_a^2$	51,97	51,78	40,2286	6,34	12	36,62	76,36	(51,58 - 52,36)
	$\sigma_e^2$	132,33	132,42	21,7492	4,66	4	118,11	146,92	(132,05 - 132,62)
	$h^2_d$	0,28	0,28	0,0009	0,03	11	0,20	0,38	(0,280 - 0,283)
PC21	$\sigma_a^2$	122,18	121,57	211,9820	14,56	12	84,86	169,79	(121,27 - 123,08)
	$\sigma_e^2$	252,49	252,62	99,7981	9,99	4	221,66	284,92	(251,87 - 253,11)
	$h^2_d$	0,33	0,33	0,0011	0,03	10	0,24	0,42	(0,324 - 0,328)
PC28	$\sigma_a^2$	228,43	227,57	697,6571	26,41	12	161,61	310,62	(226,80 - 230,07)
	$\sigma_e^2$	377,95	377,76	286,0394	16,91	4	328,44	430,06	(376,91 - 379,00)
	$h^2_d$	0,38	0,38	0,0013	0,04	10	0,28	0,47	(0,374 - 0,378)
PC35	$\sigma_a^2$	376,58	376,65	1526,1524	39,07	10	279,39	493,12	(374,16 - 379,00)
	$\sigma_e^2$	428,80	428,69	534,1045	23,11	5	360,13	498,83	(427,37 - 430,23)
	$h^2_d$	0,47	0,47	0,0014	0,04	8	0,36	0,57	(0,464 - 0,469)
PC42	$\sigma_a^2$	439,68	438,24	3346,5762	57,85	13	295,30	626,78	(436,10 - 443,27)
	$\sigma_e^2$	679,29	680,02	1300,7708	36,07	5	549,85	780,70	(677,05 - 681,52)
	$h^2_d$	0,39	0,39	0,0018	0,04	11	0,27	0,53	(0,389 - 0,395)

202 \*  $\sigma_k^2$ : variância das amostras do parâmetro k, onde k é um dos parâmetros  $\sigma_a^2$  (variância genética aditiva direta),  
 203  $\sigma_e^2$  (variância residual) ou  $h^2_d$  (herdabilidade direta); CV: coeficiente de variação; IC: intervalo de credibilidade.  
 204

205 Os componentes de variância e as correspondentes herdabilidades obtidos para as 56  
 206 análises realizadas estão apresentados na Tabela 3. Para cada uma das 28 análises com  
 207 correspondência nas duas abordagens - clássica (via WOMBAT) e bayesiana (via

208 INTERGEN) - são destacadas a variância genética aditiva ( $\sigma_a^2$ ), a variância residual ( $\sigma_e^2$ ), a  
 209 variância fenotípica ( $\sigma_p^2$ ) e a herdabilidade direta ( $h_d^2$ ) e são indicadas as quantidades de  
 210 dados disponíveis para cada idade de pesagem. Observe-se que as informações de *pedigree*  
 211 são disponíveis para quase a totalidade dos animais das populações consideradas. Pode ser  
 212 observado, pelas correspondências entre as colunas REML e bayesiana (Tabela 3), que as  
 213 diferenças entre os resultados são desprezíveis, indicando uma efetiva equivalência entre as  
 214 duas metodologias de obtenção dos parâmetros genéticos para aplicações similares. Esta  
 215 equivalência pode ser observada nos gráficos correspondentes às estimativas dos parâmetros  
 216 obtidas para as 10 gerações (Figura 1), onde ocorre uma sobreposição quase perfeita das  
 217 representações das duas abordagens. O erro de amostragem aproximado determinado por  
 218 REML e representado para a herdabilidade na figura (quadro à direita) também é equivalente  
 219 ao desvio padrão do parâmetro obtido pela abordagem bayesiana, correspondentemente.



220  
 221 Figura 1 - Gráficos comparativos dos componentes de variância e das herdabilidades,  
 222 demonstrando similaridade plena entre resultados das duas abordagens.  
 223

224 O mesmo perfil de relacionamento entre as abordagens repete-se para os demais  
 225 grupos analisados. Observa-se também um elevado grau de persistência das variâncias em  
 226 todos os grupos analisados, o que indica uma estabilidade nestes parâmetros mesmo com a

227 inclusão de novas gerações. As estimativas das herdabilidades apresentam uma leve redução  
 228 entre os grupos formados pelas 7 gerações iniciais e pelas 10 gerações, para todas as idades.

229 Tabela 3 - Componentes de variância e herdabilidades para os pesos corporais de codornas  
 230 em 7 idades nas populações crescentes dos 4 grupos cumulativos de gerações  
 231 obtidos pelos dois métodos comparados (REML e bayesiano)

	NA/ NR*	REML				Bayesiana			
		$\sigma_a^2$	$\sigma_e^2$	$\sigma_p^2$	$h^2_d$	$\sigma_a^2$	$\sigma_e^2$	$\sigma_p^2$	$h^2_d$
	4617	7 gerações							
PC1	4586	0,54	0,54	1,08	0,500	0,54	0,54	1,08	0,500
PC7	3972	13,49	27,10	40,59	0,332	13,52	27,13	40,65	0,333
PC14	3823	54,68	121,81	176,49	0,310	54,79	121,98	176,77	0,310
PC21	3729	117,42	227,32	344,74	0,341	118,23	227,20	345,43	0,342
PC28	3707	215,71	320,60	536,31	0,402	216,38	320,74	537,12	0,403
PC35	3653	336,07	366,24	702,31	0,479	337,60	365,98	703,58	0,480
PC42	3607	439,96	587,67	1027,63	0,428	449,31	583,49	1032,80	0,435
	4952	8 gerações							
PC1	4921	0,55	0,56	1,11	0,496	0,55	0,56	1,11	0,495
PC7	4278	14,42	27,67	42,09	0,343	14,41	27,68	42,09	0,342
PC14	4124	55,51	122,93	178,44	0,311	55,85	122,80	178,65	0,313
PC21	4028	117,53	227,83	345,36	0,340	117,96	227,70	345,66	0,341
PC28	3999	222,52	321,36	543,88	0,409	221,91	321,82	543,73	0,408
PC35	3921	365,35	365,95	731,30	0,500	366,69	365,01	731,70	0,501
PC42	3876	447,29	627,45	1074,74	0,416	453,51	623,77	1077,28	0,421
	5362	9 gerações							
PC1	5331	0,54	0,56	1,10	0,493	0,55	0,56	1,11	0,495
PC7	4637	13,75	28,41	42,16	0,326	13,77	28,36	42,13	0,327
PC14	4464	55,49	127,67	183,16	0,303	55,74	127,39	183,13	0,304
PC21	4365	123,07	237,43	360,50	0,341	124,11	236,44	360,55	0,344
PC28	4334	227,30	341,60	568,90	0,400	230,13	339,31	569,44	0,404
PC35	4256	371,28	391,06	762,34	0,487	377,43	386,39	763,82	0,494
PC42	4200	441,54	650,14	1091,68	0,404	457,48	639,75	1097,23	0,417
	5726	10 gerações							
PC1	5695	0,53	0,56	1,09	0,487	0,53	0,56	1,09	0,486
PC7	4969	13,13	29,33	42,46	0,309	13,11	29,28	42,39	0,309
PC14	4779	51,79	132,67	184,46	0,281	51,97	132,33	184,3	0,282
PC21	4679	121,08	253,85	374,93	0,323	122,18	252,49	374,67	0,326
PC28	4643	225,05	381,21	606,26	0,371	228,43	377,95	606,38	0,377
PC35	4563	371,81	433,80	805,61	0,462	376,58	428,80	805,38	0,468
PC42	4499	430,22	687,08	1117,30	0,385	439,68	679,29	1118,97	0,393
		Aproximações dos erros de amostragem <sup>#</sup>							
PC1	(5966)	0,03254	0,02008		0,022				
PC7	(5217)	1,36346	1,00489		0,028				
PC14	(5012)	6,06378	4,59585		0,029				
PC21	(4915)	13,5079	9,60198		0,031				
PC28	(4879)	24,1327	16,0449		0,033				
PC35	(4795)	35,4009	21,4377		0,034				
PC42	(4727)	48,2244	31,2942		0,036				

232 \* NA: número total de animais do grupo (em linhas de identificação); NR: número de animais com peso corporal  
 233 registrado (em linhas de resultados). # Incluídas apenas para o grupo completo (10 gerações); os números  
 234 efetivos de genótipos usados na determinação da variância aditiva são destacados entre parêntesis.

235

236            Numa abordagem preliminar, considerando apenas as 7 gerações iniciais deste  
237 programa de melhoramento (DIONELLO et al., 2012), foi observada a similaridade entre as  
238 estimativas produzidas pelo REML e pela inferência bayesiana (naquela oportunidade,  
239 realizada com o apoio do programa MTGSAM). Adicionalmente, foi destacada a ocorrência  
240 de pequenas diferenças nas herdabilidades, cujas estimativas na inferência bayesiana, mais  
241 elevadas em 5,88% e 7,14%, respectivamente para PC21 e PC42, sinalizavam uma vantagem  
242 para a aplicação desta metodologia. Entretanto, com a consolidação e ampliação da base de  
243 dados (e uso do INTERGEN), estas diferenças desapareceram e a máxima diferença nesse  
244 grupo não passa de 1,2%, sendo de 2,1%, quando se considera as 10 gerações, ou mesmo no  
245 caso mais amplo, 3,2%, considerando somente as 9 gerações iniciais, sempre para PC42.  
246 FALCÃO et al. (2004), avaliando comparativamente componentes de variância para produção  
247 de leite de vacas da raça Holandesa, com modelo similar, porém incluindo efeito de ambiente  
248 permanente, também encontraram diferenças apenas "ligeiramente superiores" pró abordagem  
249 bayesiana para as estimativas de variância genética e herdabilidade. Em outro estudo  
250 comparativo (FALCÃO et al., 2009), também sobre bovinos de leite, com perfis ambientais e  
251 disponibilidade de dados diferenciados, foi observado que nem sempre ocorre a concordância  
252 entre as inferências obtidas pelas duas abordagens. A discordância é mais evidente nos casos  
253 de amostras pequenas. Também ficou evidenciado que o método bayesiano é mais sensível  
254 que o REML à quantidade de animais na matriz de parentesco.

255            Em ambas abordagens, as estimativas das herdabilidades variam dentro de uma faixa  
256 considerada de média a alta com valores de 0,28 a 0,50. A representação gráfica para o grupo  
257 completo (Figura 1) sinaliza que na proximidade dos 35 dias de idade há um máximo de  
258 influência genética aditiva na variabilidade do peso corporal ( $h_a^2 = 0,46$ ). Isto indica que a  
259 seleção dos animais poderia ser antecipada para 35 dias, para melhor aproveitamento da  
260 expressão genética, com relação à característica em estudo. Esta é a idade de início da vida

261 reprodutiva e já teria estabilizado o índice de viabilidade das aves, o qual apresenta-se mais  
262 baixo nos primeiros dias de vida. TEIXEIRA et al. (2013) estimaram herdabilidades um  
263 pouco mais baixas, variando de 0,26 a 0,44 no período dos 7 aos 42 dias, para dois outros  
264 grupos genéticos, usando apenas 5 gerações. As estimativas ao nascimento também foram  
265 mais altas e, nesses grupos, ainda mais elevadas (0,64 e 0,68) do que nos grupos do presente  
266 estudo. Entretanto, a variação da herdabilidade no período não reproduz a tendência  
267 observada neste estudo, embora em um dos grupos tenha ocorrido alguma similaridade com  
268 provável máximo de influência genética entre os 28 e os 35 dias de idade.

269         As aproximações dos erros de amostragem (*approximate sampling error*) fornecidos  
270 pelo WOMBAT (MEYER, 2007) para as estimativas dos componentes de variância obtidas  
271 por REML (exemplificadas na Tabela 3), permitem determinar os correspondentes intervalos  
272 de confiança, no nível de 95%. Comparando-se estes intervalos com os intervalos de  
273 credibilidade da média obtidos na abordagem bayesiana (exemplificados na Tabela 2),  
274 observa-se que os primeiros apresentam amplitudes menores, da ordem de 38% a 50%, em  
275 relação aos últimos, correspondentemente. Isto representa um alto grau de confiabilidade das  
276 estimativas obtidas pela abordagem clássica (via WOMBAT). Basta observar que estas  
277 estimativas reproduzem plenamente aquelas produzidas pela abordagem bayesiana (via  
278 INTERGEN), cujos intervalos de credibilidade, mais largos, são determinados em função das  
279 distribuições a posteriori diretamente relacionadas aos parâmetros (uma das qualidades dos  
280 métodos bayesianos). Embora os primeiros refiram-se a probabilidades relativas aos dados, as  
281 estimativas obtidas com a disponibilidade de uma grande quantidade de dados, como é o caso,  
282 são altamente confiáveis, conforme relatado por BLASCO (2001). Isto indica que é  
283 dispensável a aplicação de uma abordagem bayesiana, reconhecidamente mais complexa e  
284 onerosa, para casos similares de análise univariada. Resta como vantagem para a solução  
285 bayesiana a elegância de permitir a obtenção dos intervalos sem a necessidade de

286 aproximações e de pressuposições de normalidade. Invertendo o sentido de observação, fica  
287 bem caracterizada a adequação da abordagem bayesiana como alternativa para a obtenção de  
288 estimativas similares às produzidas pela abordagem clássica. Dessa forma, confirma-se a  
289 conclusão de Blasco, que associa a escolha por uma das abordagens à preferência do cientista,  
290 com exceção dos casos em que uma das abordagens oferece solução não disponível na outra.

291 O uso do aplicativo GUAIACA trouxe uma visão nova para as ferramentas  
292 empregadas nas análises (WOMBAT e INTERGEN), oferecendo uma interface amigável e  
293 integradora, facilitando as operações relacionadas à organização de dados e à execução  
294 controlada de múltiplas análises e comparações, envolvendo as diferentes ferramentas de  
295 software em um único ambiente operacional. Seus registros descritores demonstram que o  
296 tempo de execução de cada análise realizada pela abordagem bayesiana (via INTERGEN) é  
297 de 180 a 270 vezes maior que o tempo necessário para executar a análise correspondente pela  
298 abordagem clássica (via WOMBAT). Para a primeira, este tempo varia de 6 a 9 minutos e,  
299 para a última, é de aproximadamente 2 segundos, nas condições e características das análises  
300 programadas.

301

## 302 CONCLUSÕES

303 A idade de 35 dias é a que melhor salienta a herdabilidade para a característica peso  
304 corporal em codornas de corte e sua estimativa ( $h^2 = 0,46$  após 10 gerações) é muito próxima  
305 dos mais altos valores obtidos ao nascimento.

306 A abordagem bayesiana comprova plenamente os resultados da abordagem clássica  
307 (REML). Para análises unicaracterísticas com conhecimento pleno de informações de  
308 *pedigree*, não é possível evidenciar diferenças entre as duas abordagens quanto aos valores  
309 dos parâmetros genéticos estimados.

310 A abordagem clássica usando a metodologia REML via WOMBAT é suficiente para  
311 a obtenção de boas estimativas dos parâmetros, pois o custo computacional e operacional é  
312 bastante inferior ao exigido pela abordagem bayesiana via INTERGEN e os resultados são  
313 equivalentes.

314 Um possível limiar entre os modelos de análise, a partir do qual se demonstre  
315 vantagens efetivas para a abordagem bayesiana, pode ser perseguido pela continuidade deste  
316 estudo com aplicações de análises multivariadas e de regressão aleatória.

317

## 318 **AGRADECIMENTOS**

319 Ao CNPq pela concessão de financiamento via Edital Universal, permitindo a  
320 manutenção da linhagem de codornas de corte, fonte primária de dados para esta pesquisa.

321

## 322 **REFERÊNCIAS**

323 BLASCO, A. The Bayesian controversy in animal breeding. **Journal of Animal Science**,  
324 v.79, p.2023-2046, 2001.

325 BONAFÉ, C.M. et al. Modelos de regressão aleatória considerando heterogeneidade de  
326 variâncias residuais na descrição de curva de crescimento de codornas européias. In:  
327 REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 45., 2008, Lavras.  
328 **Anais...** Lavras, 2008. CD-ROM.

329 BONAFÉ, C.M. et al. Modelos de regressão aleatória para descrição da curva de crescimento  
330 de codornas de corte. **Revista Brasileira de Zootecnia**, v.40, n.4, p.765-771, Abr. 2011.

331 CARDOSO, F.F. **Manual de utilização do programa INTERGEN – Versão 1.0 em**  
332 **estudos de genética quantitativa animal**. Bagé: Embrapa Pecuária Sul, 2008. 74p.

333 (Documentos, 74) ISSN 1982-5390

- 334 DIONELLO, N.J.L. et al. Estimativas da trajetória genética do crescimento de codornas de  
335 corte utilizando modelos de regressão aleatória. **Arquivo Brasileiro de Medicina**  
336 **Veterinária e Zootecnia**, v.60, n.2, p. 454-460, 2008.
- 337 DIONELLO, N.J.L. et al. Avaliação de pesos corporais em codornas de corte ajustados pelo  
338 uso de modelos de regressão aleatória. In: REUNIÃO ANUAL DA SOCIEDADE  
339 BRASILEIRA DE ZOOTECNIA, 46., 2009, Maringá. **Anais...** Maringá, 2009. CD-ROM.
- 340 DIONELLO, N.J.L. et al. Estimação de parâmetros genéticos de pesos corporais de codornas  
341 de corte por meio de inferência bayesiana. In: REUNIÃO ANUAL DA SOCIEDADE  
342 BRASILEIRA DE ZOOTECNIA, 49., 2012, Brasília. **Anais...** Brasília: Sociedade Brasileira  
343 de Zootecnia, 2012. CD-ROM.
- 344 FALCÃO, A.J.S.; MARTINS, E.N.; COSTA, C.N.; SAKAGUTI, E.S.; MAZUCHELI, J.  
345 Aplicação de métodos REML e Bayesiano via amostrador de Gibbs na estimação de  
346 componentes de variância para produção de leite no estado do Paraná. In: SIMPÓSIO DA  
347 SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 5., 2004, Pirassununga.  
348 **Anais eletrônicos...** Pirassununga: Sociedade Brasileira de Melhoramento Animal, 2004.  
349 Disponível em: <<http://sbmaonline.org.br/anais/v/trabalhos/>>. Acesso em: 20 jul. 2011.
- 350 FALCÃO, A.J.S. et al. Efeitos do número de animais na matriz de parentesco sobre as  
351 estimativas de componentes de variância para produção de leite usando os métodos de  
352 Máxima Verossimilhança Restrita e Bayesiano. **Revista Brasileira de Zootecnia**, v.38, n.8,  
353 p.1478-1487, 2009.
- 354 HENDERSON, C.R. **Applications of linear models in animal breeding**. University of  
355 Guelph, Guelph. 462p, 1984.
- 356 HIDALGO, A. M. et al. Genetic characterization of egg weight, egg production and age at  
357 first egg in quails. **Revista Brasileira de Zootecnia**, v.40, n.1, p.95-99, 2011.

- 358 MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by  
359 restricted maximum likelihood (REML). **Journal of Zhejiang University SCIENCE B**, v.8,  
360 n.11, p.815–821, 2007.
- 361 MEYER, K. **WOMBAT - a program for mixed model analyses by restricted maximum**  
362 **likelihood**. User Notes. Armidale, Australia: Animal Genetics and Breeding Unit, University  
363 of New England, 2006-2014. 107p. Disponível em:  
364 <<http://didgeridoo.une.edu.au/km/download.php?file=WombatManual.pdf>>. Acesso em: 18  
365 mar. 2015.
- 366 SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC Methods in**  
367 **Quantitative Genetics**. New York: Springer-Verlag New York, 2002. 740p.
- 368 TEIXEIRA, B.B. et al. Herdabilidade de características de produção e postura em matrizes de  
369 codornas de corte. **Ciência Rural**. v.43, n.2, p.361-365, 2013.
- 370 WANG, C.S.; RUTLEDGE, J.J.; GIANOLA, D. Marginal inferences about variance  
371 components in a mixed linear model using Gibbs sampling. **Genetic Selection Evolution**,  
372 v.25, p.41-62, 1993.
- 373 WINTER, E.M.W. et al. Aplicação do método Bayesiano na estimação de correlações  
374 genéticas e fenotípicas de peso em codornas de corte em várias idades. **Revista Brasileira de**  
375 **Zootecnia**, v.35, n.4, p.1684-1690, 2006.
- 376

#### 4 Considerações Finais

O aplicativo Guaiaca resultante, modelado como um "cinto de utilidades", apresenta seus recursos distribuídos em dois módulos - um Assistente de Dados e um Assistente de Execução. Ele implementa uma interface operacional configurável, mas baseada em um padrão único para qualquer aplicativo de análise que seja integrado. A versão inicial para Windows está disponível gratuitamente para uso acadêmico. Esta versão foi configurada e testada para aplicação integrada com os dois aplicativos de análises utilizados neste trabalho - WOMBAT e INTERGEN. Seu desenvolvimento será continuado para a inclusão de mais facilidades.

O ambiente integrado foi utilizado para o estudo de caso proposto, com dados cumulativos de pedigree e peso corporal de 7, 8, 9 e 10 gerações de codornas de corte, resultando nas seguintes conclusões:

- a) a idade de 35 dias é a que melhor salienta a herdabilidade para essa característica ( $h^2 = 0,46$  após 10 gerações);
- b) não é possível evidenciar diferenças entre os resultados obtidos pelas abordagens Bayesiana e clássica, ou seja, são equivalentes os valores dos parâmetros genéticos estimados, para o caso estudado;
- c) a abordagem clássica usando a metodologia REML é suficiente para a obtenção de boas estimativas dos parâmetros, considerando os aplicativos utilizados e o custo computacional adicional associado à abordagem Bayesiana.

A generalidade adotada como meta no desenvolvimento da ferramenta Guaiaca representa uma contribuição adicional, pois sua aplicação não fica limitada à área de melhoramento genético de animais, focada neste trabalho, podendo ser útil em qualquer área em que os mesmos padrões de manipulações e análises de dados sejam necessários.

## Referências

- BLASCO, A. The Bayesian controversy in animal breeding. **Journal of Animal Science**, v.79, p.2023-2046, 2001.
- BOLDMAN, K.G.; KRIESE, L.A.; Van VLECK, L.D. et al. **A manual for use of MTDFREML. A set of programs to obtain estimates of variance and covariance**. Lincoln: USDA/ARS, 1993. 120p.
- BONAFÉ, C.M.; TEIXEIRA, R.B.; SOUSA, M.F. et al. Modelos de regressão aleatória considerando heterogeneidade de variâncias residuais na descrição de curva de crescimento de codornas européias. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 45., 2008, Lavras. **Anais...** Lavras: SBZ, 2008. CD-ROM.
- BONAFÉ, C.M.; TORRES, R.A.; SARMENTO, J.L.R. et al. Modelos de regressão aleatória para descrição da curva de crescimento de codornas de corte. **Revista Brasileira de Zootecnia**, v.40, n.4, p.765-771, Abr. 2011.
- CARDOSO, F.F. **Manual de utilização do programa INTERGEN – Versão 1.0 em estudos de genética quantitativa animal**. Bagé: Embrapa Pecuária Sul, 2008. 74p. (Documentos, 74).
- CRUZ, C.D. GENES - a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum. Agronomy**, v. 35, n. 3, p.271-276, 2013.
- DELWICHE, L.D.; SLAUGHTER, S.J. **The Little SAS Book: A Primer**. 5a. ed. Cary, NC: SAS Institute Inc, 2012.
- DIONELLO, N.J.L.; CORREA, G.S.S.; SILVA, M.A et al. Estimativas da trajetória genética do crescimento de codornas de corte utilizando modelos de regressão aleatória. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.60, n.2, p. 454-460, 2008.
- DIONELLO, N.J.L.; GOTUZZO, A.G.; DALLMANN, H. et al. Avaliação de pesos corporais em codornas de corte ajustados pelo uso de modelos de regressão aleatória. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 46., 2009, Maringá. **Anais...** Maringá: SBZ, 2009. CD-ROM.
- DIONELLO, N.J.L.; VAYEGO, S.A.; MEDEIROS, G.C.R. et al. Estimação de parâmetros genéticos de pesos corporais de codornas de corte por meio de inferência bayesiana. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 49., 2012, Brasília. **Anais...** Brasília: SBZ, 2012. CD-ROM.
- FALCÃO, A.J.S.; MARTINS, E.N.; COSTA, C.N.; SAKAGUTI, E.S.; MAZUCHELI, J. Aplicação de métodos REML e Bayesiano via amostrador de Gibbs na estimação de componentes de variância para produção de leite no estado do Paraná. In:

SIMPÓSIO DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 5., 2004, Pirassununga. **Anais eletrônicos...** Pirassununga: SBMA, 2004. Disponível em: <<http://sbmaonline.org.br/anais/v/trabalhos/>>. Acesso em: 20 jul. 2011.

FALCÃO, A.J.S.; MARTINS, E.N.; COSTA, C.N.; MAZUCHELI, J. Efeitos do número de animais na matriz de parentesco sobre as estimativas de componentes de variância para produção de leite usando os métodos de Máxima Verossimilhança Restrita e Bayesiano. **Revista Brasileira de Zootecnia**, v.38, n.8, p.1478-1487, 2009.

FERREIRA, D.F. **Estatística Computacional em Java**. 1. ed. Lavras: Ed. UFLA, 2013. v. 1. 695p.

GARRICK, D.J.; SAATCHI, M. Opportunities and challenges for genomic selection of beef cattle. **Revista Brasileira de Zootecnia**, v.40, p.310-316, 2011. ISSN 1806-9290.

GIANOLA, D.; FERNANDO, R.L. Bayesian Methods in Animal Breeding Theory. **Journal of Animal Science**, v.63, n.1, p.217-244, 1986.

GILMOUR, A.R.; GOGEL, B.J.; CULLIS, B.R.; THOMPSON, R. **ASReml User Guide Release 3.0**. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, 2009. Disponível em: <<http://www.vsnl.co.uk>>. Acesso em: 24 nov. 2014.

HENDERSON, C.R. **Applications of linear models in animal breeding**. University of Guelph, Guelph. 462p, 1984.

HIDALGO, A.M.; MARTINS, E.N.; SANTOS, A.L. et al. Genetic characterization of egg weight, egg production and age at first egg in quails. **Revista Brasileira de Zootecnia**, v.40, n.1, p.95-99, 2011.

LEGARRA, A.; AGUILAR, I.; MISZTAL, I. A relationship matrix including full pedigree and genomic information. **Journal of Dairy Science**, v.92, n.9, p.4656-4663, 2009.

LUSH, J.L. **Animal Breeding Plans**. 3a. ed. Ames: Iowa State University Press, 1945. 443p.

MADAN, A.; DUBEY, S.K. Usability evaluation methods: a literature review. **International Journal of Engineering Science and Technology**, v.4, n.2, p.590-599, 2012.

MEDEIROS, G.C.R.; BARROS, W.S.; DIONELLO, N.J.L. Avaliação da influência da endogamia no peso corporal de aves em programa de melhoramento de codornas de corte. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 49., 2012, Brasília. **Anais...** Brasília: SBZ, 2012. CD-ROM.

MEYER, K. DFREML - A set programs to estimate variance components under an individual animal model. **Journal of Dairy Science**, v.71, suppl. 2, p.33-34, 1988.

MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). **Journal of Zhejiang University SCIENCE B**, v.8, n.11, p.815-821, 2007.

- MEYER, K. **WOMBAT - a program for mixed model analyses by restricted maximum likelihood**. User Notes. Armidale, Australia: Animal Genetics and Breeding Unit, University of New England, 2006-2014. 107p. Disponível em: <<http://didgeridoo.une.edu.au/km/download.php?file=WombatManual.pdf>>. Acesso em: 18 mar. 2015.
- MISZTAL, I.; TSURUTA, S.; LOURENÇO, D. et al. **Manual for BLUPF90 family of programs**. Athens, USA: University of Georgia, 2014. Disponível em: <<http://nce.ads.uga.edu/wiki/doku.php>>. Acesso em: 24 nov. 2014.
- MRODE, R.A. **Linear models for the prediction of animal breeding values**. 2a. ed. Wallingford: CABI Publishing, 2005. 344p.
- PAULINO, C.D; TURKMAN, A.A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003. 446p.
- RESENDE, M.D.V. **Selegen-Reml/Blup : Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos**. 1. ed. Colombo: Embrapa Florestas, 2007. 360p .
- SAS INSTITUTE INC. **SAS 9.3 Companion for Windows**. Cary, NC: SAS Institute Inc, 2010.
- SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics**. New York: Springer-Verlag New York, 2002. 740p.
- TEIXEIRA, B.B.; EUCLYDES, R.F.; TEIXEIRA, R.B.; SILVA, L.P.; TORRES, R.A.; SILVA, F.G.; LEHNER, H.G.; CAETANO, G.C. Herdabilidade de características de produção e postura em matrizes de codornas de corte. **Ciência Rural**. v.43, n.2, p.361-365, 2013.
- VAN TASSEL, C.P.; VAN VLECK, L.D. Multiple-trait Gibbs sampler for animal models: flexible programs for bayesian and likelihood based (co)variance component inference. **Journal of Animal Science**, v.74, p. 2586-2597, 1996.
- WANG, C.S.; RUTLEDGE, J.J.; GIANOLA, D. Marginal inferences about variance components in a mixed linear model using Gibbs sampling. **Genetic Selection Evolution**, v.25, p.41-62, 1993.
- WINTER, E.M.W.; ALMEIDA, M.I.M.; OLIVEIRA, E.G. et al. Aplicação do método Bayesiano na estimação de correlações genéticas e fenotípicas de peso em codornas de corte em várias idades. **Revista Brasileira de Zootecnia**, v.35, n.4, p.1684-1690, 2006.