

UNIVERSITY OF GEORGIA  
College of Agricultural & Environmental Sciences

# Best practices in BLUPF90

Daniela Lourenco  
BLUPF90 TEAM – 02/2023

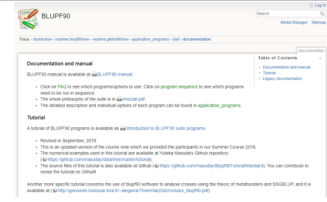
1

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

1. Read the manual
2. Understand what the programs can do for you
3. Understand the theory
4. Read the output of the programs
 

```
blupf90+ renf90.par | tee blup.log
```



1. Understand the data
  - How many genotypes?
  - How many phenotypes?
  - How many geno with pheno?
  - Outliers?
  - Population structure (PCA)

2

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- Steps to estimate breeding values using BLUPF90 family of programs

1. renumf90 to renumber the data
2. Estimate variance components if they are not available
  1. If need to estimate: blupf90+ with OPTION method VCE  
The solutions file has EBV computed with the estimated VC
  1. If available: put the values into renf90.par and run blupf90+ to estimate (G)EBV

3

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- Remove animals from the data file that have missing phenotypes
  - They are just increasing the number of equations
  - They will get predictions if they are in the pedigree or genotype file
  - In renumf90: OPTION remove\_all\_missing
- Dealing with genomic information
  - Run preGSf90 to perform quality control and save the clean files
  - OPTION saveCleansNP
  - OPTION createGInverse 0
  - OPTION createA22Inverse 0
  - OPTION createGimA22i 0
  - The 3 last options avoid extensive computations

4

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- Dealing with genomic information
  - Work with the clean file: `snp.txt_clean`
  - `OPTION SNP_file snp.txt_clean`
  - `OPTION no_quality_control`
- Investigate the output
  - What is the correlation between **G** and **A<sub>22</sub>**?
  - Diagonal = inbreeding correlation
  - All → ideal range [0.5 – 0.9]
    - Lower = incompatibility of genomic and ped information (at least one is wrong)
    - Lower = admixed population
    - Higher = limited advantage of genomic information

5

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- Parent-progeny conflicts?
  - `preGSf90`: removes genotypes for progeny with conflict
  - `Seekparentf90` can search for possible parents among genotyped animals
    - Corrects the pedigree
  - Which one is better?
    - Pedigree and genotypes may be wrong
    - Knowledge about the data

6

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- Running any software that takes a parameter file besides `renumf90`?
  - Use `OPTION use_yams`
  - Speeds up several computations, especially with multi-trait models
- Estimating variance components with REML or AIREML in `blupf90+`?
  - Use `OPTION use_yams`
  - Do not use UPG or metafounders
    - Weird behavior if inverting the LHS of MME as in VCE
  - Need to know if convergence was reached?
    - Try `blupf90+` and `gibbsf90+`

7

UNIVERSITY OF GEORGIA

## Best practices – BLUPF90

- `blupf90+` and `gibbsf90+` can compute SE for estimates
  - `OPTION se_covar_function <label> <function>`

```

Sampling variances of covariances function of random effects (n=10000)
h2d - Function: g_2_2_1_1/(g_2_2_1_1+r_1_1)
Mean: 0.31113
Sample Mean: 0.31103
Sample SD: 0.82903E-02
elapsed time 2.1800995E-03
  
```

- Is it SD or SE?
  - Sample SD = SE

8

## Best practices – BLUPF90

- The output depends on the input

$$f(\text{trash}) = \text{trash}$$

$$f(\text{trash}) = \text{trash}^3$$

Plus  
genomics