Exercises Day 3

Part 1

Exercise 1   Optimizing breeding Programs

Imagine a breeding program for beef cattle. You can use *truncsel.xls* ('optimize age structure in a herd'-sheet) to estimate the amount of genetic gain that can be achieved per year.
The a breeder decides to select on an index that aims for a specified objective (e.g. the Japanese export market, with a standard deviation of the breeding objective of $20.

The accuracy of bulls' EBVs for the selection index increases with age as follows;

Age:	1	2	3	4	5
Male Accuracy	0.35	0.45	0.78	0.80	0.82
Female Accuracy	0.35	0.45	0.52	0.54	0.55

Assume a breeding program where the breeder has a closed stud (nucleus) of 400 breeding females. The weaning rate is 1. He mates 40 cows per bull annually. The age structure of the cows is such that of all cows mated annually, the distribution over age classes 2, 3, 4, and 5 is 40%, 30%, 20% and 10%, respectively.

- Predict the annual rate of genetic gain if bulls are used only once (progeny dropped when age of sire is 2 yrs) and females are selected as heifers and drop their first calves at 2 years. After 2 years, some cows are culled at random.

Response = selection superiority/generation interval = $(S_m + S_f)/(L_m + L_f)$
S_m = im x acc_m x sigma_H.
S_f = if x acc_f x sigma_H
Need 10 bulls per year. 200 weaned: select 10/200 weaned = 5% → i_m= 2.06
Need 40%=160 cows every year, select 160/200 weaned → i_f= 0.351
sigma_H is SD of breeding objective (in $) is not given.
Lm = 2
Lf = 0.4 x 2 + 0.3 x 3 + 0.2 x 4 + 0.1 x 5 = 3
Selection accuracy is 0.35 as selection is at 1 year of age when progeny are dropped at 2 yrs of age, after that they are not selected.
Hence, response per year = [(2.06 x 0.35 + 0.351 *0.35)/(2+3)] . σ_H
=0.17 σ_H (so the progress is about one fifth of the standard deviation of the breeding objective)

σ_H = $20, so the gain is $3.37 per year

It is handy top set up a Table with the age structure as t makes it easier to see how many need to be selected.

Exercise Day3

own performance test scenario											
age structure	2	3	4	5	6	total	L (gen interv)	sel.prop	sel.int	sel.acc	SDH
males	10	0	0	0	0	10	2.00	5%	2.06	0.35	20
females	160	120	80	40	0	400	3.00	80%	0.35	0.35	20
	resp	Sm	14.42	Sf		2.45					
	3.374	Lm	2.00	Lf		3.00					

- Answer question 2, but now assuming bulls are used only once and the first time at the age of 4 years, after being progeny tested.

There are 2 changes: L_m= is now, as if used at 4 years they would drop progeny at 5 years of age 5 and acc_m = 0.78. Response = 0.22 σ_H = \$4.43

progeny test scenario											
age structure	2	3	4	5	6	total	L (gen interv)	sel.prop	sel.int	sel.acc	SDH
males	0	0	0	10	0	10	5.00	5%	2.06	0.8	20
females	160	120	80	40	0	400	3.00	80%	0.35	0.35	20
	resp	Sm	32.96	Sf		2.45					
	4.4263	Lm	5.00	Lf		3.00					

SO the progeny test scenario seems very good. However, one might ask which progeny are being teste din the scheme. Not the progeny in the nucleus, as we only use progeny tested sires in the nucleus, and we probably would not progeny test 200 young males. More realistic is to test, for example, 40 young males, outside the nucleus. The selection intensity would then be only 10/40 is 25%, and the selection intensity would only be 1.27. The response would only be 2.84. Much lower than the own performance test scenario, and much more costly!

progeny test scenario											
age structure	2	3	4	5	6	total	L (gen interv)	sel.prop	sel.int	sel.acc	SDH
males	0	0	0	10	0	10	5.00	25%	1.27	0.8	20
females	160	120	80	40	0	400	3.00	80%	0.35	0.35	20
	resp	Sm	20.32	Sf		2.45		need to progeny test 40 bulls here			
	2.84625	Lm	5.00	Lf		3.00					

- Predict (calculate) the annual response when selection is optimized across age classes.

(truncsel does this for you, assume the folowing

Nr of breeding females	400
Mating ratio (females/male)	40
Weaning rate	100%
mortality rate (%)	0
Nr of males used per year	10
Nr of females used per year	400
Nr progeny born per year	400
Genetic Standard Deviation	20
Earliest possible Year of first drop	2
Maximum nr of age classes	6

Exercise Day3

males accuracy	females accuracy
0.350	0.350
0.350	0.350
0.450	0.450
0.780	0.520
0.800	0.540
0.820	0.550

Note that I have moved up the accuracies one age class. For example, first drop of progeny is at 2 years, so for age class 2 we should really use a year 1 accuracy, as selection takes place almost a year before the progeny are born.

Note that the group means need not to be given as input, they are calculated by the program as a difference of dG between age classes

males Nr Selected	females
0.0	0.0
0.7	170.2
0.6	112.9
5.2	67.6
2.4	34.7
1.0	14.6
0.0	0.0
0.0	0.0
0.0	0.0
0.0	0.0
10.00	400.0
Total	Total

mean of next year	34.8807			
progress per year	5.8135	is	29.1	%of GeneticSD

	males	females	average		gain per year
parent superiority	35.61	6.60	21.11	=	5.81
genertion interval	4.23	3.03	3.63		

So with optimised age structure we have more gain

Exercise Day3

One important trait in the index is marbling, which is not directly measured. What would happen to the rate of genetic gain, and the underlying components such as age structure) if we had genomic selection in place? Assume that the accuracy of selection improves as follows

Age:	1	2	3	4	5
Male Accuracy		0.50	0.55	0.80	0.81	0.83
Female Accuracy	0.50	0.55	0.58	0.58	0.58

ageclass	N in group	mean	males accuracy	females accuracy
1	200	24.46	0.500	0.500
2	190	18.35	0.500	0.500
3	181	12.23	0.550	0.550
4	171	6.12	0.800	0.580
5	163	0.00	0.810	0.650

males	females
Nr Selected	
0.0	0.0
3.5	160.1
1.3	114.6
3.4	70.6
1.3	40.8
0.5	13.9
0.0	0.0
0.0	0.0
0.0	0.0
0.0	0.0
10.00	400.0
Total	Total

mean of next year	38.3829			
progress per year	6.3972	is	32.0	%of GeneticSD

	males	females		average		gain per year
parent superiority	33.19	8.29		20.74	=	6.40
genertion interval	3.40	3.08		3.24		

So with optimised genomic selection we have a gain of $6.40
Generation intervals are lower than in previous, because young animals have relative more accuracy, more gain from extra genomic information.

Exercise 3.2   Optimal Contributions Selection

The exercise demonstrates the idea of balancing selection and long-term inbreeding in an index of genetic merit and parental co-ancestry.

Use the *xGxAx.xls* program.

This allows you to optimally select from 4 males and 4 females. The X-matrix contains the 'optimal contributions', i.e. the proportion of genes in the next generation that will be contributed by this parent.

Notice that the contributions of males sum to 0.5, and similarly for females. Notice that some males and females are genetically superior to the competitors. With no relationships among the four candidates, matrix A is the identity matrix. The criterion (Crit) for judging this solution is equal to x'G + λx'Ax. Be sure that you understand this before continuing.

For a give dataset (i.e. Index values in G and Additive genetic relationships in the Relationships matrix), you can find optimal contributions for a single lambda (λ) value. Lambda is the penalty on inbreeding. You can also draw the graph, which loops through all lambda values between 0 and - (a big number). A graph will be different each time you change values in either breeding values or relationships among candidates.

Put all relationships to 0, and find optimal contributions for $\lambda = 0$ and $\lambda = -9999$.

Before clicking the 'optimal contributions' button, try to predict them by reasoning, and subsequently check them by the program.

Change the relationships by making the two best males full sibs.

1.   Why does the graph change (after clicking 'draw graph')?

2.   Predict and find optimal contributions for $\lambda = 0$ and $\lambda = -9999$.

3.   For 2) and 3) make a note of dF and dG.

4.   Find a best value for λ (i.e "what's your favourite point on the graph?")

Answer questions 1-4 after changing more relationships by also making the two best females full sibs.

Exercise 3.3   Open Nucleus Schemes

Use *truncsel.xls* to work out the degree of openness of a nucleus breeding scheme.

Imagine a herd with 400 breeding females, breeding bulls for about 15,000 cows in the second tier (is this feasible?). IN the 'optimize age structure in a herd'-sheet, you can verify that the herd needs around 8 male and 140 females replacements per year.

Now use the 'optimize selection across age classes' sheet to work out the degree of openness. Use two classes, one for the nucleus and one for the second tier. Make assumptions for each class about
1) the number of selection candidates
2) the SD, i.e. accuracy/, of the EBV / selection criterion)
3) the difference in mean between the two tiers.

Now based on these assumptions, work out the optimal number selected across these tiers, first for males and then for females. Also, estimate the extra gain that can be achieved by a certain level of trait recording in the commercial tier.

Exercise 3.4   Effect of reproductive technologies

Still using the same seedstock herd with 400 breeding females, and the 'optimize age structure in a herd'-sheet in *truncsel*, look at options to boost reproductive rates (Look at MOET and JIVET) and give an estimate of how much extra genetic gain this could give.

Are these reasonable prediction of the effect of such technologies? What are the shortcomings of a simple modeling via this *truncsel* program?

Exercise 5   Economic Evaluation of Breeding Programs

The bulls from the stud in Exercise 2 that are sold to the commercial farmer have on average 150 progeny of 3 years that they are used (weaning rate 100%, mated to 50 dams). Work out the size of the commercial population that can be served by this stud.

Make a prediction of the net present value of the genetic gain due genetic improvement in this flock (for the case without genomic selection.

How much can the flock afford to invest in genetic improvement to at least break even on its breeding effort.

How much can the flock afford to invest in genomic selection.

What if only bulls were genotyped?

Assuming 50 daughters per bull per year, so if a bull gets used for 3 years it has 150 progeny. The stud has 200 males born per year, needs 10 for nucleus replacements so can sell 190 bulls per year (this is assumed but in reality, not all bulls are usually sold). 190 bulls can sire 190*50=9500 cows per year, but they are used for 3 years hence in total the nucleus can serve 28,500 commercial cows

Assume now the same genetic gain as in the previous exercise with genomic selection: $6.12 per cow per year

After an initial lag of 2 generations (about 6 years), this value gets added to the genetic mean year over year. Value in each year is the genetic mean multiplied by the animals expressing it (N=28,500/2 (divide by 2 as only males express the beef trait)).

Value expressed in year t is worth in today's value $1/(1+r)^t$ where r is the discount rate.

Assume r=0.05 (interest rate).

Hence the genetic mean and value over the next 20 years (assuming a 6 year lag) could look like

year	disc fact	genetic mean	cum benefit	cost	disc retruns
1	1.000	0	$ -		$0
2	0.952	0	$ -		$0
3	0.907	0	$ -		$0
4	0.864	0	$ -		$0
5	0.823	0	$ -		$0
6	0.784	0	$ -		$0
7	0.746	$6.04	$ 86,083		$64,236
8	0.711	$12.08	$ 172,166		$122,355
9	0.677	$18.12	$ 258,248		$174,793
10	0.645	$24.16	$ 344,331		$221,959
11	0.614	$30.20	$ 430,414		$264,237
12	0.585	$36.25	$ 516,497		$301,985
13	0.557	$42.29	$ 602,580		$335,539
14	0.530	$48.33	$ 688,663		$365,212
15	0.505	$54.37	$ 774,745		$391,299
16	0.481	$60.41	$ 860,828		$414,073
17	0.458	$66.45	$ 946,911		$433,791
18	0.436	$72.49	$ 1,032,994		$450,692
19	0.416	$78.53	$ 1,119,077		$465,000
20	0.396	$84.57	$ 1,205,160		$476,923
				NPV	$4,482,094

Showing a cumulative NPV of genetic improvement for the population of 28,500 cows (expressing the trait in only male progeny) of $4.48M.

Exercise Day3

Using GFLOW

Use the *gflow.xls* sheet to determine the GFLOW of improved genetic material from the stud to the commercial tier.

What is the value of one unit of difference in EBV between two stud bulls?

What is this value for a difference between two bulls sold to the commercial tier?

In the scenario with genomic selection the numbers used from each age class were

age class	1	2	3	4	5	total
nr males selected	0.0	3.8	1.4	3.5	1.3	10.0
nr females selected	0.0	162.0	118.1	75.5	44.4	400.0
male contrib to progeny	0.000	0.190	0.069	0.174	0.067	0.5
female contrib to progeny	0.000	0.203	0.148	0.094	0.056	0.5

The last rows are inserted in gflow.xls (rows 10, 15 and 25 (last one not shown below)

P matrix	Sires of Nucleus					Dams of Nucleus				
	1	2	3	4	5	1	2	3	4	5
1	0.00	0.19	0.07	0.17	0.07	0.00	0.20	0.15	0.09	0.06
2	1	0	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0
1	0.00	0.19	0.07	0.17	0.07	0.00	0.20	0.15	0.09	0.06
2	0	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	0	1	0	0	0

Note that this gives the following in gflow (same as truncsel)

Nucleus	Generation Interval
Lm	3.234504
Lf	3.005625
1/(Lm+Lf)	0.160253

The bulls sold from the stud to the commercial pass on genes over 3 years, so we insert in rows 30 and 35 columns S-U

Exercise Day3

row 30	0.166667	0.166667	0.166667
	0	0	0
	0	0	0
	0	0	0
	0	0	0
row 35	0.166667	0.166667	0.166667

and the female reproduction in the commercial is assumed the same (keep cows for 3 years.

Now consider expression of a beef trait only expressed in males (steers) at the age of 2, For this we put a value of 0.5 in cell AN31. We consider one unit of genetic superiority in the stud males (cell AM10).
Using an interest rate of 5% gives a CDE of 0.506. The CDE for females (putting a 1 in column AM15) is 0.518

The value of the genetic improvement calculated with gflow is the sum of all expressions discounted over years (CDE) times the superiority time the number of total animals expressing it (use N=28 now as we already accounted for the fact that only 0.5 express it.

28500 *(0.506*$31.55 + 0.518*$6.12) = $545,253

This is the value of one selection round over the next 20 years
If we do another round next year, we get the same, but a discounted, as it all happens one year later. The sum of all round over the next 20 years is $3.56M. This is quite a bit lower than the $4.48 as calculated with the simple dG methods in the previous exercise. The reason is that GFLO accounts better for the delay. Due to the long generation intervals, the future selection rounds will also not be fully completed over the next 20 years, as seen below

year	disc fact		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.000	$ -	$ -																			
2	0.952	$ -	$ -	$ -																		
3	0.907	$ -	$ -	$ -	$ -																	
4	0.864	$ -	$ -	$ -	$ -	$ -																
5	0.823	$ 13,471	$ 13,471	$ -	$ -	$ -	$ -															
6	0.784	$ 31,146	$ 18,316	$ 12,829	$ -	$ -	$ -	$ -														
7	0.746	$ 64,179	$ 34,516	$ 17,444	$ 12,218	$ -	$ -	$ -	$ -													
8	0.711	$ 93,540	$ 32,417	$ 32,873	$ 16,614	$ 11,636	$ -	$ -	$ -	$ -												
9	0.677	$ 129,791	$ 40,706	$ 30,874	$ 31,307	$ 15,822	$ 11,082	$ -	$ -	$ -	$ -											
10	0.645	$ 160,846	$ 37,235	$ 38,767	$ 29,403	$ 29,817	$ 15,069	$ 10,555	$ -	$ -	$ -	$ -										
11	0.614	$ 193,895	$ 40,708	$ 35,462	$ 36,921	$ 28,003	$ 28,397	$ 14,351	$ 10,052	$ -	$ -	$ -	$ -									
12	0.585	$ 224,482	$ 39,821	$ 38,769	$ 33,774	$ 35,163	$ 26,670	$ 27,045	$ 13,668	$ 9,573	$ -	$ -	$ -	$ -								
13	0.557	$ 253,999	$ 40,206	$ 37,924	$ 36,923	$ 32,165	$ 33,489	$ 25,400	$ 25,757	$ 13,017	$ 9,117	$ -	$ -	$ -	$ -							
14	0.530	$ 280,564	$ 38,660	$ 38,292	$ 36,119	$ 35,165	$ 30,634	$ 31,894	$ 24,190	$ 24,530	$ 12,397	$ 8,683	$ -	$ -	$ -	$ -						
15	0.505	$ 305,233	$ 38,030	$ 36,819	$ 36,468	$ 34,399	$ 33,490	$ 29,175	$ 30,375	$ 23,038	$ 23,362	$ 11,807	$ 8,270	$ -	$ -	$ -	$ -					
16	0.481	$ 327,520	$ 36,822	$ 36,219	$ 35,066	$ 34,732	$ 32,761	$ 31,896	$ 27,786	$ 28,929	$ 21,941	$ 22,250	$ 11,245	$ 7,876	$ -	$ -	$ -	$ -				
17	0.458	$ 347,550	$ 35,626	$ 35,068	$ 34,494	$ 33,396	$ 33,078	$ 31,201	$ 30,377	$ 26,462	$ 27,551	$ 20,896	$ 21,190	$ 10,709	$ 7,501	$ -	$ -	$ -	$ -			
18	0.436	$ 365,231	$ 34,231	$ 33,929	$ 33,399	$ 32,852	$ 31,806	$ 31,503	$ 29,715	$ 28,930	$ 25,202	$ 26,239	$ 19,901	$ 20,181	$ 10,199	$ 7,144	$ -	$ -	$ -	$ -		
19	0.416	$ 380,761	$ 32,922	$ 32,601	$ 32,314	$ 31,808	$ 31,287	$ 30,291	$ 30,002	$ 28,300	$ 27,553	$ 24,002	$ 24,990	$ 18,954	$ 19,220	$ 9,714	$ 6,804	$ -	$ -	$ -	$ -	
20	0.396	$ 394,206	$ 31,576	$ 31,355	$ 31,049	$ 30,775	$ 30,293	$ 29,797	$ 28,849	$ 28,574	$ 26,952	$ 26,241	$ 22,859	$ 23,800	$ 18,051	$ 18,305	$ 9,251	$ 6,480	$ -	$ -	$ -	$ -
		$ 3,566,413	$ 545,263	$ 489,226	$ 436,068	$ 385,733	$ 338,055	$ 293,106	$ 250,770	$ 211,354	$ 174,077	$ 140,118	$ 108,455	$ 81,520	$ 54,971	$ 35,162	$ 16,055	$ 6,480	$ -	$ -	$ -	$ -

In the following Table we see the geneflow expressions (undiscounted) adding up to the same total. At the end, the equilibrium values as $N.(Sm+Sf)(DE_m + DE_f)$ where DE is a discounted expression

Exercise Day3

	calculations based on GFLOW				
	sire selection	dam selection			
superiority	31.5800	6.1160		GFLOW	
year	Expr_SS	Expr_DS	cum benefit		disc retruns
1	0.000	0.000	$ -	$	-
2	0.000	0.000	$ -	$	-
3	0.000	0.000	$ -	$	-
4	0.000	0.000	$ -	$	-
5	0.015	0.016	$ 16,374	$	13,471
6	0.021	0.028	$ 39,751	$	31,146
7	0.043	0.044	$ 86,006	$	64,179
8	0.042	0.045	$ 131,620	$	93,540
9	0.057	0.053	$ 191,761	$	129,791
10	0.053	0.057	$ 249,525	$	160,846
11	0.062	0.062	$ 315,834	$	193,895
12	0.063	0.064	$ 383,941	$	224,482
13	0.067	0.067	$ 456,146	$	253,999
14	0.068	0.069	$ 529,044	$	280,564
15	0.070	0.070	$ 604,341	$	305,233
16	0.071	0.072	$ 680,891	$	327,520
17	0.072	0.072	$ 758,657	$	347,550
18	0.073	0.073	$ 837,116	$	365,231
19	0.074	0.074	$ 916,348	$	380,761
20	0.074	0.074	$ 996,138	$	394,206
			NPV	$	3,566,413

Exercise Day3

The expression of each pathway converges to 1/(Lm+Lf) = 0.16. However, the values are halved as expression is ny in males, and the equilibrium I not fully reached after 20 years, as we can see in the GFLOW spreadsheet. The value is 0.074 which is a bit less than half of 0.16.
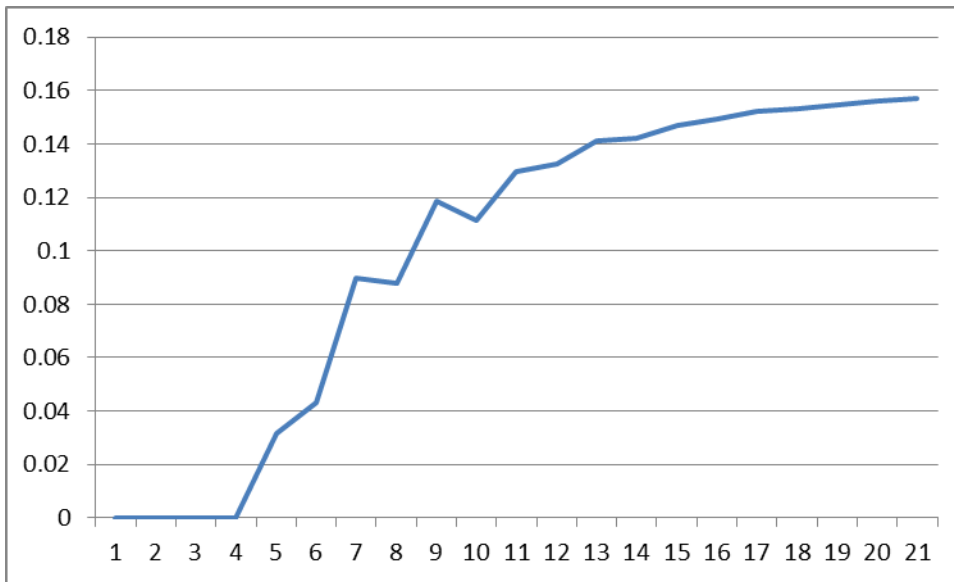
Comparing the dG method with GFLOW gives
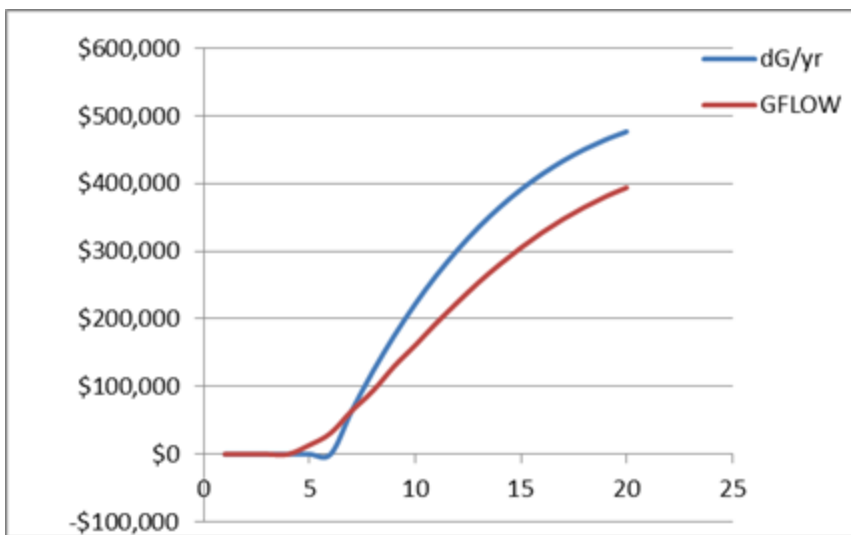
See the spreadsheet GENEFLOWanddGv2.xls

CONCLUSION: Gflow adds up all benefits which is similar to adding up dG per year, accumulating it and multiplying it by the number of animals expressing it. Gflow does a better job in modeling the initial la, which is especialy important in the lag (generation interval) is rather long

Exercise Day3