Improvements in the Manchega genetic breeding programme

Jurado J.J., Serrano M., Pérez-Guzman M.D., Montoro V.

in

Gabiña D. (ed.).
Strategies for sheep and goat breeding

Zaragoza : CIHEAM
Cahiers Options Méditerranéennes; n. 11

1995
pages 133-141

Jurado J.J., Serrano M., Pérez-Guzman M.D., Montoro V. **Improvements in the Manchega genetic breeding programme.** In : Gabiña D. (ed.). *Strategies for sheep and goat breeding* . Zaragoza : CIHEAM, 1995. p. 133-141 (Cahiers Options Méditerranéennes; n. 11)

Improvements in the Manchega genetic breeding programme

J.J. JURADO
M. SERRANO
DEPARTAMENTO DE MEJORA ANIMAL
CIT-INIA
MADRID
SPAIN

M.D. PEREZ-GUZMAN
V. MONTORO
CERSYRA
VALDEPEÑAS, CIUDAD REAL
SPAIN

SUMMARY - In this paper we present the last works applied on the Manchega sheep breed genetic program. Those works deal with improvements of the evaluation model quality, estimation of variance components and selection response and extending part-lactation.

The inclusion of the effect herd-mobile period, which supposes about a 45% of the phenotypic traits variation, was the most important improvement of the evaluation model.

Results about estimation of variance components are not definitive and depends on the method used.

Projection factors for extending part-lactation were calculated. Results provides the possibility to extend incomplete lactation records if at least first and second test-day are known.

Key words: Animal breeding, Genetics, Manchega sheep breed.

RESUME - On présente les derniers travaux appliqués sous le programme d'amélioration génétique de la race ovine Manchega. Ces travaux versent sur l'optimisation du modèle d'évaluation, l'estimation des composantes de la variance, la réponse à la sélection et sur l'extension des lactations partielles.

L'inclusion de l'effet troupeau- période mobile, ce qui représente environ le 45% de la variation phénotypique, a été la modification plus importante du modèle d'évaluation.

Les résultats sur l'estimation des composantes de la variance ne sont pas définitifs et ils en dépendent de la méthode utilisée.

Les facteurs de projection pour étendre les lactations partielles ont été obtenus. Ces résultats montrent la possibilité d'appliquer les facteurs de projection si on connaît au moins les rendements obtenus pendant les deux premiers contrôles.

Mots-clés: Amélioration génétique animale, Race ovine Manchega.

INTRODUCTION

In this communication we present a summary of the last works applied on the Manchega breed genetic program. These works were or will be published in other scientific meetings. This communication consist on three works:

- Improvement of evaluation model quality (Serrano *et al.*, 1994a)
- Estimation of variance components and selection response (Serrano *et al.*, 1994b, Jurado *et al.*, 1995).
- Extending part-lactation to achieve more data (Serrano *et al.*, 1995)

Breed description:

Manchega breed sheep are raised for milk production in the semiarid, range region of central Spain and number 1,3 million adults in about 3300 herds.

Average annual rainfall is 350 mm and temperatures go down to -5°C in the winter and up to 42°C in the summer are normal. Such conditions conform it as an arid region in where the Manchega breed shows its main characteristic: the adaptation to the background.

Average milk production of the breed is 126.9 l. in 120 lactation days. Milk has a 7.3% average content of fat, 5.6% of protein and 18.5% of dry extract. The milk is mainly used in the production of the "Manchego cheese", a highly appreciated product for consumers with a legally registered name. The cheese production is about 3000 MT/year.

Breeders association called (AGRAMA) keeps and manages the pedigree records and the "racial standard". There is also a technical center (CERSYRA) where artificial insemination (AI) is provided and production controls are kept.

In 1986 a selection program (ESROM) was initiated to improve the breed genetic level. The objective is to increase the quantity of milk produced using as criterion the milk production standardized to 120 days of lactation and 6% of fat (Montoro *et al.*, 1993).

Because the high number of herds, with an average size of 260 heads/herd, and without a controlled interchange of animals among them, a reference sire scheme using AI to connect herds was set up. Actually there are 120 connected herds, with 15000 inseminations made per year.

Genetic evaluation is made with the BLUP methodology using a repeatability animal model. The sires to connect herds are evaluated by their daughters milk production, and those with greater breeding values are selected as sires. These males are mated with the best females in each herd. Some of the males-progeny of these matings are tested in the technical center. The best of these tested males will replace the reference sires with smaller breeding values. So reference sires are used in a double purpose: to connect herds and to improve average genetic performance.

IMPROVEMENT OF THE EVALUATION MODEL QUALITY

Material and Methods:

In this work we have used the accumulated lactations until 1994. The files consist in 26.668 lactations and 23.043 animals with and without data for the following traits: normalized milk (standardized to 120 days and 6% fat, LN), whole milk (standardized to 120 days, LT), fat percentage (standardized to 120 days, PG) and protein percentage (standardized to 120 days, PP).

The quality of the model was determined by calculating the determination coefficient (CD), using the GLM procedure of SAS statistical package (SAS,1990) for the fixed models, and a FORTRAN program, BLUPAM-VERROR (Jurado and Hernandez, 1990), for the mixed models. The original fixed model is denominated GLM1. The improvement of this model consist of:

- Inclusion of the effect "number of days between lambing and the first control date" as a fixed effect with three levels (int).
- Inclusion of the effect "lambing number-lambing age", (nped).
- Definition of "mobile period" inside the effect herd-year-season, in order to make the factor "herd-mobile period", (rem).

The fixed effect "herd-mobile period", (rem), (Wiggans *et al.*, 1988) was implemented by the AMV101 program (Rozzi, 1992) modified for our conditions. Groups of at least six observations were formed for periods of one or two months. Periods of three months needed at least ten observations per group, (Ugarte, 1990). Thus, we avoid making small subclasses, that would result in progeny group comparisons ineffective (Chauhan and Hill, 1986).

Results and Discussion:

The GLM1 test belongs to the 1993 genetic evaluation fixed model, original model (Serrano *et al.*, 1993). In this model "lambing season", (est), was included as an effect distributed in four levels ((1) December to February, (2) March to May, (3) June to August and (4) September to November) while "lambing age", (ed), and "lambing-first production control date interval", were not included. However, these two last effects were added to model GLM6.

GLM1: $y = reb + ap + est + tp + an + np + e.$ Original model.
GLM6: $y = rem + tp + nped + int + e.$

reb = herd (54 levels).
ap = lambing year (15 levels).
est = lambing season (4 levels).
tp = number of lambs at lambing (4 levels).
an = birth year (23 levels).
np = lambing number (4 levels).
rem = herd-mobile period (679 levels).
nped = lambing number-lambing age (102 levels).
int = lambing-first control date interval (3 levels).
e = residual.

In model GLM6, "birth year" effect, (an), was taken out, in order to include this as genetic group in future evaluations. The effect "lambing number-lambing age" (nped) was made grouping the levels with few observations. In this model all fixed effects were highly significant ($p > 0.0001$).

Table 1 shows the percentage of the total variation explained by the fixed effects included in model GLM6.

GLM1 and GLM6 showed a CD value for milk production of 0.32 and 0.43, respectively.

Table 1. Percentage of variation explained by the fixed effects included in model GLM6. LN = Normalized milk; LT= Whole milk; PG= Fat percentage; PP = Protein percentage.

MODEL	EFFECT	LN	LT	PG	PP
MODEL GLM6	rem	39.99	48.63	54.15	37.26
	nped	2.18	1.36	0.35	0.63
	tp	0.86	1.11	0.11	0.02
	int	0.04	0.09	0.09	0.49

In model GLM6, the effect herd-mobile period, (rem), explain the greatest percentage of the total variation for all characters analyzed.

Table 2 shows BLUPAM-VERROR (5000 iterations of Gauss-Seidel process) determination coefficients, (CD), and residual variances, (VE), for the mixed models which include the fixed effects of GLM1 (MOD1) and GLM6 (MOD6), and animal genotype (u) and permanent environmental effect (c) as random effects.

It has been observed an increase of the Determination Coefficient (CD) and a decrease of the Residual Variance (VE) from MOD1 to MOD6 for all traits.

Changes in the ranking of the males in MOD1 with respect to MOD6, for all characters studied, are observed.

Females were also affected by models in their breeding values, this effect being less important, due to a still very small selection pressure of ESROM on them.

These results show the importance to choose the model of evaluation to obtain the breeding values of animals and to make the ranking of selected males.

The accuracy of the genetic predictions was not affected substantially by changes in the model of evaluation. However, it increases systematically as the number of daughters to evaluate a male augments.

Table 2. Determination Coefficients (CD) and Residual variances (VE) for the mixed models:

MOD(1) : $y = reb + ap + est + tp + an + np + u + c + e.$
MOD(6) : $y = rem + tp + nped + int + u + c + e.$

CHARACT.	CD		VE	
	MOD1	MOD6	MOD1	MOD6
LN	0.58	0.65	1468.67	1244.14
LT	0.65	0.70	943.53	812.64
PG	0.65	0.77	0.79	0.52
PP	0.65	0.72	0.12	0.09

LN = Normalized milk.
LT = Whole milk.
PG = Fat percentage.
PP = Protein percentage.
u = genotype.
c = permanent environmental effect.
Fixed effects specified previously.

Conclusions:

The model of evaluation affects breeding values predictions of animals for all traits studied, mainly for the fixed effects and interactions considered.

The inclusion of the effect "herd-mobile period" (rem) improve the quality of the model (determination coefficient increment). As average, this effect explain a 44.7% of the total variability of fixed effects for the traits studied. Also the effects "lambing-first production control date" (int) and "lambing number-lambing age" (nped) improve the fixed and mixed models.

Model change not only affects the determination coefficient but also, and in a more important way, animal position changes in the genetic ranking which is essential to choose breeding animals for the future.

ESTIMATION OF VARIANCE COMPONENTS AND SELECTION RESPONSE

In this study we compare first results of a Gibbs Sampling method (GS) in a Bayesian application with the Restricted Maximum Likelihood method (REML) to estimate variance components and genetic trends of several traits in Manchega sheep breed.

Material and Methods.

28099 lactations records of 18622 ewes collected from 1986 to 1992 were used in this study. The traits analyzed were Normalized milk (LN), Whole milk (LT) and Fat (PG) and Protein (PP) percentages

The model used was that described by Serrano *et al.* (1994a) and previously denoted as MOD6. In this case, rem effect was distributed in 731 levels, nped in 161 levels, int in 3 levels (<44, 45-52,>52 days), u in 18622 levels and c in 9477 levels.

Mixed model matrix notation was:

$$y = X\beta + Z_1 u + c + e$$

where:

y is a nx1 vector of observations (lactations); X, Z_1, and Z_2 are incidence matrices relating location parameters in ß and variables in u and c to y; respectively; ß is the vector of fixed effects; u is a vector of random additive genetic effects; c is a vector of random permanent environmental effects and e is a random vector of residual.

The distribution assumptions were:

$$u \sim N(0, A\sigma^2_u),\ c \sim N(0, I_c\sigma^2_c) \text{ and } e \sim N(0, I_e\sigma^2_e)$$

where σ^2_u ,σ^2_c and σ^2_e are variance components and **A** the numerator of Wrights's relationship matrix. Rank of X was 899. u, c and e were supposed to be independent variables.

-REML analysis (DFREML).

Estimates of the variance components were obtained by REML using DFREML programs (Meyer, 1993) for each character. Then variance component values were used in Henderson's mixed model equations to obtain BLUP estimates of animal breeding values. Finally selection response was estimated by constructing appropriate function of the predicted breeding values.

-Bayesian analysis via Gibbs Sampling (GS).

Methodology from Wang *et al.*, (1994) and Sorensen *et al.* (1994) was used for GS analysis. Inferences about fixed effects estimates, genetic predictions and variance component estimates were based on its marginal posterior distributions. The marginalization of the functions were achieved by means of the GIBBS Sampler.

Two sets of priors were used in this study: Flat priors for the three variance components (G1) and informative inverted χ^2 (G2) independents for u,c y e.

A FORTRAN77 program was written to generate the samples and subroutines from NUMERICAL RECIPES (1990) were used to generate random numbers.

Results and Discussion.

DFREML variance component estimates show different values according to the starting value used. This was more evident in the case of σ^2c (variance of permanent effects).

The results averaged from several ejecutions of DFREML program and the modal values obtained by GS in the two cases, G1 and G2, are summarized in Table 3.

Table 3. Heritability (h^2) and repeatability (r) estimates for Normalized milk (LN), Whole milk (LT) and Fat (PG) and Protein (PP) percentages obtained by DFREML and GIBBS Sampling with different priors (G1 and G2).

TRAIT	DFREML		GIBBS SAMPLING			
			MODE G1		MODE G2	
	$h^2 \pm$ SD	r	h^2	r	h^2	r
LN	0.29 ± .024	0.48	0.34	0.34	0.23	0.37
LT	0.33 ± .024	0.50	0.30	0.38	0.27	0.40
PG	0.23 ± .005	0.36	0.24	0.24	0.19	0.28
PP	0.35 ± .020	0.47	0.27	0.40	0.29	0.41

In a Bayesian context, information contained in both priors has more influence over variance components estimates than the information provided by the actual data.

Genetic trend per year and total response in the whole period of time are shown in Table 4.

Table 4. Genetic trend per year (b) and total response in the whole period of time (tr) for Normalized milk (LN) and Whole milk (LT) obtained by DFREML and GIBBS Sampling with different priors (G1 and G2).

PARAMETER	DFREML	GIBBS SAMPLING			
		MEAN G1	MEAN G2	MODEG1	MODEG2
b(LN)	0.842 ± 0.62	1.239 ± 0.59	1.008 ± 0.52	1.232	0.864
tr(LN)	2.37	2.479 ± 1.19	2.016 ± 1.03	2.464	1.728
b(LT)	0.805 ± 0.71	1.138 ± 0.46	1.076 ± 0.44	1.063	0.936
tr(LT)	2.50	2.276 ± 0.93	2.153 ± 0.88	2.127	1.873

DFREML estimated values are smaller than those obtained by GS (without differences between priors). Total selection response means were similar with both priors (G1 and G2) with GS method, however mode values were different. Although these trends and response values are not significantly different from zero, the analysis suggests that they might be positive.

Conclusions

- Differences between DFREML variance component estimates depended on initial parameter values that would be explained because likelihood function is not well defined.
- The Bayesian approach shows that information contained in the data is less informative than that of priors, due to the poor quality of the records.
- Genetic trend and total response showed slightly positive values.
- GS was found to be easy to implement once BLUP code is developed, but, computational requirements were extremely high.

EXTENDING PART-LACTATION OF SEVERAL TRAITS IN MANCHEGA SHEEP BREED.

The objective of this work is to estimate the extension coefficients for several traits to achieve more data for sire evaluation.

Material and Methods.

In this work we have used the accumulated lactations until 1995. The files consist in 47989 records from the recording milk control.

Projection factors for extending milk, fat, protein and dry extract yields from part-lactation records were calculated. Methodology from Auran (1976) and Danell (1982) was applied to 63225 monthly records for each trait.

Monthly extension factors (c_i) for all characters included in the analysis were calculated through the expression:

$$c_i=T_i/L_i$$

Extended lactations to 120 days (LT_{120}) were calculated as:

$$LT_{120}=LT_i+b(120-d_i)L_i,$$

where:

T_i is the cummulative yield since the last known test-day till 120 days.
L_i is the last known test-day.
LT_i is the cumulative yield till the last known test-day (L_i).
d_i the number of days between lambing and the last known test-day (L_i).
b is the regression coefficient of c_i on $(120-d_i)$.

Results and Discussion.

b values for milk, fat, protein and dry extract were estimated through a linear model including herd production level (≤122 kg, >122 kg), lambing season (1= Dec, Jan, Feb; 2= Mar, Abr, May; 3= Jun, Jul, Aug and 4= Sep, Oct, Nov) and lambing number (1,2,≥3) as fixed effects, the number of days between last test-day and 120 as a covariable nested to the previous effects and c_i as independent term.

The average b value from the twenty-four values obtained for milk, fat, protein and dry extract yields ($p<0.0001$) with SAS statistical package (SAS, 1990) were 0.78, 0.90, 0.85 and 0.82, respectively. Table 5 show b values for production level, lambing-season and lambing-number in each trait.

Table 5. Production level (PN), lambing-season (LS) and lambing-number (LA) b values for milk, fat, protein and dry extract yields.

Character	PN		LS				LA		
	≤122	>122	1	2	3	4	1	2	≥3
Milk	0.79	0.76	0.83	0.79	0.75	0.74	0.85	0.77	0.71
Fat	0.91	0.89	0.95	0.95	0.84	0.84	0.98	0.89	0.82
Protein	0.87	0.83	0.88	0.86	0.86	0.80	0.94	0.85	0.78
D.extract	0.84	0.81	0.87	0.84	0.79	0.78	0.90	0.82	0.75

To test the goodness of the extension method, correlation coefficients between real and first, first-second and first-second-third test-day extended yield were calculated. Table 6 shows these coefficients for all the traits.

Table 6. Correlation coefficients between real and first, first-second and first-second-third test-day extended yield for milk, fat, protein and dry extract.

Character	First	First-second	First-second-third
Milk yield	0.94	0.98	0.99
Fat yield	0.61	0.82	0.94
Protein yield	0.93	0.97	0.99
Dry extract yield	0.93	0.97	0.99

Correlation coefficients for milk, protein and dry extract yields were very high for the three situations analyzed, however those for fat were lower, as we expected, because the poor quality of such records.

Conclusions.

Such results reveal the possibility of extending part-lactation if at least first and second test-day are known. This extension would be more accuracy in the cases of milk, protein and dry extract yields.

In the Manchega breeding program is possible to extend 3848 lactations when the third control is known and 2144 when the second one is known. This suppose a total of 5992 extended lactations. Thus there are 34295 normalized and 5992 extended lactations.

REFERENCES:

AURAN, T. 1976: Studies on monthly and cumulative monthly milk yield records. IV. Estimating total lactation from part-lactation. Act Agric. scand. 26:10-17.

CHAUHAN, V. P. S. and HILL, N. G., 1986: Seasonal grouping in herd-year-season model of sire evaluation. Anim. Prod. 43: 63-71.

DANELL, B. 1982: Studies on lactation yield and individual test-day yields of Swedish dairy cows. IV. Extension of part-lactation records for use in sire evaluation. Act agric. scand. 32:103-113.

JURADO, J.J. and HERNANDEZ, D., 1990: Catálogo de Software de Interés en Agricultura. M.A.P.A.. Programa nº 113, pag: 81.

JURADO, J.J.; SERRANO, M.; PEREZ-GUZMAN, M.D. and MONTORO, V. 1995: Estimation of variance components and selection progress in Manchega sheep breed using the GIBBS sampler. ITEA (Submitted).

MEYER, K. 1991: Program DFREML version 2.1.08.

MONTORO, V.; PEREZ.GUZMAN, M.D.; GARZON, A.I.; GARDE, J.J. and PEREZ, S.S., 1993: Evolución y problemas técnicos en el desarrollo del esquema de selección de la raza ovina Manchega. I.T.E.A. Vol. 89(A), (2), Pag: 172-180.

NUMERICAL RECIPES, 1990: FORTRAN version. Cambridge University Press.

ROZZI, P., 1992: Personal communication.

SAS LANGUAGE, 1990: Reference V.G. (1042 pp.) Cary, N.C. SAS Institute Inc.

SERRANO, M.; JURADO,J.J.; PEREZ-GUZMAN, M.D. and MONTORO, V. 1993: Estudio de la tendencia genética y fenotípica de diversos caracteres en la raza ovina Manchega. ITEA. 12, Vol. 1:233-235.

SERRANO, M.; PEREZ-GUZMAN, M.D.; MONTORO, V. and JURADO,J.J. 1994(a): Mejora del modelo de evaluación genética en el esquema de selección de la raza ovina Manchega. VII Reunión Nacional sobre Mejora Genética Animal. Lérida. Pags: 145-147.

SERRANO, M.; PEREZ-GUZMAN, M.D.; MONTORO, V. and JURADO, J.J. 1994(b): Genetic parameters estimation and selection progress in the selection program of ovine Manchega breed. 45th Annual Meeting of the European Association of Animal Production. Edimburg, 5-8 September.

SERRANO, M.; PEREZ-GUZMAN, M.D.; MONTORO, V. and JURADO, J.J. 1995: Extending part-lactation of several traits in Manchega sheep breed. 46th Annual Meeting of the European Association for Animal Production. Praha. (Submitted).

SORENSEN, D.; WANG, C.S.; JENSEN, J. and GIANOLA, D. 1994: Bayesian analysis of genetic change due to selection using GIBBS Sampling. Genet. Sel. Evol (in press).

UGARTE, E., 1990: Estudio de los grupos contemporáneos en la valoración genética de reproductores. Análisis de datos obtenidos por simulación en el ordenador. Tesis Doctoral. Falcultad de Biológicas, Universidad Complutense. Madrid. España.

WANG, C.S.; GIANOLA, D.: SORENSEN, D.A. and JENSEN, 1994: Response to selection for litter size in Danish Landrace pigs: a Bayesian analysis. Theor. and Appl. Genet. 88: 220-230.

WIGGANS, G.R.; MISZTAL, I. and VAN VLECK, L.D., 1988: Implementation of an Animal Model for genetic evaluation of dairy cattle in the United States. Proceeding of the Animal Model Workshop. J. Dairy. Sci., Vol 71. Suppl. 2: 54-69.