

Across Country Comparison of Dairy Cattle

Lawrence R. Schaeffer

2021

©L. R. Schaeffer, 2021

Contents

1	Introduction	5
2	Conversion Methods	7
2.1	Example	8
2.1.1	Predicting Bulls From A in B and C	10
2.1.2	Misuses of Prediction Equations	12
2.1.3	Weighted Regressions	14
2.2	References	17
3	Interbull	19
3.1	Topics of Workshops	20
4	Linear Model of 1984	23
4.1	Example	26
4.2	Application to Real Data	30
4.2.1	Rozzi et al. 1990	30
4.2.2	Banos et al. 1991	30
4.3	References	30

5	MACE - 1993	33
5.1	The Model	34
5.2	Example Calculations	35
5.3	Application	37
5.4	References	39
6	Joint Data Evaluations	41
6.1	Jamrozik et al. 2002	42
6.2	Barrett et al. 2005	42
6.3	References	43
7	The Rise of Genomics	45
7.1	Using SNP Genotypes	46
7.2	SNP Effects Model	46
7.3	International Considerations	51
7.4	References	53
8	Competition Model	55
8.1	Introduction	55
8.2	Model and Methods	56
8.3	Discussion	62
8.4	References	64

Chapter 1

Introduction

Cattle were domesticated in Africa between 8000 and 6000 BC. Thus, they have been a food source for the human population for many centuries. How long have humans been selectively breeding cattle to produce progeny with desirable characteristics is unknown. There were 3,000 different species identified in 1985 by FAO. Thus, humans have had an influence for many centuries. Some species of cattle have become extinct too.

Robert Bakewell (1725-1795) was an early pioneer in animal breeding creating the Longhorn cattle breed as well as a horse and sheep breed too. He was known for matings that created high degrees of inbreeding. His attempts were to consolidate his desired traits into all of his animals. He did progeny testing to rank his bulls. He hired out his bulls, for a fee, to other breeders. Herdbooks appeared after Bakewell. Dutch Friesians began in 1875. Herdbooks only recorded the pedigrees, but later on some began recording conformation and production data.

The Holstein-Friesian breed of dairy cattle originated in the Netherlands (Holland). The first established herd in the United States appeared in 1869, and the first herd in Ontario, Canada was founded in 1881 (LEWINGTON, 1983). The United States placed much emphasis on improved milk production and were successful at it due to the work of J. L. Lush and his many students during the 1930's and 1940's at Iowa State University. Statistical analyses became popular for dairy cattle milk production through W. R. Harvey and C. R. Henderson. Lush's students were also from Europe, for example, Harold

Skjervold (Norway), Alan Robertson, Charles Smith, Franz Pirchener, R. D. Politiek, and others. Ideas and practices were spread quickly around the world. Lush's techniques were copied and applied primarily to dairy cattle.

The success of US breeding programs meant that in the 1960's US Holstein cattle were being exported back to the Netherlands, Europe, and elsewhere in the world. Canadian cattle, bred more for conformation, also started to be exported around the globe.

Although the Holstein breed composes 90% or more of the cattle in North America, many red breeds are the majority in other countries, like Sweden and Norway. However, the Holstein breeds drove the need for international sire comparisons. The US wanted to sell semen, and European buyers wanted to compare US bull potential to their populations. Due to increased movement of Holstein genetics during the 1970's from North America to other parts of the world, it became important to be able to compare bulls from the US and Canada to bulls from the importing countries. One problem was that each country had its own systems of milk recording and genetic evaluations, and more importantly, each country had different standards and methods of expressing EBVs of bulls. Importers of bull semen faced the challenge of selecting sires from several exporting countries. Producers understood their own country's EBV system, but did not know or trust the EBV system in other countries, and therefore, they did not know how to rank foreign bulls compared to their own. By the same token, it was important for the semen exporter to make sure their bulls were ranked highly, by some means, in the importing countries or that producers in the importing country knew how to interpret foreign EBVs.

These notes outline the evolution of techniques to compare dairy sires and cows across countries from the 1970's to the present. While there was some scientific effort behind the scenes, the comparison procedures were driven more by market pressures than by academics. The techniques described are based on many assumptions, and most have been violated, but tolerated or ignored. With so many countries involved in dairy cattle exchange, agreement among them has been difficult to achieve.

Lastly, many people have contributed to dairy cattle comparisons, but the most consistent of those has been Dr Jan Philipsson of Sweden. Without his dedication, the chaos may have turned into mayhem.

Chapter 2

Conversion Methods

In 1981 the International Dairy Federation (IDF) sanctioned the use of prediction equations to convert a sire's genetic merit assessed in one country to the genetic base and scale of another country. The approach was to use a simple regression model,

$$EBV_I = a_I + b_I(EBV_E) + e$$

where EBV_I is the EBV in the Importing country, and EBV_E is the EBV in the Exporting country. Thus, there had to be a number of bulls that had enough daughters in both countries in order to obtain estimates of a_I , the intercept, and b_I , the slope. Then those a_I and b_I values were applied to all EBV_E to convert them to the Importing country mean and scale.

Assumptions are always part of a model. In this case the following assumptions were implied.

1. Bulls should rank identically in both countries. That means there does not exist a genotype by environment (GxE) interaction.
2. EBVs in both the importing and exporting countries are assumed to be unbiased, from random mating to cows in each country, and a good random sample of daughters from those cows.
3. EBVs in both countries should be accurate, meaning they were based on a large number of progeny. 'Large number' was never defined, but

was taken to mean 10, 20, or 50 daughters or more, depending on the countries involved.

4. A large number of bulls should have EBVs in both countries. Again, large was never defined, but was understood to be 10 or more.

These simple assumptions were hardly ever valid. GxE interactions were thought to be likely, and thus, bulls do not necessarily have to rank the same in both countries. Imported bulls, due to the generally higher prices for their semen, were never randomly mated to cows in the importing country. The daughters that survived to make lactation records were generally the better daughters, or daughters which were in well managed herds. Thus, EBVs in the importing country were hardly unbiased and often not accurate. In genetic evaluation systems imported bulls were often treated differently from indigenous bulls, and could be lumped together as ‘foreign’. The foreign group could have bulls from several countries. Thus, EBVs in the importing country could be highly biased upwards or downwards depending on how they were treated in the genetic evaluation system, and what other bulls were being imported at that time.

In essence, the simple regression model was too simplistic to be of good value, but it was the easiest to apply, and there were no other proposals, at the time.

2.1 Example

Below is a table of EBVs of 30 bulls from country A that also had daughters in countries B and C. GxE interactions were built into the true breeding values.

Table 2.1 Example Bull EBVs.

Bull	Country A		Country B		Country C	
	N_A	EBV_A	N_B	EBV_B	N_C	EBV_C
1	63	52	42	97	18	37
2	94	42	36	119	16	34
3	45	38	12	87	35	27
4	21	36	49	88	13	22
5	65	34	34	98	19	15
6	20	34	33	52	24	22
7	30	34	13	90	37	21
8	35	28	21	63	30	23
9	27	28	22	70	17	16
10	22	26	45	34	30	4
11	39	26	20	64	12	14
12	36	26	36	71	34	10
13	23	25	14	44	34	15
14	35	24	46	48	19	11
15	25	24	11	45	13	5
16	81	23	14	51	26	17
17	64	22	28	63	34	16
18	35	21	49	32	15	9
19	53	21	26	11	17	6
20	22	20	38	32	21	6
21	46	19	12	17	38	8
22	34	17	34	7	22	8
23	30	17	31	52	23	3
24	53	15	34	32	19	11
25	33	14	40	78	20	11
26	94	13	44	35	14	8
27	77	13	27	27	22	11
28	54	13	29	8	39	10
29	26	13	16	33	38	17
30	44	13	18	59	21	16

Please note that

- the number of daughters in each country is different,

- the EBVs do not rank the bulls in the same order in each country, and
- the mean and variance of the EBVs is different for each country.

These features are typical of international EBVs. The methods of recording milk yields could differ, and the models used for genetic evaluation could differ from country to country. The differences suggest that EBVs of bulls should not be compared across countries. Economics, however, and the desire to sell and buy semen from superior bulls wherever they originate mean that comparisons will inevitably be made, and so the best method possible should be used.

Politics come to bear on the methods, as every country wants their bulls to compare favourably with those of other countries. Another issue that comes into play are the health regulations involved in moving bull semen across borders. The best bulls are not always permitted to be sold in some countries. Health restrictions could limit some countries from selling any bull semen, and as such, act as a trading sanction. Sometimes a bull from country A might be banned from being sold in Country B, but might be sold to country C. Then the same bull's semen might be able to move into country B from country C.

2.1.1 Predicting Bulls From A in B and C

Using the data from Table 2.1, select bulls from country A with 20 or more daughters, then include them if they have 10 or more daughters in country B. That gives 30 bulls. Regress country B EBVs on country A EBVs. The resulting prediction equation is

$$PEBV_{B,A} = -0.5157 + 2.2195(EBV_A).$$

Bull 22, for example, would have

$$PEBV_{B,A} = -0.5157 + 2.2195(+17) = 37.22.$$

This is greater than the actual proof in country B of +7. The equation would be applied to all bulls from country A that were eligible to be imported to country B even though they were not included in the calculations to derive the prediction equation. The converted PEBV would rank the bulls of country A

in exactly the same order as in country A. The R^2 was 0.56. Most prediction equations had an R^2 of less than 0.50, in practice. The greater the R^2 the better are the predictions.

Using the same criteria of 20 or more daughters in country A and 10 or more in country C, the regression equation was based on 30 bulls.

$$PEBV_{C,A} = -1.7006 + 0.6621(EBV_A)$$

Using Bull 22 again,

$$PEBV_{C,A} = -1.7006 + 0.6621(+17) = 9.56,$$

with an R^2 of 0.60.

Equations could be derived in the other direction too. To predict country A EBV from country B, require all bulls in country B to have 20 daughters and all those in country A to have at least 10. That gives 22 bulls for deriving the prediction equation.

$$PEBV_{A,B} = 11.2928 + 0.2486(EBV_B),$$

with an $R^2 = 0.56$.

Similarly, for country C,

$$PEBV_{A,C} = 12.4997 + 0.7009(EBV_C),$$

with an $R^2 = 0.36$ based on 17 bulls.

Finally, predictions between countries B and C based on bulls of country A used in both countries.

$$PEBV_{C,B} = 2.5771 + 0.2119(EBV_B),$$

with an $R^2 = 0.53$ based on 22 bulls, and

$$PEBV_{B,C} = 19.3060 + 2.0821(EBV_C),$$

with an $R^2 = 0.33$ based on 17 bulls.

The question becomes which country or organization should derive the prediction equations. Each country would include only the bulls that give their country's bulls the greatest advantage in the importing country. However, dairy producers in the importing country may not believe the predictions they receive from the exporting country. Thus, they want their scientists to derive the prediction equations and to perform the conversions. Dairy producers were soon faced with many converted EBVs and did not trust any of them. Some producers had traveled to see the daughters and records of bulls in their country of origin, and only then would they decide to buy semen from foreign bulls. There was a great deal of uncertainty and conflict during 1970 through 1980's. Even so, USA bull semen was widely exported around the world, being known for their high milk producing daughters. Around 1990, Holstein bulls from the Netherlands, Italy, and Germany started to be used in Canada and the USA.

2.1.2 Misuses of Prediction Equations

Reciprocal Property

Some producers believed that if EBVs of country A were converted to their country's scale, then applying the prediction equation to convert their country's EBVs to those of country A, then one should get back the EBV from country A.

For example, converting Bull 22 from country A to country B gave

$$PEBV_{B,A} = -0.5157 + 2.2195(+17) = 37.22.$$

The prediction equation to go from B to A was

$$PEBV_{A,B} = 11.2928 + 0.2486(EBV_B),$$

then for bull 22,

$$PEBV_{A,B} = 11.2928 + 0.2486(37.22) = 20.55.$$

This is not equal to the actual EBV in country A of +17. Obviously, the prediction equations can not be used in both directions because of many reasons such as, different and few bulls in the derivation of prediction equations, the low R^2 of the prediction equations, and the possible existence of GxE interactions.

Mathematically, if

$$\begin{aligned} PEBV_{B,A} &= a_B + b_{BA} \cdot EBV_A \\ PEBV_{A,B} &= a_A + b_{AB} \cdot EBV_B \end{aligned}$$

when you try to convert $PEBV_{B,A}$ back to $PEBV_{A,B}$, you get

$$PEBV_{A,B} = a_A + b_{AB} \cdot (a_B + b_{BA} \cdot EBV_A)$$

which only works if

$$a_A + b_{AB} \cdot a_B = 0$$

and if

$$b_{AB} \times b_{BA} = 1$$

. In the above example,

$$\begin{aligned} a_A + b_{AB} \cdot a_B &= 11.2928 + 0.2486(-0.5157) \\ &= 11.16, \end{aligned}$$

which is not zero, and

$$b_{BA} * b_{BA} = 2.2195 \cdot 0.2486 = 0.5517,$$

which is not 1. Clearly the prediction equations are not reciprocal, and therefore, should not be used to predict based on a predicted EBV back into the country of origin. In general, this should not be attempted.

Third Country Predictions

Suppose there was Country D with no bulls in common with Country A, thus, there could be no prediction equation to convert EBV_A into $PEBV_{D,A}$. Suppose there was a prediction equation between Country C to D, let it be

$$PEBV_{D,C} = 6.34 + 1.5(EBV_C).$$

The equation to convert EBV_A to $PEBV_{C,A}$ was

$$PEBV_{C,A} = -1.7006 + 0.6621(EBV_A).$$

Then,

$$\begin{aligned} PEBV_{D,CA} &= 6.34 + 1.5(-1.7006 + 0.6621(EBV_A)) \\ &= (6.34 + 1.5(-1.7006)) + 1.5(0.6621)(EBV_A) \\ &= 3.7891 + 0.9931(EBV_A). \end{aligned}$$

There is no R^2 that can be computed, and similar to the reciprocal property, the prediction equation is probably very biased and subject to errors. The problem could become more complicated if there was a prediction equation from Country B to Country D, then another prediction equation could be derived for A to D.

2.1.3 Weighted Regressions

One problem in the prediction equations was that each bull's EBV were assumed to be known quantities of fixed constant value. Unfortunately, EBVs have variable reliability based on number of daughters. The more daughters there are, the more accurate the EBV. If the number of daughters could be taken into account, then perhaps all of the bulls with proofs in each country could be used in the prediction equation calculations. This was the reason for the WILMINK (1987) and GODDARD (1987) proposed regression models.

Wilmink's Method

A simple regression is

$$Y = a + b \cdot X + e,$$

where Y and X are assumed to be known without error, but in our context both Y and X have errors associated with them. Wilmink decided to weight the exporting EBV by their reliability, then to estimate the regression of importing EBV on exporting EBV, and finally to use that regression to estimate the intercept. The R-script for this is as follows:

```

# Adjust EBV in exporting country for mean,
# weight by reliability (repA)
ETAA=(BVA - meanA)*repA
Y=BVB
X=ETAA
M=X*0 + 1
W=cbind(M, X)
WW=t(W)%*%W
WY=t(W)%*%Y
C=ginv(WW)
ahat=C%*%WY
# Estimate the intercept
ahatBA=meanB - ahat[2]*meanA
ahat[1]=ahatBA
yhat=W%*%ahat
p=cor(yhat,Y)
p*p

```

The resulting formula is

$$PEBV_{B,A} = -17.6 + 2.9207(EBV_A),$$

with an $R^2 = 0.57$.

Goddard's Method

Goddard's method chooses to weight the importing country's EBVs by their reliabilities, rather than the exporting country's.

```

# Weight Importing Country EBV
ETAB = (BVB - meanB)/repB
Y=ETAB
X=BVA
M=X*0+1
W=cbind(M,X)
WW=t(W)%*%W
WY=t(W)%*%Y

```

```

C=ginv(WW)
ahatG=C%*%WY
# Estimate the intercept
ahatBA=meanB - ahatG[2]*meanA
ahatG[1]=ahatBA
yhat=W%*%ahatG
p=cor(yhat,Y)
p*p

```

The resulting formula is

$$PEBV_{B,A} = -14.61 + 2.7980(EBV_A),$$

with an $R^2 = 0.55$.

Weighted LS Method

Weighted least squares weights both sides of the regression equation for the number of daughters of the importing country EBV. The assumption is made that the exporting country EBV do not have any error.

```

Y=BVB
X=BVA
M=X*0+1
W=cbind(M,X)
DI=diag(nB)
WW=t(W)%*%DI%*%W
WY=t(W)%*%DI%*%Y
C=ginv(WW)
ahatR=C%*%WY
yhat=W%*%ahatR
p=cor(yhat,Y)
p*p

```

The resulting formula is

$$PEBV_{B,A} = 1.2521 + 2.1427(EBV_A),$$

with an $R^2 = 0.56$.

Other methods were also put forward. In essence, the prediction equation market was wide open and chaotic. Everyone was free to produce their own prediction equation. Some order had to be imposed.

2.2 References

- GODDARD, M. E.** 1985. A method of comparing sires evaluated in different countries. *Livest. Prod. Sci.* 13:321-331.
- WILMINK, J. B. M.** , A. MEIJERING, B. ENGEL. 1986. Conversion of breeding values for foreign populations. *Livest. Prod. Sci.* 14:223-229.

Chapter 3

Interbull

In 1983 a permanent subcommittee of ICAR (International Committee on Animal Recording) was established with the help of the European Association of Animal Production (EAAP) and the International Dairy Federation (IDF). The goal was to obtain some control over the comparison of dairy cattle across countries by establishing standards that all countries would agree to abide by. The leading pioneers of this movement were a group of scientists from several main countries. They were Josef Lederer (Austria), Edward Burnside (Canada), Claude Gaillard (Switzerland), Hans Gravert (Germany), Hermann Schulte-Coerne (Germany), Lars Gjol Christensen (Denmark), Jean Claude Mocquot (France), Sandy McClintock (Great Britain), Patrick Cunningham (Ireland), Jan Dommerholt (Netherlands), Erling Fimland (Norway), and Jan Philipsson (Sweden).

A home office was established in Sweden because of its neutrality in exporting and importing of Holstein cattle, and thus it has remained in Uppsala. Today it has a Steering Committee composed of members from the main exporting countries which monitors the progress of the Interbull Centre. Additionally, there is a scientific advisory committee and a Technical Committee which conduct workshops on new developments and research within member countries. In addition there are annual Business Meetings. The workshops and Business meetings are held in different countries from year to year.

In the beginning the workshops attracted around twenty people, and discussions took one day or two at the most. Now the workshops have up to 200

attendees and sometimes more than two days of presentations. Not all of the presentations are concerned with international dairy comparisons, but include the latest developments in dairy cattle genetics research.

3.1 Topics of Workshops

In 1993 Arhus, Denmark, Jan Philipsson presented the projects and publications from INTERBULL during 1992-1993. The projects were

1. Survey on sire evaluation procedures for production traits (28 countries) and non-production traits (in 19 countries).
2. COPA/EC project on ranking Black and White bulls in Europe.
3. NBC project on ranking Nordic bulls.
4. Feasibility study on global ranking of Guernsey bulls (World Guernsey Cattle Federation).
5. Feasibility study on global ranking of Ayrshire bulls.
6. Research on measures of connectedness between populations.

There were 10 publications related to the above projects from members of the INTERBULL staff. The following presentations were made at the 1993 meeting. This should give you a good idea of the truly international flavour of these meetings.

SCHAEFFER, L. R. Canada. Multi-trait, across country evaluation of dairy sires.

BANOS, G. INTERBULL Centre. Linear model comparisons of Black-and-White dairy bulls from Nordic countries.

POWELL, R. USA. Effect of direction of gene flow on conversion equations.

BONAITI, B. France. Problems arising with genetic trend estimation in dairy cattle.

- MATTALIA, S.** France. Use of full-sib families to estimate the a- coefficients of conversion formulas between countries.
- ROBINSON, J. A. B.** Canada. Accuracy of conversion of proofs using Wilmink's and Goddard's methods.
- DIERS, H.** Germany. European and world-wide harmonization of linear type classification - Definition of traits and estimation of breeding values.
- LOHUIS, M.** The possibilities of producing reciprocal conversion formulae for conformation traits.
- HAMMOND, K.** FAO. Conservation and evaluation of indigenous breeds of cattle.
- GABOR, G.** Hungary. Physiological opportunity in selection of breeding bulls for reproduction traits.
- BEARD, K.** Australia. Genetic evaluation for milking speed, temperament, likeability and survival in Australia.
- BAGNATO, A.** Italy. Herdlife in the Italian Holstein Friesian.
- POWELL, R.** USA. Genetic evaluations for somatic cell score.
- REENTS, R.** Germany. Estimation of breeding values for type traits in Germany.
- CASANOVA, L.** Switzerland. Genetic evaluation of linear type traits for Swiss Braunvieh.
- PHILIPSSON, J.** Sweden. Use of total merit index in bull selection.
- GEORGOUDIS, A.** Greece. Objectives and present activities in the BovMap - The Bovine Gene Mapping Project.
- LOHUIS, M.** Canada. The effect of continuous evaluation on genetic response in progeny testing programs.
- CARABANO, M.** Spain. A study on heterogeneity of variances adjustment in genetic evaluations in Spain.
- SWALVE, H.** Germany. Genetic parameters for test-day models.

LEROY, P. Belgium. Estimation of breeding values of Belgian Black and White, Red and White, White Red, and Red of Flanders breeds with the animal model.

EZRA, E. Israel. Animal model in Israel.

DJEMALI, M. Tunisia. Genetic parameters of milk yield and days in milk of Friesian cattle in Tunisia.

ABE, T. Japan. Recent developments in the progeny testing of dairy bulls in Japan.

Chapter 4

Linear Model of 1984

In 1984, the first Interbull Workshop took place in Prague, Czechoslovakia. One of the limitations of regression formulas for predicting EBV in other countries was that only two countries were involved at a time. Was it possible to find a method that used EBVs from several countries simultaneously to derive an international ranking? I proposed the following linear model method at this meeting. The equation of the model was

$$\mathbf{y} = \mathbf{Xc} + \mathbf{ZQg} + \mathbf{Zs} + \mathbf{e},$$

where

- The observation vector is

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_{USA} \\ \mathbf{y}_{CAN} \\ \mathbf{y}_{NLD} \\ \mathbf{y}_{FRA} \end{pmatrix}$$

is an array of de-regressed sire daughter averages from each country going into the analysis. Let the number of countries be n_c . Some bulls would appear in more than one country.

- \mathbf{c} is an $n_c \times 1$ vector of country means for the bulls in \mathbf{y} .
- \mathbf{g} is a vector of genetic groups indicating country and year of birth of each bull. \mathbf{Q} would be a matrix indicating the group (or fraction of groups)

to which each bull belongs. For example, if there were three countries and 3 genetic groups,

$$\mathbf{g} = \begin{pmatrix} g_A \\ g_B \\ g_C \end{pmatrix},$$

then if a bull was born in country A, then the corresponding row of \mathbf{Q} for bull j would be

$$\mathbf{q}'_j = (1 \ 0 \ 0).$$

If a bull's sire was born in country B and the dam born in country C, then

$$\mathbf{q}'_j = (0 \ 0.5 \ 0.5).$$

- \mathbf{s} is a vector of length equal to the total number of bulls with data in \mathbf{y} and also ancestor sires.
- \mathbf{e} is a vector of residual effects with a different residual variance for each country divided by the number of daughters for that bull in that country.

Also,

$$E \begin{pmatrix} \mathbf{s} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix},$$

and

$$Var \begin{pmatrix} \mathbf{s} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{A}\sigma_s^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}\sigma_e^2 \end{pmatrix},$$

where the ratio of residual to sire variances is assumed constant over countries and scaled to one common residual variance, σ_e^2 . Thus, \mathbf{D} is diagonal with elements equal to

$$(\sigma_{ec}^2/\sigma_e^2)/n_{ci},$$

where n_{ci} is the number of daughters of bull i in country c , and σ_{ec}^2 is the residual variance for country c . \mathbf{A} is a sire additive genetic relationship matrix built using sire and maternal grandsire pedigree information, if available.

A big problem at this time was the unique identification of dairy bulls. When bulls were exported to another country they often lost their country of origin ID. Thus, one bull could have more than one ID depending how

many countries its semen had been exported. Even the regression model work required unique identification. ICAR eventually came up with a system where each animal has letters to describe the breed, such as HO, AY, or GU, for example, and three letters to indicate the country of birth, e.g. USA, CAN, GDR, NLD, or ITA. Another letter would indicate gender of the animal, M or F, followed by up to 16 numbers for their country of birth ID.

Some assumptions were

1. All ETA (estimated transmitting abilities) within one country had to have the same genetic base. That means that the ETA had to be the latest ones available.
2. ETAs had to be generated by an accurate method of genetic evaluation which accounted for the major factors within a country, such as BLUP or MCC.
3. All ETAs had to be for the same trait, e.g. milk yield, in kilograms, adjusted to a mature equivalent basis, and based on a single trait methodology. ETAs from a multiple trait method could not be included.
4. Covariances between residual effects in different countries were assumed to be zero. This is logical because daughters will have been raised in completely different environments.
5. Covariances between residual effects within a country were also assumed to be zero. This is not strictly correct because daughters of different sires might be raised together in the same herds. Accounting for the correct structure would be very cumbersome.
6. Countries have to be connected, meaning that each country has to have daughters of bulls from each of the other countries.
7. Heritability was the same in each country for that trait.
8. Heterosis from crossing lines of Holsteins from different countries was assumed not to exist.
9. Genotype by environment(countries) interactions were assumed to not exist.

10. Sires were randomly mated within countries.
11. No preferential treatment of daughters of any bulls.

ETAs of all bulls evaluated in each country were needed for this model, and not just the bulls that had daughters in both countries.

4.1 Example

Consider the data in Table 4.1 to illustrate the linear model approach. The example comes from SCHAEFFER (1985).

Table 4.1 Example data for Linear Model method.

Country of proof	Bull ID	Sire ID	Country of birth	Year of birth	ETA (kg)	n_{ci}	$(n+k)/n$	y_{ci}
A	1	77	A	1980	0	100	1.1500	0
A	2	88	A	1980	+140	60	1.2500	+175
A	4	77	A	1981	-210	70	1.2143	-255
A	6	88	B	1980	+580	80	1.1873	+689
B	1	77	A	1980	+40	20	1.7500	+70
B	3	82	A	1981	-470	150	1.1000	-517
B	5	60	B	1981	+70	100	1.1500	+80
B	6	88	B	1980	+625	200	1.0750	+666
B	7	77	B	1981	-280	60	1.6250	-350

The last column of Table 4.1 are the de-regressed ETA, namely equal to

$$y_{ci} = ETA_{ci} * (n+k)/n,$$

where $k = 15$ for both countries. There are 4 genetic groups, i.e. A-80, A-81, B-80, and B-81.

Assume that the residual variances had been estimated and were 100,000 in country A and 134,000 in country B. That gives an average of 117,000. So the residuals can be scaled to the average. In country A divide 100,000 by 117,000 = .8547, and in country B divide 134,000 by 117,000 = 1.1453.

The R-script for the example data is as follows:

```

resA=100000
resB=134000
ares = (resA + resB)/2
Ra=c(100, 60, 70, 80, 0, 0, 0, 0, 0) #/(resA/ares)
Rb=c(0, 0, 0, 0, 20, 150, 100, 200, 60) #/(resB/ares)
Ri=Ra+Rb
RI=diag(Ri)
# Design matrix for country of proof effects
X=matrix(data=c(1,0, 1,0, 1,0, 1,0, 0,1, 0,1,
               0,1, 0,1, 0,1),
         byrow=TRUE,ncol=2)
# Assigning sires to genetic groups
Q=matrix(data=c(1,1,0,0,0,0,0,0, 0,0,0,0,0,
               0,0,1,1,0,0,0,0, 0,0,0,0,0,
               0,0,0,0,0,0,1,0, 0,0,0,0,0,
               0,0,0,0,0,1,0,1,0,0,0,0),byrow=TRUE,ncol=11)
Q=t(Q)
Z=matrix(data=c(1,0,0,0,0,0,0,0,0,0,0,0,
               0,1,0,0,0,0,0,0,0,0,0,0,
               0,0,0,1,0,0,0,0,0,0,0,0,
               0,0,0,0,0,1,0,0,0,0,0,0,
               1,0,0,0,0,0,0,0,0,0,0,0,
               0,0,1,0,0,0,0,0,0,0,0,0,
               0,0,0,0,1,0,0,0,0,0,0,0,
               0,0,0,0,0,1,0,0,0,0,0,0,
               0,0,0,0,0,0,1,0,0,0,0,0,
               0,0,0,0,0,0,0,1,0,0,0,0),byrow=TRUE,ncol=11)
ZQ=Z%*%Q

```

The relationship matrix inverse is based only on the sire of the bull in this example, but in practice one should use sire and maternal grandsire, and perhaps maternal granddam too. Henderson's rules for this example would be to add $4/3$ to the diagonal for bulls with the sire known, $-2/3$ to the off-diagonal between the bull and its sire, and $1/3$ to the diagonal of the sire. If the sire is unknown, then add 1 to the bull's diagonal. The result is shown below for bulls 1 to 7 followed by 77, 88, 82, and 60.

```

AI=matrix(data=c(4,0,0,0,0,0,0,-2,0,0,0,0,

```

```

0,4,0,0,0,0,0,0,-2,0,0,
0,0,4,0,0,0,0,0,-2,0,
0,0,0,4,0,0,0,-2,0,0,
0,0,0,0,4,0,0,0,0,-2,
0,0,0,0,0,4,0,0,-2,0,0,
0,0,0,0,0,0,4,-2,0,0,0,
-2,0,0,-2,0,0,-2,6,0,0,0,
0,-2,0,0,0,-2,0,0,5,0,0,
0,0,-2,0,0,0,0,0,0,4,0,
0,0,0,0,-2,0,0,0,0,0,4),byrow=TRUE,ncol=11)
AI=AI/3
AIk=AI*15 # k = 15 assumed for both countries. h2=0.25
AIk

```

The mixed model equations are formed as follows:

```

Y=matrix(data=c(0,175,-255,689,70,-517,80,666,-350),ncol=1)
W=cbind(X,ZQ,Z)
WRW=t(W)%*%RI%*%W
WRY=t(W)%*%RI%*%Y
BD=c(0,0,0,0,0,0)
BDi=diag(BD)
BC=matrix(data=c(0),nrow=6,ncol=11)
HIa=cbind(BDi,BC)
HIb=cbind(t(BC),AIk)
HI=rbind(HIa,HIb)
MME=WRW+HI
# removing equation 3, restriction on solutions
ka=c(1,2,4:17)
MMEr=MME[ka,ka]
RHS=WRY[ka, ]
C=ginv(MMEr)
soln=C%*%RHS

```

The solutions are given in Table 4.2.

Table 4.2 Solutions to Example Data
From Mixed Model Equations.

	Solution		Solution
\hat{c}_A	97.58702	\hat{s}_1	-77.33863
\hat{c}_B	66.46312	\hat{s}_2	63.78393
\hat{g}_{A80}	0.00000	\hat{s}_3	-127.15696
\hat{g}_{A81}	-441.74288	\hat{s}_4	65.86704
\hat{g}_{B80}	580.75600	\hat{s}_5	181.37573
\hat{g}_{B81}	-198.99827	\hat{s}_6	15.94598
		\hat{s}_7	-165.53248
		\hat{s}_{77}	-5900135
		\hat{s}_{88}	31.89197
		\hat{s}_{82}	-63.57848
		\hat{s}_{60}	90.68787

To obtain ETAs, add together the genetic group solutions to the sire solutions for sires with ETAs in one or more countries.

$$ETA_{LM} = \mathbf{Q}\hat{\mathbf{g}} + \hat{\mathbf{s}}$$

The results are in the last column of Table 4.3.

Table 4.3 Within Country ETAs and Combined ETAs
for Example Data.

Bull	ETA_A	ETA_B	ETA_{LM}
1	0	+40	-77
2	+140	-	+64
3	-	-470	-569
4	-210	-	-376
5	-	+70	-18
6	+580	+625	+597
7	-	-280	-365

The end product is one international ETA for each bull comprised of all the ETAs from each country in which the bull had daughters. Please remember all of the assumptions associated with this model, most of which are violated.

4.2 Application to Real Data

4.2.1 Rozzi et al. 1990

ROZZI et al (1990) applied this model to bulls from Italy, Canada, and the US. Give the basics of the study.

4.2.2 Banos et al. 1991

BANOS et al. (1991) combined bull ETAs for the Ayrshires and the Jerseys between Canada (CAN) and the USA. The first step of the study was to quantify the direct ties (bulls having daughters in both countries) and relationship ties through pedigrees between countries. Relationships among bulls between countries was similar to relationships within countries.

In Ayrshires there were 1,772 bulls evaluated in the USA and 2,229 bulls evaluated in CAN. In Jerseys the numbers were 7,910 in USA and 2,067 in CAN. All bulls were born between 1950 and 1984 and had to have a minimum of 5 daughters. Heritability was assumed to be 0.29 for both breeds, both countries and both traits, i.e. milk yields and fat yields.

Results from the linear model were compared to predictions based on Wilmlink's regression model. The correlations between the two methods was 0.99 for both directions and both traits. The linear model method was preferred if more than two countries were to be involved.

4.3 References

- BANOS, G.** , SCHAEFFER, L. R., BURNSIDE, E. B. 1991. Genetic relationships and linear model comparisons between United States and Canadian Ayrshire and Jersey bull populations. *J. Dairy Sci.*74:1060-1068.
- ROZZI, P.** , L. R. SCHAEFFER, E. B. BURNSIDE, W. SCHLOTE. 1990. International evaluation of Holstein-Friesian dairy sires from three countries. *Livest. Prod. Sci.* 24:15.

- SCHAEFFER, L. R.** 1985. Model for international evaluation of dairy sires. *Livest. Prod. Sci.* 12:105-115.

Chapter 5

MACE - 1993

By 1993 the regression methods for pairs of countries were still being calculated, and countries were still arguing about which prediction equations were correct or fair. Many countries were now involved in exporting bull semen around the world. Semex Canada, with an office in Guelph, put pressure on many scientists in Guelph and around Canada to come up with evidence that one country was being unfair in its regression methodology. Instead of showing someone was not being fair, I tried to think of something that was better than the regression method. The linear model approach of 1984 was not being used by any group of countries or INTERBULL. The existence of GxE interactions was believed to be true by nearly everyone, and neither the regression methods nor the linear model of 1984 acknowledged this belief.

One night in a dream it came to me that GxE interactions meant that bulls do not rank the same in each country, and that milk yield of daughters in different countries were actually different traits. A multiple trait linear sire model was the answer. Estimate the genetic correlations between countries. Obtain genetic evaluations for all bulls in each country. I awoke eager to start working on this idea. A meeting at Semex Canada with their exporters and scientists from Guelph and Agriculture Canada, everyone agreed with my idea. Wencan Zhang of Semex was tasked with helping me to obtain data for a small project. We agreed that the methods had to be tested before presenting them to INTERBULL. Also to give people an idea of how the results will look.

5.1 The Model

The model that describes Holstein sire daughter yield deviations (DYD) for the i^{th} country is

$$\mathbf{y}_i = \mu_i \mathbf{1} + \mathbf{Z}_i \mathbf{Q} \mathbf{g}_i + \mathbf{Z}_i \mathbf{s}_i + \mathbf{e}_i,$$

where

\mathbf{y}_i is the vector of sire average DYD from country i for protein yield.

μ_i is the overall average DYD for country i , which reflects the definition of genetic base in that country.

\mathbf{g}_i is a vector of phantom parent genetic group effects which are defined across countries and by year of birth within country of birth.

\mathbf{s}_i is a vector of sire genetic effects (transmitting abilities) for country i . The length of the vector is equal to the total number of sires (including ancestor sires) in all countries combined.

\mathbf{e}_i is a vector of random mean residual effects (i.e. each residual term is based on a different number of daughters).

\mathbf{Z}_i is the matrix that relates elements of \mathbf{y}_i to elements in \mathbf{s}_i , and

\mathbf{Q} is a matrix that associates sires with their genetic groups. This matrix is the same for each country.

Suppose there are t countries, then

$$Var \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_t \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1 \sigma_{e1}^2 & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \sigma_{e2}^2 & & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{D}_t \sigma_{et}^2 \end{pmatrix},$$

and

$$Var \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_t \end{pmatrix} = \begin{pmatrix} \mathbf{A} \sigma_{s1}^2 & \mathbf{A} \sigma_{s1s2} & & \mathbf{A} \sigma_{s1st} \\ \mathbf{A} \sigma_{s2s1} & \mathbf{A} \sigma_{s2}^2 & & \mathbf{A} \sigma_{s2st} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A} \sigma_{sts1} & \mathbf{A} \sigma_{sts2} & & \mathbf{A} \sigma_{st}^2 \end{pmatrix}.$$

\mathbf{D}_i are diagonal matrices whose diagonals are reflective of the accuracy of the DYD for each sire. This can be a number derived from the reliability of a sire's ETA in country i , or it can be one over the number of daughters in a bull's ETA, or it can be one over the number of effective daughters (takes into account the distribution of daughters in small and larger herds. Whatever is used should be consistent across countries.

The relationship matrix, \mathbf{A} , was initially based on sire-maternal grandsire relationships, but later maternal granddam information was added.

5.2 Example Calculations

Consider two countries, C1 and C2, and the bull proofs given in Table 5.1 for protein yields. The units are not important for the example, but note that ETA have different magnitudes in the two countries.

Table 5.1 Example Bull DYDs.

Bull ID	Country of Proof	Number of daus.	DYD
1	C1	10	56
2	C1	20	-23
3	C1	50	8
1	C2	100	9
4	C2	40	3
5	C2	20	-11

Let

$$\begin{aligned}\sigma_{s1}^2 &= 100 \\ \sigma_{s2}^2 &= 5 \\ \sigma_{s1s2} &= 20 \\ \sigma_{e1}^2 &= 1000 \\ \sigma_{e2}^2 &= 80\end{aligned}$$

which gives heritabilities of 0.36 in C1, and 0.235 in C2, and a genetic correlation of 0.89 between C1 and C2.

The pedigrees and phantom group assignments are in Table 5.2.

Table 5.2 Pedigree for Example Bulls.

Bull	Sire	MGS	MGD	m
1	6	7	P5	0
2	8	9	P5	0
3	10	8	P5	0
4	10	11	P6	0
5	2	6	P6	0
6	P1	P2	P6	5
7	P1	P2	P6	5
8	P1	P2	P6	5
9	P3	P4	P6	5
10	P3	P4	P6	5
11	P3	P4	P6	5

$m = 0$ if both sire and MGS are known,

$m = 1$ if MGS is unknown,

$m = 4$ if sire is unknown, and

$m = 5$ if sire and MGS are unknown.

Henderson's rules for forming the inverse of the relationship matrix in this example are given below. Add the following elements to the appropriate places in the inverse.

	bull	sire	MGS	MGD
bull	x	-.5x	-.25x	-.25x
sire	-.5x	.25x	.125x	.125x
MGS	-.25x	.125x	.0625x	.0625x
MGD	-.25x	.125x	.0625x	.0625x

where x equals $16/(m + 11)$ for the m in Table 5.2.

The solutions for sires from the MME (mixed model equations) are equal to $\mathbf{Q}\hat{\mathbf{g}} + \mathbf{s}$. The results rounded off are given in Table 5.3.

Table 5.3 ETAs for Example Data from MME.

Bull	Country 1	Country 2
1	31	7
2	-27	-6
3	-2	0
4	4	1
5	-29	-7
6	11	2
7	9	2
8	-13	-3
9	-8	-2
10	1	0
11	0	0

5.3 Application

Bull proofs were obtained from Canada (CAN), the USA, Italy (ITA), and the Netherlands (NLD) for spring 1993. A summary is in Table 5.4.

Table 5.4 Number of Bulls.

Country	With DYD	Ancestors
CAN	2494	418
ITA	1607	414
NLD	3269	674
USA	8329	576
Total	15,699	1577

Unknown sires and MGS were assigned to 43 phantom parent groups. Description of the phantom groups are in Table 5.5.

Table 5.5 Phantom Group Information.

Country	Birth Year	Sire Groups		MGS Groups	
		Group	Bulls	Group	Bulls
CAN	-1975	1	93	2	93
	1976-78	3	111	4	134
	1979-81	5	66	6	79
	1982-	7	37	8	38
ITA	1975-77	9	67	10	170
	1978-81	11	13	12	261
	1982-	13	7		
	1982-84			14	107
	1985-			15	12
NLD	-1975	16	161	17	161
	1976-78	18	187	19	187
	1979-81	20	60	21	60
	1982-	22	14	23	14
USA	-1975	24	162	25	162
	1976-78	26	326	27	343
	1979-81	28	170	29	176
	1982-84	30	55	31	55
	1985-	32	10	33	10
DEU	-1981	34	16	35	16
	1982-	36	14	37	14
ISR	1985-86	38	2	39	2
GBR	1976-77	40	3	31	3
DNK,NZL	1982-84	42	4	43	4

The heritabilities used by each country were 0.25 (ITA), 0.25(USA), 0.30(NLD) and 0.33 (CAN). The estimated sire variances are shown in Table 5.6.

Table 5.6 Covariances and Correlations for Sire Effects for Protein Yields. Covariances above diagonal, correlations below.

	CAN	ITA	NLD	USA	Residual
CAN	38.9	71.7	53.4	150.1	427.3
ITA	.973	139.3	103.1	290.2	1946.3
NLD	.906	.924	89.3	227.3	1036.9
USA	.948	.969	.948	644.3	7683.0

Correlations of MACE ETA with within country proofs were all greater than 0.99.

The advantages of MACE are

1. GxE interactions are allowed.
2. Different units of measure between countries are allowed.
3. Additive relationships among all bulls across countries are included.
4. All bulls evaluated can be included for each country.
5. Sire and residual variances can differ between countries.
6. Bulls are evaluated in each country even though they may not have daughters in that country.

Bull proofs within countries are assumed to be accurately evaluated. Methods to validate genetic evaluation systems were derived and tests must be conducted on the systems before the proofs can be included. Each country is responsible for their own validation, but INTERBULL does a check as well. Random matings of bulls within countries is still important, but everyone realizes this may not hold.

5.4 References

- SCHAEFFER, L. R., ZHANG, W. 1993. Multi-trait, Across Country Evaluation of Dairy Sires. INTERBULL Bulletin 8, Arhus, Denmark.
- SCHAEFFER, L. R. 1994. Multiple-Country Comparison of Dairy Sires. J. Dairy Sci. 77:2671-2678.

Chapter 6

Joint Data Evaluations

The MACE approach to international bull comparisons was adopted by INTERBULL in 1995 with about 8 countries included. In 2021, nearly 30 countries participate in MACE. It is striking that MACE has not been replaced with other methods in all that time (26 years), although enhancements have been made over that time.

In 1994 the random regression test-day model was introduced at the World Congress of Genetics Applied to Livestock Production. By 2000 the model was adopted in Canada, and many other countries were researching it. In this model, the records are each cow's individual test day performances during the lactation, of which there are 1 to 9 records per cow. In the USA with the largest Holstein population of any country there are too many test day records to process so that a test day model has not been adopted.

There were questions about the adequacy of MACE, and scientists thought about how to test it out. WEIGEL et al. (2001) combined 305-d lactation yields of cows from 17 countries into one multiple trait sire model. The results were similar to MACE.

6.1 Jamrozik et al. 2002

JAMROZIK et al. (2002a,b) took first lactation test-day records of Holsteins from Australia, New Zealand, Italy, and Canada and performed single trait and multiple trait analyses on those test day records using a random regression test day animal model. Milk production in Canada and Italy is based mainly on intensive management systems. Australia and New Zealand are largely based on grazing. Thus, there was most likely a GxE interaction among the countries even within hemispheres. Estimated genetic correlations in total lactation yield ranged from 0.65 between ITA and NZL to 0.83 between AUS and NZL.

There were 173 sires that had daughters in all four countries. To estimate genetic parameters samples of data were taken, starting with the sires represented in all countries. Between 14,000 and 15,000 test day records per country were sampled. Combined there were 316,967 TD records on 51,489 cows. Bayesian methods with Gibbs sampling were used to estimate posterior means, using a single chain of 105,000 samples and a burn-in period of 10,000.

The same model was assumed for all four countries. Sampling of records limited the scope for modelling as well. Order 4 Legendre polynomials were used for genetic and permanent environmental effects. Twenty nine residual groups were formed. Lastly, the genetic ties between the four countries were low.

Genetic evaluations were calculated using the estimated parameters both with countries and between countries. Correlations between EBV from single and multiple country evaluations were 0.95 for sires and close to 1 for cows. Under the multiple trait model, correlations of bull EBV between countries ranged from 0.93 to 0.96. A multiple country random regression animal model would not be practical for more than four countries as a routine genetic evaluation for international comparisons.

6.2 Barrett et al. 2005

Some breeds, such as the Milking Shorthorn, are small in size of population to make an international evaluation using a multi-country random regression test-day animal model. BARRETT et al. (2005) had 1,018,528 TD records on

68,653 cows of five countries. There were 66 common sires. Genetic correlations between countries were low from 0.08 to 0.46. The genetic ties between countries were lacking in this study. Better connections should be planned and carried out if this model of international comparisons is to be used in this breed. International comparisons of bulls would offer a greater choice of genetic material to producers in each country, and would increase the number of ties. Outside sources of genetics would also reduce levels of inbreeding within countries.

6.3 References

- BARRETT, R.** , MIGLIOR, F., JANSEN, G., JAMROZIK, J., SCHAEFFER, L. R. 2005. Joint international evaluation of Milking Shorthorn dairy cattle for production traits. *J. Dairy Sci.* 88:3326-3336.
- JAMROZIK, J.** , SCHAEFFER, L. R., WEIGEL, K. A. 2002a. Genetic evaluation of bulls and cows with single- and multiple-country test-day models. *J. Dairy Sci.* 85:1617-1622.
- JAMROZIK, J.** , SCHAEFFER, L. R., WEIGEL, K. A. 2002b. Estimates of genetic parameters for single- and multiple-country test-day models. *J. Dairy Sci.* 85:3131-3141.
- WEIGEL, K. A.** , REKAYA, R., ZWALD, N. R., FIKSE, W. F. 2001. International genetic evaluation of dairy sires using a multiple-trait model with individual animal performance records. *J. Dairy Sci.* 84:2789-2795.

Chapter 7

The Rise of Genomics

Single nucleotide polymorphisms (or SNPs) are one base pair that occur throughout the genome. Some are with a gene, many are near a gene, but there are millions of them on all the chromosomes. SCHAEFFER (2006) presented a strategy for utilizing SNPs to evaluate dairy bulls and to achieve greater genetic change over the traditional progeny testing scheme. His work was based on MEUWISSEN et al.(2001) and HAYES and GODDARD (2001).

That strategy and various forms of it were adopted in the Netherlands and France. In the beginning there were 6K SNP panels, then 10K, and so on. More SNPs for the same amount of money. How many SNPs were needed? More always sounds better. Soon there were 100K and 200K panels. Imputation was a statistical means of going from a panel of 50K SNPs to ‘guess’ with high reliability what the SNP genotypes would be on a 100K or 200K panel. Thus, cows could be genotyped using a cheaper 50K panel, and through imputation you could also have results for 100K.

A consortium grew to genotype 10,000 dairy sires from many countries using the 50K panel. Countries provided DNA of their bulls in exchange for a copy of the genotypes of all the bulls. Over time, however, the SNP chips became cheaper and dairy producers did not want to be left behind. They wanted their good cows to be genotyped too. Now there are many more cows than bulls with SNPs available. Countries have also become more possessive of their DNA information.

7.1 Using SNP Genotypes

STRANDEN and GARRICK(2009) and VANRADEN (2008) showed that the thousands of SNPs could be used to derive a better, enhanced additive genetic relationship matrix among the animals that were genotyped. Let \mathbf{A}^+ be the improved relationship matrix. The SNPs indicated that some animals were more related to each other than the values in the \mathbf{A} matrix would normally predict, and others were less related (i.e. fewer SNP genotypes in common). The resulting genomic EBV (GEBV) would be more accurate than EBV based only on data (DEBV). The more accurate relationship matrix meant that GEBV allowed people to pick the better animals amongst a group of full-sibs more accurately than using DEBV. SNPs allowed a better estimation of the Mendelian sampling effects of full-sibs. By genotyping calves, or embryos, those individuals could be more accurately evaluated before they could ever have progeny or even make a lactation record. Thus, it was no longer valid that progeny were a random sample of all possible progeny, a key assumption for the animal model to function properly. This was known as a ‘preselection bias’. The animal model as a tool for unbiased prediction of genetic merit became obsolete, but everyone is still using it, ignoring the bias being caused by preselection. Eventually something will break.

At the moment countries are utilizing genomic information to obtain GEBV within each country. Initially there was sharing of SNP genotypes, but AI companies have become more possessive of their genomic data. Many countries use single-step genomic BLUP methods to calculate GEBVs, combining genomic data on both sires and cows. The GEBV are sent to INTERBULL and put through a GMACE series of programs. The details of this are complex, and I have not kept up to date on the changes.

7.2 SNP Effects Model

Instead of single-step GBLUP, I propose the following SNP effects model, where the direct SNP effects are estimated and not genetic values of animals. GEBV are obtained after the analysis. All animals must be genotyped or imputations need to be available. GENGLER et al. (2007) described how to obtain estimates of SNP genotypes using an animal model. Then, all animals

with data can have either an actual genotype (-1, 0, or 1) or an estimated genotype (any number between -1 and 1).

Data on 20 dairy cows in five herd-year-seasons, and two age groups are given in Table 7.1. Imagine each cow has been genotyped for 50K SNPs, but for the sake of demonstration only six SNPs are utilized. For the SNP effects model to work there has to be more phenotypic records than there are SNPs. Thus, for 50K SNPs there would need to be 60K observations or more. Assume that these are first lactation cows. The SNP genotypes are -1, 0, and 1 for aa, Aa, and AA genotypes, respectively, where A and a are the two alleles for a SNP location. Cows are assumed randomly distributed over herd-year-seasons and age groups, as are the SNP genotypes.

Table 7.1 Cows with SNP genotypes.

Cow	HYS	Age	SNP Genotypes						Record
81	2	2	0	-1	0	1	0	0	22
82	5	2	0	-1	-1	-1	0	0	14
83	3	2	0	1	0	0	0	1	25
84	2	2	0	0	0	0	0	0	25
85	1	2	1	1	-1	0	-1	0	16
86	1	2	0	-1	0	0	-1	0	18
87	5	1	-1	0	0	1	-1	1	20
88	4	2	0	-1	0	1	0	0	25
89	2	2	0	-1	0	0	-1	0	18
90	1	2	1	0	0	1	-1	-1	15
91	5	2	0	0	1	0	0	0	30
92	4	1	1	0	0	0	-1	1	21
93	2	2	0	0	0	-1	-1	-1	18
94	4	2	-1	-1	0	-1	1	1	22
95	3	2	0	0	-1	0	0	0	22
96	3	2	-1	1	-1	0	0	-1	25
97	1	1	1	1	1	0	1	0	18
98	4	1	1	0	1	0	0	0	29
99	1	1	-1	1	1	1	0	1	23
100	3	1	0	0	1	0	0	0	24

All animals are assumed to be from the same country. The frequencies of SNP alleles could differ between countries and their effects could also differ

between countries, giving the GxE interactions. The linear model is

$$y_{ijk} = A_i + H_j + \sum_{m=1}^s S_{km}b_m + e_{ijk},$$

where

y_{ijk} is a production record on cow k made in age group i within herd-year-season j ,

A_i is a fixed effect of the i^{th} age group on production,

H_j is a random effect of the j^{th} herd-year-season on production,

b_m is the additive effect of the m^{th} SNP, and S_{km} is the genotype of cow k at SNP m (either -1, 0, or 1), and

e_{ijk} is a random residual effect.

The SNP effects are a fixed factor, but to get good estimates, one is added to the diagonals of the mixed model equations (MME). For this example the ratio of residual to HYS variances was assumed to be 4. In practice, these variances would need to be estimated.

The mixed model equations would be as follows:

$$\begin{pmatrix} 6 & 0 & 2 & 0 & 1 & 2 & 1 & 1 & 2 & 4 & 2 & -1 & 3 \\ 0 & 14 & 3 & 4 & 3 & 2 & 2 & 0 & -3 & -3 & 0 & -4 & -1 \\ 2 & 3 & 9 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 2 & -2 & 0 \\ 0 & 4 & 0 & 8 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & -2 & -1 \\ 1 & 3 & 0 & 0 & 8 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 8 & 0 & 1 & -2 & 1 & 0 & 0 & 2 \\ 1 & 2 & 0 & 0 & 0 & 0 & 7 & -1 & -1 & 0 & 0 & -1 & 1 \\ 1 & 0 & 2 & 0 & -1 & 1 & -1 & 10 & 1 & 1 & 0 & -2 & -2 \\ 2 & -3 & 2 & -2 & 2 & -2 & -1 & 1 & 12 & 1 & 1 & 1 & 0 \\ 4 & -3 & 1 & 0 & -1 & 1 & 0 & 1 & 1 & 10 & 2 & 2 & 2 \\ 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 9 & -2 & 1 \\ -1 & -4 & -2 & -2 & 0 & 0 & -1 & -2 & 1 & 2 & -2 & 10 & 1 \\ 3 & -1 & 0 & -1 & 0 & 2 & 1 & -2 & 0 & 2 & 1 & 1 & 9 \end{pmatrix} \begin{pmatrix} \widehat{A}_1 \\ \widehat{A}_2 \\ \widehat{H}_1 \\ \widehat{H}_2 \\ \widehat{H}_3 \\ \widehat{H}_4 \\ \widehat{H}_5 \\ \widehat{b}_1 \\ \widehat{b}_2 \\ \widehat{b}_3 \\ \widehat{b}_4 \\ \widehat{b}_5 \\ \widehat{b}_6 \end{pmatrix}$$

and the RHS (right hand sides) are

$$= \begin{pmatrix} 135 \\ 295 \\ 90 \\ 83 \\ 96 \\ 97 \\ 64 \\ 9 \\ -12 \\ 47 \\ 51 \\ -86 \\ 53 \end{pmatrix}.$$

The solution vector is

$$\begin{pmatrix} \hat{A}_1 \\ \hat{A}_2 \\ \hat{H}_1 \\ \hat{H}_2 \\ \hat{H}_3 \\ \hat{H}_4 \\ \hat{H}_5 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \hat{b}_4 \\ \hat{b}_5 \\ \hat{b}_6 \end{pmatrix} = \begin{pmatrix} 19.88339409 \\ 22.51593555 \\ -2.39648734 \\ -0.21382561 \\ 1.01237132 \\ 1.51885495 \\ 0.07908668 \\ -0.81744936 \\ 1.38447175 \\ 3.11656417 \\ 1.16360560 \\ 1.16201446 \\ 0.26016488 \end{pmatrix}.$$

To get genomic EBV, (GEBV) for the cows having observations, calculate

$$GEBV_k = \sum_{m=1}^s S_{km} \hat{b}_m$$

or

$$GEBV = \begin{pmatrix} -0.22 \\ -5.66 \\ 1.64 \\ 0.00 \\ -3.71 \\ -2.55 \\ 1.08 \\ -0.22 \\ -2.55 \\ -1.08 \\ 3.12 \\ -1.72 \\ -2.59 \\ -0.31 \\ -3.12 \\ -1.17 \\ 4.85 \\ 2.30 \\ 6.74 \\ 3.12 \end{pmatrix},$$

for cows 81 to 100.

Cows with the same genotypes, which happens frequently with only 6 SNPs, their GEBV would be identical. For 50K SNPs, it is less likely that two animals would share the exact same genotypes unless they were fraternal twins.

If there were a new group of 10 female calves with genotypes, then GEBV could be calculated for these animals, in the same way. Below are ten calves, their SNP genotypes and GEBV.

Table 7.2 Ten calves with SNPs and GEBV.

Cow ID	SNP Genotypes						GEBV
1	-1	1	0	0	0	0	2.20
2	-1	-1	-1	1	0	0	-2.52
3	1	-1	0	0	0	0	-2.20
4	0	-1	0	1	0	0	-0.22
5	-1	0	1	0	0	1	4.19
6	-1	0	-1	0	-1	1	-3.20
7	1	-1	-1	-1	0	1	-6.22
8	-1	-1	0	1	-1	0	-0.56
9	0	1	1	-1	-1	1	2.44
10	0	-1	0	1	0	0	-0.22

Note that there was no need to know the pedigrees of the animals or to use the additive relationship matrix. If there had been preselection of cows prior to making a record, the selection would have been made on GEBV. However, the randomness of SNP genotypes with respect to herd-year-season effects and age effects should still apply. Therefore, the SNP effects model should be unbiased.

With 50K SNPs, the order of the MME would be equal to the number of herd-year-seasons plus number of age groups plus 50,000 SNP effects. One could even evaluate 200K SNPs, but there would have to be more than 200K animals with records. If cows had more than one record, then cow permanent environmental effects would need to be in the model. Thus, a SNP effects model should be computationally easier to solve than an animal model, which would have at least one equation for every animal in the pedigrees.

As an exercise, what would an analogous test-day model look like? How many observations would be needed?

7.3 International Considerations

If every country used a SNP effects model, then conceptually there would be a set of estimates of the SNP effects for each country, assuming the same SNP chip or SNPs were used in each country. Below is a table of hypothetical SNP effect estimates from 3 hypothetical countries.

Table 7.3 SNP effect estimates from 3 countries.

SNP ID	Countries		
	1	2	3
1	-0.8174	-0.5071	-1.2934
2	1.3845	0.4783	1.5654
3	3.1166	2.9016	3.0069
4	1.1636	0.2836	1.2951
5	1.1620	1.0261	0.9872
6	0.2602	0.4377	0.2053

Using the same ten calves from Table 7.2, then applying the genotypes of the calves to the SNP effects from each country, then each calf gets a different GEBV for each country, as shown below.

Table 7.4 GEBV of ten calves for three countries.

Calf ID	Countries		
	1	2	3
1	2.2019	0.9854	2.8588
2	-2.5201	-2.5892	-1.9838
3	-2.2019	-0.9854	-2.8588
4	-0.2209	-0.1947	-0.2703
5	4.1942	3.8464	4.5056
6	-3.2010	-2.9829	-2.4954
7	-6.2219	-3.7329	-6.9555
8	-0.5655	-0.7137	0.0359
9	2.4357	2.5079	2.4953
10	-0.2209	-0.1947	-0.2703

The rankings of calves differs only slightly between countries. The correlations among the three countries was 0.96 between 1 and 2, 0.99 between 1 and 3, and 0.92 between 2 and 3.

Either SNP genotypes need to be shared among countries or the estimates of SNP effects. Exporters will have the genotypes of their animals and will need the SNP effect estimates for the country in which they wish to sell. Importers have the SNP effect estimates, but may not have the SNP genotypes of the bulls they wish to buy. INTERBULL would be a good place to store both and to provide GEBV to countries on their country scale for all bulls eligible to be sold or which already have progeny in that country.

7.4 References

- GENGLER, N.** , **MAYERES, P.**, **SZYDLOWSKI, M.** 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal* 1:21-27.
- HAYES, B. J.** , **GODDARD, M. E.** 2001. The distribution of the effects of genes affecting quantitative traits in livestock. *Genetics Selection Evolution* 33:209-229.
- MEUWISSEN, T. H. E.** , **HAYES, B. J.**, **GODDARD, M. E.** 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.
- SCHAEFFER, L. R.** 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123(4):218-223.
- STRANDEN, I.** , **GARRICK, D. J.** 2009. Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *J. Dairy Sci.* (submitted).
- VANRADEN, P.** 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414-4423.

Chapter 8

Competition Model

8.1 Introduction

MACE uses a linear model where de-regressed proofs from each country are treated as separate, but correlated, traits. Compared to earlier methods that assumed the genetic correlation between countries was one, MACE relaxed that assumption plus it handled more than two countries at a time. The correlations among countries were less than unity because of differences in genetic evaluation methods and because of true genotype by environment interactions, although both causes are difficult to quantify. De-regressed proofs were weighted by the number of effective daughters in each country where a bull has a proof. The additive genetic relationship matrix was built on sire, maternal grandsire, and maternal granddam information for all bulls from all countries.

Since 1994, MACE has undergone many changes and improvements by numerous researchers as well as increasing the number of traits by which bulls are evaluated. Part of the de-regression method involved the effective number of daughters, and several methods for calculating effective numbers and for de-regressions were derived over the years. Multiple Trait MACE was theoretically possible, but much more computationally demanding.

The genomic era hit the dairy industry around 2010, and GEBVs and regular EBVs were available from many countries. In some countries, genomics

has led to pre-selection of dairy bulls (in the embryo to 4 week old stages), such that bulls that become available as young sires were no longer a random group of progeny from a sire-dam mating. This led to bias in both regular EBVs and GEBVs (DUCROCQ and LEGARRA, 2011). If the proofs were biased, then using them in MACE could also cause bias in that system too. Bias was caused by the additive genetic relationship matrix inverse which assumes all progeny are random samples within a sire-dam mating pair. If sons of a particular sire-dam pair were pre-selected on the basis of their Mendelian sampling effects through genomics, then the sire and dam can be over-evaluated. Subsequently, the grandparents would also be biased, and cohorts of a sire within a herd-year-season would be biased by being compared to a bull that was over-evaluated. The bias may be small, but it permeates throughout the system of mixed model equations essentially affecting all animals, to some degree, through the relationship matrix.

Sequence data, runs of homozygosity, very large density SNP panels, are becoming proprietary data of artificial insemination organizations. If that data does not go to INTERBULL, then how are international genetic evaluations to be generated? What is the future of INTERBULL in this light?

HARVILLE (1977) presented a simple model, call it a competition model, for ranking high school football teams within one state. The observations were the differences in points scored between the home and away teams. Some teams performed better at home and others were better away, so that each team had a home team effect and an away team effect.

The purpose of this paper is to describe a competition model with the aim of ranking bulls across countries, giving one international ranking rather than attempting to estimate genetic values for bulls, which could suffer from pre-selection bias. The competition model is based on HARVILLE (1977) and is similar to social interaction models (MUIR 2005). The idea is to compare bulls pairwise within countries and to combine results across countries.

8.2 Model and Methods

At first thought, the milk proofs of bulls, for example, could be compared within country, but the units for milk proofs could differ a little or a lot between

countries depending on the evaluation methods and the level of production within each country. However, percentile rankings are of the same magnitude in every country, even though they do not reflect genetic differences in units of the trait. For example, the difference between 95 and 90 percentiles may not be the same number of kilograms of milk as the difference between 65 and 60 percentiles. Percentiles just provide a means of ranking bulls. Percentiles will be used in the competition model.

Table 8.1 contains example data on percentile rankings of ten bulls evaluated in three countries for milk production and the number of effective daughters in each country. Only four bulls were used in more than one country. This table contains only bulls that are currently available for exporting or which have lots of progeny in two or more countries (to provide linkages between countries). There is no need to rank bulls that are dead or not available for sale. INTERBULL currently requires all bulls' proofs from each country for a given trait, including all culled bulls and those not of interest to anyone.

Table 8.1

Example data for competition model.
ED = number of effective daughters.

Bull	Countries					
	%-ile	ED	%-ile	ED	%-ile	ED
1	98	44	92	20		
2	95	55			90	30
3	66	82				
4	20	15	30	40		
5			71	50		
6			84	60	80	48
7			96	51		
8					77	36
9					62	53
10					49	68

No pedigree information is required in the competition model. However, bulls still have to be uniquely identified across countries. The assumption is made that percentile rankings are a standard, straight-forward methodology that anyone can apply and provide to INTERBULL. The highest bull can only be 99 and the lowest bull can not be lower than 1. If country A uses

GEBVs and countries B and C do not, percentile rankings can be based on either GEBVs or EBVs depending on the country. No validation of genetic evaluation procedures is needed, although this should be of major importance within each country. There is no need to have the genotypes, or genomic data, of bulls stored at INTERBULL.

Let p_{ti} be the percentile ranking of bull i in country t , then within each country, differences among all pairs of bulls need to be calculated, as shown in Table 8.2. Hence if there are N animals provided by a country there would be $C = N * (N - 1)/2$ pairwise comparisons possible. If N is large, then C could be in the billions. Thus, N should be kept as low as possible, and this could be restricted by INTERBULL or by the country providing the percentile rankings. For example, if $N = 5000$, then $C = 12,497,500$, and if $N = 50,000$, then $C = 1,249,975,000$. Hence the need to possibly limit the number of bulls submitted per country.

Table 8.2
Percentile differences within country.

Country	i	j	$p_i - p_j$	$n_{ti} + n_{tj}$	$n_{ti} \cdot n_{tj}$	$(n_{ti} \cdot n_{tj}) / (N_{ti} + n_{tj})$
A	1	2	3	99	2420	24.44
A	1	3	32	126	3608	28.63
A	1	4	78	59	660	11.19
A	2	3	29	137	4510	32.93
A	2	4	75	70	825	11.79
A	3	4	46	97	1230	12.68
B	1	4	62	60	800	13.33
B	1	5	21	70	1000	14.29
B	1	6	8	80	1200	15.00
B	1	7	-4	71	1020	14.37
B	4	5	-41	90	2000	22.22
B	4	6	-54	100	2400	24.00
B	4	7	-66	91	2040	22.42
B	5	6	-13	110	3000	27.27
B	5	7	-25	101	2550	25.25
B	6	7	-12	111	3060	27.57
C	2	6	10	78	1440	18.46
C	2	8	13	66	1080	16.36
C	2	9	28	83	1590	19.16
C	2	10	41	98	2040	20.82
C	6	8	3	84	1728	20.57
C	6	9	18	101	2544	25.19
C	6	10	31	116	3264	28.14
C	8	9	15	89	1908	21.44
C	8	10	28	104	2448	23.54
C	9	10	13	121	3604	29.79

The model equation is then

$$(p_{ti} - p_{tj}) = \mu_t + b_i - b_j + e_{tij}$$

where

$(p_{ti} - p_{tj})$ is the difference in percentile rankings between bull i and bull j within country t ,

μ_t is a country mean percentile difference for the bulls provided,

b_i is a bull effect (fixed) for bull i , and

e_{tij} is a residual error.

If $y_{tij} = (p_{ti} - p_{tj})$, and \mathbf{y} is the vector of all pairwise percentile differences, then

$$\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \mathbf{W}\mathbf{b} + \mathbf{e},$$

where

\mathbf{X} is a design matrix relating country of ranking to the percentile differences,

$\boldsymbol{\mu}$ is a vector of country means,

\mathbf{W} is a design matrix relating observations to the pair of bulls being compared. Each row of \mathbf{W} has one 1 and one -1 .

\mathbf{b} is a vector of bull effects for all bulls in all countries, in this case ten bulls.

\mathbf{e} is a vector of residual effects.

Thus, \mathbf{b} ranks bulls across countries, similar to rankings of professional tennis players around the world, or the FIFA world rankings of country soccer teams.

Assume that each comparison between two bulls, y_{tij} has a different residual variance. Let n_{ti} be the number of effective daughters of bull i in country t , then

$$\text{Var}(y_{tij}) = \frac{n_{ti} + n_{tj}}{n_{ti}n_{tj}}\sigma_e^2$$

is a good approximation if there are lots of bulls.

$$\text{Var}(\mathbf{y}) = \mathbf{R}.$$

The least squares equations to be solved are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{W} \\ \mathbf{W}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{I} \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\mathbf{b}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{W}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}.$$

Note that \mathbf{I} was added to $\mathbf{W}'\mathbf{R}^{-1}\mathbf{W}$ to provide a ridge regression-like estimator of \mathbf{b} . This does not have to be done, but it readily forces the sum of $\hat{\mathbf{b}}$ to be zero.

Notice that there are no genetic correlations involved among countries. Only one set of solutions are obtained for bulls, i.e. their international rankings, just sort from high to low. The order of the equations is equal to the number of countries plus the total number of bulls in all countries being analyzed. Currently, MACE has equations greater than the total number of animals in the pedigrees multiplied times the number of countries. Thus, if there are 100,000 bulls and 30 countries that is 3 million equations, and 30 solutions per bull. With the competition model there are only 100,030 equations and one solution per bull.

The results from analyzing the example data from Table 8.2 are presented in Table 8.3. The solutions for country means are not zero, but they are not important in comparing bulls or countries. They depend entirely on which bulls are provided by each country.

Table 8.3
International Bull Rankings for example data.

Bull	Rating
7	21.4
2	20.3
1	19.0
6	10.0
8	7.4
5	-3.0
3	-3.7
9	-7.3
10	-19.9
4	-44.3

Suppose INTERBULL and the international community have suspicions

about the genetic evaluation system and data coming from country C, in the example. One option is to exclude the data from country C. Another option is to reduce the weights used in \mathbf{R}^{-1} . Consider multiplying the diagonals of \mathbf{R}^{-1} by 0.5 (50% less reliable than information from countries A and B). Re-doing the analysis with these new weights give the results in Table 8.4. The international rankings do not change substantially. Determining the relative importance of percentile rankings from each country may not be straight-forward. Alternatively, the weights may be based on genetic correlations between countries that have already been estimated by INTERBULL for several years.

Table 8.4

International Bull Rankings for example data.
Results assuming Country C 50% less reliable.

Bull	Rating	Re-Analysis
7	21.4	21.1
2	20.3	20.2
1	19.0	18.9
6	10.0	9.7
8	7.4	7.5
5	-3.0	-3.2
3	-3.7	-3.8
9	-7.3	-6.9
10	-19.9	-19.2
4	-44.3	-44.4

8.3 Discussion

The competition model simplifies international comparisons

- by providing only one ranking of bulls across all countries,
- by only using connecting bulls and bulls that are available for export from each country,
- by not needing bulky pedigree files,
- by avoiding calculation of de-regressed proofs, and

- by not needing SNP genotype data on bulls. Thus, maintaining the proprietary nature of a country's data.

Not all countries collect data on or evaluate certain traits. MACE for mastitis uses actual proofs for mastitis from some countries, but proofs for somatic cell score were only available for other countries. Percentile rankings, on the other hand, always have the same units and range of values, even if one is based on actual mastitis and the other on an indirect trait.

Every country has a merit index, and using the percentile ratings on these indexes in every country, they can be combined into an international ranking with the competition model, even though the weights in each index might be different. However, bulls tend to rank similarly for different indexes because the primary weight is on profitability, and therefore, lactation production. The merit index becomes another trait to be analyzed.

Percentile rankings are easily calculated and provided. The competition model does not need heritabilities or genetic correlations. Each trait would be analyzed separately. The competition model should save money and time for each country and especially for Interbull. The competition model will have fewer equations to solve, but more pairwise comparisons. Fewer checks need to be made on the percentile values themselves. Direct connectedness between countries needs to be assured, but should not be a problem these days as the effective population size is around 50. Interbull will not need to store genomic data, and not worry about regular EBVs versus GEBVs, just percentile rankings.

The competition model has drawbacks from the current MACE methodology, which need to be weighed against the advantages already listed.

- The competition model is a major change from MACE and will necessitate changes to the traditional thinking that has been established over the last 25 years. In some ways it is a step backwards.
- Genotype by environment interactions are assumed not to exist or to be negligible.
- The international rankings of bulls from the competition model are not genetic evaluations.

- The international rankings can not be merged into the EBV system within a country for any trait to assist the accuracy of EBVs in that country. This could have major consequences for some countries that include MACE evaluations in their within country systems. On the other hand it could greatly simplify their within country system and dealings with INTERBULL.
- Each country will need to decide what to do with the international rankings. Countries whose bulls are at the top will readily advertise this fact, while countries whose bulls are lower down the rankings, may not publish the news. Countries that are buying bulls, however, would like access to the entire list. The list should be available to every country even if they are not a paying participant in INTERBULL.

Conceptually, competition model analyses could be run whenever any country submits new percentile rankings of its bulls. Hence the international ranking would always be up to date, just like tennis players are rated after each tournament is completed. This would remove deadlines and schedules from each country's activities and those of INTERBULL.

The stability of the competition model needs to be monitored. As with the tennis player rating system, tennis players move up and down the ratings very quickly, especially if they are injured and stop playing for a few months. New dairy bulls are continuously appearing.

In conclusion, a competition model has been suggested to replace the MACE system for international livestock comparisons. The change-over would be simple mathematically, but the current structures and functioning of dairy industries in many countries, as well as to INTERBULL, may need to be overhauled.

8.4 References

- BANOS, G.** 2010. Past, present and future of genetic evaluations of dairy bulls. Proceedings of 9th WCGALP. Leipzig, Germany.
- DUCROCQ, V. , LEGARRA, A.** 2011. Interbull Meeting. Stavanger, Norway. August 2011.

HARVILLE, D. A. 1977. The use of linear-model methodology to rate high school or college football teams. *J. Amer. Stat. Assn.* 72:278-289.

MUIR, W. M. 2005. Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* 170:1247-1259.