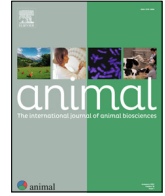




Animal

The international journal of animal biosciences



A comparison of genetic and genomic breeding values in Saanen and Alpine goats



A. Negro^{a,d}, A. Cesarani^{b,c}, M. Cortellari^d, A. Bionda^{d,*}, P. Fresi^a, N.P.P. Macciotta^b, S. Grande^a, S. Biffani^e, P. Crepaldi^d

^a Ufficio Studi, Associazione Nazionale della Pastorizia, 00187 Rome, Italy

^b Dipartimento di Scienze Agrarie, Università degli Studi di Sassari, 07100 Sassari, Italy

^c Department of Animal and Dairy Science, University of Georgia, Athens, GA, USA

^d Dipartimento di Scienze Agrarie e alimentari, Università degli studi di Milano, 20133 Milan, Italy

^e Istituto di Biologia e Biotecnologia, Consiglio Nazionale delle Ricerche, 20133 Milan, Italy

ARTICLE INFO

Article history:

Received 8 September 2023

Revised 17 February 2024

Accepted 20 February 2024

Available online 27 February 2024

Keywords:

Accuracy

Estimated breeding values

Genomic selection

Goat

Single-step genomic best linear unbiased prediction

ABSTRACT

Nowadays, several countries are developing or adopting genomic selection in the dairy goat sector. The most used method to estimate breeding values is Single-Step Genomic Best Linear Unbiased Prediction (**ssGBLUP**) which offers several advantages in terms of computational process and accuracy of the estimated breeding values (**EBVs**). Saanen and Alpine are the predominant dairy goat breeds in Italy, and both have similar breeding programs where EBVs for productive traits are currently calculated using BLUP. This work describes the implementation of genomic selection for these two breeds in Italy, aligning with the selection practices already carried out in the international landscape. The available dataset included 3 611 genotyped animals, 11 470 lactation records, five traits (milk, protein and fat yields, and fat and protein percentages), and three-generation pedigrees. EBVs were estimated using BLUP, GBLUP, and ssGBLUP both with single and multiple trait approaches. The methods were compared in terms of correlation between EBVs and genetic trends. Results were also validated with the linear regression method excluding part of the phenotypic data. In both breeds, EBVs and GEBVs were strongly correlated and the trend of each trait was similar comparing the three methods. The average increase in accuracy across traits and methods amounted to +13 and +10% from BLUP to ssGBLUP for Alpine and Saanen breeds, respectively. Results indicated higher prediction accuracy and correlation for GBLUP and ssGBLUP compared to BLUP, implying that the use of genotypes increases the accuracy of EBVs, particularly in the absence of phenotypic data. Therefore, ssGBLUP is likely to be the most effective method to enhance genetic gain in Italian Saanen and Alpine goats.

© 2024 The Author(s). Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Implications

Selection for productivity in Italian Alpine and Saanen goats currently relies on genetic evaluations. More recently, an extensive genotyping of these populations has started with the aim of introducing genomic selection and developing new breeding values. In this context, our work investigates the potential for enhancing the existing selection scheme in these breeds by introducing single-step genomic best linear unbiased predictions. Implementing genomic selection would bring significant improvements in the economic, environmental, and ethical sustainability of Italian goat farming. Moreover, the results presented here will add

valuable data to the broader international understanding of goat selection practices.

Introduction

In the past few years, there has been a notable increase in the worldwide goat population. Its significance as a valuable resource has grown, driven by shifts in human incomes and food preferences, as well as the impact of climate change (Miller and Lu, 2019). Currently, public opinion and climate change led to the need to develop new breeding systems and methods that can improve goat dairy production while also ensuring animal welfare, resilience, efficiency, and health (Biffani et al., 2020; Mucha et al., 2022).

The introduction of the genomic selection theorised by Meuwissen et al. (2001) is nowadays recognised as one of the most

* Corresponding author.

E-mail address: arianna.bionda@unimi.it (A. Bionda).

important tools to improve the accuracy of genetic evaluations and it is now widely practiced across the breeding industry (Misztal et al., 2020). Genomic selection is based on the prediction of the animals' genomic value through the association of single nucleotide polymorphisms (SNPs) and canonical production data in a reference population previously genotyped and phenotyped. Then, considering the additive effects of each single SNP allele on a particular trait altogether, it is possible to estimate the so-called Genomic Estimated Breeding Values (GEBVs) (Teissier et al., 2020). Despite the considerable decrease in the genotyping costs, those are still expensive especially if compared to the economic value of a dairy goat (Cortellari et al., 2022). Therefore, the genotyping of the whole Italian goat population is not currently feasible. In this context, the application of genomic selection would require the multiple-step procedures that have some disadvantages: i) biased or inaccurate predictions for genotyped animals; ii) absence of gain in accuracy for non-genotyped animals; iii) incompatibility between estimated breeding values for genotyped and non-genotyped animal (Bermann et al., 2022). To overcome these problems, the Single Step Genomic Best Linear Unbiased Predictions (ssGBLUP) has been proposed. This method allows to simultaneously consider genotyped and non-genotyped animals improving the accuracy of their EBVs. Nowadays, ssGBLUP is used by researchers, private companies, and breeders' associations in genetic evaluations of different species (Misztal et al., 2020; Bermann et al., 2022).

Genomic selection is nowadays the gold standard in the most productive dairy cattle populations, such as the Holstein and Jersey cows, but in recent years, its application has started to spread also in other species. Indeed, several countries all over the globe, including France (Carillier et al., 2014; Teissier et al., 2018; 2019; 2020; Arnal et al., 2019; 2023), Spain (Molina et al., 2018), Canada (Massender et al., 2022) and New Zealand (Scholtens et al., 2021), have started to use genomic selection in different dairy goat populations, such as Saanen and Alpine (which is referred as Camosciata delle Alpi in Italy).

In Italy, routine genetic evaluation for milk yields and morphological traits has been in place for both Saanen and Alpine since 2000. Official genetic evaluations for milk yields use 210-day lactation records, pedigree data, and BLUP method to estimate EBVs. Currently, the selection is oriented to enhance the yields (kg) of protein and fat.

In the framework of the Italian National Rural Development Plan – sub-measure 10.2, two national projects, namely “Conservation, Health and Efficiency Empowerment of Small Ruminant (CHEESR)” and “Sustainability Health Environment Economy Profitability & Genomic Organisation Animal (pheno)Typing (SHEEP&GOAT)”, have boosted the introduction of genomic selection in Italian dairy goats (<https://www.assonapa.it>). Using data collected within these two projects, Cortellari et al. (2022) described the breeding systems and evaluated the possible impact of the use of genomic data in the management of pedigrees and inbreeding in these populations, underlining their positive impact on genetic evaluations.

The objective of the present work was to develop a genomic breeding evaluation comparing three different methods BLUP (as benchmark), GBLUP, and ssGBLUP for production traits in Alpine and Saanen breeds using the Italian official data and model.

Material and methods

Genomic data

With the aim of developing a first national genomic evaluation, 2 472 Alpine and 1 139 Saanen goats born from 1995 to 2021 were

genotyped with the Illumina medium density SNP Bead Chip (IGGC 65 K v2). The quality control was performed using PLINK 1.9 software (Purcell et al., 2007) applying the following thresholds: minor allele frequency greater than 0.0001, missing genotype lower than 0.05, missing per individuals lower than 0.05, and Hardy-Weinberg Equilibrium equal 0.000001. The SNPs located on the sexual chromosome were excluded. SNPs were mapped according to the ARS 1.2 goat genome version.

Phenotypic data

The phenotypic data used in the present work were supplied by the Italian Sheep and Goat Breeders Association (Asso.Na.Pa). The analysed phenotypes were standard 210-d lactation traits: milk yield (MY, expressed in litres), protein and fat yield (PY and FY, expressed in kilograms), and protein and fat percentages (PP and FP, expressed in %). These traits, belonging to genotyped Alpine (8 881 records) and Saanen (2 589 records) goats, were collected according to the International Committee for Animal Recording rules from 2010 and 2022. The main descriptive statistics were calculated with R base language for all considered traits and breeds.

Pedigree data

Pedigree data came from the official herdbooks shared by Asso.Na.Pa. They were corrected by genomic data using seekparentf90 software (Misztal et al., 2014) with default settings. Paternity and maternity records were checked within genotyped animals, based on mendelian conflicts, and not compatible sires or dams were set to missing. Then, if missing parents were genotyped, the likeliest ones were assigned. The corrected pedigree was analysed through *optiSel* (R-package); using the function “*summary.Pedig*”, the following pedigree parameters were estimated: pedigree completeness (index of pedigree completeness, which is the harmonic mean of the pedigree completeness of the parents (MacCluer et al., 1983)), number of fully traced generations and inbreeding of the genotyped animals. The truncation of pedigree, i.e., using only two or three generations, can alleviate biases due to missing information and reduce computation time without compromising the prediction accuracy for selection candidates, in agreement with several studies (Lourenco et al., 2014; Pocrnic et al., 2016; Cesarani et al., 2021c). Thus, only three generations of relatives for each animal were retrieved for this study.

Breeding values evaluation

Phenotypes, genotypes, and pedigree data were used to develop a new genomic breeding evaluation in Alpine and Saanen breeds. The official data were analysed using the official model applied in the Italian selection schemes. The model can be expressed as:

$$y = Xb + Za + W_{pe} + e$$

where y is the vector of production observation (MY, PY, FY, PP, and FP); b is the vector of fixed effect (Herd – Year – Season); a is the vector of random animal effect, pe is the vector of the random effect of the permanent environment which was included because of the repeated measurements, and e is the vector of random residuals. X , Z and W are, respectively, the incidence matrices of fixed, animal, and permanent environment effects.

Three different methods were compared: i) Best Linear Unbiased Prediction (BLUP); ii) Genomic Best Linear Unbiased Prediction (GBLUP); and iii) Single – Step Genomic Best Linear Unbiased Prediction (ssGBLUP). For each method, single-trait (ST) and multiple-trait (MT, which simultaneously consider all the traits thus accounting for their covariance) analyses were carried

out. Table 1 reports the number of levels for each effect and method in both breeds. Across methods, the animal effect had different levels number due to the combination of animals with phenotypes, with genotypes, and with both phenotypes and genotypes.

The analyses were performed with blupf90 + software (Misztal et al., 2014) used to estimate the variance components, and the estimated breeding values (EBVs) for ST approach; and gibbs2f90 and Blupf90 + to estimate respectively the variance components and EBVs for MT approach. Gibbs2f90 was also used to estimate genetic correlations. We used the following gibbs2f90 parameters: chains equal to 200 000, burn-in equal to 30 000, and thinning interval equal to 100. We estimated the variance components, as well as genetic correlations, only in BLUP, and we used these estimates in GBLUP and ssGBLUP to make them comparable. Furthermore, the traits' heritability (h^2) was calculated as

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$$

where σ_a^2 is the additive genetic variance, σ_{pe}^2 is the variance associated to *pe* effect, and σ_e^2 is the residual variance.

The comparison among BLUP, GBLUP and ssGBLUP was done through the correlations between (G)EBVs estimated with different methods. The correlation was performed with *GGally* (R-package) and the function "ggpairs". To assess the traits' genetic trends for all the analysed methods, (G)EBVs were adjusted for a genetic base, defined as the average (G)EBV of animals born in 2016 (median of years of birth). In this way, all average (G)EBVs were plotted as departure from the same value (i.e., the average (G)EBV value of animals born in 2016). The values were then standardised using the animal genetic variance for comparison. The genetic trends of each method and traits were plotted using "ggplot2" R package for all animals born from 2010 and with both genomic and phenotypic data.

Validation

To compare BLUP, GBLUP, and ssGBLUP in terms of accuracy and bias, the linear regression method, proposed by Legarra and Reverter (2018) and used by Macedo et al. (2020a, b), was applied. The method's objective is to evaluate how the models work on newborn animals that do not have records. To this purpose, two phenotypic datasets were created: (i) whole (**W**), with all lactations recorded from 2010 to 2022; (ii) partial (**P**), where lactations recorded ≥ 2019 were excluded. EBVs were estimated twice, first using the W dataset and then the P dataset. Validation animals included genotyped females (569 and 296 animals respectively for Alpine and Saanen breeds) with phenotypic data in the W but not in the P dataset. According to Legarra and Reverter (2018), the following linear regression parameters were calculated on validation females' EBVs.

Dispersion:

$$b_{w,p} = \frac{cov(\widehat{u}_w, \widehat{u}_p)}{var(\widehat{u}_p)}$$

where \widehat{u}_w and \widehat{u}_p are the vectors of EBVs for the validation animals in the whole and partial datasets, respectively.

Table 1
Effects and number of levels per method and goat breed.

Breed	BLUP			GBLUP			ssGBLUP		
	HYS	Animal	pe	HYS	Animal	pe	HYS	Animal	pe
Alpine	1 058	5 669	5 669	1 058	2 422	2 422	1 058	6 324	6 324
Saanen	466	2 798	2 798	466	1 114	1 114	466	3 409	3 409

Abbreviations: BLUP = Best Linear Unbiased Prediction; GBLUP = Genomic Best Linear Unbiased Prediction; ssGBLUP: Single-Step Genomic Best Linear Unbiased Prediction; HYS = Herd - Year - Season; pe = permanent environment.

Prediction accuracy:

$$\widehat{acc} = \sqrt{\frac{cov(\widehat{u}_w, \widehat{u}_p)}{(1-\bar{F})\sigma_u^2}}$$

where \bar{F} is the average pedigree inbreeding of the validation animals and σ_u^2 is the additive genetic variance.

Correlation:

$$\rho_{w,p} = cor(\widehat{u}_w, \widehat{u}_p)$$

Finally, the predictivity was computed according to Misztal et al. (2014) as:

$$Pred = cor(\widehat{y}_w; \widehat{u}_p);$$

where \widehat{y}_w is the vector of predicted y calculated with predictf90 software (Misztal et al., 2014).

Results

Genomic, pedigree, and phenotypic data

After quality control, 2 422 (292 males and 2 130 females) Alpine, 1 114 (219 males and 895 females) Saanen, and a total of 48 262 SNPs out of the initial 50,058 were retained for further analyses. According to the pedigree correction procedure, progeny-sire attribution was correct for 83% of Alpine's pedigrees and 77% of Saanen's, whereas progeny-dam attribution matched 86 and 74% of pedigrees, respectively. Using *seekparentf90* and genomic data, it was possible to correct 44 (34% of the mismatches) and 22 (31%) progeny-sire and 21 (26%) and 26 (44%) progeny-dam mismatches in Alpine and Saanen breeds, respectively. For the remaining mismatching subjects, the sire/dam was considered missing.

Pedigree files of the 2 422 Alpine and 1 114 Saanen goats were analysed. The mean number of fully traced generations, the maximum number of traced generations, and the index of pedigree completeness were respectively 2.1 ± 1.6 , 10.4 ± 5.4 , and 0.7 ± 0.4 for Alpine and 2.2 ± 1.5 , 12.1 ± 5.0 , and 0.8 ± 0.4 for Saanen breeds.

The pedigree-based inbreeding coefficient of the genotyped animals was, on average, 0.02 ± 0.04 both for Alpine and Saanen goats.

Descriptive statistics for all analysed traits are reported in Table 2. More than one lactation was available for each animal: 4 ± 2 (mean \pm sd) in the Alpine and 3 ± 2 (mean \pm sd) in the Saanen breed.

Genetic parameters and breeding values

Genetic parameters were estimated using the BLUP method. The analysed traits were characterised by a moderate h^2 (Tables 3 and 4) in both breeds with the exception of PY in Alpine goats, which presented a moderate-to-high h^2 around 0.50. As shown in Table 5, we found a positive correlation among MY, PY, and FY (> 0.82) and between PP and PF (> 0.40). PP was positively correlated with PY and FP with FY.

For each trait, the correlation between the different methods was estimated (Supplementary Material S1 for Alpine breed and Supplementary Material S2 for Saanen breed). The correlations between breeding values estimated using GBLUP and ssGBLUP

Table 2
Descriptive statistics for the investigated production traits in Alpine and Saanen goat breeds.

Trait ¹	Alpine ²	Saanen ³
MY (l)	701.3 ± 214.9 (8 181)	777.1 ± 258.4 (2 589)
PY (kg)	24.1 ± 7.7 (8 100)	26.3 ± 8.6 (2 554)
FY (kg)	25.2 ± 8.2 (8 161)	26.8 ± 9.8 (2 578)
PP (%)	3.4 ± 0.3 (8 100)	3.3 ± 0.3 (2 554)
FP (%)	3.5 ± 0.5 (8 161)	3.4 ± 0.5 (2 578)

¹ Trait: MY = milk yield (l); PY = protein yield (kg); FY = fat yield (kg); PP = protein percentage; FP = fat percentage.

² Mean ± SD and number of data (in parenthesis) of the investigated production traits in Alpine breed.

³ Mean ± SD and number of data (in parenthesis) of the investigated production traits in Saanen breed.

were equal to 1 for all the traits, both using ST and MT approaches. The correlations between BLUP and GBLUP or ssGBLUP were high as well, all being over 0.91 in Alpine and 0.87 in Saanen goats. In particular, PP trait always showed the highest correlation.

The trends of (G)EBVs estimated for the genotyped goats born from 2010 to 2020 were evaluated for all the traits (Supplementary Materials S1 and S2 for Alpine and Saanen breeds, respectively). (G)EBV trends estimated using BLUP, GBLUP, and ssGBLUP were very similar. In both breeds and approaches, an almost constant increase in MY, PY and FY-related EBVs from 2011 was observed.

Table 3
Single-trait and multitrait BLUP variance components in Alpine goat breed.

Trait	Single-trait model			Multitrait model		
	heritability (CI95%)	σ_c^2	σ_{pe}^2	heritability (CI95%)	σ_c^2	σ_{pe}^2
Milk yield	0.30 (0.29–0.30)	6 487.5	1 489.1	0.28 (0.23–0.33)	6 190.0	1 837.0
Protein yield	0.33 (0.32–0.32)	8.1	0.7	0.29 (0.24–0.34)	7.3	1.7
Fat yield	0.28 (0.28–0.28)	9.1	0.5	0.25 (0.19–0.30)	8.1	1.8
Protein percentage	0.52 (0.52–0.53)	330.7	55.5	0.50 (0.43–0.57)	319.3	75.2
Fat percentage	0.34 (0.34–0.34)	610.4	110.7	0.31 (0.26–0.37)	570.9	173.6

Abbreviations: BLUP = Best Linear Unbiased Prediction; CI95% = confidence interval (95%); σ_c^2 = additive genetic variance; σ_{pe}^2 = permanent environment variance.

Table 4
Single-trait (ST) and multitrait (MT) BLUP variance components in Saanen goat breed.

Trait	Single trait model			Multitrait model		
	heritability (CI95%)	σ_c^2	σ_{pe}^2	heritability (CI95%)	σ_c^2	σ_{pe}^2
Milk yield	0.31 (0.30–0.32)	10 159	3 933.1	0.28 (0.16–0.39)	9 212.9	5 568.7
Protein yield	0.33 (0.33–0.33)	11.9	3.2	0.31 (0.20–0.42)	11.3	4.6
Fat yield	0.28 (0.28–0.28)	14.6	6.7	0.27 (0.18–0.37)	14.9	8.1
Protein percentage	0.33 (0.33–0.33)	208.6	182.9	0.35 (0.23–0.47)	231.2	185.5
Fat percentage	0.20 (0.20–0.20)	377.3	475.5	0.25 (0.14–0.35)	499.7	468.8

Abbreviations: BLUP = Best Linear Unbiased Prediction; CI95% = confidence interval (95%); σ_c^2 = additive genetic variance; σ_{pe}^2 = permanent environment variance.

Table 5
Genetic correlations estimated in Alpine (above diagonal) and Saanen (below diagonal) goat breeds.

Trait	MY	PY	FY	PP	FP
MY		0.94 (0.90–0.97)	0.83 (0.73–0.91)	–0.13 (–0.43–0.17)	–0.23 (–0.54–0.09)
PY	0.88 (0.85–0.91)		0.87 (0.80–0.92)	0.18 (–0.10–0.50)	–0.04 (–0.36–0.26)
FY	0.82 (0.77–0.87)	0.90 (0.86–0.93)		0.03 (–0.24–0.34)	0.32 (0.002–0.60)
PP	–0.35 (–0.47– –0.21)	0.12 (–0.03–0.25)	0.04 (–0.10–0.20)		0.41 (0.16–0.62)
FP	–0.35 (–0.51–0.23)	–0.06 (–0.20–0.09)	0.22 (–0.05–0.36)	0.64 (0.55–0.36)	

Abbreviations: MY = milk yield (l); PY = protein yield (kg); FY = fat yield (kg); PP = protein percentage; FP = fat percentage. The confidence interval (95%) for the correlations is reported in parentheses.

In a 10-year-interval, these EBVs increased of about 0.5 and 1 SD in Alpine and Saanen breeds, respectively. The trends of PP and FP were rather steady in Saanen goats, whereas in the Alpine breed, they showed an increase from 2018.

Validation

The results of the linear regression validation parameters and predictivity are reported in Table 6 and Table 7 for Alpine for Saanen, respectively.

Prediction accuracy ranged from 0.19 to 0.58 in Alpine and from 0.18 to 0.34 in Saanen. BLUP always showed lower values, whereas the two genomics methods had superimposable values but were slightly higher in ssGBLUP compared to GBLUP. Prediction accuracy tended to be slightly higher using ST approach in yield traits, whereas the opposite was true for percentage traits, especially FP in Saanen. Alpine showed greater prediction accuracy for all the traits and methods compared to Saanen. Our results showed a positive prediction accuracy gain using GBLUP and ssGBLUP instead of BLUP with a very small difference between ST and MT approach (on average GBLUP = +9% and ssGBLUP = +11% for Alpine, and GBLUP = +8% and ssGBLUP = +8% for Saanen).

Ideally, it is expected to have a dispersion parameter ($b_{w,p}$) equal to 1. If $b_{w,p}$ is < 1, an over-dispersion occurs and the model will give over-estimated results, and vice versa (Macedo et al.,

Table 6
Linear regression parameter results from validation of the estimates (Alpine goat breed).

Linear regression parameter	Single-trait model			Multitrait model		
	BLUP	GBLUP	ssGBLUP	BLUP	GBLUP	ssGBLUP
Milk yield						
\widehat{acc}	0.29	0.32	0.33	0.26	0.29	0.30
$b_{w,p} \pm SE$	0.73 ± 0.08	0.68 ± 0.07	0.68 ± 0.07	0.70 ± 0.08	0.66 ± 0.07	0.66 ± 0.07
$\rho_{w,p}$	0.35	0.38	0.39	0.33	0.37	0.38
Pred	0.70	0.71	0.71	0.69	0.70	0.70
Protein yield						
\widehat{acc}	0.31	0.38	0.39	0.29	0.37	0.38
$b_{w,p} \pm SE$	0.73 ± 0.08	0.74 ± 0.06	0.74 ± 0.06	0.73 ± 0.08	0.78 ± 0.06	0.78 ± 0.06
$\rho_{w,p}$	0.35	0.43	0.44	0.35	0.45	0.46
Pred	0.68	0.72	0.72	0.68	0.72	0.72
Fat yield						
\widehat{acc}	0.30	0.38	0.38	0.28	0.35	0.36
$b_{w,p} \pm SE$	0.76 ± 0.08	0.80 ± 0.07	0.80 ± 0.07	0.76 ± 0.08	0.80 ± 0.08	0.80 ± 0.06
$\rho_{w,p}$	0.35	0.44	0.45	0.36	0.45	0.45
Pred	0.64	0.71	0.71	0.67	0.71	0.71
Protein percentage						
\widehat{acc}	0.41	0.55	0.57	0.41	0.56	0.58
$b_{w,p} \pm SE$	0.91 ± 0.07	1.01 ± 0.05	1.00 ± 0.05	0.93 ± 0.07	1.01 ± 0.05	1.01 ± 0.05
$\rho_{w,p}$	0.48	0.63	0.64	0.50	0.65	0.66
Pred	0.76	0.81	0.82	0.76	0.82	0.82
Fat percentage						
\widehat{acc}	0.19	0.35	0.36	0.23	0.37	0.38
$b_{w,p} \pm SE$	0.54 ± 0.11	0.84 ± 0.07	0.83 ± 0.07	0.65 ± 0.10	0.83 ± 0.06	0.82 ± 0.06
$\rho_{w,p}$	0.20	0.44	0.44	0.27	0.48	0.49
Pred	0.67	0.73	0.73	0.69	0.75	0.75

Abbreviations: BLUP = Best Linear Unbiased Prediction; GBLUP = Genomic Best Linear Unbiased Prediction; ssGBLUP: Single-Step Genomic Best Linear Unbiased Prediction; \widehat{acc} = prediction accuracy; $b_{w,p}$ = dispersion; $\rho_{w,p}$ = correlation between estimated breeding values calculated on whole and partial datasets; Pred = predictivity.

Table 7
Linear regression parameter results from validation of the estimates (Saanen goat breed).

Linear regression parameter	Single-trait model			Multitrait model		
	BLUP	GBLUP	ssGBLUP	BLUP	GBLUP	ssGBLUP
Milk yield						
\widehat{acc}	0.21	0.27	0.27	0.20	0.27	0.27
$b_{w,p} \pm SE$	0.95 ± 0.17	0.94 ± 0.12	0.94 ± 0.12	1.00 ± 0.17	0.97 ± 0.12	0.97 ± 0.12
$\rho_{w,p}$	0.31	0.40	0.40	0.32	0.43	0.43
Pred	0.65	0.68	0.68	0.65	0.69	0.69
Protein yield						
\widehat{acc}	0.21	0.28	0.29	0.20	0.29	0.29
$b_{w,p} \pm SE$	0.91 ± 0.17	0.89 ± 0.11	0.89 ± 0.11	0.92 ± 0.17	0.87 ± 0.12	0.86 ± 0.11
$\rho_{w,p}$	0.29	0.39	0.39	0.29	0.40	0.41
Pred	0.64	0.68	0.68	0.64	0.69	0.69
Fat yield						
\widehat{acc}	0.22	0.33	0.33	0.21	0.34	0.34
$b_{w,p} \pm SE$	0.87 ± 0.16	0.86 ± 0.10	0.85 ± 0.10	0.77 ± 0.16	0.84 ± 0.10	0.83 ± 0.10
$\rho_{w,p}$	0.30	0.45	0.45	0.26	0.45	0.45
Pred	0.62	0.68	0.68	0.61	0.68	0.68
Protein percentage						
\widehat{acc}	0.18	0.26	0.26	0.19	0.27	0.27
$b_{w,p} \pm SE$	0.84 ± 0.20	1.22 ± 0.16	1.24 ± 0.16	0.93 ± 0.20	1.14 ± 0.16	1.16 ± 0.16
$\rho_{w,p}$	0.28	0.44	0.45	0.26	0.38	0.39
Pred	0.66	0.72	0.73	0.61	0.66	0.66
Fat percentage						
\widehat{acc}	0.22	0.25	0.26	0.26	0.32	0.32
$b_{w,p} \pm SE$	0.81 ± 0.11	1.11 ± 0.10	1.10 ± 0.10	0.62 ± 0.11	0.95 ± 0.11	0.95 ± 0.10
$\rho_{w,p}$	0.38	0.47	0.47	0.31	0.46	0.46
Pred	0.72	0.74	0.74	0.72	0.77	0.77

Abbreviations: BLUP = Best Linear Unbiased Prediction; GBLUP = Genomic Best Linear Unbiased Prediction; ssGBLUP: Single-Step Genomic Best Linear Unbiased Prediction; \widehat{acc} = prediction accuracy; $b_{w,p}$ = dispersion; $\rho_{w,p}$ = correlation between estimated breeding values calculated on whole and partial datasets; Pred = predictivity.

2020a). The model tested in our work slightly over-estimated breeding values in Alpine ($b_{w,p}$ was always < 1, except for PP where $b_{w,p}$ was close to 1 both for ST and MT). Instead, all models seemed to work better in Saanen, for which the $b_{w,p}$ values were close to 1

for most of the traits. For each trait in both breeds and approaches, we obtained similar $b_{w,p}$ for BLUP, GBLUP, and ssGBLUP methods, except for PP and PF where BLUP showed markedly higher dispersion than the other methods.

The correlation ($\rho_{w,p}$) was moderate for almost all the traits and methods; the highest correlation was obtained for ssGBLUP and GBLUP for PP in Alpine goats (respectively equal to 0.66 and 0.65), whereas the weakest correlation (0.20) was found for FP estimated using BLUP in the same breed. Genomic models outperformed BLUP with a markedly higher correlation, on average equal to GBLUP = +12% and ssGBLUP = +12% for Alpine and GBLUP = +12% and ssGBLUP = +13% for Saanen. The difference between BLUP and the genomic methods was particularly evident for the PP and FP traits (on average GBLUP = +19% and ssGBLUP = +19% for Alpine breed, GBLUP = +12% and ssGBLUP = +13% for Saanen breed). While in Alpine, overall, MT had the highest correlation coefficients, the opposite was true for the Saanen breed. Lastly, correlations were higher when calculated on Alpine than Saanen samples for all the investigated traits.

Predictivity was calculated both on male and female animals. For all the traits, ssGBLUP and GBLUP showed superimposable values, which were higher than those calculated for BLUP EBVs (on average GBLUP = +4% and ssGBLUP = +4% for Alpine breed, and GBLUP = +5% and for ssGBLUP = +5% for Saanen breed). Moreover, the differences between the two approaches were minimal, except for PP trait in Saanen breed, where ST had higher predictivity compared with MT. Alpine had a higher predictivity for all the traits with the exception of FP.

Discussion

With the aim to develop a new genomic evaluation in Alpine and Saanen breeds, this work compared three methods, BLUP, GBLUP, and ssGBLUP analysing how they predict breeding values.

It is known that the success of the genomic selection can be affected by the architecture of the traits, the structure of the reference population, the number of the phenotyped animals, and the relationship between training and validation (i.e., candidate to selection) populations (Massender et al., 2022; Wicki et al., 2023). The availability of both genotype and pedigree data allowed us to check the pedigree correctness or complete the pedigree information of the animals included in our analyses. This is especially important in sheep and goat breeding, where assessing the correct parentage can be challenging due to several factors, such as the high missing information rate, the limited use of artificial insemination, the suboptimal use of mating groups with the simultaneous presence of more males, and the extensive farming system (Cortellari et al., 2022). Our results showed that the pedigree of the two breeds presented mismatches (on average 20% for progeny-sire and 25% progeny-dam attributions), which, if both parents are genotyped, can be corrected by using genomic data.

Heritability was calculated for all the investigated traits using the BLUP method. The three investigated yield-related traits (milk, fat, and protein) showed moderate values consistently with Mucha et al. (2022), where heritability ranged from 0.27 to 0.30. Instead, for FP and PP in Saanen, we obtained lower values compared to the mean h^2 (0.48 for PP and PF) reported by Mucha et al. (2022). Additionally, previous research on French goats (Carillier et al., 2014; Teissier et al., 2019) found heritability values close to our results for both the involved breeds. Our estimates on the Alpine breed are also consistent with the Canadian ones (Massender et al., 2022).

We found very high correlations among EBVs across traits and methods. Indeed, the three tested methods showed very similar genetic trends within breed and trait. The trends suggest that the selection for PY and FY has produced good results increasing the yields without compromising the quality of milk. Moreover, in the Alpine, an increase of PP and FP, and thus the quality of milk, was observed in the last three years.

To understand how different predictive models work in the absence of phenotype, the linear regression method was used to validate our results obtained on the whole dataset, with all phenotypes, and the partial dataset, where phenotypes from 2019 were excluded. We calculated different linear regression parameters such as dispersion, correlation, and prediction accuracy. Furthermore, we calculated the predictivity of the estimates. According to Tsuruta et al. (2011), dispersion values within 15% from the unity are acceptable. For Alpine breed, $b_{w,p}$ was always below 0.85 suggesting that the models returned over-estimated EBVs. In this breed, ssGBLUP and GBLUP had higher values of $b_{w,p}$ than BLUP, except for MY. Instead, in Saanen breed, $b_{w,p}$ ranged from 0.85 to 1.15 for most of the traits. Carillier et al. (2014) used the regression coefficients of daughter yield deviations (DYD) on GEBV as an indicator of dispersion and found similar results, i.e., less than 0.85 for most of the traits in Alpine goats and higher values for Saanen goats. In both breeds, $\rho_{w,p}$ between whole and partial datasets increased from BLUP to the genomic methods. Furthermore, across all traits, GBLUP and ssGBLUP showed a higher predictivity than BLUP, on average equal to +4% for Alpine goat and +5% for Saanen goat. Compared to Carillier et al. (2014), we obtained a higher predictivity for all traits. Lastly, the prediction accuracy for ssGBLUP was higher than GBLUP and BLUP for both ST and MT approaches. The average accuracy gain across traits and approaches resulted equals to +13 and +10% from BLUP to ssGBLUP for Alpine and Saanen breeds, respectively. The linear regression results, in particular the prediction accuracy, showed that adding genotypes to the traditional BLUP leads to larger accuracy of the (G)EBVs. Similar results were obtained in other studies where the increase of accuracy, moving from BLUP to ssGBLUP, resulted between 0– +22% for dairy goats (Molina et al., 2018; Teissier et al., 2018; Scholtens et al., 2021; Massender et al., 2022; Mucha et al., 2022), between +9 and +16% for dairy sheep (Baloche et al., 2014; Macedo et al., 2020a, b), between +6 and +18% for dairy cows (Cesarani et al., 2021a; Mancin et al., 2021; Sungkhapreecha et al., 2021), and between +7 and +49% for buffaloes (Cesarani et al., 2021b).

In conclusion, this study, through a validation process using the linear regression method, shows that genomic methods can increase the accuracy of the breeding values especially when animals do not have phenotypes, and allow to check the paternity and maternity fixing mismatches in the pedigree. Owing to the solid reference population of Italian Alpine and Saanen goats, built by the genotyping effort of CHEESR and SHEEP&GOAT projects, the implementation of a genomic evaluation based on ssGBLUP method in the selection scheme of these breeds seems feasible and warranted to reduce the generation interval and improve the genetic gain, as well as to manage the inbreeding minimising the risk of inbreeding depression. Dissemination and communication of the potentiality of genomic selection will be crucial to enhance the profitability and the sustainability of this sector, as well as to encourage the development of a common international selection scheme.

Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.animal.2024.101118>.

Ethics approval

Datasets were obtained from pre-existing data based on routine animal recording procedures; moreover, DNA sampling for all individuals was conducted using nasal swabs and no invasive procedures were applied. Thus, in accordance with the 2010/63/EU guide and the adoption of the Law D.L. 04/03/2014, n.26 by the Ital-

ian Government, an ethical approval was not required for our study.

Data and model availability statement

The data presented in this study are not publicly available and are part of the official herd book information managed by the Italian Sheep and Goat Breeders Association (Asso.Na.Pa., Rome, Italy). For further information, please contact S.G. (direzione@assonapa.it).

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) did not use any AI and AI-assisted technologies.

Author ORCIDs

- A. Negro:** <https://orcid.org/0000-0002-2433-0210>.
A. Cesarani: <https://orcid.org/0000-0003-4637-8669>.
M. Cortellari: <https://orcid.org/0000-0002-5161-0648>.
A. Bionda: <https://orcid.org/0000-0001-9116-3208>.
P. Fresi: <https://orcid.org/0000-0002-2211-5306>.
N.P.P. Macciotta: <https://orcid.org/0000-0001-5504-9459>.
S. Biffani: <https://orcid.org/0000-0001-5559-3630>.
S. Grande: <https://orcid.org/0000-0001-8685-6202>.
P. Crepaldi: <https://orcid.org/0000-0002-6526-2162>.

CRedit authorship contribution statement

A. Negro: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **A. Cesarani:** Validation, Methodology, Investigation, Formal analysis, Conceptualization. **M. Cortellari:** Writing – review & editing, Writing – original draft. **A. Bionda:** Writing – review & editing, Writing – original draft. **P. Fresi:** Writing – review & editing, Project administration, Investigation, Formal analysis. **N.P.P. Macciotta:** Writing – review & editing, Methodology, Conceptualization. **S. Grande:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **S. Biffani:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **P. Crepaldi:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgements

The authors would like to thank the Italian Sheep and Goat Breeders Association (Asso.Na.Pa., Rome, Italy) for providing data supporting this study. Data were shared within CHEESR and SHEEP&GOAT projects.

Financial support statement

This research was funded by the Italian National Rural Development Plan (PSRN) – sub-measure 10.2, through the projects “Conservation, Health and Efficiency Empowerment of Small Ruminant (CHEESR – CUP: J89H18000210005)” and

“Sustainability Health Environment Economy Profitability & Genomic Organisation Animal (pheno)Typing (SHEEP&GOAT – CUP: J82C21000800005).

References

- Arnal, M., Larroque, H., Leclerc, H., Ducrocq, V., Robert-Granié, C., 2019. Genetic parameters for first lactation dairy traits in the Alpine and Saanen goat breeds using a random regression test-day model. *Genetics, Selection, Evolution: GSE* 51, 43. <https://doi.org/10.1186/s12711-019-0485-3>.
- Arnal, M., Robert-Granié, C., Ducrocq, V., Larroque, H., 2023. Validation of single-step genomic BLUP random regression test-day models and SNP effects analysis on milk yield in French Saanen goats. *Journal of Dairy Science* 106, 4813–4824. <https://doi.org/10.3168/jds.2022-22550>.
- Baloche, G., Legarra, A., Sallé, G., Larroque, H., Astruc, J.M., Robert-Granié, C., Barillet, F., 2014. Assessment of accuracy of genomic prediction for French Lacaune dairy sheep. *Journal of Dairy Science* 97, 1107–1116. <https://doi.org/10.3168/JDS.2013-7135>.
- Bermann, M., Cesarani, A., Misztal, I., Lourenco, D., 2022. Past, present, and future developments in single-step genomic models. *Italian Journal of Animal Science* 21, 673–685. <https://doi.org/10.1080/1828051X.2022.2053366>.
- Biffani, S., Tiezzi, F., Fresi, P., Stella, A., Minozzi, G., 2020. Genetic parameters of weeping teats in Italian Saanen and Alpine dairy goats and their relationship with milk production and somatic cell score. *Journal of Dairy Science* 103, 9167–9176. <https://doi.org/10.3168/jds.2020-18175>.
- Carillier, C., Larroque, H., Robert-Granié, C., 2014. Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. *Genetics, Selection, Evolution: GSE* 46, 67. <https://doi.org/10.1186/s12711-014-0067-3>.
- Cesarani, A., Garcia, A., Hidalgo, J., Degano, L., Vicario, D., Macciotta, N.P.P., Lourenco, D., 2021a. Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. *Journal of Dairy Science* 104, 5719–5727. <https://doi.org/10.3168/jds.2020-19838>.
- Cesarani, A., Biffani, S., Garcia, A., Lourenco, D., Bertolini, G., Neglia, G., Misztal, I., Macciotta, N.P.P., 2021b. Genomic investigation of milk production in Italian buffalo. *Italian Journal of Animal Science* 20, 539–547. <https://doi.org/10.1080/1828051X.2021.1902404>.
- Cesarani, A., Masuda, Y., Tsuruta, S., Nicolazzi, E.L., VanRaden, P.M., Lourenco, D., Misztal, I., 2021c. Genomic predictions for yield traits in US Holsteins with unknown parent groups. *Journal of Dairy Science* 104, 5843–5853. <https://doi.org/10.3168/jds.2020-19789>.
- Cortellari, M., Negro, A., Bionda, A., Grande, S., Cesarani, A., Carta, A., Macciotta, N., Biffani, S., Crepaldi, P., 2022. Using Pedigree and genomic data toward better management of inbreeding in Italian dairy Sheep and Goat breeds. *Animals* 12, 2828. <https://doi.org/10.3390/ani12202828>.
- Legarra, A., Reverter, A., 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genetics Selection Evolution* 50, 53. <https://doi.org/10.1186/s12711-018-0426-6>.
- Lourenco, D.A.L., Misztal, I., Tsuruta, S., Aguilar, I., Lawlor, T.J., Forni, S., Weller, J.L., 2014. Are evaluations on young genotyped animals benefiting from the past generations? *Journal of Dairy Science* 97, 3930–3942. <https://doi.org/10.3168/jds.2013-7769>.
- MacCluer, J.W., Boyce, A.J., Dyke, B., Weitkamp, L.R., Pfening, D.W., Parsons, C.J., 1983. Inbreeding and pedigree structure in Standardbred horses. *Journal of Heredity* 74, 394–399. <https://doi.org/10.1093/oxfordjournals.jhered.a109824>.
- Macedo, F.L., Reverter, A., Legarra, A., 2020a. Behavior of the linear regression method to estimate bias and accuracies with correct and incorrect genetic evaluation models. *Journal of Dairy Science* 103, 529–544. <https://doi.org/10.3168/jds.2019-16603>.
- Macedo, F.L., Christensen, O.F., Astruc, J.-M., Aguilar, I., Masuda, Y., Legarra, A., 2020b. Bias and accuracy of dairy sheep evaluations using BLUP and ssGBLUP with metafounders and unknown parent groups. *Genetics Selection Evolution* 52, 47. <https://doi.org/10.1186/s12711-020-00567-1>.
- Mancin, E., Tuliozi, B., Sartori, C., Guzzo, N., Mantovani, R., 2021. Genomic Prediction in Local Breeds: the Rendena Cattle as a case study. *Animals* 11, 1815. <https://doi.org/10.3390/ani11061815>.
- Massender, E., Brito, L.F., Maignel, L., Oliveira, H.R., Jafarikia, M., Baes, C.F., Sullivan, B., Schenkel, F.S., 2022. Single-step genomic evaluation of milk production traits in Canadian Alpine and Saanen dairy goats. *Journal of Dairy Science* 105, 2393–2407. <https://doi.org/10.3168/jds.2021-20558>.
- Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Miller, B.A., Lu, C.D., 2019. Current status of global dairy goat production: an overview. *Asian-Australasian Journal of Animal Sciences* 32, 1219–1232. <https://doi.org/10.5713/ajas.19.0253>.
- Misztal, I., Tsuruta, S., Lourenco, D.A.L., Aguilar, I., Legarra, A., and Vitezica Z. 2014. Manual for BLUPF90 family of programs.
- Misztal, I., Lourenco, D., Legarra, A., 2020. Current status of genomic evaluation. *Journal of Animal Science* 98, 1–14. <https://doi.org/10.1093/jas/skaa101>.

- Molina, A., Muñoz, E., Díaz, C., Menéndez-Buxadera, A., Ramón, M., Sánchez, M., Carabaño, M.J., Serradilla, J.M., 2018. Goat genomic selection: impact of the integration of genomic information in the genetic evaluations of the Spanish Florida goats. *Small Ruminant Research* 163, 72–75. <https://doi.org/10.1016/j.smallrumres.2017.12.010>.
- Mucha, S., Tortereau, F., Doeschl-Wilson, A., Rupp, R., Conington, J., 2022. Animal board invited review: meta-analysis of genetic parameters for resilience and efficiency traits in goats and sheep. *Animal* 16. <https://doi.org/10.1016/j.animal.2022.100456> 100456.
- Pocrnic, I., Lourenco, D.A.L., Masuda, Y., Legarra, A., Misztal, I., 2016. The dimensionality of genomic information and its effect on genomic prediction. *Genetics* 203, 573–581. <https://doi.org/10.1534/genetics.116.187013>.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81, 559–575. <https://doi.org/10.1086/519795>.
- Scholtens, M., Lopez-Villalobos, N., Lehnert, K., Snell, R., Garrick, D., Blair, H.T., 2021. Advantage of including genomic information to predict breeding values for lactation yields of Milk, Fat, and Protein or Somatic Cell Score in a New Zealand dairy Goat Herd. *Animals* 11, 24. <https://doi.org/10.3390/ani11010024>.
- Sungkhapreecha, P., Misztal, I., Hidalgo, J., Lourenco, D., Buaban, S., Chankitisakul, V., Boonkum, W., 2021. Validation of single-step genomic predictions using the linear regression method for milk yield and heat tolerance in a Thai-Holstein population. *Veterinary World* 14, 3119–3125. <https://doi.org/10.14202/vetworld.2021.3119-3125>.
- Teissier, M., Larroque, H., Robert-Granié, C., 2018. Weighted single-step genomic BLUP improves accuracy of genomic breeding values for protein content in French dairy goats: a quantitative trait influenced by a major gene. *Genetics Selection Evolution* 50, 31. <https://doi.org/10.1186/s12711-018-0400-3>.
- Teissier, M., Larroque, H., Robert-Granié, C., 2019. Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in French dairy goats. *Journal of Dairy Science* 102, 3142–3154. <https://doi.org/10.3168/jds.2018-15650>.
- Teissier, M., Larroque, H., Brito, L.F., Rupp, R., Schenkel, F.S., Robert-Granié, C., 2020. Genomic predictions based on haplotypes fitted as pseudo-SNP for milk production and udder type traits and SCS in French dairy goats. *Journal of Dairy Science* 103, 11559–11573. <https://doi.org/10.3168/jds.2020-18662>.
- Tsuruta, S., Misztal, I., Aguilar, I., Lawlor, T.J., 2011. Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. *Journal of Dairy Science* 94, 4198–4204. <https://doi.org/10.3168/jds.2011-4256>.
- Wicki, M., Raoul, J., Legarra, A., 2023. Effect of subdivision of the Lacaune dairy sheep breed on the accuracy of genomic prediction. *Journal of Dairy Science* 106, 5570–5581. <https://doi.org/10.3168/jds.2022-23114>.