Selection and Grouping of Herds in International Genetic Evaluation of Daughter Performance Records

K. A. Weigel and N. R. Zwald
Department of Dairy Science
University of Wisconsin
Madison, WI 53706 USA

Introduction

Although international dairy sire evaluations have traditionally been calculated by reanalyzing national EBV data, recent research has focused on direct analysis of individual animal performance records from participating countries. The latter approach offers several theoretical advantages, including relaxation of the assumption that all herds within a country use the same production system and freedom from "phantom" genotype by environment interactions that are actually artifacts of differences in genetic evaluation systems between countries. Previous research in our laboratory has focused on estimation of genetic correlations between countries using individual performance records in a sire model, as well as methodology for grouping herds across country borders (according to management variables) for the purpose of genetic evaluation.

The objectives of the current study were to investigate alternative plans for grouping herds into "clusters" or production systems, and to consider options for selecting well-connected subsets of herds for genetic parameter estimation.

Materials and Methods

Data from the current study included first lactation milk yield records from more than 16 million cows in 17 Interbull member countries: Australia, Austria, Belgium, Canada, Czech Republic, Switzerland, Germany, Estonia, Finland, Hungary, Ireland, Israel, Italy, Netherlands, New Zealand, United States of America, and South Africa. Records from Holstein-sired cows calving in 1990-1997 were considered.

Herd clusters were formed by calculating within-herd means for peak milk yield, temperature (region-specific), herd size (1^{st} parity cows only), days to peak yield, percent North American Holstein genes, and standard deviation for milk yield. These six variables were chosen as the most important of seventeen herd management variables used in our earlier studies. A factor analysis was performed, and factors were calculated for each herd prior to clustering. Herds were divided into clusters such that the cubic clustering criterion (a measure of distance between clusters) was maximized.

In our investigation of genetic connectedness between herds and countries, any bull with ≥ 10 progeny born after 5 years of age in ≥ 2 countries was considered as an "international proven sire", and any bull with ≥ 10 progeny born before 5 years of age in ≥ 2 countries was considered as an "international young sire".

Discussion

As shown below, the optimal number of clusters in this study was four. The first cluster, characterized by high average milk yield, consisted mainly of herds in Australia, Canada, Italy, and the United States. Cluster 2, characterized by large herd size, consisted primarily of herds in Czech Republic, Germany, Hungary, Italy, New Zealand, and the United States. Cluster 3, characterized by low peak milk yield, low percentage of North American Holstein genes, and low days to peak yield, consisted primarily of herds in Australia, Czech Republic, Germany, and New Zealand. Cluster 4, characterized by small herds with a high percentage of North American Holstein genes, consisted mainly of herds in Canada, Germany, and Netherlands. Thus, it seems possible to group herds logically into a relatively small number of production systems for the purpose of international genetic evaluation. Countries with high diversity in management and climate conditions may be represented in multiple clusters, while small countries with uniform conditions will have most herds in a single cluster.

The former group of countries may wish to publish separate sire rankings for each type of herd, while national sire lists may persist in the latter group.

When estimating genetic parameters between countries or between clusters using individual animal performance records, computational feasibility is a concern. Use of a simplified model (e.g., a sire model) may lead to underestimation of heritability and genetic correlation parameters, although bias in sire EBV (e.g., due to merit of mates) could be avoided by using a more complex model with fixed genetic parameters when estimating breeding values. Use of a complicated model for parameter estimation will necessitate sampling of herds. However, one must be careful not to introduce bias during the sampling process. The practice of selecting well-connected subsets of data is well accepted within the current MACE system, but in the case of daughter performance records we would be sampling certain herds, rather than certain sires.

As shown in the accompanying tables, there is great heterogeneity between countries in the extent of international sire usage in individual herds. Countries such as Australia, Belgium, Canada, Germany, Hungary, Italy, Netherlands, and United States have a strong representation of international proven sires in most herds. On the other hand, countries such as Austria, Estonia, Finland, and Switzerland have few genetic ties, due to small herd size and/or limited use of foreign semen. For young sires, the situation is more extreme. One can hypothesize that data from young sire herds would give parameter estimates that are less influenced by selection (as compared with estimates based on a handful of highly selected bulls with imported semen). However, only Australia, Germany, United States, and (to a lesser extent) New Zealand have adequate data for this purpose. In nearly all countries, herds with a heavier use of international genetics have a higher production level than other herds, and production generally increases as the usage of international sires increases. This may be due to a tendency for better-managed herds to seek top foreign genetics, or it may be an artifact of higher production due to heavier use of elite local bulls (with semen exported later). In summary, well-connected herds may not be representative of all national herds but better-managed herds that have used international sires in the paste may be a more appropriate target audience for international sire EBV.

Table 1. Mean number of progeny/herd of international proven sires.

	Total 1st Parity Cows	No. from Intl. Proven Sires	% from Intl. Proven Sires
AUS	89.3	36.6	41%
AUT	6.4	4.2	72%
BEL	24.3	16.9	70%
CAN	70.4	52.4	74%
CHE	17.2	9.6	56%
CSK	14.8	4.0	27%
DEU	40.5	23.3	57%
EST	24.6	0.2	1%
FIN	9.7	0.2	2%
HUN	306.7	126.0	41%
IRL	25.1	10.5	42%
ISR	201.6	98.1	49%
ITA	71.3	41.7	58%
NLD	56.4	44.1	78%
NZL	84.3	56.4	67%
USA	102.6	70.2	69%
ZAF	96.7	29.1	30%

Table 2. Mean number of progeny/herd of international young sires.

	Total 1st Parity Cows	No. from Intl. Young Sires	% from Intl. Young Sires
AUS	89.3	9.2	10%
AUT	6.4	1.1	18%
BEL	24.3	5.1	21%
CAN	70.4	5.1	7%
CHE	17.2	1.9	11%
CSK	14.8	0.8	5%
DEU	40.5	1.2	3%
EST	24.6	0.3	1%
FIN	9.7	0.0	0%
HUN	306.7	3.1	1%
IRL	25.1	0.7	3%
ISR	201.6	0.0	0%
ITA	71.3	1.6	2%
NLD	56.4	1.6	3%
NZL	84.3	1.7	2%
USA	102.6	5.3	5%
ZAF	96.7	2.1	2%

Table 3. Number of herds with progeny of international proven sires.

	Total Herds	= 10 Intl. Proven Sire Dtrs.	= 50 Intl. Proven Sire Dtrs.
AUS	9515	5923	2795
AUT	2167	293	6
BEL	3040	1256	408
CAN	14,400	11,669	7497
CHE	2719	1084	10
CSK	12,488	1428	266
DEU	63,014	30,385	12,936
EST	1732	8	0
FIN	14,191	54	1
HUN	1342	811	596
IRL	6324	1975	287
ISR	1103	867	481
ITA	17,307	8523	4292
NLD	26,292	18,879	10,658
NZL	18,452	14,636	7491
USA	41,990	30,370	18,210
ZAF	1263	454	192

Table 4. Number of herds with progeny of international young sires.

	Total Herds	= 10 Intl. Young Sire Dtrs.	= 50 Intl. Young Sire Dtrs.
AUS	9515	2385	261
AUT	2167	25	0
BEL	3040	567	4
CAN	14,400	2420	27
CHE	2719	32	0
CSK	12,488	274	8
DEU	63,014	1658	100
EST	1732	16	0
FIN	14,191	0	0
HUN	1342	97	25
IRL	6324	88	0
ISR	1103	0	0
ITA	17,307	788	23
NLD	26,292	912	13
NZL	18,452	984	53
USA	41,990	6880	464
ZAF	1263	81	2

Table 5. Production of herds with progeny of intl. proven sires.

	All Herds	= 10 Intl. Proven Sire Dtrs.	= 50 Intl. Proven Sire Dtrs.
AUS	4664	4772	5024
AUT	5856	6272	6680
BEL	6293	6489	6884
CAN	7029	7068	7358
CHE	5966	6273	6794
CSK	4733	5500	5801
DEU	6083	6302	6407
EST	3651	5959	
FIN	6161	7417	
HUN	5003	5192	5387
IRL	5149	5401	5854
ISR	8266	8297	8556
ITA	6656	7201	7750
NLD	6592	6651	6859
NZL	3600	3600	3597
USA	7854	7928	8164
ZAF	6056	6783	7091

Table 6. Production of herds with progeny of intl. young sires.

	All Herds	= 10 Intl. Young Sire Dtrs.	= 50 Intl. Young Sire Dtrs.
AUS	4664	4955	5285
AUT	5856	6166	
BEL	6293	6562	7011
CAN	7029	7588	7851
CHE	5966	6583	
CSK	4733	5507	5724
DEU	6083	6234	5734
EST	3651	5010	
FIN	6161		
HUN	5003	6072	6144
IRL	5149	5902	
ISR	8266		
ITA	6656	8304	8211
NLD	6592	7107	6785
NZL	3600	3615	3545
USA	7854	8404	8919
ZAF	6056	7501	9213

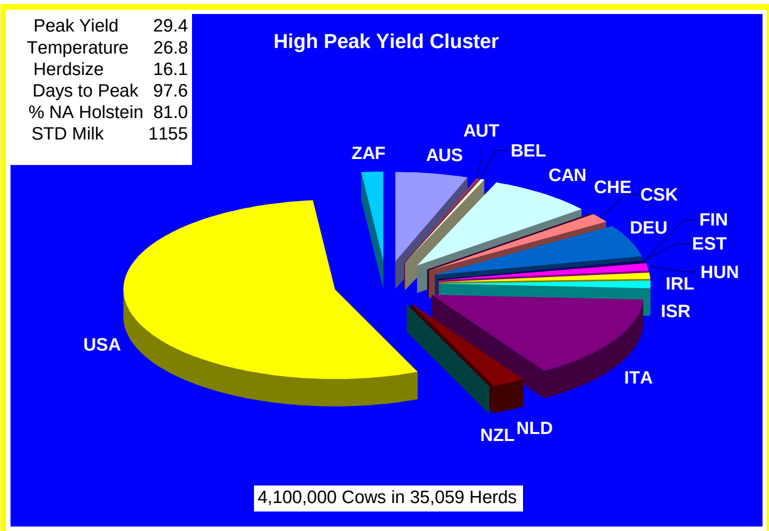

Figure 1. Percentage of cows from each country in cluster 1.

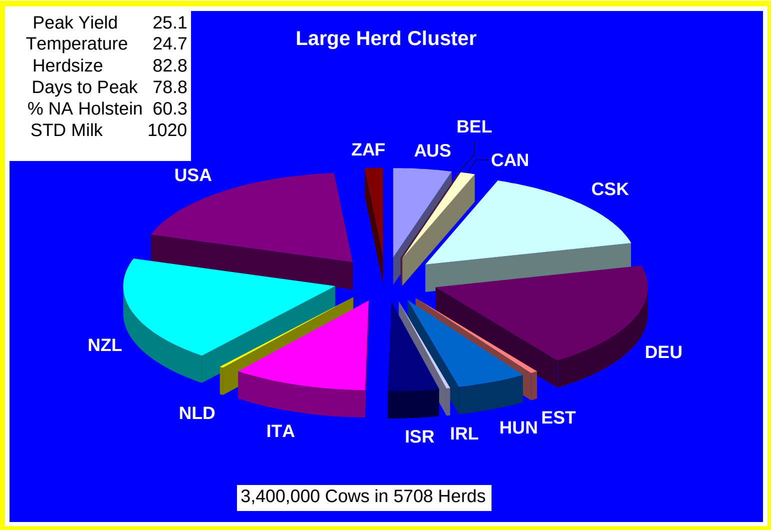

Figure 2. Percentage of cows from each country in cluster 2.

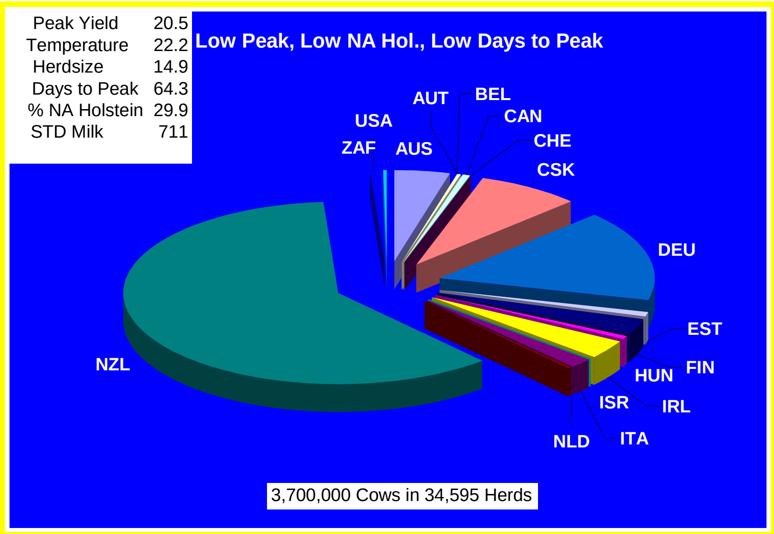

Figure 3. Percentage of cows from each country in cluster 3.

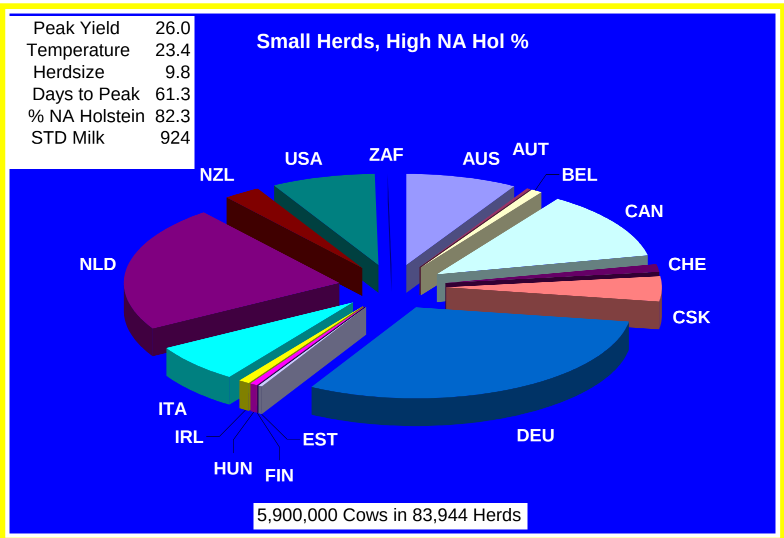

Figure 4. Percentage of cows from each country in cluster 4.