

# Evaluating the accuracy of the genetic ranking of rams in the selected population of the Sarda dairy sheep breed

**Sotero Salaris<sup>1</sup>, Sara Casu<sup>1</sup>, Mario Graziano Usai<sup>1</sup>,  
Pancrazio Fresi<sup>2</sup>, and Antonello Carta<sup>1</sup>**

*<sup>1</sup>Research Unit: Genetics and Biotechnology, DIRPA-AGRIS Sardegna, S.S. 291 Sassari-Fertilia, Km. 18.6, 07100 - Sassari, Italy*

*<sup>2</sup>Associazione Nazionale della Pastorizia – ASSONAPA, Via Togliatti 1587, 00155 - Rome, Italy*

The main feature of the **Sarda breeding program** is the large application of the **controlled natural mating (CNM)** combined with **low rates of AI**.

It is **difficult to extend the use of AI** to the whole selected population due to specific features of the breeding system.

The **current rate of AI (8 to 10% of replacement ewes)** is not sufficient by itself to assure the genetic connectedness between all selected flocks.

The large use of **CNM** with an important exchange of males, mainly sons of AI rams, may counterbalance the low spread of AI.

The combining of CNM with AI allowed to reach **a high number of direct genetic links** between rams, i.e. direct comparisons through lactating daughters in the same management unit (Salaris et al., 2008)

*“A more precise evaluation of the genetic connectedness of the registered Sarda breed population was needed to assess the accuracy of the genetic ranking of animals”*

Lewis et al. (1999) proposed the **correlation of breeding value prediction errors** ( $r_{ij}$ ) as a measure of the connectedness between pairs of sires.

Khuen et al. (2009) used the **average  $r_{ij}$**  to cluster flocks in different groups.

Salaris et al. (2010) proposed to use  $r_{ij}$  as a criterion to assign rams to different genetic subpopulation and the square of the correlation between estimated and true differences between breeding values (**CD**; Laloë et al., 1993) as a measure of the accuracy of the contrast between EBV.

$r_{ij}$  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 aim of this study was to assess the accuracy of the genetic ranking of AI and CNM rams obtained by BLUP method in the selected population of the Sarda dairy sheep breed by using objective measures of connectedness.

A dataset including **308,503 first lactation records** of ewes born from known sires (**9,330**) between **1995** to **2007** was extracted from the 2008 genetic evaluation database managed by the Italian Association of Sheep 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** as fixed effect and the random **sire** effect.

The **connectedness correlation** ( $r_{ij}$ ) and the **coefficient of determination (CD)** of the **contrast** for all possible pairs of 1,404 sires of 21,331 ewes born in 2007 were calculated as:

$$r_{ij} = \frac{PEC(\hat{u}_i, \hat{u}_j)}{PEV(\hat{u}_i) + PEV(\hat{u}_j)} \quad CD(x_{ij}) = 1 - \frac{(PEV(\hat{u}_i) + PEV(\hat{u}_j) - 2PEC(\hat{u}_i, \hat{u}_j))/\sigma_e^2}{a_{ii} + a_{jj} - 2a_{ij}}$$

Where:

- **PEV** is the **prediction error variance** and **PEC** is the **prediction error covariance** between EBV of sire i and j
- PEV and PEC were calculated multiplying the **residual variance** ( $\sigma_e^2$ ) by the diagonal and the off diagonal elements of  $[Z'Z - Z'X(X'X)^{-1}X'Z + \lambda A^{-1}]^{-1}$ , the **inverse of the fixed effect absorption matrix**
- **X** and **Z** are the incidence matrices of the management unit fixed factor and the sire random factor
- $\lambda$  was calculated as the ratio of  $\sigma_e^2$  and the **sire variance** ( $\sigma_s^2$ ) with  $h^2$  equal to 0.28
- $x$  is the contrast between sire i and j
- $a$  is the element of the numerator **relationship matrix A**



A **clustering iterative process** was used to **group sires into connected subpopulations** using  $r_{ij}$  as a measure of distance between sires.

1. The process **begins setting each sire in a cluster** by itself
2. Then the iterative procedure **joins the 2 sires (or clusters) with the highest  $r_{ij}$**  in a new cluster
3. When a cluster includes more than one sire, **the average  $r_{ij}$**  of the included sires with the sires of other clusters **was calculated**
4. The procedure **stops** when the **highest  $r_{ij}$**  between two clusters is below **a given threshold**.

Accordingly to Khuen et al. (2009), two thresholds were chosen:

- ❑ **0.010** as “*superior*” level of connectedness
- ❑ **0.005** as “*good*” level of connectedness

A further *intermediate* threshold of **0.007** was introduced.

At each threshold, the number and the size of identified clusters were calculated.

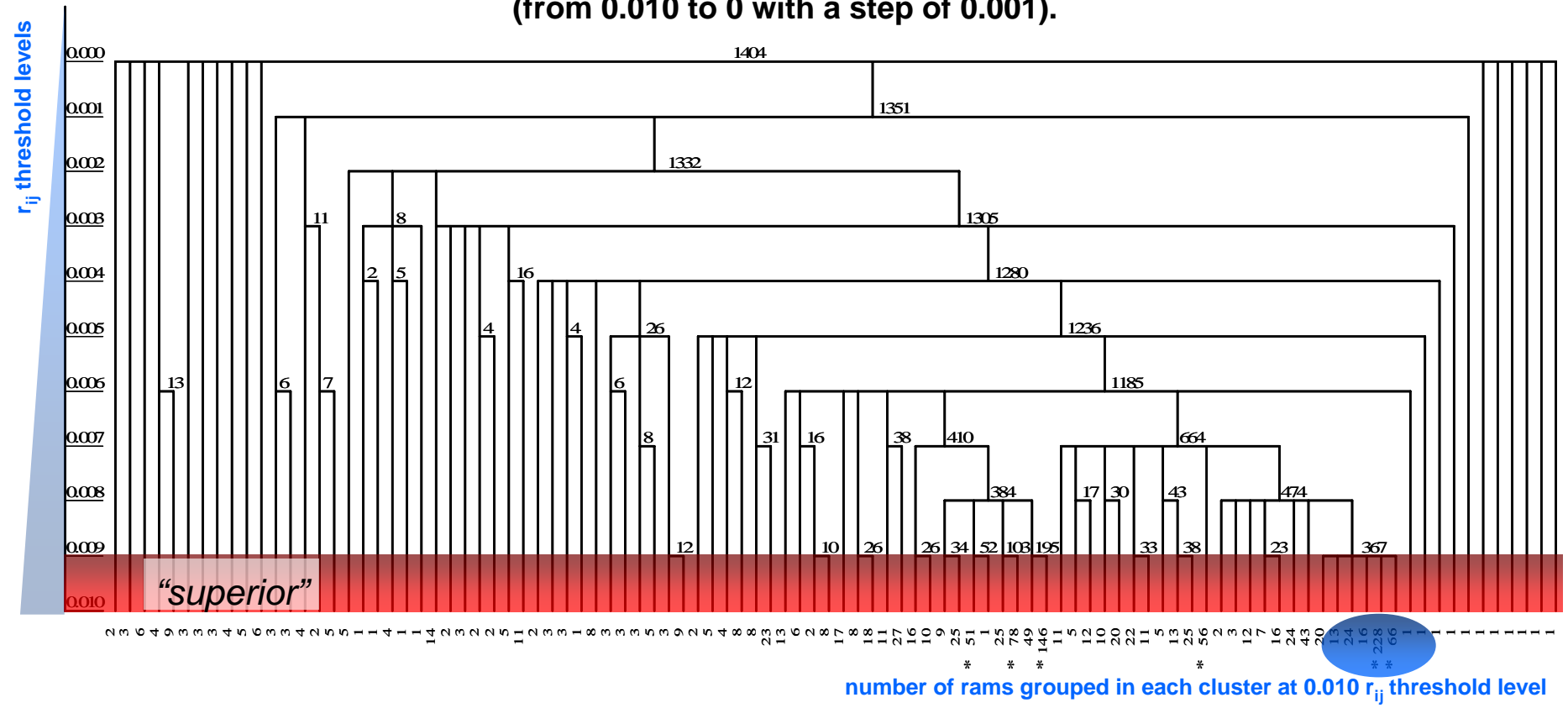
The following parameters were averaged for relevant clusters (more than 50 sires):

- ❑ connectedness correlation ( $r_{ij}$ ) and the coefficient of determination (**CD**) of the contrasts
- ❑ **NF**: the total number of daughters per sire
- ❑ **NFW**: the total number of effective daughters per sire
- ❑ **NCG**: the total number of CG in which a sire had daughters
- ❑ **NP**: the total number of directly compared sires per sire
- ❑ **IC**: the inbreeding coefficient
- ❑ **PW**: the average relationship of a sire within CG calculated as the average relationship coefficient with the other sires in each CG weighted for the effective number of daughters.

**NF**, **NFW**, **NCG** and **NP** **affect** the accuracy of EBV (Tosh and Wilton, 1994)  
increasing **PW** and **IC** **reduce** CD of contrast (Kennedy and Trus, 1993)

# 3.0 Results and Discussion

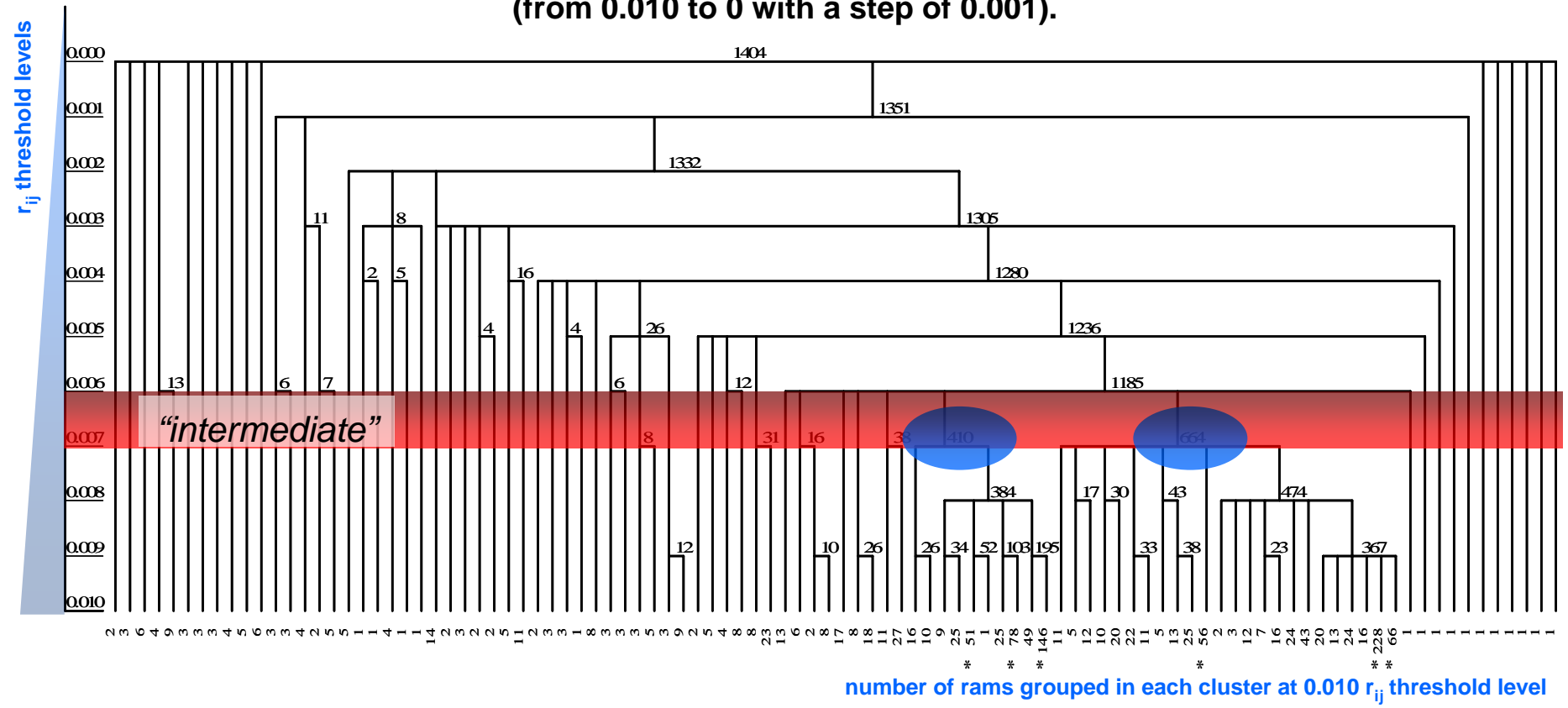
Diagram of clustering procedure results at successively decreasing connectedness correlation ( $r_{ij}$ ) thresholds (from 0.010 to 0 with a step of 0.001).



- 83 different subpopulations including on average  $17 \pm 31$  sires ranging from 2 to 228 sires were detected at the 0.010  $r_{ij}$  threshold
- The percentage of sires included in the largest cluster was 16.2% (228/1,404)

# 3.0 Results and Discussion

Diagram of clustering procedure results at successively decreasing connectedness correlation ( $r_{ij}$ ) thresholds (from 0.010 to 0 with a step of 0.001).



Two large separate subpopulations of rams were identified. The first included AI sires whereas the second included many rams born in few flocks, located in the central-west part of Sardinia, which represents the historical region where selection of the Sarda breed started. Breeders of these regions usually raise own homebred rams and are the main rams sellers to the registered flocks.

The percentage of sires included in the largest cluster was 47.3% (664/1,404)



### 3.0 Results and Discussion

Number of sires (**NS**) and AI sires (**NAIS**) grouped in clusters with more than 50 rams at 0.010, 0.007 and 0.005  $r_{ij}$  threshold levels (**TL**) and means<sup>1</sup> per cluster of: **prediction error variance of EBV (PEV)**, **number of daughters (NF)** and **number of effective daughters (NFW)** per sire, **number of management units in which the sire has daughters (NCG)**, **number of other sires directly compared to the considered sire (NP)**, **average additive relationship coefficient within CG (PW)**, **inbreeding coefficient (IC)** per sire.

TL	NS	NAIS	$r_{ij}$	CD	PEV	NF	NFW	NCG	NP	PW	IC
0.010	51	0	0.028	0.540	40.3	42	17	3	9	0.063	0.158
	56	1	0.040	0.502	39.7	33	18	3	9	0.060	0.071
	66	5	0.021	0.564	35.3	38	22	5	24	0.046	0.064
	78	0	0.024	0.500	41.9	35	15	2	5	0.058	0.162
	146	1	0.023	0.518	39.3	38	20	2	9	0.106	0.180
	228	74	0.016	0.586	33.2	42	25	11	64	0.036	0.044
0.007	410	1	0.014	0.519	41.1	39	18	2	7	0.076	0.166
	664	119	0.012	0.543	37.4	37	21	7	39	0.038	0.052
0.005	1236	128	0.010	0.524	39.3	37	19	5	24	0.052	0.089

*Significant higher value of NFW, NCG, and NP and smaller values of PW and IC means were found in clusters with high rate of AI sires.*

*These differences are in the direction of a better accuracy of EBV as confirmed by a significant lower PEV mean.*

## 4.0 Conclusion

- Accepting 0.005  $r_{ij}$  as good level of connectedness, 90% of sires of the last generation belong to the same genetic population and their EBV can be compared.
- The use of the other sires should be limited for selection within flock only.
- Changes of Flock-book rules are in progress to take into account these results.

This study showed that the current breeding management of the Sarda dairy sheep breed allows to reach a sufficient level of accuracy of the genetic ranking of rams involved in the breeding program, including CNM rams which still play a crucial role in the selection program.