Paper G6.14 - *55th Annual Meeting of the EAAP, 5-9 September, 2004. Bled, Slovenia. Corresponding author's e-mail:mjc@inia.es*

Comparing alternative definitions of the contemporary group effect in Avileña Negra Ibérica cattle in a Bayesian analysis

M.J. Carabaño[1*], A. Moreno[1], P. López-Romero[2], C. Díaz[1]

[1]Depto Mejora Genética Animal. INIA. Apdo 8111. 28080 Madrid.[2]Ud. de Bioinformática. CBMSO-CSIC 28049 Madrid, Spain

---

Abstract

Data on weaning weight from 12,740 animals were used to compare different contemporary group (CG) definitions to be applied in the current genetic evaluation in the Avileña Negra Ibérica beef cattle breed. Several classical 'ad-hoc' statistics and criteria for statistical model comparison within a Bayesian analysis were used. Six alternative definitions for the CG effect were compared: Herd- year-season of calving (HYS), with seasons defined according to the four natural seasons; herd-year-month of calving (HYM); herd clusters with length of 30 days (HC30-30) or 90 days (HC90-90), and, adaptive clusters with two time limits, 30 and 90 days (HC30-90), and 30 and 180 days (HC30-180). Classical statistics pointed at HYM and HC30-30 as the models showing smaller within CG variation and smaller residual variance, with slightly worse performance in terms of accuracy of prediction for breeding values for direct genetic effects. Bayes factors and cross-validation predictive densities, allowed for a better discrimination among models. Models including CG spanning 30 days were more plausible and showed better predicting ability than models spanning 90 days. Adaptive CG showed intermediate results. Model definition had a relatively large impact on the variance components estimates. Overall, HYM showed best results but implied the largest loss of data (14%). HC30-90 might represent a compromising solution for this population.

Abbreviation Key: BF = Bayes Factor; CG = Contemporary groups; HC30-30: Herd cluster spanning 30 d; HC30-90: Herd cluster spanning 30 d minimum and 90 d maximum; HC30-180: Herd cluster spanning 30 d minimum and 90 d maximum; HC90-90: Herd cluster spanning 90 d; HYM = Herd-Year-Month; HYS: Herd-Year-Season; MD = Marginal density of the data

---

1. Introduction

Genetic evaluations for weaning weight in the Avileña-Negra Ibérica beef cattle population use predetermined seasons to form contemporary groups (CG). A problem associated with this definition is the arbitrary assignment of seasons, which does not respond either to maximum accuracy or minimum bias criteria. Alternative approaches have been proposed to try to account for this problem (Wiggans et al., 1988; Schmitz et al., 1991; Crump et al., 1997). Those procedures require establishing the numerical value of the parameters, size and time span of the CG, that will result in an optimized definition.

Several 'ad-hoc' criteria have been used to compare alternative definitions of CG (Schmitz et al., 1991, Sivarajasingam, 1993, Crump et al., 1997, Van Bebber et al., 1998). These criteria include estimates of within CG variance, residual variances, effective number of progeny and accuracy of genetic evaluations. Other criteria based on the likelihood have been explored to a less extent. The Bayesian analysis provides the tools for model selection in a more general framework.

The goal of this study was to compare alternative definitions of CG for weaning weight of Avileña Negra Ibérica beef cattle using several criteria that include classical and Bayesian criteria for comparison of models.

2. Material and methods

2.1. Data

Weaning weight records from 12,740 purebred animals from the Avileña-Negra Ibérica beef cattle breed, born between 1984 and 1999 in 83 herds were used in this study. These records formed the genetic evaluation data base for this breed. The Avileña Negra Ibérica is a Spanish local breed managed under very extensive conditions. Artificial insemination is not commonly used, except for generating genetic ties among herds for the genetic evaluations. As a consequence, a relatively low degree of connectedness among herds is expected in this data set.

The pedigree file included 21,483 animals. Totals of 325 bulls and 5773 cows were sires and dams of animals with records. Of those, 77 bulls and 966 cows also had records as calves. A total of 131 bulls were maternal grandsires of cows having own record, which also were dams of calves with records.

2.2. Models- Definition of CG

The following model, currently used in genetic evaluations for this breed, was used:

$$y_{ijklno} = CG_i + sx_j + dage_k + sf_l + b(age_n) + u_{dn} + u_{mo} + p_o + e_{ijklno} \qquad [1]$$

where: y_{ijklno} is the weaning weight of animal n; CG_i is the effect of the ith contemporary group, sx_j is the effect of the sex of the animal; $dage_k$ is the effect of the kth class of the age of the dam; sf_l is the effect of the lth class for supplement feed; b is the linear regression coefficient of weight on age of calf at weaning (age_n); u_{dn} is the direct additive genetic effect of the animal; u_{mo} is the maternal additive genetic effect of o, dam of animal n; p_o is the maternal permanent environmental effect of o, and, e_{ijklno} is the error term.

Six alternative definitions of CG were considered. A minimum number of five observations per CG class were required in all cases. Contemporary groups were formed within herd according to 'conventional' seasons (HYS) or months (HYM), or, following the 'natural' calving pattern. The CG classes for the 'natural' clusters were obtained in two steps. In a first step, animals were sorted by birth date within herd. Starting from the first date of birth, a CG was formed if adding 30 or 90 d to the first birth date in the group resulted in groups that included five or more animals. Two data sets resulted from this first step, HC30-30 and HC90-90, for 'natural' fixed period herd clusters, spanning up to 30 or 90 d. In the second step, adaptive strategies involving two time limits were developed. Animals not attached to any CG for the HC30-30 definition were assigned to the previous/subsequent CG if the difference between its date of birth and the date of birth of the first/last animal in that CG did not exceed 90 d, for the HC30-90 definition, or 180 d, for the HC30-180 definition.

The six data sets with alternative definitions of CG were analyzed with model equation [1]. Both, classical approaches and Bayesian procedures were used. For the classical analyses, BLUP procedures were employed. Variance components used in the current evaluation system were utilized to obtain the 'classical' model comparison criteria for all definitions of CG.

For the Bayesian analyses, variances and other parameters involved in the model were considered unknown. Data were assumed to be generated from a multivariate normal distribution (MVN). MVN prior distributions were also assumed for the location parameters. For the dispersion parameters, scaled inverse chi-square (χ^{-2}) and inverse Wishart (IW) distributions were used

Posterior marginal inferences on parameters of interest were drawn using a Gibbs sampling scheme. A burn-in period of 20,000 iterates were carried out for all analyses and a total of 100,000 iterates were carried out after burn-in.

2.3. Comparison of models

The between and within CG variances, the effective number of progeny for direct effects, and, the accuracy of breeding values for groups of animals were computed. The between and within CG variance were obtained using the SAS proc Varcomp (SAS Inst. Inc., Cary, NC) under the type I option for the following model:

$$y^*_{ij} = GC_i + e_{ij} \qquad [2]$$

where, y^*_{ij} is the observation corrected by the corresponding BLUE and BLUP solutions for other than the CG effect in model [1].

Within the Bayesian framework, the Bayes factor (BF) (Newton and Raftery, 1994; Kass and Raftery, 1995) and the cross-validation predictive densities of the data (Gelfand et al., 1992) were computed. The BF for two competing models was computed as the ratio of the corresponding marginal densities of the data (MD) under each model. The MD for each model was obtained from the Newton and Raftery (1994) estimator. Since the value of the MD varies with the size of the data set, the total data set containing all 12,740 initial observations was used to compute the log of the MD for each alternative definition of CG.

The expected values of the difference between the observed value (y_r) and its corresponding prediction (Y_r) with respect to its univariate cross validation predictive density, $d_r = E_{Y_r|y_{(r)}}[Y_r - y_r]$, were also computed. The best model would be the one having minimum $D = \frac{1}{n}\sum_{i=1}^{n} d_r^{\,2}$. An importance sampling, with the joint posterior density of the parameters given the data as an importance distribution, was implemented to evaluate every d_r. The D statistic was computed for both the total data set (D_T) and the specific data sets of different size associated to the alternative CG models (D_s).

The log of the MD and D statistics were obtained for each data within the Gibbs sampling process (details of the implementation can be found in López-Romero et al., 2003).

3. Results and discussion

Table 1 presents information about the weighted averages for size and time span of the CG under the six alternative definitions. Definitions of CG with 90 d periods (HYS and HC90-90) resulted in CG of largest size (more than 20 observations per contemporary group, on average), but also larger average time span (around 60 d). Loss of information was low for these strategies (4.7 and 2.7 % for HYS and HC90-90, respectively), but not the lowest. Definitions with 30 d, with one (HYM, HC30-30) or two time limits (HC30-90 and HC30-180), yielded CG of similar average size (ranging from 13.2 to 14.1 observations) but quite different average time spans (from 21.4 for HYM to 43.0 for HY30-180). The loss of information was also variable among these strategies. HC30-180 yielded the lowest loss (0.7% of the data and 7.2% of the herds) over all definitions.

Table 2 shows the residual variance estimate under different definitions of CG obtained in the Bayesian analysis under model [1], as well as the between and within CG variances under each definition for model [2]. The residual variance estimate and the within CG variance became progressively larger as the period of time involved in the definition of the CG increased. As expected, HYS and HC90-90 had a worse performance than the other models. Models HYM and HC30-30 had the smallest residual and within CG variance, but the ratio of the within to the total variance was very close for all models, except for HYS and HC90-90. For the between CG variance, HYS and HY90-90 also had the worst performance. The between CG variance was similar for the other models.

Table 3 shows the accuracy of prediction for direct and maternal genetic values for groups of animals that should have the largest accuracies for each type of genetic value.Average accuracies for predicting genetic values were small, partially due to the relatively low degree of connectedness among herds in this population. As expected, accuracies were larger for sires of progeny with data. Given that the variance components used to compute accuracies were the same for all models, differences among accuracies should be due to differences in amount of information available for each animal and/or to differences in the structure of the data. The HYS and HC90-90 definitions had the largest accuracies, particularly for predicting direct genetic values of sires of calves with records.

Table 1. Number of herds, records and number of CG, and weighted averages for CG size and time span (d) for different definitions of CG

	HYM	HC30-30	HC30-90	HC30-180	HYS	HC90-90
No. herds	67	70	75	77	70	75
No. records	10936	11529	12481	12652	12144	12375
No. CG	827	838	899	895	557	513
CG size	13.2	13.7	13.9	14.1	21.8	24.1
Time span	21.4	25.2	36.0	43.0	56.2	68.6

Table 2. Posterior mean and 95% high posterior density interval (HPD) for the residual variance (σ_e^2) in model [1] and between (σ_b^2) and within (σ_w^2) CG variances for model [2] (kg^2)

	HYM	HC30-30	HC30-90	HC30-180	HYS	HC90-90
σ_e^2						
Mean	450.22	452.81	466.05	473.57	476.30	479.39
HPD	[419.4,479.8]	[418.2,485.2]	[446.0,499.1]	[443.4,509.1]	[443.4,509.1]	[440.1,516.2]
σ_b^2	690.64	686.53	690.80	683.27	650.28	652.82
σ_w^2	316.65	318.37	321.46	325.08	341.79	348.26

Also, the average effective number of progeny was very similar, approximately 15.6, for definitions with shorter periods of time (HYM, HC30-30, HC30-90, HC30-180) and approximately 17.6 for the other definitions. As expected, given the smaller number of levels for the CG effect, the two definitions with 90 d periods (HYS and HC90-90) had larger effective number of progeny, in agreement with the slightly larger accuracy observed for this group of animals. Even though the alternative definitions resulted in relatively large differences in the number of discarded records, the definitions did not substantially change the structure of the information available.

The use of the same parameters in the BLUP analyses with all models allowed inferences about effects of the amount and structure of available information on accuracy of genetic value prediction and comparison of between and within CG variation from different models on the same scale. However, the comparison with the same parameters is not strictly correct because true variance components are not expected to be equal for all models. In particular, the residual variance would be expected to decrease as more environmental variation is accounted for by the CG effect, i.e., for CG spanning shorter periods of time, with no differences expected for the other components. In this study, estimates of the residual variance followed the expected trend but the genetic components were also significantly affected by the definition of CG (results not shown).

Criteria for comparison of models from the Bayesian analysis are shown in Table 4. Predictive ability of excluded observations of the alternative models measured in the total (D_T) and specific data sets (D_s) ranked models in the same order. Low values of both statistics indicate better predictive ability of the model. The D_T values were variably larger than the corresponding D_s values, probably because in the total data set, where no observations are discarded, some CG have very few observations (less than five). The HYS and HC90-90 models showed lower predictive ability of missing observations than HYM and HC30-30 models or models with two time limits. Ranking of models was the same when the log of the MD statistic was considered. In all cases, the BF showed strong evidence favoring definitions with CG spanning shorter periods of time.

Table 3. Number of animals (No.) and means (standard deviation in parenthesis) of accuracies (%) of estimated direct genetic effects for sires of animals with records (SIRES), and calves with records (CALVES) and maternal effects for grandsires of animals with records (MGS) and dams of calves (DAM), and effective number of progeny for the direct effect (N_{ed}) for different definitions of CG

	HYM	HC30-30	HC30-90	HC30-180	HYS	HC90-90
Direct-SIRES: No.	283	300	314	317	304	314
Accuracy.	32.9 (17.6)	32.8 (17.8)	30.9 (16.7)	30.6 (16.7)	37.1(18.4)	38.6 (19.0)
N_{ed}	15.59	15.63	15.69	15.67	17.57	17.64
Direct-CALVES:No.	9,945	10,508	11,457	11,579	10,508	11,315
Accuracy	25.7 (6.3)	25.8 (6.3)	25.8 (6.4)	25.9 (6.4)	26.6 (6.4)	26.8 (6.4)
Maternal-MGS: No.	426	444	457	462	452	453
Accuracy	28.9 (13.6)	29.1 (13.7)	30.0 (13.7)	30.0 (13.9)	30.0 (13.9)	30.5 (13.8)
Maternal-DAM: No.	4983	5259	5652	5724	5493	5617
Accuracy	24.2 (7.8)	24.2(8.0)	24.4 (8.0)	24.4 (8.0)	24.8 (8.0)	24.8 (8.0)

Table 4. D statistics used to measure the predictive ability of unknown observations computed with specific data for each definition of CG (D_s) and with the total data set (D_T), and log marginal densities of the data (LMD_T) computed with the total data set with different definitions of CG

	HYM	HC30-30	HC30-90	HC30-180	HYS	HC90-90
D_s	631.80	635.49	642.44	649.25	663.81	675.59
D_T	642.43	644.08	647.27	650.74	674.18	679.40
LMD_T	-57489.2	-57638.5	-57792.3	-57884.8	-58001.8	-58096.5

4. Conclusions

Overall, it seems that according to the evaluation criteria used in this study, average time span rendered by the different definitions had a major effect in the way the definitions were ranked. Thus, definitions involving shorter periods of time, such as HYM or HC30-30, would be preferred for this population. This is due to the fact that those definitions yield a better adjustment of the environmental changes over time without a large loss of accuracy. However, from the breeders point of view, the loss of information associated with these definitions might be unacceptable. Definition HC30-90, which provided intermediate results but relatively close to the optimum definitions with a much lower loss of information, might provide a compromising solution in this population. Bayesian criteria have proved to have some advantages (in terms of providing measures that summarize model plausibility in a general framework) that are difficult to be overcome in the classical approach.

References

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

Gelfand, A. E., D. K. Dey, and H. Chang. 1992. Model determination using predictive distributions with implementation via sampling-based methods. Pages 147-167 in Bayesian Statistics 4, Bernardo, J. M., J. O. Berger, A. P. David, and A. F. M. Smith, ed. Oxford University Press.

Kass, R. E., and A. E. Raftery. 1995. Bayes Factors. J. Am. Stat. Assoc. 90:773-795.

López-Romero, P, R. Rekaya, and M. J. Carabaño. 2003. Assessment of homogeneity vs. heterogeneity of residual variance in random regression test day models in a Bayesian analysis. J. Dairy Sci. 86:3374-3385.

Newton, M. A., and A. E. Raftery. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. J. Roy. Stat. Soc. B 56:3-48.

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

Sivarajasingam, S. 1993. Comparison of alternative methods of handling contemporary group effects in animal model prediction. J. Anim. Breed. Genet. 110:401-411.

Van Bebber, J., N. Reinsch, W. Junge, and E. Kalm. 1998. A Kalman filter based procedure to cluster adjacent herd-test-days into comparison groups. 49th EAAP Meeting, Warsaw (Poland).

Wiggans, G. R., I. Misztal, and L. D. Van Vleck. 1988. Implementation of an animal model for genetic evaluation of dairy cattle in the United States. In: Proc. of the Animal Model Workshop, Edmonton (Canada). J. Dairy Sci. (Suppl2) 71:54-69.