

Evaluating The Connectedness Of A Dairy Sheep Selected Population With Low Rates Of Artificial Insemination And Controlled Natural Mating

S. Salaris^{*}, Sara Casu^{*}, M.G. Usai^{*}, P. Fresi[†], and A. Carta^{*}

Introduction

In breeding programs the use of BLUP methodologies allows to perform genetic evaluations across management units. However, for providing a reliable ranking of selection candidates, management units should be connected and animals belong to a unique genetic population. The evaluation of connectedness is crucial in selection programs of dairy species where high rates of artificial insemination (AI) are not feasible. The size of the selected population of the Sarda dairy sheep breed represents approximately 6-7% of the whole population. The main feature of the breeding program is the large application of the controlled natural mating (NM), *i.e.* grouping ewes with a single ram during the reproduction period (“mating group”), in order to assign the correct sire on the basis of the lambing date. AI started in the early nineties at experimental level and has been applied to the whole selected population since 1995. AI is mainly used to create genetic links between flocks either directly through rams used by AI or rams born from AI and used by NM. Over the last years, from 8 to 10% of replacement ewes are born by AI (Salaris et al. 2008). Several methods have been proposed to evaluate connectedness of a selected population (Laloë et al. 1996; Huismann et al. 2006; Khuen et al. 2007; and Fouilloux et al. 2008). Kennedy and Trus (1993) stated that the mean squared error of prediction of difference between couples of selection candidates (PEVD) is the most appropriate measure of connectedness. It represents the difference between the sum of prediction error variances (PEV) of EBVs of two animals and their prediction error covariance (PEC). PEC increases for pairs of animals evaluated in the same management unit or connected through common pedigree ties. When connectedness between pairs of animals does not exist, PEC equals zero and their PEV will be independent. Laloë (1993) proposed the square of the correlation between the predicted and true EBV differences (CD). When animals are unrelated, PEVD and CD are equivalent. Then an increase in CD corresponds to a decrease in PEVD, and the use of both measures leads to the same conclusions. Moreover, he argued that contrasts between disconnected animals, even if theoretically estimable, must be considered null and not predictable. Lewis et al. (1999) proposed the correlation of breeding value prediction errors (r_{ij}). Khuen et al. (2008) concluded that r_{ij} was better than CD to evaluate the accuracy of contrasts since it takes into account the difference in genetic means between subpopulations. In this study we measured the connectedness of the Sarda breed selected population by using the described objective

^{*} Agris Sardegna, DIRPA, Research Unit: Genetics and Biotechnology, 07040 Olmedo Italy

[†] Associazione Nazionale della Pastorizia – ASSONAPA, 00155 Rome Italy

measures with the aim of evaluating the accuracy of the genetic ranking of rams and the strategy of combining AI with NM.

Material and methods

308,503 first lactation records of yearlings born from known sire (9,330 sires) between 1995 and 2007 were extracted from the database used for the genetic evaluation of 2008 provided by the Italian Association of Sheep and Goat Breeders (ASSONAPA). The ewes yielded in 10,034 management units (flock-year of production combinations). Data were analyzed with a mixed model including the management unit fixed effect and the random sire effect.

Two measures were calculated:

- the coefficient of determination of the contrast (CD, Laloë 1993)

$$CD(x) = \frac{\mathbf{x}'(\mathbf{A} - \lambda\mathbf{C})\mathbf{x}}{\mathbf{x}'\mathbf{A}\mathbf{x}}$$

where \mathbf{x} is the vector of coefficients defining the contrast, \mathbf{C} is the inverse of the fixed effect absorption matrix, $[\mathbf{Z}'\mathbf{Z} - \mathbf{Z}'\mathbf{X}(\mathbf{X}'\mathbf{X})\mathbf{X}'\mathbf{Z} + \lambda\mathbf{A}^{-1}]^{-1}$, \mathbf{X} and \mathbf{Z} are the incidence matrices of the management unit fixed effect and the sire random effect respectively, \mathbf{A} is the numerator relationship matrix, λ was calculated as the ratio of the residual (σ_e^2) and the sire (σ_s^2) variances. σ_s^2 and σ_e^2 were imposed equal to 80 L² and 1,040 L² following Carta et al. (1998).

- the connectedness correlation (r_{ij} , Lewis et al. 1999)

$$r_{ij} = \frac{PEC(\hat{u}_i, \hat{u}_j)}{\sqrt{PEV(\hat{u}_i)PEV(\hat{u}_j)}}$$

where PEV is the prediction error variance and PEC is the prediction errors covariance of the EBV (\hat{u}_i and \hat{u}_j) of the i^{th} and j^{th} sires calculated as the diagonal and the off diagonal elements of \mathbf{C} multiplied by σ_e^2 respectively.

These two measures were calculated for all possible pairs of 1,404 sires with daughters born in 2007 (984,906 pairs). Descriptive statistics and Pearson correlation of the 2 measures were calculated.

A clustering iterative process was used to group sires into connected subpopulations using r_{ij} as a measure of distance between sires. At the start of each process each sire begins in a cluster by itself. The iterative process continues joining at each step the 2 sires (or clusters) with the highest r_{ij} in a new cluster. When a cluster includes more than one sire, r_{ij} with the other sires (or clusters) is the average of r_{ij} of the included sires with the other sires (or clusters). The process stops when the highest r_{ij} equals a given threshold. Successive decreasing r_{ij} thresholds were considered from 0.020 to 0 with a step of 0.001. At each threshold, the number and the size of identified clusters were evaluated.

Results and discussion

On average CD was equal to 0.514 and ranged between 0 and 0.945. The average r_{ij} was 0.010 and ranged between 0 and 0.641 with 95% of values below 0.020. The lowest CD value was calculated for a pair of highly related sires (additive relationship coefficient of 0.507) with 2 and 22 daughters respectively yielding in two disconnected management units

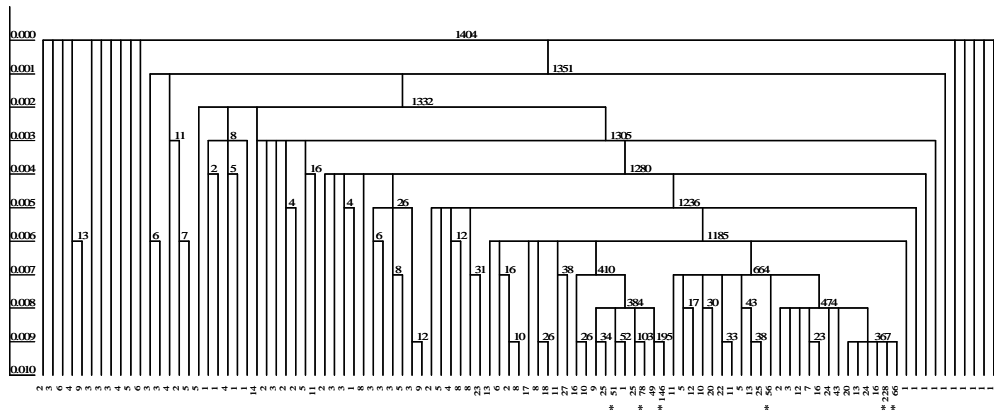


Figure 1: Diagram of clustering procedure results at successively decreasing connectedness correlation (r_{ij}) thresholds (from 0.010 to 0 with a step of 0.001). Horizontal axis represents the number of rams grouped in each cluster at 0.010 r_{ij} threshold level (clusters highlighted with asterisk included more than 50 sires). Vertical axis represents r_{ij} threshold levels. The number of rams grouped in a cluster are reported within the diagram.

without contemporaries. Their PEVs resulted $66.6 L^2$ and ranked in the worst 4% of PEV distribution. PEC was $26.7 L^2$ in the highest 2.5% of PEC values and r_{ij} was 0.40 in the top 0.1% of r_{ij} values. The highest value of CD was detected for a pair of no related sires, with low PEV (3.9 and $5.8 L^2$, respectively), PEC in the highest 13% of PEC values ($0.4 L^2$), and r_{ij} of 0.084 in the highest 0.3% of r_{ij} values. The 3.5% of all possible pairs of sires showed a r_{ij} of 0 with CD values ranging from 0.017 to 0.759.

The overall analysis of these instances showed that r_{ij} measures the degree of connectedness of two sires but it is not useful to estimate the accuracy of their contrast. On the other hand, CD should be considered a reliable measure of the accuracy of a contrast only if the connectedness between two sires measured by r_{ij} is above a sufficient threshold. The different nature of CD and r_{ij} was confirmed by the Pearson correlation value (0.33). According to these assumptions, we used r_{ij} to group sires in clusters on the basis of their connectedness and to identify genetic subpopulations where the use of CD to evaluate the accuracy of the ranking of animals is reliable..

Some authors affirmed that the choice of the r_{ij} threshold at which a population can be considered connected is a subjective matter (Huisman et al. 2006; Fouilloux et al. 2008). Kuehn et al. (2008) proposed two r_{ij} values as benchmarks to use in monitoring risks of bias associated with animals' comparison across flocks: 0.005 for a "good" connectedness and 0.010 for a "superior" connectedness. Results of the clustering process applied in this study are reported in Figure 1. At the lowest threshold a large cluster including 1,236 sires and other 26 clusters including more than 2 sires were formed. On the other hand, at the 0.010 threshold, 83 different subpopulations including on average 17 ± 31 sires (from 2 to 228 sires) were detected. The result obtained at the 0.005 threshold can be ascribed to the strategy of combining a low rate of AI with an important exchange of NM sires. Actually, Salaris et al. (2008) showed that the average number of natural mating groups per flock was 1.9, with 52% of flocks with only 1 sire, 38% with 2 or 3 sires and 10% with more than 3 sires. Over

the last twenty years the percentage of NM sires born by AI increased from 2% to 32%, connecting 19% of flocks not directly involved in the AI program. As far as the genetic links created by NM rams are concerned, the average annual percentage of sires born in one flock and used in another one was 62% and the average annual percentage of NM sires with daughters in more than one flock was 23%.

Conclusion

This study showed that the evaluation of the reliability of a genetic ranking of animals based on BLUP methodologies lies on the preliminary detection of genetic subpopulations by using a measure of connectedness between pairs of animals. As far as the Sarda breed selection scheme is concerned, our results confirmed that the current rates of AI combined with a large use of controlled natural mating allows to reach a sufficient level of reliability of the genetic ranking of animals for most flocks. However, in order to increase the connectedness of the selected population, the selection scheme should be modified either increasing the AI rate or improving the management of natural mating rams.

Acknowledgments

This work was realized in the framework of the research program “APQ per la ricerca scientifica e l’innovazione tecnologica - Progetto P5a - Attivazione del Centro di biodiversità animale per la valorizzazione del patrimonio animale con riferimento alla produzione e alla ricerca al servizio dell’allevamento” of the Regional Government of Sardinia (Italy).

References

- Carta, A., Sanna, S.R., Rosati, A., and Casu, S. (1998). *Ann. Zootech.*, 47: 59-66.
- Fouilloux, M.N., Clément, V., and Laloë, D. (2008). *Genet. Sel. Evol.*, 40:145–159.
- Huisman, A.E., Tier, B., and Brown, D.J. (2006). *Livest. Sci.*, 104: 254–267.
- Kennedy, B.W., and Trus, D. (1993). *J. Anim. Sci.*, 71: 2341–2352.
- Kuehn, L.A., Notter, D.R., Nieuwhof, G.J., and Lewis, R.M. (2008). *J. Anim. Sci.*, 86: 536-544.
- Laloë, D. (1993). *Genet. Sel. Evol.*, 25: 557–576.
- Lewis, R.M., Crump, R.E., Simm, G., and Thompson, R. (1999). *Proc. Brit. Soc. Anim. Sci.* 121.
- Salaris, S., Casu, Sara, Fresi, P., and Carta, A. (2008). In *Proc. ICAR 36th Session*, Technical Series n. 13, pages 113-121.