

Documentos

196

Campinas, SP / Julho, 2025

Software BRBLUP Manual do Usuário



***Empresa Brasileira de Pesquisa Agropecuária
Embrapa Agricultura Digital
Ministério da Agricultura e Pecuária***

e-ISSN 2764-2488

Documentos 196

Julho, 2025

**Software BRBLUP
Manual do Usuário**

Roberto Hiroshi Higa

***Embrapa Agricultura Digital
Campinas, SP
2025***

Embrapa Agricultura Digital

Av. Dr. André Tosello, 209 - Cidade Universitária
Campinas, SP, 13083-886

www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações

Presidente

Júlio Cesar Dalla Mora Esquerdo

Secretária-executiva

Sônia Ternes

Membros

Adauto Luiz Mancini, Alan Massaru Nakai,
Carla Cristiane Osawa, Geraldo Magela
de Almeida Cançado, Graziella Galinari,
Joice Machado Bariani, Juliana Erika de
Carvalho Teixeira Yassitepe, Luiz Manoel
Silva Cunha, Magda Cruciol e Paula
Regina Kuser Falcão.

Revisão de texto

Graziella Galinari

Projeto gráfico

Leandro Sousa Fazio

Diagramação

Ana Caroline Bianchi

Foto da capa

Fulano de Tal

Publicação digital: PDF

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte,
constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Agricultura Digital

Higa, Roberto Hiroshi

Software BRBLUP: manual do usuário / Roberto Hiroshi Higa. – Campinas :
Embrapa Agricultura Digital, 2025.

PDF (32 p.) : il. color. – (Documentos / Embrapa Agricultura Digital, e-ISSN 2764-
2488; 196)

1. Software. 2. Software brblup. 3. Avaliação genética. I. Título. II. Embrapa
Agricultura Digital. III. Série

CDD (21. ed.) 005.1068

Carla Cristiane Osawa (CRB-8/10421)

© 2025 Embrapa

Autor

Roberto Hiroshi Higa

Engenheiro Eletricista, doutor em Engenharia Elétrica, pesquisador da
Embrapa Agricultura Digital, Campinas, SP

Apresentação

Atualmente, a maioria dos programas nacionais de melhoramento genético animal dependem da utilização de softwares importados, via de regra desenvolvidos por universidades ou institutos de pesquisa estrangeiros, para realização de suas avaliações genéticas. Isso resulta em uma dependência tecnológica do setor em relação a esses fornecedores, além de um custo anual com licenciamento de software que pode alcançar 30 mil dólares/ano.

Desde 2014, a Embrapa Agricultura Digital, anteriormente Embrapa Informática Agropecuária, e a Embrapa Gado de Corte estão envolvidas no desenvolvimento de uma solução computacional, no escopo do programa de melhoramento genético Embrapa/Geneplus, capaz de realizar avaliações genéticas de grandes populações de animais. Resultado dessa colaboração, o software brblup tem por objetivo prover esta funcionalidade aos programas nacionais de melhoramento genético animal.

O brblup é escrito em linguagem Python e integra um pacote python denominado brbv, que disponibiliza um conjunto de softwares (interface por linha-de-comando) para suporte às atividades de avaliação genética de animais. Este manual apresenta aos usuários do brblup as informações necessárias para a sua aplicação na realização de avaliações genéticas.

Stanley Robson de Medeiros Oliveira

Chefe-Geral da Embrapa Agricultura Digital

Sumário

Introdução	9
Formato de arquivos	10
Avaliação genética com brblup	15
Exemplos	17
Modelo animal univariado	17
Modelo animal com covariáveis	18
Modelo de repetibilidade	19
Modelo animal bivariado com a mesma matriz de delineamento	21
Modelo animal bivariado com dados faltantes	22
Modelo animal univariado com efeito maternal	25
Single step GBLUP	27
Avaliação genética de grandes populações	29
Comentários finais	31
Referências	31

Introdução

Programas de melhoramento genético animal têm por objetivo promover o aumento de alelos desejáveis em uma população, identificando os indivíduos superiores para características de interesse da cadeia produtiva, particularmente aquelas de importância econômica. Softwares de avaliação genética apoiam esse objetivo, também servindo como ferramentas na pesquisa e desenvolvimento em genética e melhoramento. Sua funcionalidade principal consiste em utilizar dados de animais com genotipagem, fenotipagem e a genealogia completa da população de animais em avaliação para, com base em um modelo genético-estatístico pré-definido, calcular a DEPG (diferença esperada na progênie aprimorada pela genômica) para os animais dessa população.

Segundo Menezes (2015, 2018), a maioria dos programas nacionais de melhoramento genético animal dependem da utilização de softwares importados, via de regra desenvolvidos por universidades ou institutos de pesquisa estrangeiros, para realização de suas avaliações genéticas. Isso resulta em uma dependência tecnológica do setor em relação a esses fornecedores que, em termos econômicos, podem representar um custo anual com licenciamento de software da ordem de 30 mil dólares/ano.

Nesse sentido, desde 2014, a Embrapa Agricultura Digital, anteriormente Embrapa Informática Agropecuária, e a Embrapa Gado de Corte estão envolvidas no desenvolvimento de uma solução computacional, no escopo do programa de melhoramento genético Embrapa/Geneplus, capaz de realizar avaliações genéticas de grandes populações de animais. Resultado dessa colaboração, o software brblup disponibiliza um conjunto de funcionalidades para apoio às tarefas de avaliação genética de animais. Ele é escrito em linguagem Python, versão 3.8.10, para o sistema operacional Linux (distribuição Ubuntu 22.04.4 LTS) e integra um pacote Python denominado brbv, que disponibiliza um conjunto de softwares (interface por linha-de-comando) para suporte às atividades de avaliação genética de animais. Atualmente, ele é distribuído como um pacote que contém o software

e todas as bibliotecas necessárias à sua execução, gerado pelo software PyInstaller (Cortesi, 2025).

O objetivo deste documento é apresentar ao usuário do brblup, versão 3.6.9, o conjunto de informações necessárias para sua utilização no processo de avaliação genética de animais. As demais seções do presente documento encontram-se organizadas da seguinte forma: a) inicialmente, na seção "Formato de arquivos", são apresentados os formatos dos diferentes arquivos utilizados pelos softwares disponibilizados no pacote brbv; b) na sequência, são apresentadas suas respectivas opções de execução; c) segue-se uma série de exemplos ilustrativos de sua utilização; d) os comentários finais; e e) a bibliografia relevante.

Formato de arquivos

O arquivo de observações segue o formato csv, assim como a maior parte dos arquivos utilizados pelos softwares do brbv. Isso significa dizer que as observações possuem um número fixo de campos, separados por vírgula (',') e que dados faltantes são indicados por campos vazios ou contendo apenas caracteres 'branco'. Adicionalmente, convencionou-se que campos que representam classes ou categorias são representados por n números inteiros e sequenciais, variando de 0 a n-1. O exemplo da Figura 1 ilustra um trecho de um arquivo de observações com cinco campos. Note que na primeira observação, o quarto campo possui um valor faltante.

```
...  
10,2,7,, 6.0  
11,4,3,8,7.2  
...
```

Figura 1. Exemplo de arquivo de observações.

O arquivo de pedigree também segue o formato csv, mas com a restrição de que três das suas colunas possuem conteúdos específicos: o ID do animal, o ID do pai e o ID da mãe. Esses IDs devem variar de 0 a n-1, onde n é o número de animais no pedigree. Além disso, essa sequência deve ser tal que todo animal possua ID maior que o do seu pai e o de sua mãe, quando estes são conhecidos. Já quando um animal possui ambos os pais desconhecidos, essa linha pode ser omitida do arquivo de pedigree. O exemplo da Figura 2 corresponde a pedigree em que a primeira coluna contém o ID do animal, a segunda o ID do pai e a terceira o ID da mãe. Note que apesar do arquivo possuir apenas cinco entradas, o pedigree contém oito animais (os animais 0, 1 e 2 possuem pai e mãe desconhecidos e os correspondentes registros foram omitidos no arquivo).

```
3,2,1
4,0,1
5,0,
6,5,3
7,2,4
```

Figura 2. Exemplo de arquivo de pedigree.

Para os dados de genotipagem, dois arquivos são utilizados:

- Um arquivo com duas colunas com campos de tamanho fixo para os genótipos; e
- um arquivo csv com três colunas que relacionam os animais no pedigree com os respectivos genótipos.

O exemplo da Figura 3 ilustra o conteúdo típico de um arquivo de genótipos:

```

12 2011000212
13 1000520210
14 2000012212
...
23 1011020100
24 0051010221
25 0112000212

```

Figura 3. Exemplo de arquivo de genótipos.

A primeira coluna do arquivo contém o ID do genótipo, enquanto a segunda contém o genótipo propriamente dito, codificado com base no número de um dos alelos no SNP: 0, 1, 2 (homozigoto, heterozigoto e homozigoto alternativo). Por convenção, o brblup assume que o arquivo de genótipos não possui marcadores faltantes, usualmente codificados por 5, tendo sido submetido a processos prévios de controle de qualidade (QC) e imputação.

Para mapear os IDs dos animais no pedigree aos IDs dos genótipos, utiliza-se um arquivo csv com três colunas, conforme ilustra o exemplo da Figura 4. A primeira coluna contém o ID do animal, ordenado de acordo com o arquivo de pedigree, e a segunda o ID do correspondente genótipo. A terceira coluna é utilizada apenas no caso de utilização do algoritmo APY (Misztal et al., 2014), que interpreta o genótipo como 'core' se essa coluna possui valor 1 e 'não core' em caso contrário (usualmente, valor 0).

```

1,12,1
2,15,1
14,13,1
...
23,23,0
24,24,0
25,25,1

```

Figura 4. Exemplo de arquivo de mapeamento de ID no pedigree para ID de genótipos.

Os arquivos com a solução do modelo de equações mistas e as correspondentes acurácias também possuem formato csv, conforme ilustram as Figuras 5 e 6.

```
Trait,Effect,Level,Solution
0,1,0,4.3585023298549554
0,1,1,3.404430005906674
0,0,0,0.09844457570387921
0,0,1,-0.018770099100872666
0,0,2,-0.04108420292708503
0,0,3,-0.008663122661940758
0,0,4,-0.18573209949465097
0,0,5,0.17687208768130258
0,0,6,-0.24945855483362797
0,0,7,0.18261468793069588
```

Figura 5. Exemplo de arquivo com solução de equações de modelo misto.

O arquivo com a solução do modelo de equações mistas possui quatro campos: a característica, o efeito, o nível do efeito e o correspondente valor da solução.

```
Trait,Effect,Level,PEV,Rel,Acc,Sep,BIF
0,0,0,18.37321586616228,0.08133920669188599,0.28520029223667703,4.2863989392218595,0.04153205932169335
0,0,1,19.358726710176622,0.03206366449116893,0.17906329744302413,4.399855305595472,0.016162444552541455
0,0,2,17.880343158665056,0.10598284206674724,0.32555006076907317,4.2285154792982675,0.054475194437896945
0,0,3,16.67246095295381,0.16637695235230954,0.4078933100117107,4.083192495211781,0.08697040154894731
0,0,4,18.79229239769769,0.06038538011511552,0.2457343690148277,4.335007773660584,0.030662793510491282
0,0,5,17.21243778693954,0.13937811065302308,0.3733337791481278,4.148787508048531,0.07230291078015294
0,0,6,16.399229341686915,0.1800385329156542,0.4243094777584566,4.0495961948892389,0.0944827626796132
0,0,7,16.130625361014673,0.1934687319492664,0.43985080646654084,4.016294979332901,0.10192914085205196
```

Figura 6. Exemplo de arquivo com acurácia da solução de equações de modelo misto.

O arquivo com a acurácia da solução de equações de modelo misto possui oito campos: a característica, o efeito, o nível do efeito e os correspondentes valores para a variância do erro de predição, a confiabilidade, a acurácia, o erro quadrático predito e a acurácia BIF (Beef Improvement Federation, 2024). Note que os valores nos campos Trait e Effect referem-se à correspondente ordem de especificação no arquivo de configuração xml, enquanto que o nível do efeito refere-se ao correspondente valor na coluna do arquivo de observações associada ao efeito.

Por fim, um arquivo em formato xml, denominado arquivo de configuração, é utilizado para especificar o modelo de avaliação. Para fins de ilustração, o exemplo da Figura 7 apresenta um modelo animal univariado, onde a característica de interesse é o ganho de peso dos animais antes da desmama e o sexo é considerado como efeito fixo.

```
<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_3.1/data.csv" type="obs"/>
    <file filename="examples/mrode_3.1/pedigree.csv" type="ped"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="2" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="8"/>
    <effect type="alpha" columns="1" levels="2"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1"/>
    <!-- residual covariances -->
    <covariances>40</covariances>
    <group effects="0" type="animal" inbreeding="no">
      <!-- genetic covariances. -->
      <covariances>20</covariances>
    </group>
  </model>
</brblup>
```

Figura 7. Especificação de modelo animal univariado.

É possível observar que o arquivo de configuração possui duas seções principais, delimitadas pelas tags XML <files> e <model>. Na primeira, são especificados os arquivos de entrada, cada um, por sua vez, especificado por meio de uma tag XML <file>. Já na segunda seção, são especificados os elementos do modelo de avaliação propriamente dito, utilizando o seguinte conjunto de tags XML:

- a tag XML <trait> é utilizada para especificar a coluna no arquivo de observações correspondente ao valor observado para a característica;

- a tag XML <effect> é utilizada para especificar o tipo do efeito, a correspondente coluna no arquivo de observações e o número de níveis do efeito, quando for o caso;

- a tag XML <equation> é utilizada para especificar a equação que associa a característica aos efeitos. Note que a equação faz referência à característica e aos efeitos utilizando uma numeração que corresponde à ordem (0 para o primeiro trait/efeito, 1 para o segundo, etc) em que as correspondentes tags aparecem no arquivo de configuração xml.

- a tag XML <group> especifica um conjunto de efeitos aleatórios dependentes, por exemplo efeitos genético animal e maternal;

- a tag XML <covariances> é utilizada para especificar as matrizes de covariâncias; quando ela ocorre no escopo da tag XML <model>, assume-se que a matriz de covariâncias é a residual. Enquanto que quando ela ocorre no escopo da tag XML <group>, assume-se que a matriz de covariâncias refere-se a efeitos especificados no grupo, por exemplo, se a matriz está no escopo do grupo de efeitos genéticos a matriz é entendida como sendo a matriz covariâncias genéticas.

Elementos adicionais do arquivo XML, utilizados para especificar diferentes modelos de avaliação, serão descritos na subseção “Exemplos”, quando uma série de exemplos será apresentada.

Avaliação genética com brblup

O comando `brblup` calcula a solução BLUP (ou ssGBLUP) e, opcionalmente, as correspondentes acurácias para um modelo de avaliação especificado em um arquivo de configuração xml. Sua sintaxe e opções disponíveis são apresentadas na Figura 8.

```
usage: brblup [-h] [-g {yes,no}] [-o OUTPUT] [-f FACTOR] [-m {pcg,chol}] [-p {yes,no}] [-t TOL] [-n NITER] [-z TOL1] [-s STEPS1] [-v TOL3] [-u NITER3] input
```

Mixed model solver (BLUP). Version 3.6.9.

positional arguments:

input path to a xml file describing the evaluation model.

optional arguments:

-h, --help show this help message and exit

-g {yes,no}, --cfg {yes,no}

print selected options only. Default: no

-o OUTPUT, --output OUTPUT

directory to write output file(s). Default: .

-f FACTOR, --factor FACTOR

maximum number of rows/cols of intermediary matrices in memory. Default: 20000

-m {pcg,chol}, --method {pcg,chol}

method to solve the Mixed Model Equations: interactive (PCG) or direct (Cholesky decomposition). Default: pcg

-p {yes,no}, --pev {yes,no}

calculate accuracy (yes/no). Default: no

-t TOL, --tol TOL threshold for convergence (absolute error tolerance). Default: 1e-08 (PCG only).

-n NITER, --niter NITER

threshold for convergence (maximum number of iterations). Default: 5000 (PCG only).

-z TOL1, --tol1 TOL1 threshold for convergence (relative error tolerance) for accuracy calculation. Default: 1e-03.

-s STEPS1, --steps1 STEPS1

frequency of feedback message while calculating accuracy. Default: 10.

-v TOL3, --tol3 TOL3 threshold for convergence (absolute error tolerance) for accuracy calculation. Default: 1e-06.

-u NITER3, --niter3 NITER3

threshold for convergence (maximum number of iterations) for accuracy calculation. Default: 2000.

Figura 8. Lista de opções para o comando `brblup`.

Note que o único parâmetro é o arquivo de configuração que especifica o modelo de avaliação (vide seção “Formato de arquivos”).

Exemplos

Esta seção ilustra a utilização do software brblup para solução de diferentes modelos de avaliação. Para cada exemplo é apresentado o arquivo de configuração XML que define o modelo de avaliação e a linha de comando utilizada para invocar o brblup. Para todos os exemplos, considera-se que os conjuntos de dados estão organizados sob o diretório './examples' e os arquivos correspondentes à solução, solution.csv, e à acurácia, pev.csv, são gerados no diretório corrente.

Modelo animal univariado

Este é o exemplo 3.1 de Mrode (2014) e avalia o ganho de peso dos animais na desmama como característica de interesse, tendo o sexo como efeito fixo. A variância genética é igual a 20 e a variância residual igual a 40. O arquivo de configuração é apresentado na Figura 7.

Para resolver o modelo, executa-se o comando brblup passando como parâmetro o arquivo de configuração examples/mrode_3.1/conf.xml:

```
> ./brblup examples/mrode_3.1/conf.xml
```

A solução das equações do modelo misto pode ser verificada no arquivo solution.csv, descrito na seção "Formato de arquivos". Também é possível calcular a acurácia associada à solução utilizando a opção -p yes.

```
> ./brblup -p yes examples/mrode_3.1cov/conf.xml
```

Neste caso, além do arquivo solution.csv também é gerado um arquivo pev.csv, descrito na seção "Formato de arquivos". A menos de diferenças de precisão numérica, ambos os resultados estão em correspondência com aqueles reportados em Mrode (2014). Por fim, note que o método de solução utilizado para resolver o sistema de equações do modelo misto é o gradiente conjugado pré-condiciona-

do, que corresponde à opção padrão. Utilizando a opção `-m chol` é possível obter os mesmos resultados, a menos de diferenças de precisão numérica, resolvendo o sistema de equações do modelo misto por meio do método direto baseado na decomposição de Cholesky. Em geral, esta opção demanda mais recursos computacionais (memória RAM e processamento), mas os resultados obtidos são mais precisos.

Modelo animal com covariáveis

Este exemplo é uma adaptação do exemplo anterior para ilustrar a especificação de covariáveis no modelo de avaliação. O arquivo de observações deste exemplo difere do anterior por possuir uma coluna a mais correspondendo a uma covariável. O arquivo de configuração é apresentado na Figura 9. Nele, observam-se duas diferenças em relação ao exemplo 'Modelo animal univariado': um efeito adicional do tipo 'real' associado à coluna adicional no arquivo de observações e a sua inclusão na equação que relaciona os efeitos à característica. Além disso, a especificação do arquivo de pedigree ilustra como indicar explicitamente as colunas correspondentes aos IDs do animal e de seus pais.

```
<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_3.1cov/data.csv" type="obs"/>
    <file filename="examples/mrode_3.1cov/pedigree.csv" type="ped" id="0" sire="1" dam="2"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="2" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="8"/>
    <effect type="alpha" columns="1" levels="2"/>
    <effect type="real" columns="3" levels="1"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1,2"/>
    <!-- residual covariances -->
    <covariances>40</covariances>
    <group effects="0" type="animal">
      <!-- genetic covariances. -->
      <covariances>20</covariances>
    </group>
  </model>
```

Figura 9. Especificação de modelo animal univariado com covariável.

Para resolver o modelo, o comando `brblup` é invocado de forma análoga à apresentada no exemplo ‘Modelo animal univariado’, passando como parâmetro o arquivo de configuração ‘examples/mrode_3.1cov/conf.xml’. Por ser uma adaptação, este exemplo não possui correspondência com Mrode (2014), entretanto, é possível verificar o desvio introduzido na solução do exemplo original pela introdução da covariável no modelo.

Modelo de repetibilidade

Este é o exemplo 4.1 de Mrode (2014) e corresponde à aplicação do modelo animal univariado a um conjunto de dados contendo mais de uma observação por animal, conhecido como modelo de repetibilidade. O exemplo considera o teor de gordura no leite em bovinos como característica e, além do efeito genético aditivo, também considera o efeito ambiental permanente e dois efeitos fixos, o número de partos e classe de rebanho-ano-temporada (do inglês `herd+year+season`, HYS). A variância genética considerada é igual a 20, a ambiental permanente é igual a 12 e a residual igual a 28. O arquivo de configuração é apresentado na Figura 10 e o modelo pode ser resolvido invocando o comando `brblup`, passando como parâmetro o arquivo de configuração ‘examples/mrode_4.1/conf.xml’. Adicionalmente, a especificação do grupo com efeito animal ilustra como indicar a utilização do coeficiente de consanguinidade no cálculo da matriz A^{-1} . Neste exemplo, a opção `inbreeding="no"` indica que o coeficiente de consanguinidade não deve ser utilizado.

```

<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_4.1/data.csv" type="obs"/>
    <file filename="examples/mrode_4.1/pedigree.csv" type="ped"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="4" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="8"/>
    <effect type="alpha" columns="1" levels="5"/>
    <effect type="alpha" columns="2" levels="2"/>
    <effect type="alpha" columns="3" levels="4"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,2,3,1"/>
    <!-- residual covariances -->
    <covariances>28</covariances>
    <group effects="1" type="identity">
      <!-- permanent environmental covariances. -->
      <covariances>12</covariances>
    </group>
    <group effects="0" type="animal" inbreeding="no">
      <!-- genetic covariances. -->
      <covariances>20</covariances>
    </group>
  </model>
</brblup>

```

Figura 10. Especificação de modelo animal com medidas repetidas.

A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

Modelo animal bivariado com a mesma matriz de delineamento

Este é o exemplo 5.1 de Mrode (2014) e corresponde a um modelo animal bivariado (ganho de peso antes da desmama e ganho de peso após a desmama) com a mesma matriz de delineamento e nenhum valor faltante. O modelo assume a característica sexo como efeito fixo e o efeito animal.

A covariância genética é dada pela matriz $\begin{pmatrix} 20 & 18 \\ 18 & 40 \end{pmatrix}$

e a matriz de covariância residual pela matriz $\begin{pmatrix} 40 & 11 \\ 11 & 30 \end{pmatrix}$

Para resolver o modelo, o comando `brblup` é invocado de forma análoga à apresentada no exemplo ‘Modelo animal univariado’, passando como parâmetro o arquivo de configuração ‘examples/mrode_5.1/conf.xml’. A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

```

<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_5.1/data.csv" type="obs"/>
    <file filename="examples/mrode_5.1/pedigree.csv" type="ped"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="2" levels="1"/>
    <trait type="real" column="3" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="8"/>
    <effect type="alpha" columns="1" levels="2"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1"/>
    <equation trait="1" effects="0,1"/>
    <!-- residual covariances -->
    <covariances>
      40,11,
      11,30
    </covariances>
    <group effects="0" type="animal">
      <!-- genetic covariances. -->
      <covariances>
        20,18,
        18,40
      </covariances>
    </group>
  </model>
</brblup>

```

Figura 11. Especificação de modelo animal bivariado.

Modelo animal bivariado com dados faltantes

Este é o exemplo 5.3 de Mrode (2014). Ele utiliza o mesmo modelo apresentado na seção anterior, tal que o arquivo de configuração é o

mesmo, menos a especificação dos arquivos de observações e pedigree, que encontram-se no diretório 'examples/mrode_5.3'.

Para resolver o modelo, o brblup é invocado de forma análoga à apresentada no exemplo 'Modelo animal univariado', passando como parâmetro o arquivo de configuração 'examples/mrode_5.3/conf.xml'.

A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

Modelo animal bivariado com diferentes matrizes de delineamento

Este é o exemplo 5.5 de Mrode (2014) e é construído a partir do exemplo 'Modelo de repetibilidade', utilizando o número de partos para definir as diferentes características e os correspondentes efeitos fixos.

A covariância genética é dada pela matriz
$$\begin{pmatrix} 35 & 28 \\ 28 & 30 \end{pmatrix}$$

e a matriz de covariância residual pela matriz
$$\begin{pmatrix} 65 & 27 \\ 27 & 70 \end{pmatrix}$$

O arquivo de configuração é apresentado na Figura 12.

Para resolver o modelo, o comando brblup é invocado de forma análoga à apresentada no exemplo 'Modelo animal univariado', passando como parâmetro o arquivo de configuração 'examples/mrode_5.5/conf.xml'. A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

```

<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_5.5/data.csv" type="obs"/>
    <file filename="examples/mrode_5.5/pedigree.csv" type="ped"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="3" levels="1"/>
    <trait type="real" column="4" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="8"/>
    <effect type="alpha" columns="1" levels="2"/>
    <effect type="alpha" columns="2" levels="2"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1"/>
    <equation trait="1" effects="0,2"/>
    <!-- residual covariances -->
    <covariances>
      65,27,
      27,70
    </covariances>
    <group effects="0" type="animal" inbreeding="no">
      <!-- genetic covariances. -->
      <covariances>
        35,28,
        28,30
      </covariances>
    </group>
  </model>
</brblup>

```

Figura 12. Especificação de modelo animal bivariado com diferentes delineamentos.

Modelo animal univariado com efeito maternal

Este é o exemplo 6.1 de Mrode (2014) e ilustra a utilização do brblup com um modelo que inclui efeito animal direto e maternal. O modelo considera a característica peso de bezerros ao nascimento e como efeitos fixos a classe de rebanho-ano-temporada (HYS), o sexo do animal e estima os efeitos genéticos direto e maternal, considerando como efeitos ambientais permanentes as mães com observações da característica alvo.

A covariância genética é dada pela matriz $\begin{pmatrix} 150 & -40 \\ -40 & 90 \end{pmatrix}$

a variância ambiental permanente é igual a 40 e a variância residual é igual a 350.

O arquivo de configuração é apresentado na Figura 13.

Para resolver o modelo, o comando brblup é invocado de forma análoga à apresentada no exemplo 'Modelo animal univariado', passando como parâmetro o arquivo de configuração 'examples/mrode_6.1/conf.xml'. A menos de diferenças de precisão numérica, o resultado está em correspondência com aquele reportado em Mrode (2014).

```

<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/mrode_6.1/data.csv" type="obs"/>
    <file filename="examples/mrode_6.1/pedigree.csv" type="ped"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="6" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="0" levels="14"/>
    <effect type="alpha" columns="2" levels="14"/>
    <effect type="alpha" columns="3" levels="4"/>
    <effect type="alpha" columns="4" levels="3"/>
    <effect type="alpha" columns="5" levels="2"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1,2,3,4"/>
    <!-- residual covariances -->
    <covariances>350</covariances>
    <group effects="2" type="identity">
      <!-- permanent environmental covariances. -->
      <covariances>40</covariances>
    </group>
    <group effects="0,1" type="animal" inbreeding="no">
      <!-- genetic covariances: direct and maternal -->
      <covariances>
        150,-40,
        -40,90
      </covariances>
    </group>
  </model>
</brblup>

```

Figura 13. Especificação de modelo com efeito genético direto e maternal.

Single step GBLUP

Este exemplo ilustra a utilização da metodologia ssGBLUP implementada no brblup. Ele assume um modelo animal univariado com dois efeitos fixos e uma covariável, além do efeito animal, para explicar a característica de interesse. O parentesco entre os animais, utilizado para determinar as covariâncias genética entre os indivíduos, utiliza informações tanto do pedigree quanto de genotipagem com marcadores SNPs para subgrupos de animais, sendo as variância genética e residual iguais a 0.5 e 2.0, respectivamente. O arquivo de configuração é apresentado na Figura 14.

Note que:

1. Na seção <files>, a estrutura de parentesco é definida pela especificação do arquivo de pedigree (ped) em conjunto com as especificações dos arquivos de snps (snps) e de mapeamento animal-genótipo (map).
2. Na seção <group>, os atributos inbreeding, scale, alpha, tau, omega e apy especificam os parâmetros utilizados para obtenção da inversa da matriz de covariâncias. De fato, as especificações desses atributos poderiam ter sido omitidas, uma vez que os valores especificados correspondem aos valores padrão de cada uma.

```

<?xml version="1.0"?>
<brblup>
  <files>
    <file filename="examples/qblupf90_p34/data.sorted.csv" type="obs" idcol="4"/>
    <file filename="examples/qblupf90_p34/pedigree.sorted.csv" type="ped" id="0" sire="1" dam="2"/>
    <file filename="examples/qblupf90_p34/snps.txt" type="snps"/>
    <file filename="examples/qblupf90_p34/snps.map.sorted.csv" type="map"/>
  </files>
  <model type="animal">
    <!-- traits -->
    <trait type="real" column="0" levels="1"/>
    <!-- effects -->
    <effect type="alpha" columns="1" levels="3"/>
    <effect type="alpha" columns="2" levels="2"/>
    <effect type="real" columns="3"/>
    <effect type="alpha" columns="4" levels="15"/>
    <!-- one equation for each trait -->
    <equation trait="0" effects="0,1,2,3"/>
    <!-- residual covariances -->
    <covariances>
      2.
    </covariances>
    <group effects="3" type="animal">
      <!-- genetic covariances. -->
      <covariances>
        .5
      </covariances>
    </group>
  </model>
</brblup>

```

Figura 14. Especificação de modelo com alguns animais genotipados (ssGLUP).

Para resolver o modelo, o comando brblup é invocado de forma análoga à apresentada no exemplo ‘modelo animal univariado’, passando como parâmetro o arquivo de configuração ‘examples/qblupf90_p34/conf.xml’. Este exemplo não possui correspondência com Mrode (2014).

Avaliação genética de grandes populações

Para realizar a avaliação genética de populações muito grandes (milhões de animais no pedigree) utilizando a metodologia ssGBLUP, três diferentes comandos são invocados: `brsspp`, `brblup2` e `bracc`. Em relação ao `brblup`, esses comandos utilizam algumas estratégias para, durante o processamento, otimizar o consumo de memória RAM e/ou processamento: a) armazenando matrizes intermediárias em arquivo; b) utilizando o algoritmo `apy` (Misztal et al., 2014) para obtenção da matriz H^{-1} ; e c) utilizando uma estratégia (algoritmo) alternativa para aproximar a acurácia da solução.

O comando `brsspp` realiza o processamento dos arquivos de pedigree e genótipos gerando arquivos intermediários correspondentes às matrizes A^{-1} , G^{-1} , obtidas por meio do algoritmo APY (Misztal et al., 2014), e A^{-1} , que são utilizadas pelo comando `brblup2` para construção e resolução das equações de modelo misto. Sua sintaxe e opções disponíveis são apresentadas na Figura 15:

```
usage: brsspp [-h] [-o OUTPUT] [-t {ainv,nrm,grm0,grm,ginv,all}] [-f FACTOR] input
```

Pre-processing for single step analysis. Version 1.0.0.

positional arguments:

input path to a xml file describing the evaluation model.

optional arguments:

```
-h, --help show this help message and exit
-o OUTPUT, --output OUTPUT directory to write output file(s). Default: .
-t {ainv,nrm,grm0,grm,ginv,all}, --task {ainv,nrm,grm0,grm,ginv,all} target pre-processing. Dependencies: grm: nrm, grm0; ginv: grm. Default: all
-f FACTOR, --factor FACTOR memory usage factor for matrix in file. Default: 50000
```

Figura 15. Lista de opções para o comando `brsspp`.

O comando `brblup2` constrói e resolve o sistema de equações de modelo misto utilizando o algoritmo PCG, a partir dos arquivos intermediários gerados pelo comando `brppss`. A sintaxe e opções disponíveis para este comando são apresentadas na Figura 16.

```
usage: brblup2 [-h] [-g {yes,no}] [-i INPUT] [-o OUTPUT] [-t TOL] [-n NITER] xml

Mixed model solver (ssGBLUP). Version 0.9.8.

positional arguments:
  xml                  path to a xml file describing the evaluation model.

optional arguments:
  -h, --help            show this help message and exit
  -g {yes,no}, --cfg {yes,no}
                        print selected options only. Default: no
  -i INPUT, --input INPUT
                        directory containing pre-calculated file(s). Default: .
  -o OUTPUT, --output OUTPUT
                        directory to write output file(s). Default: .
  -t TOL, --tol TOL     threshold for convergence (absolute error tolerance). Default: 1e-08.
  -n NITER, --niter NITER
                        threshold for convergence (maximum number of iterations). Default: 5000.
```

Figura 16. Lista de opções para o comando `brblup2`.

Finalmente, para obter aproximações da acurácia utiliza-se o comando `bracc`. A sintaxe e opções disponíveis para este comando são apresentadas na Figura 17.

```
usage: bracc [-h] [-o OUTPUT] [-v INVG] input

Multivariate animal model accuracy approximation. Version 0.9.0.

positional arguments:
  input                path to a xml file describing the evaluation model.

optional arguments:
  -h, --help            show this help message and exit
  -o OUTPUT, --output OUTPUT
                        directory to write output file(s). Default: .
  -v INVG, --inv INVG   path to inverse of G obtained by using the APY algorithm. Default:
```

Figura 17. Lista de opções para o comando `bracc`.

Comentários finais

Neste documento foram apresentados ao usuário do brblup o conjunto de informações necessárias para sua utilização no processo de avaliação genética de animais. Foram apresentadas a sintaxe das interfaces por linha-de-comando dos softwares brblup, brsspp, brblup2 e bracc e um conjunto de exemplos ilustrativos aplicados a diferentes modelos. Espera-se, com isso, que o usuário do brblup esteja plenamente habilitado para utilizá-lo.

Por fim, cabe observar que os softwares apresentados neste documento encontram-se em contínua evolução, com a disponibilização de novas funcionalidades e/ou correções podendo ocorrer a qualquer momento.

Referências

BEEF IMPROVEMENT FEDERATION. **BIF guidelines wiki**. 2024.

Disponível em: <https://guidelines.beefimprovement.org/index.php/Accuracy>.

Acesso em: 10 fev. 2025.

CORTESI, D. **PyInstaller manual**. 2025. Disponível em: <https://pyinstaller.org/>. Acesso em: 5 fev. 2025.

MENEZES, G. R. de O. **Desenvolvimento e implementação de metodologias genético-estatísticas em avaliações genéticas de gado de corte**. Campo Grande, MS: Embrapa Gado de Corte, 2015. Projeto 02.13.14.004.00.00. Projeto concluído.

MENEZES, G. R. de O. **Otimização do fluxo genético em rebanhos sob seleção, de características econômicas relacionadas à eficiência alimentar, carcaça e reprodução, compondo índices de seleção e ferramental, visando ao melhoramento e uso da raça nelore**. Campo Grande, MS: Embrapa Gado de Corte, 2018. Projeto 20.18.01.013.00.00. Projeto concluído.

MISZTAL, I.; LEGARRA, A.; AGUILLAR, I. Using recursion to compute the inverse of the genomic relationship matrix. **Journal of Dairy Science**, v. 97, n. 6, p. 3943-3952, 2014. DOI: <https://doi.org/10.3168/jds.2013-7752>.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. 3rd ed. Oxfordshire: CABI, 2014.

