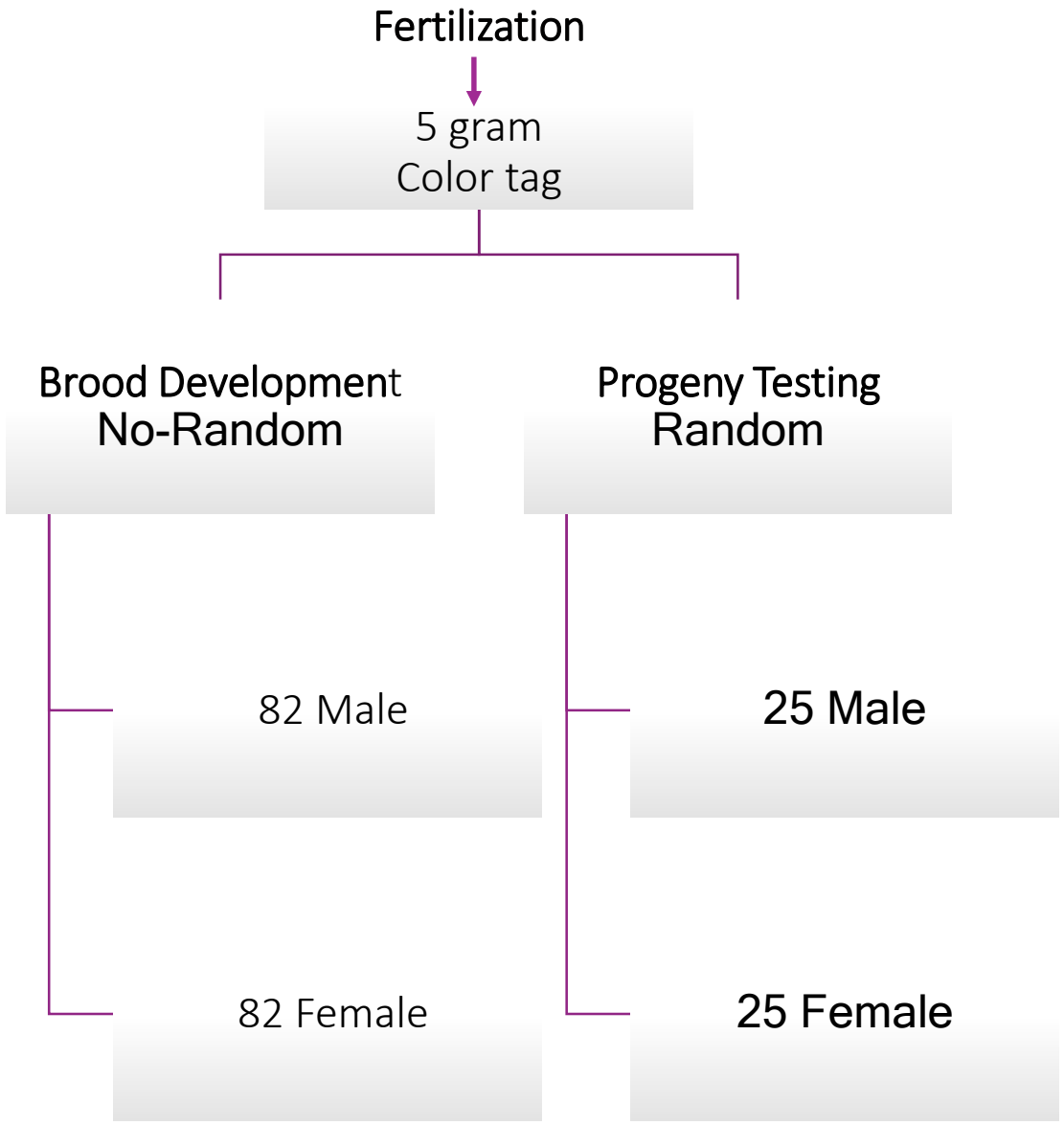
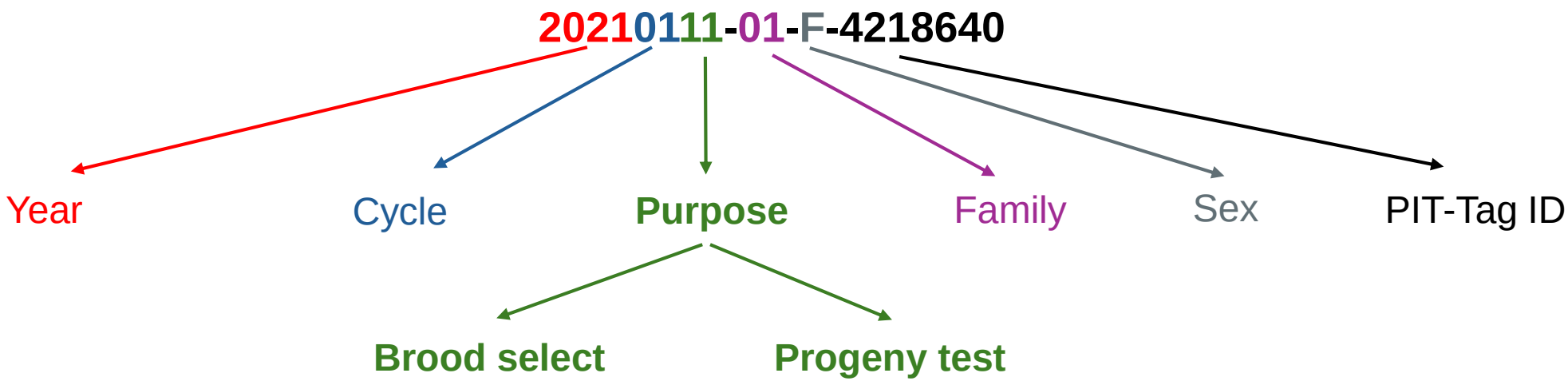
Overview of BLUPF90 Models for Genetic Evaluation

DATA COLECTION



ID_RECORD	CG	FAM_ID	BD_SEX	BD_AGE1	BD_Wt1	BD_AGE2	BD_Wt2	BD_AGE3	BD_Wt3	PT_SEX	PT_AGE	PT_Wt
20210111-01-F-4218640	20210111001	2101-01	F	213	43	283	66	332	80	-9999	-9999	-9999
20210111-01-F-4230091	20210111001	2101-01	F	213	37	283	57.5	332	67	-9999	-9999	-9999
20210111-01-F-4231261	20210111001	2101-01	F	213	49	283	74	-9999	-9999	-9999	-9999	-9999
20210111-01-F-4218650	20210111001	2101-01	F	213	42.5	283	62.5	332	75	-9999	-9999	-9999
20210111-01-F-4218765	20210111001	2101-01	F	213	48.5	283	83	332	97.5	-9999	-9999	-9999
20210111-01-F-4216907	20210111001	2101-01	F	213	41.5	283	66	-9999	-9999	-9999	-9999	-9999
20210111-01-F-4231704	20210111001	2101-01	F	213	47.5	283	67	332	78	-9999	-9999	-9999
20210111-01-F-4216539	20210111001	2101-01	F	213	45.5	283	68.5	332	79	-9999	-9999	-9999
20210111-01-F-4217079	20210111001	2101-01	F	213	42.5	283	68	332	80	-9999	-9999	-9999
20210111-01-F-4213121	20210111001	2101-01	F	213	44.5	283	63.5	332	68	-9999	-9999	-9999
20210111-01-F-4218615	20210111001	2101-01	F	213	42.5	283	67	332	72	-9999	-9999	-9999
20210111-01-F-4231675	20210111001	2101-01	F	213	50	283	73	332	85.5	-9999	-9999	-9999
20210111-01-F-4218606	20210111001	2101-01	F	213	42.5	283	75	332	89.5	-9999	-9999	-9999



Contemporary Group (CG) :

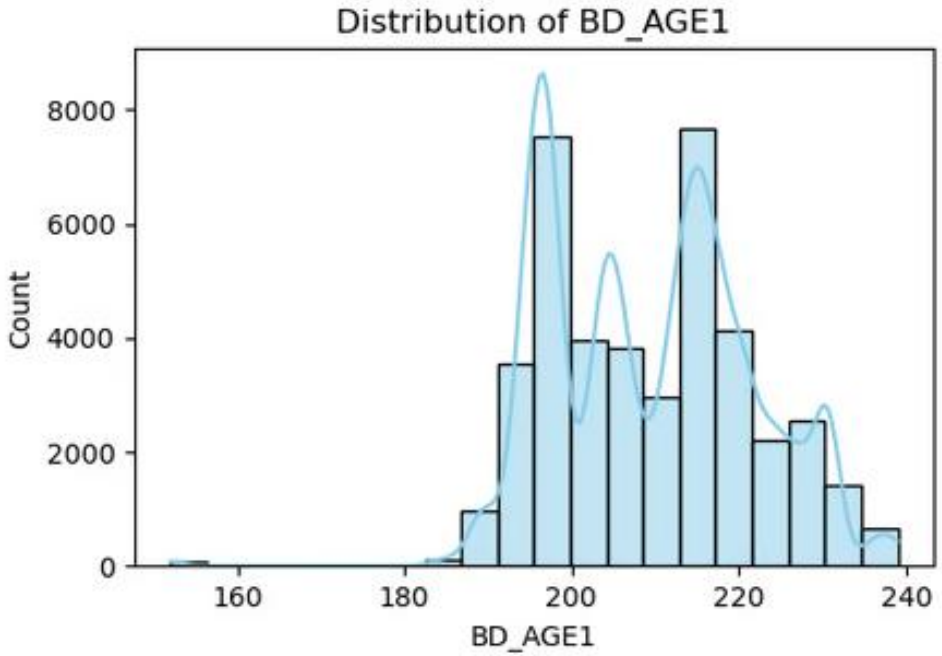
- Là nhóm cá thể cùng điều kiện môi trường nuôi (ví dụ: cùng ao nuôi, cùng trang trại, ….).
- Mục đích: Loại bỏ biến thiên không di truyền, tăng độ chính xác của EBV.

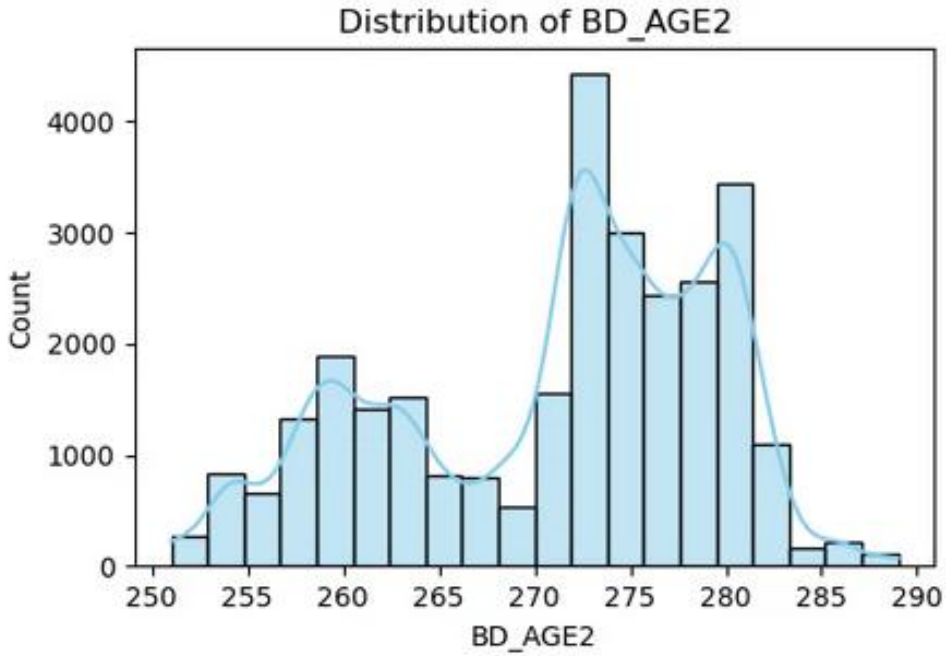
FAM_ID: Permanent Environment (PE)

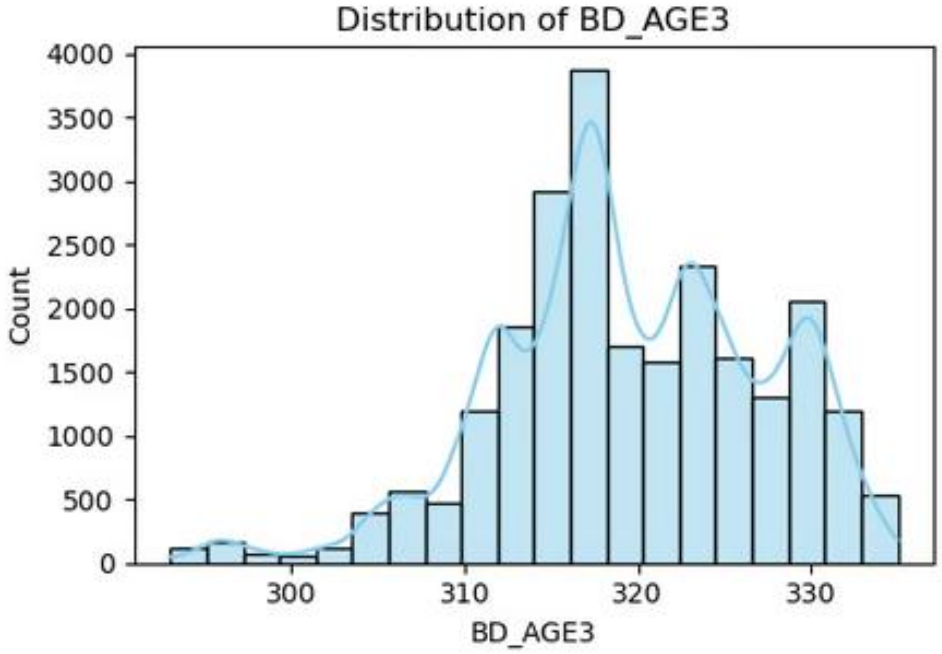
- Phản ánh ảnh hưởng môi trường vĩnh viễn lên cá thể (ví dụ: cùng bố mẹ).
- Tách biệt với di truyền, giảm sai lệch trong ước lượng EBV.

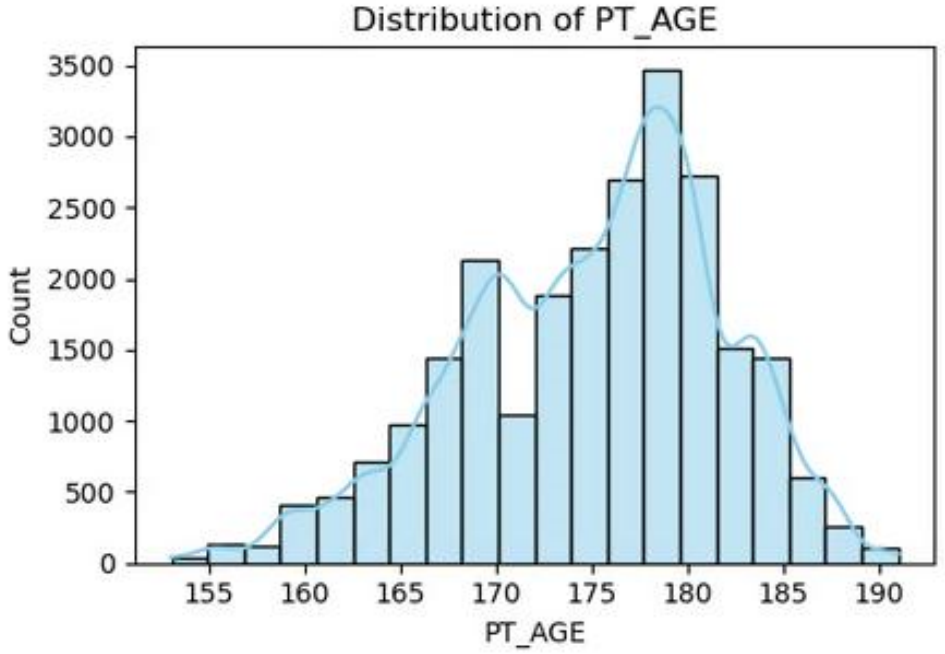
AGE = FERT – MEASURE DAY

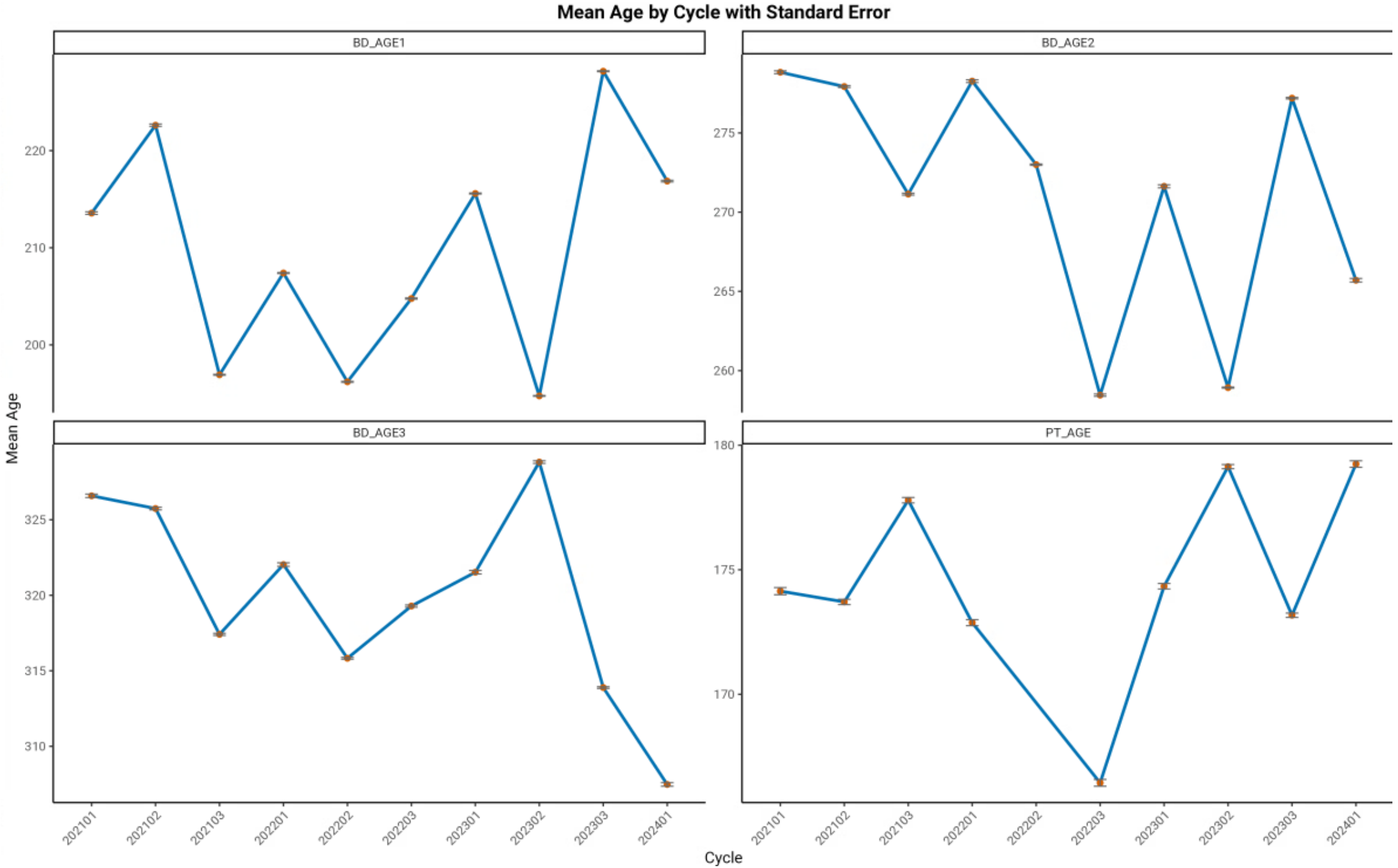
Overview of the input data



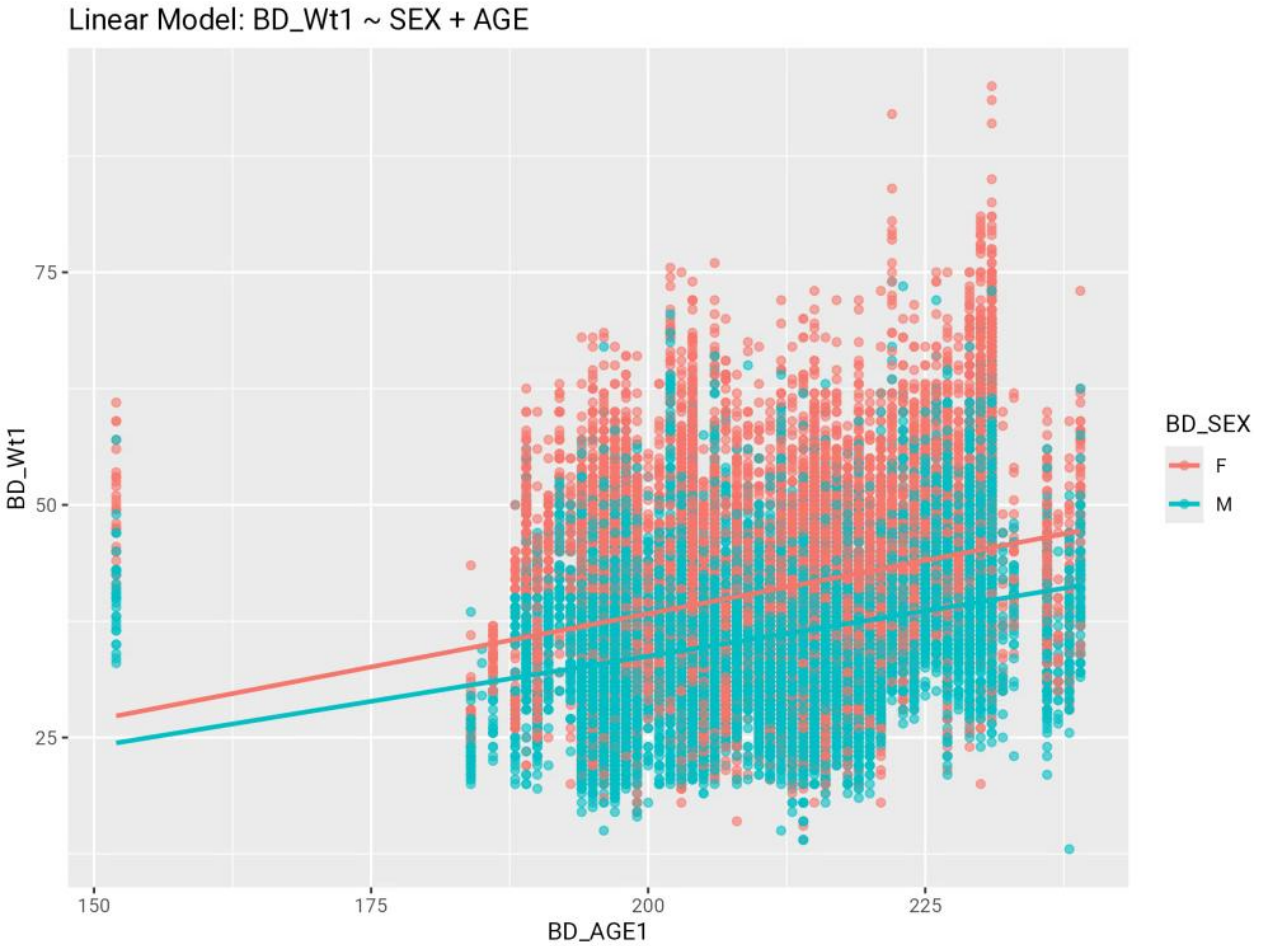


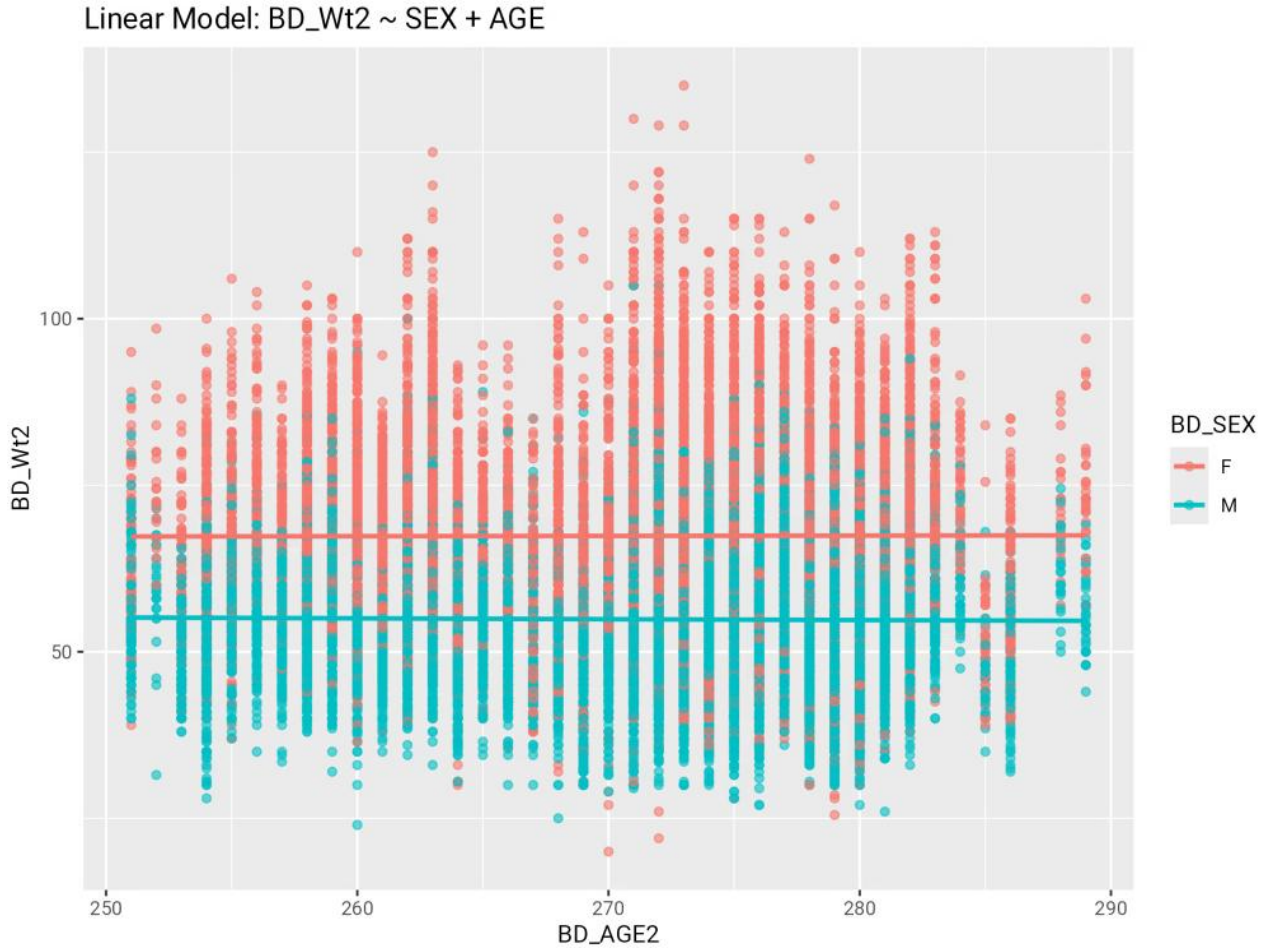


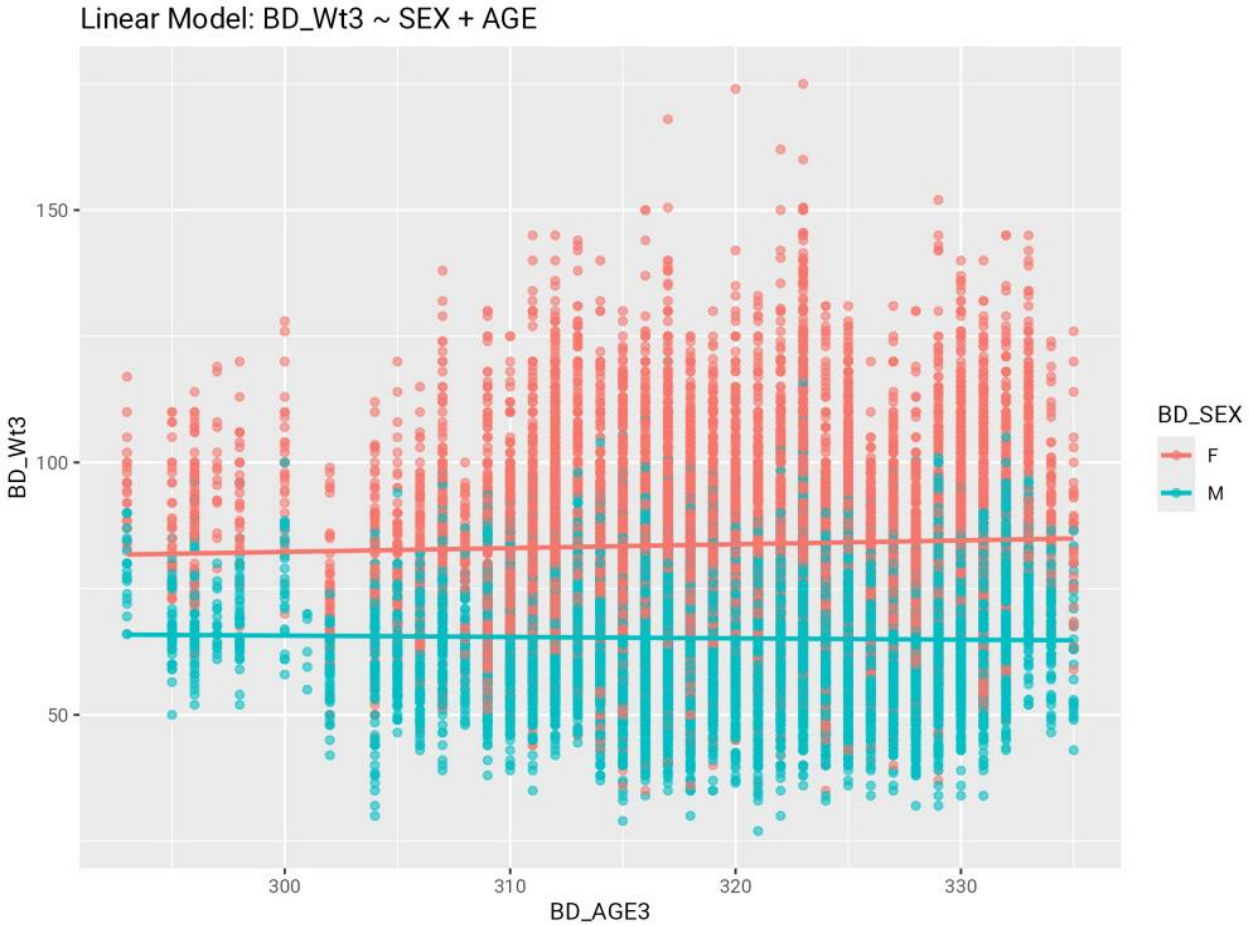


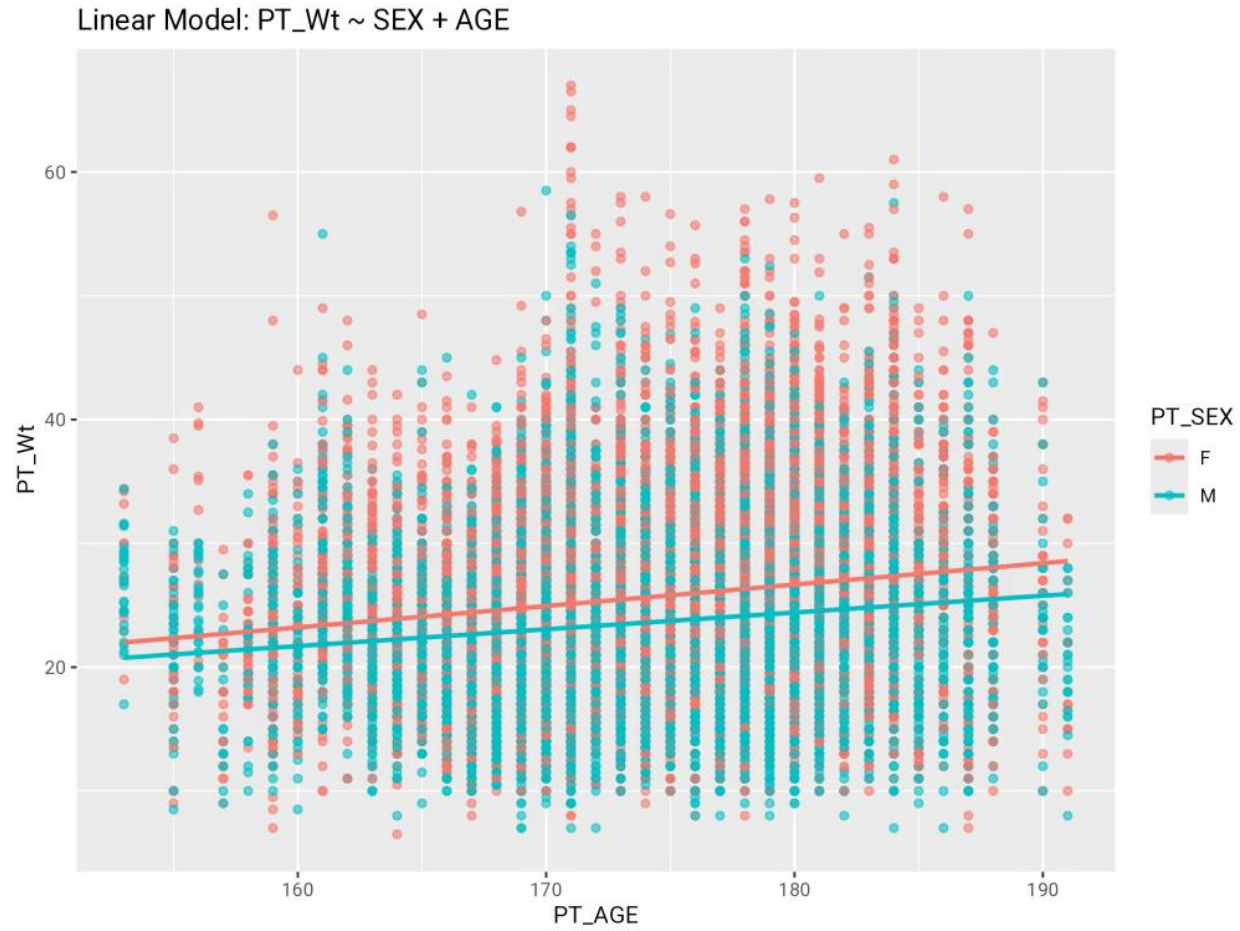

Mean Age by Cycle with Standard Error
BD_AGE1
BD_AGE2
BD_AGE3
PT_AGE
Mean Age
Cycle
220
210
200
275
270
265
260
325
320
315
310
180
175
170
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401

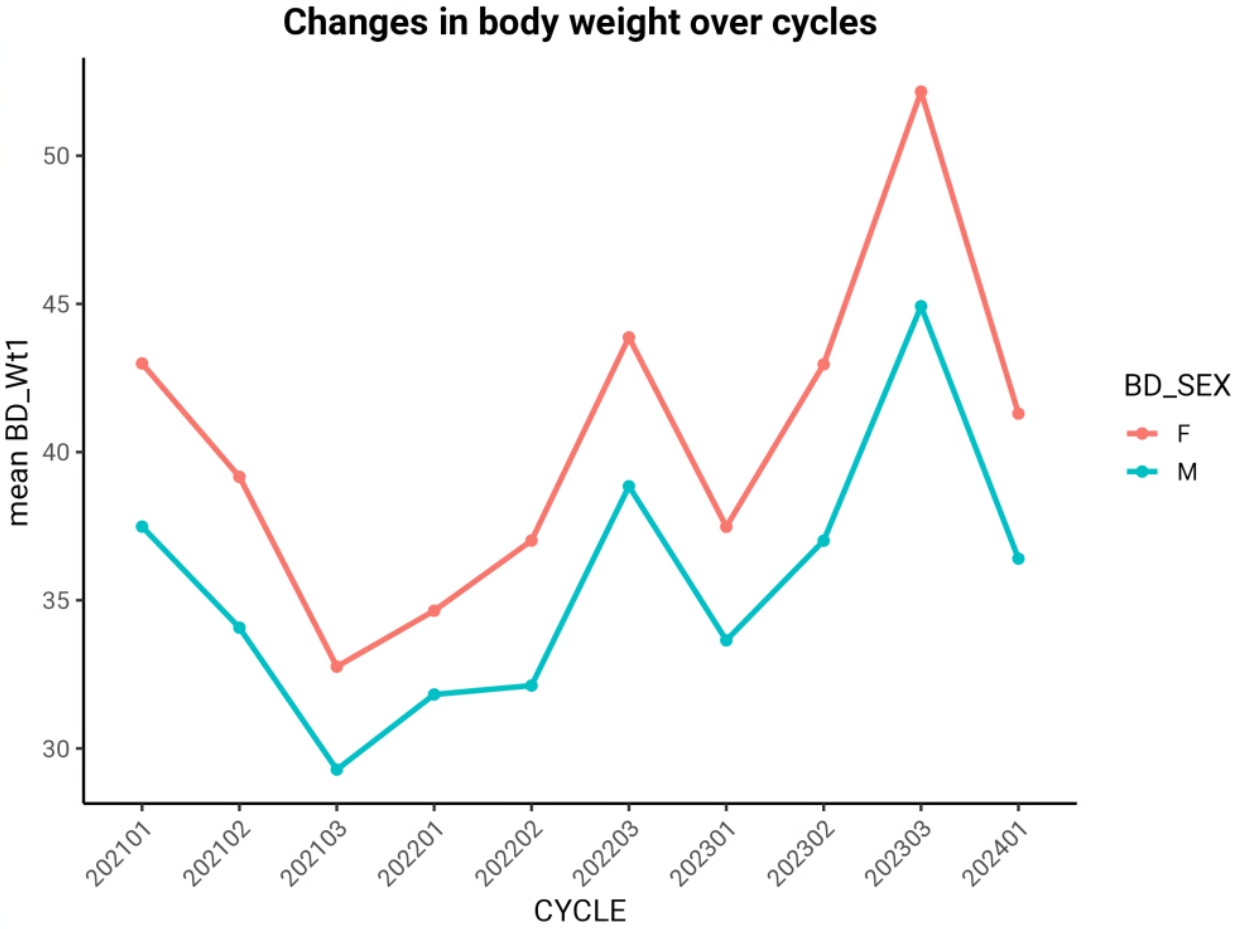
Weight distribution by age and sex

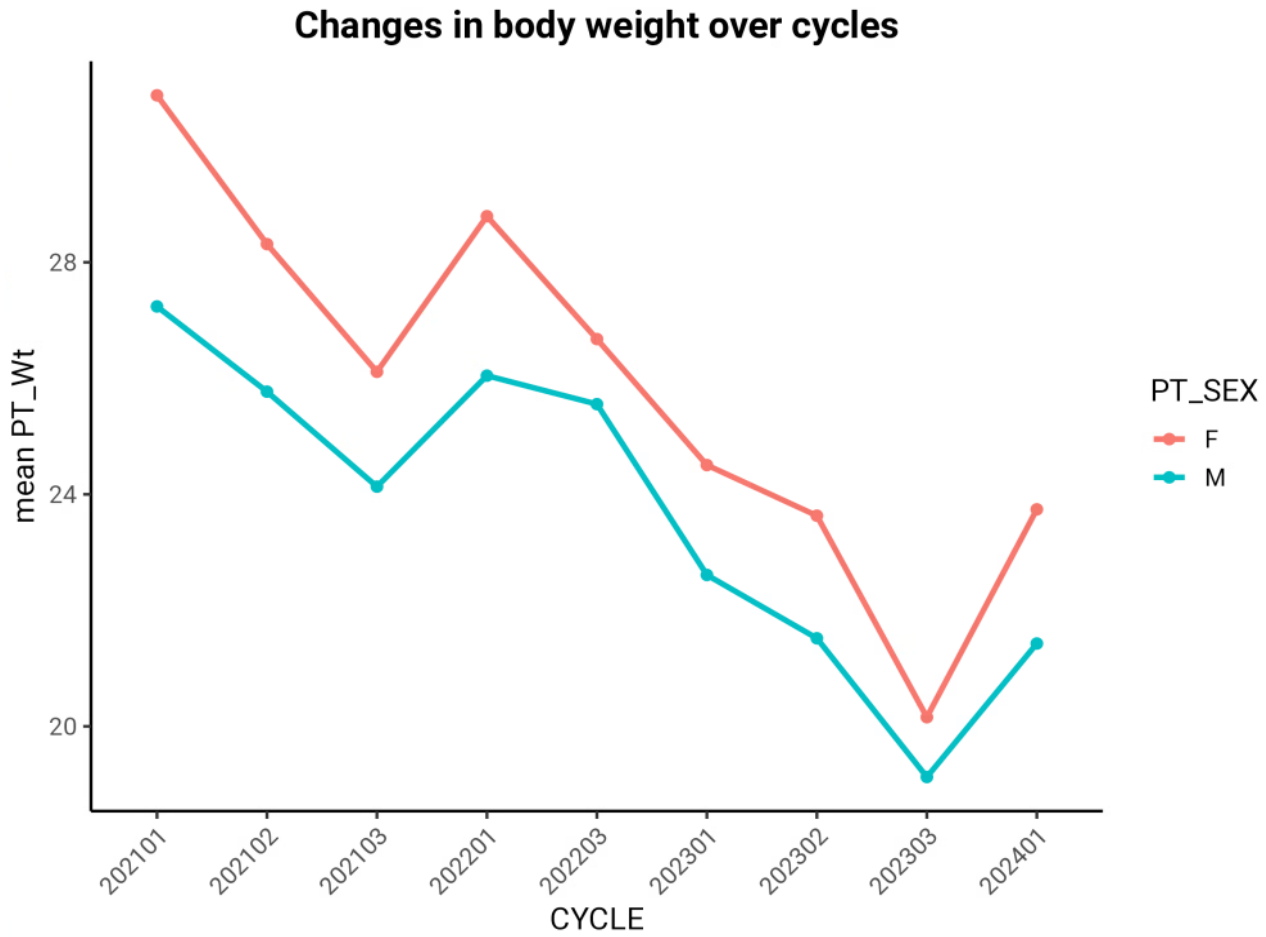


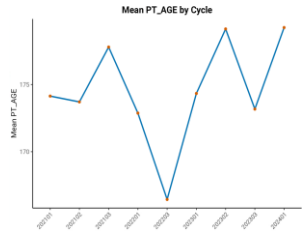


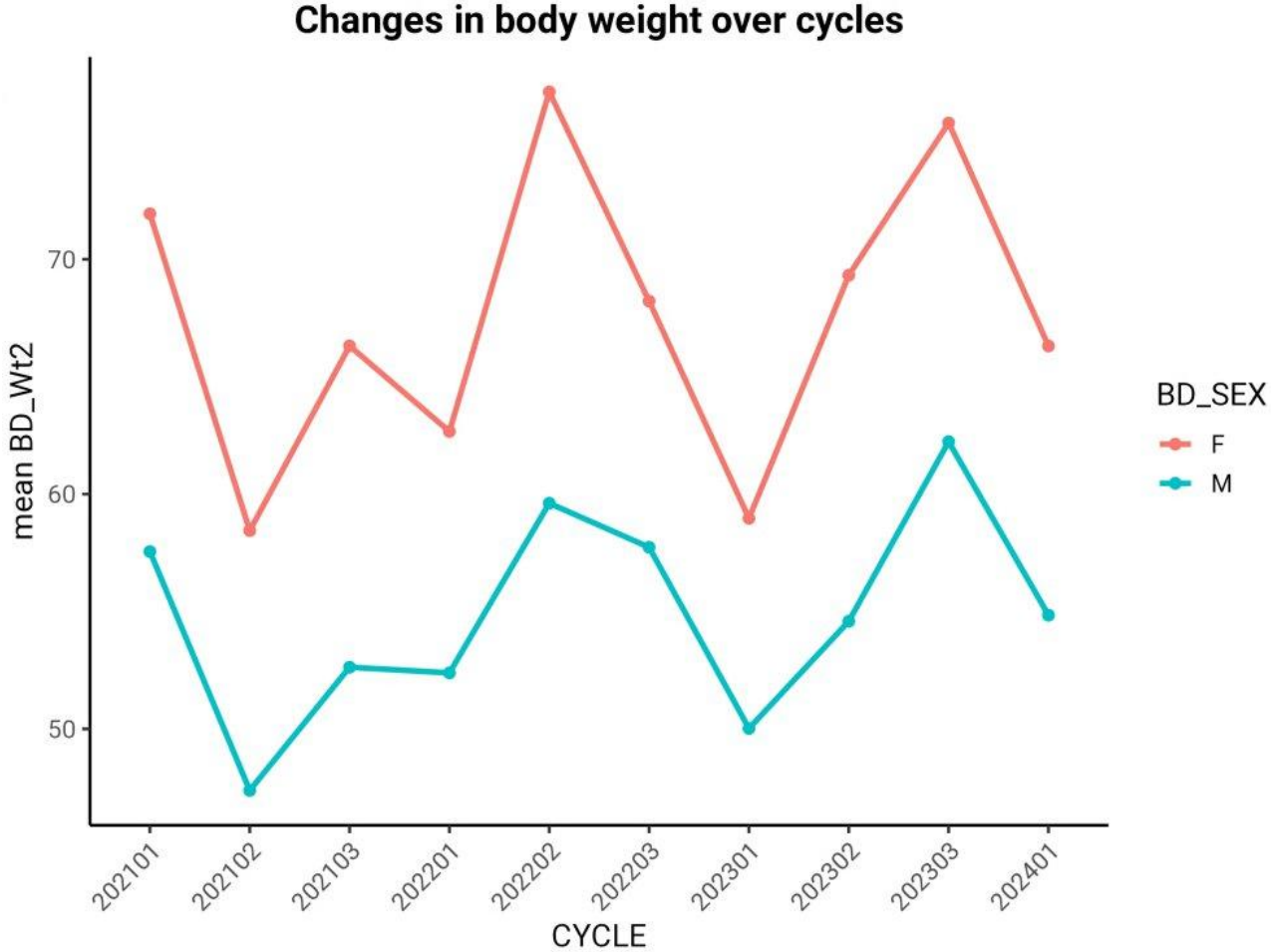

Linear Model: BD_Wt3 ~ SEX + AGE
BD_Wt3
150
100
50
300
310
320
330
BD_AGE3
BD_SEX
F
M

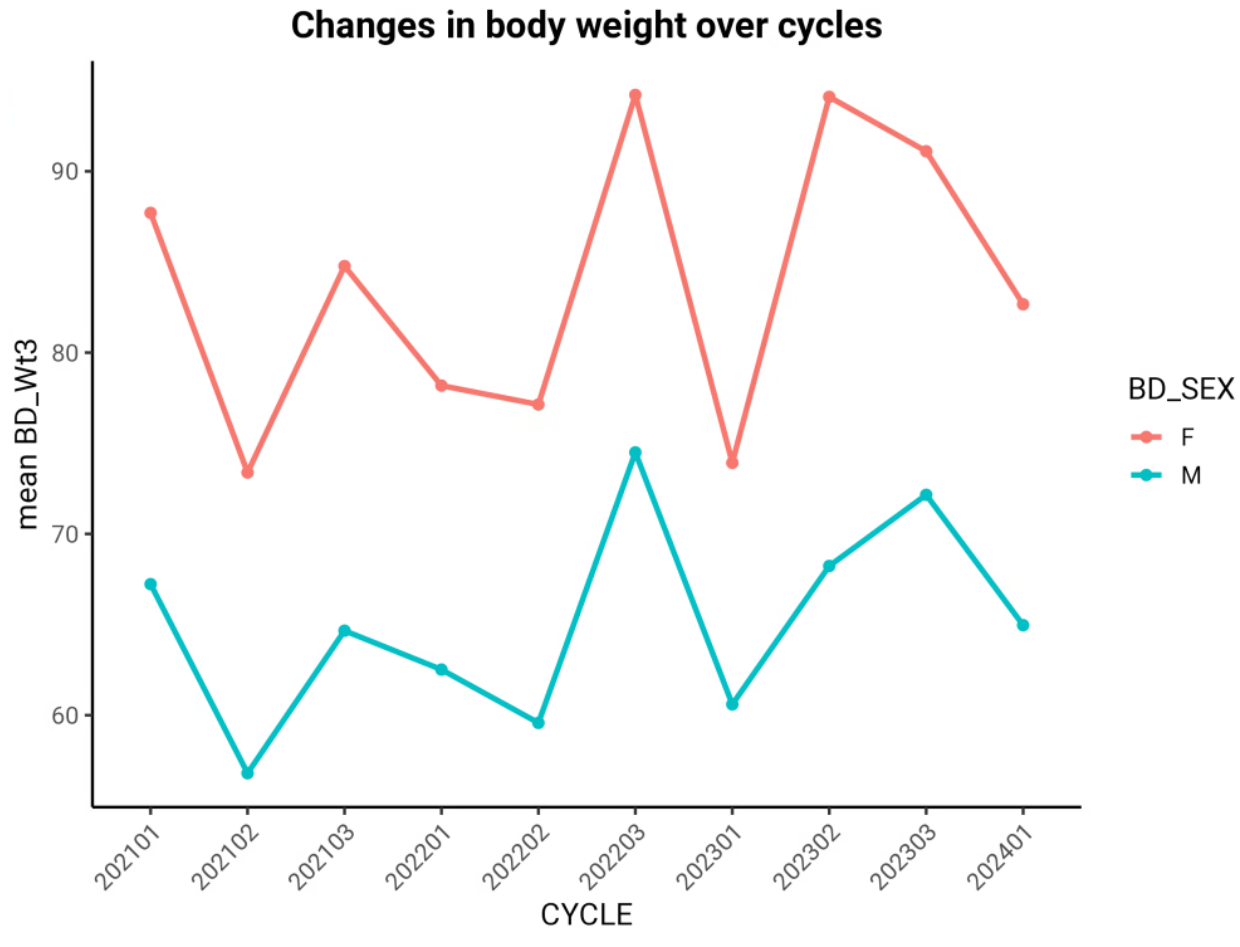

Linear Model: PT_Wt ~ SEX + AGE
PT_Wt
60
40
20
160
170
180
190
PT_AGE
PT_SEX
F
M


Changes in body weight over cycles
mean BD_Wt1
50
45
40
35
30
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
CYCLE
BD_SEX
F
M

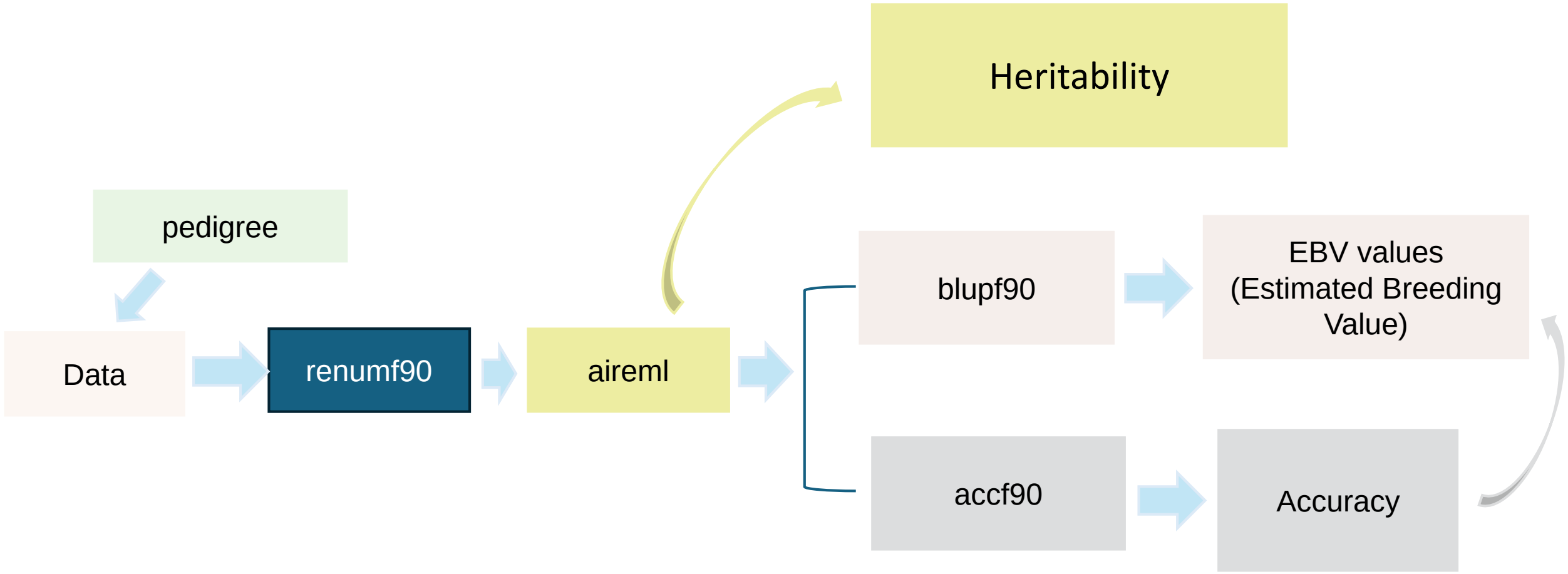

Changes in body weight over cycles
mean PT_Wt
28
24
20
202101
202102
202103
202201
202203
202301
202302
202303
202401
CYCLE
PT_SEX
F
M


Mean PT_AGE by Cycle


Changes in body weight over cycles
mean BD_Wt2
70
60
50
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
CYCLE
BD_SEX
F
M

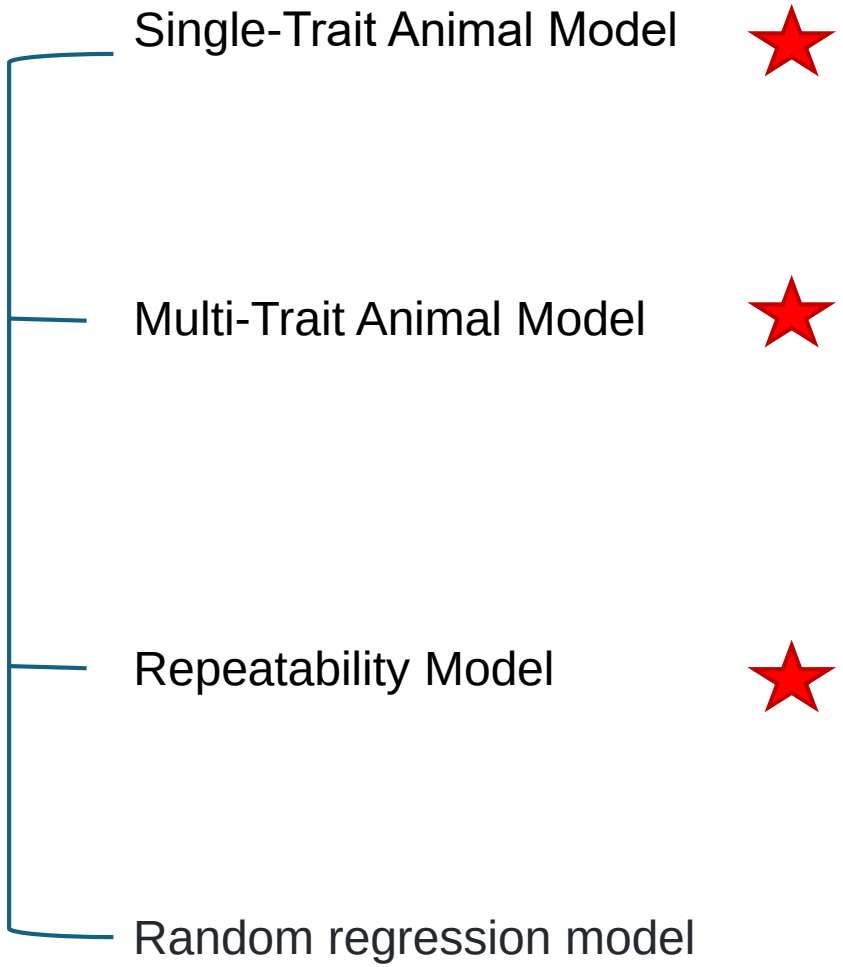

Changes in body weight over cycles
mean BD_Wt3
90
80
70
60
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
CYCLE
BD_SEX
F
M

Workflow of animal breeding program



BLUPF90

blupf90



renumf90

Chuẩn hóa, kết hợp data và pedigree để chạy các tính toán tiếp theo

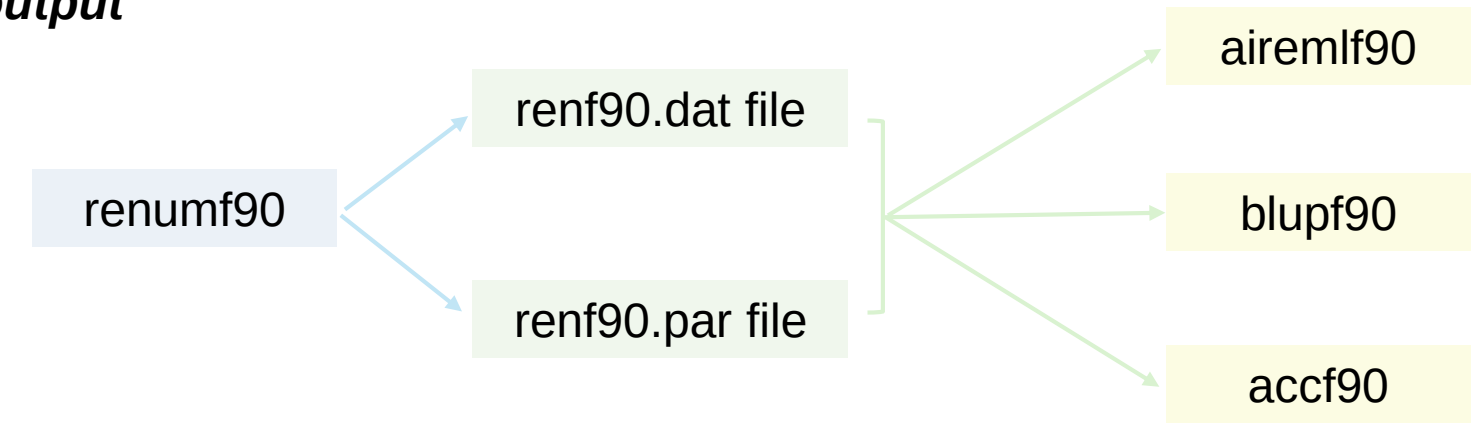
Fix effect
không thay đổi theo cá thể

- Contemporary Group
- SEX
- AGE

Random effect
Hiệu ứng ngẫu nhiên có phân bố xác suất

- Animal effect
- Residual
- Permanent Environment (yếu tố gia đình)

renumf90 output



aireml

VCE: Variance Component Estimation

Method: Restricted Maximum Likelihood (REML)

$$y = X\beta + Z\alpha + e$$

y: Vector of observations

X: Design matrix of fixed effects including sex, age and CG,

β: Vector of fixed effects;

Z: Incidence matrix of animal effects,

α: Random additive genetic effects

e: Random residual.

Animal Effect

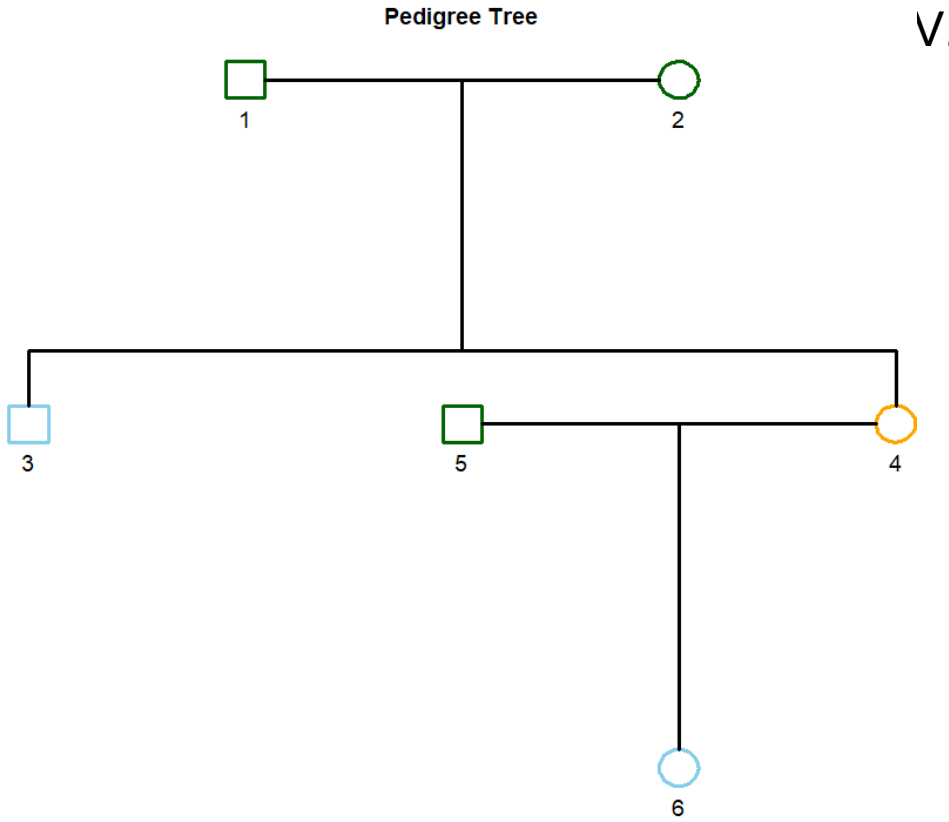
A random effect in the mixed model representing the additive genetic value (breeding value) of each individual animal.

Using the A matrix (*pedigr*

V.

```
ped

  id dadid momid sex
1  1    NA    NA   1
2  2    NA    NA   2
3  3     1     2   1
4  4     1     2   2
5  5    NA    NA   1
6  6     5     4   2
```




```
A matrix

     1    2    3    4    5    6
1 1.00 0.00 0.50 0.50 0.00 0.25
2 0.00 1.00 0.50 0.50 0.00 0.25
3 0.50 0.50 1.00 0.50 0.00 0.25
4 0.50 0.50 0.50 1.00 0.00 0.50
5 0.00 0.00 0.00 0.00 1.00 0.50
6 0.25 0.25 0.25 0.50 0.50 1.00
```

Additive Genetic Effect

The portion of genetic variation that is due to the additive effects of genes, passed from parents to offspring.

Heritability

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2}$$

σ_A^2 genetic variance.

σ_E^2 residual variance

$\sigma_A^2 + \sigma_E^2$ phenotypic variance

```
Final Estimates
Genetic variance(s) for effect 4
38.538
Residual variance(s)
62.691
inverse of AI matrix (Sampling
Variance)
13.471 -6.7551
-6.7551 6.5499
Correlations from inverse of AI
matrix
1.0000 -0.71914
-0.71914 1.0000
SE for G
3.6703
SE for R
2.5593
```

Heritability = 0,38

Results of VU data

	BD_Wt1	BD_Wt2	BD_Wt3	PT_Wt
single trait	0.4	0.43	0.42	0.1

	BD_Wt1 and PT_Wt	BD_Wt2 and PT_Wt	BD_Wt3 and PT_Wt
2_traits	0.336 0.99987 0.155	0.466 0.99882 0.16	0.45 0.96201 0.15

	BD_Wt1 and BD_Wt2	BD_Wt2 and BD_Wt3
2_traits	0.75 0.99998 0.738	0.88 0.9999 0.9

	BD_Wt1; BD_Wt2 and BD_Wt3		
3_traits	0.748	0.99978	0.97325
	0.99978	0.74	0.96858
	0.97325	0.96858	0.185

Note:

BD_Wt1 **PT_Wt**

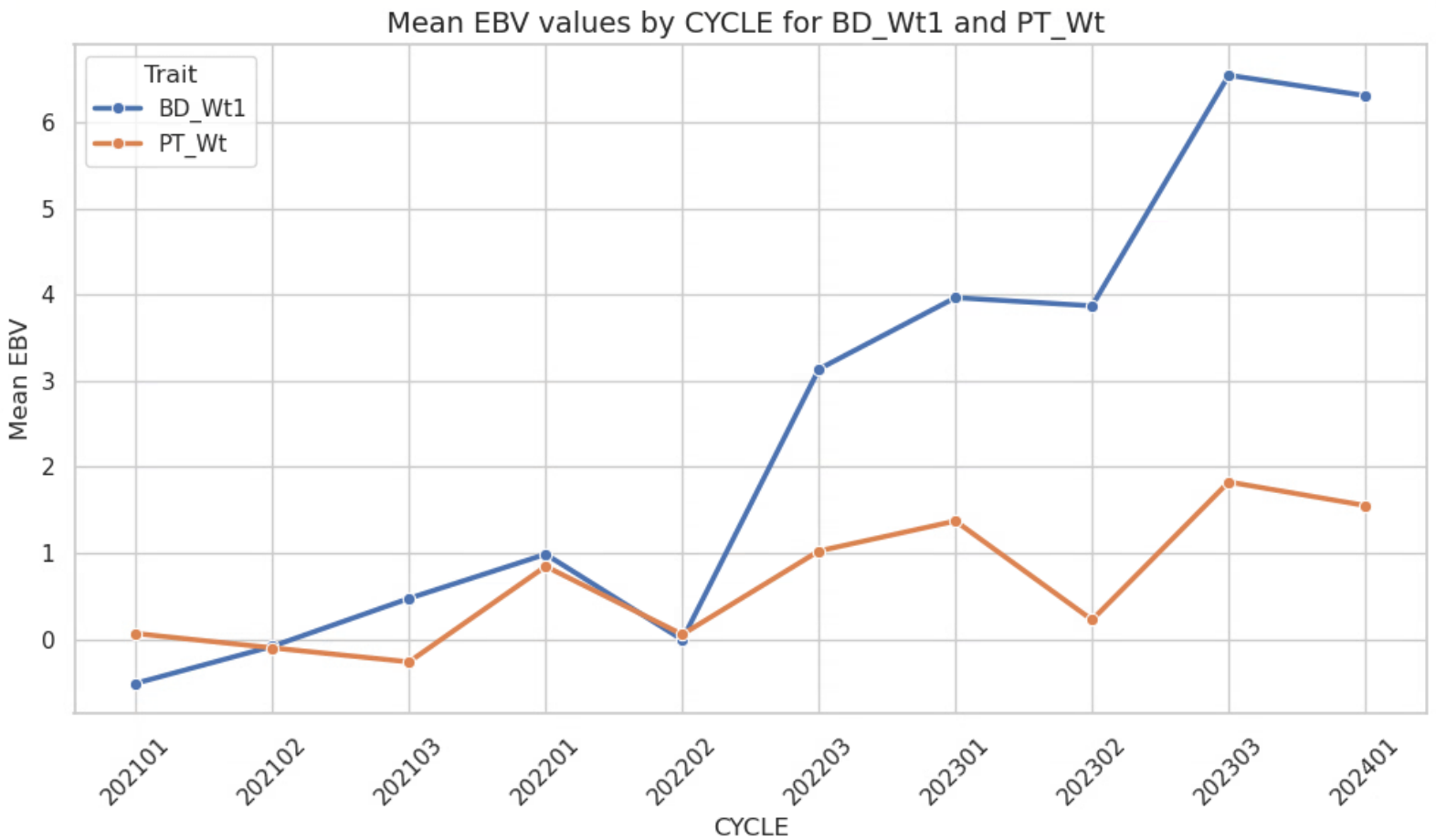
0.336

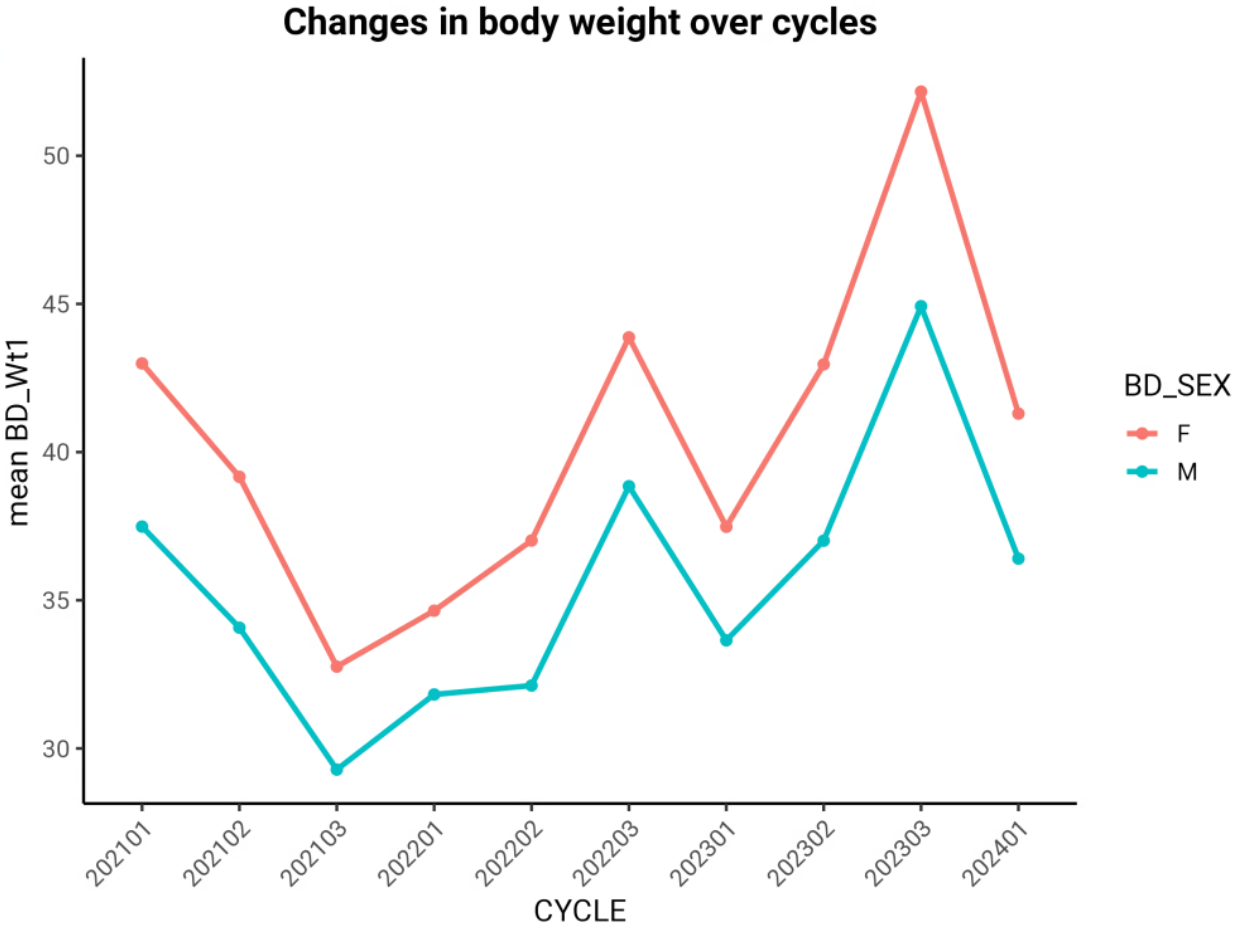
0.99987 _0.155

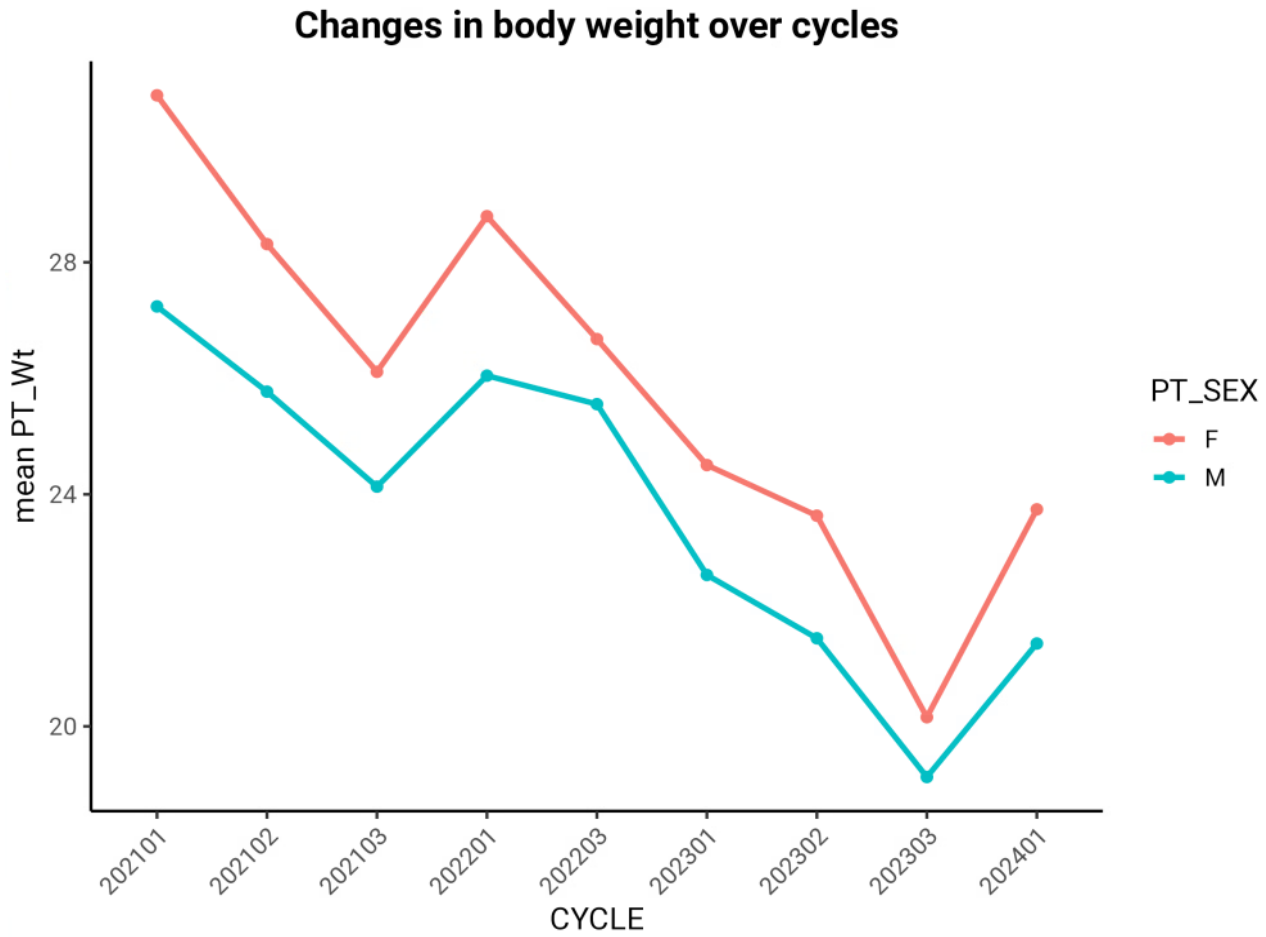
↓ ↓

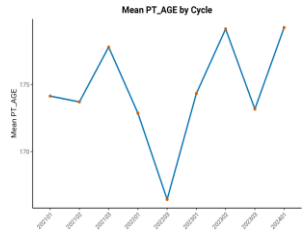
correlations heritability

EBV value of VU'data




Changes in body weight over cycles
mean BD_Wt1
50
45
40
35
30
202101
202102
202103
202201
202202
202203
202301
202302
202303
202401
CYCLE
BD_SEX
F
M


Changes in body weight over cycles
mean PT_Wt
28
24
20
202101
202102
202103
202201
202203
202301
202302
202303
202401
CYCLE
PT_SEX
F
M


Mean PT_AGE by Cycle

accf90: Accuracy of Estimated Breeding Values (EBV)

The correlation between true breeding value (TBV) and estimated breeding value (EBV).

$$Accuracy = \sqrt{1 - \frac{PEV}{\sigma_\alpha^2}}$$

PEV: prediction error variance = square(s.e)

σ_α^2 *additive genetic variance*

Ranges from *0 to 1* (higher is better)

trait/effect	level	solution	acc
1	5	16774	1.04996848	0.6465
2	5	16774	-0.32818595	0.4629
1	5	16775	8.18234062	0.9444
2	5	16775	1.27146697	0.7643
1	5	16776	3.62566495	0.6467
2	5	16776	2.16748476	0.4642
1	5	16777	9.98571205	0.6500
2	5	16777	2.03272080	0.4800
1	5	16778	-3.56862426	0.6453
2	5	16778	0.55143082	0.4625
1	5	16779	-3.50841045	0.6452
2	5	16779	2.68363714	0.4608
1	5	16780	8.08479786	0.6479
2	5	16780	1.89306855	0.4709
1	5	16781	0.67373127	0.6358

Conclusion

The chosen model: **Two traits** for BD_Wt1 and PT_Wt

Heritability (h^2):

- BD_Wt1: **0.33**
- PT_Wt: **0.15**

No significant differences observed **compared to another models**

Reasons for selection:

- **High genetic correlations** between traits – *correlation ≈ 1*
- **Cost-effective**: earlier selection reduces **feeding and maintenance costs**

Thank you
for listening

```
# Parameter file for program renf90; it is translated to parameter
# file for BLUPF90 family f programs.
DATAFILE
/data/KhanLe_Data/Monodon/DataFromCan/CSIRO_PT/renum6/growth_data_1.dat
TRAITS
6
FIELDS_PASSED TO OUTPUT

WEIGHT(S)

RESIDUAL_VARIANCE
   5.987
EFFECT
2 cross alpha #CG
EFFECT
4 cross alpha #Sex
EFFECT
5 cov #Age
EFFECT
3 cross alpha #PE
RANDOM
diagonal
(CO)VARIANCES
   8.8881
EFFECT
1 cross alpha
RANDOM
animal
#OPTIONAL
#pe
FILE
/data/KhanLe_Data/Monodon/DataFromCan/CSIRO_PT/renum6/pedigree.dat
FILE_POS
1 2 3
#SNP_FILE
#geno_ready.txt
PED_DEPTH
5
```

1	2	3	4	5	6	7	8
ID_RECORD	CG	FAM_ID	BD_SEX	BD_AGE1	BD_Wt1	BD_AGE2	BD_W
20210111-01-F-4218640	20210111001	2101-01	F	213	43	283	
20210111-01-F-4230091	20210111001	2101-01	F	213	37	283	
20210111-01-F-4231261	20210111001	2101-01	F	213	49	283	
20210111-01-F-4218650	20210111001	2101-01	F	213	42.5	283	
20210111-01-F-4218765	20210111001	2101-01	F	213	48.5	283	
20210111-01-F-4216907	20210111001	2101-01	F	213	41.5	283	
20210111-01-F-4231704	20210111001	2101-01	F	213	47.5	283	
20210111-01-F-4216539	20210111001	2101-01	F	213	45.5	283	
20210111-01-F-4217079	20210111001	2101-01	F	213	42.5	283	
20210111-01-F-4213121	20210111001	2101-01	F	213	44.5	283	
20210111-01-F-4218615	20210111001	2101-01	F	213	42.5	283	
20210111-01-F-4231675	20210111001	2101-01	F	213	50	283	
20210111-01-F-4218606	20210111001	2101-01	F	213	42.5	283	
20210111-01-F-4218483	20210111001	2101-01	F	213	38	-9999	-9
20210111-01-F-4217163	20210111001	2101-01	F	213	44.5	283	

ID_RECORD	SIR_RECORD	DAM_RECORD
20210111-01-F-4218640	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4230091	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4231261	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4218650	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4218765	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4216907	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4231704	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4216539	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4217079	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4213121	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4218615	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4231675	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4218606	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4218483	20200111-50-M-4220397	20200111-04-F-4219125
20210111-01-F-4217163	20200111-50-M-4220397	20200111-04-F-4219125

```
-2 logL =    33423.4510659485   : AIC =   33427.4510659485
 In round              8 convergence = 3.476428650201818E-
015
 delta convergence = 1.835758207312910E-006
new R
   62.691
new G
   38.538
solutions stored in file : "solutions"

Final Estimates
Genetic variance(s) for effect 4
  38.538
Residual variance(s)
  62.691
inverse of AI matrix (Sampling Variance)
  13.471       -6.7551
 -6.7551        6.5499
Correlations from inverse of AI matrix
  1.0000     -0.71914
 -0.71914     1.0000
SE for G
  3.6703
SE for R
  2.5593
```

-2logL: Negative twice the restricted likelihood

Smaller is better: Indicates better model fit

AIC = −2logL+2p (p is the number of unique variance components)

Convergence: Indicates algorithm stability

Smaller is better

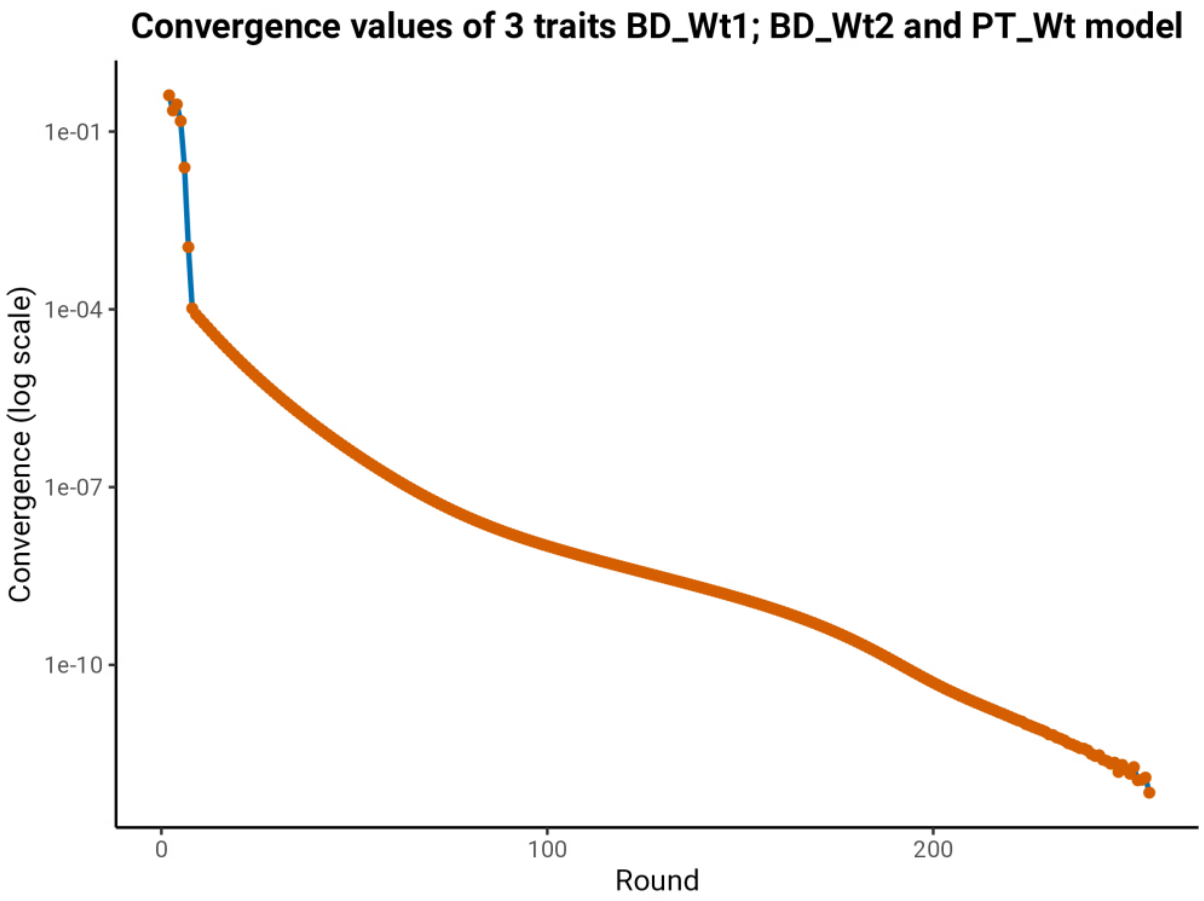
new R: Updated residual variance (Reflects environmental variation)

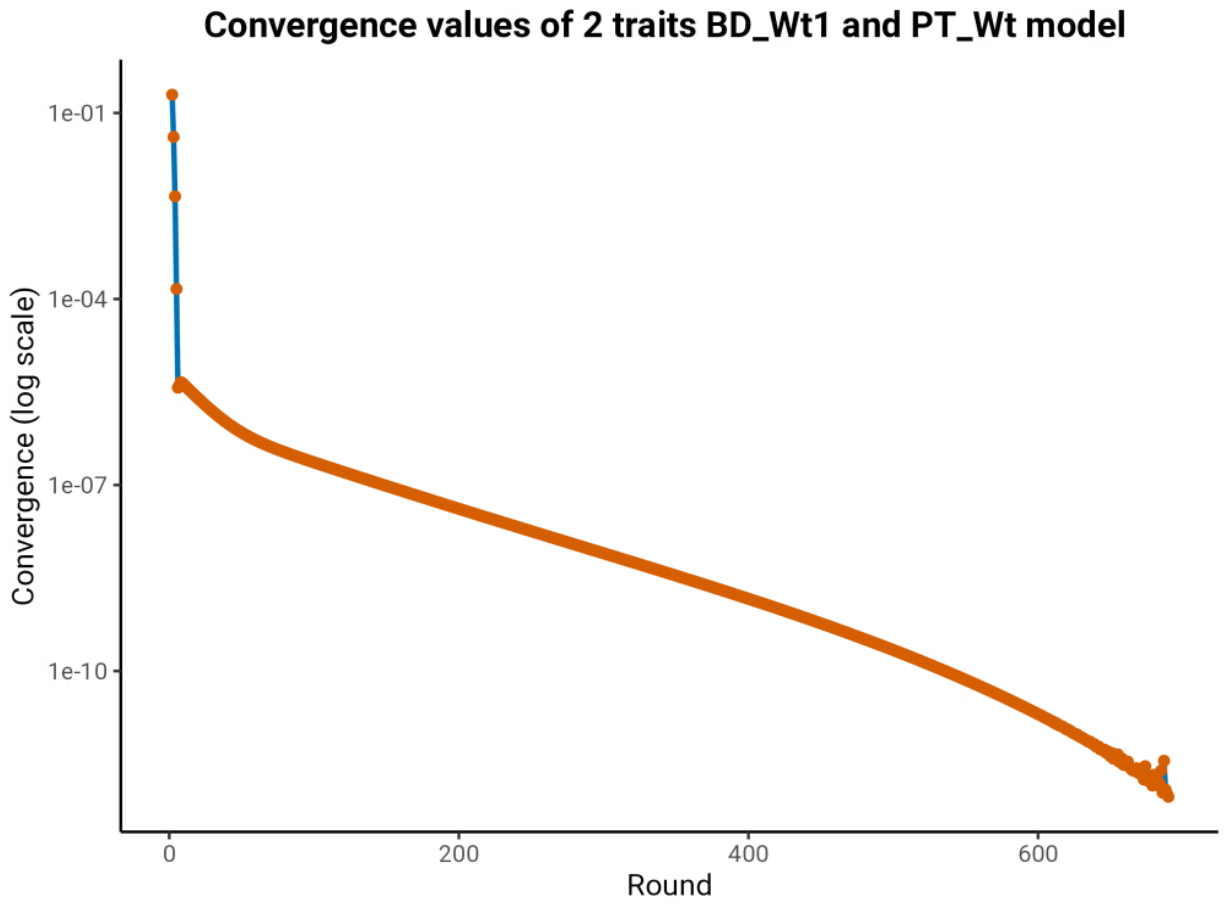
new G: Updated genetic variance (Represents additive genetic variation)

Correlations: Measures the degree of genetic dependency between traits

High values (close to 1 or -1) indicate a strong shared genetic influence between traits

Convergence plot

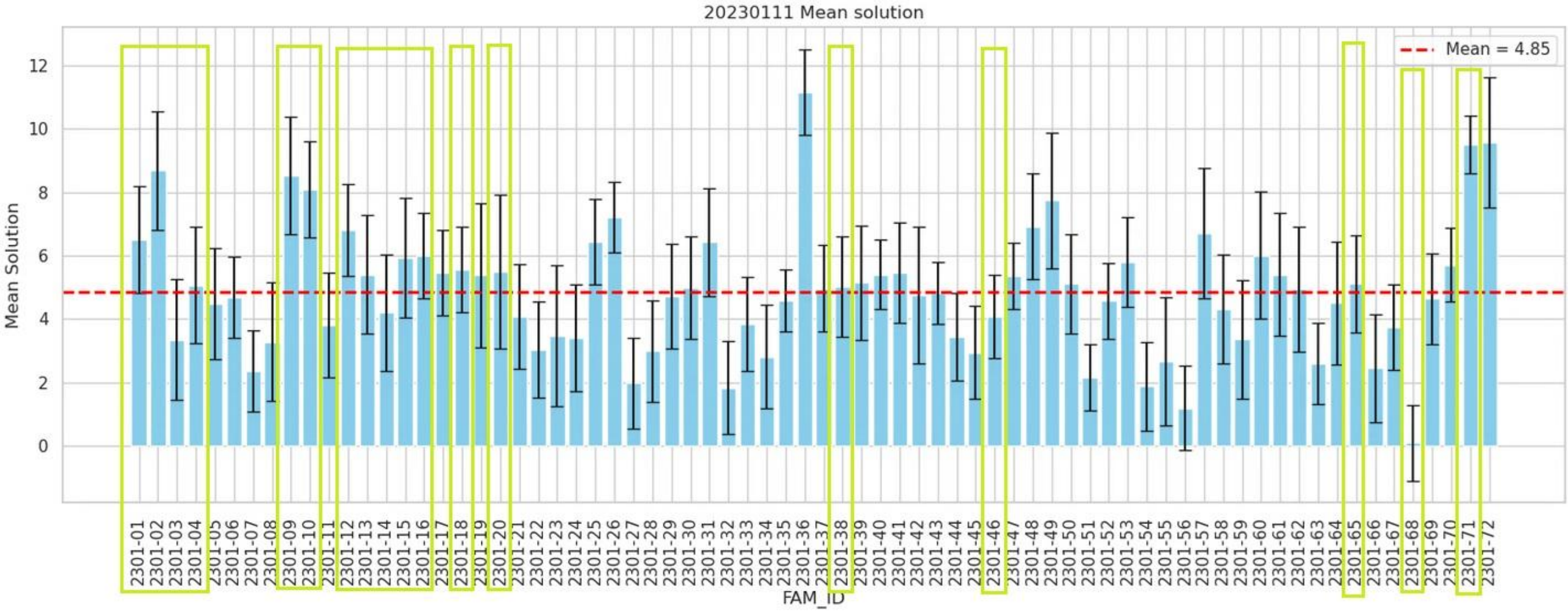


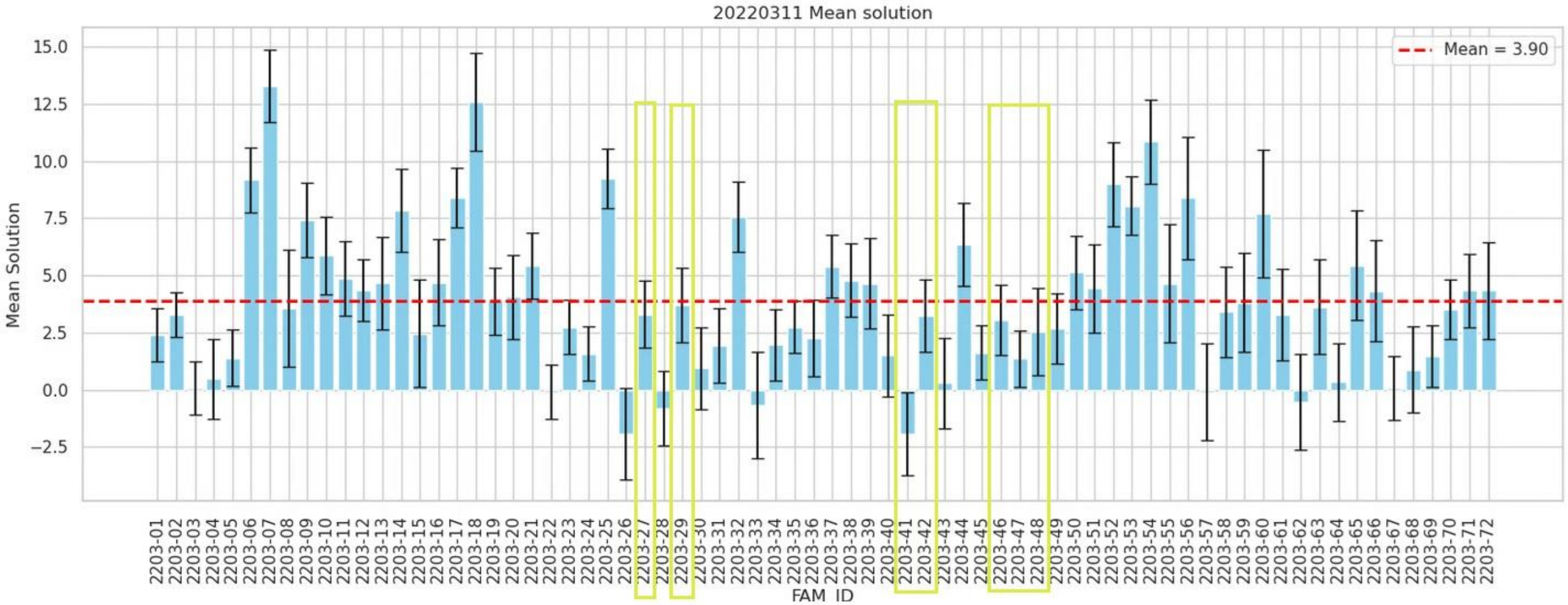


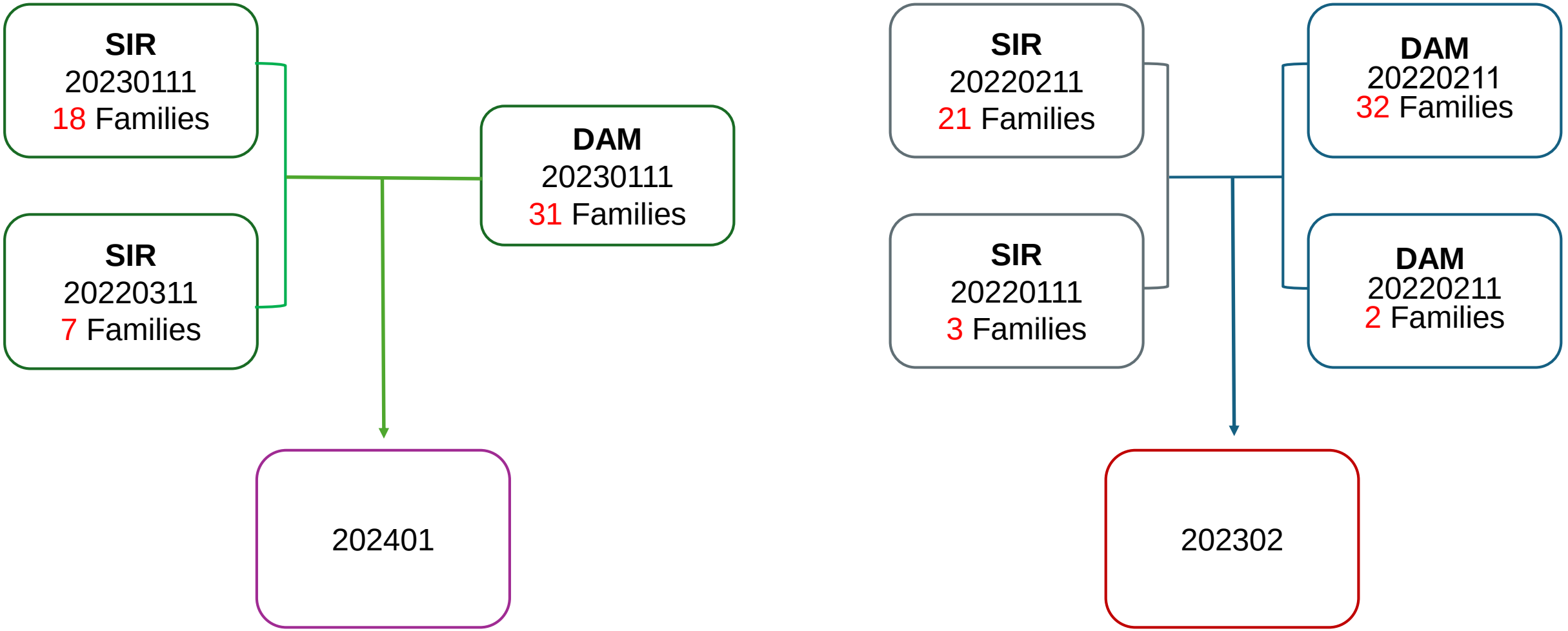
Correlations

2 traits	BD_Wt1 and PT_Wt		BD_Wt2 and PT_Wt		BD_Wt3 and PT_Wt	
correlations	1.0000	0.99987	1.0000	0.99882	1.0000	0.96201
	0.99987	1.0000	0.99882	1.0000	0.96201	1.0000

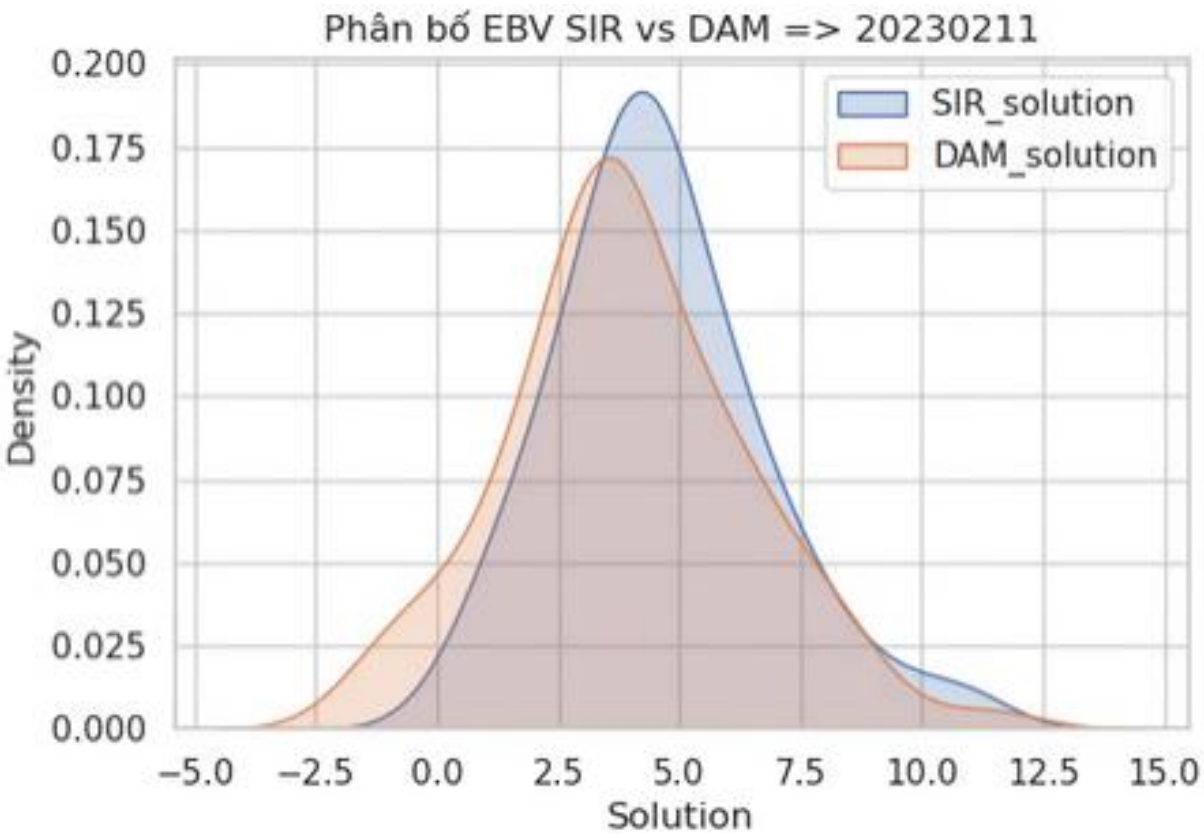
3 traits	BD_Wt1; BD_Wt2 and PT_Wt		
correlations	1.0000	0.99978	0.97325
	0.99978	1.0000	0.96858
	0.97325	0.96858	1.0000

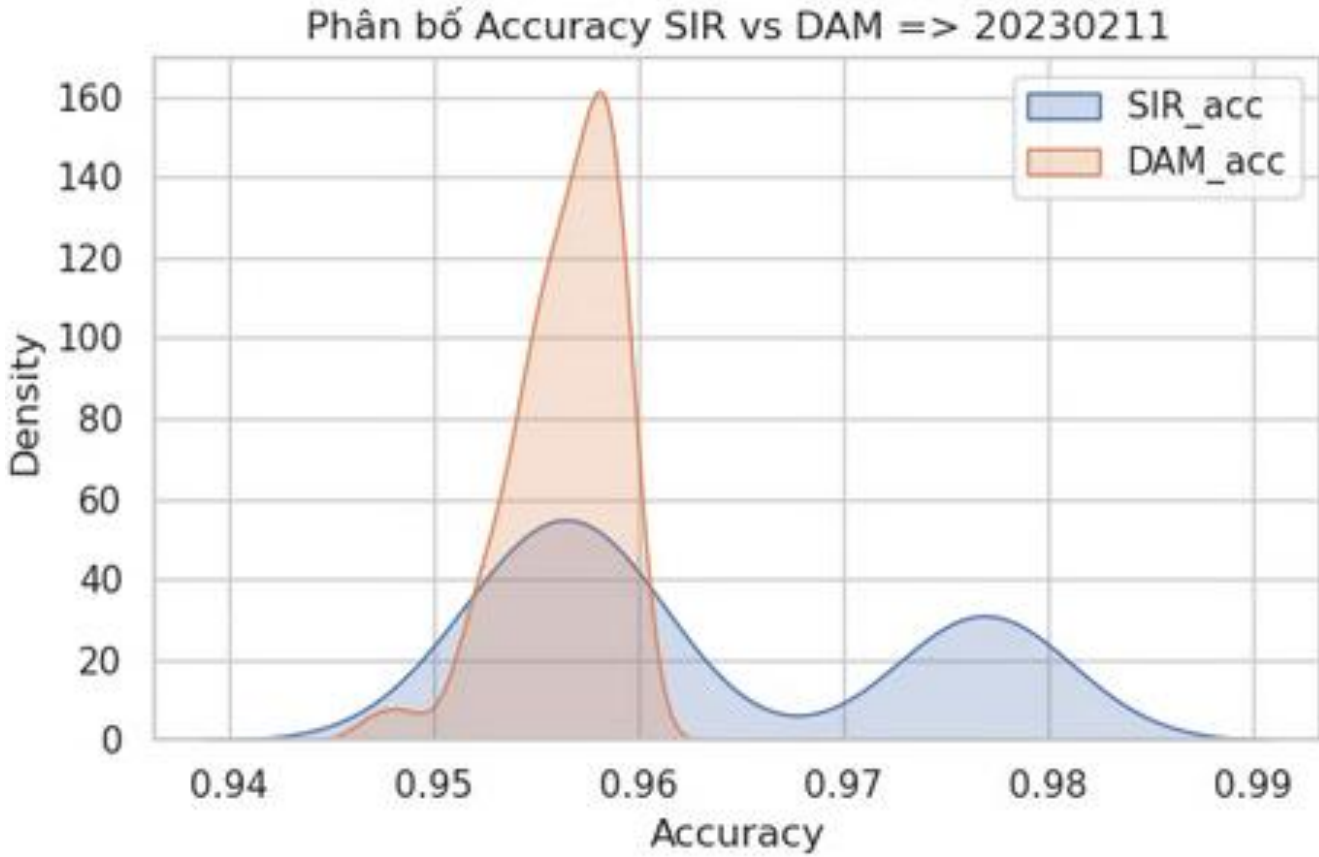

20230111 Mean solution
Mean = 4.85
Mean Solution
0
2
4
6
8
10
12
2301-01
2301-02
2301-03
2301-04
2301-05
2301-06
2301-07
2301-08
2301-09
2301-10
2301-11
2301-12
2301-13
2301-14
2301-15
2301-16
2301-17
2301-18
2301-19
2301-20
2301-21
2301-22
2301-23
2301-24
2301-25
2301-26
2301-27
2301-28
2301-29
2301-30
2301-31
2301-32
2301-33
2301-34
2301-35
2301-36
2301-37
2301-38
2301-39
2301-40
2301-41
2301-42
2301-43
2301-44
2301-45
2301-46
2301-47
2301-48
2301-49
2301-50
2301-51
2301-52
2301-53
2301-54
2301-55
2301-56
2301-57
2301-58
2301-59
2301-60
2301-61
2301-62
2301-63
2301-64
2301-65
2301-66
2301-67
2301-68
2301-69
2301-70
2301-71
2301-72
FAM_ID

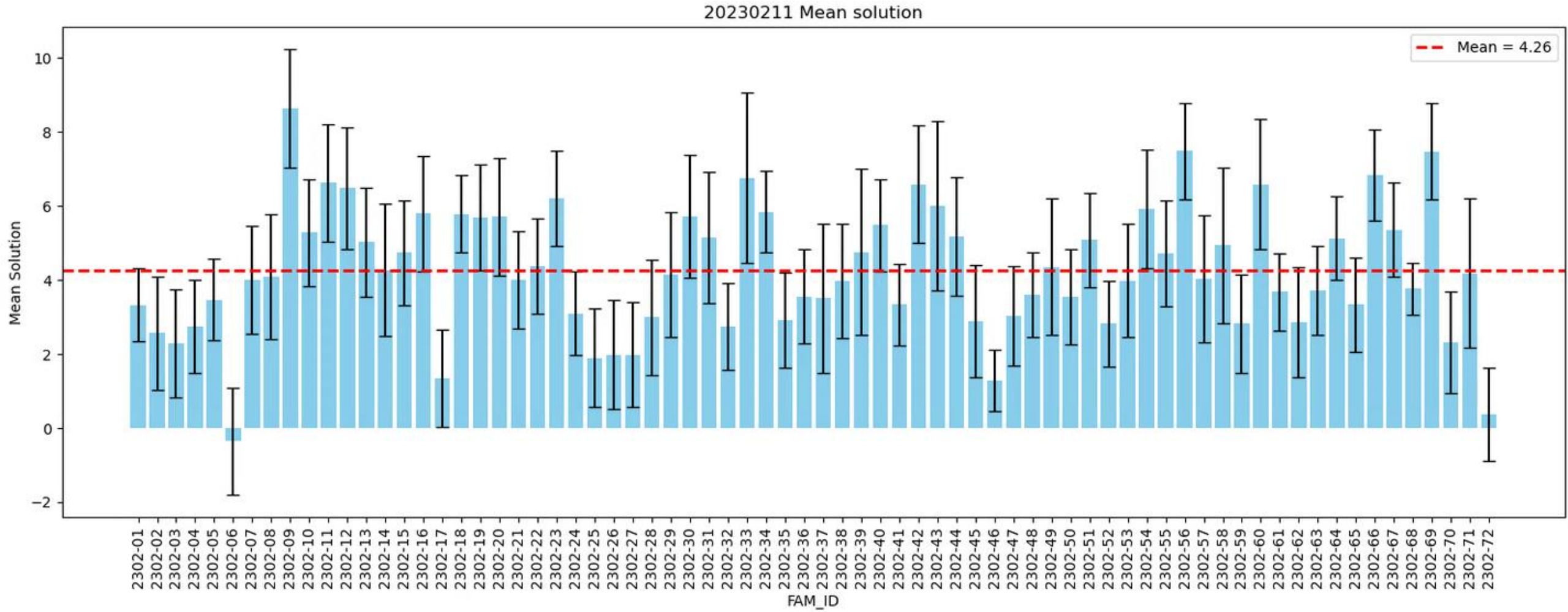

20220311 Mean solution
Mean = 3.90
Mean Solution
15.0
12.5
10.0
7.5
5.0
2.5
0.0
−2.5
2203-01
2203-02
2203-03
2203-04
2203-05
2203-06
2203-07
2203-08
2203-09
2203-10
2203-11
2203-12
2203-13
2203-14
2203-15
2203-16
2203-17
2203-18
2203-19
2203-20
2203-21
2203-22
2203-23
2203-24
2203-25
2203-26
2203-27
2203-28
2203-29
2203-30
2203-31
2203-32
2203-33
2203-34
2203-35
2203-36
2203-37
2203-38
2203-39
2203-40
2203-41
2203-42
2203-43
2203-44
2203-45
2203-46
2203-47
2203-48
2203-49
2203-50
2203-51
2203-52
2203-53
2203-54
2203-55
2203-56
2203-57
2203-58
2203-59
2203-60
2203-61
2203-62
2203-63
2203-64
2203-65
2203-66
2203-67
2203-68
2203-69
2203-70
2203-71
2203-72
FAM ID

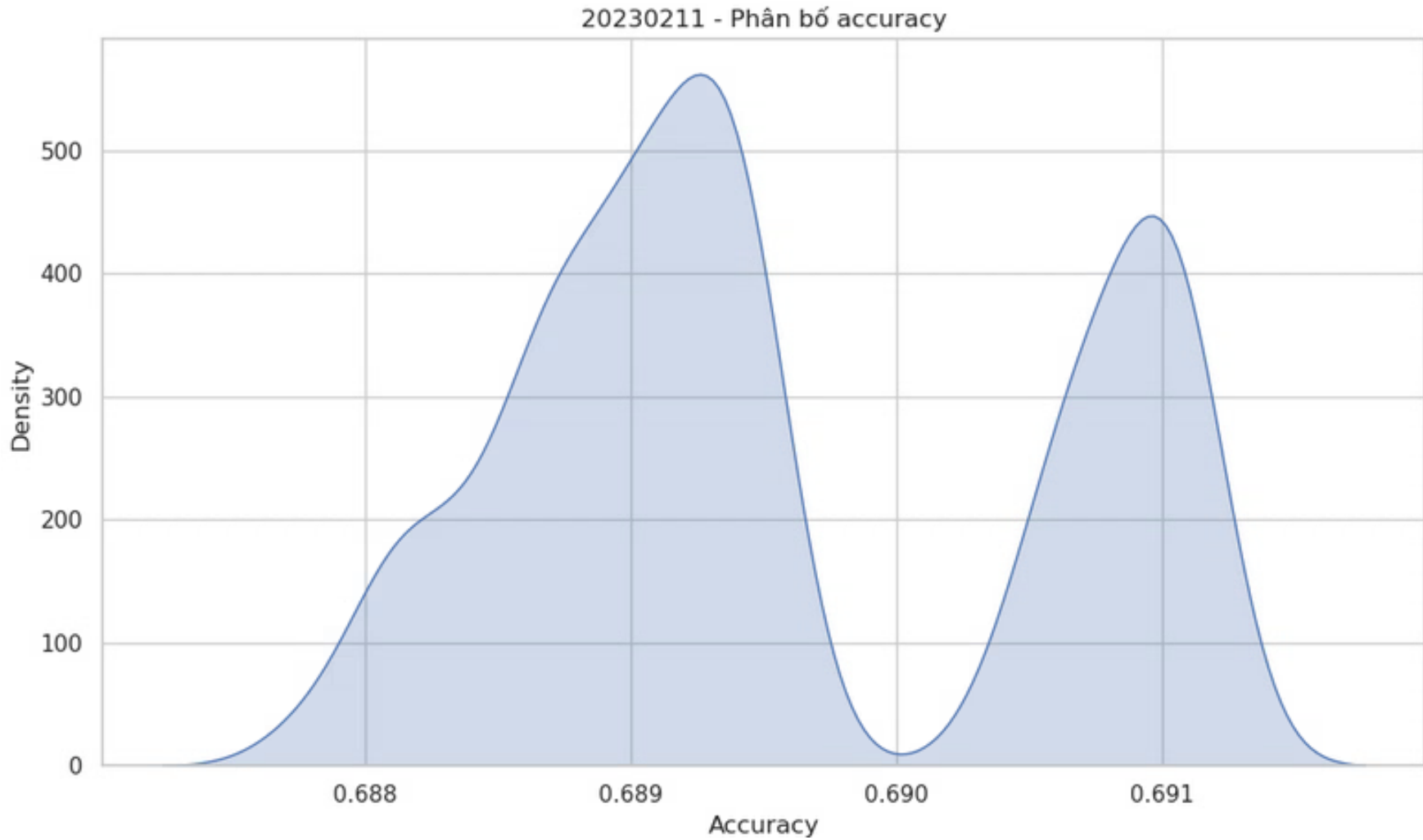

SIR
20230111
18 Families
SIR
20220311
7 Families
DAM
20230111
31 Families
202401
SIR
20220211
21 Families
SIR
20220111
3 Families
DAM
20220211
32 Families
DAM
20220211
2 Families
202302

COHORT: 20230211

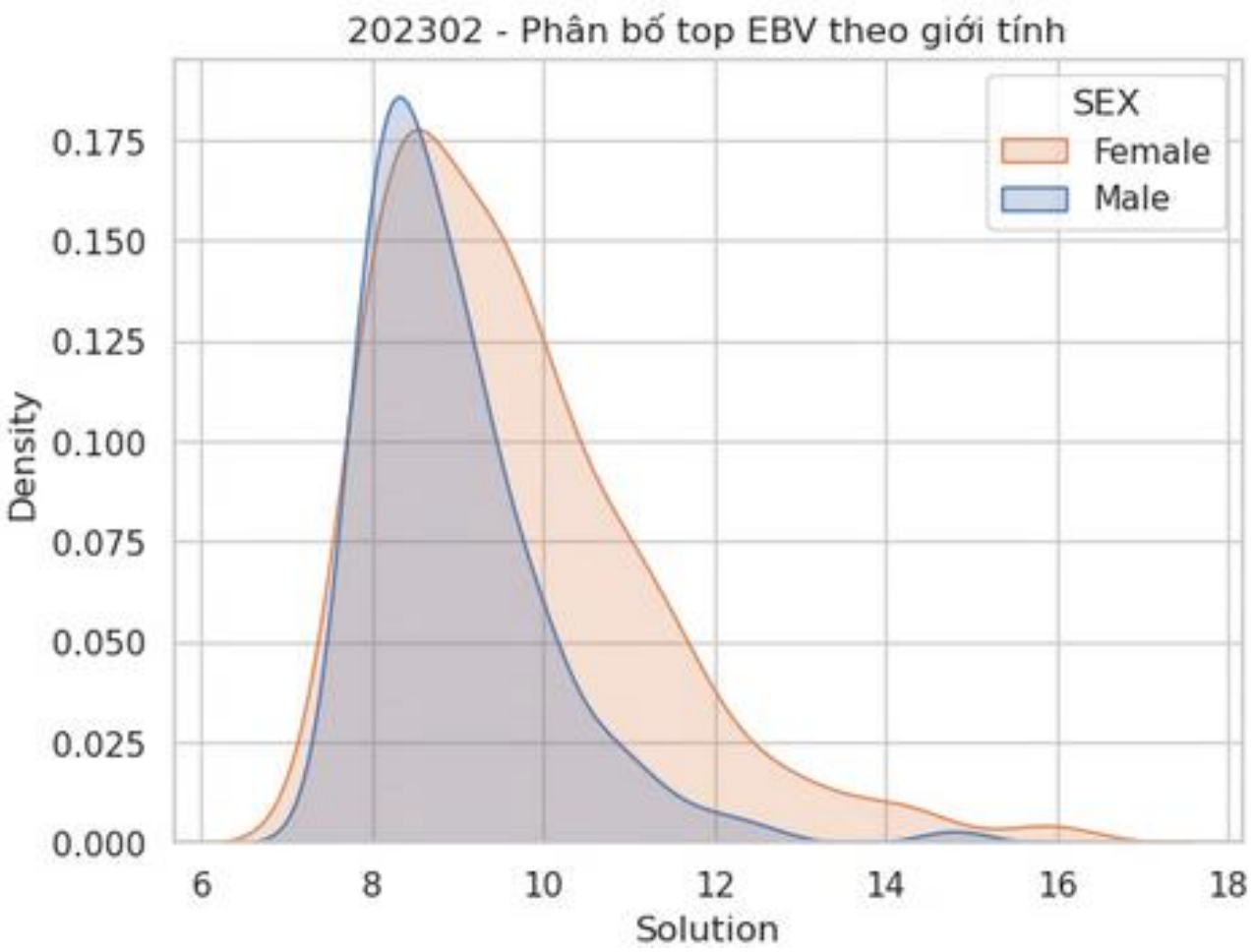


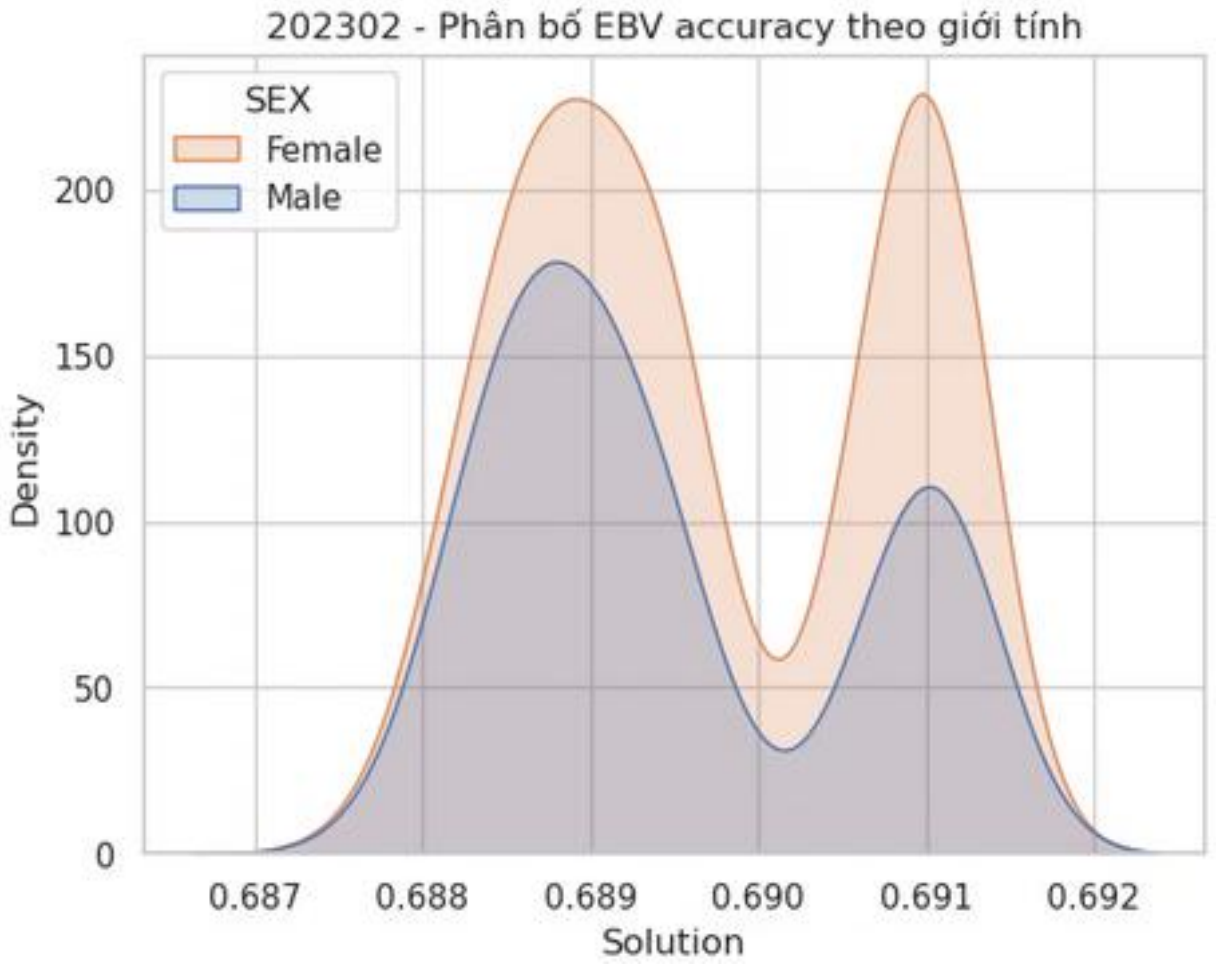



20230211 Mean solution
Mean = 4.26
Mean Solution
FAM_ID
10
8
6
4
2
0
−2
2302-01
2302-02
2302-03
2302-04
2302-05
2302-06
2302-07
2302-08
2302-09
2302-10
2302-11
2302-12
2302-13
2302-14
2302-15
2302-16
2302-17
2302-18
2302-19
2302-20
2302-21
2302-22
2302-23
2302-24
2302-25
2302-26
2302-27
2302-28
2302-29
2302-30
2302-31
2302-32
2302-33
2302-34
2302-35
2302-36
2302-37
2302-38
2302-39
2302-40
2302-41
2302-42
2302-43
2302-44
2302-45
2302-46
2302-47
2302-48
2302-49
2302-50
2302-51
2302-52
2302-53
2302-54
2302-55
2302-56
2302-57
2302-58
2302-59
2302-60
2302-61
2302-62
2302-63
2302-64
2302-65
2302-66
2302-67
2302-68
2302-69
2302-70
2302-71
2302-72

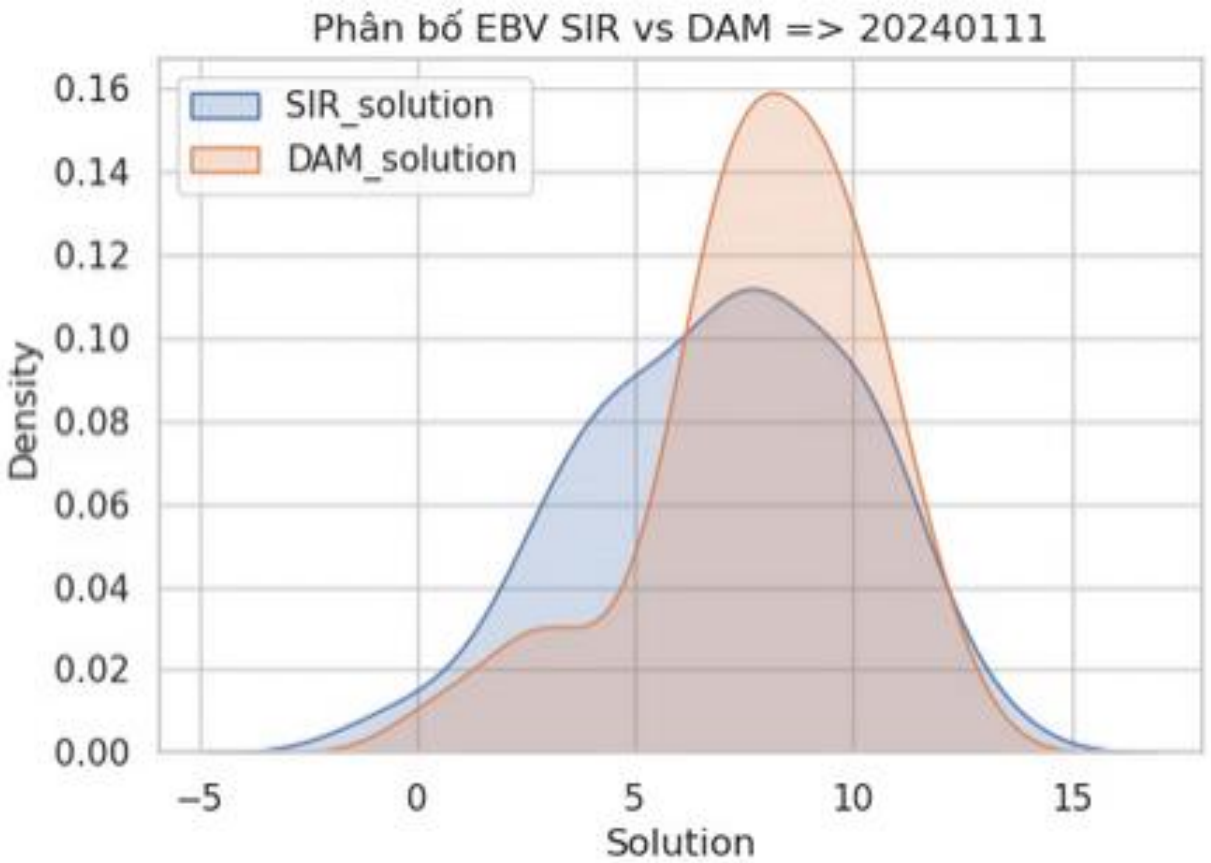

20230211 - Phân bố accuracy
Density
500
400
300
200
100
0
0.688
0.689
0.690
0.691
Accuracy

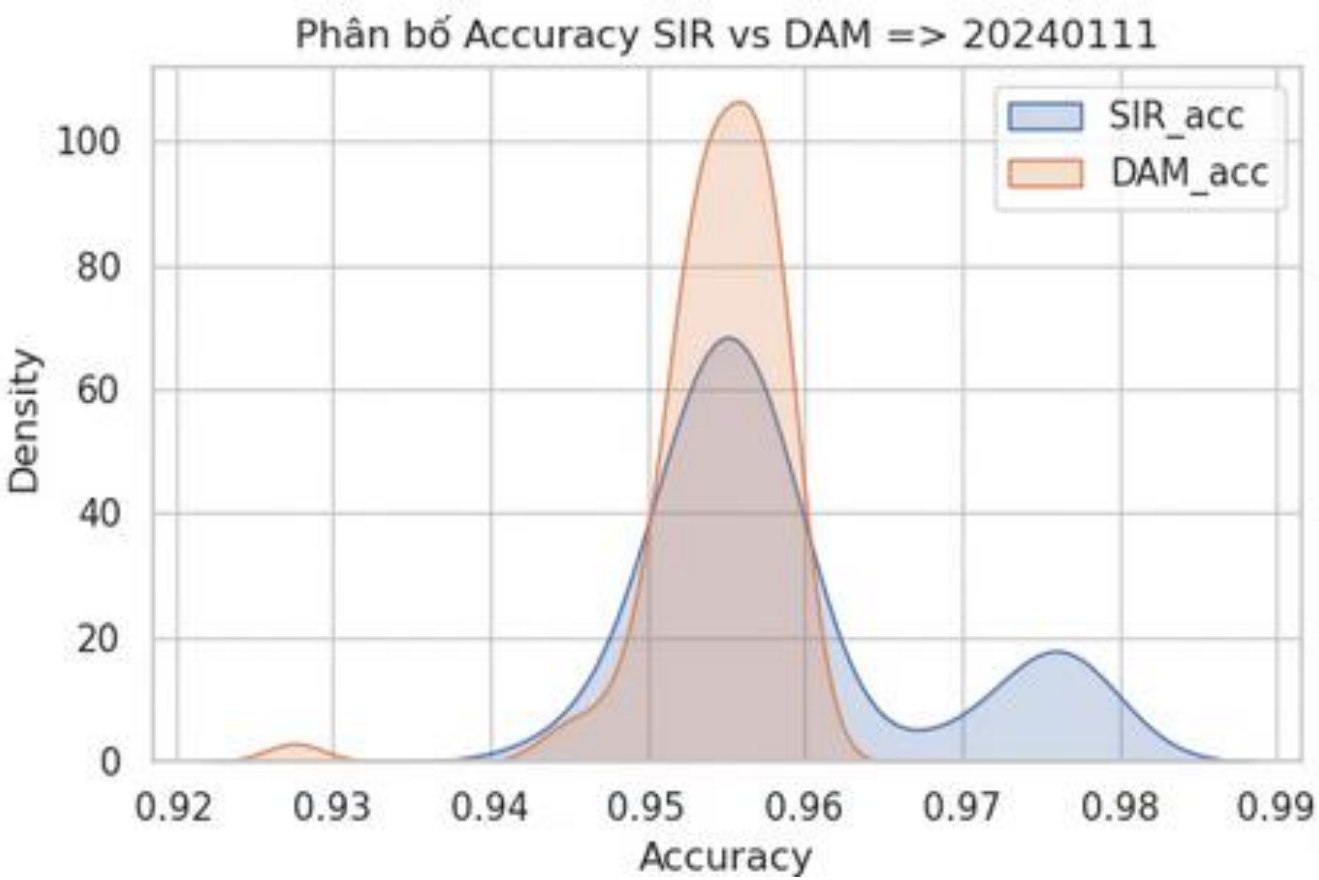
Top-ranked EBV

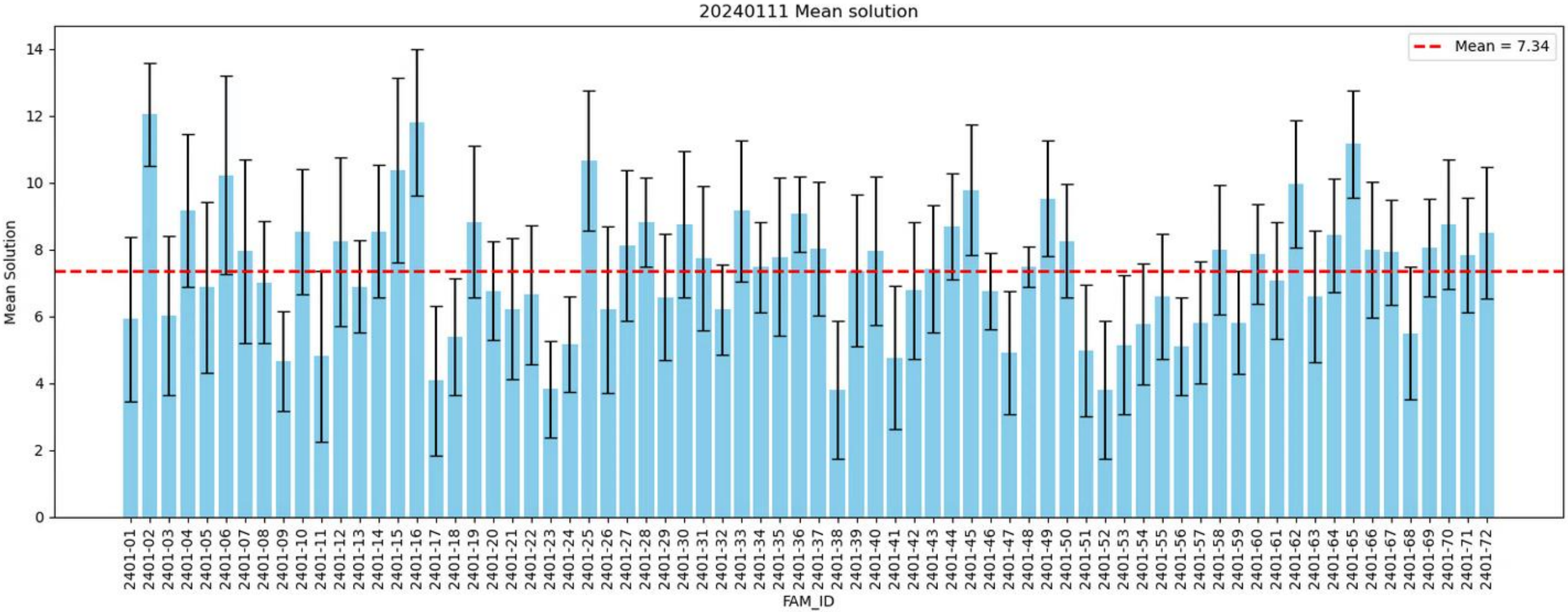


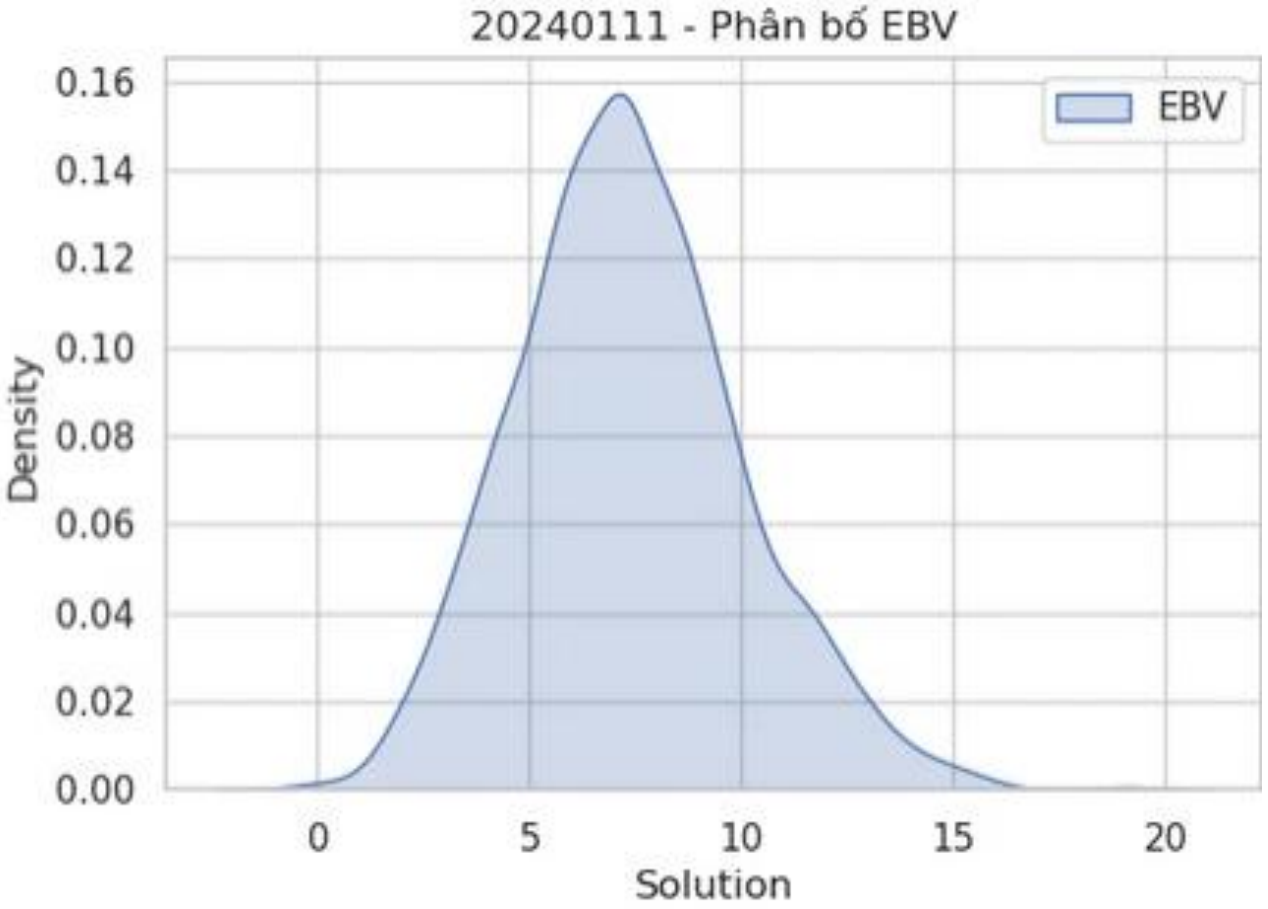


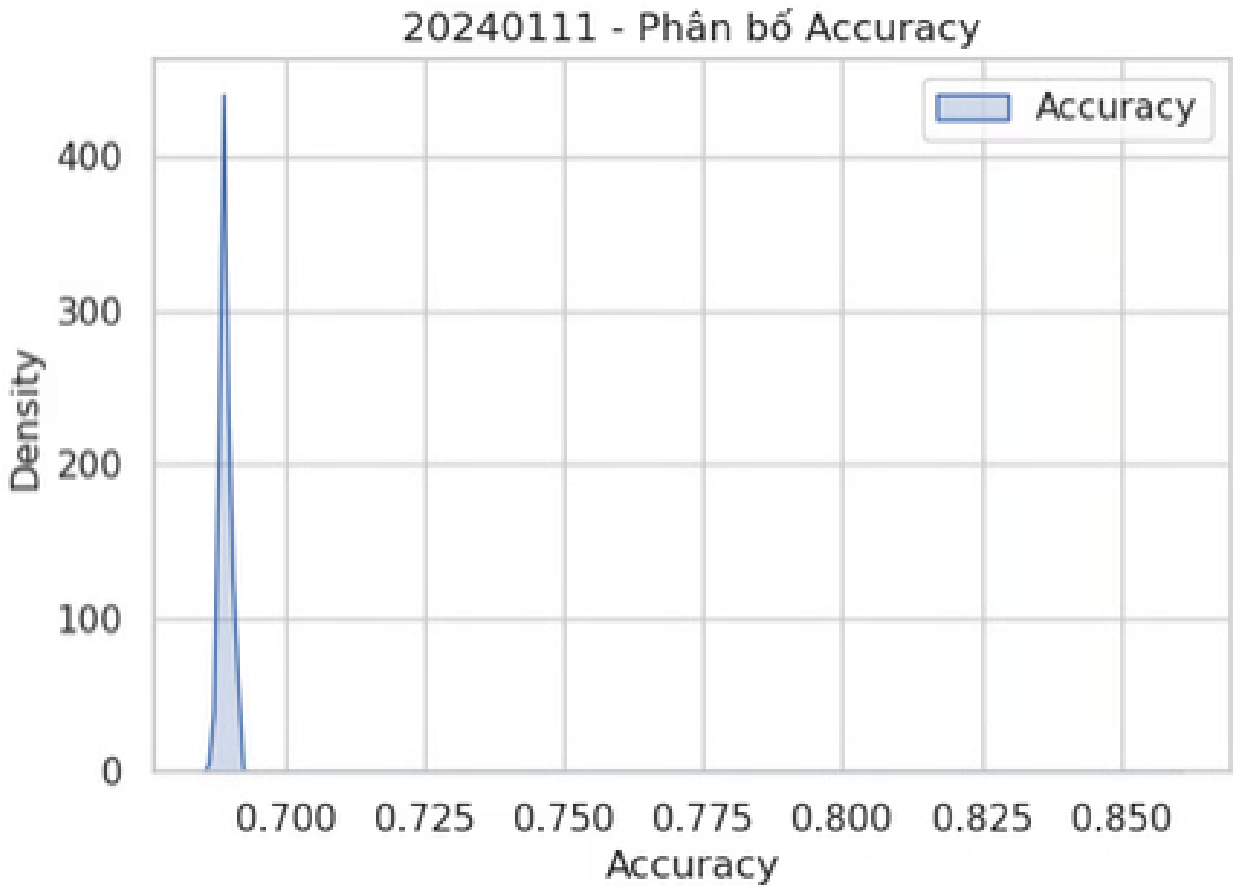
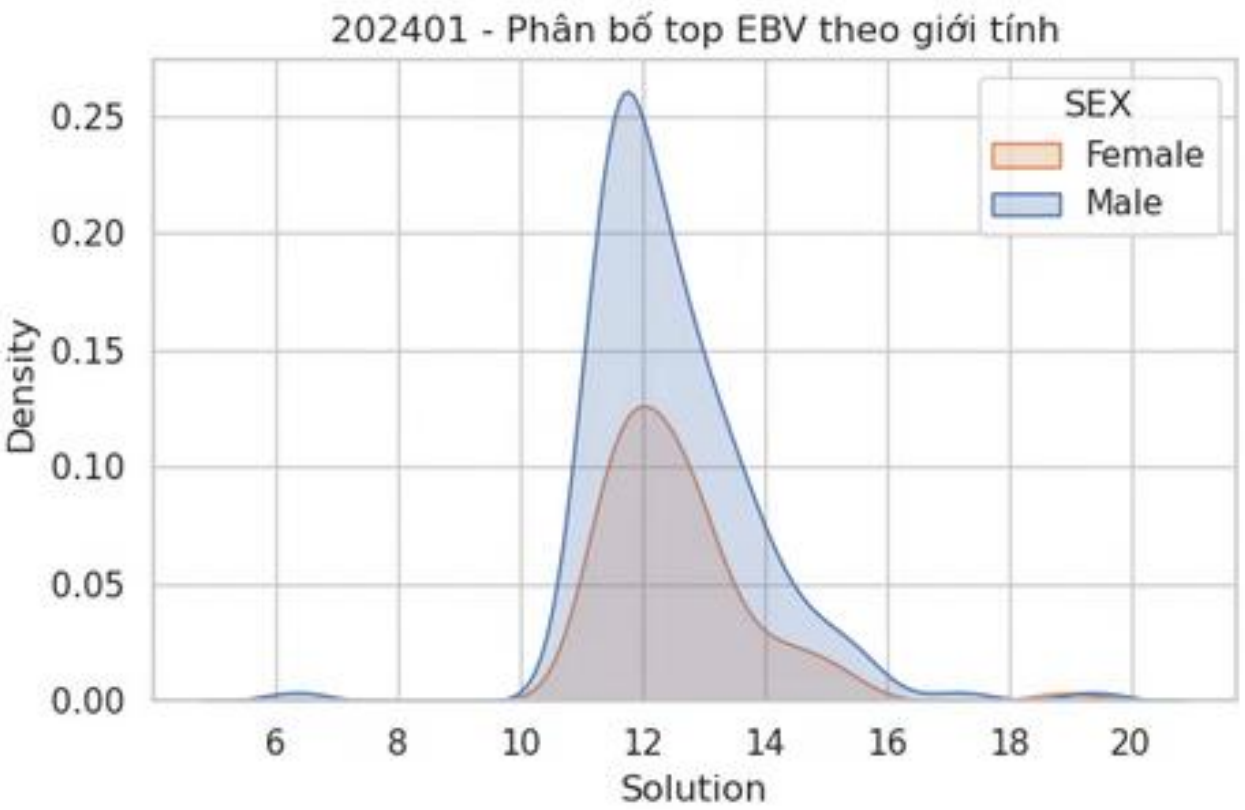
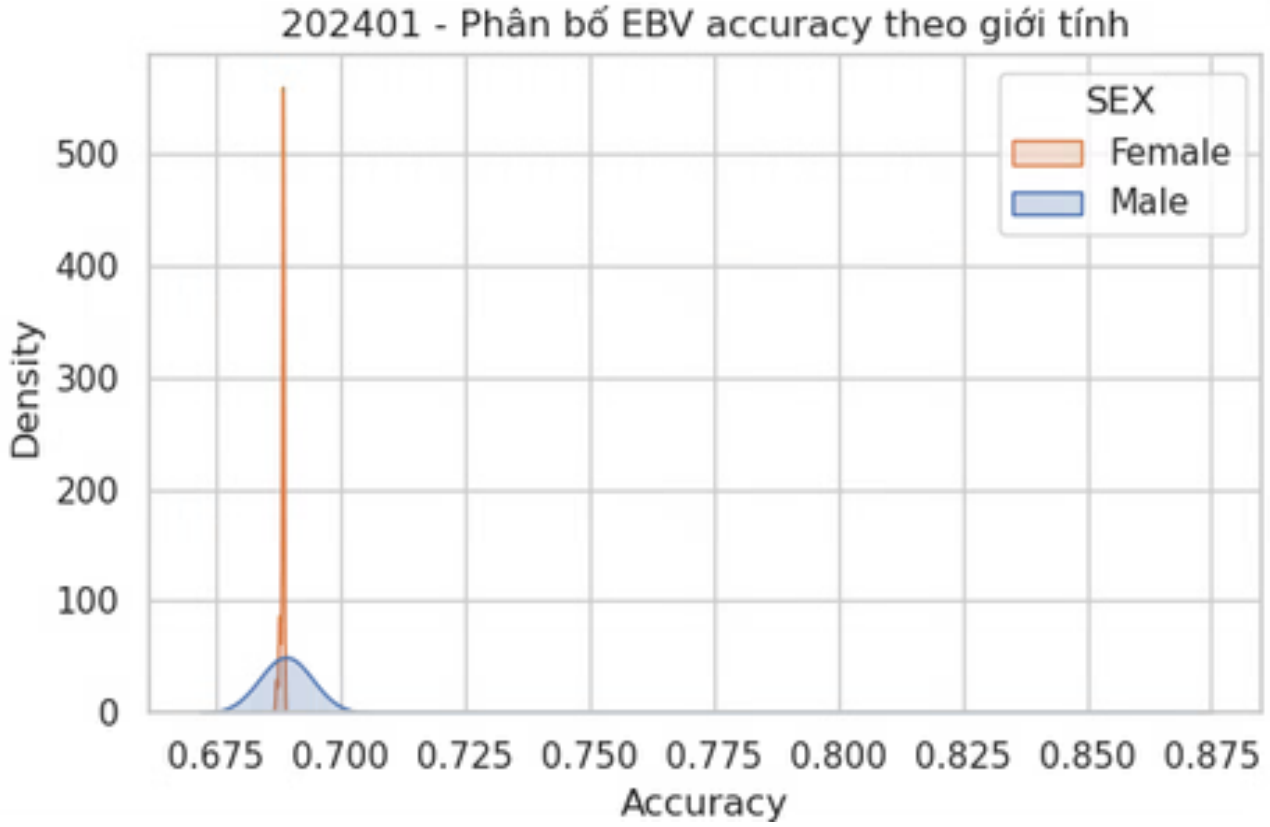
COHORT: 20240111






20240111 Mean solution
Mean = 7.34
Mean Solution
FAM_ID
0
2
4
6
8
10
12
14
2401-01
2401-02
2401-03
2401-04
2401-05
2401-06
2401-07
2401-08
2401-09
2401-10
2401-11
2401-12
2401-13
2401-14
2401-15
2401-16
2401-17
2401-18
2401-19
2401-20
2401-21
2401-22
2401-23
2401-24
2401-25
2401-26
2401-27
2401-28
2401-29
2401-30
2401-31
2401-32
2401-33
2401-34
2401-35
2401-36
2401-37
2401-38
2401-39
2401-40
2401-41
2401-42
2401-43
2401-44
2401-45
2401-46
2401-47
2401-48
2401-49
2401-50
2401-51
2401-52
2401-53
2401-54
2401-55
2401-56
2401-57
2401-58
2401-59
2401-60
2401-61
2401-62
2401-63
2401-64
2401-65
2401-66
2401-67
2401-68
2401-69
2401-70
2401-71
2401-72


20240111 - Phân bố EBV
EBV
Density
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
Solution


20240111 - Phân bố Accuracy
Accuracy
Density
400
300
200
100
0
0.700
0.725
0.750
0.775
0.800
0.825
0.850
Accuracy

Top-ranked EBV