TÜBİTAK

Research Article

Turk. J. Vet. Anim. Sci.
2010; 34(2): 129-135
© TÜBİTAK
doi:10.3906/vet-0710-18

Contemporary grouping in mixed-size dairy herds experiencing four seasons

Mohammad Ali NILFOROOSHAN

Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala - SWEDEN

Received: 08.10.2007

Abstract: Countries with different herd structures and climates use different combinations of herd, year, and season to define the effect of their climatic conditions and management. Estimates of variance components and Akaike information criterion values resulting from 8 analytical models differentiated by different combinations of herd, year and season effects were compared for milk and fat yields. Data consisted of pedigree and records of 11,850 Holsteins freshening in 38 dairy herds of Isfahan, Iran, between January 1994 and October 2001. Heritability estimates were lowest for the models that used herd-year combination as a random effect. Moreover, the use of these models is not recommended due to their low goodness of fit, at least for countries that have too much variety in herd size. The models that considered herd-year or herd-year-season as a fixed effect to define contemporary groups fit the data best. It is worthwhile rechecking the frequency of animals in herd-year-season groups in order to control the minimum number of animals in contemporary groups. Countries with 4 seasons and small herds may consider herd-year and 4-month season as 2 separate fixed effects to observe contemporary groups sufficiently and to make a better correction for seasonal changes or considering herd-year-season as a random effect.

Key words: Herd, year, season, contemporary, Akaike information criterion

Introduction

One aim of breeding programs is to separate the effects of genes and environment to select animals with high genetic merit (1). While the adjustment of performance records can be done for some known environmental effects (e.g., age, milking frequency, and days in milk) before BLUP evaluations are made, the effects of some environmental factors (e.g., herd, year, and season) can be estimated simultaneously with breeding values while solving BLUP mixed-model equations. Comparing a bull's daughters with their contemporaries, milked in the same herd, year, and season still forms the basis of most evaluation methods. The statistical models used in BLUP evaluations account for differences in management between and within herds by assigning animals to different contemporary groups. An animal's milk record is usually assigned to a contemporary group depending on the herd, year, and season of calving in which it occurred (1). Countries have a variety of population structures, herd sizes, climatic conditions, and seasonal changes within and between years. These differences can provide a basis for categorizing animals that seem to come from similar

* E-mail: Mohammad.Nilforooshan@hgen.slu.se

environmental conditions. In Iran, as a subtropical country, the pattern of seasonal changes is generally four 3-month periods, and herd sizes vary considerably from small to very large.

Based on the data structure and the aims of the investigation, studies apply different models including different combinations of herd, year, and season effects. In comparative studies, where the aim is mainly to study the direct effects of herd, year, and season on phenotypic performances instead of the accuracy of genetic evaluations, these effects may be considered separately (2). Boettcher et al. (3) considered herd-year as a random effect and season as a fixed effect in the model. They considered herd-year groups that had at least 12 cows in each to ensure that contemporary groups were of sufficient size. Others had both herd-year and season as fixed effects (4,5). Khan and Shook (6) and Suzuki et al. (7) included herd-year and the month of calving separately in the model. Wiggans et al. (8) considered herd-year and year-season in the model of analysis as random and fixed effects, respectively. They considered year-season combination to allow seasonal differences to change over time. In addition, many studies have used a herd-year-season (HYS) combination. Different studies have considered various seasonal definitions including 6 (9), 4 (10), 3 (4,11), or 2 (8) seasons. Five birth season definitions (2, 3, 4, 6, and 12 seasons/year) were investigated by Fatehi et al. (12) on 2 heifer fertility traits. Crump et al. (13) studied various fixed-time seasons (3, 4, and 6 months), as well as a rolling- based season that took account of the calving patterns within herds as a way to combine animals to form groups of sufficient size. A season definition may attempt to include a similar number of observations in each season within a herd-year (14). However, in countries with considerable seasonal changes, it is also important as a climate correction factor. Based on the population structure, climatic conditions, and the traits concerned, countries apply various combinations of herd, year, and season for different traits (15).

The objective of this study was to find the fittest combination of herd, year, and season effects, based on the models' goodness of fit (Akaike information criterion) for genetic evaluation of production traits in dairy cattle, especially for Iran as a subtropical country, with a high variability in herd sizes.

Materials and methods

The data included pedigree, milk, and fat yield records adjusted to 305 DIM, mature equivalent (age), and twice daily milking, which were previously used by Nilforooshan (10). After excluding herd-years with fewer than 6 observations (animals with records), the final data comprised the first lactation of 11,850 Holstein cows freshening in 38 dairy herds of Isfahan province in Iran. Isfahan province covers an area of approximately 107,027 km^2, situated in the center of Iran (+32°N +51°E). It includes 18 townships, 38 counties, 67 cities, and 2470 villages. The province experiences a moderate and dry climate.

Eight different animal models were applied to study the efficiency of various combinations of herd, year, and season. The models were not made more complex because only the differences in the combinations of herd, year, and season were of interest. Table 1 shows the differences among the 8 animal models studied. The linear model and the variance structure for models 1, 2, 3, 4, and 8 were as follows:

$$Y = Xb + Z_1a + e$$

$$\mathrm{Var}\begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 \\ 0 & I\sigma_e^2 \end{bmatrix}$$

The linear model and the variance structures for models 5, 6, and 7, which had an additional uncorrelated random effect, herd-year, were as follows:

$$Y = Xb + Z_1a + Z_2c + e$$

$$\mathrm{Var}\begin{bmatrix} a \\ c \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & 0 \\ 0 & I\sigma_c^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

where Y, b, a, c, and e are the vectors of observations, fixed effect(s), random additive genetic, additional uncorrelated random effect, and random residual effect, respectively. X, Z_1, and Z_2 are the incidence matrices relating records to fixed, additive genetic, and additional uncorrelated random effects, respectively. σ_a^2, σ_c^2, and σ_e^2 are Var (a), Var (c), and Var (e), respectively. A and I are the numerator relationship matrix among all animals and the identity matrix with the order of number of records, respectively.

Milk yield and fat yield traits were analyzed separately using the univariate procedure (DFUNI) of the DFREML program. Estimation of heritabilities and variance components was carried out using the derivative-free restricted maximum likelihood method developed by Meyer (16) through Simplex way with the convergence criterion of 10^{-8}. The starting value for both h^2 (σ_a^2/σ_p^2) and C^2 (σ_c^2/σ_p^2) was 0.25 (σ_p^2 is the total or phenotypic variance). The chosen criterion for comparing models was Akaike Information Criterion (AIC) and all models were compared to model 1.

Following the analysis of data with various models, including model 8, the number of animals in HYS groups was studied, and the analysis of model 8 was repeated excluding HYS groups containing fewer than 5 observations.

Results

A description of the differences among the models studied is presented in Table 1. Various models had different combinations of herd, year, and season effects. Model 1 considered these effects separately with no interaction among them. Model 2 contained herd-year and season, model 3 contained herd-season and year, and model 4 considered year-season and herd effects all as fixed; each model had a different hidden interaction. Models 5, 6, and 7 had herd-year as a random effect. These models had season, herd-season, and year-season, respectively, as a fixed effect. For models 6 and 7, it was impossible to implement both effects as fixed because the incidence of fixed effects was not independent from each other and the animal model could not proceed. Finally, model 8, which is the preferred model for comparison of contemporary groups, contained 1 fixed effect of HYS that took into account 4 possible interactions of herd × year, herd × season, year × season, and herd × year × season. Table 2 shows the estimated variance components and AIC ratios (relative to model 1) for milk yield and fat yield traits. Milk yield heritabilities ranged from 0.176 (model 5) to 0.245 (model 2), and fat yield heritabilities from 0.157 (model 5) to 0.261 (model 1).

Model 1 did not have a high goodness of fit because it did not take into account any interaction and contemporary comparisons. For both traits, model 8 (the most complete model) and then model 2, which considered both herd-year and season as fixed effects, had the highest goodness of fit (the lowest AIC). Generally, models that considered herd-year as a random effect (models 5, 6, and 7) could not appropriately fit the data. This may be a reason for contemporary groups to be usually considered as a fixed effect. For both traits, the heritabilities estimated by these models were lower than the other models and the expectations, which may be a result of the estimation of the additional variance component (V_C) and overestimation of the phenotypic variance (V_P). Comparing models 5, 6, and 7, it was observed that model 6 resulted in higher heritability and lower AIC values (better fit). In these models, as C^2 increased, V_P increased and h^2 decreased. Due to the high number

Table 1. A description of the models studied with respect to different combinations of herd, year, and season.

Model	Effect						
	H	Y	S	HY	HS	YS	HYS
1	F	F	F				
2			F	F			
3		F			F		
4	F					F	
5			F	R			
6				R	F		
7				R		F	
8							F
Levels	38	8	4	188	152	31	645

F = Fixed effect; R = Random effect; H = Herd; Y = Year; S = Season of animals in herd-year groups (63 ± 81), it is not recommended to use herd-year as a random effect for Iranian Holstein data.

A comparison among models 2, 3, and 4 shows that using herd-year was preferable to herd-season and year-season. This may be due to the fact that herd-year with 188 levels can better fit the phenotypic changes than the herd-season with 152 and the year-season with 31 levels.

After running model 8, the size of HYS groups, the number of animals, and the number of HYSs in various group sizes were studied (Figure). In this study, even by removing herd-years with fewer than 6 cows, there were still 48 HYSs with a single observation. Overall, 509 cows were from 197 HYSs with fewer than 5 observations (Figure). After removing this part of data as uninformative contemporary groups, the analysis of model 8 was repeated (Table 2: R8). In this way, both the heritability and AIC estimates of milk and fat yields were improved.

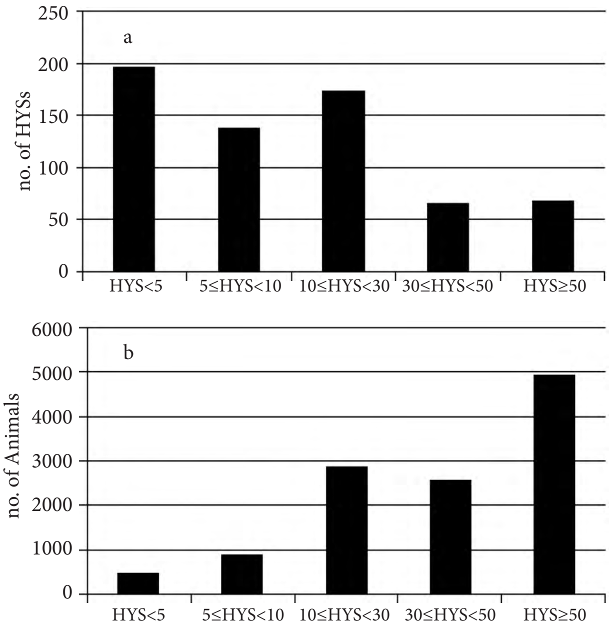

Figure. The number of HYSs (a) and the number of animals (b) in different HYS group sizes.

Table 2. Estimates of heritabilities and variance components for milk and fat yields (kg) applying different models.

Trait	Model	V_A	V_C	V_E	V_P	h^2 ± S.E.	C^2 ± S.E.	AIC_1
Milk	1	407453		1443797	1851250	0.220 ± 0.027		100.00
	2	438396		1347343	1785739	0.245 ± 0.028		98.76
	3	406159		1430403	1836560	0.221 ± 0.027		99.18
	4	391080		1448476	1839556	0.213 ± 0.026		99.85
	5	448193	754610	1340377	2543180	0.176 ± 0.021	0.297 ± 0.024	100.27
	6	445507	293874	1331341	2070722	0.215 ± 0.025	0.142 ± 0.019	99.12
	7	433511	647679	1347603	2428794	0.178 ± 0.022	0.267 ± 0.023	100.08
	8	413716		1348970	1762686	0.235 ± 0.029		95.27
	R8	424704		1355033	1779737	0.238 ± 0.029		92.60
Fat	1	346.42		980.15	1326.57	0.261 ± 0.024		100.00
	2	302.84		908.73	1211.57	0.250 ± 0.025		98.14
	3	342.12		971.99	1314.11	0.260 ± 0.024		99.25
	4	332.31		982.70	1315.01	0.253 ± 0.024		99.85
	5	305.64	731.67	906.66	1943.97	0.157 ± 0.018	0.376 ± 0.026	99.68
	6	302.31	326.65	899.63	1528.59	0.198 ± 0.022	0.214 ± 0.022	98.57
	7	297.05	569.58	905.99	1772.62	0.168 ± 0.019	0.321 ± 0.025	99.46
	8	288.41		887.98	1176.39	0.245 ± 0.026		94.73
	R8	296.31		889.14	1185.45	0.250 ± 0.026		91.99

V_A = additive genetic variance; V_C = variance due to the additional uncorrelated random effect (herd-year); V_E = random residual variance; V_P = phenotypic variance; h^2 = heritability; C^2 = the proportion of the additional uncorrelated random effect variance to the total variance; AIC_1 = Akaike information criterion relative to model 1 (smaller values are better); R8 = repeated model 8 with 448 HYSs instead of 645.

Discussion

Assigning animals to contemporary groups is a usual procedure, which is based on the differences among individual performances, used to remove herd, year, and seasonal effects from the average performance of animals milking in the same herd at the same time. The accuracy of selection will be improved by ensuring that animals in a contemporary group get as a similar treatment as possible. Preferably, contemporary groups can be as large as possible to allow the best possible separation of genetic and environmental effects. In practice, a compromise has to be reached between making the groups large but including animals which did not get a similar treatment, and accepting smaller groups of animals that really do share a similar environment (1). Both Crump et al. (13) and Schmitz et al. (14) studied contemporary grouping strategies providing an ability of joining (clustering) animals into the desired contemporary group sizes, using various time spans. Schmitz et al. (14) based their HYS clustering study on various maximum season intervals and a desired minimum number of observations in a cluster (contemporary group). They concluded that the accuracy of sire evaluation is more dependent on the number of contemporaries than the procedure of grouping animals (fixed or variable season definition (clustering)). When the number of contemporaries is small, using HYS may not make efficient use of all information within and between HYS groups. Although the accuracy of HYS solutions is not of high importance, the effects of some low frequency HYS groups may not be estimated accurately. Moreover, clustering cannot completely overcome the problem with small herds, but it contributes to a better use of information available (14).

Although model 8 was superior to all prior models, the potential problem with this model, especially with the current or similar data including small herds, is the low number of observations in some contemporary groups. In this situation, applying a different model, removing animals in small contemporary groups, joining small contemporary groups (increasing seasonal intervals or joining small herds into regions), or using HYS as a random effect may be options. Usually, contemporary groups are treated as fixed, because treating them as random when a non-random association exists between sires and herds may lead to biased predictions by increasing the contribution of information or the effective number of daughters from which an animal is being evaluated (17). However, in some countries with many small herds, due to less certainty about the effect of contemporary groups, treating them as random is inevitable.

It makes sense to merge some contemporary groups with a very low number of observations to make them eligible to join a group. However, in countries experiencing 4 seasons, increasing seasonal intervals for the whole or a large set of data may reduce the accuracy of genetic evaluations by ignoring some seasonal differences. Therefore, in these situations, joining herds into regions or the use of large contemporary groups (herd-year) with full correction for seasonal changes (4 seasons) is recommended (model 2). Considering different seasonal patterns or joining small herds into regions is up to the countries, depending on how important the management or seasonal differences are. For instance, USA uses a flexible herd-year-season with a rolling base season (2-12 months), Turkey uses region-herd (joining small herds into regions) and year-month, Australia and Portugal use herd-year-season (2 seasons), Argentina uses herd-year-season and month of calving, and Canada uses region-year-season and herd-test day in its national genetic evaluation system for production traits (15). In lactation models, contemporary group effect is fixed in animal's whole lactation, whereas in test-day models there is an opportunity to account for different environmental effects on different test dates by assigning test-day records to different groups and considering lactation curves. In this way, more animals from different test-days will be included in a group. Therefore, applying test-day models can overcome the problem of small contemporary groups to some extent. In test-day models, considering herd-year and season effects, all test dates of an animal are assigned to 1 or 2 herd-years and the season effect would be considered the same over all herd-years, which is not usually the case. Furthermore, to take into account the interaction between the test dates and the management, it is recommendable to combine herd and test date into one effect (HTD), which is used in the model of national genetic evaluations in many countries using test-day models such as Italy, Germany-Austria, Canada, and Belgium (15).

Schmitz et al. (14) studied 5 dairy breeds in the northeast U.S. and concluded that the accuracy of sire evaluations increases with the number of observations in contemporary groups. However, they also reported that a bias is expected to occur in Holsteins using long time intervals to make contemporary groups from animals within herds. In the current data, there were many small herds as well as very large herds, which contributed to the high variability in contemporary group sizes. As it can be seen from the Figure, a small proportion of the data (4%) came from a large number of small (with fewer than 5 observations) HYS groups (30% of the contemporary groups). Therefore, the chosen strategy was to exclude this part of the data and to re-run model 8 (R8). As a result, both heritabilities and AIC ratios improved compared to model 8 for both traits. However, excluding a part of the data is not always a good solution, especially with a higher number of animals in small contemporary groups (e.g., more than 10%), because there would be a possibility of over-sampling from large herds and neglecting information from small herds as a part of the real population. Thus, in this situation, merging small contemporary groups may be recommended. Moreover, in the case of dealing with data from small herds and high proportion of animals in small contemporary groups (e.g., more than one-third) joining them may lead to a negligence of herd or season effects (depending on the way of joining), especially for the countries experiencing 4 seasons with considerable differences in herd management.

Applying model 8, it is recommended to study the frequency of animals by contemporary group size prior to the analysis. It is also important to consider the climate and the system of seasonal changes in the specific country, herd sizes, and the percentage of replacement heifers (the number of animals in contemporary groups increases as the percentage of replacement heifers increases). Overall, the best combination of herd, year, and season, together with season definitions, depends on the herd structure and climatic conditions of the specific country, the breed, and the trait concerned, since the number of animals with records in contemporary groups may differ from breed to breed and trait to trait.

References

1. Simm, G.: Genetic Improvement of Cattle and Sheep. Farming Press Books and Videos, UK. 1998.
2. Kaya, I., Uzmay, C., Kaya, A., Akbaş, Y.: Comparative analysis of milk yield and reproductive traits of Holstein-Friesian cows born in Turkey or imported from Italy and kept on farms under the Turkish-ANAFI project. Ital. J. Anim. Sci., 2003; 2: 141-150.
3. Boettcher, P.J., Jairath, L.K., Dekkers, J.C.M.: Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. J. Dairy Sci., 1999; 82: 1034-1044.
4. Chongkasikit, N., Vearasilp, T., ter Meulen, U.: Heritability estimates of protein %, fat %, lactose %, non fat solids and total solids of dairy cattle in northern Thailand. Conference on International Agricultural Research for Development, Witzenhausen, Germany. 2002.
5. Peixoto, M.G.C.D., Verneque, R.S., Teodoro, R.L., Penna, V.M., Martinez, M.L.: Genetic trend for milk yield in Guzerat herds participating in progeny testing and MOET nucleus schemes. Genet. Mol. Res., 2006; 5: 454-465.
6. Khan, M.S., Shook, G.E.: Effects of age on milk yield: time trends and method of adjustment. J. Dairy Sci., 1996; 79: 1057-1064.
7. Suzuki, M., Pereira, J.A., Kuchida, K., Saito, Y., Ikeuchi, Y.: Interaction due to genetic and environmental effects on milk production in Japanese Holstein. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia. 1998.
8. Wiggans, G.R., Van Tassell, C.P., Philpot, J.C., Misztal, I.: Comparison of dystocia evaluations from sire and sire-maternal grandsire threshold models. Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. 2002.
9. Gengler, N., Tijani, A., Wiggans, G.R., Van Tassell, C.P., Philpot, J.C.: Estimation of (co)variances of test day yields for first lactation Holsteins in the United States. J. Dairy Sci., 1999; 82: 225.
10. Nilforooshan, M.A.: Genetic evaluation and comparison between domestic and imported (Canada, USA, Europe) sperms for milk yield, fat yield, and fat percentage in some dairy farms of Isfahan province. MSc Thesis. Isfahan (Khorasgan) Azad University, Isfahan, Iran. 2003. (MSc thesis in Farsi, with an abstract in English)
11. Márquez, A.P., Cobos, S., Bueno, F., Guerrero, J.N.: Genetic parameters and breeding values for milk yield of Holstein sires in a commercial Holstein dairy herd in Mexicali, Mexico. Proceedings of the Western Section Meeting, American Society of Animal Science, Oregon, USA. 2004; 55: 106-109.

12. Fatehi, J., Schaeffer, L.R., Jamrozik, J.: Research report to the GEB of Canada. Models with herd by year interaction and different definitions of season of birth for two heifer fertility traits. 2005; http://cgil.uoguelph.ca/dcbgc/Agenda0503/March2005_JFReprot.pdf

13. Crump, R.E., Wray, N.R., Thompson, R., Simm, G.: Assigning pedigree beef performance records to contemporary groups taking account of within-herd calving patterns. Anim. Sci., 1997; 65: 193-198.

14. Schmitz, F., Everett, R.W., Quaas, R.L.: Herd-year-season clustering. J. Dairy Sci., 1991; 74: 629-636.

15. Interbull Centre: National information, description of GES as applied in member countries. http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm. Accessed August 2008.

16. Meyer, K.: DFREML: Programs to estimate variance components by restricted maximum likelihood using a derivative-free algorithm. User notes, Version 3.1. University of New England, Armidale, Australia. 1998.

17. Fırat, M.Z.: Bayesian analysis of test day milk yields in an unbalanced mixed model assuming random herd-year-month effects. Turk. J. Vet. Anim. Sci., 2001; 25: 327-333. (article in Turkish, with an abstract in English).