Estimating breeding values with BLUPF90

Slides adapted from Gabriela Mafra Fortuna, Raphael Mrode, Gregor

Gorjanc

ILRI

Feb 2026

@HighlanderLab

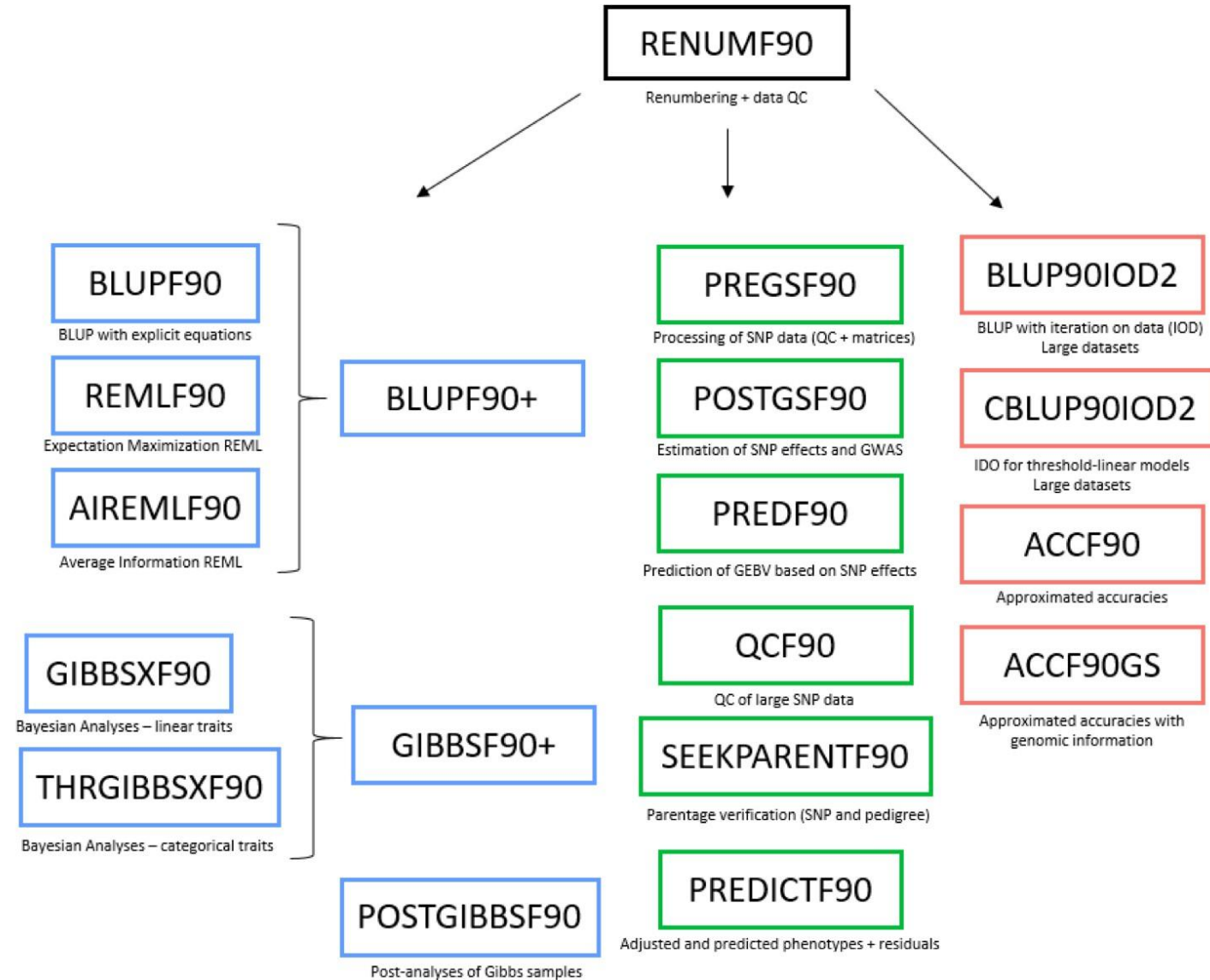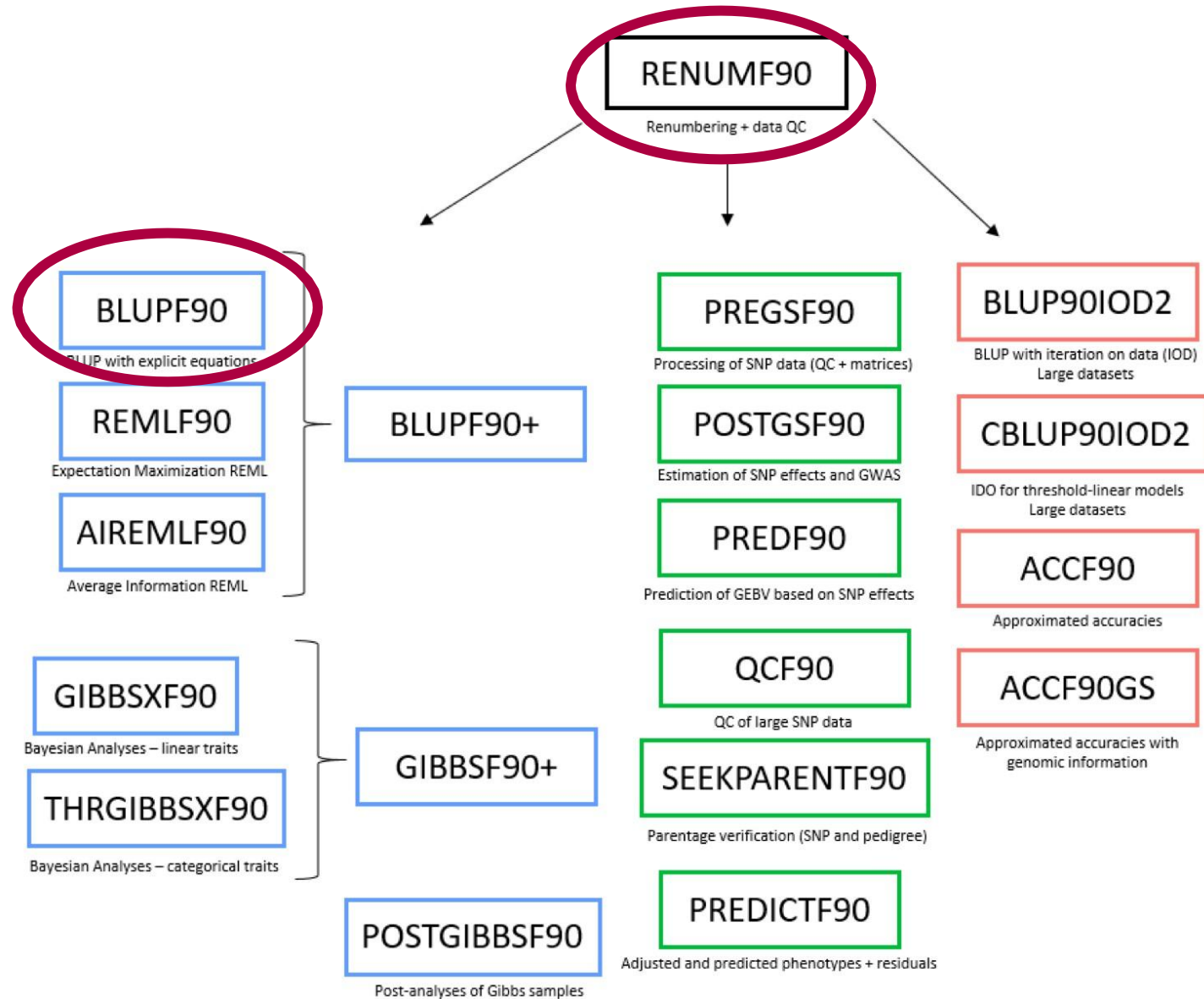# What is BLUPF90

- Programme for solving mixed-models
- Estimate variance components
- Estimate approximated model accuracy
- Perform Genome-wide association studies (GWAS)

What is BLUPF90

RENUMF90
Renumbering + data QC

BLUPF90 — BLUP with explicit equations
REMLF90 — Expectation Maximization REML
AIREMLF90 — Average Information REML
BLUPF90+

GIBBSXF90 — Bayesian Analyses – linear traits
THRGIBBSXF90 — Bayesian Analyses – categorical traits
GIBBSF90+
POSTGIBBSF90 — Post-analyses of Gibbs samples

PREGSF90 — Processing of SNP data (QC + matrices)
POSTGSF90 — Estimation of SNP effects and GWAS
PREDF90 — Prediction of GEBV based on SNP effects
QCF90 — QC of large SNP data
SEEKPARENTF90 — Parentage verification (SNP and pedigree)
PREDICTF90 — Adjusted and predicted phenotypes + residuals

BLUP90IOD2 — BLUP with iteration on data (IOD) Large datasets
CBLUP90IOD2 — IDO for threshold-linear models Large datasets
ACCF90 — Approximated accuracies
ACCF90GS — Approximated accuracies with genomic information

What is BLUPF90

For this course

- PBLUP
  - Pedigree
  - Phenotype
- GBLUP
  - SNP
  - Phenotype
- ssGBLUP
  - Pedigree
  - SNP
  - Phenotype

Input files

- Pedigree format:
  - Iid fid mid
  - No header
  - Cannot contain #
  - Space delimited
  - Missing data is coded as 0
  - No need to be ordered

id1	0	0
id2	0	0
id3	0	id2
id4	id1	id2
id5	id3	id4
id6	id3	id4

Input files

- Data file:
  - No header
  - Cannot contain #
  - Space delimited
  - Missing is coded 0

```
head sheep.dat

1011001 1 2001 1 172 2 3 1 1 37.9 2840 43.02 0.00782018 -0.079224108 0.032544493 0.4282 10.5 1.478
1011002 1 2001 1 172 2 3 1 1 37.9 2972 44.29 0.01781932 -0.078670819 0.03151353 0.4391 142.6 2.723
1011003 1 2001 1 173 2 3 1 1 30.6 2732 44.11 0.015338981 -0.004697449 -0.012611516 -7.083 -61.77 2.292
1011004 2 2001 1 173 2 3 1 1 33 2476 41.51 0.01243738 -0.006046247 -0.000384095 -1.513 -12.37 1.778
1011005 1 2001 1 174 2 3 1 3 40.7 3316 40.13 0.006278096 -0.12210416 0.075196885 3.402 467.6 -1.202
1011006 2 2001 1 174 2 3 1 3 31.5 2524 37.32 -0.000254866 -0.127165519 0.078914 -2.654 -21.85 -1.943
1011007 1 2001 1 175 2 4 1 3 43 3132 41.63 -0.008974213 -0.093302187 -0.202486965 4.73 282 -0.8902
1011008 2 2001 1 175 2 4 1 3 43.1 2920 41.58 -0.02036864 -0.100279829 -0.201677079 7.961 371.3 1.132
1011009 2 2001 2 176 2 2 1 3 34.5 2240 39.12 0.018447894 -0.117857625 0.094535854 1.833 -299.5 0.036
1011011 1 2001 2 177 2 3 1 3 36.4 2652 42.82 -0.17754614 0.066065662 0.03260203 -1.32 -97.97 1.361
```

Pedigree Blup (PBLUP)

Getting started: RENUMF90

- Renumbering programme
- Performs pedigree checks
- Perform data summaries
- Supports unknown parent groups

* It is convenient and reduces the chances of further error

Parameter files

```
base ~

~/bin/renumf90

 RENUMF90 version 1.162 with zlib
* START job    Wall time: 02-18-2024  16h 06m 26s 454
 name of parameter file?
```

Parameter files

- Datafile
- Pedigree/SNP file
- What are the covariates
- What are the random effects
- Variances
- …

All the information that you used to build and solve the mixed-models so far

RENUMF90 parameter file

```
base ~ (0.055s)

~/bin/renumf90 --show-template
```

RENUMF90 parameter file

```
base ~ (0.045s)

~/bin/renumf90 --show-template

# parameter file for renumf90
DATAFILE

TRAITS

FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE

EFFECT

#RANDOM
#
#OPTIONAL
#
#FILE
#
#FILE_POS
#
#SNP_FILE
#
```

RENUMF90 parameter file

- Save the template by piping the output of the renumf90 command to a file:

```
base ~

~/bin/renumf90 --show-template > blupf901.par
```

RENUMF90 parameter file

- Save the template by piping the output of the renumf90 command to a file:

```
base ~

~/bin/renumf90 --show-template > blupf901.par
```

- We will use a text editor to adapt that template to our dataset

RENUMF90 parameter file

- Model:

FIXED

y = sex + year + litter_size + dam_age + group + line +

dob + litter + animal + e

covariate

RANDOM

RENUMF90 parameter file: fixed effects

```
 1  # parameter file for renumf90¬
 2  DATAFILE¬
 3  sheep.dat¬
 4  TRAITS¬
 5  10 # trait of interest: LW¬
 6  FIELDS_PASSED TO OUTPUT¬
 7  ¬
 8  WEIGHT(S)¬
 9  ¬
10  RESIDUAL_VARIANCE¬
11  10.00¬
12  EFFECT¬
13  #RANDOM¬
14  #OPTIONAL¬
15  #pe¬
16  #FILE¬
17  #¬
18  #PED_DEPTH¬
19  #¬
20  #INBREEDING¬
21  #¬
22  #(CO)VARIANCES¬
23  #_
```

$\mathbf{y}$ = sex + year + litter_size + dam_age + group + line + dob + litter + animal + $\mathbf{e}$

RENUMF90 parameter file: fixed effects

```
 1  # parameter file for renumf90
 2  DATAFILE
 3  sheep.dat
 4  TRAITS
 5  10 # trait of interest: LW
 6  FIELDS_PASSED TO OUTPUT
 7
 8  WEIGHT(S)
 9
10  RESIDUAL_VARIANCE
11  10.00
12  EFFECT
13  2 cross alpha # data @ col 2: sex
14  #RANDOM
15  #OPTIONAL
16  #pe
17  #FILE
18  #
19  #PED_DEPTH
20  #
21  #INBREEDING
22  #
23  #(CO)VARIANCES
24  #
25  #OPTION use_yams
26  #OPTION origID
```

y = **sex** + year + litter_size + dam_age
+ group + line + dob +
litter + animal + e

RENUMF90 parameter file: fixed effects

```
 1  # parameter file for renumf90
 2  DATAFILE
 3  sheep.dat
 4  TRAITS
 5  10 # trait of interest: LW
 6  FIELDS_PASSED TO OUTPUT
 7
 8  WEIGHT(S)
 9
10  RESIDUAL_VARIANCE
11  10.00
12  EFFECT
13  2 cross alpha # data @ col 2: sex
14  EFFECT
15  3 cross alpha # data @ col 3: Year
16  #RANDOM
17  #OPTIONAL
18  #pe
19  #FILE
20  #
21  #PED_DEPTH
22  #
23  #INBREEDING
24  #
25  #(CO)VARIANCES
26  #
```

$$y = sex + \textbf{year} + litter\_size + dam\_age$$
$$+ \ group + line + dob +$$
$$litter + animal + e$$

RENUMF90 parameter file: fixed effects

```
1   # parameter file for renumf90¬
2   DATAFILE¬
3   sheep.dat¬
4   TRAITS¬
5   10 # trait of interest: LW¬
6   FIELDS_PASSED TO OUTPUT¬
7   ¬
8   WEIGHT(S)¬
9   ¬
10  RESIDUAL_VARIANCE¬
11  10.00¬
12  EFFECT¬
13  2 cross alpha # data @ col 2: sex¬
14  EFFECT¬
15  3 cross alpha # data @ col 3: Year¬
16  EFFECT¬
17  5 cross alpha # data @ col 5: Litter size¬
18  #RANDOM¬
19  #OPTIONAL¬
20  #pe¬
21  #FILE¬
22  #¬
23  #PED_DEPTH¬
24  #¬
25  #INBREEDING¬
26  #¬
```

y = sex + year + **litter_size** + dam_age + group + line + dob + litter + animal + e

RENUMF90 parameter file: fixed effects

```
1   # parameter file for renumf90¬
2   DATAFILE¬
3   sheep.dat¬
4   TRAITS¬
5   13 # trait of interest: LW ¬
6   FIELDS_PASSED TO OUTPUT¬
7    ¬
8   WEIGHT(S)¬
9    ¬
10  RESIDUAL_VARIANCE¬
11  10.00¬
12  EFFECT¬
13  2 cross alpha # data @ col 2: sex¬
14  EFFECT¬
15  3 corss alpha # data @ col 3: Year¬
16  EFFECT¬
17  5 cross alpha # data @ col 5: Litter sise¬
18  EFFECT¬
19  7 cross alpha # data @ col 7: dam age¬
20  #RANDOM¬
21  animal¬
22  #OPTIONAL¬
23  #¬
24  #FILE¬
25  #¬
26  #FILE POS¬
```

$$y = \text{sex} + \text{year} + \text{litter\_size} + \mathbf{dam\_age} + \text{group} + \text{line} + \text{dob} + \text{litter} + \text{animal} + e$$

RENUMF90 parameter file: fixed effects

```
1   # parameter file for renumf90
2   DATAFILE
3   sheep.dat
4   TRAITS
5   10 # trait of interest: LW
6   FIELDS_PASSED TO OUTPUT
7
8   WEIGHT(S)
9
10  RESIDUAL_VARIANCE
11  10.00
12  EFFECT
13  2 cross alpha # data @ col 2: sex
14  EFFECT
15  3 cross alpha # data @ col 3: Year
16  EFFECT
17  5 cross alpha # data @ col 5: Litter size
18  EFFECT
19  7 cross alpha # data @ col 7: dam age
20  EFFECT
21  8 cross alpha # data @ col 8: Group
22  #RANDOM
23  #OPTIONAL
24  #pe
25  #FILE
26  #
```

$$y = sex + year + litter\_size + dam\_age$$
$$+ \mathbf{group} + line + dob +$$
$$litter + animal + e$$

RENUMF90 parameter file: fixed effects

```
1   # parameter file for renumf90¬
2   DATAFILE¬
3   sheep.dat¬
4   TRAITS¬
5   10 # trait of interest: LW¬
6   FIELDS_PASSED TO OUTPUT¬
7   ¬
8   WEIGHT(S)¬
9   ¬
10  RESIDUAL_VARIANCE¬
11  10.00¬
12  EFFECT¬
13  2 cross alpha # data @ col 2: sex¬
14  EFFECT¬
15  3 cross alpha # data @ col 3: Year¬
16  EFFECT¬
17  5 cross alpha # data @ col 5: Litter size¬
18  EFFECT¬
19  7 cross alpha # data @ col 7: dam age¬
20  EFFECT¬
21  8 cross alpha # data @ col 8: Group¬
22  EFFECT¬
23  9 cross alpha # data @ col 9: Line¬
24  RANDOM¬
25  #OPTIONAL¬
26  #pe¬
```

$$y = sex + year + litter\_size + dam\_age + group + \mathbf{line} + dob + litter + animal + e$$

RENUMF90 parameter file: fixed effects

```
1   # parameter file for renumf90
2   DATAFILE
3   sheep.dat
4   TRAITS
5   10 # trait of interest: LW
6   FIELDS_PASSED TO OUTPUT
7
8   WEIGHT(S)
9
10  RESIDUAL_VARIANCE
11  10.00
12  EFFECT
13  2 cross alpha # data @ col 2: sex
14  EFFECT
15  3 cross alpha # data @ col 3: Year
16  EFFECT
17  5 cross alpha # data @ col 5: Litter size
18  EFFECT
19  7 cross alpha # data @ col 7: dam age
20  EFFECT
21  8 cross alpha # data @ col 8: Group
22  EFFECT
23  9 cross alpha # data @ col 9: Line
24  EFFECT
25  4 cov # data @ col 4: day of birth
26  EFFECT
```

$$y = \text{sex} + \text{year} + \text{litter\_size} + \text{dam\_age}$$
$$+ \ \text{group} + \text{line} + \textbf{dob} +$$
$$\text{litter} + \text{animal} + e$$

RENUMF90 parameter file: random effects

```
  4  TRAITS
  5  10              # trait of interest: LW
  6  FIELDS_PASSED TO OUTPUT

  7
  8  WEIGHT(S)

  9
 10  RESIDUAL_VARIANCE
 11  10.0
 12  EFFECT
 13  2 cross alpha # data @ col 2: sex
 14  EFFECT
 15  3 cross alpha # data @ col 3: Year
 16  EFFECT
 17  5 cross alpha # data @ col 5: Litter size
 18  EFFECT
 19  7 cross alpha # data @ col 7: dam age
 20  EFFECT
 21  8 cross alpha # data @ col 8: Group
 22  EFFECT
 23  9 cross alpha  # data @ col 9: Line
 24  EFFECT
 25  4 cov          # data @ col 4: day of birth
 26  EFFECT
 27  5  cross alpha # data @ col 5: litter
 28  RANDOM
 29  diagonal
```

y = sex + year + litter_size + dam_age + group + line + **litter** + animal + e

RENUMF90 parameter file: random effects

```
21  8 cross alpha # data @ col 8: Group
22  EFFECT
23  9 cross alpha  # data @ col 9: Line
24  EFFECT
25  4 cov          # data @ col 4: day of birth
26  EFFECT
27  5  cross alpha # data @ col 5: litter
28  RANDOM
29  diagonal
30  EFFECT
31  1 cross alpha
32  RANDOM
33  animal
34  OPTIONAL
35
36  FILE
37  newped.txt # name of pedigree file
38  PED_DEPTH
39  0
40  # how far on pedigree to look: 0 = all
41  INBREEDING
42  pedigree
43  (CO)VARIANCES
44  0.1
```

$$y = sex + year + litter\_size + dam\_age + group + line + \mathbf{animal} + e$$

RENUMF90 parameter file: optional arguments

```
21   8 cross alpha # data @ col 8: Group
22   EFFECT
23   9 cross alpha  # data @ col 9: Line
24   EFFECT
25   4 cov           # data @ col 4: day of birth
26   EFFECT
27   5  cross alpha # data @ col 5: litter
28   RANDOM
29   diagonal
30   EFFECT
31   1 cross alpha
32   RANDOM
33   animal
34   OPTIONAL
35
36   FILE
37   newped.txt # name of pedigree file
38   PED_DEPTH
39   0
40   # how far on pedigree to look: 0 = all
41   INBREEDING
42   pedigree
43   (CO)VARIANCES
44   0.1
45   OPTION use_yams
46   OPTION origID
```

- **OPTION use_yams**

Uses the programme YAMS.

Improves computational time.

OPTION origID

To keep original ids
(it makes it easier when retrieving results)

Running RENUMF90

- Renumf90 <parameter_file>

```
base ~/Documents/2024/ILRI_training/pedigree_eval (0.865s)

~/bin/renumf90 pblup_renum.par
 RENUMF90 version 1.162 with zlib
* START job    Wall time: 02-20-2024  09h 14m 01s 856
 pblup_renum.par
 datafile:sheep.dat
 traits:           10
 number of weights=              1
 in columns:           0 (0 means no weight)
 R
   10.00

 Processing effect  1 of type cross
 item_kind=alpha

 Processing effect  2 of type cross
 item_kind=alpha

 Processing effect  3 of type cross
 item_kind=alpha

 Processing effect  4 of type cross
 item_kind=alpha

 Processing effect  5 of type cross
 item_kind=alpha

 Processing effect  6 of type cross
 item_kind=alpha
```

Running BLUPF90

- Renumf90 will generate the parameter file for BLUPF90 called **renf90.par**

- We can run blupf90
  as we done with renumf90:
  blupf90+ renf90.par

- All files generated by both programmes have standard names and structures

```
cat renf90.par
# BLUPF90 parameter file created by RENUMF90
DATAFILE
 renf90.dat
NUMBER_OF_TRAITS
          1
NUMBER_OF_EFFECTS
          9
OBSERVATION(S)
    1
WEIGHT(S)

EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT[EFFECT NESTED]
 2        2 cross
 3        3 cross
 4      397 cross
 5        4 cross
 6        2 cross
 7        5 cross
 8 1 cov
 9      397 cross
 10     4968 cross
RANDOM_RESIDUAL VALUES
   10.000
 RANDOM_GROUP
    8
 RANDOM_TYPE
 diagonal
```

Understanding the solutions

- R again!!!