



XVI Simpósio Brasileiro de Melhoramento Animal
Piracicaba, SP – 04 e 05 de agosto de 2023

Brblup: utilizando python para avaliação genética aprimorada pela genômica em animais

Roberto Hiroshi Higa^{*1}, Ivan Carvalho Filho², Marcos Jun-Iti Yokoo³

¹Embrapa Agricultura Digital, Campinas/SP, Brasil; ²Geneplus Consultoria, Campo Grande/SP; ³Embrapa Pecuária Sudeste, São Carlos/SP.

*Autor correspondente – roberto.higa@embrapa.br

Python é uma linguagem de programação ao mesmo tempo simples e altamente expressiva, o que contribui para a produtividade tanto na construção de softwares quanto em sua manutenção ao longo do tempo. Além disso, a possibilidade de integrar bibliotecas, tais como NumPy e Scipy, facilita a adoção de implementações de algoritmos superiores. O software brblup é escrito em linguagem Python (versão 3.12.3) com foco na realização de avaliações genéticas aprimoradas pela genômica em um passo único. Este estudo teve por objetivo avaliar a sua adequabilidade na realização desta tarefa, ao estimar valores genéticos. Para isso, dois conjuntos de dados foram utilizados: O primeiro (CJ1) foi obtido por simulação utilizando o pacote AlphaSimR. Os valores iniciais de pesos e os componentes de variâncias foram fornecidos pelo Programa Embrapa Geneplus com base no banco de dados reais da raça Nelore. A população simulada foi composta por quinze gerações, totalizando 310.501 animais no pedigree, dos quais 20.000 animais das últimas 5 gerações possuíam genótipo. O banco de dados continha as informações de peso à desmama (PD) e peso ao sobreano (PS), além de sexo, grupo de contemporâneos e os respectivos valores genéticos verdadeiros. No conjunto de dados final, foram considerados apenas 10.000 animais genotipados escolhidos aleatoriamente. O segundo conjunto de dados (CJ2) foi obtido junto ao programa de melhoramento genético Brangus+, sendo composto por 112.794 animais, 41.095 deles com observações para os fenótipos peso ao nascimento (PN) e peso à desmama (PD), e 10.118 animais com genótipos. Em ambos os casos, os valores genéticos foram obtidos por meio do método ssGBLUP, conforme implementado no brblup. No caso do CJ1, os valores genéticos obtidos foram comparados com os valores genéticos verdadeiros, enquanto que no caso do CJ2, os valores genéticos obtidos foram comparados com os valores genéticos obtidos utilizando o software Blupf90. Em ambos os casos, as métricas utilizadas para comparação foram as correlações de Pearson e Spearman. As avaliações foram executadas em um desktop HP com 12 processadores com velocidade de 3.40GHz, 32GB de memória RAM com sistema operacional Ubuntu 20.04.6 LTS. O tempo de processamento no caso do CJ1 foi de 1,24 horas, sendo que 1,22 horas para processamento das informações de parentesco, enquanto que no caso do CJ2 o tempo de processamento foi de 5,36 minutos, sendo que 3,16 minutos para processamento das informações de parentesco. No caso do CJ1, para a característica PD as correlações foram de 0,963 (Pearson) e 0,968 (Spearman) e para a característica PS as correlações foram de 0,965 (Pearson) e 0,969 (Spearman). Já no caso do CJ2, as correlações foram de 1,000 (Pearson) e 0,999 (Spearman) para ambas as características PD e PN. Esses resultados demonstram que o brblup, uma tecnologia baseada em Python, é um software adequado para realização de avaliações genéticas de populações de animais contendo até centenas de milhares de animais em seu pedigree e dezenas de milhares de animais genotipados.

Palavras-chave: avaliação genética, python.

Agradecimentos: Aos parceiros de pesquisa, Embrapa (SEG:20.23.08.001.00.00), Geneplus e a Associação Brasileira de Brangus pelo suporte financeiro na coleta de dados e realização dos experimentos.