

LONGEVITY ANALYSIS IN SPANISH HOLSTEIN-FRIESIAN CATTLE

Z. Chirinos¹, M.J. Carabaño¹ and D. Hernández²

¹ Departamento de Mejora Genética Animal, INIA Apdo. 8111 28080 Madrid, Spain

² Departamento Técnico, CONAFE, Ctra. Andalucía, Km. 23,6 28340 Madrid, Spain

INTRODUCTION

Functional traits have become more and more important in dairy cattle, particularly in the Holstein breed, highly selected for production in the past. Functional longevity (Ducrocq, 1987) has become a common way of measuring fitness of the cows. This is particularly the case when secondary traits are not fully recorded, such as in the Spanish Holstein population. Survival analysis techniques are considered the method of choice for the analysis of longevity data due to their ability of managing incomplete records and allowing a more proper modelling of this type of data (Strandberg and Sölkner, 1996 ; Essl, 1998 ; Boettcher *et al.*, 1998). The aim of this paper was to analyse longevity data from the Spanish Holstein cattle using survival analysis techniques in order to propose a genetic evaluation model for this population.

MATERIAL AND METHODS

Data. Five data sets from different regions in Spain, Andalusia (AN), Basque country (BC), Catalonia (CA), Galicia (GA) and Navarre (NA), have been used to investigate the possible regional or sampling effect on the estimated parameters. Longevity of a cow was computed from the date of the first and last day of milk recording. Data were complete when the difference between the date of last recording of the cow and the date of last recording in the herd was over 7 months. Otherwise, data were censored.

Statistical Analysis. Data from AN and CA were analysed in first place under a Cox model to investigate the definition of environmental factors to be included in the future genetic evaluation model and to check the validity of assuming proportionality of hazards and Weibull distribution for subsequent analyses. Secondly, all five data sets were used to study the genetic components of the model under a Weibull model. The Survival Kit package (Ducrocq and Sölkner, 1998) was used to analyse the data.

For the analysis under a Cox model, the hazard function was defined as :

$$h(t) = h_0(t) * \exp \{ AFC_i + L_j(t) + S_k(t') + M_l(t) + F_m(t) + P_n(t) + HYS_o(\tau) + HS_p(\tau') \}$$

where, $h(t)$ is the hazard of a cow at t days after first calving, $h_0(t)$ is the baseline hazard, assumed arbitrary, AFC is the age at first calving effect, $L(t)$ is the lactation number effect, changing t days after the first calving at the date of each new calving (lactations 1, 2, 3, 4, 5, 6, 7 +), $S(t')$ is the stage of lactation effect changing in each lactation at t' days after the current calving ($t' = 30, 60, 90, 180, 240$ and 300), $M_l(t)$, $F_m(t)$ and $P_n(t)$ are the effects of milk, fat and protein percent deviation with respect to the herd mean in the year of calving, respectively, changing at the day t of each calving (8 classes from -30% to $+30\%$ deviations), $HYS_o(\tau)$ is

the effect of the herd-year-season of calving, changing at calendar date τ , which is the first day of months 3, 6, 9 and 12 in each year ; $HS_p(\tau')$ is the p^{th} class of the percent change in herd size, which is measured the first of January of each year (11 classes ranging from -30% to $+40\%$, including an unknown change class).

For the Cox analysis, likelihood ratio tests were obtained for the AN and CA data sets. An analysis considering a stratified baseline according to the AFC classes was also performed in order to check the validity of proportional hazards. AFC was the only effect not changing with time and stratification implies a less restrictive assumption than the proportional hazard with a common baseline assumed. The estimated survivor functions for each strata i ($\hat{S}_i(t)$) were obtained and the $\log[-\log(\hat{S}_i(t))]$ plotted against $\log t$ to also check the validity of assuming a Weibull distribution for subsequent analyses. Straight lines would indicate that the Weibull assumption is correct and parallel lines would imply that the proportional hazards assumption holds for this effect and a common baseline could be used.

For the analyses under a Weibull model, the model previously shown was slightly modified to include the additive genetic effects of the sire and maternal grandsire of the animal. In addition, the contemporary group effect, HYS, was treated as random and the interaction between lactation number and stage of lactation was included in the model. Heritabilities in the log and original scales were computed following expressions in Vollema *et al.*(2000).

RESULTS AND DISCUSSION

The first columns in table 1 show the number of data, average longevity, percentage of censored information and number of bulls used in the Weibull analyses with the model including genetic effects.

Table 1. Number of data, average longevity in days (LON), percentage of censored information and number of bulls. Estimates of Weibull parameters (ρ and intercept), parameters associated to the contemporary group (γ) and sire effects (V_s) and heritabilities in the log (h^2 log) and original scale (h^2 orig)

| Data set | No. data | LON | Cens. | No. bulls | ρ | Intercept | γ | V_s | h^2 log | h^2 orig |
|-----------|----------|---------|-------|-----------|--------|-----------|----------|-------|-----------|------------|
| Andalusia | 21058 | 835.14 | 52.66 | 3761 | 2.058 | -13.43 | 2.20 | 0.012 | 0.022 | 0.048 |
| Catalonia | 46828 | 940.77 | 48.33 | 4326 | 1.693 | -11.27 | 2.99 | 0.016 | 0.030 | 0.073 |
| Galicia | 78842 | 917.25 | 64.59 | 7563 | 1.265 | - 8.63 | 3.12 | 0.015 | 0.029 | 0.094 |
| Navarre | 21629 | 1003.84 | 45.23 | 2864 | 2.177 | -14.516 | 2.69 | 0.020 | 0.038 | 0.076 |
| Basque C. | 42827 | 1054.25 | 42.69 | 3397 | 2.284 | -15.245 | 2.41 | 0.016 | 0.029 | 0.059 |

A likelihood ratio test under the Cox model indicated that all factors included in the model had a significant effect on the hazard, for both data sets analysed. Maddala's R^2 values were 0.56 and 0.57 for the AN and CA data sets, respectively. Figure 1 shows the plot of $\log[-\log(\hat{S}_i(t))]$ against $\log t$ for the AN data set when a AFC stratified baseline is considered. A similar pattern was observed for the CA data set (not shown). Lines corresponding to the different strata were near parallel for most part of the time, indicating that the assumption of a common baseline for all AFC levels might be reasonable. The lines followed a nearly constant slope with two exceptions, one at the beginning of the productive life and another at around the first part of the second lactation. The first may be due to the fact that very little information exists on culled animals in the first stages of lactation. The second change in slope might be due to the fact that culling seems to be more intense during the first lactation and, probably, at the beginning of the second lactation. All data sets showed similar trends for all the factors (not shown). Given that, except for the first part of the productive life, it seems that a Weibull distribution might not be an unrealistic assumption, a Weibull model was applied to further investigate the model to be used in the genetic evaluations.

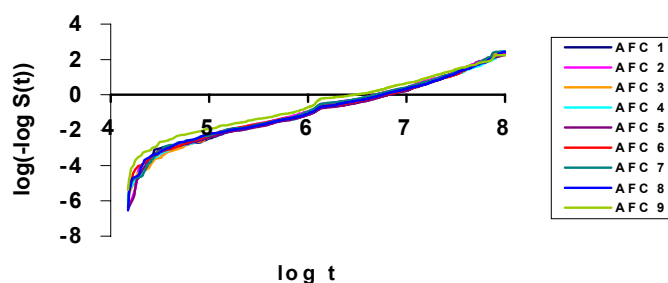


Figure 1. Plot of $\log[-\log(\hat{S}_i(t))]$ ($i = 1, 9$) against $\log t$ for the Andalusia data set when the baseline is stratified according to age at first calving (AFC)

Figure 2 shows results for the effect of different factors on the risk of being culled under a Weibull model. Risk of culling was higher for older cows at first calving, for first and second lactation, for herds decreasing their size over a 30 % and for low milk and protein producers, being fat production less important. The factors with an associated higher risk were an extreme decrease in herd size and a very low production level.

The last columns of table 1 show the estimates of the Weibull parameters, variances associated to random effects and heritabilities in the log and original scale. Values for the estimated parameters are within the values found by other authors (see e.g. Vollema, 1998), ranging between 0.05 in AN and 0.09 in GA for the heritability in the original scale. Given that the regions involved in the study represent a wide range of production systems within the country and the relatively small differences in the estimated heritabilities, it would seem reasonable to assume a common set of parameters for the joint national evaluation.

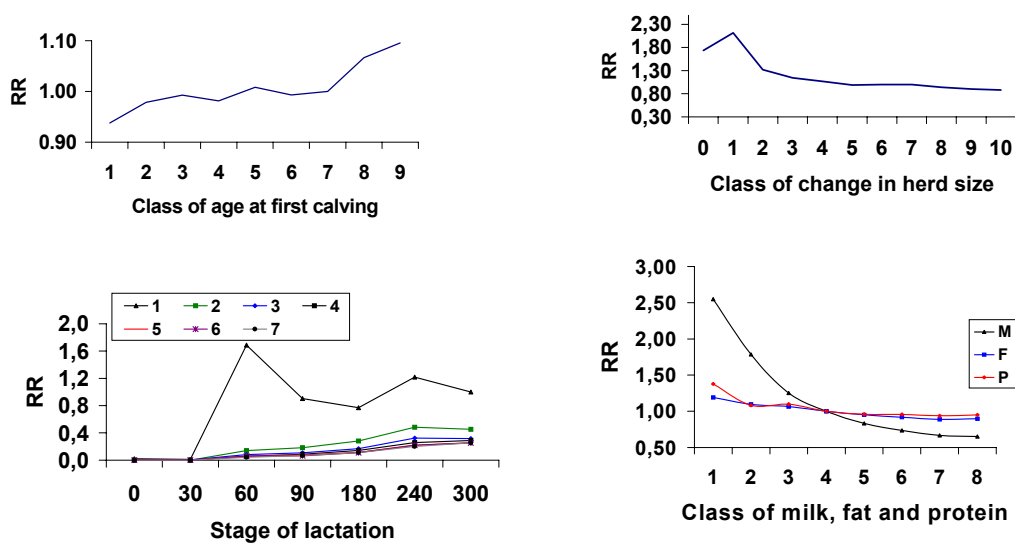


Figure 2. Relative risk (RR) of environmental factors on the hazard of being culled

CONCLUSION

The environmental factors proposed in this study showed a significant influence on the hazard of a cow to be culled. Stratification according to the AFC effect does not seem necessary and a Weibull model with a unique baseline might be appropriate for the national evaluation. Collecting more information on culling during the first part of the productive life would be needed in order to investigate the large deviation from the expected pattern from a Weibull model in that period of time. The consistency of the results among regional samples indicate that the proposed model could be applied at the national level.

REFERENCES

- Boettcher, P.J., Jaraith, L.K. and Dekkers, J.C.M. (1998) *Proc. 6th WCGALP*. **23** : 363-366.
 Ducrocq, V.P. (1987) PhD Thesis, Cornell University, Ithaca, New York, USA.
 Ducrocq, V.P. and Sölkner, J. (1998) *Proc. 6th WCGALP*. **27** : 447-448.
 Essl, A. (1998) *Livest. Prod. Sci.* **57** : 79-89.
 Strandberg, E. and Sölkner, J. (1996) *Interbull Bull.* **12** : 111-119.
 Vollema, A.R. (1998) *Anim. Breed. Abstr.* **66** : 781-802.
 Vollema, A.R., Van der Beek, S., Harbers, A.G.F. and De Jong, G. (2000) *J. Dairy Sci.* **83** : 2629-2639.