

ANTONIOS P. KOMINAKIS

Effect of unfavourable population structure on estimates of heritability, systematic effects and breeding values

In memory of the beloved colleague Alexandros Argyrokastritis

Abstract

Empirical estimations of heritability, systematic effects and predictions of sires' breeding values (BVs) were obtained under various population structures for simulated populations consisted of $n=400$ animals in 5 herds for a trait of medium heritability ($h^2=0.30$). An infinitesimal additive genetic animal model was assumed while simulating data. Population structure was varied to allow for good and poor connectedness across herds and (non)random association between the genetic and the environmental effects. The impact of the various population structures on the parameter estimation(s) was assessed using Mean Squared Error (MSE) and Pearson's correlations. Allowing sires to have progenies in more than one herd (good herd connectedness) and random use of sires across herds generally resulted in good parameter estimations. Poor connectedness significantly affected herd effects estimation and BV prediction but not heritability estimation as long as random usage of sires across environments was guaranteed. Selective use of the best sires in the best herds along with poor connectedness resulted in poorest estimations of all parameters examined. In the latter case, heritability was seriously underestimated ($h^2=0.06$) while highest error, lowest accuracies for the BVs and a remarkable underestimation of the genetic gain were observed. Use of reference sires on a natural mating basis to create genetic links between herds has served a good solution for both heritability and BVs estimation under unfavorable structure. Mating 0.25 of the herd ewes with reference sires resulted in a heritability estimate close to the simulated one. Significantly better estimates of systematic effects and BVs were, however, obtained when 0.5 of the herd ewes were mated by reference sires.

Keywords: heritability, estimation, fixed effects, breeding values, connectedness, population structure

Zusammenfassung

Titel der Arbeit: Auswirkungen ungünstiger Populationsstrukturen auf Heritabilitäts-, systematische Effekte- und Zuchtwertschätzungen

Es wurden bei unterschiedlichen simulierten Populationsstrukturen mit kleinen Populationen, bestehend aus 400 Tieren in fünf Herden, empirische Heritabilitäts-, fixe Effekte- und Zuchtwerte für ein Merkmal mittlerer Heritabilität ($h^2=0,30$) geschätzt. Die Simulationen erfolgten mit verschiedenen Populationsstrukturen sowohl mit ausreichender, das heißt Väter mit Nachkommen in mehreren Herden, als auch geringer Herdenverbundenheit sowie bei Abhängigkeit zwischen genetischen und umweltbedingten Effekten. Der Einfluss der verschiedenen Populationsstrukturen auf die Parameterschätzungen wurde mittels der mittleren quadratischen Abweichung und der Pearson Korrelation quantifiziert. Zuverlässigere Parameterschätzungen ergaben sich wenn die Väter in mehr als einer Herde Nachkommen aufwiesen oder wenn eine zufällige Väterverteilung vorlag. Bei zufälligen Zuchtväternutzung oder geringerer Herdenverbundenheit ergaben sich keine Auswirkungen auf die Heritabilitätsschätzungen wohl aber Beeinträchtigungen bei der Schätzung systematischer Effekte und Zuchtwerte. Für die schlechtesten Schätzwerte sorgte eine selektive Nutzung der besten Väter in den besten Herden in Verbindung mit ungünstiger Populationsstruktur, d.h. geringer Herdenverbundenheit. In diesem Fall wurde, in Verbindung mit höchstem Schätzfehler, geringster Genauigkeit der Zuchtwerte sowie geringstem Zuchterfolg, die Heritabilität erheblich unterschätzt ($h^2=0,06$). Bei ungünstigen Populationsstrukturen führte die Nutzung von Referenzzuchtvätern im Natursprung zu akzeptableren Parameterwerten sowohl bei der Schätzung der Heritabilitäts- als auch der Zuchtwerte. Eine Anpaarung von 25% der Schafmüttern an Referenzväter ergab Heritabilitätsschätzwerte, die nahe den simulierten Werten lagen. Noch bessere Schätzungen der fixen Effekte und Zuchtwerte sind erreichbar, wenn dieser Anteil von Müttern, die mit Referenzv Vätern angepaart werden, bei 50% liegt.

Schlüsselwörter: Heritabilität, Schätzung, Fixe Effekte, Zuchtwerte, Verbundenheit, Populationsstruktur

Introduction

Sheep and goat as well as beef cattle herds are traditionally, small, closed populations of limited number of animals. In such populations, artificial insemination (AI) is practically non-existent, increasing their regional isolation and the potential of genetic drift. Moreover, animals' performance is recorded across highly diverse environments with poor data quality due to recording errors, misclassification or manipulation of contemporary groups (e.g. herds) and non-random use, selection and mating of breeding animals (SCHAEFFER et al., 1998). Such data are less favourable from an estimation point of view, because of unbalancedness, low connectedness and dependence of the genetic and the environmental structure. Although the effect of data structure in the context of genetic parameters estimation (CLEMENT et al., 2001) and genetic evaluation (HANOCQ et al., 1996; TONG et al., 1980) has been studied, the supremacy of REML and BLUP in analysis of data with unfavourable data structure in small populations still needs to be verified.

Aim of the present study was thus the estimation of heritability, systematic effects and breeding values (BVs) under unfavourable population structures by means of stochastic simulations. Population structure was varied to allow for good and poor connectedness and/or (non)random association between the random and the systematic effects. The effect of creating genetic links between herds through usage of reference sires was also investigated and quantified.

Materials and methods

Data simulation

Simulated data were generated using a stochastic procedure that is commonly used (e.g. HARDER et al., 2005). The genetic model assumed a large number of unlinked loci contributing to the genetic variance of a single hypothetical metric trait. The base population consisted of 20 males and 400 females, which were assumed to be unrelated, unselected and randomly sampled from a conceptually infinite population. The base animals were mated at random (20 females per male) to produce 400 progeny. Each progeny record was simulated including the animals' breeding value, a population mean, a herd effect and the residual term. The model for simulation data was thus as follows:

$$y_{ij} = \mu + h_i + a_{ij} + e_{ij}$$

where y_{ij} =the phenotypic observation of animal j in herd i , μ =the population mean ($\mu=100$), h_i =the herd effect i ($i=1, \dots, 5$), a_{ij} =the additive breeding value of animal j ($i=1, \dots, 400$) in herd i , and e_{ij} =the random residual term.

Values for e_{ij} were independently drawn from a normal distribution with mean zero and variance σ_e^2 . The additive genetic variance for the base population was σ_a^2 . Breeding values of base animals were independently drawn from a $N(0, \sigma_a^2)$. Breeding values in progeny were simulated as $a_{ij} = 0.5(a_{sj} + a_{dj}) + m_{ij}$, where a_{sj} and a_{dj} are the breeding values of the sire (s) and dam (d) of individual j , and m_{ij} is the Mendelian sampling term of individual j . The Mendelian sampling term was drawn from a $N(0, 0.5\sigma_a^2)$ and it was assumed to be independent of the breeding values of the sire and dam. Herd effects were drawn from a uniform distribution ranging from -30 to $+30$, assuming five levels: $+30$, $+20$, 0 , -20 and -30 , respectively. Variances, i.e. σ_a^2

and σ_e^2 were set to 30 and 70, respectively, to reflect a trait of medium heritability ($h^2=0.30$) such as many production traits (e.g. milk yield) in dairy sheep and goats. One hundred repeats of simulating records were created to describe the randomness in the system. Each repeat included 400 animals with phenotypic records plus 420 base population animals without records. Herd size and structure(s) were chosen to reflect a structure commonly observed in small sheep and goat (and in some beef cattle) populations where natural service predominates.

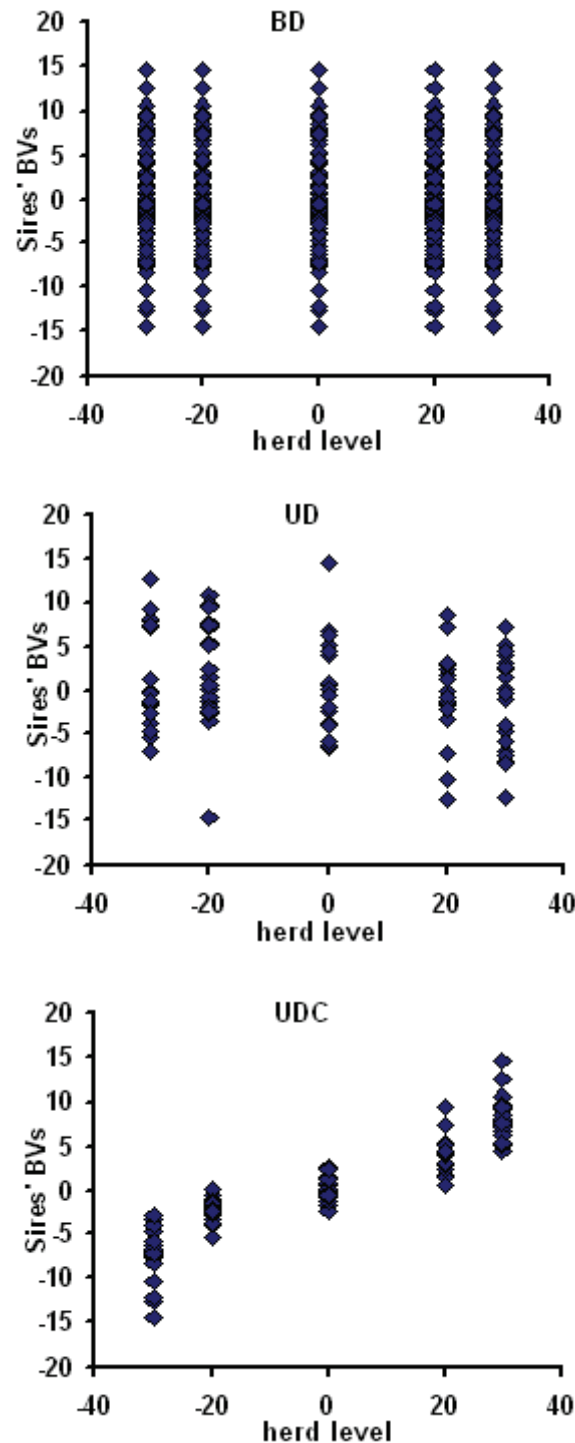


Fig. 1: Scatter plot of sires' simulated BVs across herd levels under three population structures (data from 5 repeats): balanced design (BD), unbalanced design (UD), unbalanced design with genetic-environmental dependence (UDC) (Streubild der simulierten Zuchtwerte von Vätern beider Herdenklassen mit drei Populationsstrukturen (Daten von 5 Wiederholungen): balanciertes Design [BD], unbalanciertes Design [UD], unbalanciertes Design mit genetisch-umweltbedingter Abhängigkeit [UDC])

Population structure(s)

Initially, there were three population structures simulated. In all these structures, the number of daughters per sire was 20 and the number of animals per herd was 80. Across the three structures, daughters per sire were either randomly or non-randomly assigned to herd levels. In the first structure, sires were assigned randomly across the five herd levels, each sire having equal ($n=4$) number of daughters in each of the five herd levels. Random use of sires and herd connectedness was thus guaranteed in this case (Figure). This structure will be referred as the balanced design (BD). In the second structure, each sire had $n=20$ of daughters in only one herd. In this case, sires were randomly assigned across herd levels. This structure did not allow for connectedness between herds but random distribution of the sire's BVs across herds was guaranteed (Figure 1). This structure will be referred as the unbalanced design (UD). In the third structure, sires had daughters only in one herd while sires with highest/lowest breeding values were used in herds with highest/lowest levels, respectively. This structure (UDC) will be referred as the unbalanced design with genetic environmental dependence (Figure). Three more population scenarios under unfavourable structure were simulated by assuming usage of reference sires (RS) to create genetic links between herds. RS usage was assumed on a natural service since AI is costly and difficult in sheep. The three scenarios were based on the number of RS used per herd: 1, 5 and 10. Reference sires were assumed to be selected at random from top animals based on true BVs.

Measures of connectedness and genetic-environmental association

Connectedness is variously defined in the literature. Sometimes, it is the existence of genetic ties between levels of fixed effects (TOSH and WILTON, 1994) and other times it is defined as the estimability of contrasts between levels of genetic effects (KENNEDY and TRUS, 1993; FOULLEY et al, 1992). All measures belonging to the second category along with others proposed (LALOE, 1993; LEWIS et al. 1999, MATHUR et al., 2002) are difficult to calculate and follow. For this reason, description of connectedness was attempted here using measures proposed by TOSH and WILTON (1994): effective number of progeny per sire (n_{e_j}), number of herds in which a sire has progeny (M) and direct connections, DC =number of sires with progeny in the same herd. Let n_{ij} be the number of progeny of the j^{th} sire in the i^{th} herd, then the effective number of progeny (n_{e_j}) of the j^{th} sire is:

$$n_{e_j} = n_{.j} - \sum (n_{ij}^2 / n_{i.})$$

Dependence between sires' BVs (genetic) and herd levels was quantified using three non-parametric statistical measures: the Kendall's tau-b (τ), the Hoeffding's measure of dependence (D) and the Spearman correlation (ρ). The Kendall's τ is a measure of association based on the number of concordances and discordances in paired observations. Hoeffding's measure of dependence (D) is a measure of association that detects more general departures from independence. All above association statistics were estimated by procedure CORR in SAS (2002).

Analyses

The operational model was defined to be the same as the true model used for simulation of the data sets in all population structures. The univariate linear mixed model (animal model) used to analyze the simulated data was:

$$y = 1\mu + Xb + Za + e$$

The distributional assumptions were: $a \sim N(0, A\sigma_a^2)$ and $e \sim N(0, I\sigma_e^2)$, where a = the vector of additive breeding values and e = the vector of random residual effects, and A = the numerator relationship matrix that included the base population animals. REML estimates of variance components, BLUE of fixed effects and BLUP of BVs per repeat for all the analyses were obtained using an average information algorithm and the ASREML computer program (GILMOUR et al., 1999). Because emphasis was put on sires' use across the herd levels, prediction of BVs was focused on only sires ($n=20$).

Comparison criteria

Estimated heritability coefficients and fixed effects as well as predicted BVs were compared with respect to their simulated (true) values using Mean Squared Error (MSE) and Pearson's correlations. MSE is the average squared deviation of predicted values from their corresponding true values and represents the mean distance between the simulated and true values. The lower the value of MSE the closer the estimated values are to the simulated (true) ones:

$$MSE_p = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

where n = the number of sires or the number of herd levels, and $\hat{}$ refers to the predicted or estimated value of the parameter p . In the case of BVs, the Pearson's correlation coefficient is by definition a measure of accuracy of prediction. Means comparisons of estimated or predicted parameters were performed between the three initial structures i.e. BD, UD and UDC as well as between the UDC case and the three sire reference schemes. Comparisons were carried out by application of the Ryan-Einot-Gabriel-Welsch Multiple Range multiple range test in SAS (2002).

Results

Connectedness and genetic-environmental association

Measures of herd connectedness and genetic-environmental association under the various population structures are in Table 1. Effective number of progeny per sire (n_{e_j}) was maximum in the balanced ($n_{e_j} = 19$) and lowest in the unbalanced design ($n_{e_j} = 15$). This parameter gradually improved as the number of RSs per herd increased. The same trend was also observed for the rest two connectedness measures, i.e. number of herds in which a sire has progeny (M) and direct connections (DC). Genetic-environmental dependence for the UDC design was confirmed by high values for Kendall's τ , Hoeffding' D and Spearman correlation. However, increasing the number of RSs per herd resulted in diminishing levels of association (Table 1). Note that non integer values for the three connectedness measures in the case of using one reference sire per herd implies an unbalanced design: apart from one sire used in all herds, two more sires are necessarily used in two herds each having 16.4 and 12.8 progeny, respectively.

Table 1

Measures of herd connectedness and of genetic-environmental association across the various population structures: balanced design (BD), unbalanced design (UD), unbalanced design with genetic-environmental dependence (UDC) and sire reference schemes (Maßzahl der Herdenverbundenheit und der genetisch-umweltbedingten Assoziation bei verschiedenen Populationsstrukturen: balanciertes Design [BD], unbalanciertes Design [UD], unbalanciertes Design mit genetisch-umweltbedingter Abhängigkeit [UDC] und den Väter-Referenzschemata)

	BD	UD	UDC	No. of reference sires per herd		
				1	5	10
Effective no. of progeny per sire (n_{e_j})	19	15	15	14.85	16	17
No. of herds (M)	5	1	1	1.4	2	3
No. of direct connections (DC)	20	4	4	5.6	8	12
Kendall's τ	0	0.02	0.92	0.82	0.52	0.23
Hoeffding's D	0	0.03	0.58	0.49	0.24	0.07
Spearman (ρ)	0	0.01	0.98	0.88	0.55	0.24
Proportion of ewes in herd mated with reference sires				0.05	0.25	0.50

Heritability estimation

Simulated as well as estimated heritability estimates under the various population structures are in Table 2. The average simulated heritability was 0.297, very close to the desired value ($h^2=0.30$). Mean estimated heritability was comparable across cases BD and UD but it was remarkably lower in the UDC design ($h^2\sim 0.06$). This result was due to a significant underestimation of the additive genetic variance (results not shown). Using one RS under the most unfavorable population structure, improved heritability estimation ($h^2=0.15$). Using 5 or 10 RSs resulted in a heritability estimate close to the simulated value. A total number of 2, 37, 18 and 2 (out of 100) of heritability estimates were found equal (or restricted because of REML) to zero in the UD, UDC, 1RS and 5RS cases, respectively (Table 2).

Table 2

Descriptive statistics and mean squared error (MSE) of simulated (S) and estimated heritability under the various population structures: balanced design (BD), unbalanced design (UD), unbalanced design with genetic-environmental dependence (UDC) and sire reference schemes (Beschreibende Statistik und mittleres Fehlerquadrat [MSE] der simulierten und der geschätzten Heritabilität unter verschiedenen Populationsstrukturen: balanciertes Design [BD], unbalanciertes Design [UD], unbalanciertes Design mit genetisch-umweltbedingter Abhängigkeit [UDC] und den Väter-Referenzschemata)

		Mean (CI, 95 %)	Min	Max	MSE $\times 100$ (CI, 95 %)
Population structure	S	0.297 ^a (0.292-0.303)	0.258	0.354	
	BD	0.285 ^a (0.257-0.314)	0.048	0.665	1.87 ^a (1.41-2.32)
	UD	0.280 ^a (0.252-0.308)	0 (n=2)	0.604	1.84 ^a (1.33-2.35)
	UDC	0.059 ^{b1} (0.046-0.073)	0 (n=37)	0.219	6.18 ^{b1} (5.60-6.77)
Number of reference sires per herd	1	0.138 ² (0.118-0.159)	0 (n=18)	0.369	2.61 ² (2.09-3.14)
	5	0.308 ³ (0.277-0.339)	0 (n=2)	0.689	1.94 ² (1.36-2.53)
	10	0.292 ³ (0.262-0.323)	0.009	0.633	2.04 ² (1.55-2.53)

Means with different letters or numbers are statistically significant different ($p<0.05$). Letters are pertaining to comparisons among the S, BD, UD and UDC cases, while numbers are pertaining to comparisons among UDC and the sire reference schemes.

CI: confidence intervals

Estimation of fixed effects and prediction of BVs

Table 3 shows MSE for the fixed effects and BVs, Pearson's correlations for the BVs as well as means of the estimated BVs of the 25 % best sires as a measure of genetic gain (GG) across the various scenarios. No correlation coefficients are presented for fixed effects because of limited number of observations (classes). As reasonably expected, MSE for fixed effects and BVs were lowest in the balanced design (BD), were statistically significant higher in the unbalanced case (UD) and highest in UDC. Prediction accuracy of BVs was highest in the balanced design ($r=0.76$) close to the theoretically expected value estimated by the classical formula (FALCONER, 1989): $r = \sqrt{(n/n+k)}$ with $k = (4-h^2)/h^2$. For $n=20$ and $h^2=0.285$ this formula gives $r=0.778$. Prediction accuracy of BVs decreased in the unbalanced design ($r=0.68$) and it was remarkably low ($r=0.12$) in the extremely unfavourable structure (UDC). MSE for fixed effects and BVs did not discriminate between 1 and 5 RS schemes and it was significantly lower when using 10 RSs. Using more than one RS gradually improved accuracy of prediction of the BVs. Genetic gain was similar across cases BD, UD and it was significantly lowest in the UDC case. Using 5 or 10 RSs per herd resulted in genetic gains similar to those estimated in the fully balanced design (Table 3).

Table 3

Mean Squared Error (MSE) for fixed effects (f) and breeding values (BV), accuracy of BVs estimation (r) as well as genetic gain (GG) under the various population structures: balanced design (BD), unbalanced design (UD), unbalanced design with genetic-environmental dependence (UDC) and sire reference schemes (Mittleres Fehlerquadrat [MSE] für fixe Effekte [f] und für Zuchtwerte [BV], Genauigkeit der Zuchtwertschätzung [r] und Zuchtfortschritt [GG]) unter verschiedenen Populationsstrukturen: balanciertes Design [BD], unbalanciertes Design [UD], unbalanciertes Design mit genetisch-umweltbedingter Abhängigkeit [UDC] und den Väter-Referenzschemata)

		MSE _f (95 % CI)	MSE _{BV} (95 % CI)	r (95 % CI)	GG (95 % CI)
Population structure	BD	1.69 ^a (1.51-1.88)	15.09 ^a (13.89-16.30)	0.76 ^a (0.73-0.78)	5.11 ^a (4.70-5.53)
	UD	3.18 ^b (2.70-3.65)	19.00 ^b (17.54-20.35)	0.68 ^b (0.65-0.71)	4.43 ^a (4.07-4.80)
	UDC	7.64 ^{c1} (6.95-8.32)	33.28 ^{c1} (31.28-35.28)	0.12 ^{c1} (0.10-0.15)	1.81 ^{b1} (1.55-2.08)
Number of reference sires per herd	1	8.63 ¹ (7.92-9.54)	35.02 ¹ (32.72-37.32)	0.07 ¹ (0.4-0.10)	3.08 ² (2.76-3.40)
	5	9.91 ¹ (9.06-10.78)	32.95 ¹ (30.70-35.19)	0.25 ² (0.22-0.27)	5.20 ³ (4.82-5.60)
	10	5.02 ² (4.58-5.46)	20.42 ² (19.11-21.72)	0.61 ³ (0.59-0.65)	5.07 ³ (4.63-5.50)

Means with different letters or numbers are statistically significant different ($p<0.05$). Letters are pertaining to comparisons among the BD, UD and UDC cases, while numbers are pertaining to comparisons among UDC and the sire reference schemes. CI: confidence intervals

Discussion

This study has confirmed supremacy of the REML method even when herds are disconnected. Herd connectedness did not affect the accuracy of heritability estimation (compare cases BD and UD, Table 2). However, this is not the case for BLUP, where disconnectedness significantly affected estimation of BVs and particularly of the systematic effects. In the latter case, MSE in the unbalanced design was twofold when compared to the balanced design while error and accuracy for BVs were 0.20 higher and 0.10 lower, respectively. It was, however, the selective

use of sires in combination with disconnectedness that seriously affected the estimation of all parameters examined. Under this extreme scenario, a remarkably low heritability, highest estimation error for fixed effects and BVs, lowest accuracy for the BVs and lowest genetic gain were observed. Apparently, in absence of connectedness there is no way to estimate genetic and/or environmental effects when these effects are confounded. CLEMENT et al. (2001) have also reported underestimated heritabilities for additive and maternal genetic effects under disconnectedness when herds differ in genetic merit. Only the scenario of assortative assignment of sires across environments was simulated in the present study, which may reflect the results of selection and/or culling bias as well. Selective use of sires should be thus avoided and random distribution of sires across environments should be pursued in any breeding program. If not so, the use of reference sires serves as a good solution. Apart from creating genetic ties between herds, the usage of reference sires removes any association between the genetic and the environmental effects ultimately allowing for better parameter estimations. Mating 0.25 of the herd ewes with reference sires resulted in a fairly good approximate of the true heritability value. However, when the estimation of fixed effects and of BVs is coming in question, the proportion of herd ewes mated with RSs should increase up to 0.5. Even in the latter case, the level of accuracy of BVs remained significantly lower than that of the fully balanced design. By simulating other scenarios and population structures, KUHN (2005) also reported reduced bias of comparisons between homebred sires when more progeny were allocated to reference sires. Reduction in bias was markedly lower with 20% linking sire progeny, but differences in bias reductions when 50 vs. 33% of progeny were from linking sires were minor. HANOCQ et al. (1996) also reported a non linear increase for gains in individual accuracies of BVs estimation under a small level of connection in contrast to higher levels of connections.

When compared to the balanced design, genetic gain in the unbalanced design was lower, however, not statistically significant. Interestingly, mating 0.25 of herd ewes with reference sires under unfavourable population structure resulted in similar genetic gain to that of the balanced design. Other simulation studies by assuming AI sire referencing have shown improved genetic gain by 30 to 35% compared to within-unit selection as well as improved accuracy and lower inbreeding coefficients (HANOCQ et al., 1996; LEWIS and SIMM, 2000; RODEN, 1996). Although referencing is usually through AI, applications of such schemes on a natural mating basis have shown comparable to AI-SR genetic gains while being superior in terms of inbreeding (KUHN, 2005). AI may create extensive connections, but AI is currently not widely used in sheep, goat and beef cattle populations because it is either costly, or difficult to apply.

Methodological approach

A problem often encountered when analysing data with unfavourable structure is how to treat contemporary groups (CG). HENDERSON (1973) showed that by treating CGs as fixed effects renders genetic evaluations invariant to CG effects. Other studies (UGARTE et al., 1992; VISSCHER and GODDARD, 1993) suggested treating CGs as random effects under non-random distribution of sires across CGs. Accuracy of BV estimation was improved in the random model in all cases examined by UGARTE et al. (1992). VISSCHER and GODDARD (1993) also found higher

accuracies for the BVs when the best sires were used in the best CGs using a random model. However, when the best sires were represented only in the worst CGs, the correlation between true and predicted breeding values for the random model became negative. The authors finally concluded that when “a non-random association exists between sires and CGs, the groups should be treated as fixed effects for practical genetic evaluations”. In dairy cattle, Interbull (2001) recommends considering an effect as fixed if the association between the effect and the main random effect (animal or sire) is non-random. Practically, an effect with limited number of levels cannot be fitted as random and this is the main reason why herd effect was fitted here as fixed effect. In the present study the effect of a single use of reference sires within of one generation was evaluated. Connections by time are reasonably expected to result in more accurate genetic evaluations and higher genetic responses when herds have different genetic levels. Connections allow for gene flow from herds having high genetic level to herds with low genetic level. As a result, differences in the average genetic level gradually decrease across herds (HANOCQ et al., 1996). Further studies are needed to examine the effect on parameter estimation and genetic evaluation of (dis)continual connectedness by time.

Acknowledgements

Professor Dr. H. Simianer of the Department of Animal Breeding and Genetics in the Goettingen University (Germany) is gratefully acknowledged for providing valuable suggestions on a previous version of the manuscript.

References

- CLÉMENT, V.; BIBÉ, B.; VERRIER, E.; ELSEN, J.M.; MANFREDI, E.; BOUIX, J.; HANOCQ, E.:
Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. *Genet. Sel. Evol.* **33** (2001), 369-395
- FALCONER, D.S.:
Introduction to Quantitative Genetics. 3rd Ed., London (1989)
- FOULLEY, J.L.; HANOCQ, E.; BOICHARD, D.:
A criterion for measuring the degree of connectedness in linear models of genetic evaluation. *Genet. Sel. Evol.* **24** (1992), 315-330
- GILMOUR A.R.; CULLIS B.R.; WELHAM S.J.; THOMPSON R.:
ASREML Reference Manual. NSW Agriculture, Orange, Australia (1999)
- HARDER, B.; BENNEWITZ, J.; REILISCH, N.; MAYER, M.; KALM, E.:
Effect of missing sire information on genetic evaluation. *Arch. Tierz.* **48** (2005), 219-232
- HANOCQ, E.; BOICHARD, D.; FOULLEY J.L.:
A simulation study of the effect of connectedness on genetic trend. *Genet. Sel. Evol.* **28** (1996), 67-82
- HENDERSON, C.R.:
Sire evaluation and genetic trends. In: *Proc. Anim. Breed. Genet. Symp*, in Honor of Dr. J. L. Lush, ASAS and ADSA, Champaign, Illinois, (1973), 10-41
- INTERBULL:
Guidelines for national & international genetic evaluation systems in dairy cattle with focus on production traits. Bulletin no. **28** (2001)
- KENNEDY, B.W.; TRUS, D.:
Considerations on genetic connectedness between management units under an animal model. *J. Anim. Sci.* **71** (1993), 2341-2352
- KUHN, L.R.:
Implications of Connectedness in the Genetic Evaluation of Livestock. PhD Dissertation, Virginia Polytechnic Institute and State University, Blackbourg, Virginia (2005)
- LALOE, D.:
Precision and information in linear models of genetic evaluation. *Genet. Sel. Evol.* **25** (1993), 557-576

- LEWIS, R.M.; CRUMP, R.E.; SIMM, G., THOMPSON, R.:
Assessing connectedness in across-flock genetic evaluations, in: Proc. Brit. Soc. Anim. Sci., 22-24 March 1999, Scarborough, 121
- LEWIS, R.M.; SIMM, G.:
Selection strategies in sire referencing schemes in sheep. Livest. Prod. Sci. **67** (2000), 129-141
- MATHUR, P.K.; SULLIVAN, B.P.; CHESNAIS J.P.:
Measuring connectedness: concept and application to a large industry breeding program. In: Proc. 7th World Congr. Genet. Appl. to Livest. Prod., 19.-23. August 2002, Montpellier, Communication No. 20-13
- RODEN J.A.:
A comparison of alternative nucleus breeding systems and a sire referencing scheme for sheep improvement. Anim. Sci. **62** (1996), 265-270
- SAS (2002). SAS Institute Inc., Cary, NC, USA.
- SCHAEFFER L.R.; SCHENKELLY F.S.; FRIES, L.A.:
Selection bias on animal model evaluation. In: Proceedings of the 6th World Congress of Genetics Applied to Livestock Production., Armidale, Australia **25** (1998), 501-508
- TONG, A.K.W.; KENNEDY, B.W.; MOXLEY J.E.:
Potential errors in sire evaluation from regional genetic differences. J. Dairy Sci. **63** (1980), 627-633
- TOSH, J.J.; WILTON, J.W.:
Effects of data structure on variance of prediction error and accuracy of genetic evaluation. J. Anim. Sci. **72** (1994), 2568-2577
- UGARTE, E.; ALENDA, R.; CARABANO, M. J.:
Fixed or random contemporary groups in genetic evaluations. J. Dairy Sci. **75** (1992), 269-278
- VISSCHER, P.M.; GODDARD, M.E.:
Fixed and random contemporary groups. J. Dairy Sci. **76** (1993), 1444-1454

Received: 2007-12-09

Accepted: 2008-08-27

Author:

Dr. ANTONIOS P. KOMINAKIS
Department of Animal Breeding & Husbandry
Faculty of Animal Science
Agricultural University of Athens
Iera Odos 75, 11855
Athens
Greece

email: acom@aua.gr